



**Carla Sofia
Brites de Oliveira**

**Identification of molecules and pathways regulating
mistranslation in *Candida albicans***

**Identificação de moléculas e processos que
regulam os erros de tradução em *Candida albicans***

DECLARAÇÃO

Declaro que este relatório é integralmente da minha autoria, estando devidamente referenciadas as fontes e obras consultadas, bem como identificadas de modo claro as citações dessas obras. Não contém, por isso, qualquer tipo de plágio quer de textos publicados, qualquer que seja o meio dessa publicação, incluindo meios eletrônicos, quer de trabalhos académicos.



**Carla Sofia
Brites de Oliveira**

Identification of molecules and pathways regulating mistranslation in *Candida albicans*

Identificação de moléculas e processos que regulam os erros de tradução em *Candida albicans*

Dissertação apresentada à Universidade de Aveiro para cumprimento dos requisitos necessários à obtenção do grau de Mestre em Biologia Molecular e Celular, realizada sob a orientação científica da Doutora Ana Rita Macedo Bezerra, investigadora em Pós-Doutoramento do Departamento de Biologia da Universidade de Aveiro

o júri

Presidente

Doutora Maria Paula Polónia Gonçalves

Professor associada do Departamento de Biologia da Universidade de Aveiro

Arguente

Doutora Raquel Monteiro Marques da Silva

Investigadora do Instituto de Biomedicina da Universidade de Aveiro

Orientador científico

Doutora Ana Rita Macedo Bezerra

Investigadora Pós-Doutoramento do Departamento de Biologia da Universidade de Aveiro

agradecimentos

Em primeiro lugar gostaria de agradecer ao Prof. Doutor Manuel Santos pelos ensinamentos e por me ter dado a oportunidade de desenvolver a minha tese no laboratório de Biologia do RNA.

Este trabalho não teria sido possível sem a preciosa ajuda da minha orientadora Doutora Ana Rita Bezerra e do “meu orientador não oficial”, que me ensinou a trabalhar em série, Edgar Lopes. Agradeço a amizade, paciência e todo o tempo que disponibilizaram para me ajudar neste trabalho. Tive muita sorte em tê-los como orientadores, são das pessoas mais generosas que conheci no mundo académico.

A todo o grupo do Laboratório de Biologia do RNA por me terem acolhido tão bem desde o primeiro dia, pelos conselhos e bons momentos passados neste ano.

Aos meus amigos pelo carinho e motivação, que apesar da distância sempre que preciso estão lá para mim. Em especial às amigas que ganhei também por realizar esta tese, Rita, Jéssica, Lia, Sara e ao meu “braço direito” Ana.

Por último, e mais importante, à minha família em especial aos meus pais e ao meu irmão por acreditarem em mim e me apoiarem incondicionalmente.

palavras-chave

Candida albicans, erros de tradução, alteração genética, aminoacil-tRNA sintetases, infecção.

resumo

Candida albicans é o fungo patogénico mais predominante em humanos, causando doenças que podem variar entre ligeiras infeções de pele a infeções sistémicas severas em pacientes imunodeprimidos. A natureza patogénica deste organismo deve-se principalmente à capacidade de proliferação em vários locais do corpo humano e à sua capacidade de adaptação a mudanças drásticas no seu ambiente. *Candida albicans* exibe um sistema de tradução único, descodificando o codão de leucina CUG ambiguamente como leucina (3% dos codões) e serina (97%). Para tal usa um tRNA híbrido de serina ($\text{tRNA}_{\text{CAG}}^{\text{Ser}}$) que é aminoacilado por duas aminoacil-tRNA sintetases (aaRSs): leucil-tRNA sintetase (LeuRS) e seril-tRNA sintetase (SerRS). Trabalhos anteriores mostraram que a exposição de *C. albicans* a macrófagos, stress oxidativo, pH e antifúngicos aumenta os níveis de ambiguidade de 3% a 15%, sugerindo que *C. albicans* tem a capacidade de regular os níveis de erros de tradução em resposta às defesas do hospedeiro, antifúngicos e stress ambiental. Desta forma, a hipótese testada neste trabalho é a de que a variável incorporação de Leu e Ser nos codões CUG é dependente da competição entre LeuRS e SerRS pelo $\text{tRNA}_{\text{CAG}}^{\text{Ser}}$. Para testar esta hipótese, os níveis de SerRS e LeuRS foram indiretamente quantificados em diferentes condições fisiológicas, usando um sistema repórter fluorescente que determina a atividade dos respetivos promotores. Os resultados sugerem que o aumento de incorporação de leucina em codões CUG está associado a um aumento da expressão de LeuRS, sendo mantidos os níveis de SerRS.

Na segunda parte do trabalho, pretendeu-se identificar possíveis reguladores da expressão da SerRS e LeuRS. Para tal, uma coleção de estirpes de *C. albicans* com fatores de transcrição deletados foram transformadas com o sistema repórter fluorescente, de forma a quantificar a expressão das duas aaRSs. Alterações no rácio LeuRS/SerRS em estirpes deletadas relativamente à estirpe não deletada permitiram identificar 5 fatores de transcrição como possíveis reguladores da expressão destas duas aaRSs: *ASH1*, *HAP2*, *HAP3*, *RTG3* e *STB5*. Globalmente, este trabalho constitui o primeiro passo para elucidar o mecanismo molecular de regulação de erros de tradução em *C. albicans*.

keywords

Candida albicans, mistranslation, genetic code alteration, aminoacyl-tRNA synthetases, infection.

abstract

Candida albicans is the major fungal pathogen in humans, causing diseases ranging from mild skin infections to severe systemic infections in immunocompromised individuals. The pathogenic nature of this organism is mostly due to its capacity to proliferate in numerous body sites and to its ability to adapt to drastic changes in the environment. *Candida albicans* exhibit a unique translational system, decoding the leucine-CUG codon ambiguously as leucine (3% of codons) and serine (97%) using a hybrid serine tRNA (tRNA_{CAG}^{Ser}). This tRNA_{CAG}^{Ser} is aminoacylated by two aminoacyl tRNA synthetases (aaRSs): leucyl-tRNA synthetase (LeuRS) and seryl-tRNA synthetase (SerRS). Previous studies showed that exposure of *C. albicans* to macrophages, oxidative, pH stress and antifungals increases Leu misincorporation levels from 3% to 15%, suggesting that *C. albicans* has the ability to regulate mistranslation levels in response to host defenses, antifungals and environmental stresses. Therefore, the hypothesis tested in this work is that Leu and Ser misincorporation at CUG codons is dependent upon competition between the LeuRS and SerRS for the tRNA_{CAG}^{Ser}. To test this hypothesis, levels of the SerRS and LeuRS were indirectly quantified under different physiological conditions, using a fluorescent reporter system that measures the activity of the respective promoters. Results suggest that an increase in Leu misincorporation at CUG codons is associated with an increase in LeuRS expression, with levels of SerRS being maintained.

In the second part of the work, the objective was to identify putative regulators of SerRS and LeuRS expression. To accomplish this goal, *C. albicans* strains from a transcription factor knock-out collection were transformed with the fluorescent reporter system and expression of both aaRSs was quantified. Alterations in the LeuRS/SerRS expression of mutant strains compared to wild type strain allowed the identification of 5 transcription factors as possible regulators of expression of LeuRS and SerRS: *ASH1*, *HAP2*, *HAP3*, *RTG3* and *STB5*. Globally, this work provides the first step to elucidate the molecular mechanism of regulation of mistranslation in *C. albicans*.

Contents

List of Figures	VII
List of Tables	XI
List of Abbreviations	XIII
Chapter I – Introduction.....	1
1. Genetic code	3
1.1.1 Standard genetic code.....	3
1.1.2 Genetic code alterations	4
1.2 Genetic code components.....	5
1.2.1 tRNA	5
1.2.2 aaRSs	7
1.3 Translation	9
1.3.1 Mechanism of mRNA translation.....	9
1.3.2 Mistranslation	11
1.4 <i>Candida albicans</i>	13
1.4.1 <i>C. albicans</i> biology	13
1.4.2 <i>C. albicans</i> genetic code.....	15
1.4.3 Mistranslation in <i>C. albicans</i>	16
1.4.4 <i>C. albicans</i> interactions with host	18
1.5 Working hypothesis and objectives	20
Chapter II – Material and Methods.....	21
2.1 Strains and growth conditions	23
2.1.1 Strains	23
2.1.2 Standard growth conditions	24
2.2 Primers.....	24

2.3 Plasmid construction	25
2.3.1 Plasmids for LeuRS/SerRS quantification	25
2.3.2 Plasmids for quantification of Leu misincorporation	26
2.4 Expression of host transformation	27
2.4.1 Preparation of <i>E.coli</i> competent cells	27
2.4.2 Transformation of <i>E. coli</i>	27
2.4.3 Plasmidic DNA purification from <i>E.coli</i>	28
2.5 Manipulation of <i>C. albicans</i> strains	28
2.5.1 Transformation of <i>C. albicans</i>	28
2.5.2 DNA extraction from <i>C. albicans</i>	29
2.5.3 Integration confirmation	29
2.6 Choice of transcription factors mutants to analyse	30
2.7 Stress conditions	32
2.8 Epifluorescence Microscopy	32
2.9 Statistical Analysis	33
Chapter III - Results	35
3.1 SerRS and LeuRS expression and CUG ambiguous decoding	37
3.1.1 Overview	37
3.1.2 SerRS and LeuRS expression in different physiological conditions	37
3.1.3 Comparison of LeuRS/SerRS levels with Leu/Ser misincorporation levels .	40
3.2 Identification of potential transcription factors that control LeuRS and SerRS expression	41
3.2.1 Overview	41
3.2.2 Screening of TF knock-out collection	42

3.2.3 Comparison of Leu misincorporation and LeuRS/SerRS ratio expression in TF knock-out collection	45
Chapter IV – Discussion, conclusions and future perspectives	53
4.1 General discussion.....	55
4.2 Main conclusions and future perspectives.....	61
References.....	63
Annexes.....	75
A.1 – Plasmids described in the chapter 2, previously constructed for quantification of SerRS and LeuRS expression in <i>C. albicans</i> strains.....	75
A.2 – Plasmids described in the chapter 2, previously constructed for quantification of Leu misincorporation in <i>C. albicans</i> strains.	76
A.3 – Plasmids described in the chapter 2, used for quantification of Leu misincorporation in <i>C. albicans</i> strains.....	77

List of Figures

Figure 1.1 – Standard genetic code table.....	3
Figure 1.2 – The standard code and its variations.	5
Figure 1.3 – Structure of tRNA.	6
Figure 1.4 – Aminoacylation reaction	7
Figure 1.5 – mRNA translation.	10
Figure 1.6 – Causes of errors in protein synthesis in eukaryotes.	11
Figure 1.7 – Major morphologies of <i>Candida albicans</i>	15
Figure 1.8 – Secondary and tertiary structures of the <i>C. albicans</i>	16
Figure 3.1 – Schematic representation of the reporter systems used to quantify LeuRS and SerRS expression.	38
Figure 3.2 – LeuRS/SerRS expression ratio in <i>C. albicans</i> WT strain SN152 grown at different physiological conditions.	39
Figure 3.3 – LeuRS and SerRS expression in <i>C. albicans</i> WT strain SN152 grown at different physiological conditions.	40
Figure 3.4 – Comparison between Leu and Ser mistranslation at CUG sites and LeuRS/SerRS expression.....	41
Figure 3.5 – Heat maps of the LeuRS/SerRS ratio expression in <i>C. albicans</i> TF KO strains normalized to WT strain in several physiological conditions.	43
Figure 3.6 – Schematic representation of the reporter system used to quantify Leu incorporation at CUG codons.....	46
Figure 3.7 – Comparison of Leu misincorporation and LeuRS/SerRS ratio expression in WT strain SN152 and deletion strains <i>ASH1</i> , <i>HAP3</i> , <i>HAP2</i> , <i>RTG3</i> , and <i>STB5</i> obtained at control conditions.	47

Figure 3.8 – Comparison of Leu misincorporation and LeuRS/SerRS ratio expression in WT strain SN152 and deletion strains ASH1, HAP3, HAP2, RTG3, and STB5 obtained at 37°C.	48
Figure 3.9 – Comparison of Leu misincorporation and LeuRS/SerRS ratio expression in WT strain SN152 and deletion strains ASH1, HAP3, HAP2, RTG3, and STB5 obtained under osmotic stress.	49
Figure 3.10 – Comparison of Leu misincorporation and LeuRS/SerRS ratio expression in WT strain SN152 and deletion strains ASH1, HAP3, HAP2, RTG3, and STB5 obtained at pH6.	50

List of Tables

Table 2.1 – List of the <i>C. albicans</i> strains used in this study.	23
Table 2.2 – List of primers used in this study.....	24
Table 2.3 – Plasmids used in this thesis that were previously constructed by the host laboratory.....	25
Table 2.4 – Constructed plasmids during this study.	27
Table 2.5 – List of the <i>C. albicans</i> TFs selected for this study.	30
Table 2.6 – Conditions used in the stress experiments.	32
Table 4.1 – Resume of the validation results of the identified transcription factors	60

List of Abbreviations

aa-AMP	aminoacyl-adenylate
aaRS	aminoacyl-tRNA synthetase
AAS	amino acid accepting stem
ASL	anticodon stem loop
aa-tRNA	aminoacylated tRNA
AMP	adenosine monophosphate
ATP	adenosine triphosphate
CaCl₂	calcium chloride
CMT	Charcot-Marie-Tooth
DNA	deoxyribonucleic acid
DNTP's	deoxynucleotides
DSL	D-Stem loop
EDTA	Ethylenediamine tetraacetic acid
eIF	eukaryotic translation initiation factors
eRF	eukaryotic release factors
g (mg, µg)	gram (milligram, microgram)
KCl	potassium chloride
KO	knock-out
L (ml, µl)	liter (mililiter, microliter)
Leu	leucine
LeuRS	leucyl tRNA synthetase
LiAc	lithium acetate
M (mM, µM)	molar (milimolar, micromolar)
MAPK	mitogen activated protein kinase
Met	methionine
MM	minimal medium
MnCl₂	manganese chloride
MOPS Na	3-(N-Morpholino)propanesulfonic acid sodium salt
mRNA	messenger ribonucleic acid
NaCl	sodium chloride
OD	optical density
ORF	open reading frame

PAMPs	pathogen associated molecular patterns
PCR	polymerase chain reaction
PEG	polyethylene glycol
PheRS	phenylalanyl-tRNA synthetase
PPi	pyrophosphate
RbCl₂	rubidium chloride
RNA	ribonucleic acid
Rpm	revolutions per minute
SAP	Shrimp Alkaline Phosphatase
SAPK	stress-activated protein kinase
Ser	serine
SerRS	seryl tRNA synthetase
ThrRS	treonyl-tRNA synthetase
TF	transcription factor
Tris	tris(hydroxymethyl)aminomethane
tRNA	transfer ribonucleic acid
TSL	T-stem loop
ValRS	valyl-tRNA synthetase
yEGFP	yeast enhanced green fluorescent protein
YPD	yeast extract peptone dextrose
WT	wild type

Other abbreviations will be explained when used in the text.

Chapter I – Introduction

1. Genetic code

1.1.1 Standard genetic code

In 1968, Crick defined the genetic code as a universal non-overlapping triplet code that translates the nucleotide sequences into amino acid sequences (1). This code consists of 64 codons (Figure 1.1) comprising all possible three nucleotide combinations (from a set of four nitrogenated bases, A, T, C and G). Sixty one of those codons encode the canonical 20 amino acids, the building blocks of the proteins. The other three (TAG, TAA, TGA) lack a complementary tRNA and act as stop codons to end protein translation (2-4). Most amino acids, except methionine and tryptophan, are specified by more than one codon, showing that the genetic code is highly redundant (5, 6). For example, leucine (Leu) and serine (Ser) are codified by 6 different codons that are called synonymous codons.

By the time it was established, the genetic code was hypothesized as a “frozen accident”, a random event that originated the strict codon-amino acid association, whose structure was shared by organisms across all domains of life (1, 6). However, the discovery of two additional amino acids (selenocysteine and pyrrolysine) and the fact that the genetic code is not strictly universal contradict the “frozen accident” hypothesis and its origin and evolution are still open to debate (5, 7, 8).

		Second nucleotide					
		U	C	A	G		
First nucleotide	U	UUU Phe UUC Phe UUA Leu UUG Leu	UCU Ser UCC Ser UCA Ser UCG Ser	UAU Tyr UAC Tyr UAA STOP UAG STOP	UGU Cys UGC Cys UGA STOP UGG Trp	U	U
	C	CUU Leu CUC Leu CUA Leu CUG Leu	CCU Pro CCC Pro CCA Pro CCG Pro	CAU His CAC His CAA Gln CAG Gln	CGU Arg CGC Arg CGA Arg CGG Arg	C	C
	A	AUU Ile AUC Ile AUA Ile AUG Met	ACU Thr ACC Thr ACA Thr ACG Thr	AAU Asn AAC Asn AAA Lys AAG Lys	AGU Ser AGC Ser AGA Arg AGG Arg	A	A
	G	GUU Val GUC Val GUA Val GUG Val	GCU Ala GCC Ala GCA Ala GCG Ala	GAU Asp GAC Asp GAA Glu GAG Glu	GGU Gly GGC Gly GGA Gly GGG Gly	G	G

Figure 1.1 – Standard genetic code table.

The standard genetic code contains 64 codons of which 61 codify 20 different amino acids and 3 codify stop codons. Adapted from Clancy and Brown, 2008.

1.1.2 Genetic code alterations

The genetic code is highly conserved among all species and it was initially postulated by Crick as immutable, since any change would produce altered proteins, leading to proteome chaos that would be lethal to the cell (1, 9).

However, the discovery in 1979 that the stop codon UGA was decoded as tryptophan in human mitochondria has called the universality of the genetic code into question (10, 11). Since then, several genetic deviations have been found both in nuclear and organellar genomes (plastids and mitochondria) (12). Alterations in mitochondrial DNA have been described both in nonsense and sense codons in multicellular and unicellular eukaryotes (10, 13). For instance, the UAG codon is translated as tyrosine in calcareous sponges (14) but encodes leucine or alanine in some green algae (15). Arginine codons (AGA and AGG) have been reassigned to serine in Bilateria, to glycine in Urochordates and became stop codons in vertebrates (16). Also, mitochondria of the yeast species *Saccharomyces*, *Nakaseomyces* and *Vanderwaltozyma* decode the four Leu-CUN codons as threonine (17).

In contrast, in prokaryotic and eukaryotic nuclear genomes only nonsense codons reassignments have been found and part of them are found in mitochondria (13). The rationalization is that mitochondrial genomes are particularly tolerant to reassignments due to their reduced genome size and complexity, when compared to nuclear genomes (18). Examples of nuclear reassignments include the UGA translation as tryptophan in *Mycoplasma* spp. and cysteine in *Euplotes* spp., rather than a termination codon (9). The genera *Oxytricha*, *Paramecium* and *Tetrahymena* decode UAA and UAG codons as glutamine (19). Also, in *Bacillus subtilis* the UGA stop codon is decoded as tryptophan, but also retained its ability to be used for translation termination (20). Currently, the only exception is the decoding of the leucine CUG codon as serine in six species of the genus *Candida* and *Debaryomyces* (10, 13). In total, there are 23 known deviations from the genetic code in nuclear or mitochondrial codes (Figure 1.2)(12).

The genetic code has also suffered an expansion from 20 to 22 amino acids in the last years. The 21st amino acid is selenocysteine (Sec) and is incorporated in all three domains of life at UGA stop codons (7, 10, 21), while the 22nd amino acid is pyrrolysine

(Pyl) and is encoded by the UAG termination codon, mostly, in the *Methanosarcinaceae* family (22). Incorporation of Sec and Pyl in response to an in-frame stop codon is achieved by a complex recoding machinery to inform the ribosome not to stop at this position on the mRNA. Combined, genetic code expansion and deviations encountered contradict the “frozen accident” theory and show that the genetic code may be still evolving.

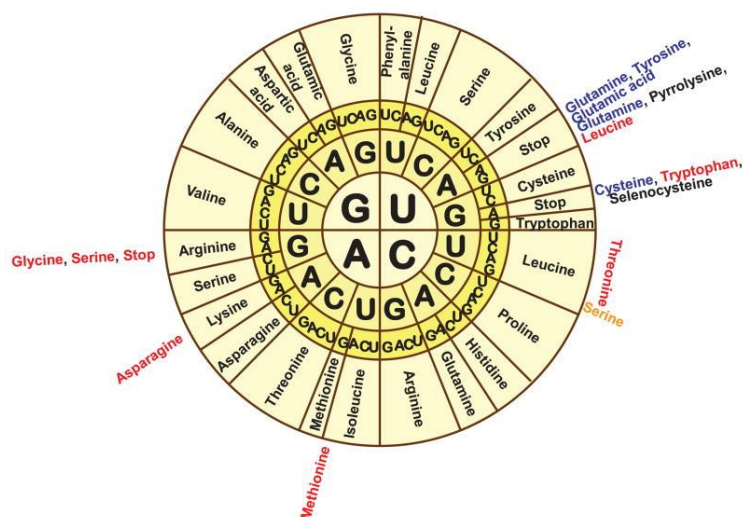


Figure 1.2 – The standard code and its variations.

The genetic code is shown in a circular form. Known differences to the standard genetic code are represented outside of the circle, where: red stands for mitochondrial, blue for some bacteria and eukaryotes and orange for the ambiguous yeast nuclear code. Adapted from Lobanov *et al.*, 2010.

1.2 Genetic code components

1.2.1 tRNA

Transfer RNAs (tRNAs) are short non-coding RNAs of approximately 70-100 nucleotides, which principal function is to recognize the codon and transfer the cognate amino acid to the growing polypeptide chain in the ribosome (23). Thus, tRNAs function as adaptors between the decoding of the genetic information and protein synthesis (24). Besides this function, tRNAs are also responsible for functions such as cell wall biosynthesis, protein labelling for degradation, apoptosis and precursors of small regulatory RNAs (25, 26).

Canonically, tRNAs have a clover-leaf secondary structure that is divided in 4 domains: an amino acid accepting stem (AAS) containing 7 base pairs followed by an unpaired nucleotide at position 73 and the 3'-CCA terminal, a D-Stem loop (DSL) with 4 base pairs and a D-loop of 8-11 bases, an anticodon stem loop (ASL) with a 5 base pairs stem and a 7 nucleotide loop, a T-stem loop (TSL) containing a stem of 5 base pairs and a variable region of 4 to 24 nucleotides. The AAS is where the amino acid is attached and the DSL and TSL are named by the presence of the conserved dihydrouridine (D) and ribothymidine (T) residues, respectively. The ASL detains the anticodon located between positions 34-36 in the centre of the loop and the variable arm connects the ASL and TSL stems. The presence of conserved and semi-conserved residues shapes the tRNA into its three-dimensional L-like structure (Figure 1.3) (27, 28).

In eukaryotes, tRNAs are transcribed as precursor molecules by RNA polymerase III and are submitted to a series of post-transcriptional alterations to generate mature tRNAs. These maturation steps occur in the nucleus and include the removal of the 5' and 3' nucleotide extensions and addition of a CCA sequence to the 3' terminal that is required for aminoacylation. Next, end-processed tRNAs are exported to the cytoplasm where introns are removed by the splicing machinery (29, 30)

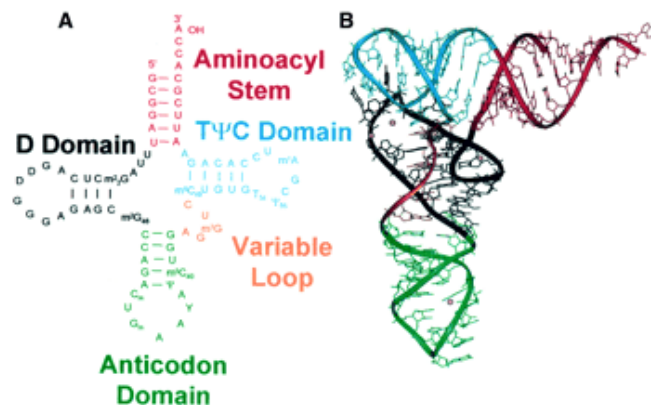


Figure 1.3 – Structure of tRNA.

(A) Cloverleaf secondary structure. (B) L-shaped tRNA tertiary structure. In both figures, red is the aminoacyl stem; blue is the t-stem-loop; orange is the variable loop; green is the anticodon stem-loop; and black is the d-stem-loop. Adapted from Giege *et al.*, 2012.

Apart from end processing and splicing, tRNA maturation also comprises a number of nucleotide modifications (30). In fact, tRNAs are the most modified type of RNAs. It is estimated that 15% to 25% of all nucleotides in eukaryotic tRNAs contain modifications.

These modifications include different reactions such as deaminations, glycosylation and methylations. These modifications play an important role in tRNA stability, tRNA discrimination and translation fidelity itself (31). Particularly, alterations at the wobble 34 position and position 37 in the anticodon loop. For instance, an anticodon with an U₃₄ decodes both A and G nucleotides while I₃₄ recognizes C, A and U bases, allowing tRNAs to read more than one codon (32). Modifications in the position 37 facilitate the maintenance of the loop conformation also important for an efficient anticodon-codon base pairing (33, 34). tRNA changes that affect codon/anticodon base pairing or recognition by aaRSs are the causes of the majority of the genetic code alterations (9, 21).

1.2.2 aaRSs

Aminoacyl-tRNA synthetases (aaRSs) are a family of enzymes responsible for charging tRNAs with their cognate amino acid, thus providing a relationship between the codon, anticodon and the correspondent building block of proteins (35). The attachment of the amino acid to its correspondent tRNA or aminoacylation occurs in two reactions. Firstly, the amino acid is activated with ATP at the aaRS active site generating aminoacyl-adenylates (AA-AMPs), and secondly, the intermediate AA-AMP is transferred to the 3' end of tRNA, releasing the AMP (Figure 1.4) (36, 37).

In eukaryotes, there are two groups of aaRSs, 20 cytoplasmic enzymes and 20 nuclear-encoded mitochondrial enzymes, 1 for each standard amino acid (38).

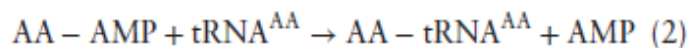
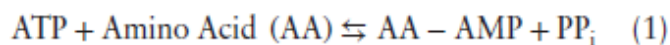
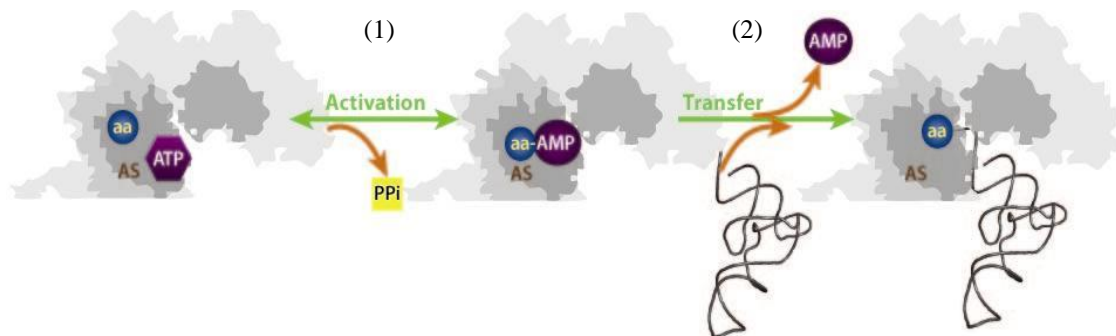


Figure 1.4 – Aminoacylation reaction.

(1) The amino acid is activated with ATP at the aaRS active site to form aminoacyl adenylate (aa-AMP), with the release of pyrophosphate (PPi). (2) The intermediate aaAMP is then transferred to the tRNA 3' end. Adapted from Ling *et al.*, 2009 and Pang *et al.*, 2014.

Synthetases can be separated in two groups based on the structural differences between the domains (catalytic and tRNA recognition sites). Class I enzymes have an active-site domain that forms a *Rossmann* fold (five-handed parallel β -sheet) that promotes the aminoacylation reaction, binds the minor groove of the tRNA acceptor branch, and normally aminoacylates the 2'-OH group of the terminal tRNA nucleotide. In contrast, class II enzymes aminoacylation sites have an antiparallel β -sheet, bind the major groove of the tRNA acceptor stem, and aminoacylate the 3'-OH position of the terminal ribose (38, 39).

Accurate recognition of tRNAs by aaRSs is vital for the fidelity of the aminoacylation reaction and, consequently, to the fidelity of protein synthesis. tRNAs share the same secondary and tertiary structures. So, aaRS distinguish tRNAs by a set of domains, called identity elements (37, 40). Identity elements include the position 73 (N73) (tRNAs for chemically similar amino acids have the same nucleoside at position 73) and nucleotide variations at the anticodon loop and acceptor stem of the tRNA, regions directly involved with aaRSs interaction. These are grouped either as determinants, elements that promote aminoacylation, or anti-determinants, elements that prevent mischarging (37). Interesting examples of anti-determinants are the G73 nucleotide on yeast tRNA^{Ser} that hinders LeuRS recognition and the A73 nucleotide on tRNA^{Leu} that prevents SerRS recognition (37, 41).

In addition, aaRSs possess quality control mechanisms to assure a correct match between the amino acid and their cognate tRNA. It is estimated that aaRSs have an error rate of 10^{-4} in the amino acid selection and 10^{-6} in tRNA discrimination (42). Some synthetases have the ability to hydrolyse AA-AMPs at the catalytic site (*pretransfer editing*), while others carry a specific editing domain to deacylate tRNAs incorrectly charged (*posttransfer editing*) or some of them use both of the editing mechanisms (42, 43). SerRS, MetRS and class II LysRS are the only three enzymes that do not possess domains for posttransfer editing. So, these three aaRSs only perform pretransfer editing in

the catalytic site. Essentially, incorrect AA-AMPs or mischarged AA-tRNAs are hydrolysed by the nucleophilic attack of water to release the amino acid (42, 44).

Although aaRSs play an important role in the translation machinery, as mentioned, several non-canonical functions have been discovered in the last years. Among them are RNA splicing, transcriptional and translational regulation through binding with transcription factors and involvement in signalling responses, such as apoptosis and inflammation (45). Considering the range of activities and uncovered protein-protein interactions, it is thought that these synthetases are key players in the response to the various cellular stresses to maintain homeostasis (46). Moreover, altered expression of aaRSs has been associated with human disorders such as cancer, neuronal and autoimmune diseases (47, 48). For this reason, aaRSs are an attractive target for the development of therapies (49), including approaches that aim their active sites in pathogenic microorganisms (47).

1.3 Translation

1.3.1 Mechanism of mRNA translation

The genetic information flows from DNA to RNA to protein. Messenger RNA molecules (mRNAs) are produced by transcription from DNA and then are processed. In eukaryotic cells, mRNAs resulting from transcription are interrupted by introns that are removed by a splicing reaction. Additionally, both ends of the mRNAs are modified: a m⁷G cap is added to the 5'-end and the 3'-end is polyadenylated with a poly A tail (50). Once the mRNAs are processed, their *open reading frames* (ORF) are translated into the amino acid sequence of a protein following the rules of the genetic code. The mRNA sequence is translated into an amino acid sequence until a stop codon is encountered. Then, the ribosome releases the finished protein (50, 51).

Therefore, the translation cyclical process can be divided into three steps: initiation, elongation and termination (Figure 1.5). To initiate translation in eukaryotic mRNAs, besides the ribosome, at least 12 proteins are needed. They are called eukaryotic initiation factors (eIFs). The translation initiation phase starts with the formation of a ternary complex consisting of GTP, a methionyl tRNA (initiator tRNA) and the initiation factor

eIF2. The ternary complex and additional initiation factors associate with the ribosomal 40S subunit to form a 43S pre-initiation complex. These factors attach to the capped 5' end of the mRNA and scan the mRNA in the 3' direction, until the AUG initiation codon is encountered, and pair with the initiator tRNA at the ribosomal P-site. Once it happens, the 60S ribosomal subunit is joined to begin the polypeptide elongation (52-54). The next codon of the ORF is present in the A (acceptor) site of the ribosome until the binding of the correspondent aminoacyl-tRNA (aa-tRNA) (55). Elongation proceeds and the peptide chain is assembled step-by-step in accordance with the sequence of the mRNA. This phase is less complex than initiation, as mainly only two eukaryotic elongation factors (eEFs) are required: eEF1A and eEF2 (51, 54), with the exception of yeast and higher fungi that use an additional eEF3 (55). eEF1A helps the delivery of the charged tRNA to the ribosome, eEF2 promotes the movement of the tRNAs present in the A and P sites to P and E (exit) sites respectively, and it is thought that eEF3 may aid the release of deacylated tRNA from the E site (55, 56). When the ribosome encounters a stop codon, the release factor eRF1 (eukaryotic elongation release factor 1) recognizes it, and in association with eRF3 (eukaryotic elongation release factor 3), stimulates the release of the peptide chain and the ribosomal subunits (51, 54). Subunits of the ribosome dissociate, mRNA and deacylated tRNA are released, originating the necessary components for the next translation cycle (55, 57).

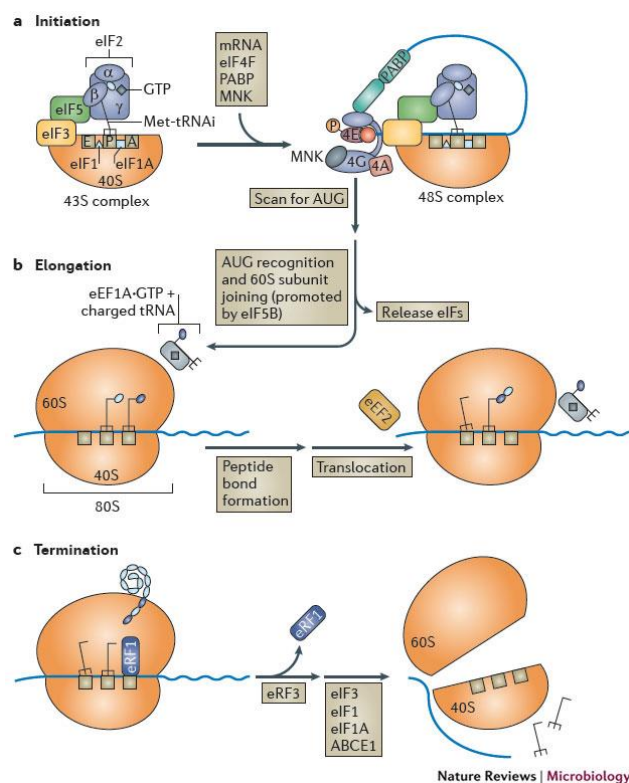
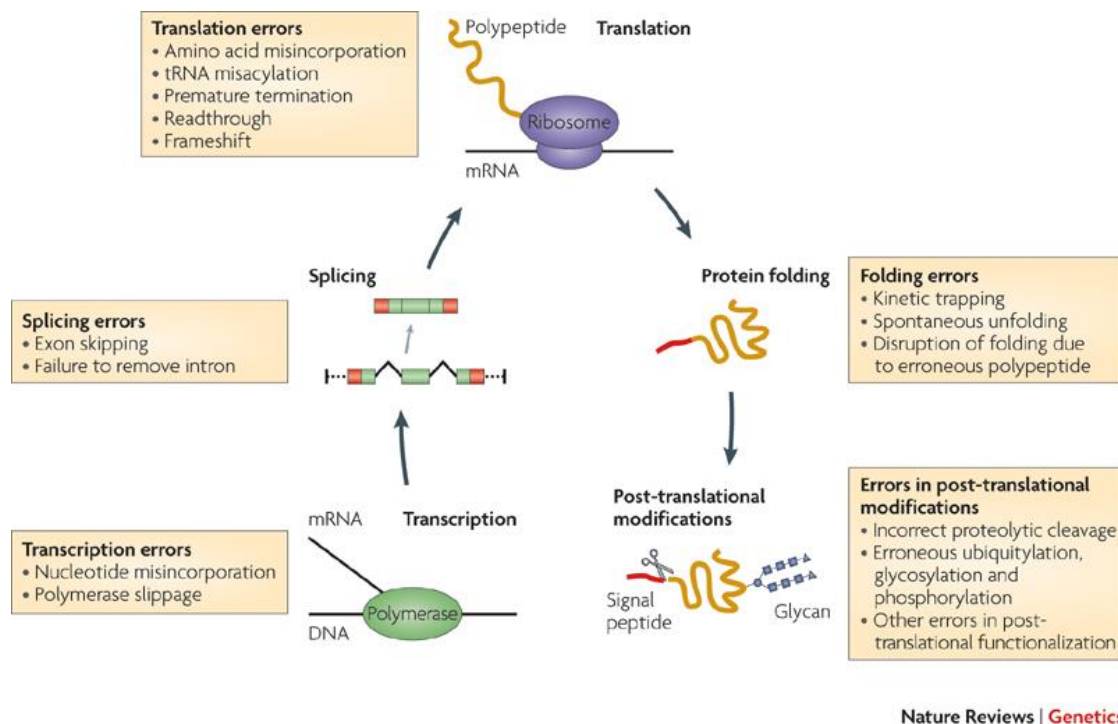


Figure 1.5 – mRNA translation.

Overview of the three stages of mRNA translation in eukaryotes: **(a)** initiation, **(b)** elongation and **(c)** termination. **a)** Assembly of the 43S pre-initiation complex and scan of the mRNA until the AUG initiation codon is encountered. **b)** Peptide chain elongation according with the mRNA sequence. **c)** Stop codon is recognized, activating the release of the peptide. The complex is dissociated and the ribosomal subunits are recycled. Adapted from Walsh and Mohr 2011.

1.3.2 Mistranslation

In spite of the surveillance mechanisms, errors in protein synthesis are inevitable. It is estimated that during translation 1 mistake per 10^4 codons translated is made (58). In eukaryotes, these errors can occur at all stages of the flow of the genetic information: transcription, splicing and translation. Even in cases where proteins have the correct amino acid sequence, post-translational modifications errors as well as folding errors may alter their function (Figure 1.6) (58, 59). The error frequency in protein synthesis is not fully understood, but evidence suggests that the translation process is the most error prone and that error rates are similar in prokaryotes and eukaryotes (60).

**Figure 1.6 – Causes of errors in protein synthesis in eukaryotes.**

Altered proteins can result from errors in different phases, such as, transcription, splicing mechanisms, translation, protein folding and protein modifications after the translation process. Adapted from Drummond and Wilke 2009.

Translation errors can occur during tRNA charging by aaRSs or during mRNA decoding by the ribosome. Mischarging errors are mainly caused by failure of the aaRSs to distinguish similar amino acids or by the incorrect recognition of tRNAs. These errors can be prevented by aaRSs editing mechanisms, which clear the incorrect bound amino acids and by specific tRNA-aaRS interaction (60, 61). Codon decoding errors fall into three categories: nonsense errors, missense errors and frameshifting. Nonsense errors occur when the ribosome prematurely terminates the translation of a coding sequence. Missense errors occur when the wrong amino acid is incorporated. Finally, frameshifting results from the loss of the mRNA reading frame, leading to premature termination (58, 62). These errors contribute to the production of misfolded proteins (58), which can be refolded with the aid of molecular chaperones or can be degraded through the ubiquitin-proteasome system. When these unfolded proteins exceed the capacity of cells to remove them, toxic aggregates can accumulate (63, 64).

Normally, mistranslation is viewed as deleterious because it produces proteins with altered function that can reduce growth rate and fitness (59, 65). Several alterations in aaRSs are linked with disease, namely cancer, type 2 diabetes mellitus and neuronal disorders (47, 48). One good example is the mouse sticky mutation which is an alteration in the editing domain of alanyl-tRNA synthetase that compromises the removal of mischarged tRNAs, introducing genetic code ambiguity. Ultimately, it causes accumulation of misfolded proteins that leads to neurodegeneration (66). Following this discovery, it was investigated if editing defects in another aaRS (Valyl-tRNA synthetase-ValRS) in mammalian cells would have similar consequences. It was observed that mutations in the editing domain of ValRS also induced an increase of amino acid misincorporation. Additionally, the editing alteration in ValRS is sufficient to modify cell morphology and initiate caspase-dependent apoptosis (67). Mutations in at least five different aaRS genes (namely the glycyl-tRNA synthetase gene) were found in patients with peripheral neuropathy Charcot-Marie-Tooth (CMT). The direct link between aaRS mutations and the disease remains elusive, but evidence suggests that impaired tRNA charging in peripheral nerve axons may be in play (68). These investigations demonstrate that mistranslation in mammalian cells may be related to cell pathology.

On the other hand, there is evidence that in numerous cases mistranslation originated altered proteins that contribute to an improved cellular stress response and adaptation (69, 70). It was reported that *Mycoplasma* spp. have editing defective LeuRS, PheRS (phenylalanyl-tRNA synthetase) and ThrRS (treoninyl-tRNA synthetase). These synthetases have mutations and deletions in their editing domains, resulting in mischarging of the corresponding tRNAs. These mischarged tRNAs increase mistranslation and ultimately lead to the production of statistical proteins. The increase of mistranslation levels could be the source of antigenic diversity that *Mycoplasma* uses to escape host immune system (71). Another example is the mysacylation of Met to non-methioninyl-tRNAs (Met-tRNAs) in *E.coli*, yeast and mammalian cells, under oxidative stress. It was proposed that Met reacts with ROS (reactive oxygen species) produced by oxidative stress, protecting the proteome from oxidative damage (59, 72). In *Mycobacteria*, high rates of substitution of glutamate for glutamine and aspartate for asparagine is also beneficial. Misincorporation of glutamate and aspartate generate protein variants important for phenotypic tolerance to the antibiotic rifampicin (73).

The mistranslation phenomenon can also be observed in several species of the genus *Candida*, particularly in *Candida albicans*, where the leucine CUG codon is read by a tRNA (tRNA_{CAG}^{Ser}) as leucine (~3%) and serine (~97%) (74). Strains that misincorporate increasing levels of leucine at CUG codons show remarkable phenotypic diversity and have a better response to stressful conditions such as presence of metals (75), antifungal drugs, alteration of pH and temperature (76) and macrophage phagocytosis (77).

Altogether, these examples indicate that genetic code alterations are not mere abnormalities and can, in fact, represent a potential to adaptation, allowing species to colonize new ecological niches.

1.4 *Candida albicans*

1.4.1 *C. albicans* biology

Candida albicans is an eukaryotic yeast of the fungal kingdom (78). *C. albicans* genome is divided in eight pairs of chromosomes, which genome size is 14.3-14.4 Mb encoding 6,107-6,159 genes (79). For a long time, it was thought that *C. albicans*

was an obligate diploid (2N) but recently tetraploid and haploid cells have been detected (80).

This fungus is an opportunistic commensal which colonizes skin, genital and/or intestinal mucosa of 30-70% of healthy individuals and, under normal circumstances, does not cause significant disease (81). However, under certain conditions such as treatment with antibiotics, aggressive chemotherapy, diabetes and immune suppression, *C. albicans* is capable of causing infection in the host (82). *C. albicans* can be responsible for a range of infections, from mild skin and mucosal infections to severe systemic infections in a number of organs (81, 83, 84). Although infections from other species of *Candida* have become more widespread, *C. albicans* is still the most prevalent human pathogen (85).

Interestingly, *C. albicans* cells have the ability to present three different morphologies: yeast (blastospores) and the filamentous forms pseudohyphae and hyphae (Figure 1.7). Cells with unicellular yeast morphology are similar to yeast *Saccharomyces cerevisiae*. Pseudohyphal cells are in the intermediary form and consist of attached elongated cell buds that keep constrictions at the septa formation, while hyphal cells have a long shape with no visible constrictions (86, 87). This form of morphological switching depends upon extracellular conditions and is considered important to stress adaptation and virulence (88).

Another important growth feature of *C. albicans* is the “white-opaque” transition. *C. albicans* can reversibly change from the normally white and domed-shaped colonies (containing round-shaped cells – white cells) to opaque and flat colonies (containing elongated cells) (86, 89). The “white-opaque” switch is one of the key regulators of mating in *C. albicans* (90). To date, no complete sexual cycle has been described in *C. albicans* although an elaborate mechanism for mating does exist. The described cycle for *C. albicans* is a parasexual cycle. In this case, mating of diploid cells is followed by mitosis and concerted chromosome loss instead of meiosis (91). Diploid cells (2n) of *C. albicans* are typically heterozygous at the mating type locus (MTL), *MTLa* and *MTLa*, but cells may lose their heterozygosity by loss of one copy of Chromosome 5, eliminating the a or α allele. The resulting α/α or a/a diploid strains (2n, opaque cells) undergo pheromone signalling between cells of opposite genotypes. When diploid a/a and α/α cells mate, they form tetraploid aa/ $\alpha\alpha$ cells (4n, white cells) which undergo mitosis and non-

meiotic reduction in the number of chromosomes to return to diploid a/α cells ($2n$, white cells). The parasexual cycle is completed with no recognized meiosis (89, 92). Apparently, haploid cells are also formed via a non-meiotic reduction in chromosome number, similar to that seen in the tetraploid-to-diploid transition (80). The existence of semi-stable, non-diploid *C. albicans* cell types highlights the flexibility of the *C. albicans* genome (93).

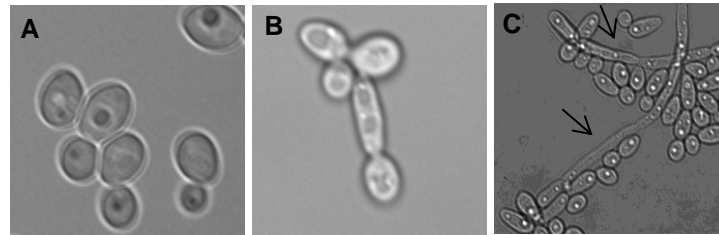


Figure 1.7 – Major morphologies of *Candida albicans*.

A- Yeast cells (also called blastospores). **B-** Pseudohyphal cells. **C-** Hyphal cells. (Cells photographed at 630x magnification).

1.4.2 *C. albicans* genetic code

Several *Candida* species, *Pichia stipites*, *Debararyomyces hansenii* and *Lodderomyces elongisporus* (species belonging to the CTG clade), have the exceptional capability to translate the Leu-CUG codon as serine using the $\text{tRNA}_{\text{CAG}}^{\text{Ser}}$ in the cytoplasm (94). It is estimated that the reassignment process began approximately 275 million years ago with the appearance of the hybrid $\text{tRNA}_{\text{CAG}}^{\text{Ser}}$ (95, 96). The $\text{tRNA}_{\text{CAG}}^{\text{Ser}}$ was able to compete with the natural $\text{tRNA}_{\text{CAG}}^{\text{Leu}}$ decoder for the CUG codon, thus generating an ambiguous CUG codon. This enabled CUG reassignment through selection of the mutant Ser-tRNA and elimination of the cognate Leu-tRNA. The Ser-tRNA_{CAG} was maintained in the lineage that originated the genus *Candida*, but was lost in the lineage leading to the genus *Saccharomyces*. This separation occurred 170 million years ago, indicating that the yeast ancestor was ambiguous for at least 100 million years. CUG codon ambiguity imposed strong negative selection against old CUG codons, which mutated to UUG and UUA codons and this resulted in low CUG codon usage (94). Currently in the genus *Candida*, some species still have CUG ambiguity (as in the case of *C. albicans*), while others have achieved complete reassignment of the CUG codon (such as *C. cylindracea*) (95, 96).

In *C. albicans*, the CUG is ambiguously decoded as Ser (~97%) and Leu (~3%) by this novel tRNA_{CAG}^{Ser}, which is charged by two aaRSs: SerRS (seryl-tRNA synthetase) and LeuRS (leucyl-tRNA synthetase) (74, 97). This dual recognition by SerRS and LeuRS originates two aminoacyl-tRNAs (Leu-tRNA_{CAG}^{Ser} and Ser-tRNA_{CAG}^{Ser}), which insert either Leu or Ser at CUG sites during mRNA translation (95).

The hybrid nature of tRNA_{CAG}^{Ser} is due to the presence of identity elements for both aaRSs. For example, nucleotides A₃₅ and G₃₇ in the anticodon loop are directly identified by LeuRS. Contrarily, SerRS recognizes three G-C base pairs in the variable loop and the Ser-tRNA discriminator base G₇₃. In addition, the tRNA_{CAG}^{Ser} contains a guanosine at position 33 (G₃₃), located to the 5' base of the anticodon, in substitution of the regular uridine (U₃₃). Since this mutation induces a modification of the tRNA_{CAG}^{Ser} anticodon stem that lowers the efficiency of LeuRS binding, it is considered a key element in CUG reassignment (Figure 1.8). (13, 95, 98).

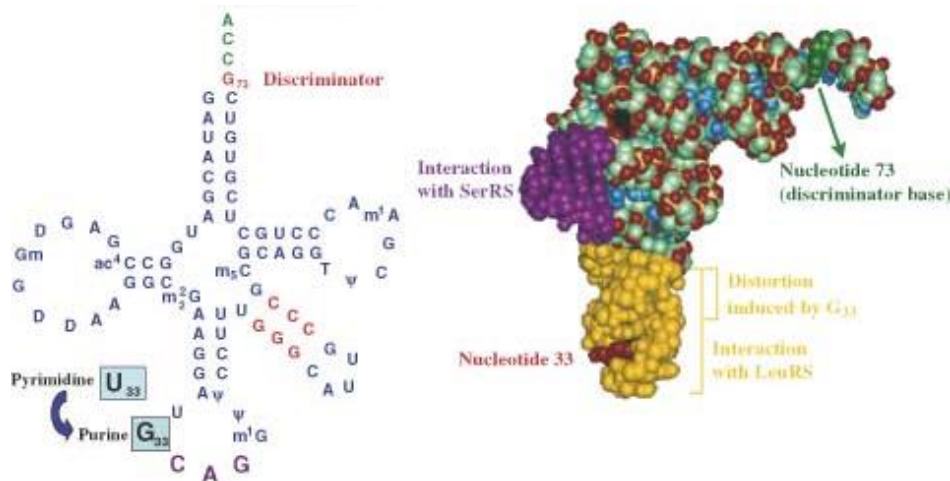


Figure 1.8 – Secondary and tertiary structures of the *C. albicans* tRNA_{CAG}^{Ser}.

Representation of identity elements for both SerRS and LeuRS. Adapted from Miranda *et al.*, 2007.

1.4.3 Mistranslation in *C. albicans*

Considering that serine is hydrophilic, whereas leucine is hydrophobic, the variable incorporation of these two amino acids into a protein has the potential to create proteins with altered function (77, 99). Previous study from Gomes *et al.* demonstrated that levels of leucine misincorporation at the CUG codons increase in response to different growth

conditions in *C. albicans*: oxidative stress, temperature and pH. Cells grown at 30°C presented 2.96% of Leu misincorporation, while at 37°C the values increased to 3.9%, to 4.03% in presence of hydrogen peroxide (H₂O₂) and to 4.95% at pH4. This study also described that there are 13,074 CUG codons distributed over 66% of *C. albicans* genes, which suggests that CUG ambiguity contributes to the expansion of the proteome and phenotypic diversity (74). Interestingly, the distribution of the CUG codons in the genome is non-random. Rocha and colleagues performed an extensive structural analysis of *C. albicans* proteins containing CUG-encoded residues (99) and revealed that 90% of the CUG codons are located in nonconserved positions where both leucine and serine can be introduced without major disruption of protein structure and function. Data was reinforced with the crystal structures of the two isoforms of *C. albicans* SerRS. Leucine or serine incorporation at the CUG position in *C. albicans* SerRS induced only local structural changes. Consequently, the Leu/Ser ambiguity cause minimal protein misfolding and it explains the tolerance of *C. albicans* cells to mistranslation (99).

High tolerance to CUG ambiguity was unequivocally demonstrated by Bezerra *et al.* In this study, *C. albicans* strains were constructed by engineering Ser-tRNAs to misincorporate increasing levels of Leu at CUG codons. These recombinant strains tolerated levels ranging from 20% to 99% of Leu at CUG sites. Surprisingly, the misincorporating strains grew faster than the control strain in the presence of oxidative stressors (menadione and H₂O₂) and protein misfolding agents (guanidine hydrochloride and urea). In addition, ambiguous cells were more tolerant against the antifungals, fluconazole and itraconazole, than the control. Moreover, CUGs are prevalent in genes associated with membrane and cell wall processes, which suggest that CUG ambiguity may influence drug resistance and the way *C. albicans* interacts with the immune system (76).

The transcriptional response to mistranslation has been investigated by gene profiling in yeast. To do so, *Saccharomyces cerevisiae* strains were transformed with a plasmid containing one copy of the hybrid *C. albicans* tRNA_{CAG}^{Ser} (100). This manipulation increased mistranslation in yeast to 1.4%, which triggered up-regulation of genes related to oxidative and general stress, carbohydrate metabolism and molecular chaperones. In contrast, genes involved with the protein synthesis machinery were down-

regulated (100, 101). This group of genes are usually included in a core transcriptional response to stress, known as ESR (approximately 220 genes in *S. cerevisiae*). These results showed how mistranslation could induce resistance to stress and induce phenotypic variability. Cell tolerance to external stressors can be viewed as a secondary effect of the activation of the environmental stress response (ESR) caused by mistranslation (101).

However, the classical response to stress found in *S. cerevisiae* is not observed in *C. albicans*. *C. albicans* only has a small number of genes (about 20) implicated in the core stress response. The common stress response is mainly regulated by Hog1 SAPK (stress-activated protein kinase) and the transcription factor Cap1 (102, 103). Cap1 possess three conserved cysteine residues in the C-terminal region that become oxidised under oxidative stress (104). This induces its nuclear accumulation and subsequent activation of the target genes, which have antioxidant functions (e.g. catalase, glutathione reductase). In accordance, Cap1 inactivation decreases the expression of these genes (103). Thus, Cap1 is considered a key player in the transcriptional response to the reactive oxygen species (ROS) produced by the host immune cells (102-104). Hog1 is an element of the evolutionary conserved mitogen activated protein kinase (MAPK) family, which is involved in the response to osmotic stress in yeast. In presence of cationic and osmotic stressors, such as NaCl and KCl, Hog1 is activated by phosphorylation and it accumulates in the nucleus, leading to the activation of the target genes (e.g. a glycerol phosphatase) (102, 105). Additionally, Hog1 SAPK is also activated in response to heavy metals, heat and oxidative stress (106). Cells with defective Hog1 SAPK function presented impaired virulence in mouse models (107).

1.4.4 *C. albicans* interactions with host

The ability of *C. albicans* to survive and infect different host locations is supported by a number of virulence factors and other factors that promote the virulence without interacting directly with the host (fitness attributes). Virulence factors include expression of adhesins and invasins on the cell surface, which allow the complex adherence to the surface of the host and then permit the penetration in epithelial tissue (81, 108). Phenotypic switching between yeast to hyphal form is considered essential for invasion and pathogenicity. Biofilm formation is another important virulence factor, considering that

these are constituted by groups of cells embedded in an extracellular matrix material that protects cells from antifungals and immune defense (109, 110). Additionally, fitness attributes comprise adaptation to changes in environmental pH, temperature and robust stress response (102, 110). These properties and the genomic flexibility that has been previously described suggest that *C. albicans* can undergo specific genomic changes in order to survive stresses in the human host (111).

It is thought that there is a limit of *C. albicans* cells that the host can tolerate and that immune system must distinguish not threatening from invasive fungal cells in order to keep homeostasis. The first barrier of defence is the epithelial cells and its produced cytokines (81). When the microbial flora is unstable or the immune system is compromised, *C. albicans* cells can penetrate the epithelium and invade bloodstream. Following that, the host innate immunity recognizes and initiates a response (112). The initial response to the infection is determined by the recognition of fungal cell wall components, known as pathogen associated molecular patterns (PAMPs), pattern recognition receptors present on the surface of the nonspecific immune system cells (81, 113). These cells are mainly phagocytes, such as neutrophils, monocytes, macrophages and dendritic cells (114). Macrophages are one of the most important in the response to *C. albicans*. Once macrophages recognize the yeast PAMPs, it is initiated the process of phagocytosis. After being phagocytized, pathogens are exposed to low pH and cytotoxic products, namely reactive oxygen and reactive nitrogen species. After processing the pathogen antigens, macrophages display them on the surface to be recognized by T cells and trigger adaptive immunity (112).

As mentioned above, *C. albicans* has developed mechanisms to survive and escape the host immune system. Thus, elucidating the pathways that regulate stress in *C. albicans* may be important to better understand host-pathogen interactions.

1.5 Working hypothesis and objectives

Previous studies by the host laboratory showed that within the *C. albicans* cytoplasm there are two charged forms of the $\text{tRNA}_{\text{CAG}}^{\text{Ser}}$: the Leu- $\text{tRNA}_{\text{CAG}}^{\text{Ser}}$ (3%) and the Ser- $\text{tRNA}_{\text{CAG}}^{\text{Ser}}$ (97%), and both compete for the CUG codon decoding at the ribosome A-site. However, it has also been demonstrated that the leucine misincorporation at CUG codons varies under different physiological conditions. Therefore, the main objective of this Master's thesis was to identify molecules and pathways involved in the regulation of mistranslation, and ultimately contribute to a better knowledge of this *C. albicans* unique feature. Since the $\text{tRNA}_{\text{CAG}}^{\text{Ser}}$ has been identified as being responsible for CUG ambiguity, we hypothesized that Leu and Ser misincorporation is dependent upon competition between the LeuRS and SerRS for that tRNA_{CAG} . To test this hypothesis, this study has the following specific objectives:

1. Quantification of SerRS and LeuRS expression in different physiological conditions;
2. Correlate LeuRS/SerRS levels with Leu/Ser misincorporation levels;
3. Identification of potential transcription factors that control LeuRS and SerRS expression.

Chapter II – Material and Methods

2.1 Strains and growth conditions

2.1.1 Strains

Escherichia coli strain JM109 (*recA1 SupE44 endA1 hsdR17 gyrA96 relA1 thi* Δ [*Lac-proAB*] *F'*[*traD36 proAB-lacI lacZ* Δ M15]) was used as a host for all DNA manipulations. *Candida albicans* strain SN152 (*arg4* Δ / *arg4* Δ *leu2* Δ / *leu2* Δ *his1* Δ / *his1* Δ *URA3/ura3* Δ ::*imm*⁴³⁴ *IRO1/iro1* Δ ::*imm*⁴³⁴) was used by Homann *et al* (115) to construct the transcription factor (TF) deletion library used in this study. All *C. albicans* strains are described in table 2.1.

Table 2.1 – List of *C. albicans* strains used in this study.

Strain	Description	Deletion
SN152	Control strain from the TF deletion library (ref Homann)	-----
PHO4	Knock-out strain of the bHLH transcription factor of the myc-family	Double
19.2730	Has domain(s) with predicted zinc ion binding activity	unknown
STB5	Putative transcription factor with zinc cluster DNA-binding motif	Double
BAS1	Putative transcription factor with zinc cluster DNA-binding motif	Double
GLN3	GATA transcription factor	Double
RTG1	Sequence-specific DNA binding RNA polymerase II transcription factor activity	Double
CUP2	Putative copper-binding transcription factor	Double
CWT1	Zn2Cys6 transcription factor	Double
ACE2	Transcription factor; similar to <i>S. cerevisiae</i> Ace2 and Swi5	Double
MRR1	Putative Zn(II)2Cys6 transcription factor	Double
ZCF39	Zn(II)2Cys6 transcription factor	Double
HAP31	CCAAT-binding transcription factor; regulates CYC1	Double
HAP43	CCAAT-binding factor-dependent transcription factor	Double
SKN7	Predicted to be a response regulator protein in a phosphorelay signal transduction pathway	Double
HAP2	CCAAT-binding transcription factor	Double
OPI1	Leucine zipper transcription factor	Double
HAP5	Component of CCAAT-binding transcription factor	Double
HAP3	Predicted CCAAT-binding transcription factor that regulates respiration	Double

MSN4	Zinc finger transcription factor	Double
ASH1	GATA-like transcription factor	Double
TEC1	TEA/ATTS transcription factor	Double
19.173	C ₂ H ₂ transcription factor.	Double
CAPI	AP-1 bZIP transcription factor	Double
19.2315	Putative transcription factor with bZIP DNA-binding motif	Double
ZCF5	Zn(II) ₂ Cys ₆ transcription factor	Double
19.1150	GATA-like transcription factor	Double
STP2	Amino-acid-regulated transcription factor	Double

2.1.2 Standard growth conditions

E. coli cells were grown at 37°C in LB (1% peptone from casein, 0.5% yeast extract, 1% sodium chloride; *Formedium*) broth medium or LB 2% agar (*Formedium*). When necessary media was supplemented with ampicillin (75µg/ml; *Sigma-Aldrich*). *C. albicans* strains were grown at 30°C in YPD (2% glucose, 1% yeast extract, 1% peptone; *Formedium*). *C. albicans* strains from the TF deletion library, carrying plasmids pUA563 and pUA564 (table 2.3), were grown in minimal medium (MM) lacking arginine (0.67% yeast nitrogen base without amino acids, 2% glucose, and 0.2% drop-out mixture with all the required amino acids; *Formedium*).

2.2 Primers

Primers used in this study were purchased from IDT[®] (*Integrated DNA Technologies*) and were diluted in milli-Q water to a final stock concentration of 100µM. The primers design was made with *OligoCalc* software to check their properties (<http://www.basic.northwestern.edu/biotools/oligocalc.html>) (116). All primers used in this study are listed in table 2.2.

Table 2.2 – List of primers used in this study.

Primer name	Primer sequence (5' → 3')	T _m (°C)
oUA 1515	URA3 amplification GCCCCTTTTACAGTTGAA	54
oUA 1516	URA3 amplification AGTGACACCATGAGCATT	55

oUA 1554	Confirmation of plasmid integration at RP10 locus	CGTATTCACTTAATCCCACAC	51
oUA 1555	Confirmation of plasmid integration at RP10 locus	CCAATTGGTGATGGTCC	50.5
oUA 1844	ARG4 amplification	TTTGCGGCCGCTAGAACTAGCTTGATG	65.1
oUA 1846	ARG4 amplification	TTTACTAGTAGGTATAGAAATGCTGGT	54.5

2.3 Plasmid construction

2.3.1 Plasmids for LeuRS/SerRS quantification

In order to determine the levels of the SerRS and LeuRS expression, the host laboratory previously constructed a reporter system for indirect quantification of these proteins. Plasmid pUA563 contains the reporter for determination of SerRS expression while pUA564 contains the reporter for determination of LeuRS expression (table 2.3; maps in annex A.1). Both reporters are based on the integrative vector CIp20 (117) and both rely on the yeast enhanced green fluorescent protein (*yEGFP*) expression. In pUA563, the open reading frame of this probe is fused to the promoter of the SerRS (*SES1*), so the fluorescence is proportional to the synthetase expression. In pUA564, the open reading frame of *yEGFP* is fused to the promoter of the LeuRS (*CDC60*). Additionally, both reporter systems have the mCherry gene associated with the actin promoter (*ACT1*) that works as internal control. These plasmids also had arginine (*ARG4* gene) as a selective marker and both were used to transform the TF knockout collection and the control strain listed in table 2.1.

Table 2.3 – Plasmids used in this thesis that were previously constructed by the host laboratory.

Plasmid	Description
pUA 563	Plasmid containing the <i>yEGFP</i> reporter system associated with the <i>SES1</i> promoter, allowing the determination of SerRS expression; used to transform WT and TF KO strains with ARG 4 as selective marker
pUA 564	Plasmid containing the <i>yEGFP</i> reporter system associated with the <i>CD60</i> promoter, allowing the determination of LeuRS expression; used to transform WT and TF KO strains with ARG 4 as selective marker
pUA 553	Plasmid containing the <i>yEGFP</i> reporter system with a WT TTA-leucine at position 201; URA3 gene is present as a selective marker

pUA 554	Plasmid containing the <i>yEGFP</i> reporter system with a CTG ambiguous codon at position 201; URA3 gene is present as a selective marker
pUA 555	Plasmid containing the <i>yEGFP</i> reporter system with a TCT-serine codon at position 201; URA3 gene is present as a selective marker

2.3.2 Plasmids for quantification of Leu misincorporation

To quantify Leu misincorporation levels in knock-out *C. albicans* strains, we used a gain of function fluorescent reporter system also based on *yEGFP*, described by Bezerra *et al.* (76). This reporter system consists of three different versions of the *yEGFP* gene assembled in three different plasmids previously built by the host laboratory (pUA 553, pUA 554, pUA 555). The plasmid pUA 553 contains the *yEGFP* gene with WT TTA-leucine at position 201 and functions as the positive control; plasmid pUA 554 has the codon at position 201 mutated for the ambiguous CTG codon, thus producing stable GFP only when Leu is incorporated; plasmid pUA 555 encodes the TCT-serine codon at position 201 of the *yEGFP* gene, thus producing inactive GFP (negative control). All of these three plasmids have URA3 as auxotrophic marker (table 2.3; maps in annex A.2).

However, since the strains from the TF deletion library used in this study were generated with URA3 selective marker, we had to replace the URA3 gene from plasmids pUA 553, pUA 554 and pUA 555 with another selective marker. We substituted the URA3 gene with ARG4 to transform strains from the TF knockout collection. First, the three plasmids were digested with *SpeI* and *NotI* (*Fermentas*) to remove the URA3 selective marker following the recommendations of the manufacturer. Then, ARG4 gene was amplified from plasmid pUA 564 using the forward oUA1844 and the reverse oUA1846 primers. After this, the ARG4 insert was digested with *SpeI* and *NotI* enzymes. Prior to the insertion of the selective marker into the vectors, plasmids were treated with SAP enzyme (Shrimp Alkaline Phosphatase, *Fermentas*) to prevent re-ligation of the plasmids ends. Finally, ARG4 gene was assembled between *SpeI* and *NotI* restriction sites in the plasmids pUA 553, pUA 554 and pUA 555, resulting in the plasmids pUA 567, pUA 568 and pUA 569, respectively (table 2.4; maps in annex A.3).

Table 2.4 – Constructed plasmids during this study.

Plasmid	Description
pUA 567	Plasmid derived from pUA 553 containing an ARG4 gene as selective marker inserted between <i>SpeI</i> and <i>NotI</i> restriction sites
pUA 568	Plasmid derived from pUA 554 containing an ARG4 gene as selective marker inserted between <i>SpeI</i> and <i>NotI</i> restriction sites
pUA 569	Plasmid derived from pUA 555 containing an ARG4 gene as selective marker inserted between <i>SpeI</i> and <i>NotI</i> restriction sites

2.4 Expression of host transformation

2.4.1 Preparation of *E. coli* competent cells

For the preparation of competent *E. coli* cells, we performed the TFB method (118) using cells from strain JM109. Initially, 200µl of cells from an overnight culture were inoculated in 5ml of LB medium and incubated at 37°C with 180rpm until we obtained an OD₆₀₀ of 0.3. After this, 4ml of the previous culture was inoculated in 100ml of LB medium and incubated at 37°C with 180rpm until the OD₆₀₀ was 0.3. Then, cells were collected in two 50ml falcons and kept on ice for 5 minutes. The two falcons were then centrifuged at 2500rpm for 5 minutes at 4°C. Supernatant was discarded and each pellet was resuspended in 20ml of cold TFB I solution (0.03mM potassium acetate, 0.08mM RbCl₂, 0.013mM CaCl₂, 0.08mM MnCl₂, 15.4% glycerol, pH 5.8). Tubes were then centrifuged at 2500rpm for 5min. at 4°C. Supernatant was again discarded and each pellet was resuspended in 2.5ml of cold TFB II solution (0.01mM MOPS Na, 0.01mM CaCl₂, 0.008mM RbCl₂, 13.4% glycerol, pH 6.5). Finally, cells were cooled on ice for 5 min. and then distributed in 200µl aliquots and frozen at -80°C.

2.4.2 Transformation of *E. coli*

Transformation of *E. coli* cells was performed following the Sambrook's SOC method (118). First, 20µl of ligation reactions using 1:0 to 1:5 of vector to insert ratios were prepared. In addition, the reaction contained 1µl of DNA Ligase (5U/µl) (*Thermo Scientific*), 2µl of 10x DNA Ligase Buffer (*Thermo Scientific*) and milli-Q water to complete the volume. Tubes were then incubated at 20°C for 4h, followed by an incubation

of 10 minutes at 65°C to inactivate the enzyme. Next, ligation reactions were added to 200µl aliquots of *E. coli* JM109 competent cells. Then, the reaction was incubated on ice during 30 minutes, followed by a heat shock at 42°C for 90 seconds and iced again for 2 minutes. After that, 800µl of SOC medium (for preparation of 100ml at pH 7.2g of tryptone, 0.5g of yeast extract and 0.05g of NaCl were weighted, and 1ml of KCl 250mM and 20ml of glucose 1M were added) was added to each reaction and tubes were incubated for 1h at 37°C with agitation of 180rpm. Next, tubes were centrifuged during 1 minute at 2500rpm. The resulting supernatant was discarded and the pellet was homogenized and plated into LB plates supplemented with 75µg/ml ampicillin (*Sigma-Aldrich*). Plates were incubated overnight at 37°C.

2.4.3 Plasmidic DNA purification from *E.coli*

For the plasmidic DNA purification and extraction from *E.coli*, we used the “NZYMiniprep” kit (*Nzytech*,). For this, *E. coli* cells were grown overnight in 5ml of liquid LB medium (*Formedium*) with ampicillin (75µg/ml; *Sigma-Aldrich*). The procedure was then performed according to the manufacturer’s instructions. After the purification, NanoDrop was used to quantify the yield of the purified DNA. Purified plasmids were stored at -20°C until further use.

2.5 Manipulation of *C. albicans* strains

2.5.1 Transformation of *C. albicans*

Prior to the transformation protocol, plasmids were linearized with *StuI* (*Thermo Technologies*) following the recommended reaction conditions, during 4h at 37°C.

Transformation of *C. albicans* was then carried out using an improved lithium method with minor modifications (119). *C. albicans* cells were grown overnight in falcon tubes with 10ml liquid YPD medium. These cultures were then diluted into fresh YPD medium to an OD₆₀₀ of 0.3 and grown for additional 4 hours at 30°C, 180rpm shaking until cell cultures reached an optical density of 1-1.2 at 600nm. Cells were then centrifuged at 4000rpm for 5 min, supernatants were discarded and pellets were resuspended in 150µl of LiAc-solution (10% of LiAc 1M, 10% TE buffer 10x, 80% of milliQ-water). In an

Eppendorf tube, 200µl of the cell suspension, 5µg of the plasmid DNA of interest and 50µl of carrier single strand DNA (2mg/ml) were mixed. To the transformation mixture 600µl of PEG/LiAc-solution (50% (w/v) polyethyleneglycol, 50% LiAc-solution) was added and briefly vortexed. Afterwards, the transformation tubes were incubated overnight at 30°C, followed by a heat shock of 15 min. at 44°C and another on ice for 2 min. Cells were pelleted at 4000rpm for 5 min and resuspended in appropriate minimal medium. Aliquots of 100µl of cell suspensions were plated onto selective medium plates and incubated at 30°C during 3-4 days.

2.5.2 DNA extraction from *C. albicans*

We performed DNA extraction from the selected yeast transformants, using an adaptation of the lyticase method developed by Hoffman and Winston (120). First, *C. albicans* cells were grown overnight in 5ml of appropriate minimal medium at 30°C. Cells were then centrifuged at 4000rpm for 5min and the supernatant discarded. Then, cells were resuspended in 500µl of Solution I (sorbitol 1M; EDTA-Na₂ 20mM; pH 7) and transferred to an Eppendorf tube. This step was followed by the addition of 4µl (10 mg/ml) of the lyticase enzyme (*Sigma-Aldrich*) and incubation at 37°C for 60 min. Tubes were then centrifuged for 3 min at 13000rpm and the supernatant discarded. 500µl of Solution II (Tris-HCl 50mM; EDTA-Na₂ 20mM; pH7) and 50µl of SDS were added and samples were vortexed. This was followed by incubation at 65°C for 30 min. 200µl of Potassium Acetate 5M were added and samples were incubated on ice for 60 min. The suspensions were centrifuged for 5 min. at 13000rpm and the supernatant (600µl) transferred to another Eppendorf tube. Then, 1.5V of cold ethanol 100% and 0.1V of NaCl 5M were added to each tube followed by an incubation at -30°C for 2h. After this, tubes were centrifuged for 5 min. and the supernatant was discarded. Pellets were air dried and resuspended milli-Q water. Finally, the DNA concentration was determined using *NanoDrop*[®].

2.5.3 Integration confirmation

Integration of the different plasmids was targeted to the RP10 locus of *C. albicans* strains. To confirm the correct plasmid integration in *E. coli*, transformants were tested by colony PCR. Briefly, individual colonies were picked from selective media and homogenized in 5µl of milli-Q water. Next, the suspension was submitted to 95°C during 5 min. for cell

lysis. Then, amplification of the RPS10 locus fragment was performed according to current polymerase chain reaction (PCR) protocols. The following reagents were added to the suspension: 1X Dream Taq Buffer, 0.2mM dNTP mix (*Fermentas*), 0.15mM of each specific primer (table 2.3), 0.375U of Dream Taq polymerase and milli-Q water to a final volume of 15µl. The protocol for the PCR reaction was performed in *My CyclerTM* thermal-cycler (*BIO RAD*) and consisted on a denaturation step at 95°C for 5 min, 30 cycles of denaturation at 95°C for 30s, annealing temperature of 50°C for 30s, extension at 72°C for 1 min and a final extension step at 72°C for 5 min. PCR products were then checked by gel electrophoresis, running the samples on a 1 % agarose gel.

2.6 Choice of transcription factors mutants to analyse

The identification of putative transcription factors regulators of mistranslation was previously done by the host group through a bioinformatics analysis. This approach allowed the identification of TF binding sites present in the 1000 base pairs located upstream of the AUG start codon of LeuRS (CD60) and SerRS (SES1) open reading frames. For this, the two DNA sequences were submitted to the online tool motif finder “Find TF Binding Site (s)” from the YEASTRACT database (<http://www.yeasttract.com>) (121). This tool retrieved a list of motif sequences and the corresponding transcription factors that bind to them. Since this is a yeast database, it was performed a search in the Candida Genome Database website (<http://www.candidagenome.org>) (122) for orthologs of the identified yeast TFs in *Candida albicans*. A list of 25 TFs with probable binding motifs in both promoters was obtained. Additionally, 2 TFs with no probable binding motifs in SES1 and CD60 promoters were chosen as controls. This bioinformatics analysis allowed the reduction of the screening from the initial 166 mutants in the deletion library to the 25 tested in this study (table 2.5).

Table 2.5 – List of the *C. albicans* TFs selected for this study (selected from Homann *et al.* (115)).

TF	ORF	Gene Name (<i>C. albicans</i>)	<i>S. cerevisiae</i> ortholog	Description	SerRS	LeuRS
4	19.1253	<i>PHO4</i>	<i>PHO4</i>	bHLH transcription factor of the myc-family	X	
8	19.2730		<i>GIS1</i>	Has domain(s) with predicted zinc ion binding activity		X

13	19.3308	<i>STB5</i>	<i>STB5</i>	Putative transcription factor with zinc cluster DNA-binding motif	X	X
16	19.3809	<i>BAS1</i>	<i>BAS1</i>	Putative transcription factor with zinc cluster DNA-binding motif	X	
18	19.3912	<i>GLN3</i>	<i>GLN3</i>	GATA transcription factor	X	
34	19.4722	<i>RTG1</i>	<i>RTG1</i>	Sequence-specific DNA binding RNA polymerase II transcription factor activity	X	
39	19.5001	<i>CUP2</i>	<i>CUP2</i>	Putative copper-binding transcription factor	X	
50	19.5849	<i>CWT1</i>	<i>RDS2</i>	Zn2Cys6 transcription factor	---	---
59	19.6124	<i>ACE2</i>	<i>ACE2</i>	Transcription factor; similar to <i>S. cerevisiae</i> Ace2 and Swi5	X	
69	19.7372	<i>MRR1</i>	<i>HAP1</i>	Putative Zn(II)2Cys6 transcription factor	X	X
73	19.7583	<i>ZCF39</i>		Zn(II)2Cys6 transcription factor	X	X
79	19.517	<i>HAP31</i>	<i>HAP3</i>	CCAAT-binding transcription factor; regulates CYC1	X	
80	19.681	<i>HAP43</i>	<i>HAP4</i>	CCAAT-binding factor-dependent transcription factor	X	
83	19.971	<i>SKN7</i>	<i>SKN7</i>	Predicted to be a response regulator protein in a phosphorelay signal transduction pathway	X	X
87	19.1228	<i>HAP2</i>	<i>HAP2</i>	CCAAT-binding transcription factor	X	
90	19.1543	<i>OPI1</i>	<i>OPI1</i>	Leucine zipper transcription factor	---	---
93	19.1973	<i>HAP5</i>	<i>HAP5</i>	Component of CCAAT-binding transcription factor	X	
108	19.4647	<i>HAP3</i>	<i>HAP3</i>	Predicted CCAAT-binding transcription factor that regulates respiration	X	
109	19.4752	<i>MSN4</i>	<i>MSN4</i>	Zinc finger transcription factor		X
112	19.5343	<i>ASH1</i>	<i>ASH1</i>	GATA-like transcription factor	X	X
115	19.5908	<i>TEC1</i>	<i>TEC1</i>	TEA/ATTS transcription factor		X
136	19.173		<i>AZF1</i>	C ₂ H ₂ transcription factor.	X	
140	19.1623	<i>CAP1</i>	<i>YAP1</i>	AP-1 bZIP transcription factor	X	X
142	19.2315		<i>RTG3</i>	Putative transcription factor with bZIP DNA-binding motif	X	
149	19.2315	<i>ZCF5</i>	<i>HAP1</i>	Zn(II)2Cys6 transcription factor	X	X
157	19.1150			GATA-like transcription factor	X	X
162	19.4961	<i>STP2</i>	<i>STP2</i>	Amino-acid-regulated transcription factor	X	X

2.7 Stress conditions

To determine SerRS and LeuRS expression, as well as leucine incorporation at CUG sites, 3 clones from each selected transformed knockout and control strain were cultured in several physiological conditions. Strains were grown overnight in a specific liquid MM until exponential phase, as described in table 2.6. Samples were then spotted onto microscope slides and analysed by epifluorescence microscopy.

Table 2.6- Conditions used in the stress experiments.

Assay	Stress compound	Growth Temperature
Control	pH 7	30°C
Temperature	30°C	30°C
	37°C	37°C
pH	pH 4	30°C
	pH 6	30°C
	pH 7	30°C
Osmotic stress	Sorbitol 0.2 M	30°C
	Sorbitol 1 M	30°C

2.8 Epifluorescence Microscopy

To determine LeuRS and SerRS expression, yEGFP and mCherry expression of *C. albicans* cells was observed using epifluorescence microscopy. Fluorescence was detected using a Zeiss MC80 Axioplan 2 light microscope equipped with filter sets 38 HE GFP and 63 HE mRFP. Photographs were taken using an AxionCam HRc camera, and images were then analysed with ImageJ software. LeuRS expression was calculated in at least 300 cells carrying plasmid pUA564. For that, the mean intensity of the GFP (controlled by LeuRS promoter) was divided by the mean mCherry intensity. SerRS expression was obtained from cells carrying plasmid pUA563. The mean intensity of the GFP (controlled by SerRS promoter) was divided by the mean mCherry intensity, which allowed indirect quantification of the synthetase.

To quantify leucine incorporation at CUG sites, GFP fluorescence intensity was divided by the mCherry intensity obtained in the same cell. This GFP quantification was

performed for at least 300 cells carrying each version of the reporter: pUA567 (Leu-UUA201 - positive control), pUA568 (Ser/Leu-CUG201 - reporter) and pUA569 (Ser-UCU201 - negative control). Then, values obtained in the negative strain were subtracted to the values obtained in the reporter and positive strains. After that, values resulting from the reporter strain are divided by the ones obtained in the positive strain (Equation 1).

$$\text{Leucine incorporation at CUG codons} = \frac{[\text{GFP}_{\text{CUG (Reporter)}}/\text{mCherry}_{(\text{Reporter})}] - [\text{GFP}_{\text{UCU (Negative)}}/\text{mCherry}_{(\text{Negative})}]}{[\text{GFP}_{\text{UUA (Positive)}}/\text{mCherry}_{(\text{Positive})}] - [\text{GFP}_{\text{UCU (Negative)}}/\text{mCherry}_{(\text{Negative})}]}$$

Equation 1: leucine incorporation at CUG sites

2.9 Statistical Analysis

All statistical analyses were performed using the GraphPad Prism version 5.0 software for windows. Data represent the mean (\pm standard deviation - s.d.) of three clones tested for each strain. Statistical comparisons between the deletion strains and the control WT strain were carried out using one-way ANOVA followed by a Dunnett comparison test with 95% interval with the control (**p<0.001, *p<0.01, *p<0.05). Heat maps were generated using the MeV software version 4.9 for windows.

Chapter III - Results

3.1 SerRS and LeuRS expression and CUG ambiguous decoding

3.1.1 Overview

Candida albicans has the remarkable capacity to adapt and cope with the immune system, including tolerance to extremes of pH, oxidative and nitrosative stress and stress associated with clinical therapies. In addition, this human pathogen decodes ambiguously the leucine CUG codons as serine (97%) and leucine (3%) using a hybrid tRNA_{CAG}^{Ser} (123). This serine tRNA is aminoacylated by both SerRS and LeuRS, originating the incorporation of Ser and Leu at CUG sites, respectively (123). Therefore, there is a competition between SerRS and LeuRS for this hybrid tRNA, in which LeuRS (in normal conditions) seems to be the weakest competitor, as CUG is mainly translated as serine (124). Interestingly, Leu misincorporation is flexible and increases in response to stressful conditions (74), suggesting that LeuRS activity increases under stress. It was also demonstrated that both SerRS and LeuRS have a single CUG codon in highly conserved regions (99). Consequently fluctuations in the insertion of serine or leucine at CUG positions could affect their aminoacylation activities (25, 99).

To test our hypothesis, that Leu and Ser misincorporation at CUG codons is dependent upon competition between the LeuRS and SerRS for the tRNA_{CAG}^{Ser}, we started by demonstrating that SerRS and LeuRS expression changes in conditions where fluctuation in CUG ambiguity was already proved. To do so, we used a fluorescent reporter system to indirectly quantify SerRS and LeuRS expression and established the existence of a correlation between increase in ambiguity (increase in Leu incorporation at CUG sites) and increase in LeuRS/SerRS ratio.

3.1.2 SerRS and LeuRS expression in different physiological conditions

In order to quantify SerRS and LeuRS expression we used two reporter systems based on expression of yeast-enhanced GFP (125), already available in the host laboratory. One of the reporter systems harbours the GFP open reading frame fused with the promoter of SerRS (SES1) gene and the other has GFP fused with the LeuRS (CD60) promoter. In

both cases, GFP fluorescence will be proportional to the activity of the promoter and, therefore, will indirectly measure expression of the gene that is associated with. Additionally, both reporters contain the mCherry gene fused to the actin promoter (*act1*) that was used to normalize protein expression levels (Figure 3.1).

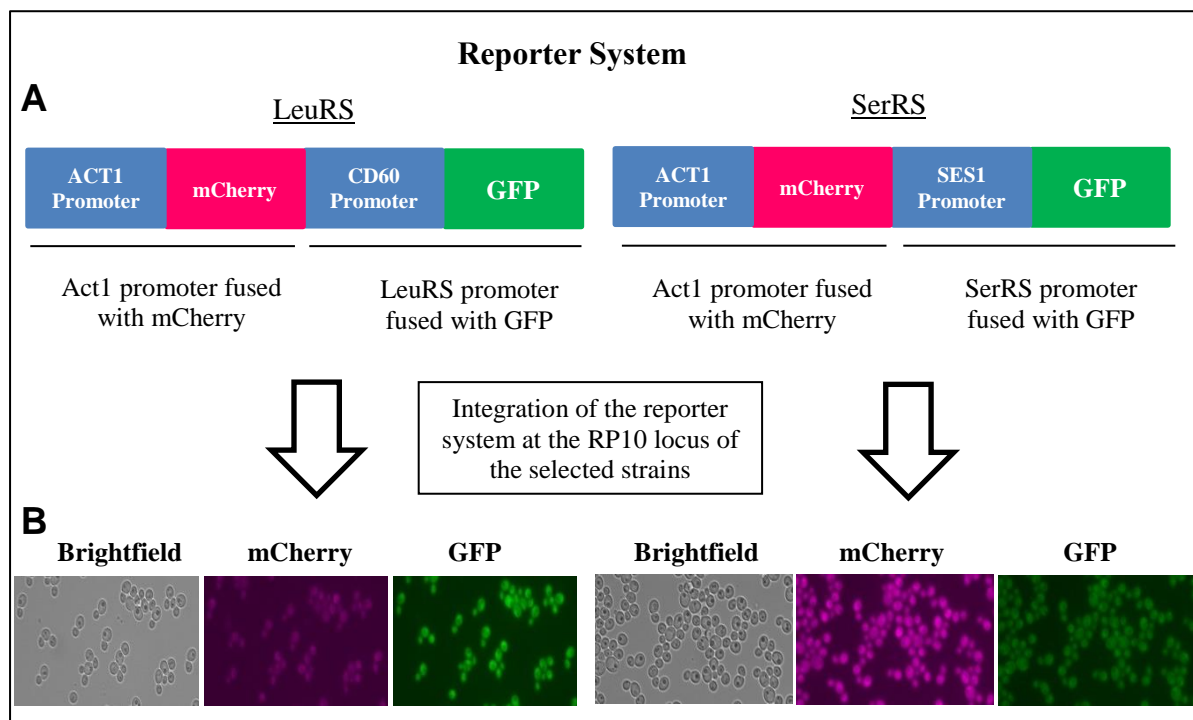


Figure 3.1 – Schematic representation of the reporter systems used to quantify LeuRS and SerRS expression.

A) These reporter systems are based on *yEGFP*, where the open reading frame of GFP was fused to the promoters of the SerRS and LeuRS genes (*SES1* and *CDC60*) and its fluorescence is proportional to their activity. A mCherry fluorophore associated with actin promoter (*ACT1*) was used as internal control. LeuRS and SerRS reporters were assembled into pUA564 and pUA563 respectively. **B)** Fluorescence and brightfield images of *C. albicans* cells obtained by epifluorescence microscopy (magnification: 630x) with the appropriate filters, using an AxionCam HRc camera (Zeiss).

These reporter systems were integrated in the RP10 locus of *C. albicans* WT strain SN152 (115) and cells were grown under different physiological conditions: temperature, pH and osmotic stress. These conditions are particularly relevant because *C. albicans* is able to colonize different host sites, where it is exposed to temperature and pH fluctuations. Also, the pathogen is exposed to the candidacidal mechanisms of phagocytes, which includes hydrolytic activity promoted by acidic pH (114). In this study, we measured GFP fluorescence of cells grown at pH 4, pH 6 and pH 7 and temperature 30°C and 37°C, considering 30°C and pH 7 as control growth conditions. A LeuRS/SerRS expression ratio

was established for all conditions. Cells grown at 37°C, pH4 and pH6 presented an increase of LeuRS/SerRS ratio, compared to the cells grown at control conditions (Figure 3.2).

Another important characteristic of *C. albicans* is the resistance to osmolarity changes. Exposure to sorbitol induces increase in cell volume and cell wall alteration in *C. albicans* (126). The selected sorbitol concentrations were 0.2M and 1M and, similarly to the results obtained for the pH and thermal stress, cells grown in presence of sorbitol had higher LeuRS/SerRS ratio expression levels than the control (0M) (Figure 3.2).

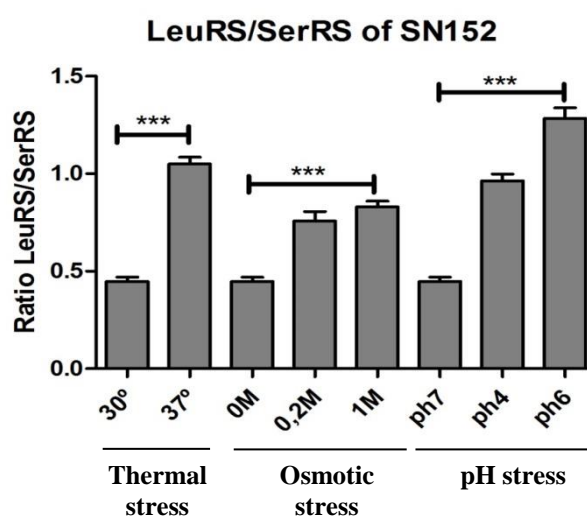


Figure 3.2 – LeuRS/SerRS expression ratio in *C. albicans* WT strain SN152 grown at different physiological conditions.

Values are presented as mean \pm s.d. of LeuRS/SerRS ratio expression at control conditions (30°C, 0M and pH7) and stress conditions (37°C, Sorbitol 0.2M and 1M, pH 6) retrieved from at least 300 cells of 3 different clones (**p<0.001 vs control).

These results raised an important question: the increase of LeuRS/SerRS expression ratio under stress conditions is caused by an increase of LeuRS or decrease of SerRS expression, or both? Therefore, an analysis of raw data concerning GFF intensity of both reporters allowed us to investigate the LeuRS and SerRS expression separately. Levels of LeuRS expression were higher in cells under all stress conditions tested, compared to control (figure 3.3A). In contrast, with the exception of the condition 0.2M sorbitol, SerRS expression was almost constant in all conditions (Figure 3.3 B). Thus, the increase of LeuRS/SerRS in response to stress seems to be mainly due to the increase of LeuRS expression.

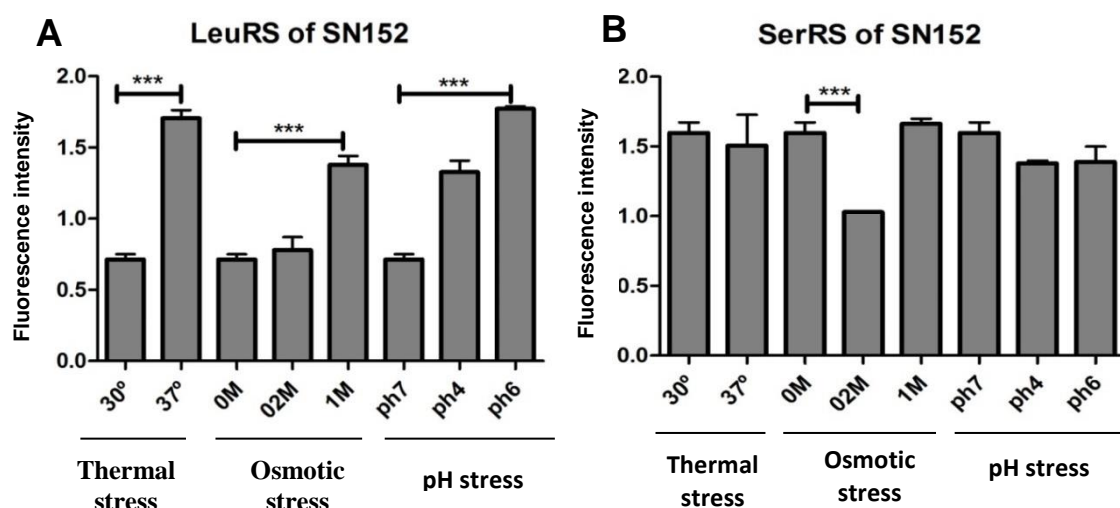


Figure 3.3 – LeuRS and SerRS expression in *C. albicans* WT strain SN152 grown at different physiological conditions.

Data represents mean \pm s.d. of LeuRS (A) SerRS (B) at control conditions (30°C, 0M and pH7) and stress conditions (37°C, Sorbitol 0.2M and 1M, pH 6) of at least 300 cells of 3 different clones (** $p < 0.001$ vs control). Fluorescence intensity (arbitrary units) was measured using a Zeiss epifluorescence microscope and GFP intensity was normalized for mCherry intensity at each cell.

3.1.3 Comparison of LeuRS/SerRS levels with Leu/Ser misincorporation levels

After the demonstration that LeuRS/SerRS levels increase under thermal, osmotic and pH stress, our second objective was to verify if it correlates with the CUG ambiguity previously measured in same conditions. To do so, we gathered the data of the quantification of Leu incorporation at CUG sites in cells grown at 30°C and 37°C from the work of Gomes *et al.* (74). Leu levels concerning the osmotic stress (0.2M, 1M) and pH (4, 6 and 7) were obtained from unpublished work from the host laboratory (João Simões PhD thesis). Then, these Leu misincorporation levels were qualitatively compared with the ratio of LeuRS and SerRS expression that were obtained in this study. A positive correlation can be observed between the increase in LeuRS/SerRS and Leu misincorporation at CUG codons in all the stress conditions tested (Figure 3.4). This was an important finding, because it indicates that Leu and Ser misincorporation at CUG sites may indeed be dependent on LeuRS and SerRS expression. Therefore, based on our hypothesis, the next objective was the identification of putative regulators controlling LeuRS and SerRS expression.

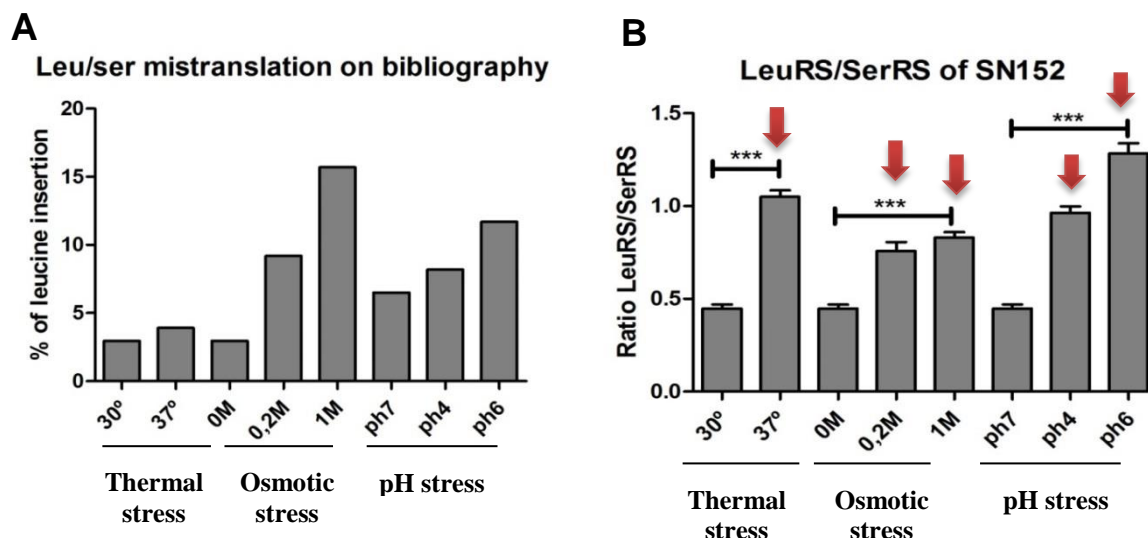


Figure 3.4 – Comparison between Leu and Ser mistranslation at CUG sites and LeuRS/SerRS expression. **A)** Data represents Leu misincorporation at CUG sites in different physiological conditions previously reported by Gomes *et. al.* (30°C and 37°C) and João Simões (Sorbitol 0.2M and 1M, pH 4, 6 and 7). No statistics is shown as the graph represents indicative data retrieved from different bibliography and unpublished data from the host laboratory. **(B)** Ratio of the LeuRS and SerRS expression at control conditions (30°C, 0M and pH7) and stress conditions (37°C, Sorbitol 0.2M and 1M, pH 6) of at least 300 cells of 3 different clones. Conditions where a positive correlation in ambiguity (increase in Leu incorporation) is accompanied by increase in LeuRS/SerRS ratio are marked with red arrows. (***) $p < 0.001$ vs control).

3.2 Identification of potential transcription factors that control LeuRS and SerRS expression

3.2.1 Overview

An adequate adaptation to stress is essential for cell surviving in severe environments. Induction of gene expression alterations is one important mechanism of stress response (127). Transcription factors (TFs) are key players in this process because they are intermediaries between the stress signals and the metabolic reprogramming needed to respond to changes in the environmental conditions (128). In recent years, many studies have been exploring the role of TFs in *C. albicans* response to stress and virulence traits (128-130).

Stress tolerance in mistranslating *C. albicans* was already widely demonstrated. The random insertion of Leu and Ser at CUG sites generates a combination of mutant proteins important to tackle environmental changes (74, 76, 131). However, little is known

about how *C. albicans* regulates CUG ambiguity. The validation of our hypothesis, that the incorporation of Ser and Leu at CUG codons is dependent on the competition between the SerRS and LeuRS for the tRNA_{CAG}^{Ser}, prompted us to study the control of CUG ambiguity through the identification of regulators of SerRS and LeuRS expression. The regulation of eukaryotic aaRSs genes is still not well studied. Nevertheless, in line with the observations in prokaryotes, it is proposed that both transcriptional and post-transcriptional mechanisms may be involved (132).

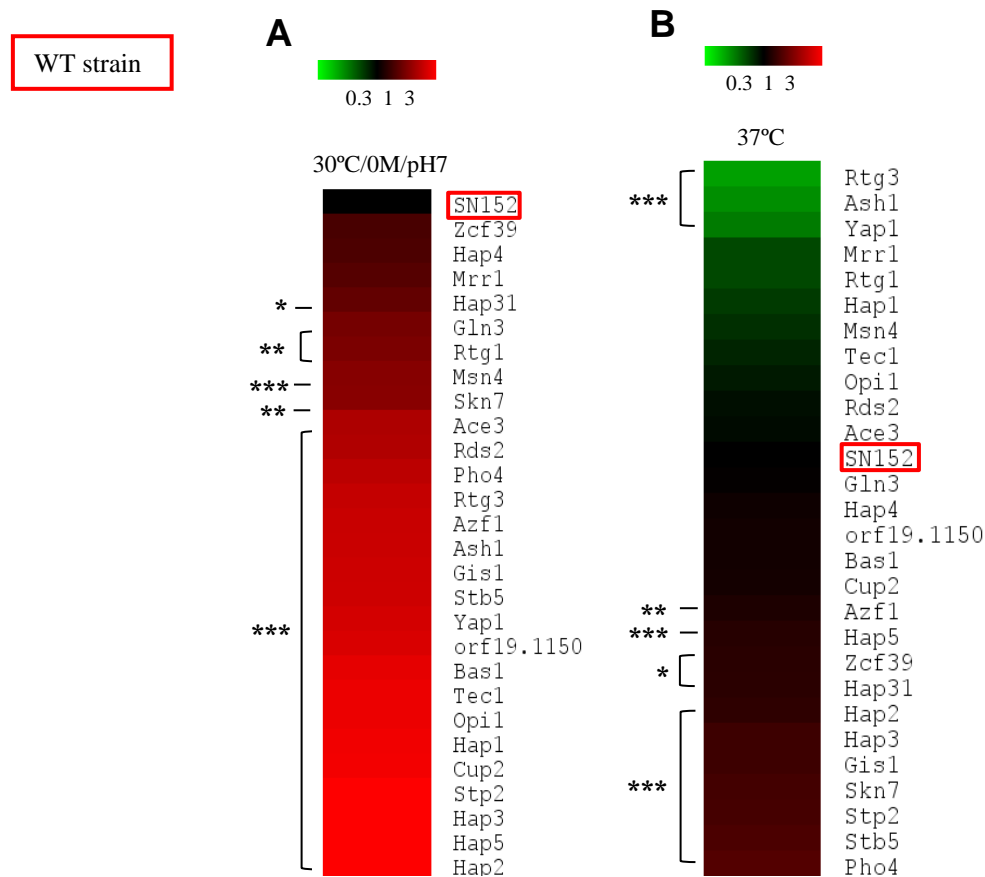
Taken all these findings into account, the study of TFs is a good starting point for the study of regulation of mistranslation. To accomplish this, we screened a *C. albicans* TF mutant collection, by transforming the strains with the fluorescent reporters to monitor the expression of SerRS and LeuRS under the physiological conditions tested in the previous section.

3.2.2 Screening of TF knock-out collection

In order to identify potential regulators of LeuRS and SerRS expression, we took advantage of a pre-existent collection of 166 genetically matched strains of *C. albicans*, each of which has been deleted for a specific transcriptional regulator (115). As mentioned in the section 2.6, a bioinformatics analysis was previously done in the host laboratory. It was performed a search for TF binding sites present in the 1000 base pairs located upstream of the AUG start codon of LeuRS (CD60) and SerRS (SES1) open reading frames. To do so, the two DNA sequences were submitted to the online tool motif finder “Find TF Binding Site (s)” from the YEASTRACT database (<http://www.yeasttract.com>) (121). Then, it was performed a search in the Candida Genome Database website (<http://www.candidagenome.org>) (122) for orthologs of the identified yeast TFs in *Candida albicans*. This analysis allowed us to narrow the screening from a collection of 166 TF mutants to a group of 27 knockout strains. The list of the strains tested in this study is depicted in table 2.1 (section 2.1).

After the transformation of these 27 strains with plasmids containing the fluorescent systems, 3 clones of each were grown under control the conditions (30°C, pH 7) and stress condition pH (pH 6) and osmotic stress (sorbitol 0.2M). At control conditions, most of the mutant strains showed a superior LeuRS/SerRS expression than WT strain

(Figure 3.5 A). Similarly, the majority of the mutant strains exposed to 0.2M concentration of sorbitol expressed a higher ratio of LeuRS/SerRS expression than the WT strain (Figure 3.5 C). Contrarily, when grown in media with pH 6 (Figure 3.5 D) almost all deletion strains exhibit a lower LeuRS/SerRS expression than the WT. Strains grown at 37°C presented a more heterogeneous behaviour. There were mutant strains that showed a lower LeuRS/SerRS expression than the WT strain, while others presented higher expression of LeuRS/SerRS expression (Figure 3.5 B).



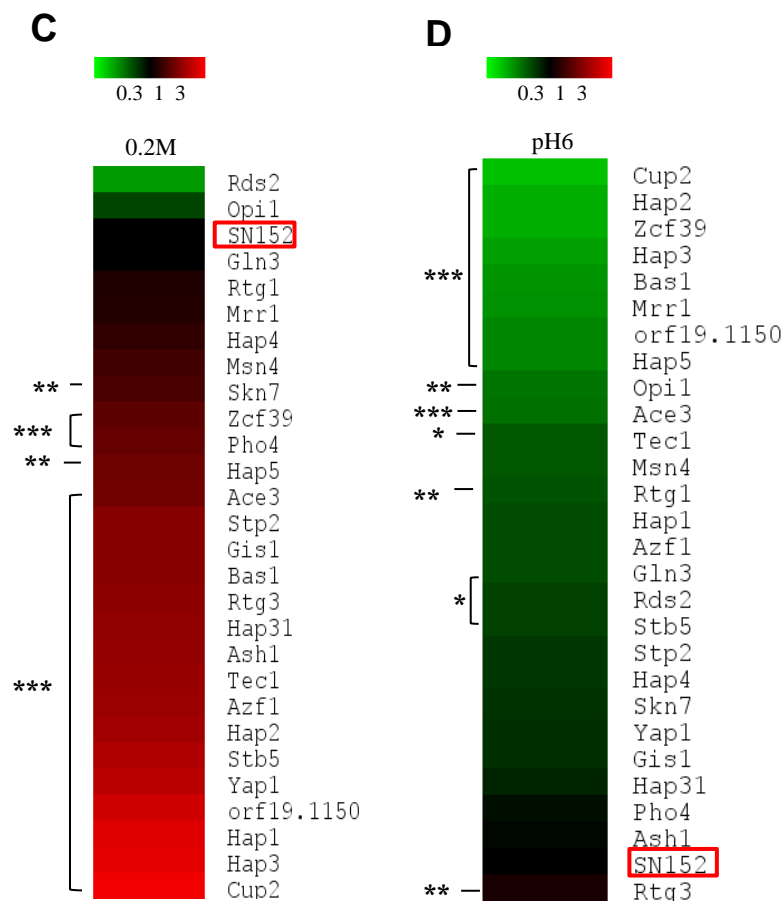


Figure 3.5 – Heat maps of the LeuRS/SerRS ratio expression in *C. albicans* TF deletion strains normalized to WT strain in several physiological conditions.

Data from the TF KO collection and WT strain SN152 at control conditions (A), at 37°C (B), sorbitol 0.2M (C), and pH6 (D). Ratio of the LeuRS and SerRS expression obtained from 300 cells of 3 clones from each KO strain were normalized to the WT strain values. Strains *HAP3*, *STB5* and *HAP2* presented high values of LeuRS/SerRS in almost all conditions. In contrast, strains *ASH1* showed a low ratio LeuRS/SerRS at 37°C and pH6 and *RTG3* in pH6. Red means higher LeuRS/SerRS expression than the control while green represents lower LeuRS/SerRS expression. (***) $p < 0.001$, ** $p < 0.01$, * $p < 0.05$ vs control).

Deletion strains lacking the TFs *HAP3*, *STB5* and *HAP2* showed high values of LeuRS/SerRS in all conditions, suggesting that these proteins could act as repressors of LeuRS expression. In contrast, mutant strain *ASH1* showed a low ratio LeuRS/SerRS at 37°C and pH 6 and *RTG3* at pH6, indicating that these genes can encode positive regulators of LeuRS expression under these environmental conditions. We reasoned that these are the most interesting mutants, because they exhibit statistically different LeuRS/SerRS expression ratios from the WT strain in the same pattern at least in two

conditions (Figure 3.5). The 2 TFs chosen as controls, with no probable binding motifs in SES1 and CD60 promoters (*OPI1* and *RDS2*), showed identical wild type LeuRS/SerRS expression at 37°C and osmotic stress, but showed statistically differences at 30°C. Although these 2 transcription factors were selected as controls from the bioinformatics analysis, we can not exclude the hypothesis that these factors may recognize motifs in the promoters of SerRS and LeuRS in *C. albicans* as the analysis was performed with tools developed for *S. cerevisiae*. Also, we can not exclude the possibility that these deletions are affecting other pathways with implications in the regulation of mistranslation beyond the transcriptional regulation of these aaRSs.

3.2.3 Comparison of Leu misincorporation and LeuRS/SerRS ratio expression in TF knock-out collection

In order to validate the regulatory potential of the identified genes, Leu misincorporation levels in these deletion strains must have a positive correlation with LeuRS/SerRS expression. The rationale behind this validation is the following: if a mutant has increased LeuRS/SerRS and this is due to an increase in LeuRS and maintenance of SerRS (as seen in section 3.1.2), then that TF is a negative regulator (repressor) of LeuRS. Therefore, the deletion strain in question must have constitutive increased Leu misincorporation at CUG sites. In the opposite scenario: if a deletion strain has a decreased LeuRS/SerRS expression and this is caused by a decrease in LeuRS and maintenance of SerRS (as seen in section 3.1.2), then that TF is a positive regulator (enhancer) of LeuRS. Thus, in this case the mutant strain must present decreased levels of Leu misincorporation at CUG codons.

To perform the quantification of Leu incorporation at CUG codons in the deletion strains, we used an established reporter system consisting of a gain of function fluorescent reporter system described by Bezerra *et al.* (76). Briefly, this reporter system includes three versions. In the positive control, at the position 201 of the GFP gene it is present the Leu-TTA codon, essential to produce stable GFP and to obtain fluorescence. The reporter was mutated for the ambiguous CUG codon at position 201 so the obtained fluorescence will be proportional to the incorporation of Leu at this position. Finally, the negative control comprises the mutation of the Leu codon at position 201 to the Ser TCT codon, originating

inactive GFP. In addition, in each version the mCherry fluorophore is associated with the actin promoter (ACT1) and functions as internal control (Figure 3.6).

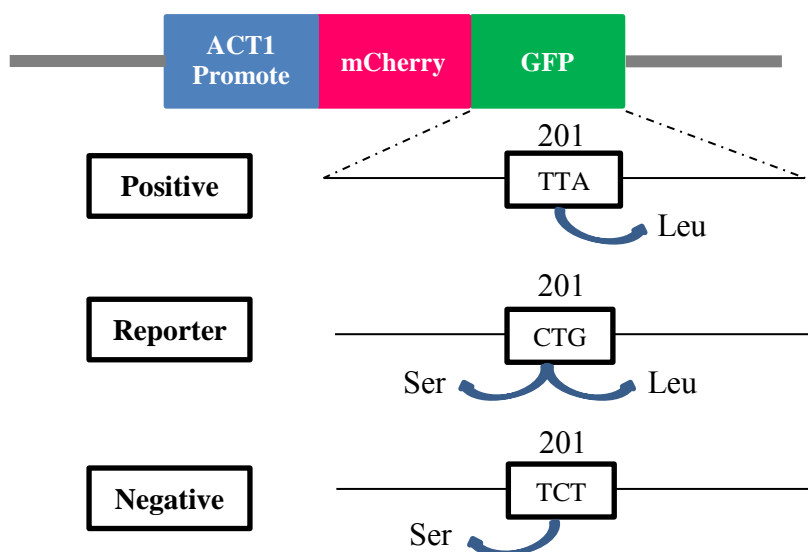


Figure 3.6 – Schematic representation of the reporter system used to quantify Leu incorporation at CUG codons.

This reporter system is based on the yeast enhanced green fluorescent protein (yEGFP) described by Bezerra *et al.* (68). The Leu-TTA codon at position 201 was mutated to the ambiguous CTG codon (reporter) and to the Ser TCT codon (negative control). The positive version was assembled into pU567, the reporter was integrated in the pUA568 and the negative control was assembled in the plasmid pUA569.

The three plasmids containing the GFP constructions were integrated in the RP10 locus of the WT and of the 5 candidate TF knockout strains. After that, 3 clones expressing each version of the fluorescent protein were submitted to the same growth conditions used to quantify SerRS and LeuRS expression.

When *C. albicans* WT (SN152) cells were cultivated at 30°C (control condition) Leu incorporation at CUG-201 was $3 \pm 0.70\%$ (Figure 3.7B). As expected, globally the selected KO strains showed higher levels of Leu incorporation in the control conditions (Figure 3.7B). Performing a qualitatively comparison between Leu misincorporation levels and LeuRS/SerRS expression ratio collected at control conditions (Figure 3.7A), it is clear that there is a positive correlation in all the selected strains. Additionally, Leu misincorporation levels are not absolutely proportional to the LeuRS/SerRS expression ratio. For instance, *HAP2* has the highest ratio of LeuRS/SerRS expression, but does not

present the highest levels of Leu misincorporation at CUG sites (Figure 3.7A, B). The increase in Leu misincorporation in strains *HAP2* and *HAP3* is due to the increase of LeuRS expression, suggesting that under control conditions these transcription factors could be repressors of LeuRS expression (Figure 3.7C). In mutants *ASH1*, *RTG3* and *STB5*, the increase in Leu misincorporation seems to be the result of increased expression of both LeuRS and SerRS (Figure 3.7C), suggesting a role of these TFs in the regulation of both enzymes.

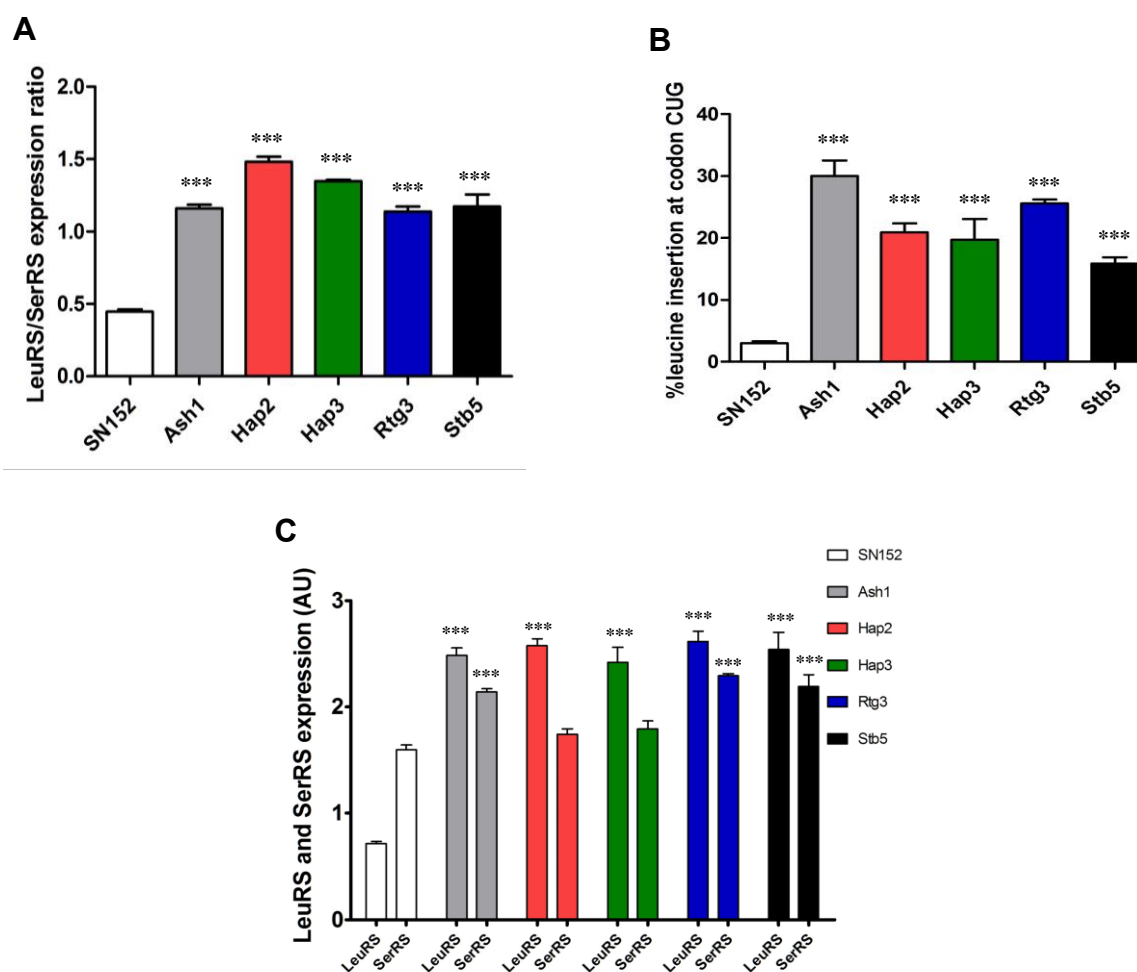


Figure 3.7 – Comparison of Leu misincorporation and LeuRS/SerRS ratio expression in WT strain SN152 and deletion strains *ASH1*, *HAP3*, *HAP2*, *RTG3*, and *STB5* obtained at control conditions.

A) LeuRS/SerRS expression ratio, **(B)** leucine misincorporation levels at CUG sites, **(C)** LeuRS and SerRS expression at control conditions (30°C, pH7) in WT strain SN152 and deletion strains *ASH1*, *HAP3*, *HAP2*, *RTG3* and *STB5*. Data retrieved from at least 300 cells of 3 different clones from each strain. (***) $p < 0.001$, ** $p < 0.01$, * $p < 0.05$ vs control).

The Leu misincorporation level in WT (SN152) cells at 37°C was $16.49 \pm 1.43\%$, which is considerably higher compared with the values obtained at 30°C. Mutant strains *HAP3*, *HAP2* and *STB5* showed higher levels of LeuRS/SerRS expression than the WT strain (Figure 3.8A). This increase of LeuRS/SerRS ratio was mainly due to enhancement of LeuRS expression (Figure 3.8C). However, the increase of LeuRS expression was not accompanied by a significantly increase of Leu misincorporation at CUG sites in these strains (Figure 3.8B). This data suggests that TFs *HAP3*, *HAP2* and *STB5* could influence LeuRS expression at temperature 37°C, but it is not sufficient to alter the mistranslation rates. The mutant strain *RTG3* that showed decreased LeuRS/SerRS expression (Figure 3.8A), did not showed a significantly decrease of Leu misincorporation compared to WT strain at 37°C (Figure 3.8B). This decrease of Leu misincorporation is accompanied by an increase of SerRS expression and decrease of LeuRS expression (Figure 3.8C). Surprisingly, the mutant strain *ASH1* showed inconsistent results. Previously, it had exhibited a decreased LeuRS/SerRS expression (Figure 3.8A), due to decrease of LeuRS expression (Figure 3.8C). However, Leu misincorporation in this strain was almost at wild type level. At 37°C, it seems none of the identified TFs were validated.

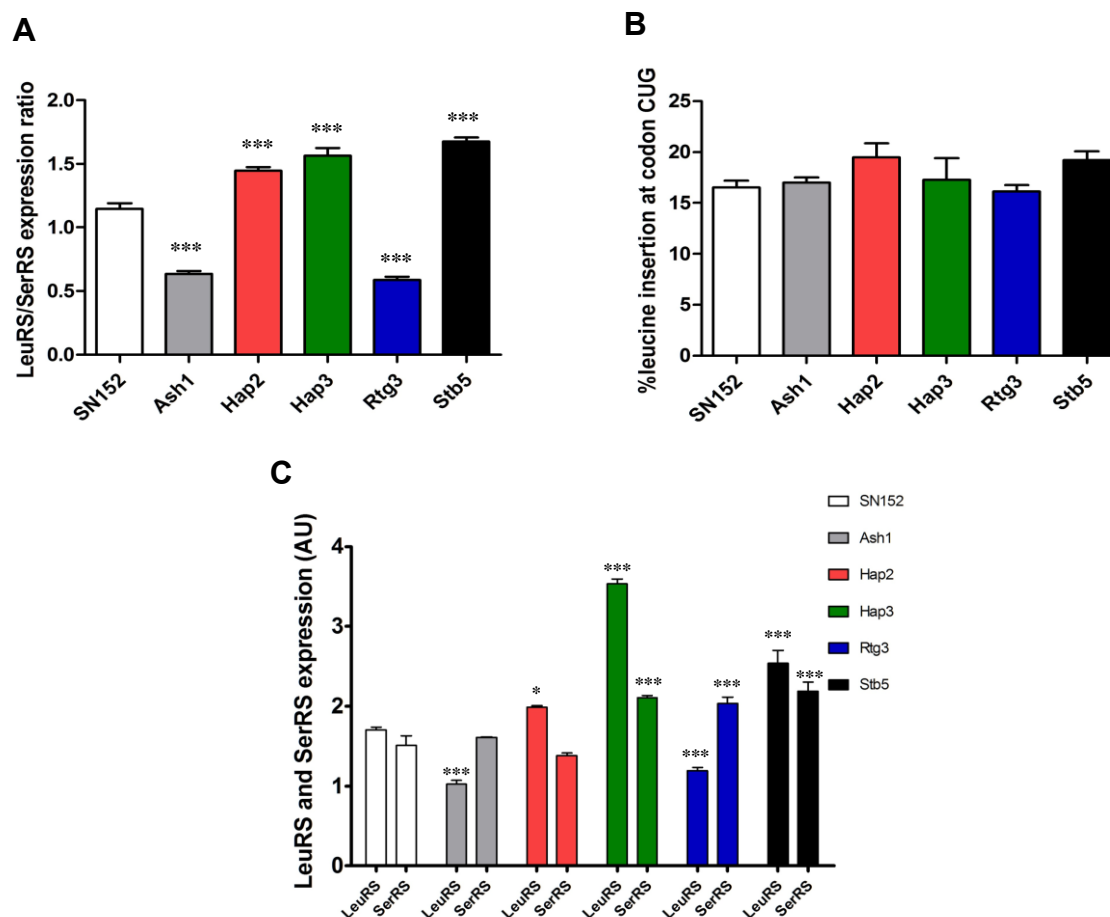


Figure 3.8 – Comparison of Leu misincorporation and LeuRS/SerRS ratio expression in WT strain SN152 and deletion strains *ASH1*, *HAP3*, *HAP2*, *RTG3*, and *STB5* obtained at 37°C.

A) LeuRS/SerRS expression ratio, **(B)** leucine misincorporation levels at CUG sites, **(C)** LeuRS and SerRS expression at 37°C in WT strain SN152 and deletion strains *ASH1*, *HAP3*, *HAP2*, *RTG3* and *STB5*. Data retrieved from at least 300 cells of 3 different clones from each strain. (***) $p < 0.001$, (**) $p < 0.01$, (*) $p < 0.05$ vs control).

WT cells under osmotic pressure induced by 0.2M sorbitol showed $4.27 \pm 1.59\%$ of Leu misincorporation at CUG sites. All the mutant strains tested displayed higher levels of Leu misincorporation than the control strain (Figure 3.9B). The increase of Leu misincorporation in the deletion strains was associated with the increase of LeuRS/SerRS expression (Figure 3.9A). In spite of the significant increase of both LeuRS and SerRS expression in response to osmotic stress in almost all strains (with the exception of *HAP3* that showed increased expression of LeuRS only), there was a clear higher LeuRS expression in relation to SerRS expression (Figure 3.9C). In line with the results from standard growth conditions, these results imply that TFs *ASH1*, *HAP3*, *HAP2*, *RTG3* and *STB5* could be repressors of LeuRS expression under osmotic stress.

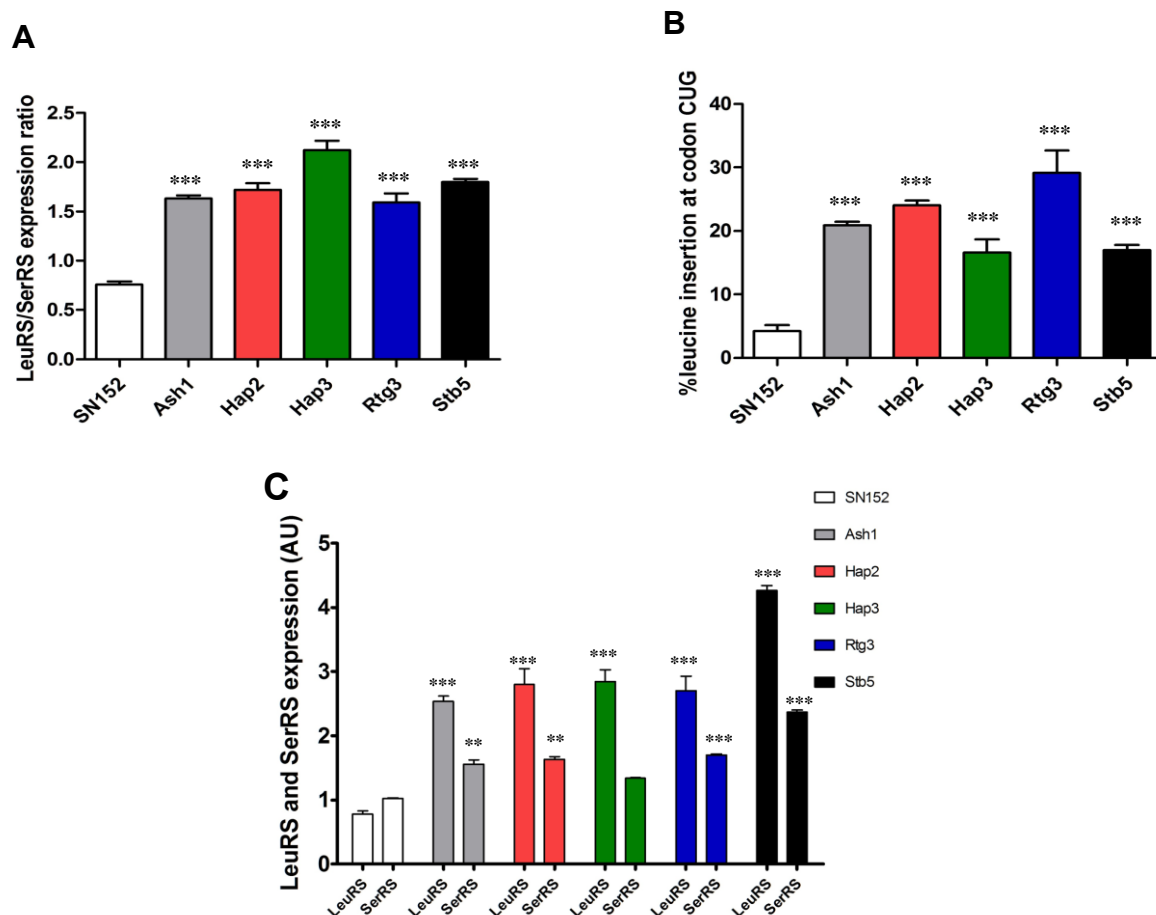
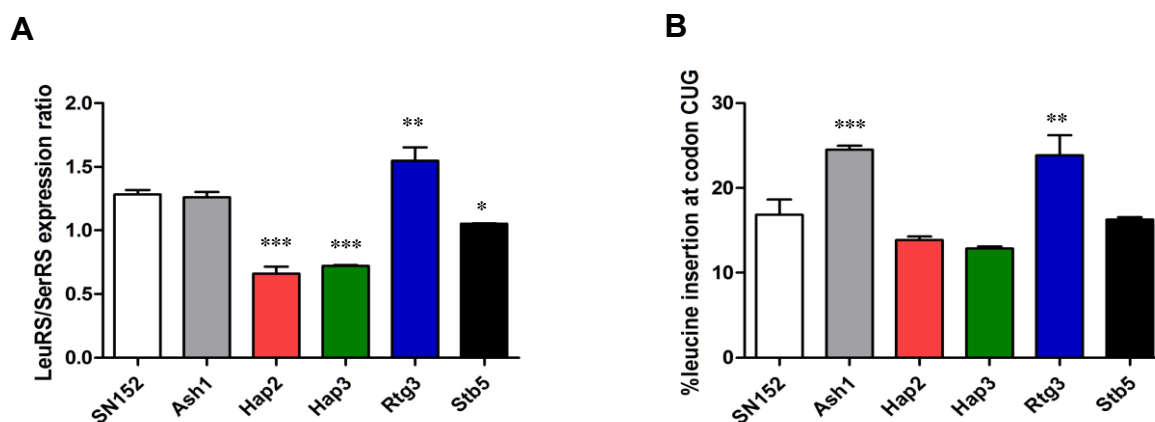


Figure 3.9 – Comparison of Leu misincorporation and LeuRS/SerRS ratio expression in WT strain SN152 and deletion strains *ASH1*, *HAP3*, *HAP2*, *RTG3*, and *STB5* obtained under osmotic stress.

A) LeuRS/SerRS expression ratio, (B) leucine misincorporation levels at CUG sites, (C) LeuRS and SerRS expression in presence of 0.2M of sorbitol in WT strain SN152 and deletion strains *ASH1*, *HAP3*, *HAP2*, *RTG3* and *STB5*. Data retrieved from at least 300 cells of 3 different clones from each strain. (*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$ vs control).

WT SN152 cells grown at pH 6 presented $16.80 \pm 3.01\%$ of Leu misincorporation at CUG codons. Deletion mutants *HAP3*, *HAP2* and *STB5* that revealed a decrease of LeuRS/SerRS ratio (Figure 3.9A), showed a slight decrease of Leu misincorporation compared to WT strain (Figure 3.9B). The decrease of LeuRS/SerRS ratio in mutant strains *HAP3*, *HAP2* was associated with an increase of SerRS expression (Figure 3.9C). Once more, mutant strain *ASH1* showed quite contradictory results. This mutant strain displayed LeuRS and SerRS expression levels similar to the WT strain when grown at media with pH 6 (Figure 3.9A, C). However, the misincorporation levels in this mutant strain were significantly higher than the WT strain (Figure 3.9B). The mutant strain *RTG3* showed an increase of LeuRS/SerRS ratio (due to a decrease of SerRS expression) and an increase of Leu misincorporation (Figure 3.9A, B). Thus, the deletion strain *RTG3* was the only knockout strain that at pH6 showed an unequivocally positive correlation between the increase of LeuRS/SerRS ratio and Leu misincorporation. The fact that the increase of LeuRS/SerRS expression was mainly due to a decrease of SerRS expression indicates that *RTG3* may act as an enhancer of SerRS expression.



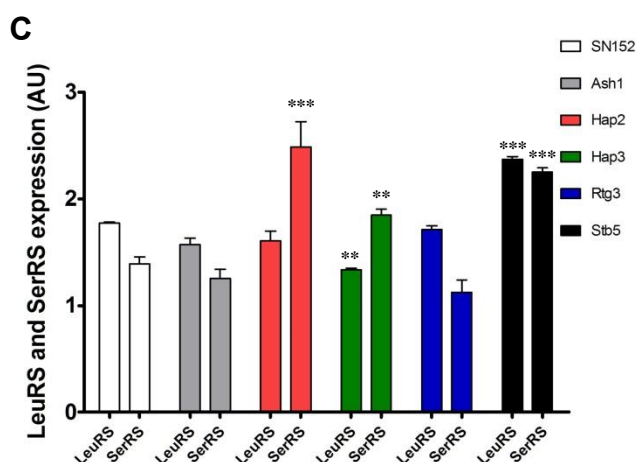


Figure 3.10 – Comparison of Leu misincorporation and LeuRS/SerRS ratio expression in WT strain SN152 and deletion strains *ASH1*, *HAP3*, *HAP2*, *RTG3*, and *STB5* obtained at pH 6.

A) LeuRS/SerRS expression ratio, **(B)** leucine misincorporation levels at CUG sites, **(C)** LeuRS and SerRS expression at pH 6 in WT strain SN152 and deletion strains *ASH1*, *HAP3*, *HAP2*, *RTG3* AND *STB5*. Data retrieved from at least 300 cells of 3 different clones from each strain. (***) $p < 0.001$, (**) $p < 0.01$, (*) $p < 0.05$ vs control).

Overall, the validation approach allowed us to exclude *ASH1* as a possible regulator of LeuRS/SerRS expression at 37°C and at pH 6. *RTG3* was validated as possible regulator of LeuRS and SerRS at 30°C, osmotic stress and pH6. This transcription factor exhibited a possible dual role in the regulation of aaRSs expression as it may be repressor at 30°C and osmotic stress but may function as enhancer at pH 6. Transcription factors *HAP2*, *HAP3* and *STB5* might be repressors of LeuRS and SerRS expression at 30°C and at 0.2 M.

Chapter IV – Discussion, conclusions and future perspectives

4.1 General discussion

In this thesis we tried to shed light in how CUG ambiguous decoding is regulated in *C. albicans*. Initially, we tested if the expression of SerRS and LeuRS was correlated with the leucine misincorporation in different physiological conditions. To accomplish this we used a fluorescent reporter system already available in the host laboratory to assess the expression of both aaRSs in the conditions we already saw that ambiguity changes. Interestingly, it was observed a significant variation in the ratio of LeuRS/SerRS expression between the control conditions and all stress conditions tested. This was in accordance with our hypothesis and was in line with non-published results from a previous assay performed in the host laboratory. In the non-published experiment, LeuRS and SerRS expression was monitored by Western-blot in *C. albicans* cells grown under different physiological conditions and an increase in LeuRS/SerRS was detected in conditions where ambiguity was higher than the control.

In this study, once the increase of LeuRS/SerRS expression under different conditions was demonstrated, we also evaluated if there was a correlation with an increase in ambiguity in the same conditions. Leu misincorporation at CUG sites has been previously measured in WT *C. albicans* cells by Gomes *et al.* (74) and João Simões (PhD thesis) in the host laboratory. In spite of the different approaches used to quantify Leu misincorporation, Gomes and colleagues performed Mass Spectrometry (MS) and João Simões built a fluorescent reporter system, both observed the same trend of increased ambiguity under stressful conditions. For example, the insertion of leucine at CUG codons in *C. albicans* cells is variable under different pH levels and the same trend was noticed in the LeuRS/SerRS ratio expression. *C. albicans* tolerates a wide range of environmental pH levels, from pH of <2 to pH of >10, yet fungi usually are more acidophilic than most of the pathogens (133). Nevertheless, our data also suggest that fluctuations of pH influence the LeuRS/SerRS balance (mostly due to LeuRS variation) and consequently, CUG ambiguity. In fact, a correlation between the LeuRS/SerRS expression and increase in Leu misincorporation was verified in all conditions tested (thermal stress, pH variation and osmotic stress), indicating that the differential expression of both enzymes may be responsible for the regulation of CUG ambiguity.

A previous study from Rocha *et al.* (99) revealed that SerRS contains a CUG codon at position 197, originating the isoforms SerRS_Ser197 and SerRS_Leu197 (99). The position 197 is located at the dimer interface of the enzyme, and substitution of the most frequent amino acid serine for leucine induces a local structural rearrangement, affecting an upstream region of the C-terminal domain which is probably involved with tRNA interactions (99). Also, an experiment carried out by João Simões (PhD thesis) in the host laboratory showed that *C. albicans* strains that only express the SerRS_Leu197 incorporated higher levels of Leu at CUG codons than the strains that only express SerRS_Ser197. Furthermore, the LeuRS also owns a single CUG codon, located at the position 919 at the C-terminal tail. A previous study revealed that the change between the insertion of Leu or Ser at the position 919 does not affect the function or structure of the aminoacylation active site. However, the most common form LeuRS_Ser919 was 30% less active than the isoform LeuRS_Leu919 (25). Altogether, these findings imply that Leu misincorporation may be regulated by a balance of LeuRS and SerRS isoforms with different affinities for the hybrid tRNA_{CAG}.

These previous results from the host laboratory and the observation of an increase of LeuRS/SerRS ratio in cells under stress conditions lead to an important question: what is causing the increase in LeuRS/SerRS ratio; increase of LeuRS expression or decrease of SerRS expression? By analysing the data of each aminoacyl-tRNA synthetase separately, we reached to the conclusion that the increase of LeuRS/SerRS ratio is mostly caused by a variation in LeuRS expression. In WT cells, values of SerRS expression were almost constant in all conditions (Figure 3.3), which suggest that this aaRS might be insensitive to environmental cues. These results oppose the proteomics analysis of *C. albicans* cells exposed to macrophages in which SerRS expression was downregulated (134). However, these results are in line with another study carried out by Gomes *et al.* (data not published). The alignment of the DNA sequence of LeuRS, SerRS and TrpRS (tryptophanyl tRNA synthetase) genes in 5 different *C. albicans* strains revealed non-silent single nucleotide polymorphisms (SNPs) in the LeuRS gene, while SerRS and TrpRS genes had silent SNPs. In each strain, those polymorphisms account for 2 LeuRS isoforms with different promoters in both alleles. Such findings imply that LeuRS gene is regulated by transcription and that leucine incorporation at CUG sites may be modulated by the differential expression and affinity for the tRNA_{CAG}^{Ser} of the 2 LeuRS isoforms. Therefore, it was proposed that in

response to stress conditions, specific transcriptions factors would be activated and would enhance the expression of the isoforms with higher affinity for the tRNA_{CAG}^{Ser}. As a result, there would be an increase of tRNA_{CAG}^{Ser} charged with Leu to compete with the Ser- tRNA_{CAG}^{Ser} for the CUG codons.

Most of the pathways regulating aaRSs in response to environmental changes involve uncharged tRNAs that act as sensors and activate the regulatory reactions. Although, the regulatory mechanisms are not the same between organisms and aaRSs it has been demonstrated that both transcriptional and post-transcriptional mechanisms are involved (132, 135). To date, little is known about the regulation of aaRSs in eukaryotes. Perhaps the most well-known case is the regulation of AspRS (Aspartyl-tRNA synthetase) in *Saccharomyces cerevisiae*. A model for AspRS feedback regulation proposed by Frugier *et al.* states that the excess of AspRS in the cytoplasm is imported in the nucleus and inhibits its own transcription by binding to the 5' untranslated region (5'UTR) and 5' extremity of its coding mRNA^{AspRS}. Moreover, the transcribed tRNA^{Asp} can compete with the mRNA^{AspRS} for the AspRS, hence releasing the AspRS from its own inhibition (132).

To investigate the regulatory mechanisms of LeuRS and SerRS, we started by screening 27 TFs (description list on section 2.6, table 2.5) from a TF deletion library constructed by Homann *et al.* (115). The LeuRS/SerRS ratio expression of the deletion mutant strains was assessed in the different physiological conditions where we previously had established a positive correlation with ambiguity. We considered promising regulatory candidates, TFs whose mutants exhibited statistically different LeuRS/SerRS expression ratios from the WT strain in at least in two tested conditions.

A global look at the screening results showed that the deletion strains do not show the same behaviour through the different stress conditions tested. This is not entirely unexpected, because TFs normally only become activated under specific conditions (136). However, this also represents one of the caveats that must be kept in mind while analysing the results from the knockout collection. If the inducing signals of the TFs are not entirely known, one cannot be sure if the lack of difference relative to the wild type is due to the environmental condition tested or if the TF actually does not affect the expression of the LeuRS or SerRS at that particular condition.

At control conditions (30°C), most mutant strains showed a higher LeuRS/SerRS than the WT strain with LeuRS being the one with a marked increased expression. In this manner, this data suggests that at 30°C, most of the TFs assessed could act as repressors of LeuRS. Another possible explanation is that the knockout of some TFs by itself can be stressful for *C. albicans* cells. Consequently, mistranslation rates are higher than in WT cells and is reflected in a higher LeuRS/SerRS ratio. Similar results were obtained in response to osmotic stress, while at 37°C there were strains whose deleted TF could function as repressors or enhancers of LeuRS expression. Strikingly, when the mutant strains were grown in media at pH 6, the LeuRS/SerRS expression tended to be lower than the WT strain. This implies that the lacking of those TFs could impair LeuRS activation under pH fluctuations. Accordingly, the analysis of LeuRS/SerRS expression individually for each physiological condition retrieved us a variety of regulators. The possibility that the LeuRS and SerRS have such an amount of regulators is quite low. For this reason, we considered a pattern across all conditions tested and detected 5 possible regulators: *ASH1*, *HAP2*, *HAP3*, *RTG3* and *STB5*.

ASH1 is a GATA-like transcription factor required for filamentous growth and virulence in mouse model (122, 137). In addition *ASH1* mRNA is often transported to daughter cells and hyphal tips for the She system, a complex responsible for the transport of mRNAs from mother to daughter cell during mitosis (138). In this screening, the mutant strain lacking *ASH1* showed increased LeuRS/SerRS expression at 30°C and in response to osmotic stress, thus could have a role as negative regulator of LeuRS and SerRS. Surprisingly, at 37°C *ASH1* mutants exhibited a lower LeuRS/SerRS expression that did not influence the level of CUG ambiguity. This finding is rather interesting, because 37°C is an inducing condition of hyphal growth (86), thus a condition where *ASH1* is required.

HAP2 and *HAP3* are CCAAT-binding TFs, which are regulatory sequences for induction of the ferric reductase *FRP1* in iron limited environments (122, 139). *FRP1* expression is induced by neutral-alkaline pH, which is the range of pH of most of the human body sites (139). *HAP2* and *HAP3* deletion strains demonstrated identical LeuRS/SerRS expression rates. Our data suggested that these TFs could act as repressors of LeuRS expression at the thermal and osmotic conditions tested.

RTG3 is one of the helix-loop-helix/leucine zipper (bHLH/Zip) transcription factors that activate the RTG pathway, a cellular stress response activated by alteration in mitochondrial functions. In addition, *RTG3* induces tolerance to azoles and terbinafine antifungals and cations (122, 140). *RTG3* was the only TF whose data implied a role as enhancer of SerRS expression at pH 6. Results concerning the thermal and osmotic stresses conveyed a repressor role of LeuRS expression.

STB5 is a TF that belongs to the fungal specific zinc cluster Zn²-Cys₆ class (141) and is thought to have a role in filamentous growth (122). The strain lacking this TF displayed similar results to the mutant strains *HAP2* and *HAP3*.

To validate the regulatory potential of these 5 candidates, we chose an already established fluorescent reporter system to quantify Leu misincorporation (76). This approach allows the quantification of Leu misincorporation at the single cell level and is less complex than mass spectrometry (MS) technique performed by Gomes *et al.* (74). The levels of Leu misincorporation in WT cells detected in this study were tendentially higher than the levels obtained by Gomes *et al.* (74) using MS. However, the proportion of Leu misincorporation among the physiological conditions is maintained. We detected 3.02% and 16.49% of Leu misincorporation at 30°C and 37°C, while Gomes *et al.* obtained at the same conditions 2.96% and 3.9% respectively (74). When WT cells were grown at pH6 and osmotic stress (0.2M of sorbitol) we obtained 16.80 % and 4.27% respectively, while João Simões (PhD thesis) obtained 11.7% and 9.2% respectively. Using the fluorescent reporter system, we validated all the TFs at the control conditions and osmotic stress. All deletions showed a positive correlation between increase of LeuRS/SerRS expression and increase of Leu misincorporation. This means, that TFs *ASH1*, *HAP2*, *HAP3*, *RTG3* and *STB5* could be repressors of LeuRS and SerRS expression. The possibility that these TFs are regulating LeuRS expression is rather intriguing, because only *ASH1* and *STB5* have predictive DNA binding motifs in LeuRS promoter in our bioinformatics analysis. *HAP2*, *HAP3* and *RTG3* only showed specificity for the SerRS promoter. We must highlight the fact that the bioinformatics tools used were designed for *S. cerevisiae*. Thus, even if the list of TFs tested in this study has orthologs in *S. cerevisiae*, the DNA binding motifs may not be the same in both organisms. Despite the proximity between these two organisms, transcriptional rewiring is common (115, 136).

However, at 37°C none of the deletion strains was validated. Despite the alterations in LeuRS/SerRS expression, there were no significant differences in Leu misincorporation levels in the knock-out strains tested. Nevertheless, at this point we cannot rule out the possibility that the high incorporation of Leu observed in the WT strain (16.49%) influenced these results.

When cells were grown at pH 6, only one mutant strain (*RTG3*) is immediately validated as possible repressor of LeuRS expression, as the expected increase of Leu misincorporation at the CUG sites was observed. However, we highlight the fact that in strains *HAP2*, *HAP3* and *STB5* the expected trend of decreased Leu misincorporation was observed (although with no statistical significance).

A summary of the validation results is depicted in table 4.1 but the regulatory role of *ASH1*, *RTG3*, *HAP2*, *HAP3* and *STB5* in LeuRS and SerRS expression should be further investigated.

Table 4.1 – Summary of the validation results of the identified transcription factors.

	30°C		37°C		Osmotic stress (0.2M)		pH6	
	LeuRS	SerRS	LeuRS	SerRS	LeuRS	SerRS	LeuRS	SerRS
<i>ASH1</i>	Repressor	Repressor	-	-	Repressor	Repressor	-	-
<i>HAP2</i>	Repressor	-	-	-	Repressor	Repressor	-	-
<i>HAP3</i>	Repressor	-	-	-	Repressor	-	-	-
<i>RTG3</i>	Repressor	Repressor	-	-	Repressor	Repressor	-	Enhancer
<i>STB5</i>	Repressor	Repressor	-	-	Repressor	Repressor	-	-

4.2 Main conclusions and future perspectives

This work provided the first insight into the transcriptional regulation of CUG mistranslation in *C. albicans*. Our data showed that an increase of Leu misincorporation is accompanied by an increase of LeuRS/SerRS expression. This data supports the hypothesis that the insertion of leucine and serine at CUG codons is dependent upon the competition between LeuRS and SerRS for the tRNA_{CAG}^{Ser}. In WT cells, the differential expression of LeuRS/SerRS seen in non-optimal conditions may be caused by fluctuations in LeuRS levels, while SerRS seems to be stress insensitive. It would be important to validate the expression LeuRS and SerRS obtained in this work using Western blot. Due to the fact that both SerRS and LeuRS have an ambiguous CUG codon, the kinetic parameters of the aminoacylation reaction of the different isoforms of LeuRS and SerRS should also be determined. It would be interesting to evaluate if the different isoforms of the enzymes have specific regulatory mechanisms.

This thesis identified the transcription factors *ASH1*, *HAP2*, *HAP3*, *RTG3* and *STB5* as possible regulators of LeuRS/SerRS expression. However, future experiments are warranted to clarify the function of each one of these proteins in the aaRSs regulation. Using the same approach of this study, it would be interesting to grow the identified knock-out strains in the condition where the highest level of Leu incorporation was observed (João Simões, PhD thesis). In this case, exposure of *C. albicans* cells to macrophage and amphotericin B causes an increase of ambiguity up to 50% and the regulatory potential of the mutant strains should accompany this considerable increase. This could be accomplished by performing a macrophage phagocytosis assay, co-culturing *ASH1*, *HAP2*, *HAP3*, *RTG3* or *STB5* mutant *C. albicans* cells with a differentiated human monocyte/macrophage cell line.

Finally, in the host laboratory, at the same time of the screening of the TF knockout strain collection, it was performed a screening of a kinase knockout strain collection for potential SerRS and LeuRS regulators. In the near future it will be interesting to cross the data from both screenings. To unveil the regulatory circuits and identify the full gene regulatory spectra of CUG mistranslation is imperative to conduct an identification of the transcription factor binding sites (of the identified 5 TFs) in the LeuRS and SerRS promoters. LeuRS and SerRS promoter recruitment of *ASH1*, *HAP2*,

HAP3, *RTG3* and *STB5* should be confirmed by Chromatin Immunoprecipitation (ChIP) followed by quantitate real time PCR (qRT-PCR).

Recent data strongly suggests a major role for CUG ambiguity on *C. albicans* pathogenesis by modulating host-pathogen interactions (76). Hence, we are hopeful that data from this study, together with follow up experiments, will contribute to uncover the regulators of CUG ambiguity and ultimately clarify the role of mistranslation on *C. albicans* virulence.

References

1. Crick FHC. The origin of the genetic code *Journal of Molecular Biology*. 1968;367-79.
2. Cropp TA, Schultz PG. An expanding genetic code. *Trends in genetics : TIG*. 2004;20(12):625-30. Epub 2004/11/04.
3. Education N. The amino acids specified by each mRNA codon. Multiple codons can code for the same amino acid. 2014 [cited 2014]; Available from: <http://www.nature.com/scitable/content/the-amino-acids-specified-by-each-mrna-6903567>.
4. Clancy S, Brown W. Translation: DNA to mRNA to Protein 2008.
5. Santos J, Monteagudo A. Study of the genetic code adaptability by means of a genetic algorithm. *Journal of theoretical biology*. 2010;264(3):854-65. Epub 2010/03/12.
6. Tlustý T. A colorful origin for the genetic code: information theory, statistical mechanics and the emergence of molecular codes. *Physics of life reviews*. 2010;7(3):362-76. Epub 2010/06/19.
7. Brocker MJ, Ho JM, Church GM, Soll D, O'Donoghue P. Recoding the genetic code with selenocysteine. *Angew Chem Int Ed Engl*. 2014;53(1):319-23. Epub 2014/02/11.
8. Koonin EV, Novozhilov AS. Origin and evolution of the genetic code: the universal enigma. *IUBMB life*. 2009;61(2):99-111. Epub 2009/01/02.
9. Hamashima K, Kanai A. Alternative genetic code for amino acids and transfer RNA revisited. *Biomolecular concepts*. 2013;4(3):309-18. Epub 2013/06/01.
10. Ohama T, Inagaki Y, Bessho Y, Osawa S. Evolving genetic code. *Proceedings of the Japan Academy Series B, Physical and biological sciences*. 2008;84(2):58-74. Epub 2008/10/23.
11. Barrell BG, Bankier AT, Drouin J. A different genetic code in human mitochondria. *Nature*. 1979;282(5735):189-94. Epub 1979/11/08.
12. Lobanov AV, Turanov AA, Hatfield DL, Gladyshev VN. Dual functions of codons in the genetic code. *Critical reviews in biochemistry and molecular biology*. 2010;45(4):257-65. Epub 2010/05/08.
13. Miranda I, Silva R, Santos MA. Evolution of the genetic code in yeasts. *Yeast*. 2006;23(3):203-13. Epub 2006/02/25.

14. Lavrov DV, Pett W, Voigt O, Worheide G, Forget L, Lang BF, et al. Mitochondrial DNA of *Clathrina clathrus* (Calcarea, Calcinea): six linear chromosomes, fragmented rRNAs, tRNA editing, and a novel genetic code. *Molecular biology and evolution*. 2013;30(4):865-80. Epub 2012/12/12.
15. Hayashi-Ishimaru Y, Ohama T, Kawatsu Y, Nakamura K, Osawa S. UAG is a sense codon in several chlorophycean mitochondria. *Current genetics*. 1996;30(1):29-33.
16. Knight RD, Freeland SJ, Landweber LF. Rewiring the Keyboard: Evolvability of the Genetic Code. *Nature Reviews Genetics* 2001;2:49-58.
17. Su D, Lieberman A, Lang BF, Simonovic M, Soll D, Ling JQ. An unusual tRNA(Thr) derived from tRNA(His) reassigns in yeast mitochondria the CUN codons to threonine. *Nucleic acids research*. 2011;39(11):4866-74.
18. Massey SE, Garey JR. A comparative genomics analysis of codon reassignments reveals a link with mitochondrial proteome size and a mechanism of genetic code change via suppressor tRNAs. *Journal of molecular evolution*. 2007;64(4):399-410. Epub 2007/03/29.
19. Sanchez-Silva R, Villalobo E, Morin L, Torres A. A new noncanonical nuclear genetic code: translation of UAA into glutamate. *Current biology : CB*. 2003;13(5):442-7. Epub 2003/03/07.
20. Lovett PS, Ambulos NP, Jr., Mulbry W, Noguchi N, Rogers EJ. UGA can be decoded as tryptophan at low efficiency in *Bacillus subtilis*. *Journal of bacteriology*. 1991;173(5):1810-2. Epub 1991/03/11.
21. Watanabe K, Suzuki T. Universal Genetic Code and its Natural Variations. *Encyclopedia of Life Sciences*. 2008.
22. Fekner T, Chan MK. The pyrrolysine translational machinery as a genetic-code expansion tool. *Current opinion in chemical biology*. 2011;15(3):387-91. Epub 2011/04/22.
23. Fujishima K, Kanai A. tRNA gene diversity in the three domains of life. *Frontiers in genetics*. 2014;5:142. Epub 2014/06/07.
24. Hopper AK, Pai DA, Engelke DR. Cellular dynamics of tRNAs and their genes. *FEBS letters*. 2010;584(2):310-7. Epub 2009/11/26.
25. Zhou X, Wang E. Transfer RNA: a dancer between charging and mis-charging for protein biosynthesis. *Science China Life sciences*. 2013;56(10):921-32. Epub 2013/08/29.
26. Raina M, Ibba M. tRNAs as regulators of biological processes. *Frontiers in genetics*. 2014;5:171. Epub 2014/06/27.

27. Giege R. Toward a more complete view of tRNA biology. *Nature structural & molecular biology*. 2008;15(10):1007-14. Epub 2008/10/07.
28. Giege R, Juhling F, Putz J, Stadler P, Sauter C, Florentz C. Structure of transfer RNAs: similarity and variability. *Wiley interdisciplinary reviews RNA*. 2012;3(1):37-61. Epub 2011/10/01.
29. Hopper AK. Transfer RNA post-transcriptional processing, turnover, and subcellular dynamics in the yeast *Saccharomyces cerevisiae*. *Genetics*. 2013;194(1):43-67. Epub 2013/05/02.
30. Wichtowska D, Turowski TW, Boguta M. An interplay between transcription, processing, and degradation determines tRNA levels in yeast. *Wiley interdisciplinary reviews RNA*. 2013;4(6):709-22. Epub 2013/09/17.
31. Wang X, He C. Dynamic RNA modifications in posttranscriptional regulation. *Molecular cell*. 2014;56(1):5-12. Epub 2014/10/04.
32. Agris PF. Decoding the genome: a modified view. *Nucleic acids research*. 2004;32(1):223-38. Epub 2004/01/13.
33. Agris PF, Vendeix FA, Graham WD. tRNA's wobble decoding of the genome: 40 years of modification. *J Mol Biol*. 2007;366(1):1-13. Epub 2006/12/26.
34. Jackman JE, Alfonzo JD. Transfer RNA modifications: nature's combinatorial chemistry playground. *Wiley interdisciplinary reviews RNA*. 2013;4(1):35-48. Epub 2012/11/10.
35. Havrylenko S, Mirande M. Aminoacyl-tRNA Synthetase Complexes in Evolution. *International journal of molecular sciences*. 2015;16(3):6571-94. Epub 2015/03/26.
36. Ling J, Reynolds N, Ibba M. Aminoacyl-tRNA synthesis and translational quality control. *Annual review of microbiology*. 2009;63:61-78. Epub 2009/04/22.
37. Pang YL, Poruri K, Martinis SA. tRNA synthetase: tRNA aminoacylation and beyond. *Wiley interdisciplinary reviews RNA*. 2014;5(4):461-80. Epub 2014/04/08.
38. Yao P, Poruri K, Martinis SA, Fox PL. Non-catalytic regulation of gene expression by aminoacyl-tRNA synthetases. *Topics in current chemistry*. 2014;344:167-87. Epub 2013/03/29.
39. Schimmel P. Development of tRNA synthetases and connection to genetic code and disease. *Protein science : a publication of the Protein Society*. 2008;17(10):1643-52. Epub 2008/09/04.
40. Shaul S, Berel D, Benjamini Y, Graur D. Revisiting the operational RNA code for amino acids: Ensemble attributes and their implications. *RNA*. 2010;16(1):141-53. Epub 2009/12/03.

41. Suzuki T, Ueda T, Watanabe K. The 'polysemous' codon--a codon with multiple amino acid assignment caused by dual specificity of tRNA identity. *The EMBO journal*. 1997;16(5):1122-34. Epub 1997/03/03.
42. Jakubowski H. Quality control in tRNA charging. *Wiley interdisciplinary reviews RNA*. 2012;3(3):295-310. Epub 2011/11/19.
43. Martinis SA, Boniecki MT. The balance between pre- and post-transfer editing in tRNA synthetases. *FEBS letters*. 2010;584(2):455-9. Epub 2009/11/28.
44. Perona JJ, Gruic-Sovulj I. Synthetic and editing mechanisms of aminoacyl-tRNA synthetases. *Topics in current chemistry*. 2014;344:1-41. Epub 2013/07/16.
45. Mirando AC, Francklyn CS, Lounsbury KM. Regulation of angiogenesis by aminoacyl-tRNA synthetases. *International journal of molecular sciences*. 2014;15(12):23725-48. Epub 2014/12/24.
46. Kim S, You S, Hwang D. Aminoacyl-tRNA synthetases and tumorigenesis: more than housekeeping. *Nature reviews Cancer*. 2011;11(10):708-18. Epub 2011/09/24.
47. Park SG, Schimmel P, Kim S. Aminoacyl tRNA synthetases and their connections to disease. *Proceedings of the National Academy of Sciences of the United States of America*. 2008;105(32):11043-9. Epub 2008/08/07.
48. Kim D, Kwon NH, Kim S. Association of aminoacyl-tRNA synthetases with cancer. *Topics in current chemistry*. 2014;344:207-45. Epub 2013/07/03.
49. Lv PC, Zhu HL. Aminoacyl-tRNA synthetase inhibitors as potent antibacterials. *Current medicinal chemistry*. 2012;19(21):3550-63. Epub 2012/06/12.
50. Alberts B, Johnson A, Lewis J, Raff M, Roberts K, Walter P. *How Cells Read the Genome: From DNA to Protein*. 5 ed. Science G, editor2008.
51. Scheper GC, van der Knaap MS, Proud CG. Translation matters: protein synthesis defects in inherited disease. *Nature reviews Genetics*. 2007;8(9):711-23. Epub 2007/08/08.
52. Preiss T, M WH. Starting the protein synthesis machine: eukaryotic translation initiation. *BioEssays : news and reviews in molecular, cellular and developmental biology*. 2003;25(12):1201-11. Epub 2003/11/25.
53. Jackson RJ, Hellen CU, Pestova TV. The mechanism of eukaryotic translation initiation and principles of its regulation. *Nature reviews Molecular cell biology*. 2010;11(2):113-27. Epub 2010/01/23.
54. Groppo R, Richter JD. Translational control from head to tail. *Current opinion in cell biology*. 2009;21(3):444-51. Epub 2009/03/17.

-
55. Dever TE, Green R. The elongation, termination, and recycling phases of translation in eukaryotes. *Cold Spring Harbor perspectives in biology*. 2012;4(7):a013706. Epub 2012/07/04.
56. Graifer D, Karpova G. Interaction of tRNA with eukaryotic ribosome. *International journal of molecular sciences*. 2015;16(4):7173-94. Epub 2015/04/02.
57. Kiel MC, Kaji H, Kaji A. Ribosome recycling: An essential process of protein synthesis. *Biochemistry and molecular biology education : a bimonthly publication of the International Union of Biochemistry and Molecular Biology*. 2007;35(1):40-4. Epub 2007/01/01.
58. Reynolds NM, Lazazzera BA, Ibba M. Cellular mechanisms that control mistranslation. *Nature reviews Microbiology*. 2010;8(12):849-56. Epub 2010/11/17.
59. Moghal A, Mohler K, Ibba M. Mistranslation of the genetic code. *FEBS letters*. 2014;588(23):4305-10. Epub 2014/09/16.
60. Drummond DA, Wilke CO. The evolutionary consequences of erroneous protein synthesis. *Nature reviews Genetics*. 2009;10(10):715-24. Epub 2009/09/19.
61. Moura GR, Carreto LC, Santos MA. Genetic code ambiguity: an unexpected source of proteome innovation and phenotypic diversity. *Current opinion in microbiology*. 2009;12(6):631-7. Epub 2009/10/27.
62. Shah P, Gilchrist MA. Effect of correlated tRNA abundances on translation errors and evolution of codon usage bias. *PLoS genetics*. 2010;6(9):e1001128. Epub 2010/09/24.
63. Rodrigo-Brenni MC, Hegde RS. Design principles of protein biosynthesis-coupled quality control. *Developmental cell*. 2012;23(5):896-907. Epub 2012/11/17.
64. Pechmann S, Willmund F, Frydman J. The ribosome as a hub for protein quality control. *Molecular cell*. 2013;49(3):411-21. Epub 2013/02/12.
65. Ryall B, Eydallin G, Ferenci T. Culture history and population heterogeneity as determinants of bacterial adaptation: the adaptomics of a single environmental transition. *Microbiology and molecular biology reviews : MMBR*. 2012;76(3):597-625. Epub 2012/08/31.
66. Lee JW, Beebe K, Nangle LA, Jang J, Longo-Guess CM, Cook SA, et al. Editing-defective tRNA synthetase causes protein misfolding and neurodegeneration. *Nature*. 2006;443(7107):50-5. Epub 2006/08/15.
67. Nangle LA, Motta CM, Schimmel P. Global effects of mistranslation from an editing defect in mammalian cells. *Chemistry & biology*. 2006;13(10):1091-100. Epub 2006/10/21.

68. Wallen RC, Antonellis A. To charge or not to charge: mechanistic insights into neuropathy-associated tRNA synthetase mutations. *Current opinion in genetics & development*. 2013;23(3):302-9. Epub 2013/03/08.
69. Bratulic S, Gerber F, Wagner A. Mistranslation drives the evolution of robustness in TEM-1 beta-lactamase. *Proceedings of the National Academy of Sciences of the United States of America*. 2015;112(41):12758-63. Epub 2015/09/24.
70. Ribas de Pouplana L, Santos MA, Zhu JH, Farabaugh PJ, Javid B. Protein mistranslation: friend or foe? *Trends in biochemical sciences*. 2014;39(8):355-62. Epub 2014/07/16.
71. Li L, Boniecki MT, Jaffe JD, Imai BS, Yau PM, Luthey-Schulten ZA, et al. Naturally occurring aminoacyl-tRNA synthetases editing-domain mutations that cause mistranslation in *Mycoplasma* parasites. *Proceedings of the National Academy of Sciences of the United States of America*. 2011;108(23):9378-83. Epub 2011/05/25.
72. Fan Y, Wu J, Ung MH, De Lay N, Cheng C, Ling J. Protein mistranslation protects bacteria against oxidative stress. *Nucleic acids research*. 2015;43(3):1740-8. Epub 2015/01/13.
73. Javid B, Sorrentino F, Toosky M, Zheng W, Pinkham JT, Jain N, et al. Mycobacterial mistranslation is necessary and sufficient for rifampicin phenotypic resistance. *Proceedings of the National Academy of Sciences of the United States of America*. 2014;111(3):1132-7. Epub 2014/01/08.
74. Gomes AC, Miranda I, Silva RM, Moura GR, Thomas B, Akoulitchiev A, et al. A genetic code alteration generates a proteome of high diversity in the human pathogen *Candida albicans*. *Genome biology*. 2007;8(10):R206. Epub 2007/10/06.
75. Santos MA, Cheesman C, Costa V, Moradas-Ferreira P, Tuite MF. Selective advantages created by codon ambiguity allowed for the evolution of an alternative genetic code in *Candida* spp. *Molecular microbiology*. 1999;31(3):937-47. Epub 1999/02/27.
76. Bezerra AR, Simoes J, Lee W, Rung J, Weil T, Gut IG, et al. Reversion of a fungal genetic code alteration links proteome instability with genomic and phenotypic diversification. *Proceedings of the National Academy of Sciences of the United States of America*. 2013;110(27):11079-84. Epub 2013/06/19.
77. Miranda I, Silva-Dias A, Rocha R, Teixeira-Santos R, Coelho C, Goncalves T, et al. *Candida albicans* CUG mistranslation is a mechanism to create cell surface variation. *mBio*. 2013;4(4). Epub 2013/06/27.
78. Gasch AP. Comparative genomics of the environmental stress response in ascomycete fungi. *Yeast*. 2007;24(11):961-76. Epub 2007/07/03.

79. Butler G, Rasmussen MD, Lin MF, Santos MA, Sakthikumar S, Munro CA, et al. Evolution of pathogenicity and sexual reproduction in eight *Candida* genomes. *Nature*. 2009;459(7247):657-62. Epub 2009/05/26.
80. Hickman MA, Zeng G, Forche A, Hirakawa MP, Abbey D, Harrison BD, et al. The 'obligate diploid' *Candida albicans* forms mating-competent haploids. *Nature*. 2013;494(7435):55-9. Epub 2013/02/01.
81. Gow NA, van de Veerdonk FL, Brown AJ, Netea MG. *Candida albicans* morphogenesis and host defence: discriminating invasion from colonization. *Nature reviews Microbiology*. 2012;10(2):112-22. Epub 2011/12/14.
82. Brunke S, Hube B. Two unlike cousins: *Candida albicans* and *C. glabrata* infection strategies. *Cellular microbiology*. 2013;15(5):701-8. Epub 2012/12/21.
83. Lim CS, Rosli R, Seow HF, Chong PP. *Candida* and invasive candidiasis: back to basics. *European journal of clinical microbiology & infectious diseases* : official publication of the European Society of Clinical Microbiology. 2012;31(1):21-31. Epub 2011/05/06.
84. Behzadi P, Behzadi E, Ranjbar R. Urinary tract infections and *Candida albicans*. *Central European journal of urology*. 2015;68(1):96-101. Epub 2015/04/29.
85. Guinea J. Global trends in the distribution of *Candida* species causing candidemia. *Clinical microbiology and infection* : the official publication of the European Society of Clinical Microbiology and Infectious Diseases. 2014;20 Suppl 6:5-10. Epub 2014/02/11.
86. Berman J, Sudbery PE. *Candida Albicans*: a molecular revolution built on lessons from budding yeast. *Nature reviews Genetics*. 2002;3(12):918-30. Epub 2002/12/03.
87. Sudbery P, Gow N, Berman J. The distinct morphogenic states of *Candida albicans*. *Trends in microbiology*. 2004;12(7):317-24. Epub 2004/06/30.
88. Cullen PJ, Sprague GF, Jr. The regulation of filamentous growth in yeast. *Genetics*. 2012;190(1):23-49. Epub 2012/01/06.
89. Soll DR. The role of phenotypic switching in the basic biology and pathogenesis of *Candida albicans*. *Journal of oral microbiology*. 2014;6. Epub 2014/01/24.
90. Bennett RJ. The parasexual lifestyle of *Candida albicans*. *Current opinion in microbiology*. 2015;28:10-7. Epub 2015/07/27.
91. Bennett RJ, Johnson AD. Completion of a parasexual cycle in *Candida albicans* by induced chromosome loss in tetraploid strains. *The EMBO journal*. 2003;22(10):2505-15. Epub 2003/05/14.

-
92. Berman J, Hadany L. Does stress induce (para)sex? Implications for *Candida albicans* evolution. *Trends in genetics : TIG*. 2012;28(5):197-203. Epub 2012/03/01.
93. Ene IV, Bennett RJ. The cryptic sexual strategies of human fungal pathogens. *Nature reviews Microbiology*. 2014;12(4):239-51. Epub 2014/03/15.
94. Massey SE, Moura G, Beltrao P, Almeida R, Garey JR, Tuite MF, et al. Comparative evolutionary genomics unveils the molecular mechanism of reassignment of the CTG codon in *Candida* spp. *Genome research*. 2003;13(4):544-57. Epub 2003/04/03.
95. Santos MA, Gomes AC, Santos MC, Carreto LC, Moura GR. The genetic code of the fungal CTG clade. *Comptes rendus biologies*. 2011;334(8-9):607-11. Epub 2011/08/09.
96. Moura GR, Paredes JA, Santos MA. Development of the genetic code: insights from a fungal codon reassignment. *FEBS letters*. 2010;584(2):334-41. Epub 2009/11/28.
97. Santos MA, Keith G, Tuite MF. Non-standard translational events in *Candida albicans* mediated by an unusual seryl-tRNA with a 5'-CAG-3' (leucine) anticodon. *The EMBO journal*. 1993;12(2):607-16. Epub 1993/02/01.
98. Santos MA, Ueda T, Watanabe K, Tuite MF. The non-standard genetic code of *Candida* spp.: an evolving genetic code or a novel mechanism for adaptation? *Molecular microbiology*. 1997;26(3):423-31. Epub 1997/12/24.
99. Rocha R, Pereira PJ, Santos MA, Macedo-Ribeiro S. Unveiling the structural basis for translational ambiguity tolerance in a human fungal pathogen. *Proceedings of the National Academy of Sciences of the United States of America*. 2011;108(34):14091-6. Epub 2011/08/10.
100. Silva RM, Paredes JA, Moura GR, Manadas B, Lima-Costa T, Rocha R, et al. Critical roles for a genetic code alteration in the evolution of the genus *Candida*. *The EMBO journal*. 2007;26(21):4555-65. Epub 2007/10/13.
101. Paredes JA, Carreto L, Simoes J, Bezerra AR, Gomes AC, Santamaria R, et al. Low level genome mistranslations deregulate the transcriptome and translate and generate proteotoxic stress in yeast. *BMC biology*. 2012;10:55. Epub 2012/06/22.
102. Brown AJ, Budge S, Kaloriti D, Tillmann A, Jacobsen MD, Yin Z, et al. Stress adaptation in a pathogenic fungus. *The Journal of experimental biology*. 2014;217(Pt 1):144-55. Epub 2013/12/20.
103. Enjalbert B, Smith DA, Cornell MJ, Alam I, Nicholls S, Brown AJ, et al. Role of the Hog1 stress-activated protein kinase in the global transcriptional response to stress in the fungal pathogen *Candida albicans*. *Molecular biology of the cell*. 2006;17(2):1018-32. Epub 2005/12/13.

104. Alarco AM, Raymond M. The bZip transcription factor Cap1p is involved in multidrug resistance and oxidative stress response in *Candida albicans*. *Journal of bacteriology*. 1999;181(3):700-8. Epub 1999/01/28.
105. Smith DA, Nicholls S, Morgan BA, Brown AJ, Quinn J. A conserved stress-activated protein kinase regulates a core stress response in the human pathogen *Candida albicans*. *Molecular biology of the cell*. 2004;15(9):4179-90. Epub 2004/07/02.
106. Smith DA, Morgan BA, Quinn J. Stress signalling to fungal stress-activated protein kinase pathways. *FEMS microbiology letters*. 2010;306(1):1-8. Epub 2010/03/30.
107. Prieto D, Roman E, Correia I, Pla J. The HOG pathway is critical for the colonization of the mouse gastrointestinal tract by *Candida albicans*. *PloS one*. 2014;9(1):e87128. Epub 2014/01/30.
108. Williams DW, Jordan RP, Wei XQ, Alves CT, Wise MP, Wilson MJ, et al. Interactions of *Candida albicans* with host epithelial surfaces. *Journal of oral microbiology*. 2013;5. Epub 2013/10/25.
109. Nobile CJ, Fox EP, Nett JE, Sorrells TR, Mitrovich QM, Hernday AD, et al. A recently evolved transcriptional network controls biofilm development in *Candida albicans*. *Cell*. 2012;148(1-2):126-38. Epub 2012/01/24.
110. Mayer FL, Wilson D, Hube B. *Candida albicans* pathogenicity mechanisms. *Virulence*. 2013;4(2):119-28. Epub 2013/01/11.
111. Selmecki A, Forche A, Berman J. Genomic plasticity of the human fungal pathogen *Candida albicans*. *Eukaryotic cell*. 2010;9(7):991-1008. Epub 2010/05/25.
112. Diez-Orejas R, Fernandez-Arenas E. *Candida albicans*-macrophage interactions: genomic and proteomic insights. *Future microbiology*. 2008;3(6):661-81. Epub 2008/12/17.
113. Tsai P-W, Chen Y-T, Hsu P-C, Lan C-Y. Study of *Candida albicans* and its interactions with the host: A mini review. *BioMedicine*. 2013;3(1):51-64.
114. Miramon P, Kasper L, Hube B. Thriving within the host: *Candida* spp. interactions with phagocytic cells. *Medical microbiology and immunology*. 2013;202(3):183-95. Epub 2013/01/29.
115. Homann OR, Dea J, Noble SM, Johnson AD. A phenotypic profile of the *Candida albicans* regulatory network. *PLoS genetics*. 2009;5(12):e1000783. Epub 2009/12/31.
116. Kibbe WA. OligoCalc: an online oligonucleotide properties calculator. *Nucleic acids research*. 2007;35(Web Server issue):W43-6. Epub 2007/04/25.
117. Murad AM, Lee PR, Broadbent ID, Barelle CJ, Brown AJ. CIP10, an efficient and convenient integrating vector for *Candida albicans*. *Yeast*. 2000;16(4):325-7. Epub 2000/02/12.

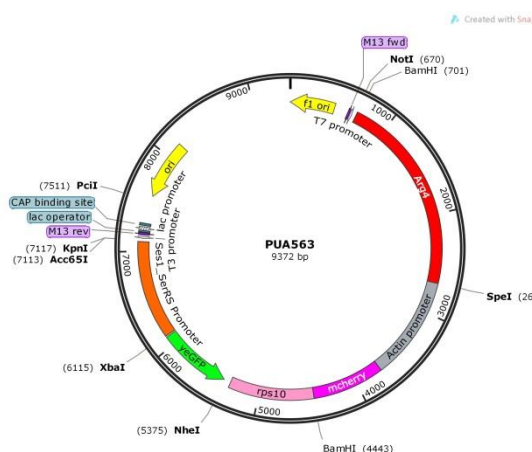
-
118. Sambrook J, Maniatis T, EF F. Molecular cloning: a laboratory manual. 3rd. ed. ed. Cold Spring Harbor, N.Y: Cold Spring Harbor Laboratory; 2001.
119. Walther A, Wendland J. An improved transformation protocol for the human fungal pathogen *Candida albicans*. *Current genetics*. 2003;42(6):339-43. Epub 2003/03/04.
120. Hoffman CS, Winston F. A ten-minute DNA preparation from yeast efficiently releases autonomous plasmids for transformation of *Escherichia coli*. *Gene*. 1987;57(2-3):267-72. Epub 1987/01/01.
121. Teixeira MC, Monteiro PT, Guerreiro JF, Goncalves JP, Mira NP, dos Santos SC, et al. The YEASTRACT database: an upgraded information system for the analysis of gene and genomic transcription regulation in *Saccharomyces cerevisiae*. *Nucleic acids research*. 2014;42(Database issue):D161-6. Epub 2013/10/31.
122. Inglis DO, Arnaud MB, Binkley J, Shah P, Skrzypek MS, Wymore F, et al. The *Candida* genome database incorporates multiple *Candida* species: multispecies search and analysis tools with curated gene and protein information for *Candida albicans* and *Candida glabrata*. *Nucleic acids research*. 2012;40(Database issue):D667-74. Epub 2011/11/09.
123. Santos MA, Perreau VM, Tuite MF. Transfer RNA structural change is a key element in the reassignment of the CUG codon in *Candida albicans*. *The EMBO journal*. 1996;15(18):5060-8. Epub 1996/09/16.
124. SÃ¡rkÃ¡ny Z, Silva A, Pereira PJB, Macedo-Ribeiro S. Ser or Leu: structural snapshots of mistranslation in *Candida albicans*. *Frontiers in Molecular Biosciences*. 2014;1.
125. Cormack BP, Bertram G, Egerton M, Gow NA, Falkow S, Brown AJ. Yeast-enhanced green fluorescent protein (yEGFP): a reporter of gene expression in *Candida albicans*. *Microbiology*. 1997;143 (Pt 2):303-11. Epub 1997/02/01.
126. Ene IV, Walker LA, Schiavone M, Lee KK, Martin-Yken H, Dague E, et al. Cell Wall Remodeling Enzymes Modulate Fungal Cell Wall Elasticity and Osmotic Stress Resistance. *mBio*. 2015;6(4):e00986. Epub 2015/07/30.
127. de Nadal E, Ammerer G, Posas F. Controlling gene expression in response to stress. *Nature reviews Genetics*. 2011;12(12):833-45. Epub 2011/11/04.
128. Amorim-Vaz S, Delarze E, Ischer F, Sanglard D, Coste AT. Examining the virulence of *Candida albicans* transcription factor mutants using *Galleria mellonella* and mouse infection models. *Frontiers in microbiology*. 2015;6:367. Epub 2015/05/23.

129. Sellam A, van het Hoog M, Tebbji F, Beaurepaire C, Whiteway M, Nantel A. Modeling the transcriptional regulatory network that controls the early hypoxic response in *Candida albicans*. *Eukaryotic cell*. 2014;13(5):675-90. Epub 2014/04/01.
130. Nobile CJ, Mitchell AP. Regulation of cell-surface genes and biofilm formation by the *C. albicans* transcription factor Bcr1p. *Current biology : CB*. 2005;15(12):1150-5. Epub 2005/06/21.
131. Miranda I, Rocha R, Santos MC, Mateus DD, Moura GR, Carreto L, et al. A genetic code alteration is a phenotype diversity generator in the human pathogen *Candida albicans*. *PloS one*. 2007;2(10):e996. Epub 2007/10/04.
132. Frugier M, Ryckelynck M, Giege R. tRNA-balanced expression of a eukaryal aminoacyl-tRNA synthetase by an mRNA-mediated pathway. *EMBO reports*. 2005;6(9):860-5. Epub 2005/08/23.
133. Vylkova S, Carman AJ, Danhof HA, Collette JR, Zhou H, Lorenz MC. The fungal pathogen *Candida albicans* autoinduces hyphal morphogenesis by raising extracellular pH. *mBio*. 2011;2(3):e00055-11. Epub 2011/05/19.
134. Fernandez-Arenas E, Cabezon V, Bermejo C, Arroyo J, Nombela C, Diez-Orejas R, et al. Integrated proteomics and genomics strategies bring new insight into *Candida albicans* response upon macrophage interaction. *Molecular & cellular proteomics : MCP*. 2007;6(3):460-78. Epub 2006/12/14.
135. Ryckelynck M, Giege R, Frugier M. tRNAs and tRNA mimics as cornerstones of aminoacyl-tRNA synthetase regulations. *Biochimie*. 2005;87(9-10):835-45. Epub 2005/06/01.
136. Schillig R, Morschhauser J. Analysis of a fungus-specific transcription factor family, the *Candida albicans* zinc cluster proteins, by artificial activation. *Molecular microbiology*. 2013;89(5):1003-17. Epub 2013/07/13.
137. Inglis DO, Johnson AD. Ash1 protein, an asymmetrically localized transcriptional regulator, controls filamentous growth and virulence of *Candida albicans*. *Molecular and cellular biology*. 2002;22(24):8669-80. Epub 2002/11/26.
138. Elson SL, Noble SM, Solis NV, Filler SG, Johnson AD. An RNA transport system in *Candida albicans* regulates hyphal morphology and invasive growth. *PLoS genetics*. 2009;5(9):e1000664. Epub 2009/09/26.
139. Baek YU, Li M, Davis DA. *Candida albicans* ferric reductases are differentially regulated in response to distinct forms of iron limitation by the Rim101 and CBF transcription factors. *Eukaryotic cell*. 2008;7(7):1168-79. Epub 2008/05/27.
140. Yan H, Zhao Y, Jiang L. The putative transcription factor CaRtg3 is involved in tolerance to cations and antifungal drugs as well as serum-induced filamentation in *Candida albicans*. *FEMS yeast research*. 2014;14(4):614-23. Epub 2014/03/13.

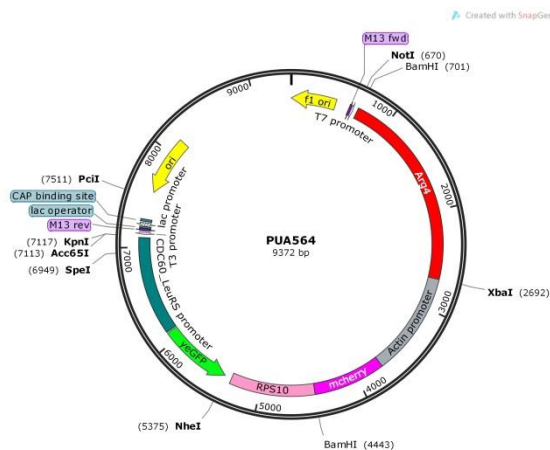
141. Vandeputte P, Ischer F, Sanglard D, Coste AT. In vivo systematic analysis of *Candida albicans* Zn2-Cys6 transcription factors mutants for mice organ colonization. *PloS one*. 2011;6(10):e26962. Epub 2011/11/11.

Annexes

A.1 – Plasmids described in the chapter II, previously constructed for quantification of SerRS and LeuRS expression in *C. albicans* strains.

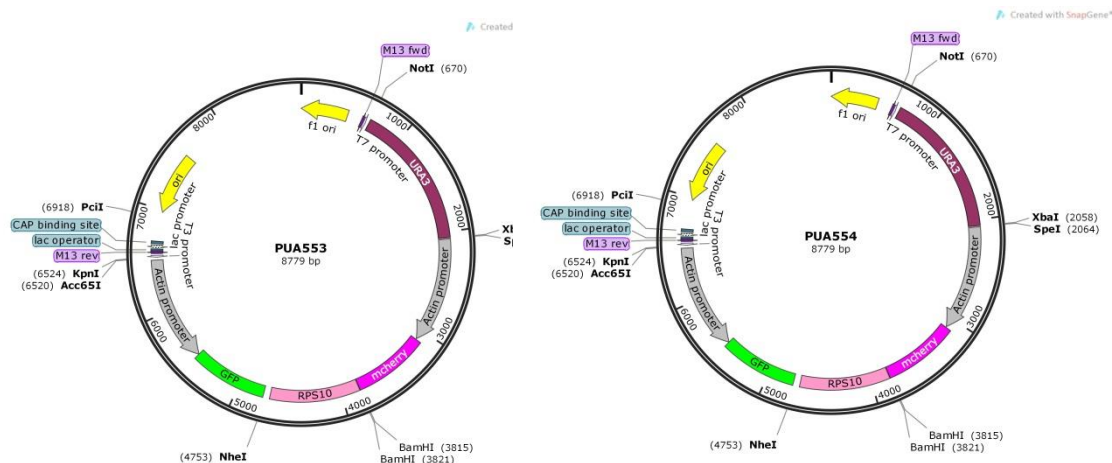


pUA563 – Plasmid containing the yEGFP reporter system associated with the SES1 promoter, allowing the determination of SerRS expression; used to transform WT and TF KO strains with ARG 4 as selective marker.



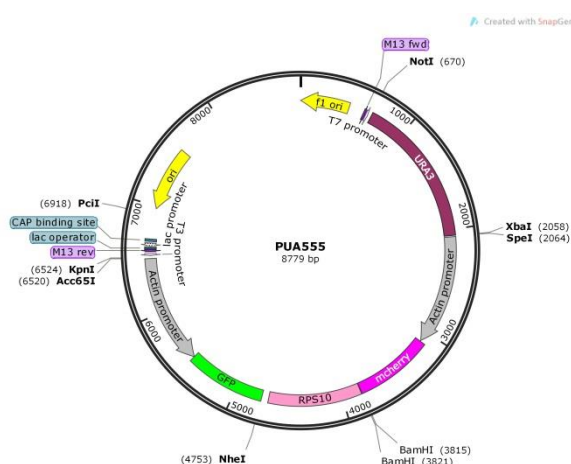
pUA564 - Plasmid containing the yEGFP reporter system associated with the CD60 promoter, allowing the determination of LeuRS expression; used to transform WT and TF KO strains with ARG 4 as selective marker.

A.2 – Plasmids described in the chapter II, previously constructed for quantification of Leu misincorporation in *C. albicans* strains.



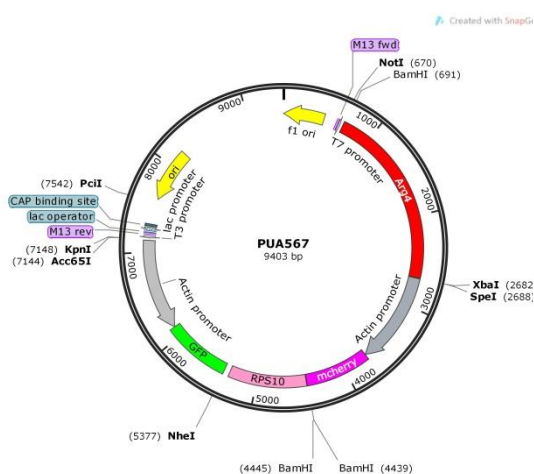
pUA 553 – Plasmid containing the yEGFP reporter system with a WT TTA-leucine at position 201; URA3 gene is present as a selective marker.

pUA 554 – Plasmid containing the yEGFP reporter system with a CTG ambiguous codon at position 201; URA3 gene is present as a selective marker.

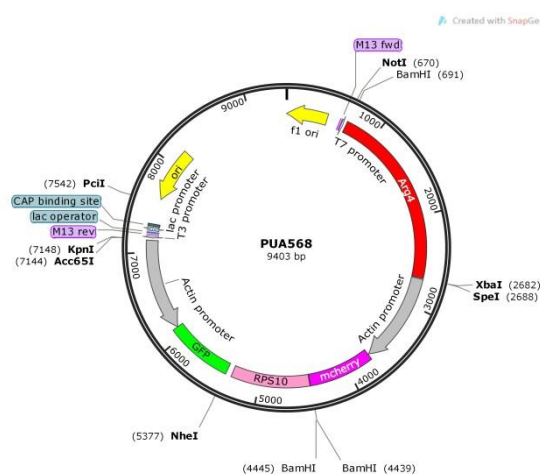


pUA 555 – Plasmid containing the yEGFP reporter system with a TCT-serine codon at position 201; URA3 gene is present as a selective marker

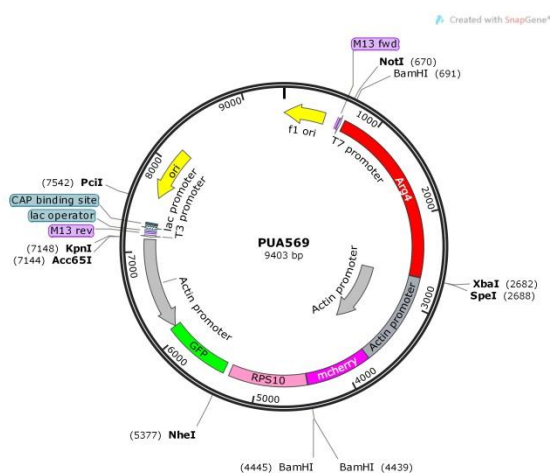
A.3 – Plasmids described in the chapter II, used for quantification of Leu misincorporation in *C. albicans* strains.



pUA567 – Plasmid derived from pUA553 containing an ARG4 gene as selective marker inserted between *SPEI* and *NOTI* restriction sites.



pUA 568 – Plasmid derived from pUA 554 containing an ARG4 gene as selective marker inserted between *SPEI* and *NOTI* restriction sites.



pUA 569 – Plasmid derived from pUA 555 containing an ARG4 gene as selective marker inserted between *SPEI* and *NOTI* restriction sites.