Universidade de Aveiro Departamento de Ciências Médicas
2017

Ana Sofia Paulo Varanda

Estudo das redes celulares que regulam o stress proteotóxico

Unravelling the cellular networks that regulate proteotoxic stress

Universidade de Aveiro Departamento de Ciências Médicas
2017

Ana Sofia Paulo Varanda

Estudo das redes celulares que regulam o stress proteotóxico

Unravelling the cellular networks that regulate proteotoxic stress

Tese apresentada à Universidade de Aveiro para cumprimento dos requisitos necessários à obtenção do grau de Doutor em Biomedicina realizada sob a orientação científica do Doutor Manuel António da Silva Santos, Professor Associado do Departamento de Ciências Médicas da Universidade de Aveiro e co-orientação da Doutora Gabriela Moura, Professora auxiliar convidada do Departamento de Ciências Médicas da Universidade de Aveiro e da Doutora Carla Oliveira, Professora associada da Faculdade de Medicina da Universidade do Porto.

Apoio financeiro da FCT e do FSE no âmbito do III Quadro Comunitário de Apoio.

“Sometimes my courage fails me and I think I ought to stop working, live in the country and devote myself to gardening. But I am held by a thousand bonds, and I don't know when I shall be able to arrange things otherwise. Nor do I know whether, even by writing scientific books, I could live without the laboratory.” Madame Curie in a Letter to her sister Bronya, September 1927

Dedico este trabalho aos meus pais.

o júri

presidente

Doutor Casimiro Adrião Pio
Professor Catedrático da Universidade de Aveiro

Doutor Manuel António da Silva Santos
Professor Associado da Universidade de Aveiro

Doutora Luisa Alejandra Helguero
Professora Auxiliar Convidada da Universidade de Aveiro

Doutor Henrique Manuel Paixão dos Santos Girão
Investigador FCT no Instituto Biomédico de Investigação de Luz e Imagem – IBILI
Universidade de Coimbra

Doutora Patrícia Joana Ferreira Morais Oliveira
Investigadora no Instituto de Investigação e Inovação em Saúde – I3S
Universidade do Porto

agradecimentos

Em primeiro lugar gostaria de agradecer ao meu orientador Professor Manuel Santos por ter acreditado em mim, pela oportunidade que me deu de fazer parte da sua equipa e por todos os ensinamentos ao longo destes anos.

À Professora Gabriela Moura, muito obrigada pelas discussões críticas ao meu trabalho.

Um agradecimento especial à Doutora Carla Oliveira por me ter acolhido no seu grupo no I3S, no Porto, pelo interesse no meu trabalho, ensinamentos e discussão de ideias.

Agradeço à Doutora Paula Gonçalves, Doutora Luisa Helguero e Doutora Daniela Ribeiro por me terem cedido espaço nas salas de cultura e criado condições para iniciar o meu projecto.

Ao Doutor Rui Vitorino, obrigada por ter acreditado na "mistranslation" e pela colaboração neste trabalho.

A todo o grupo do laboratório de Biologia do RNA de Aveiro, aos actuais e "ex-colegas", obrigada por todo o apoio, companheirismo, bom ambiente, alegria e acima de tudo por toda a paciência que tiveram de ter comigo! Quero agradecer especialmente às mentoras Ana Soares e Rita Bezerra, pelos conselhos, inspiração e boa música. À Rita, obrigada por estes cinco anos de amizade e cumplicidade. À Mafalda, um muitíssimo obrigado por tudo, pela partilha de conhecimentos, amizade, alegria, força e loucura com que sempre me fizeste enfrentar as adversidades.

A todo o grupo ERIC (I3S), obrigada pela boa disposição com que sempre me receberam. Um agradecimento especial à Patrícia Oliveira e Joana Carvalho, pelos momentos animados, paciência e colaboração neste projecto.

À FCT o meu muito obrigado por ter financiado este trabalho através de uma bolsa de Doutoramento, SFRH/BD/76417/2011 e do projecto FCT-ANR/IMI-MIC/0041/2012. Ao CESAM, ao iBiMED (UID/BIM/04501/2013) e ao I3S por terem proporcionado as condições necessárias à realização deste trabalho.

Às "meninas e meninos" do iBiMED, os "mini-cientistas", obrigada pelas discussões sobre ciência pela partilha de conhecimentos e sonhos. Foi um prazer trabalhar convosco no "open space"!

Aos Biomédicos, obrigada pela amizade ao longo destes dez anos. Em especial ao Igor e Tiago, pela alegria em forma de humor negro.

Aos Biólogos, obrigada pelos momentos de diversão e companheirismo, em especial à Vi, ao Tiago, à Dianinha e ao César.

agradecimentos (continuação)

Aos Zitos, um muito obrigado, em especial à Ana e Rute por tomarem conta de mim. À Raquel, Joãozinha e Jujuca, obrigada por colorirem a minha vida.

Às “roomies” Sara e Rocha, obrigado pelos óptimos momentos de liberdade.

À Cáti, Inês e Carina, obrigada pela amizade, carinho e por todos os sonhos tornados em viagens e aventuras.

Aos amigos da Covilhã, Fábio e Cassapo, obrigado por me trazerem sempre um gostinho de casa onde quer que esteja. À Ana, muito obrigada por estares sempre presente em tudo o que faço, pelas confidências, sábios conselhos e por alegrares a minha vida há mais de vinte anos.

Um agradecimento à minha família, em especial aos meus pais. Obrigada por me terem “solto” num laboratório quando era pequena. Obrigada por serem excelentes pais, professores e amigos. Obrigada por acreditarem sempre em mim, mas acima de tudo obrigado pela força e coragem com que me fazem encarar a vida.

palavras-chave

Incorporação de erros durante a síntese protéica; tRNA; Controlo de qualidade de proteínas; Stress proteotóxico; Evolução; Transcriptoma; Genoma

resumo

A manutenção da fidelidade durante o processo de síntese proteica é de extrema importância para a estabilidade das proteínas e o normal funcionamento dos processos celulares. Há erros que ocorrem durante este processo a uma taxa de 10^{-4}, a um nível fisiológico, sendo tolerados pelas células. Quando a taxa de erro aumenta os mecanismos de controlo de qualidade de proteínas ficam sobrecarregados, levando à acumulação de proteínas aberrantes que podem agregar. O envelhecimento e neurodegeneração têm sido associados ao aumento de erros durante a síntese de proteínas, mas até agora pouco se sabe acerca dos mecanismos causa-efeito envolvidos nestes processos.

De forma a entender como as células humanas se adaptam ao longo do tempo aos erros na síntese de proteínas, desenvolvemos uma metodologia que permite a incorporação errónea de aminoácidos de uma forma aleatória nas proteínas. Para tal, mutámos o anticodão de um tRNA de serina humano, de forma a incorpoporar serina em codões de outros aminoácidos. Para além disto, também sobre-expressámos o tRNA de Serina em células HEK293.

As células que expressam estes tRNAs, acumularam proteínas ubiquitinadas e em alguns casos agregados proteicos. Apesar disto, não se observaram grandes alterações na proliferação e viabilidade destas células após várias passagens em cultura. Os mecanismos de controlo de qualidade de proteínas, nomeadamente o sistema ubiquitina-proteasoma, o sistema de resposta a proteínas aberrantes no retículo endoplasmático e os chaperones moleculares, foram activados de forma diferencial de acordo com o tipo de aminoácidos substituídos nas proteínas. Observámos também que a evolução destas linhas em cultura levou a alterações ao nível da expressão génica, como a sobre-expressão de genes envolvidos na resposta ao stress do retículo endoplasmático e glicólise. A sequenciação do exoma das linhas celulares evoluídas permitiu-nos concluir que estas células acumularam poucas mutações comparativamente com a linha controlo. A maior parte dos genes mutados, ao longo do tempo, estão envolvidos em processos como a transcrição e ligação do RNA.

Os nossos dados indicam, assim, que as células HEK293 se adaptam aos erros que ocorrem durante a síntese proteica, activando mecanismos de controlo de qualidade de proteínas e alterando a expressão génica. A activação destes mecanismos é dependente do tipo de erros translacionais e da passagem em que as células se encontram. Este estudo revela-nos um pouco mais acerca de como as células humanas se adaptam aos erros que ocorrem durante a síntese proteica.

keywords

Protein synthesis errors; tRNA; Protein Quality Control; Proteotoxic Stress; Evolution; Transcriptome; Genome

abstract

Protein synthesis fidelity is essential for proteome stability and for functional maintenance of cellular processes. Errors occur at frequencies around 10^{-4} under normal physiological conditions and are tolerated by cells. If the frequency of these basal errors increases, the mechanisms of protein quality control become overload, leading to accumulation of misfolded proteins that may aggregate. Several conditions have been associated to this phenomenon, namely aging and neurodegeneration. But, the cause-effect mechanisms remain to be clarified in many cases.

In order to elucidate how human cells respond to the accumulation of protein synthesis errors (mistranslation) throughout time, we have developed a methodology that allows for the random serine misincorporation in proteins on a proteome wide scale. The anticodon of a human serine tRNA was mutated to read non-cognate codons. HEK293 cells expressing these tRNAs accumulate ubiquitinated proteins and in some cases protein aggregates. Despite this, only slight alterations in proliferation and viability were detected after several passages in culture. Here we show that activation of quality control pathways, namely molecular chaperones, ubiquitin-proteasome system and unfolded protein response was dependent on the type of amino acid misincorporation. In addition, we observed that evolution of these cell lines in culture deregulates gene expression, for example genes involved in endoplasmic reticulum stress and glycolysis were upregulated. Whole-exome sequencing of evolved cell lines revealed that our cells accumulate a small number of mutations comparatively to the control cell line, in genes involved in transcription and RNA binding.

Our data indicate that HEK293 cells adapt to protein synthesis errors mainly through activation of protein quality control mechanisms and other cellular processes depending on the type of amino acid substitution and cells passage number. This study reveals new insights on how human cells adapt to protein synthesis errors and the consequent accumulation of aberrant proteins.

Table of Contents

List of Figures **5**
List of Tables **9**
List of Abbreviations **11**

1. Introduction **13**

1.1 The Genetic Code: a set of rules for mRNA translation *14*
1.2 tRNAs and aaRSs establish the genetic code *15*
1.2.1 tRNAs 16
1.2.1 aaRS 18
1.3 Protein synthesis in the ribosomes *22*
1.3.1 Initiation 23
1.3.2 Elongation 26
1.3.3 Termination and Recycling 26
1.4 Maintenance of mRNA translational accuracy *29*
1.5 Errors during protein synthesis – mRNA mistranslation *30*
1.6 Protein misfolding *33*
1.7 Protein Quality Control *34*
1.7.1 Molecular Chaperones 35
1.7.2 The Ubiquitin-Proteasome System 36
1.7.3 Quality control in the ER 38
1.7.4 Autophagy 44
1.8 mRNA mistranslation and disease *45*
1.9 Mutant tRNAs as strategy to induce protein synthesis errors and proteotoxic stress *48*
1.10 Aims of the study *51*

2. Adaptation of human cells to protein synthesis errors **53**

2.1 Abstract *54*
2.2 Introduction *54*
2.3 Results *56*
2.3.1 Human cell line models of PSE 56
2.3.2 Expression and copy number of mutant tRNAs and $tRNA^{Ser}$ in HEK293 cells 58
2.3.3 Phenotypic effects of PSE 60
2.3.4 The impact of PSE in the ubiquitin-proteasome system and molecular chaperones 68
2.3.5 Effects of PSE in the UPR 77
2.3.6 Transcriptional deregulation induced by PSE 81
2.4 Discussion *84*
2.5 Conclusion *87*
2.6 Materials and Methods *87*

2.6.1 Cell culture 87
2.6.2 Construction of mutant tRNA plasmids 87
2.6.3 Generation of mistranslating cell lines 88
2.6.4 Evolution of cells in culture 89
2.6.5 Total RNA extraction 89
2.6.6 Quantification of tRNA expression and tDNA copy number 89
2.6.7 Cell fitness assessment 91
2.6.8 Protein synthesis determination 91
2.6.9 Quantification of the insoluble protein fraction 92
2.6.10 Quantification of proteasome activity 92
2.6.11 Immunoblots 93
2.6.12 Gene expression microarrays 94
2.6.13 Statistical analysis 94
2.7 Supplementary Figures *96*

3. Transcriptional alterations in mRNA mistranslating HEK239 cells 105

3.1 Abstract *106*
3.2 Introduction *106*
3.3 Results *108*
3.3.1 The impact of PSE in gene expression deregulation 108
3.3.2 Ontology of the genes deregulated in response to PSE 110
3.3.3 Long term gene expression deregulation in mistranslating cells 118
3.3.4 Gene Ontology categories deregulated during evolution 120
3.4 Discussion *126*
3.5 Conclusions *129*
3.6 Materials and Methods *130*
3.6.1 Cell culture 130
3.6.2 Mistranslating cell lines used 130
3.6.3 Total RNA extraction 130
3.6.4 Gene expression microarrays 131
3.6.5 Statistical analysis 131
3.6.6 Gene expression deregulation analysis 132

4. Genome alterations of HEK293 mistranslating cells 133

4.1 Abstract *134*
4.2 Introduction *134*
4.3 Results *136*
4.3.1 Whole-exome sequencing of evolved mistranslating HEK293 cells 136
4.3.2 Variants Lost, Shared and Gained 137
4.3.3 SNPs, INDELs and variants quality criteria 145
4.3.3 SNPs and Insertions annotated to genes 149

4.4 Discussion *154*
4.5 Conclusions *157*
4.6 Materials and Methods *157*
4.6.1 Cell Culture 157
4.6.2 Evolution of cells in culture 158
4.6.3 Mistranslating cell lines used 158
4.6.4 DNA Extraction 158
4.6.5 Exome Capture Nimblegen v3 and Illumina sequencing 158
4.6.6 Alignment and analysis of sequences 159
4.6.7 Gene Ontology analysis 159

5. General Discussion and Future Perspectives 161

5.1 General Discussion *162*
5.2 Future Perspectives *169*

6. References 171

7. Annexes 189

Annex A - Map of the plasmid 190
Annex B – List of genes with Transitions in mistranslating cell lines 191
Annex C – List of genes with Transversions in mistranslating cell lines 191
Annex D – List of genes with Insertions in mistranslating cell lines 192
Annex E – Paper (Manuscript): Mutant tRNAs are selected in tumors and accelerate tumor growth in mice ..... Erro! Marcador não definido.

List of Figures

Figure 1-1. The “Universal Genetic Code”. 15

Figure 1-2. tRNA structures. 17

Figure 1-3. The aminoacylation reaction..... 19

Figure 1-4. The two subclasses of aminoacyl-tRNA synthetases that evolved from two independent single-domain proteins..... 20

Figure 1-5. Pre-transfer and post-transfer editing of non-cognate amino acids by aaRSs. 21

Figure 1-6. The process of mRNA translation in eukaryotes showing the three main steps: Initiation, Elongation and Termination and the extra step: Recycling. 23

Figure 1-7. Scheme of the first step in eukaryotic translation – Initiation. 25

Figure 1-8. Scheme of mRNA elongation, termination and recycling in eukaryotes. 28

Figure 1-9. The two stages of quality control in the ribosome: Initial selection and Proofreading. 30

Figure 1-10. Consequences of protein synthesis errors in cells. 32

Figure 1-11. Mechanisms of adaptation to mRNA mistranslation. 33

Figure 1-12. The four major mechanisms of protein quality control. 35

Figure 1-13. The ubiquitin-proteasome system (UPS). 38

Figure 1-14. ER- associated degradation (ERAD). 39

Figure 1-15. The three UPR sensors in mammalian cells. 42

Figure 1-16. Autophagic pathways..... 45

Figure 1-17. Schematic representation of “chimeric” tRNAs that mutate the proteome. 49

Figure 1-18. Fluorescence recovery of GFP coded by a gene where serine-65 had been substituted by 10 codons recognized by the engineered tRNAs 50

Figure 2-1. Representation of the human $tRNA_{AGA}^{Ser}$ and mutant tRNAs used in the study..... 57

Figure 2-2. Schematic representation of PSE incorporation by the mutant $tRNA_{AAG}^{Ser}$ 57

Figure 2-3. Quantification of $tRNA^{Ser}$ and mutant tRNAs expression. 59

Figure 2-4. Copy number of $tRNA^{Ser}$ and mutant tRNAs. 60

Figure 2-5. Doubling time of cells, assessed by cell counting with Tripan blue. 61

Figure 2-6. Percentage of viable cells in culture determined by cell counting with Tripan blue. 62

Figure 2-7. Relative cell proliferation determined using a BrdU ELISA Kit. 63

Figure 2-8. Number of colonies formed after 12 days in culture. 63

Figure 2-9. Protein synthesis rate determined by SunSET method adapted to immunoblot with anti-puromycin. 65

Figure 2-10. Relative insoluble protein fraction 67

Figure 2-11. Relative protein ubiquitination determined by immunoblot. 69

Figure 2-12. Relative proteasome activity. 70

Figure 2-13. Relative HSP70 expression. 71

Figure 2-14. Relative Hsp27 expression. 73

Figure 2-15. Relative Hsp60 expression. 74

Figure 2-16. Relative Hsp90α expression. 75

Figure 2-17. Relative BiP expression. 76

Figure 2-18. Relative ATF6f/ATF6t expression 78

Figure 2-19. Relative eIF2αP/eIF2αt expression. 79

Figure 2-20. Relative GADD34 expression. 80

Figure 2-21. Summary of the PQC alterations identified in the different cell lines. 86

Figure 2-22. tRNA amplification with primers for the plasmid and Sanger sequencing. 96

Figure 2-23. SNaPshot analysis – Agarose gels representative of the amplification of tRNA and GAPDH. 97

Figure 2-24. SNaPshot peaks of the endogenous $tRNA^{Ser}$, mutant tRNAs and the respective control, GAPDH. Sequenced samples correspond to cDNA from P1. 98

Figure 2-25. SNaPshot peaks of the endogenous $tRNA^{Ser}$, mutant tRNAs and the respective control, GAPDH. Sequenced samples correspond to cDNA from P15. 99

Figure 2-26. SNaPshot peak of the endogenous $tRNA^{Ser}$, mutant tRNAs and the respective control, GAPDH. Sequenced samples correspond to cDNA from P30. 100

Figure 2-27. SNaPshot peaks of the endogenous $tRNA^{Ser}$, mutant tRNAs and the respective control, GAPDH. Sequenced samples correspond to DNA from P1. 101

Figure 2-28. SNaPshot peaks of the endogenous $tRNA^{Ser}$, mutant tRNAs and the respective control, GAPDH. Sequenced samples correspond to DNA from P15. 102

Figure 2-29. SNaPshot peaks of the endogenous $tRNA^{Ser}$, mutant tRNAs and the respective control, GAPDH. Sequenced samples correspond to DNA from P30. 103

Figure 3-1. Percentage of deregulated genes for each cell line comparatively with the Mock in each time point 110

Figure 3-2. Functional class enrichment analysis (GO categories) of genes upregulated (red bars) and downregulated (green bars) in mistranslating cell lines in P1. 113

Figure 3-3. Functional class enrichment analysis (GO categories) of genes upregulated (red bars) and downregulated (green bars) in mistranslating cell lines in P15. 115

Figure 3-4. Functional class enrichment analysis (GO categories) of genes upregulated (red bars) and downregulated (green bars) in mistranslating cell lines in P30. 117

Figure 3-5. Diagram showing the experimental design used to obtain the list of $tRNA^{Ser}$(S) genes used in the transcriptional analysis of cell evolution. 119

Figure 3-6. The percentage of deregulated genes in each cell line between passages. 120

Figure 4-1. Number of variants (unique variants, Q≥20, GQ>20) in each cell line (vs. *Homo sapiens*) normalized to the total number of mapped reads in each cell line (x1000). 138

Figure 4-2. Number of variants (unique variants, Q≥20, GQ>20, number of reads≥20) in each cell line vs. Mock, categorized as “Shared”, “Lost” and “Gained”..... 139

Figure 4-3. Number of variants (unique variants, Q≥20, GQ>20, number of reads≥40) in each cell line relative to Mock, normalized to the total number of mapped reads for each cell line (x10000). 141

Figure 4-4. Number of Lost and Gained HQ variants (unique variants, Q≥100, GQ≥90, with a number of reads≥40) in each cell line comparatively to Mock, normalized to the total number of mapped reads for each cell line (x10000). 142

Figure 4-5. Ratio of Gained/Lost HQ variants (unique variants, Q≥100, GQ≥90, with a number of reads≥40) in each cell line comparatively to Mock. 143

Figure 4-6. Ratio of Gained/Lost HQ variants (unique variants, Q≥100, GQ≥90, number of reads≥40) per chromosome, in each cell line, comparatively to Mock. 144

Figure 4-7. Representative scheme of Transitions and Transversions 145

Figure 4-8. Number of Transitions and Transversions (unique variants, Q≥100, GQ≥90, number of reads≥40), for each cell line, relative to Mock, normalized to the total number of mapped reads for each cell line (x10000). 147

Figure 4-9. Number of variants with no criteria (columns) or with application of the maximum quality criteria (unique variants, Q≥100, GQ≥90, number of reads≥40) (dots) that were Gained, Shared or Lost for each cell line relative to Mock, normalized to the total number of mapped reads for each cell line (x10000). 148

Figure 5-1. Summary of the main alterations detected during evolution of cells expressing the Wt $tRNA^{Ser}$ and mistranslating cell lines 168

List of Tables

Table 2-1. Deregulation of PQC genes induced by protein synthesis errors. 83

Table 3-1. Summary of the cellular processes upregulated (red arrow) and downregulated (green arrow) in the mistranslating cell lines in the three passages. 118

Table 3-2. GO categories upregulated in mistranslating cell lines during evolution in culture. 122

Table 3-3. GO Categories downregulated in mistranslating cell lines during evolution 123

Table 3-4. GO categories that are deregulated in mistranslating cell lines between P1 and P15 and at later stages (between P15 and P30). 124

Table 3-5. GO categories upregulated in mistranslating cell lines during evolution in culture (P1-P30)..... 125

Table 3-6. GO categories downregulated in mistranslating cell lines during evolution in culture (P1-P30)..... 125

Table 4-1. Number of variants in each cell line (vs. Homo sapiens) with no criteria and with the application of the first set of quality criteria..... 137

Table 4-2. Number of variants Shared, Lost or Gained relative to Mock, with a minimum of 20 or 40 reads (non-normalized values). 140

Table 4-3. Number of high quality (HQ) variants Lost or Gained in comparison with Mock, with a minimum of 40 reads or with the minimum of 40 reads with the application of the new quality criteria (Q≥100, GQ≥90) (non-normalized values). 142

Table 4-4. Number of variants (Transitions or Transversions) for each cell line, using different quality criteria..... 146

Table 4-5. Number of annotated and identified (DAVID Bioinformatic Resources 6.7) genes in our cell lines that have Transitions, Transversions or Insertions in their sequence. 149

Table 4-6. Examples of genes with Transitions in mistranslating cell lines. 150

Table 4-7. Example of genes with Transversions in mistranslating cell lines. 152

Table 4-8. Example of genes with Insertions in mistranslating cell lines. 154

List of Abbreviations

A	adenosine
aa	amino acid
aaRS	aminoacyl-tRNA synthetase
aa-tRNA	aminoacyl-tRNA
Ala	alanine
AMP	adenosine monophosphate
Arg	arginine
Asn	asparagine
Asp	aspartic acid
ATF4	activating transcription factor 4
ATF6	activating transcription factor 6
ATG	autophagy-related genes
ATP	adenosine triphosphate
BCL-2	B-cell lymphoma 2
BiP	binding immunoglobulin protein
C	cytosine
CHIP	carboxy terminus of Hsc70 interacting protein
CHOP	C/EBP homologous protein
CMA	chaperone-mediated autophagy
Cys	cysteine
D	dihydro-uridine
DNA	deoxyribonucleic acid
DUB	deubiquinating enzyme
E1	ubiquitin-activating enzyme
E2	ubiquitin-conjugating enzyme
E3	ubiquitin-ligase enzyme
eEF	eukaryotic elongation factor
eIF	eukaryotic initiation factor
ER	endoplasmic reticulum
ERAD	endoplasmic reticulum – associated degradation
eRF	eukaryotic release factor
G	guanine
GADD34	growth arrest and DNA damage-inducible protein
GCN2	general control nonderepressible 2
GDP	guanosine diphosphate
Gln	glutamine
Glu	glutamic acid
Gly	glycine
GTP	guanosine thriphosphate
His	histidine
HSP	heat shock protein
I	inosine
Ile	isoleucine
IRE-1	inositol-requiring enzyme 1
JNK	JUN N-terminal kinase
Leu	leucine
Lys	lysine
Met	methionine
mRNA	messenger ribonucleic acid

N	nucleotide
NH_2	amino group
OH	hydroxide
ORF	open reading frame
PABP	poly(A)-binding protein
PERK	protein kinase RNA-like endoplasmic reticulum kinase
Phe	phenilalanine
PPi	inorganic pyrophosphate
PQC	protein quality control
Pro	proline
RIDD	regulated IRE1-dependent decay of mRNA
RNA	ribonucleic acid
ROS	reactive oxygen species
rRNA	ribosomal ribonucleic acid
Ser	serine
sHSP	small heat shock proteins
T	tymine
Thr	threonine
tRNA	transfer ribonucleic acid
tRNAi	initiator transfer ribonucleic acid
Trp	tryptophan
Tyr	tyrosine
U	uridine
UBD	specific ubiquitin binding domains
UPR	unfolded protein response
UPS	ubiquitin-proteasome system
UTR	untranslated region
Val	valine
XBP1	X-box-binding protein 1
mTOR	mammalian target of rapamycin
AMPK	AMP-activated protein kinase
mt-tRNA	mitochondrial tRNAs

Other abbreviations will be explained when used in the text.

1. Introduction

1.1 The Genetic Code: a set of rules for mRNA translation

The genome stores the genetic information of organisms in the form of deoxyribonucleic acid (DNA). During cell division, daughter cells acquire identical copies of DNA of the mother cell, through the process of DNA replication. In this way, genetic information is faithfully transmitted to the next generation. Genomic information flows from DNA to messenger ribonucleic acid (mRNA), through the process of transcription and to proteins though protein synthesis (translation) (1–3). This flow of genetic information was incorporated in the "Central Dogma of Molecular Biology" by Francis Crick in 1958. Genomic studies in recent years unveiled many exceptions to this dogma, but the "Central Dogma of Molecular Biology" is still a framework for understanding the expression of genetic information in living organisms (3).

The genetic code establishes the rules that govern the transfer of genetic information from nucleic acids to proteins and is based on nucleotide triplets (codons) that do not overlap along the mRNA (Figure 1-1). The 4 ribonucleotides (adenosine – A, guanine – G, uridine – U and cyclidine – C) form 64 possible codons (4^3): 61 specify individual amino acids and 3 stop codons (UAA, UGA and UAG), which are used to terminate RNA translation. Some amino acids can be specified by only one codon, which is the case of methionine (Met) and tryptophan (Trp) (AUG and UGG respectively), but most amino acids are specified by more than one codon. For example, leucine (Leu), serine (Ser) and arginine (Arg) are specified by six different codons, called synonymous codons. The genetic code is, therefore, degenerated (2,4).

Synonymous codons are usually clustered in boxes rather than being randomly distributed throughout the code (Figure 1-1). Amino acids with two-codon boxes, i.e. tyrosine (Tyr) or four-codon boxes, i.e. glycine (Gly), differ only in the third position. Arg and Leu are exceptions, since they are specified by six codons, a four-codon box and two-codon box that differ in the first position of the code (4).

Codons are decoded by transfer RNAs (tRNAs) in the ribosome. In general, organisms contain fewer tRNAs codons. Since wobble decoding (5) allows the first base in the first position of the anticodon to interact with more than one nucleotide in the third position of the codon. Also, according to the wobble hypothesis, pairing between codon and anticodon at the first two codon positions follows the Watson-Crick base pairing, but the last nucleotide of the codon can move slightly from its normal position to form a non-Watson-Crick base pair with

the anticodon. This explains how multiple codons that specify the same amino acid often differ at the third base position only (2,4).

The distribution of amino acids in the genetic code is also biased toward amino acids polar properties. All codons with U in the second position specify amino acids with hydrophobic side-chains, while codons with A in the second position specify amino acids with polar side-chains (Figure 1-1) (6). Therefore, the genetic code is optimized to minimize the impact of mRNA decoding errors.

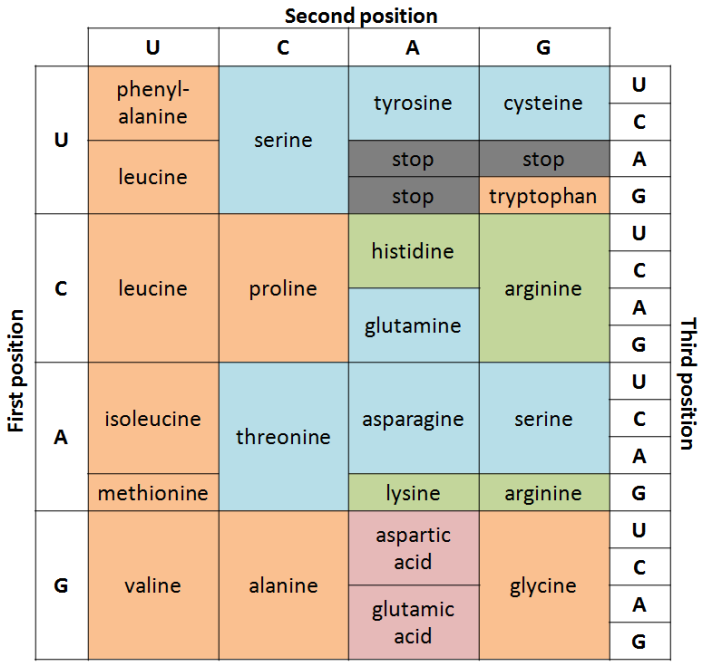


Figure 1-1. The "Universal Genetic Code". Colors in the figure indicate the chemical properties of the side chain of the amino acids. Orange indicates hydrophobic and the rest are hydrophilic aminoacids. Blue indicates polar, green basic and pink acidic amino acids. U – uridine; C – cyclidine; A – adenosine; G – guanine; Stop – stop codon [adapted from (6,7)].

1.2 tRNAs and aaRSs establish the genetic code

The accurate decoding of the four-nucleotide of the mRNA language requires two important elements: transfer RNAs (tRNAs) and aminoacyl-tRNA synthetases (aaRSs). To produce proteins, cells link amino acids to tRNAs using a group of 20 enzymes called aaRSs. This reaction results in the formation of aminoacyl-tRNAs (aa-tRNAs). The anticodon of aa-tRNA then pairs with codons in mRNA, so that the activated amino acid can be added to the growing polypeptide chain in the ribosome. It is this specific attachment of a given amino acid to a particular tRNA that establishes the genetic code (1,8,9).

1.2.1 tRNAs

tRNAs are small, ubiquitous RNAs that function as "adapter" molecules matching codons to particular amino acids. They have two crucial properties: each tRNA accepts a single amino acid in its 3'- end and contains a trinucleotide, the anticodon, complementary to the codon in the mRNA assigned to that amino acid (1,9).

In animal and plant cells there are between 50-100 different tRNAs. These numbers differ from the numbers of amino acids, because most amino acids have more than one tRNA to which they can attach. It also differs from the number of codons in the genome, because many tRNAs can pair to more than one codon (1). tRNAs with different anticodons that accept the same amino acid are called isoacceptors and their number varies between organisms. For example, minimal genomes of some bacteria have only one tRNA for each amino acid, while human cells have 49 different tRNAs isoacceptors (10,11).

tRNAs are ~ 76 nucleotides long and, despite differences in the primary structure, all have a common "cloverleaf" secondary structure (Figure 1-2) (1,11,12). This structure is characterized by three stem loops (formed by base pairing complementarity) and four helices. The acceptor stem helix is formed by seven base pairs combining the 5'and 3'- end of the molecule. At the 3'-end, there is a conserved CCA sequence, that carries the aminoacyl residue via an ester link between the ribose 2'- or 3'-hydroxil (OH) group and the carboxyl group of the amino acid. The D stem loop, is formed by a helical region of 3-4 base pairs and a loop with 8-11 nucleotides. The loop contains two hydrouridine bases, hence the name dihydrouridine-stem-loop. The anticodon stem loop located in the opposite end to the acceptor stem, contains the anticodon triplet in the center of its loop. This loop has a universal length of 7 nucleotides with a consensus sequence Py_{32}-U_{33}-XYZ-Pu(modified)-N_{38}, where Py represents pyrimidine, U uridine (conserved in all tRNAs), XYZ the anticodon, Pu purine and N any nucleotide. The TψC stem loop also contains 7 nucleotides and the TψC sequence, where T is thymine, ψ pseudouridine and C cytosine. In addition to these defined stems, there is also a variable region between the TψC stem and the anticodon stem, whose length can vary from 4 to 24 nucleotides (1,11).

tRNAs can be divided in two classes (Class I and II) according to the length of the variable loop. Class I tRNAs have a short loop, 4-5 nucleotides (most tRNAs), while Class II tRNAs have a long extra loop with at least 11 nucleotides. This class includes Leu and Ser tRNAs in eukaryotes and Tyr, Leu and Ser tRNAs in prokaryotes. The long loop may be

involved in the recognition of these tRNAs by their cognate aaRSs or in discrimination against non-cognate synthetases (11,13).

In all tRNAs, the cloverleaf structure forms a tertiary L-shaped structure created by hydrogen bonding (Figure 1-2). In the L-shaped structure there are two functional domains with independent origins, namely the acceptor stem, at one end of the tRNA and the anticodon stem, at the other end of the structure. This structure is consistent with the role of tRNAs in protein synthesis as it allows the codon/anticodon interaction at the decoding center of the small ribosomal subunit and contacts of the aminoacyl-(peptidyl) residue with the peptidyl-transferase center on the large ribosomal subunit (1,11,13).

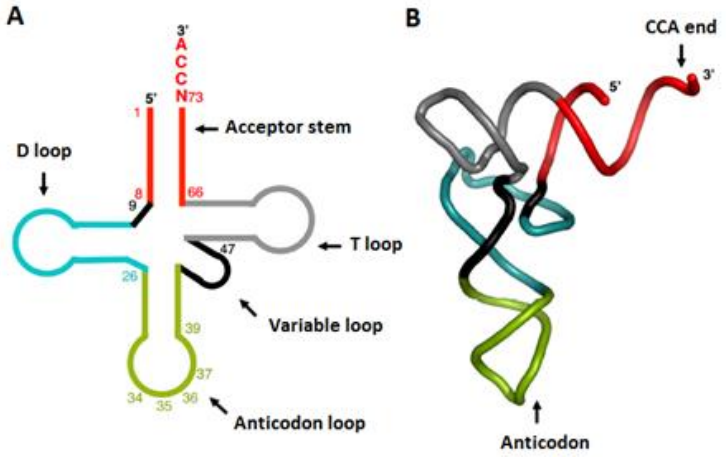


Figure 1-2. tRNA structures. A – tRNA secondary structure (cloverleaf). The acceptor stem is represented in red and has the CCA end and the 3'- terminal adenosine that accepts amino acids. Blue representes the D-loop, green the anticodon loop, black the variable loop and grey the T-loop. **B –** tRNA tertiary structure (L-shape). The cloverleaf secondary structure folds to the L-shaped tertiary structure. Both figures have corresponding colors [adapted from (12)].

aaRS recognize their cognate tRNAs through a limited number of nucleotide residues called identity determinants, which are commonly located in the tRNA anticodon, the acceptor stem and the associated "discriminator" base at position 73. Also, the variable stem can contain identity elements for tRNA recognition. Each tRNA family has its own discriminator base. For example, all Ser tRNAs have G_{73} and all Leu tRNAs have A_{73} discriminator bases. Along with tRNA identity determinants, the tRNA recognition system may also involve identity anti-determinants which are responsible for preventing their recognition by non-cognate aaRS (13,14).

tRNAs are heavily modified during their maturation process. Over 100 known post-transcriptional modifications have been reported, as well as, the enzymes responsible for catalyzing them. The number of modified bases varies among individual tRNA types, for example tRNAs of *S. cerevisiae* have an average of 11 modifications per tRNA, while mammalian tRNAs have an average of 13 to 14 modifications (15,16). Many examples of tRNA modifications include ribose/base methylations (Gm, Cm/m^5C), base isomerization (U to pseudouridine ψ), base reduction (U to D; dihydro-uridine), base thiolation (s^2C, s_2U, s^4U) and base deamination (inosine). Some modifications are conserved in all tRNAs, for example, dihydro-uridine (D) residues in the D-stem or pseudouridine in the T-stem (11,17).

Modifications in tRNAs are crucial for tRNA structure, function and stability and hypomodified tRNAs are targeted for degradation (18). Specific modifications in the anticodon loop affect directly translation efficiency. For example, modified bases at position 37 (adjacent to the anticodon) strengthen codon-anticodon interactions and prevent translational frameshifting (16,19). One example of a modified residue at position 37 is the 1-methylguanosine (m^1G) in tRNAs that decode codons starting with C (16,20). On the other hand, modifications at position 34, for example adenosine-to-inosine (A-to-I), are usually necessary for codon-anticodon wobbling to occur. These modified nucleosides in the anticodon loop and also in other positions can influence tRNA recognition of by aaRS and also function as identity or anti-identity determinants (16,18).

In addition to the roles in translation, recent studies unveil new functions for tRNAs. For example, mitochondrial tRNAs may be involved in the control of apoptosis through binding to cytocrome *c*, promoting cell survival (21). Also, specific tRNAs genes can act as insulators in the human genome where they help separate actively transcribed chromatin domains from silent ones (17,22).

1.2.1 aaRS

aaRS are responsible for the aminoacylation of tRNAs with cognate amino acids, thus establishing the amino acid – trinucleotide relationships of the genetic code (11). Each aaRS recognizes and activates a single amino acid through a universally conserved two-step mechanism. In the first step (1) the amino acid (aa) is condensed with adenosine triphosphate (ATP) forming a tightly bound aminoacyl-adenylate and releasing inorganic pyrophosphate (PPi). In the second step (2) the activated amino acid is transferred to the 3'- end of the tRNA,

originating an aminoacyl-tRNA (aa-tRNA) and releasing adenosine monophosphate (AMP) (Figure 1-3) (23,24).

(1) E+aa+ATP→E(aa-AMP)+PPi

(2) E(aa-AMP)+tRNA→E+aa-tRNA+AMP

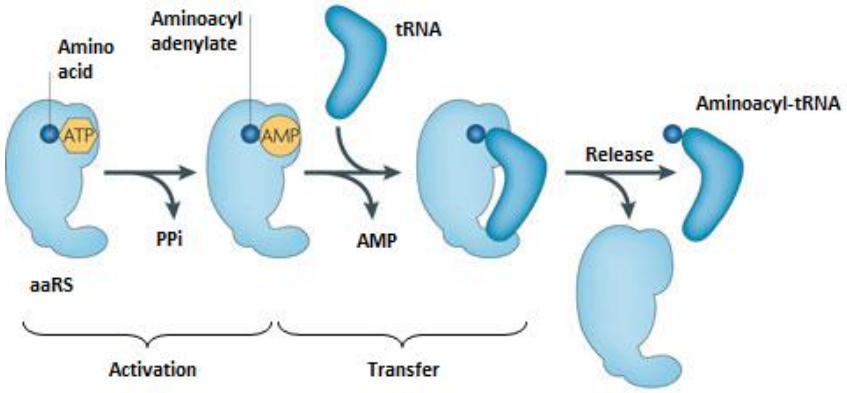


Figure 1-3. The aminoacylation reaction. In the first step there is activation of the amino acid at the aaRS active site with ATP to form aminoacyl-adenylate. The amino acid is then transferred to the 3'-end of the tRNA. ATP – adenosine triphosphate; AMP– adenosine monophosphate; PPi – pyrophosphate [adapted from (24)].

aaRS are ancient enzymes, ubiquitously expressed and universally distributed across the phylogenetic three suggesting that they are among the oldest peptide families (23,25,26). All cells contain at least 20 synthetases (eukaryotic cellular organelles use an additional set of synthetases), one for each of the 20 amino acids of the genetic code (11). These 20 aaRS are divided in two classes (Class I and Class II), each class containing 10 enzymes. This division is based on the differences in the structural topology of their active sites. All the enzymes in each class evolved from one unique single-domain protein and there is no evidence for the existence of a common ancestor of the two classes (11,25,27).

Class I aaRSs are generally monomeric and have in their catalytic domain a Rossman nucleotide binding fold composed of alternating β strands and α helices. Due to their structure, Class I aaRSs, approach tRNA molecules from the minor groove of the tRNA acceptor stem and aminoacylate the terminal adenosine (of the CCA-3' terminal acceptor stem) at the 2'-OH position. Class II aaRSs are typically multimeric enzymes and their active sites contain seven-stranded antiparallel β fold with flanking α helices. Enzymes present in Class II, approach

tRNAs from the major groove and charge the terminal adenosine at the 3'-OH position, based on studies carried out in *E. coli* (11,23,25,28).

The aaRS classes can be also divided into three subclasses, namely Ia,b,c (Class I) and IIa,b,c (Class II). Each subclass is thought to have evolved from the progenitor of the main class. The enzymes in the same subclass show a tendency to recognize amino acids that are chemically related. Subclasses Ia and IIa catalyze aminoacylation reactions of many of the hydrophobic amino acids, while subclasses Ib and IIb capture the carboxyl side-chain amino acids and the amidated (NH_2) derivates. Subclasses Ic and IIc catalyze aminoacylation reactions for aromatic amino acids (Figure 1-4) (25).

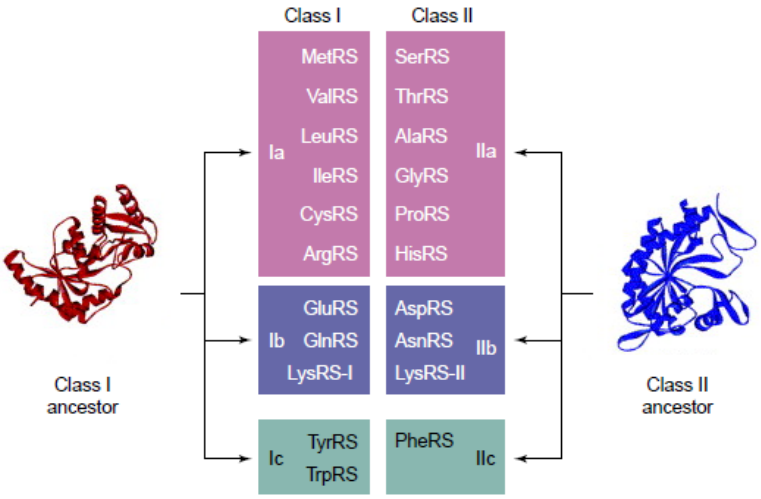


Figure 1-4. The two subclasses of aminoacyl-tRNA synthetases that evolved from two independent single-domain proteins. The ancestor three dimensional active-site domain is represented for each class [from (25)].

To ensure accurate aminoacylation, aaRSs have specific determinants that are crucial to recognize the cognate tRNAs. Generally, the conserved catalytic domain contains regions proximal to the active site responsible for binding of the acceptor stem, whereas the anticodon (when recognized) is bound by a separate divergent domain. Studies in *E. coli* revealed that the Rossman fold contains an acceptor binding domain composed of structural elements that specifically interact with a number of positions in the acceptor stem of the glutamine tRNA ($tRNA^{Gln}$). One example is the Leu-136 residue, where mutations in this site result in generally relaxed tRNA specificity leading to the mischarging of non cognate tRNAs (29,30).

aaRSs display an overall error rate of about 1 in 10,000 (10^{-4}). This specificity is achieved, as described above, by the contacts between aaRSs and tRNAs, but also between aaRSs and amino acids. The discrimination by aaRSs of amino acids is potentially more problematic, since amino acids are considerably less complex than tRNA structures. To make sure that the correct amino acid is chosen, aaRSs possess a double-sieve mechanism. The active site serves as the first sieve, where the discrimination of amino acids based on size (larger amino acids are excluded) and chemical properties occurs. Despite this, misacylation can occur when the cognate amino acid displays high structural similarity to other non-cognate amino acid (for example the amino acids that differ by a single methyl group; valine (Val) and isoleucine (Ile)). To overcome this problem, some aaRSs possess an intrinsic proofreading activity in the editing site, which is the second sieve (Figure 1-5) (31). In the editing site, misactivated or mischarged amino acids are hydrolyzed, whereas cognate amino acids are rejected on the basis of size or hydrophobicity. Editing can potentially occur either before (pre-transfer editing) or after (post-transfer editing) the misactivated amino acid is attached to the tRNA. Pre-transfer editing involves the hydrolysis of the misactivated aminoacyl-adenylate, aa-AMP, and post-transfer editing involves the hydrolysis of the aa-tRNA. All aaRSs have the capacity to hydrolyze non-cognate aa-AMPs (pre-transfer editing), but not all deacylate aa-tRNA (post-transfer editing), as methionyl-tRNA synthetase (MetRS), lysyl-tRNA synthetase (LysRS) and seryl-tRNA synthetase (SerRS) (31,32).

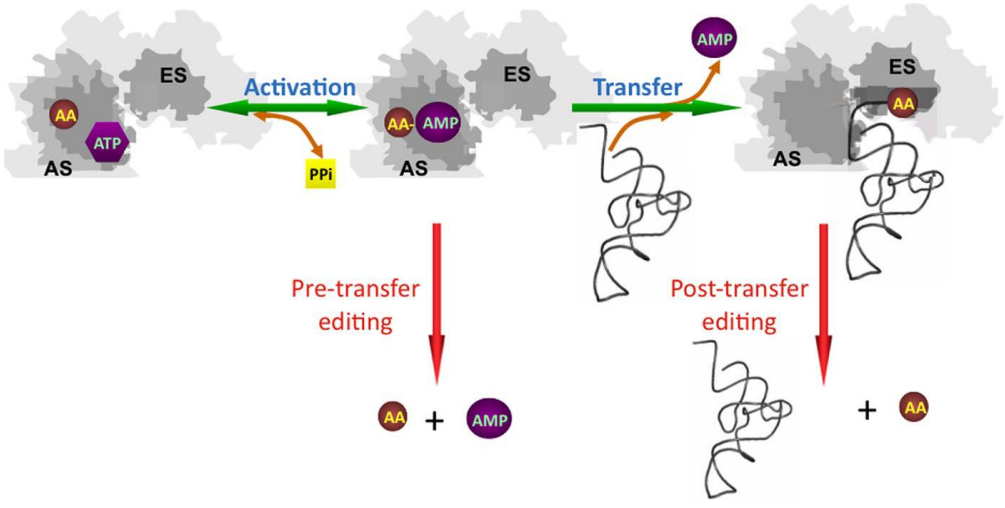


Figure 1-5. Pre-transfer and post-transfer editing of non-cognate amino acids by aaRSs. In pre-transfer editing AA-AMP is hydrolyzed directly, preventing the synthesis of a non-cognate aa-tRNA. In post-transfer editing the aa-tRNA is translocated to the editing site where the AA is removed. AS – active site; AA – amino acid; ATP – adenosine thriphosphate; ES – editing site; PPi – pyrophosphate; AMP – adenosine monophosphate [from (31)].

Editing decreases the frequency of errors during protein synthesis and is important for translational quality control, being considered an essential function (33,34). When this activity is compromised, a wrong amino acid can be incorporated into proteins (23). Several studies in bacteria, human cell cultures and mice suggest that defects in editing are toxic and can be associated to serious pathologies (33).

In addition to protein synthesis, aaRS have long been known to participate in other cellular processes. They have been implicated in the regulation of expression of their own genes (30). In both bacteria and eukaryotes, mRNAs of certain aaRSs encode short sequences that fold into the cloverleaf structure to mimic cognate tRNAs. aaRS in excess can bind to these "tRNA-mimicking elements" blocking translation of the aaRS mRNA, keeping the amount of specific aaRSs (35). Similar regulation occurs at the level of transcription. Other functions of aaRS include mediation of glucose and amino acid metabolism, regulation of the development of specific organs and tissues, triggering or silencing inflammatory responses and the control of cell death and stress responses that may lead to tumorigenesis (36–39).

1.3 Protein synthesis in the ribosomes

mRNA translation occurs in ribosomes that work in conjugation with tRNAs, amino acids, translational factors and aaRSs, in order to interpret mRNA information and catalyze the formation of peptide bonds (2). Ribosomes contain several different ribosomal RNA (rRNA) molecules and more than 50 proteins. In eukaryotes, ribosomes are composed by a small subunit (40S) and a large subunit (60S), according to their sedimentation coefficients in Svedberg units (S). The small subunit contains one molecule of 18S rRNA and 33 proteins while the large subunit contains three rRNAs (28S, 5.8S and 5S) and 47 proteins. Both subunits contain three tRNA binding sites: the aminoacyl site (A site), the peptidyl site (P site), and the exit site (E site) (1,40). The A site accepts the aa-tRNA, the P site holds the tRNA that is bound to the nascent polypeptide chain and the E site holds the deacylated tRNA before it leaves the ribosome (1).

There are three main steps in ribosomal translation: Initiation, Elongation and Termination. An extra step consists of ribosome recycling. During Initiation the 40S subunit joins the mRNA and searches for the initiation codon (AUG), using the initiator tRNA (tRNAi) as the searching tool. During elongation aa-tRNAs enter in the A site of the ribosome, where decoding takes place through base complementarity between anticodons of acylated tRNA and codons in the mRNA. The ribosome then translocates through the mRNA and catalyzes peptide

bonds between amino acids, forming a polypeptide chain. Termination occurs when the ribosome encounters a stop codon, which leads to dissociation of the newly synthesized polypeptide chain from the ribosome. In the recycling step, ribosomal subunits are dissociated, mRNA and deacylated tRNA are released and the components can be recycled for initiation of another round of translation (Figure 1-6) (41).

The process of translation is conserved among of living organisms, despite some differences between bacteria and eukaryotes (1,42). For the purpose of this work I will focus on the process of eukaryotic translation.

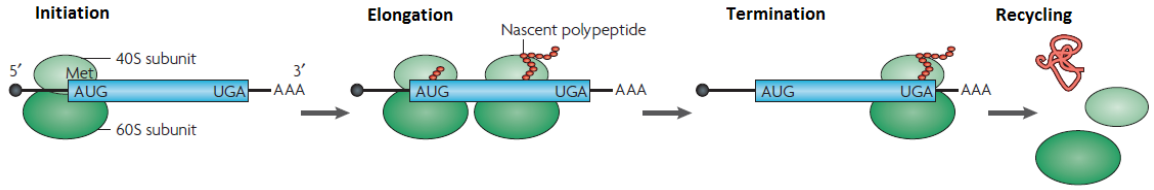


Figure 1-6. The process of mRNA translation in eukaryotes showing the three main steps: Initiation, Elongation and Termination and the extra step: Recycling. During initiation 40S subunit joins mRNA and searches AUG codon using the eukaryotic initiator tRNA (tRNAMet). During the elongation amino acids are added to the growing peptide chain. When ribosome encounters a stop codon (UAG) translation terminates. Ribosomes and other translational components dissociate to be recycled and used in another round of translation. In blue is represented the open reading frame (ORF) [adapted from (41)].

1.3.1 Initiation

Before initiation of mRNA translation two events must occur. Firstly, the amino acid is attached to the tRNAs originating aa-tRNAs. Secondly, free ribosomes must dissociate in the two ribosomal subunits after each round of translation. eIF3 and probably eIF1 and eIF1A are three eukaryotic factors that participate in this process (Figure 1-7) (42,43).

The first step of eukaryotic initiation is the assembly of the ternary complex eIF2.GTP.Met-tRNAi. The eIF2 binds GTP which is hydrolyzed to GDP during initiation. Because eIF2 has high affinity for GDP, another factor, eIF2B is needed to recycle GDP to GTP. The ternary complex binds to the small subunit 40S with the help of eIF1, eIF1A, eIF3 and probably eIF5 originating the 43S pre-initiation complex (42–44).

Natural 5'UTR possess secondary structures that require the cooperation of some factors to unwind the 5'cap of mRNA and allow ribosomal attachment. The 5'cap of mRNA is recognized by eIF4F complex (consisting of eIF4E, eIF4G and eIF4A). The eIF4E is a cap-binding protein that recognizes the terminal 7-methylguanosine of the mRNA 5'end. The

eIF4A is a DEAD-box RNA helicase. Finally, eIF4G acts as a "scaffold" that binds eIF4E, eIF4A, poly(A)-binding protein (PABP) and eIF3. The factors eIF4B and eIF4H bind to mRNA and enhance the activity of the eIF4A subunit RNA-helicase (42,43,45). The giant heteromultimeric complex, eIF3 and PABP, which recognizes the 3'-poly A tail, help the assembly of the mRNA in the 43S complex. The eIF3 serves as a scaffold to alter the conformation of the 40S subunit, allowing easier access for mRNA, while PAB interacts with eIF4G leading to circularization of the mRNA, which is thought to stimulate translation (42–44). The complex formed scans the mRNA from 5'-3' direction, and stops when the Met-tRNAi anticodon recognizes the initiation codon in the mRNA. The scanning process requires ATP hydrolysis, and the assistance of eIF4A, eIF4G and eIF4B. However, the mechanism by which these factors assist scanning is still poorly understood (42).

The selection of the initiation codon (AUG) is facilitated by specific surrounding nucleotides which form a favorable sequence context (1,42). The best context contains a purine at position -3 and a G at position +4. Then, occurs hydrolysis of GTP bound to eIF2, a reaction facilitated by eIF5 a GTPase activating protein (GAP). eIF2.GDP releases the Met-tRNAi in the P-site of the small subunit (40S) and dissociates from the complex. Some other factors, like eIFs 1, 1A, 3 and 5, also dissociate at this stage (42–44).

The large ribosomal subunit 60S joins to 40S.Met-tRNAi.mRNA complex and this requires hydrolysis of GTP bound to eIF5B. This factor stabilizes tRNAi and facilitates the joining of the two ribosomal subunits. This step is irreversible since the ribosomal subunits do not dissociate until the entire mRNA is translated and protein synthesis is terminated. eIF5B.GDP has low affinity for the ribosome so it dissociates from the complex. eIF1A also dissociates at this stage (Figure 1-7) (42,43).

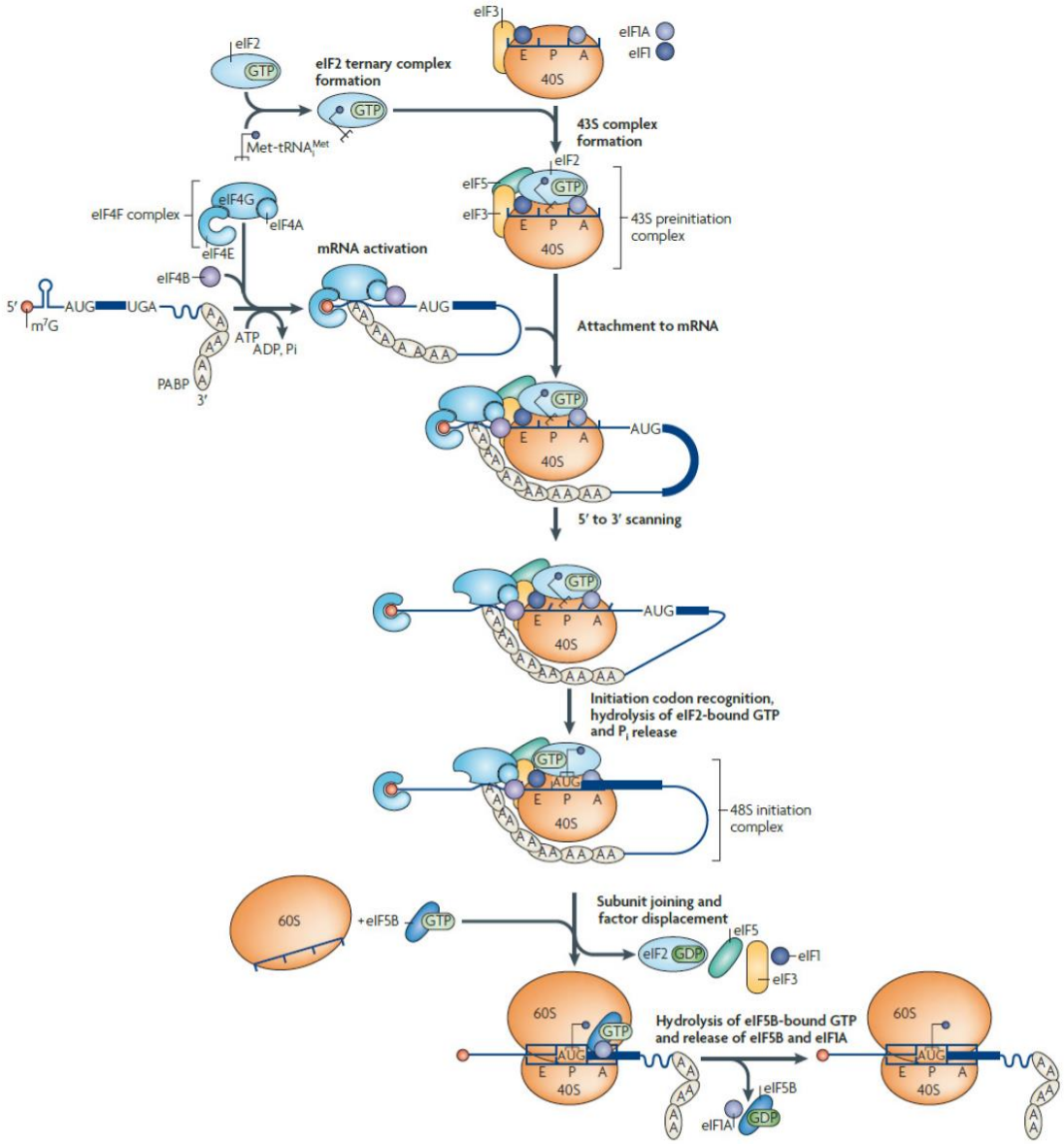


Figure 1-7. Scheme of the first step in eukaryotic translation – Initiation. After ribosome recycling, there is formation of the eIF2 ternary complex. This complex joins the 40S subunit, eIF1, eIFA, eIF3 and probably eIF5 to form 43S pre-initiation complex. During mRNA activation, the mRNA cap-proximal region is unwound in an ATP-dependent manner by the helicase eIF4F and eIF4B. 43S joins this mRNA region and scans the 5'UTR in a 5'-3' direction. When the initiation codon is recognized, eIF2-bound GTP is hydrolyzed. Finally 60S subunit joins 48S initiation complex with the concomitant displacement of eIF2-GDP and several other factors [adapted from (43)].

1.3.2 Elongation

Elongation starts with a peptidyl-tRNA in the P site and a vacant A site which accepts aa-tRNA carried in the form of a ternary complex with eEF1A.GTP (42). Only the cognate tRNA carrying the correct amino acid can enter the elongation phase to produce correct anticodon-codon interactions at the ribosomal A site. Binding of the cognate tRNA to the A site, triggers GTP hydrolysis, which promotes a conformational change in the ribosome that leads to tight binding of the aa-tRNA in the A site and the release of the eEF1A.GDP complex. The conformational change in the ribosome also positions the 3' end of aa-tRNA in the A site in proximity to the 3'- end of peptidyl-tRNA in the P site (1,46). The ribosomal peptidyl-transferase center then catalyzes the formation of the peptide bond between incoming amino acid and peptidyl-tRNA. The tRNA in the A site becomes deacylated and moves to the E site during ribosome translocation (42). This step is enhanced by eEF2 in combination with GTP hydrolysis (1,42). After translocation, the deacylated-tRNA is in the E site and the peptidyl-tRNA in the P site. The A site is free to receive another aa-tRNA complexed with eEF1A.GTP and this cycle is repeated several times until a stop codon is found at the A site and the process of termination is initiated (Figure 1-8) (1,42,46). After the hydrolysis of GTP and the release of aa-tRNA, into the ribosome, eEF1A.GDP is released and recycled to its GTP-bound form to participate in successive rounds of polypeptide elongation. eEF1B is a multifactor complex that catalyses this exchange of GDP for GTP (42).

1.3.3 Termination and Recycling

The presence of a stop codon in the ribosome A site recruits translation termination factors that activate the hydrolysis of the ester bond that links the polypeptide chain to the P site tRNA. In eukaryotes, these are two protein factors: eRF1 and eRF3. The class 1 factor, eRF1, has similar structure to that of tRNAs, recognizes the three stop codons (UAA, UAG, UGA) in the A site and also promotes the hydrolysis of peptidyl-tRNA in the P site.

The class 2 factor, eRF3, is a translational GTPase that stimulates the activity of eRF1. It also promotes the release of eRF1 from the ribosome following peptidyl-tRNA hydrolysis (Figure 1-8) (1,42,46). After termination, various components of the ribosome are recycled so they can be used in another round of translation. While in prokaryotes ribosomes appear to be recycled, in eukaryotes this step remains obscure. Little is known about how the ribosomal subunits are dissociated and how mRNA and deacylated tRNA are released. Some studies

showed that eIF3 can have an anti association activity due to induction of a conformational change in the 40S subunit, increasing the rate of subunits dissociation as well as lowering the rate of association. The close loop model of eukaryotic mRNAs suggests that the 40S subunit is not released back into the cytoplasm. Instead, it may be shuttled across or over the poly (A) tail back to the 5'- end of the mRNA allowing a subsequent round of translation in the same ORF (42,46).

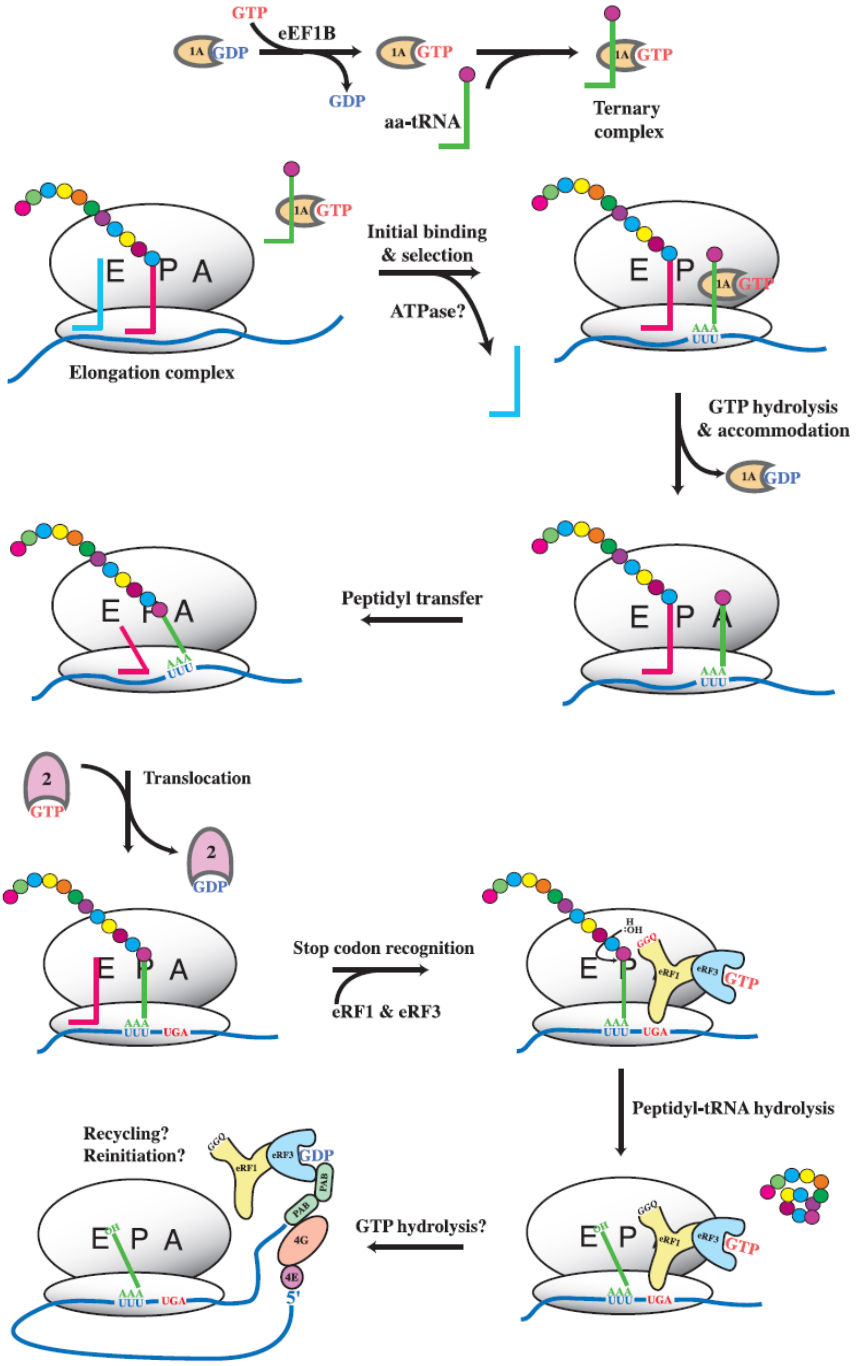


Figure 1-8. Scheme of mRNA elongation, termination and recycling in eukaryotes. During elongation the A vacant site accepts the aa-tRNA.eEF1A.GTP complex. If correct base pairing is formed in the A site, ATP is hydrolyzed and eEF1A.GDP is released. Peptide bonds are then catalyzed by ribosomes. Deacylated tRNA moves to the E site during ribosome translocation due to GTP hydrolysis by eEF2. After translocation, deacylated tRNA is in the E site while peptidyl-tRNA is in the P site. A stop codon in the A site recruits releasing factors that lead to peptidyl-tRNA hydrolysis. Finally the various components of the ribosome are recycled [from (42)].

1.4 Maintenance of mRNA translational accuracy

Translation of a particular codon depends on both the nature and abundance of the respective tRNAs, particularly on the non- random use of synonymous codons and the availability of the respective isoacceptor tRNA (47). However, in cells there are non-cognate and near-cognate tRNAs that may participate in the discrimination process by aaRS and ribosomes. Furthermore, modifications in tRNAs, especially those targeting the anticodon loop region, affect the function of the tRNA and can influence translational accuracy (48).

To facilitate accurate selection of amino acids or charged tRNAs, synthetases and ribosomes have kinetic discrimination mechanisms, driven by induced fit. In addition, to increase overall fidelity, there are quality control mechanisms in the two key events of mRNA translation: the synthesis of a correct aa-tRNA and the stringent selection of the aa-tRNA by the ribosome. The first one consists in the proofreading activity or editing in some aaRSs, a process already described in section 1.2.1 (49–51).

In the ribosomes, translational fidelity is controlled at two basic selection steps: preferential rejection of incorrect ternary complexes prior to GTP hydrolysis in the initial selection stage and preferential rejection of incorrect (mostly near-cognate) aa-tRNAs in the proofreading stage after GTP hydrolysis. In the first stage, the A site is in a low affinity state, to allow sampling of codon-anticodon interactions based on Watson-Crick base-pairing, at first and second positions. Sometimes this base-pairing is not strong enough and some deviations to the genetic code can arise that must be corrected in the second stage. The ribosome controls the differences in the stabilities of the codon-anticodon complexes and specifically accelerates the rates of GTP hydrolysis by the elongation factor. After GTP hydrolysis, the elongation factor can release aa-tRNA to enter in the large subunit A site or it can be rejected from the ribosome (proofreading) (Figure 1-9) (47,52–54). An additional mechanism can contribute to high fidelity protein synthesis after peptidyl-transfer. In this quality control step, ribosome recognizes errors by evaluating the codon-anticodon helix in the P site, leading to reduced fidelity during subsequent tRNA selection and ultimately to premature termination by release factors (49).

Over the past few years, many studies have documented the connection of ribosomes with mRNA and protein quality control processes (55–58). Ribosomes seem to be responsible for the recognition of defective mRNAs. Studies in yeast showed that the GTPase Ski7, which is related to eRF3, interacts with ribosomes and is also a component of the cytoplasmic exosome, linking the process of recognition of nonstop mRNAs (do not have a stop codon) to

its decay (55,58). Defective proteins formed in the ribosomes can be rapidly degraded by co-translational protein quality control processes (56,58).

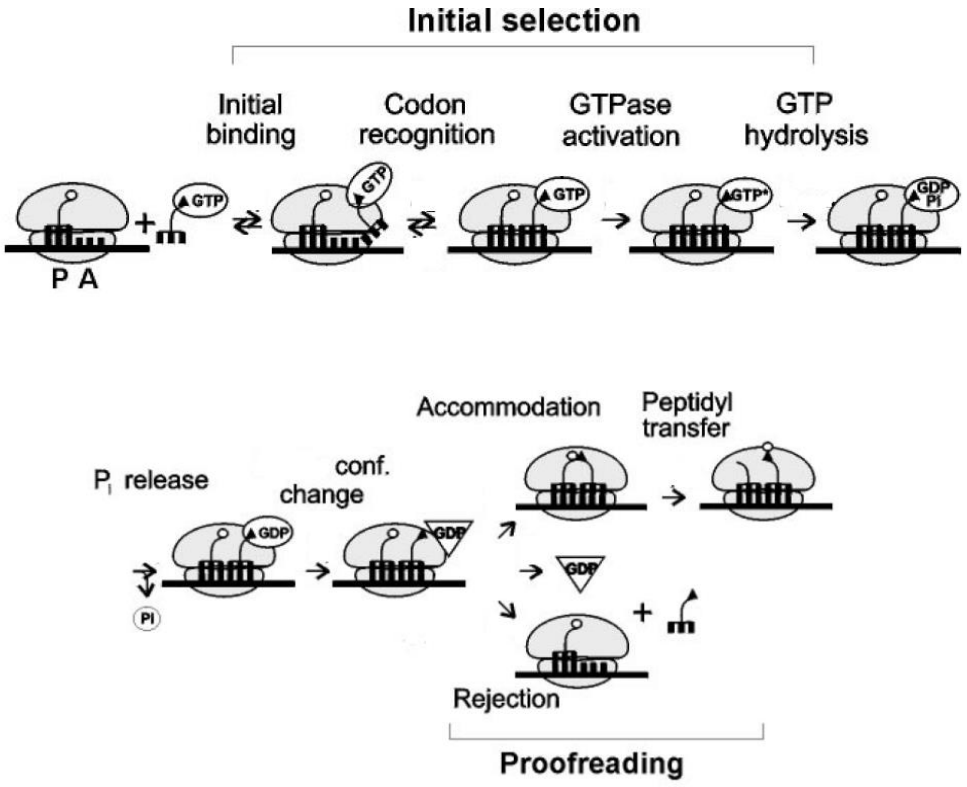


Figure 1-9. The two stages of quality control in the ribosome: Initial selection and Proofreading. Initial selection occurs by correct codon-anticodon pairing in the decoding center leading to GTP hydrolysis and conformational alterations. Once GTP is hydrolysed, the elongation factor releases aa-tRNA which can enter the A site (accommodation) form a peptidyl-tRNA or is rejected by the ribosome.
P – P site; A – A site [adapted from (53)].

1.5 Errors during protein synthesis – mRNA mistranslation

mRNA translation is a highly accurate biological process, however it is not error free. Errors can occur during the aminoacylation reaction or mRNA decoding in the ribosomes. Aminoacylation errors are mainly caused by failure of aaRS to differentiate between amino acids with similar structure or by the incorrect recognition of tRNAs. As already discussed, such errors are minimized by the aaRS editing mechanisms and by highly specific tRNA-aaRS interactions. In the ribosome, matching codons with anticodons is a tricky process that involves a certain amount of Watson-Crick wobble, to allow reading of all 64 codons with a minimal set of tRNAs. So, despite careful matchmaking, mistakes can occur. Four type of errors happen

during ribosome decoding namely; missence errors, which result in the substitution of one amino acid for another, nonsense errors, that cause readthrough of stop codons and produce proteins with extended C-termini; frameshifitng errors, that alter the mRNA reading frame, producing out-of-frame truncated proteins; and finally processivity errors, which cause premature termination (59–62).

The frequency of global translational error *in vivo*, in *E. coli* and mammalian cells is in the order of 10^{-4} (24,63–65). This physiological error is sustained due to protein quality control systems (PQC) which refold or destroy misfolded proteins after translation is finished. Environmental conditions (for example amino acid starvation) influence the fidelity of translation and the frequency of translation errors can increase significantly (60,64). Indeed, translational errors occur at a higher frequency than DNA replication errors. In *E. coli*, the typical mutation rate is approximately 10^{-9}, meaning that perfectly replicated genomes are commonplace, but perfectly synthesized proteomes never occur. Some studies suggest that the same occurs in eukaryotes (66,67). Error rates above 10^{-3} per codon originate statistical proteomes, meaning that some proteins vary stochastically in their sequence and do not faithfully represent the respective genes (67).

mRNA mistranslation leads to the synthesis of altered proteins that may not properly fold and may have the propensity to form protein aggregates. For this reason errors in protein synthesis have been assumed to have negative consequences for cells and may even cause disease (41,68). For example in mice an editing defective alanyl-tRNA synthetase (AlaRS) leads to loss of Purkinge cells and consequent ataxia (69). Surprisingly, several studies suggest that organisms can tolerate mRNA mistranslation relatively well and in specific cases errors can even be advantageous (Figure 1-10). Little is known about how cells tolerate high level of translational errors, but mistranslation may generate advantageous proteome diversity. Since these studies have focused on one type of mistranslation, it is also likely that organisms might tolerate variation in error rates for one or a limited number of amino acids (65,67). For example, *Candida spp.* contain a Ser tRNA ($tRNA_{CAG}^{Ser}$) that can be atypically recognized by both SerRS and LeuRS and for this reason is aminoacylated with both Ser and Leu, leading to the ambiguous decoding of the CUG codon (70,71). Physiologically misincorporation of Leu for Ser at this codon occurs at rates around 3% (72). Experiments with a construct tRNA from *Saccharomyces*, showed that *Candida albicans* can tolerate up to ± 28% incorporation of Leu and ± 70% incorporation of Ser at CUG sites without compromising growth rate. In certain environmental conditions this high level of mistranslation is advantageous (72,73) as ambiguous cells survive better in toxic environments containing cadmium, arsenate and

hydrogen peroxide (73,74). Increased mistranslation is also associated with major genome-wide rearrangements (75).

Another type of frequent mistranslation is the misincorporation of Met at various non-cognate codons which is beneficial in mammalian cells (76). During immune response, following exposure of mammalian cells to virus particles and Toll-like receptors, there is tenfold increase in misincorporation of Met at non-cognate-sites. Since Met residues can protect proteins from reactive oxygen species (ROS) mediated damage, the ability to increase Met levels in proteins by inducing tRNA Met-misacylation provides a protective mechanism against oxidative stress (76,77). This phenomenon was later demonstrated either in bacteria and yeast (78,79). The deliberate synthesis of proteins with alterations, in this case, an increase of Met residues, was named "adaptive translation" (77).

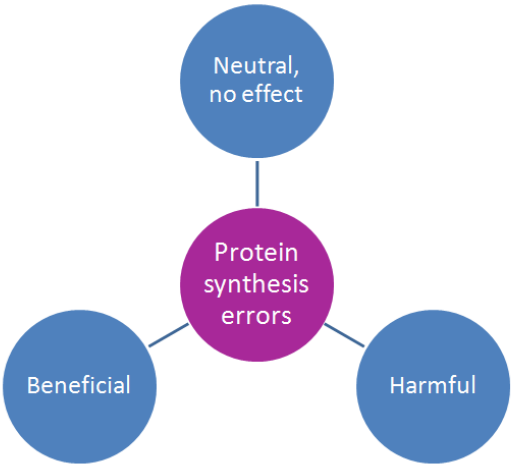


Figure 1-10. Consequences of protein synthesis errors in cells. Protein synthesis errors can be neutral having no effect (error containing proteins fold correctly), harmful (when errors lead to misfolding and aggregation) and beneficial (if the mutant protein acquires new functions). Misfolding proteins may also be beneficial if they trigger the stress response, creating pre-adaptation conditions that allow cells to survive better when exposed to stress.

The benefits of enhancing protein mistranslation rates may be due to either direct or indirect mechanisms. Indirect mechanisms include activation of the heat shock response and other stressor proteins and activation of stress-induced mutagenesis or wide-spread genome rearrangements. Direct mechanisms are related to the alteration of the proteome pool, as for example the misincorporation of Met into proteins, or the acquisition of novel protein functions (67). The last mechanism can be particularly important if the new protein variant

results in a dominant phenotype. For example in *Candida*, increase in misincorporation of Leu resulted in higher cell adhesion (80).

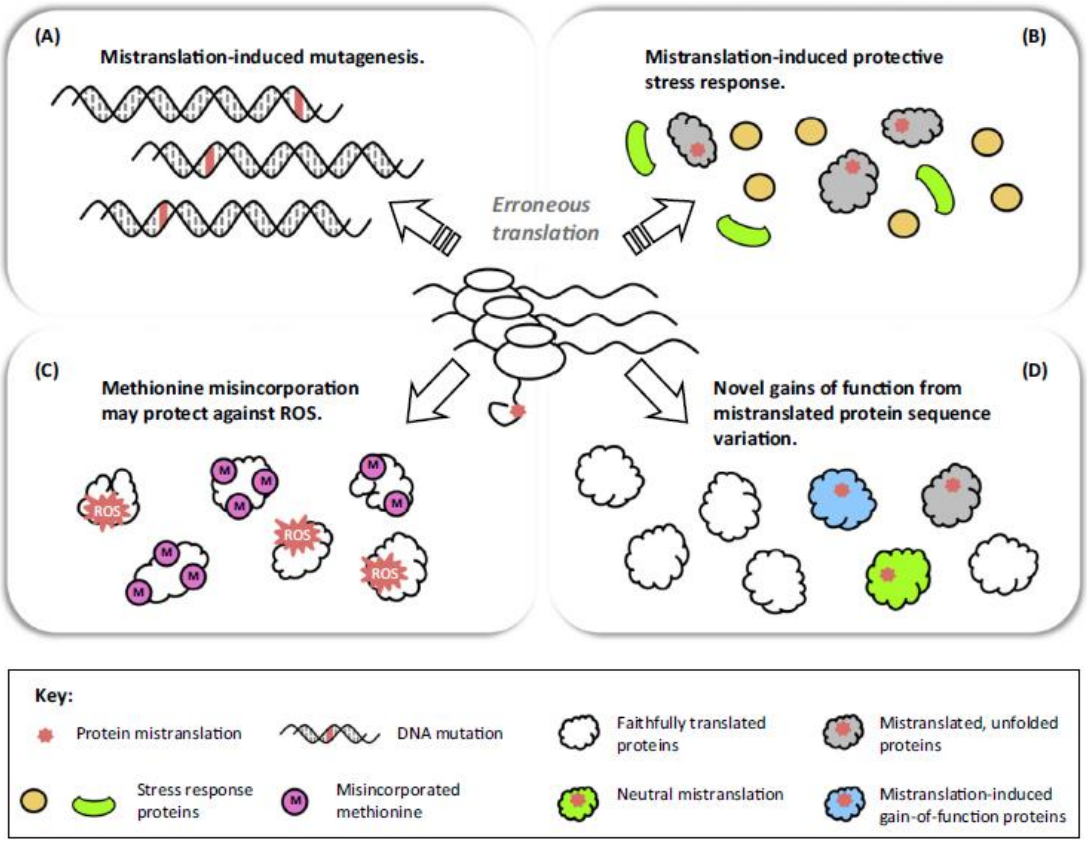


Figure 1-11. Mechanisms of adaptation to mRNA mistranslation. (A) Protein synthesis errors can cause stress-induced mutagenesis, which can increase the acquisition of drug resistance, **(B)** upregulation of stress responses such as heat shock proteins, **(C)** protection against oxidative stress by Met misincorporation and **(D)** acquisition of novel or altered functions by proteins [from (67)].

1.6 Protein misfolding

In order to carry out biological functions, proteins must fold into specific conformational states (encoded in their sequence). Folding takes place in complex and highly crowded environments, with the assistance of several auxiliary proteins known as molecular chaperones, which protect incompletely folded polypeptide chains from non-productive interactions. There are also folding catalysts that are responsible for isomerization of peptide bounds and formation of disulphide bounds. The conformation of a protein is determined by a very large number of relatively weak non-covalent interactions like hydrogen and hydrophobic

interactions. Therefore, folding depends on the "correct" (native-like) contacts between residues, which are in average more stable than "incorrect" (non-native) interactions. It also depends on the size and complexity of proteins. For larger proteins, folding is more complex and involves partially folded intermediate states (68,81,82).

Sometimes proteins cannot fold properly and are degraded or aggregate. DNA mutations, transcriptional errors, translational errors and erroneous post-translational modifications are important causes of protein misfolding. Environmental conditions and pathologies may exacerbate the errors perturbing protein folding beyond cellular repair capacity (83,84).

Misfolded proteins expose hydrophobic regions and are prone to self association and subsequent aggregation (66,81). Even though the specific mechanisms by which misfolded proteins perturb cellular functions are not entirely understood, they can disrupt cell membranes, interact with normally folded proteins, overload PQC mechanisms, generate ROS and induce apoptotic and inflammatory responses (85,86).

1.7 Protein Quality Control

To minimize toxic effects of misfolded proteins, cells have evolved a variety of PQC mechanisms that maintain proteome homeostasis (87). There are four major PQC systems in eukaryotic cells; molecular chaperones, the ubiquitin-proteasome system (UPS), Endoplasmic Reticulum associated degradation (ERAD), unfolded protein response in the Endoplasmic Reticulum (UPR) and autophagy (Figure 1-12). These folding surveillance mechanisms, counteract aggregation and eliminate misfolded proteins (82). Molecular chaperones are the first line of defense against misfolded proteins due to their capacity to refold proteins in the cytosol, endoplasmic reticulum (ER) or mitochondria. Misfolded proteins in the cytosol or ER can be degraded by the UPS (82). ER stress activates the UPR to counteract accumulation of misfolded proteins and re-establish proteome homeostasis. Finally, misfolded proteins and aggregates can also be degraded by autophagy.

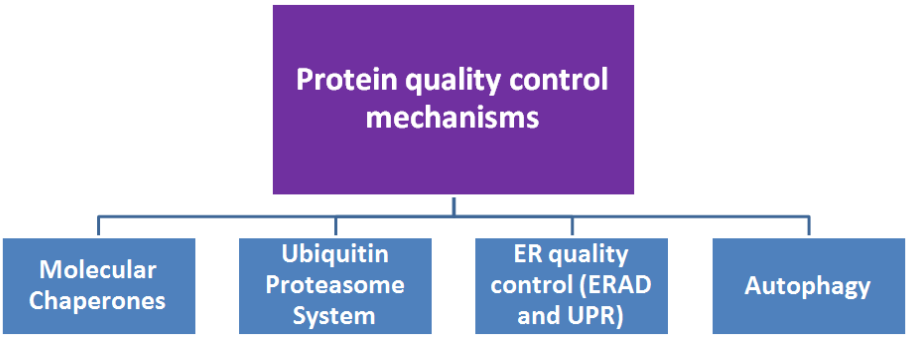


Figure 1-12. The four major mechanisms of protein quality control.

1.7.1 Molecular Chaperones

Molecular chaperones can recognize misfolded proteins, promote refolding or degradation via specialized proteases. They may also sequester proteins into specialized quality control compartments, to prevent the formation of toxic aggregates and regulate the inheritance of damage proteins (88). There are two compartments: the JUNQ (juxtanuclear quality control) and IPOD (insoluble protein deposit), that sequester different types of misfolded proteins. JUNQ compartment is associated with the cytosolic surface of the ER and contains 26S proteasomes and some chaperones. It concentrates soluble misfolded proteins that can either be degraded by the UPS or refolded by cytoplasmic chaperones. The IPOD contains insoluble aggregated proteins, and colocalizes with autophagy associated Atg8 (88). Since chaperones recognize selectively misfolded proteins, they are the first line of defense against toxic proteins (89).

There are different classes of molecular chaperones, which are usually classified according to their molecular masses: Hsp100, Hsp90, Hsp70, Hsp60 and Hsp40 and small heat shock proteins (sHsps). Each class comprises multiple chaperone isoforms (88). Some are constitutively produced while others are induced by stress (90). Some chaperones such as Hsp70 and Hsp90 interact with cofactors which influence their ATPase activity (91,92). In addition to these features, chaperones are also responsible for de novo protein folding (chaperonin TRiC/CCT and some Hsp70s), while others assist conformational maturation and degradation (Hsp70s and Hsp90). Small heat shock proteins, for example Hsp27, have been observed to tightly associate with protein aggregates (88,93).

Constitutively expressed (Hsc70) and stress-inducible forms of HSP70 are very important for proteostasis control, namely prevention of aggregation, promotion of folding to

the native state, solubilization and refolding of aggregated proteins, assistance in transport and degradation (94,95). In disease models Hsp70 is involved in preventing toxic aggregation (96–98). It interacts with hydrophobic peptide segments in an ATP dependent manner, in a Hsp40 dependent manner (91).

Hsp90, also an ATPase dependent chaperone, is involved in stress tolerance and in normal physiological conditions, is associated with a wide array of proteins (known as clients), that depend on it to acquire their active conformations. These client proteins are involved in signal transduction, protein trafficking, receptor maturation, innate and adaptive immunity. Approximately 20 co-chaperones interact with HSP90 to guide its recognition of client proteins and modulate its biochemical activity (88,92).

Chaperonins are large double-ring complexes of ~800kDa that enclose substrate proteins up to ~60kDa to accelerate folding and prevent aggregation. Hsp60 (also known as Group1 chaperonins) is largely, but not exclusively, compartmentalized in mitochondria and is responsible for folding and trafficking of pre-proteins (95,99).

Small Hsps usually form dimmers and oligomers, bind to partially denatured proteins in an unstable state and minimize irreversible aggregation. Activity of sHsps is considered energy-independent although substrate affinity binding sites are potentially affected by ATP binding (100).

1.7.2 The Ubiquitin-Proteasome System

Proteins that cannot be refolded are eliminated by specific proteases. In eukaryotic cells, the proteasome is the primary selective protease, located mainly in the cytosol, but also detectable in the nucleus. It participates in cell quality control and is essential for the rapid destruction of key regulatory proteins that control cell cycle progression, signal transduction and several other cell processes (Figure 1-13) (82,101,102).

The UPS involves the proteasome and a large number of proteins and is responsible for the identification, targeting and destruction of misfolded proteins. To be degraded by the proteasome, misfolded proteins must be covalently linked to polyubiquitin chains. This process is orchestrated by the concerted action of three enzymes; E1 (ubiquitin-activating), E2 (ubiquitin-conjugating) and E3 (ubiquitin-ligating). The first enzyme (E1) uses ATP to form a high-energy thiolester bond with the C-terminus of ubiquitin (76 amino acid protein). Next, activated ubiquitin is transferred to a ubiquitin conjugating enzyme (E2) and then is conjugated

to target proteins in a process mediated by E3 ubiquitin ligase. The E3 ligase is an adaptor molecule that interacts with both the target protein and E2. From this interaction results the formation of an isopeptide bond between the C-terminus of ubiquitin and an ε-amino group of lysine residues in the target protein (Figure 1-13). Polyubiquitin chains are then formed after successive transfers of activated ubiquitin to lysine-48 of the previously conjugated ubiquitin molecule (102,103). Polyubiquitinated proteins bind to specific ubiquitin binding domains (UBD) located in the proteasome. Prior to degradation, chains of ubiquitin are disassembled to ubiquitin monomers by cysteine-protease and metalloprotease deubiquitinating enzymes (DUBs). Proteins to be degraded are unfolded and funneled through the narrow entrance of the proteasome (102,104).

The proteasome is a macrocomplex formed by the assembly of several subunits that form the core particle (20S) and the regulatory particles (19S). The core particle exhibits a barrel-like structure in which 28 subunits form four seven-membered rings. The two outer rings are composed by seven α-subunits and the two inner rings are composed by seven β-subunits. β-rings contain the proteolytic active sites with caspase-like, trypsin-like and chymotrypsin-like activities. α-rings control the entrance of the unfolded protein into the catalytic cavity. The 19S regulatory particle is responsible for the recognition, unfolding and translocation of the protein to the 20S core. It is composed by two subcomplexes: a base (adjacent to 20S) and a lid (on top of the base). The base contains six AAA-ATPases and three non-ATPases subunits. The lid is formed by eight subunits and is critical for the recognition and deubiquitination of proteins (Figure 1-13) (101,104–106).

An active proteasome can exist in several forms; 26S (20S core plus one 19S regulatory subunit) or 30S (20S core plus two 19S regulatory subunits). Although it is still unclear whether there are functional differences between the 26S and 30S proteasomes, the latter is the most abundant proteasome in cells and is believed to be the physiologically functional unit (104,105). The accession of proteins only through 19S particles safeguards the highly specific nature of the process and prevents non-specific protein degradation (107).

Molecular chaperones cooperate directly with the UPS to redirect misfolded proteins to degradation. The mammalian cofactor of Hsp70 and Hsp90, called CHIP protein, shifts the mode of action of these two chaperones, binds to E2 and can also act as an E3 ubiquitin ligase during the ubiquitination of chaperone substrates. Several studies support the idea that CHIP is involved in quality control, in Parkinson's and Alzheimer's diseases, participating in the degradation of defective proteins (108–110).

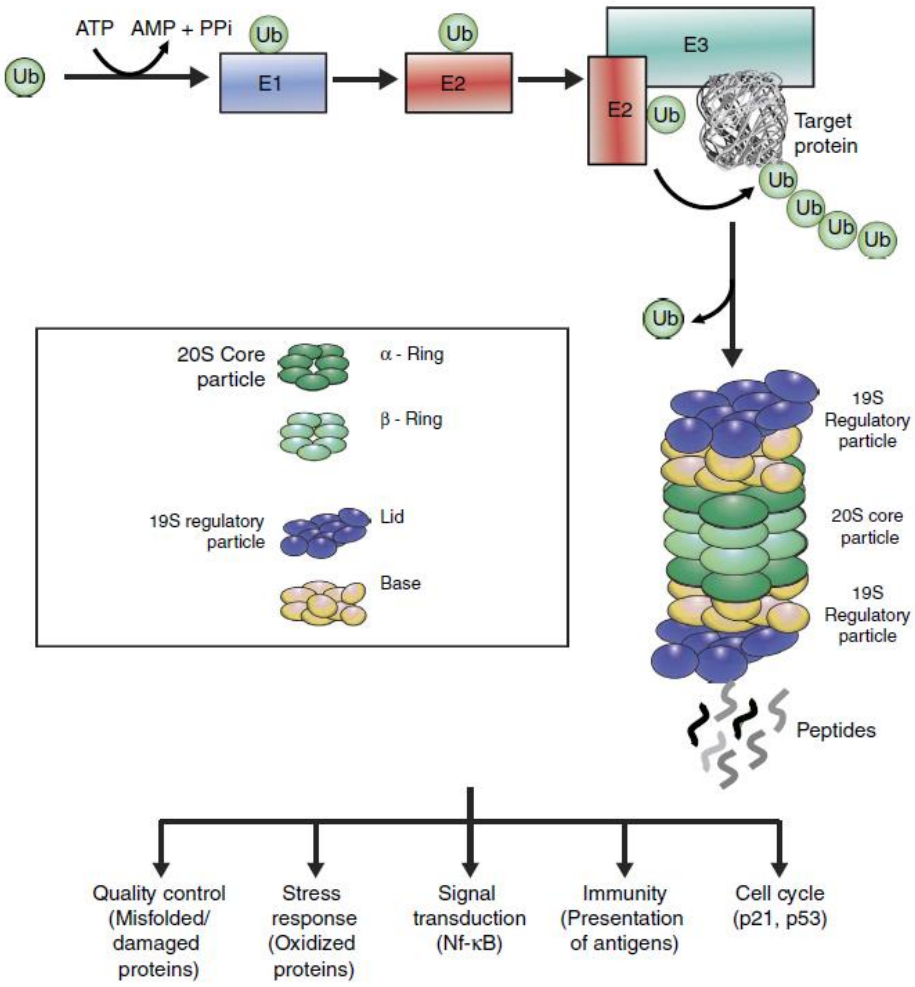


Figure 1-13. The ubiquitin-proteasome system (UPS). Degradation of proteins by the UPS is initiated with ubiquitin conjugation to the misfolded protein. First E1 (ubiquitin-activating enzyme) activates ubiquitin, in an ATP-dependent manner. Ubiquitin is then transferred to E2 (ubiquitin-conjugating enzyme). E3 (ubiquitin ligase) links ubiquitin from E2 to the target protein. This reaction is repeated to form a polyubiquitin chain. This polyubiquitin chain is recognized by the proteasome for degradation. Proteasomes are formed by a 20S particle (α-ring and β-ring) which constitutes the core and is responsible for the proteolytic activity and 19S regulatory particles (lid and base) that recognize ubiquitinated proteins. The proteasome is involved in a variety of cellular functions, such as quality control, stress response, signal transduction, immunity and cell cycle regulation [adapted from (101)].

1.7.3 Quality control in the ER

The ER is involved in several cellular functions, namely, synthesis and sorting of secretory and membrane proteins, biosynthesis of phospholipids, cholesterol, steroids, degradation of glycogen, detoxification reactions and maintenance of intracellular calcium homeostasis. About one third of the proteins synthesized in eukaryotic cells are folded in the

ER, and this process requires a myriad of chaperones and modifying enzymes. After the acquisition of native conformation, proteins leave the ER. They can be translocated to other organelles or secreted to the extracellular medium (111).

If correct folding of proteins cannot be achieved, defective proteins are usually retro-translocated to the cytoplasm to be degraded by the proteasome, a process called ER-associated degradation (ERAD). Misfolded proteins are recognized in the lumen of the ER and transported across the ER lipid bilayer into the cytoplasm. The substrate is then ubiquitinated by a membrane-associated ubiquitin ligase (E3). The ubiquitinated substrate is extracted from the membrane in an ATP-dependent manner and released in the cytosol for degradation by the proteasome (Figure 1-14). This is possible because, E3 ligase complexes are composed by proteins involved in substrate recognition and also retrotranslocation (111,112).

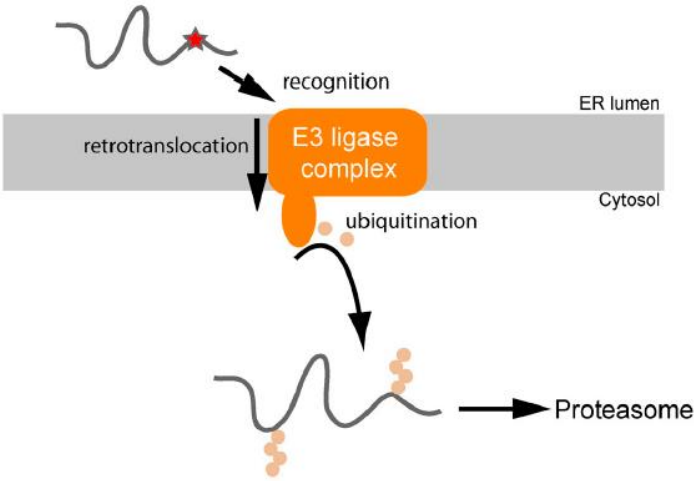


Figure 1-14. ER- associated degradation (ERAD). E3 ligase complex recognizes, retranslocates and ubiquitinates misfolded proteins, releasing them into the cytosol to be degraded by the proteasome [from (112)].

Inefficient ERAD activation results in the accumulation of misfolded proteins in the lumen and membrane of the ER, a condition known as ER stress. This is detected by three sensors that face the ER lumen: inositol-requiring enzyme 1 (IRE-1 or ERN1), a protein kinase RNA-like endoplasmic reticulum kinase (PERK, also known as PEK or EIF2AK3) and activating transcription factor 6 (ATF6) (Figure 1-15). These sensors initiate ER-to-nucleus signaling cascades to trigger an adaptive response: the UPR that restores protein folding homeostasis. This is achieved by three main mechanisms: transient reduction in protein

synthesis, increase in the folding capacity and ERAD. Initiation of programmed cell death happens when ER stress cannot be resolved (111,113).

Under non-stress conditions, binding immunoglobulin protein (BiP, also known as Grp78), the most abundant ER chaperone is associated with ER-luminal domains of the three sensors, keeping them inactive. Upon accumulation of unfolded or misfolded proteins BiP detaches from sensors to assist folding and these sensors become active. Alternatively, unfolded proteins can bind directly to IRE-1 (111,113).

IRE-1α is a bifunctional transmembrane protein, harbouring in its cytoplasmic part both a serine-threonine kinase domain and a C-terminal RNase domain. Activation of IRE1 involves its oligomerisation and *trans*-autophosphorylation of the kinase domains. Oligomerisation activates the RNase domain, responsible for the cleavage of X-box-binding protein 1 (XBP-1) mRNA at two stem-loop structures through an unconventional cytoplasmic splicing reaction. This reaction specifically removes an intron from unspliced or full length XBP-1 (XBP-1u) to generate sliced XBP-1 (XBP-1s). XBP-1s is a highly active transcription factor that activates genes that enhance ER protein-folding capacity (chaperones), degradation of misfolded proteins (ERAD) and also lipid biosynthesis (111,113). Studies showed that prolongation of IRE-1 signaling during ER stress can promote cell survival, being this adaptive mechanism important to match ER folding capacity with protein folding demand (111,114). Another mechanism proposed for IRE-1 is the degradation of mRNAs encoding proteins that transverse the secretory pathway, IRE-1 dependent decay (RIDD). Through this mechanism of protein synthesis attenuation ER stress is relieved (Figure 1-15) (115).

PERK is a transmenbrane kinase that shares with IRE-1 approximately 20% similarity in its luminal domain. It phosphorylates eukaryotic initiation factor 2α (eIF2α), leading to inhibition of the guanine-exchange factor eIF2B that recycles eIF2 to its active GTP-bound form. This results in a delay in ternary complex formation and a reduction in cap-dependent translation (116). The block of initiation of translation is not total but reduces the protein load in the ER (113), and permits translation of certain transcripts with short upstream ORF in their 5'-UTR regions, such as ATF4. ATF4 is a transcription factor responsible for induction of a set of genes involved in amino acid metabolism, antioxidant stress response and protein secretion (Figure 1-15) (113,117). It also directs the expression of a second transcription factor, C/EBP-homologous protein (CHOP) and growth arrest and DNA damage-inducible protein (GADD34). CHOP is a proapoptotic factor, while GADD34 exerts a negative feedback in eIF2α phosphorylation. GADD34 associates with protein phosphatase 1 (PP1) and promotes

dephosphorylation of eIF2α. By this means, GADD34 promotes recovery from translation repression, helping cells to return to their normal function (113,118).

ATF6 is a transmembrane protein with its carboxyl terminus facing the ER lumen and a transcription factor in its amino terminus. Upon ER stress, BiP dissociates from the luminal domain and ATF6 is transported to Golgi apparatus via coat protein COPII-covered vesicles, where is proteolytically cleaved. Site 1 and site 2 proteases (S1P and S2P, respectively) sequentially cleave ATF6 to release the amino-terminal transcription factor fragment pATF6-N from its membrane anchor. pATF6-N then moves to the nucleus and binds to promoters containing the ER stress response elements, thus inducing the expression of ER stress-induced genes such as chaperones and proteins involved in ERAD (Figure 1-15). XBP-1 expression is also induced by pATF6-N (113,119).

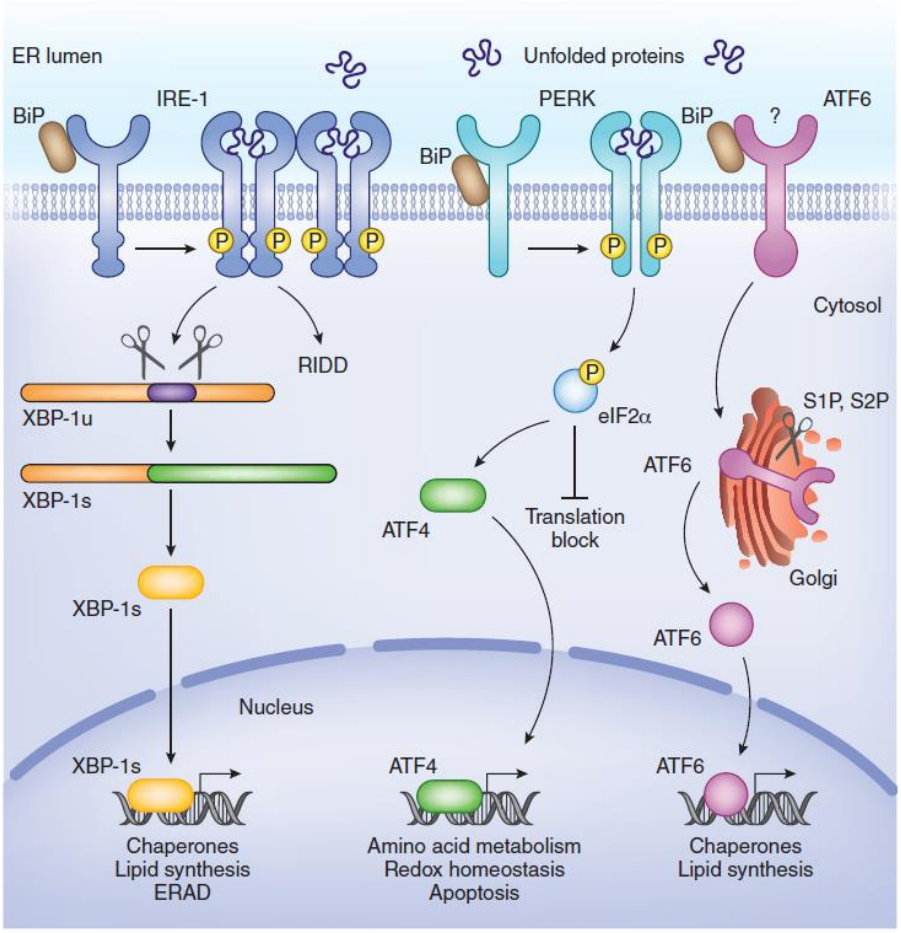


Figure 1-15. The three UPR sensors in mammalian cells. IRE-1 and PERK are activated by oligomerisation and *trans*-phosphorylation, upon release of BiP. IRE-1 cleaves XBP-1 mRNA to generate its active form, XBP-1s. This transcription factor is then translocated to the nucleus and induces expression of genes encoding chaperones, lipid synthesis and ERAD proteins. IRE-1 also contributes to RIDD, degradation of mRNAs encoding proteins from the secretory pathway. PERK phosphorylates eIF2α leading to reduction in protein synthesis. Some mRNAs with small upstream open reading frames in their 5'UTR escape the translation block, namely ATF4. ATF4 triggers expression of CHOP, GADD34 and additional factors important for amino acid metabolism and redox homeostasis. Upon BiP release, ATF6 is translocated to the Golgi, where it undergoes cleavage by S1P and S2P. The ATF6 active fragment moves to the nucleus and coordinate the expression of genes encoding chaperones and proteins involved in lipid biosynthesis [from (113)].

As mentioned before, many components of the UPR trigger apoptosis under chronic stress. Induction of AFT4 and its downstream target CHOP lead to apoptosis possibly through the transcriptional up-regulation of several pro-apoptotic proteins of the BCL-2 family, known as BH3-only proteins (for example BIM) (120). One of the isoforms of IRE-1, IRE-1α when activated interacts with the adaptor protein tumor necrosis factor receptor-associated factor 2

(TRAF2) and leads to the downstream activation of apoptosis signal-regulating kinase 1 (ASK1) and JUN N-terminal kinase (JNK), also contributing to cell death (121,122). Besides this, prolonged RIDD may reduce drastically the levels of proteins essential for survival participating in cell death (111). The molecular basis for the switch between prosurvival and proapoptotic UPR function is poorly understood. However, evidences suggest that UPR signaling can integrate stress intensity and duration information to promote cell adaptation or death (123,124). For example, in human cell lines exposed to ER stress over prolonged time, both ATF6 and IRE-1 are attenuated early, while PERK activity and CHOP expression persist until the death of the cells (114,125).

Due to the broad spectrum of activities of the UPR in organ homeostasis, ER stress plays a role in the progression of protein-misfolding disorders and also several others such as cancer, diabetes and inflammation (123,126). For example, abnormal accumulation of the amyloid-β (Aβ) peptides in the brain, a hallmark of Alzheimer's disease, leads to neuronal loss and cognitive deficits due to ER and oxidative stress (127). Elevated levels of eIF2α-P and ATF4 have also been observed in brains of people with Alzheimer's disease and in mouse models of the disease (128). In agreement with these findings a recent study also demonstrated that decreasing the expression of two of the eIF2α kinases, PERK and general control non-derepressible-2 (GCN2), improve cognitive function and synaptic plasticity in Alzheimer's disease transgenic mouse models (129,130).

ER stress and UPR activation have been documented in several human cancers (131). Because tumors arise and progress in a stressful environment, possibly transformed cells use UPR activation as a survival strategy. In fact, some studies demonstrate that UPR activation is required for oncogenic transformation, despite its known role in apoptosis. For this reason, UPR can also have a tumor-suppressing role depending on the tumor context (131,132). For example, in mice, *Perk* deletion delays *Neu*-dependent mammary tumor development and reduces lung metastases, whereas long-term PERK inactivation increases susceptibility to spontaneous mammary tumorigenesis (133). PERK activation can also promote MYC-induced cell transformation through autophagy, possibly due to the increase in ATF4, CHOP and factors that activate transcription of many autophagy genes (134,135). *IRE1A* and *XBP1* mutations have been identified in tumor cells from patients with multiple myeloma (136) and loss of function of XBP-1 promotes tumorigenesis in mouse models of intestinal cancer (137,138). On the other hand, increase *XBP1* mRNA splicing was observed in human triple negative breast cancers, possibly indicating a requirement for XBP-1 in cancer stem-like cells (130,131,139). Since ER stress and UPR pathway are involved in the etiology of many

diseases, UPR pathways could be important therapeutic targets, particularly for anticancer and anti-dementia approaches (130).

1.7.4 Autophagy

Autophagy is the generic name used for any intracellular process that results in the degradation of cytosolic fractions, organelles and macromolecules inside lysosomes. This degradation system is activated in stress conditions, such as nutrient deprivation, DNA damage, hypoxia, oxidative stress or ER stress to support energy balance. In this way, autophagy plays an important role in PQC, being responsible for the selectively degradation of misfolded and aggregated proteins (101,140).

Lysossomes are single-membrane vesicles that contain several hydrolases (proteases, lipases, nucleotidases and glycosidases), which need an acidic pH for optimal activity. In the case of protein quality control, the low pH in the lumen is responsible for the partial unfolding of substrate proteins allowing proteases to have access to internal peptides. These proteases convert proteins to di- and tripeptides and free amino acids, which are released into the cytosol and recycled in protein synthesis (101,141).

In mammalian cells, there are three main types of autophagy; macroautophagy, microautophagy and chaperone-mediated autophagy (CMA) (Figure 1-16). In macroautophagy, cytoplasmic portions are engulfed into double membrane vesicles or autophagosomes. Autophagosomes are then fused with lysosomes and sequestered material is degraded. Autophagy-related genes (ATG) encode proteins that organize into complexes, and are responsible for the regulation of the autophagic process. In microautophagy small components of the cytoplasm are directly engulfed by lysosomes. Finally, in chaperone-mediated autophagy, specific cytoplasmic proteins are recognized by chaperones through a consensus sequence and transferred to lysosomes. Macro- and microautophagy participate in the degradation of both proteins and organelles, while CMA is only responsible for protein degradation (101,141).

Several autophagic receptors respond to specific cellular insults, and recognize abnormally altered or aggregated proteins, targeting them for degradation. In macroautophagy, p62/SQSTM1, NIX and NBR1, recognize and facilitate the elimination of ubiquitinated proteins. These receptors simultaneously bind microtubule-associated protein light chain 3 (LC3) in autophagosomes and ubiquitinated proteins to allow their inclusion into autophagosomes for subsequent degradation (142). CMA mediates the lysossomal degradation

of proteins that contain a consensus pentapeptide (KFERQ), which includes a substantial number of disease-causing aggregate-prone proteins. Hsc70 escorts proteins to the receptor lysosome-associated membrane protein type 2a (LAMP2A) (140,143).

Mammalian target of rapamycin (mTOR) kinase and AMP-activated protein kinase (AMPK) are two master regulators of autophagy. mTOR integrates growth factor and nutrient signals and inhibits autophagy, whereas AMPK, senses and regulates cellular metabolism to maintain energy homeostasis and promotes autophagy. AMPK and mTOR regulate the phosphorylation of the mammalian homolog of Atg1, Ulk1 (144).

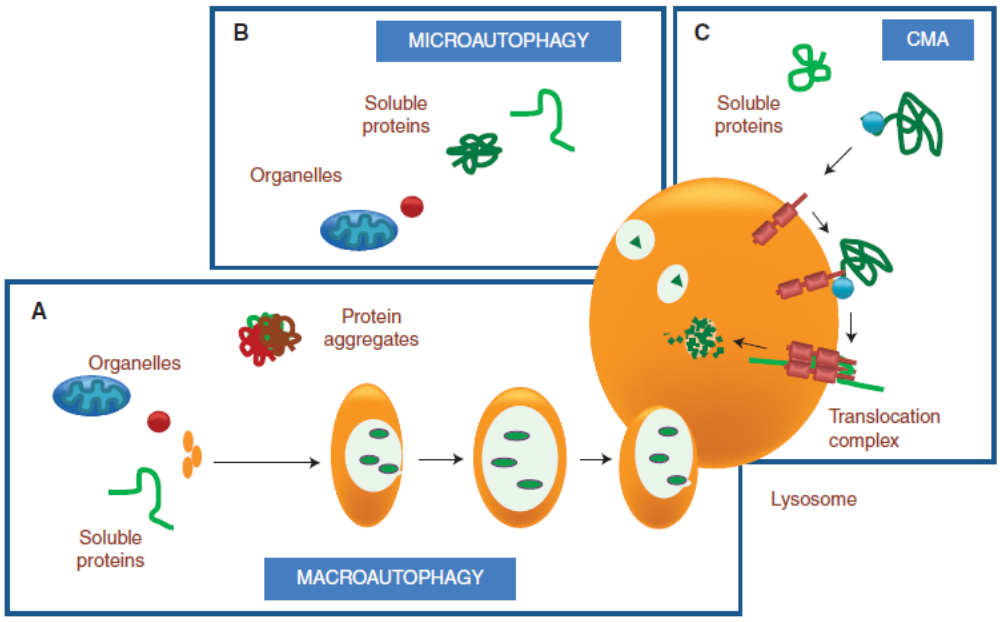


Figure 1-16. Autophagic pathways. (A) In Macroautophagy, cytoplasmic regions are sequestered into autophagosomes that are then fused with lysosomes for cargo delivery. **(B)** In Microautophagy, the lysosomal membrane invaginates to trap regions of the cytosol that are internalized into the lysosome lumen as vesicles. **(C)** In CMA a targeting motif in the substrate protein is recognized by chaperones that are responsible for the delivery to the surface of the lysosome, where it binds to specific receptors that form translocation complexes [from (141)].

1.8 mRNA mistranslation and disease

An extensive and growing catalogue of human diseases, caused by alterations in the components of protein synthesis machinery, have been discovered in recent years. Defects can be related to protein synthesis in the cytosol or can involve elements of mitochondrial protein

synthesis and have a broad spectrum of phenotypic outcomes, resulting in diverse diseases from neurodegeneration to cancer (41).

Ribosomopathies are a class of diseases caused by mutations that affect the biosynthesis and/or functionality of the ribosome. They are characterized by pleiotropic abnormalities (birth defects, heart and lung diseases, anemia and ataxia) and in middle age patients present elevated cancer risk (30- to 40- fold compared with the general population) (145,146). Mutations in the ribosomal protein gene RPL10/uL16 were recently identified in patients with T-cell acute lymphoblastic leukemia (147). This ribosomal protein plays a role in catalysis and also 60S biogenesis. Mutant ribosomes have structural defects that affect translational fidelity, promoting elevated rates of -1 programmed ribosomal frameshifting and impaired recognition of termination codons (146).

Ribosome frameshifting can also have a potential role in generation of aberrant proteins such as ubiquitin B (UBB+) and β-amyloid precursor protein (APP+) implicated in Alzheimer's and other neurodegenerative diseases. These proteins called "+1 proteins" have carboxyl-terminal amino acids encoded by an alternative reading frame of the mRNA and inhibit the function of the proteasome, forming neurofibrillary tangles and neuritic plaques (148).

Mutations in the editing domain of the mouse AlaRS, which allow the enzyme to charge $tRNA^{Ala}$ with Ser or Gly are associated with rapid Purkinge cell loss in the cerebellum and development of ataxia (23,69,86). The misincorporation of these amino acids into proteins leads to toxic aggregates, increase in protein ubiquitination, formation of autophagosomes, induction of molecular chaperones (such as members of Hsp70 family) and upregulation of the UPR (induction of BiP and CHOP) (69). Also mutations in genes that encode GlyRS and TyrRS have been associated with the Charcot-Marie-Tooth (CMT) disease, characterized by muscular weakness and atrophy of the distal extremities (149,150). The disease causing mutations are dominant and the mutant proteins are fully active for aminoacylation, but do not interact efficiently with their cognate tRNAs (151). At least 11 distinct mutant alleles in the human population have been reported in GlyRS, while two missence mutations and one de novo deletion in TyrRS have been identified in unrelated families affected with the disease (23,149,150).

Also mutations in tRNAs, particularly in mitochondrial tRNA genes, have been correlated with severe diseases, including fatal cardiopathies, encephalopathies, myophaties, diabetes and others (152). Over 200 mt-tRNA mutations have been linked to disease. The first report, in 1990, links a mutation in $tRNA^{Leu}$ gene (*MTTL1*) with Mitochondrial myopathy,

encephalopathy, lactic acidosis and stroke (MELAS) (153). Eighty percent of patients with MELAS carry a maternaly inherited A>G mutation in the nucleotide 3243 in $tRNA_{UUR}^{Leu}$ that affects its stability, impairs charging efficiency with Leu and prevents taurine modification in the anticodon (in the wobble base position). A mutation in $tRNA^{Lys}$ (A8344G) was associated with Mioclonic epilepsy with ragged-red fibers (MERF). This mutation also affects aminoacylation and taurine modification in the anticodon (154,155).

Deregulation of RNA polymerase (pol) III and its products (mainly tRNAs) have been reported in a wide range of transformed cells (156,157) and tRNAs are frequently overexpressed in breast cancer. Previous studies comparing cancer-derived versus non-cancer-derived cell lines showed that nuclear-encoded tRNAs increase up to 3 fold and mitochondrial-encoded tRNAs increase by up 5 fold. Also, in tumors versus normal breast tissue nuclear and mitochondrial-encoded tRNAs increased up to 10 fold (158). Recently, Goodarzi *et al*., identified two specific tRNAs ($tRNA_{UUC}^{Glu}$ and $tRNA_{CCG}^{Arg}$), that are upregulated in human breast cancer cells and promote breast cancer metastasis. Upregulation of these tRNAs enhances stability and ribosome occupancy of transcripts enriched for their cognate codons (159).

Finally, growing evidences indicate that tRNA modifying enzymes may play important roles in cancer, type 2 diabetes, neurological disorders and mitochondrial-linked disorders (16). Human tRNA methyltransferase 1 (*TRM1*) gene encodes a methyltransferase that dimethylates guanosines ($m^2{}_2G$) at position 26 of tRNAs and a frameshift mutation that inactivates this gene has been reported as a novel marker for recessive cognitive disorders (160). tRNA methyltransferase homolog 12 (*TRMT12*), the human homolog of yeast *TRM12*, is one of the enzymes involved in the formation of wybutosine at position 37 on $tRNA^{Phe}$. In several breast cancer cell lines and breast cancer tumors this gene was overexpressed (161). CDK5 regulatory subunit associated protein 1-like 1 (*CDKAL1*) gene encodes for a methylthiotransferase involved in the complex 2-methylthio-N6-threonyl carbamoyladenosine (ms^2t^6A) modification at position 37 in $tRNA_{UUU}^{Lys}$. Variantions in the *CDKAL1* gene are associated with type 2 diabetes (162). Evidences suggest that lack of ms^2t^6A modification leads to mistranslation of several proteins, including proinsulin, thus triggering the ER stress response. As it accumulates in an abnormal form, proinsulin cannot be converted into insulin, leading to glucose intolerance (16,163).

Despite these alterations in the protein synthesis machinery it is still unclear how the above mutations cause disease.

1.9 Mutant tRNAs as strategy to induce protein synthesis errors and proteotoxic stress

Many experimental strategies have been developed to study stress responses elicited by the accumulation of misfolded proteins. Incubation of cells with amino acid analogues, like azetidine carboxylic acid or canavanine, Pro and Arg analogues is commonly used. Amino acid analogues, due to chemical similarities to natural amino acids, can escape detection by the cellular protein synthesis machinery and become misincorporated into proteins. This leads to protein misfolding and aggregation. This strategy was extremely useful in the identification of components of the heat-shock response and UPS (164,165). Another strategy is based on proteins that are prone to aggregation, such as Aβ and polyQ, which are used as models for Alzeimer's and polyglutamine diseases, respectively. The expression of these proteins has been extensively used in human cells and animal models. For example, human cells and *C. elegans* models expressing of Aβ and polyQ were used to identify subnetworks of chaperones repressed or induced in aging and disease and also to clarify their roles in these situations (166).

In 2010, Geslain *et al*., developed a new approach to destabilize the proteome and increase the level of misfolded proteins based on mutant tRNAs (165). They introduced 10 different mutant tRNAs in a human cell line (HEK293) and in a vertebrate embryonic model (chick embryos). To construct mutagenic tRNAs, Geslain *et al*. took advantage of the recognition mechanism of $tRNA^{Ser}$ by SerRSs. Since this enzyme does not recognize the anticodon loop of Ser tRNAs, but rather the acceptor stem, D-arm and extra stem/loop (13), it was possible to alter the anticodon of this tRNAs without interfering with the serylation reaction. Mutant Ser tRNAs were aminoacylated with Ser, but translated codons complementary to the engineered anticodon, generating random proteome mutations X→S (Figure 1-17).

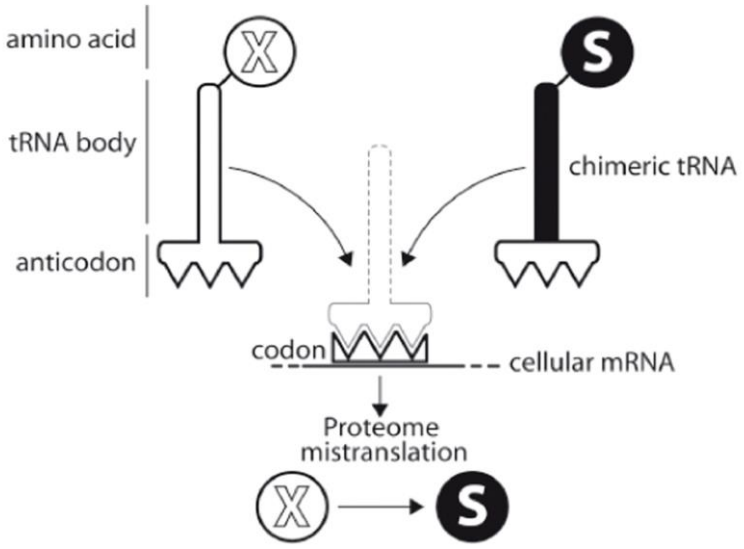


Figure 1-17. Schematic representation of "chimeric" tRNAs that mutate the proteome. The wild-type anticodon of $tRNA^{Ser}$ is altered to decode codons of different amino acids. Despite alterations in the anticodon these tRNAs are aminoacylated with the amino acid Ser. Ser is incorporated into growing polypeptide chains, producing mistranslated proteins [from (165)].

The activity of those tRNAs was demonstrated using a gain-of-function mutation introduced in the fluorescent GFP reporter. GFP contains a critical Ser residue at position 65, altering the respective codon for codons corresponding to the anticodon of the mutant tRNAs allowed for monitoring Ser incorporation at position 65 of the mutant GFP. Fluorescence measurement showed that Ser misincorporation was dependent on the tRNA used (Figure 1-18). Amino acid quantitative analysis of purified GFP and human epidermal growth factor receptor (EGFR), confirmed the misincorporation of Ser in the proteome. General parameters such as transcription, protein synthesis levels, growth rate and cell viability were also evaluated, showing different responses according to the type of error introduced in proteins (165).

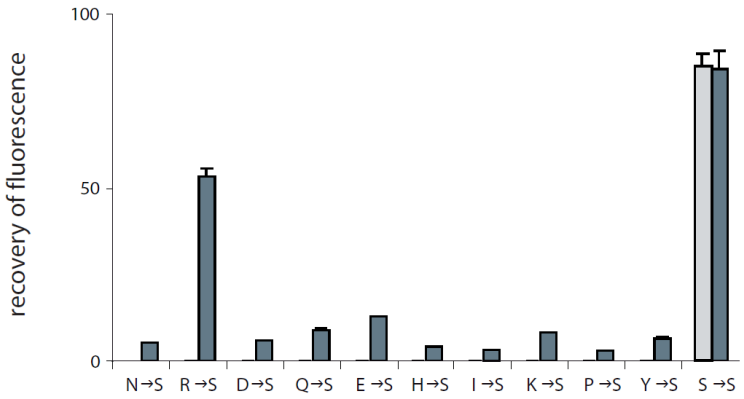


Figure 1-18. Fluorescence recovery of GFP coded by a gene where serine-65 had been substituted by 10 codons recognized by the engineered tRNAs [from (165)].

Similar experiments with mutant Ser tRNAs were carried out in zebrafish. Decoding activity of these tRNAs monitored using the GFP reporter showed that the tRNAs were functional. These Ser misincorporations into the proteome elicited a stress response in zebrafish embryos, resulting in the activation of the UPS and UPR and inhibition of protein synthesis (167).

1.10 Aims of the study

The main aim of this PhD thesis was to study the stress response and adaptation capacity of human cells to protein synthesis errors (PSE). For this, we have created human cell lines (HEK293) that mistranslate a single codon in a controlled manner and also overexpressed the Wt $tRNA^{Ser}$. Cellular responses were studied at a phenotypic and molecular level. This strategy unraveled adaptation and stress mechanisms that are important to understand the molecular basis of human diseases associated with protein synthesis errors.

The main biological questions addressed in this thesis were the following:

1 – Can human cells cope with increased level of protein synthesis errors? What are the phenotypic consequences?

2 – How do these cells adapt and counteract proteotoxic stress to thrive in culture?

3 – What are the transcriptional responses to mistranslation?

4 – What is the impact of mistranslation on the genome?

5 –Are HEK293 mistranslating cells a good model system to understand the molecular basis of human diseases associated with protein synthesis errors?

2. Adaptation of human cells to protein synthesis errors

2.1 Abstract

Protein synthesis is a highly-regulated process and maintenance of its fidelity is essential to life. Alterations in the components of the protein synthesis machinery, namely tRNAs, aminoacyl-tRNA synthetases (aaRS) or tRNA modifying enzymes, increase the level of protein synthesis errors (PSE), and are associated with several conditions, from cancer to neurodegeneration. Still, the cause-effect mechanisms remain to be elucidated in many conditions. We hypothesized that accumulation of PSE in human cells activate different protein quality control (PQC) mechanisms depending on the type of error and the duration of the stress stimulus. To address this issue, we modified the anticodon of a human serine transfer RNA ($tRNA^{Ser}$), to incorporate the amino acid serine (Ser) at various non-cognate sites and overexpressed the wild type (Wt) $tRNA^{Ser}$ to evaluate the effects of tRNA misexpression. Stable HEK293 cell lines were produced and analyzed at different time points (cells passages). As expected, mutant tRNAs and tRNA pool deregulation led to accumulation of misfolded proteins. Activation of the ubiquitin-proteasome system (UPS) and the unfolded protein response (UPR) fully protected these cells from proteotoxic stress, maintaining viability intact. Evolution of these cell lines showed differential adaptation responses to different types of tRNAs' misexpression. In some cases, adaption was mainly due to increased protein turnover, while in other cases, UPR activation with consequent protein synthesis inhibition was the main adaption mechanism. Our data provide new insight on how mammalian cells cope and adapt to proteotoxic stress induced by PSE.

2.2 Introduction

Tight control of protein synthesis is essential for cell functioning, nonetheless the high rate of ribosome decoding, which is required to maintain proteome homeostasis, affects the accuracy of mRNA translation and proteins are synthesized with some level of error. Indeed, lowering translation rate increases protein synthesis accuracy, but impacts negatively on growth rate and fitness (168,169). Protein synthesis errors (PSE) can arise during both aminoacylation of tRNAs and mRNA decoding by the ribosomes (60,66). The rate of eukaryotic protein synthesis error measured under normal experimental conditions, i.e., downstream of protein quality control (PQC) processes, is between 10^{-3} and 10^{-4} (63,66). Since most polypeptides containing erroneous amino acids are degraded, the real rate of amino acid misincorporation is much higher than those values, suggesting that defective PQC can have

catastrophic consequences for the cell. Similarly, stress conditions, in particular amino acid starvation, and metabolic deregulation associated with pathology and aging increase error frequency (88). In other words, a fraction of defective proteins is continually produced by cells making the existence of perfect proteomes an impossible task (67).

Not surprisingly, deregulation of protein synthesis factors that maintain translational accuracy has been associated with human diseases (17,41). For instance, mutations in genes encoding the glycyl-tRNA synthetase (GlyRS), have been found in patients with Charcot-Marie-Tooth neuropathy (23,151). Mutations in mitochondrial $tRNA^{Leu}$ are linked to mitochondrial encephalomyophathy, lactic acidosis, and stroke-like episodes (MELAS) (152,154). Deregulation of the tRNA pool and tRNA modifying enzymes have been observed in cancer (16,158,161,170–172). In breast tumors there is strong upregulation (> 10 fold) of nuclear and mitochondrial encoded tRNAs (158). The tRNA modifying enzyme TRMT12 (tRNA methyltransferase homolog 12, involved in the formation of wybutosine at position 37 on $tRNA^{Phe}$) is overexpressed in various breast tumors (16,161). Despite clear association with disease, how such deregulation and mutations causes disease are poorly understood. One possibility is that PSE may saturate PQC, leading to accumulation of misfolded and aggregated proteins (69,165,167,173,174). For example, molecular chaperones selectively recognize misfolded proteins and promote refolding or degradation via the ubiquitin-proteasome system (UPS) (88,175). Accumulation of misfolded proteins in the endoplasmic reticulum (ER) normally saturates ER chaperones (as GRP78/BiP), triggering the activation of the unfolded protein response (UPR) through its branches, namely ATF6, IRE1 or PERK (124). These pathways, regulate protein synthesis, decrease ER load, increase ER folding capacity and increase the degradation of misfolded proteins (176). In mice, a mutation in the editing domain of alanyl-tRNA synthetase (AlaRS), leads to the formation of toxic aggregates, an increase in protein ubiquitination, formation of autophagosomes, induction of molecular chaperones (members of Hsp70 family) and upregulation of the UPR (induction of BiP and CHOP). This ultimately leads to Purkinge cell death (69). Some cancer cells have increased growth rate, regulated by signals related to proliferation, metabolism and protein synthesis, explaining the global upregulation of tRNAs and tRNA modifying enzymes (177). This also induces ER stress and activation of UPR pathways (131,178), suggesting that protein misfolding, and eventually PSE, may be prevalent in cancer.

The objective of this study was to clarify the poorly understood consequences of PSE in human cells. In particular, we wanted to elucidate how these cells cope and adapte to such errors and whether they affect cell viability, activate the UPR and produce phenotypes that are

common in diseases associated with deregulation of protein synthesis factors. To address these questions, we have created HEK293 cell lines expressing mutant tRNAs that randomly misincorporate serine (Ser) at alanine (Ala), leucine (Leu) and histidine (His) codon sites, on a proteome wide scale. A cell line expressing extra copies of the wild type (Wt) $tRNA_{AGA}^{Ser}$ was also produced to gain insight on the cellular consequences of tRNA imbalances; often observed in tumors. Cellular responses were studied at different time points (cells passages) to clarify how these HEK293 cells adapt to PSE.

Our data show that immortalized cells cope relatively well with the presence of mutant tRNAs or overexpression of the Wt $tRNA^{Ser}$. We observed accumulation of aggregated and ubiquitinated proteins and activation of the UPS and UPR, as expected. Activation of PQC mechanisms is differential according to the type of error introduced in proteins and to the cell passage. Additionally, the expression of mutant tRNAs decreased as cell passage increase, suggesting that human cells modulate expression of mutant tRNAs to decrease their destabilizing effect on the proteome.

2.3 Results

2.3.1 Human cell line models of PSE

Previous studies from Geslain and colleagues reported a novel approach for the induction of stress responses to protein aggregation, based on engineered tRNAs in HEK293 cells (165). This strategy allows the introduction of mutations of increasing severity randomly in the proteome. The decoding sequence (anticodon) of $tRNA^{Ser}$ was altered, not influencing the recognition by the seryl-tRNA synthetase (SerRS). The altered tRNA is aminoacylated with Ser, but will be used by the ribosome to translate codons complementary to the engineered anticodon. tRNAs produced by Geslain were tested for their ability to restore fluorescence of a GFP reporter, where the essential residue serine-65 (S65) had been substituted for the codons recognized by the engineered tRNAs, proving that mutant tRNAs are fully functional in HEK293 cells. Amino acid analysis of purified GFP from cells transfected with some of these tRNAs ($tRNA^{Ser}$(Lys), $tRNA^{Ser}$(His), $tRNA^{Ser}$(Ile)) showed that the mutable residues had been replaced by serine (165).

In the current study, and to induce PSE in HEK293 cells, we mutated the anticodon of the human Wt $tRNA_{AGA}^{Ser}$ to produce a set of tRNAs: $tRNA_{AGC}^{Ser}$, $tRNA_{AAG}^{Ser}$, $tRNA_{GTG}^{Ser}$ that misincorporate Ser at Ala, Leu and His codon sites, respectively (Figure 2-1).

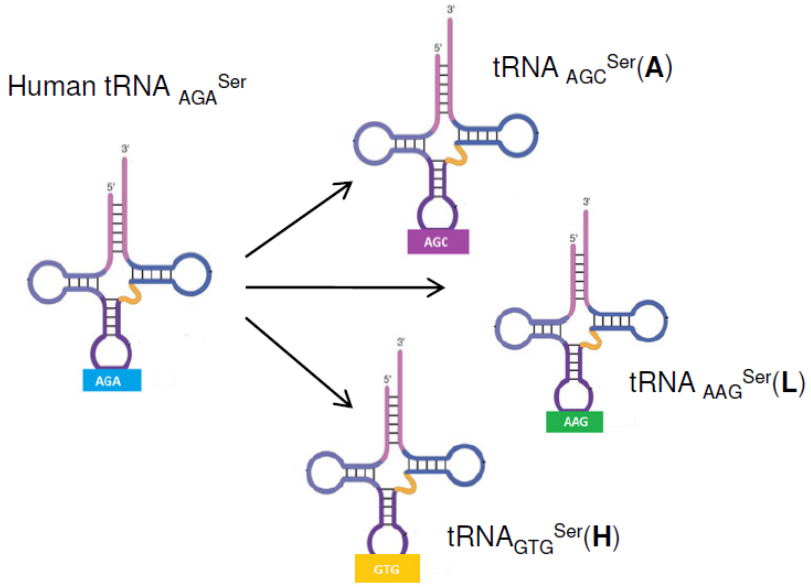


Figure 2-1. Representation of the human $tRNA_{AGA}^{Ser}$ and mutant tRNAs used in the study.

These mutations do not interfere with serylation of the tRNAs by the SerRS, because this enzyme recognizes the extra-arm and discriminator base, rather than the anticodon of $tRNA^{Ser}$ (13,179). Participation of these tRNAs in protein synthesis leads to random incorporation of Ser at non-cognate sites, synthesis, misfolding, aggregation and degradation of the faulty proteins (Figure 2-2) (165,167).

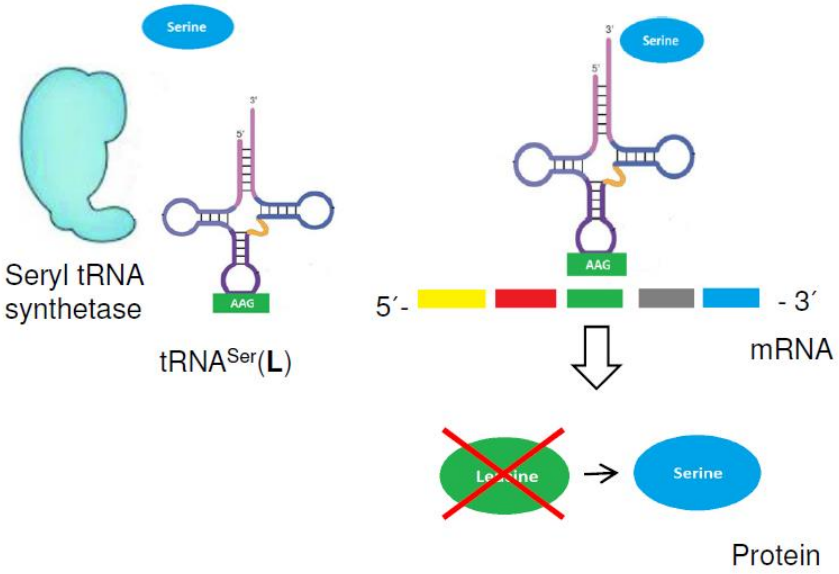


Figure 2-2. Schematic representation of PSE incorporation by the mutant $tRNA_{AAG}^{Ser}$. SerRS is not able to discriminate between endogenous $tRNA^{Ser}$ and mutant tRNAs and charges them equally, leading to Ser misincorporation at Leu codon sites.

We chose these alterations in the anticodon to have a broad spectrum of amino acid chemical differences. Ser is polar and hydrophilic and is normally present on protein surfaces, whereas Ala is nonpolar, hydrophobic and is found inside or outside proteins. Leu is hydrophobic and is generally buried in folded proteins, while His is a basic and polar residue (180). Therefore, Ser misincorporation severity should be higher in the case of misincorporation at Leu codon sites and much lower at Ala codon sites, depending also on the usage of those codons, the competition between endogenous tRNAs ($tRNA_{AGC}^{Ala}$, $tRNA_{AAG}^{Leu}$, $tRNA_{GTG}^{His}$) in the ribosome, and the competition of endogenous $tRNA_{AGA}^{Ser}$ and mutant tRNAs for the SerRS.

HEK293 cells were transfected with the plasmid pIRES2-DsRed containing one copy of each mutant tRNA. The resulting cell lines were denominated: $tRNA^{Ser}$(A), $tRNA^{Ser}$(L), $tRNA^{Ser}$(H). Two additional cell lines were also produced: one was transfected with the empty plasmid; Mock (negative control) and another misexpressing the Wt $tRNA_{AGA}^{Ser}$ ($tRNA^{Ser}$(S)). To gain insight on the long-term adaptation to PSE, three time points (P1, P15 and P30, corresponding to the number of cell passages after transfection and selection in geneticin containing media) were studied. Mutant tRNA expression in these cell lines was monitored by Sanger sequencing during evolution in culture. After P30, some of these cell lines presented additional mutations in the recombinant tRNA genes, namely in the acceptor arm, and for this reason cells were only studied till P30. Each cell line was compared with the control (Mock) of each passage, and for each cell line the values of the three passages (P1, P15 and P30) were compared among each other.

2.3.2 Expression and copy number of mutant tRNAs and $tRNA^{Ser}$ in HEK293 cells

Exogenous tRNA expression was determined using a primer extension reaction (SNaPshot analysis), which allowed to specifically detect the tRNAs with the altered nucleotide in the anticodon. We were able to detect in all cell lines, both the expression of the endogenous $tRNA^{Ser}$ (Figure 2-3 A) and each mutant tRNA (Figure 2-3 B). In the Mock cell line the expression of the endogenous $tRNA^{Ser}$ did not change from P1 to P30 (values around 5 in arbitrary units (a.u.)), but the expression of $tRNA^{Ser}$ in the $tRNA^{Ser}$(S) cell line (both the endogenous and the exogenous tRNA are detected) altered during our timeline (Figure 2-3 A). The expression of $tRNA^{Ser}$ increased slightly from passage P1 (8.9 a.u.) to P15 (12.7 a.u.) and then decreased from P15 (12.7 a.u.) to P30 (2.9 a.u.) (Figure 2-3 A), suggesting that increasing

the levels of the Wt $tRNA^{Ser}$ may be slightly advantageous at the beginning of the evolution, but becomes deleterious in the long term. Expression of all mutant tRNAs decreased gradually from P1 to P30 (Figure 2-3 B). In P1, in $tRNA^{Ser}$(A) the levels of $tRNA_{AGC}{}^{Ser}$ were 0.58 a.u. (60% of the endogenous $tRNA^{Ser}$ expression) and in P30 declined to 0.1 a.u. (10% of the endogenous $tRNA^{Ser}$ expression). The relative levels of the $tRNA_{AAG}{}^{Ser}$(in $tRNA^{Ser}$(L) cells) and $tRNA_{GTG}{}^{Ser}$(in $tRNA^{Ser}$(H) cells) in P1 were 0.3 a.u. (30% of the endogenous $tRNA^{Ser}$ expression) and declined to 0.05 a.u. and 0.03 a.u. (5% and 3% of the endogenous $tRNA^{Ser}$ expression) in P30, respectively (Figure 2-3 B).

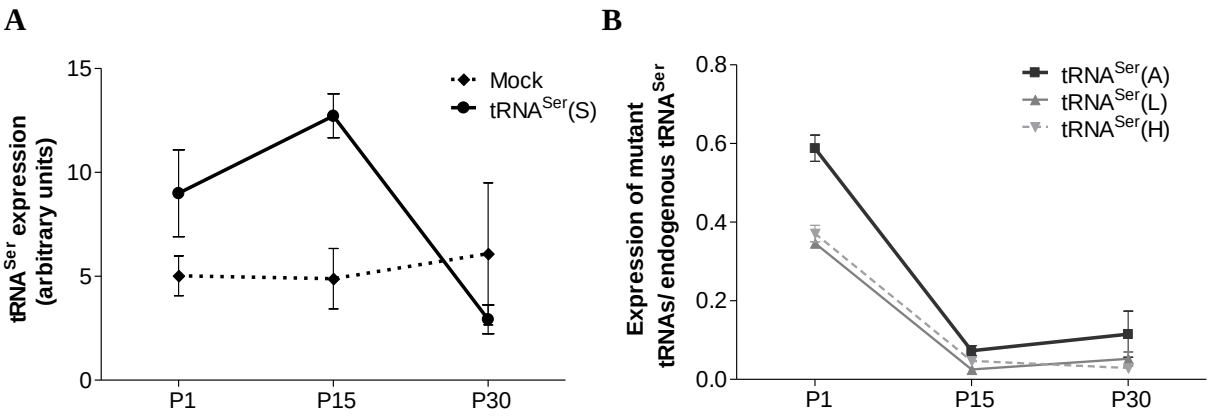


Figure 2-3. Quantification of $tRNA^{Ser}$ and mutant tRNAs expression. A – Detection of $tRNA_{AGA}{}^{Ser}$ expression in Mock and $tRNA^{Ser}$(S) cell lines, assessed by the SNaPshot assay. **B –** Expression of mutant tRNAs relative to the Wt endogenous $tRNA_{AGA}{}^{Ser}$ determined by SNaPshot in $tRNA^{Ser}$(A), $tRNA^{Ser}$(L) and $tRNA^{Ser}$(H) cell lines. tRNA values were normalized to an endogenous control, GAPDH. Data represents Average±SEM of one biological replicate and at least two technical replicates.

To clarify if the gradual decrease in tRNA levels observed during the P1-P30 evolution were caused by differences in tRNA gene copy number we have also used the SNaPshot technique to detect tDNA insertion into the genome of HEK293 cells (Figure 2-4). The number of copies of the $tRNA^{Ser}$ in the genome of Mock and $tRNA^{Ser}$(S) cell lines was not altered from P1 to P30. As expected, the copy number of $tRNA^{Ser}$ was higher in $tRNA^{Ser}$(S) cells than in the Mock cell line, confirming incorporation of the plasmid into the genome (Figure 2-4 A). In the $tRNA^{Ser}$(A) cell line the number of copies of the tDNA decreased from P1 (0.26 a.u., 26% of $tRNA^{Ser}$ copies) to P30 (0.15 a.u., 15% of $tRNA^{Ser}$copies). In $tRNA^{Ser}$(L) and $tRNA^{Ser}$(H) cell lines the number of tDNA copies was maintained. In $tRNA^{Ser}$(L) cells, tDNA values were around 0.06 a.u. (6% of $tDNA^{Ser}$ copies) and in $tRNA^{Ser}$(H) cells, tDNA showed values around

0.09 a.u. (9% of tDNA^{Ser} copies). The tRNA^{Ser}(A) cell line incorporated more copies of the respective tDNA, relative to tRNA^{Ser}(L) and tRNA^{Ser}(H) cell lines (Figure 2-4 B).

Therefore, the expression of the tRNAs and respective tDNA copy number are well correlated indicating that the former may be due to transcriptional regulation by Pol III or tRNA degradation, rather than loss of tDNA copies during evolution. tRNA^{Ser}(A) cell line, is an exception since the decreased mutant tRNA expression is accompanied by a decrease in the copy number.

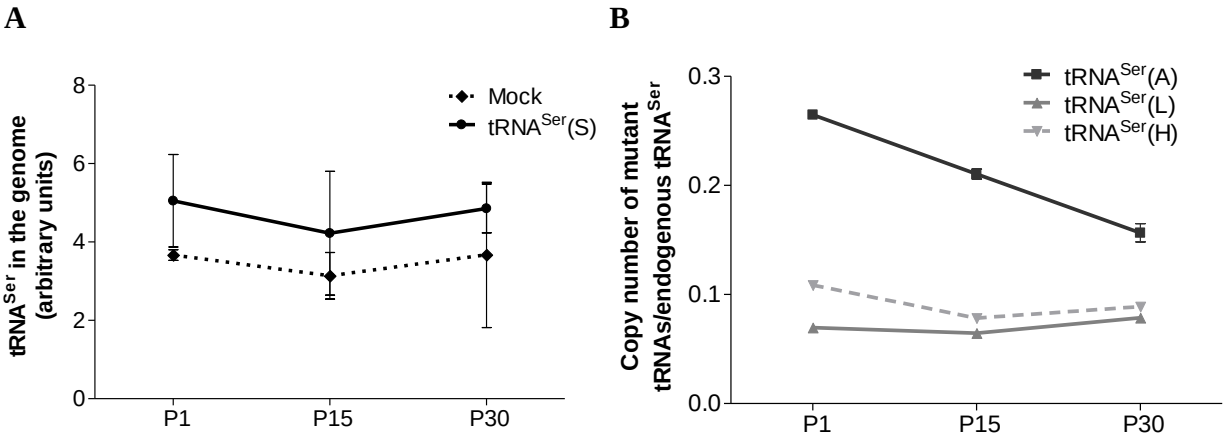


Figure 2-4. Copy number of tRNA^{Ser} and mutant tRNAs. A – tDNA$_{AGA}$^{Ser}genomic copy number in Mock and tRNA^{Ser}(S) cell lines, assessed by SNaPshot. **B –** Genomic copy number of the mutant tDNA genes relative to the Wt endogenous tRNA$_{AGA}$^{Ser}, assessed by SNaPshot, in tRNA^{Ser}(A), tRNA^{Ser}(L) and tRNA^{Ser}(H) cell lines. tRNA values were normalized to an endogenous control, GAPDH. Data represents Average±SEM of one biological replicate and at least two technical replicates.

2.3.3 Phenotypic effects of PSE

We have accessed cell viability, proliferation, number of anchorage-dependent colonies formed, protein synthesis rate and accumulation of insoluble proteins, as readouts of putative phenotypic effects produced by PSE emerging during evolution of our cell line models.

We used the number of cell passages as landmark of time points during evolution. Since overexpression of Wt tRNA^{Ser} and expression of mutant tRNAs did not affect doubling time of HEK293 cells (approximately 20h), we believed that all cells were approximately in the same generation along the experiment (Figure 2-5).

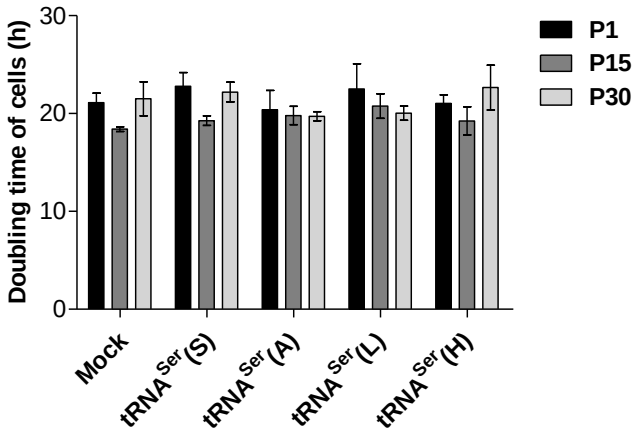


Figure 2-5. Doubling time of cells, assessed by cell counting with Tripan blue. Values represent Average±SEM of three independent experiments in triplicate. One-way ANOVA followed by Dunnett's post-test was used to assess differences between the Mock cell line and cells misexpressing the Wt $tRNA^{Ser}$ ($tRNA^{Ser}$(S)) and expressing mutant tRNAs ($tRNA^{Ser}$(A), $tRNA^{Ser}$(L), $tRNA^{Ser}$(H)) (p>0.05).

In order to test the toxicity of mutant tRNAs and increased copies of the $tRNA^{Ser}$, cell viability was assessed using Tripan blue. Overall, viability was not compromised by the expression of mutant tRNAs or alteration in the tRNA pool (Figure 2-6 A). However, $tRNA^{Ser}$(A) and $tRNA^{Ser}$(L) cell lines showed an increase in viable cells (5.14% and 6.08% respectively), relatively to the Mock in P1 (Figure 2-6 A). This was coincident with the passage in which these mutant tRNAs were most expressed (Figure 2-3 B). During evolution, $tRNA^{Ser}$(A) and $tRNA^{Ser}$(L) cells showed the same pattern of viability, which decreased from P1 and P15 (93% in P1 and a 88% in P15 of viable cells), and recovered from P15 to P30, to a level of viable cells similar to that observed in P1 (~ 93% of viable cells) (Figure 2-6 B).

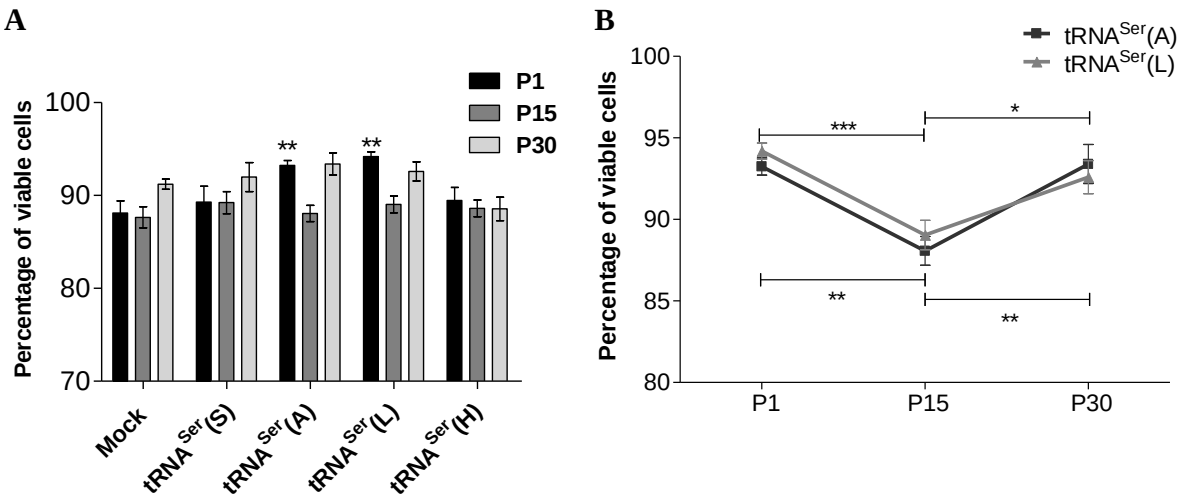


Figure 2-6. Percentage of viable cells in culture determined by cell counting with Tripan blue. A – Percentage of viable cells in comparison with the Mock in each passage (one-way ANOVA, Dunnett's post-test,**p<0.01) **B –** Percentage of viable cells in each time point in tRNA^{Ser}(A) and tRNA^{Ser}(L) cell lines (one-way ANOVA, Bonferroni's post-test, *p<0.05;**p<0.01;***p<0.001). Values were normalized to the Mock cell line of each passage and represent Average±SEM of three independent experiments in triplicate.

Regarding cell proliferation, assessed with a DNA synthesis–based cell proliferation assay (BrdU), in P1 comparatively to the Mock, there were no alterations (Figure 2-7). In P15, the tRNA^{Ser}(A) cell line showed higher proliferation than Mock (1.43 fold), while tRNA^{Ser}(H) cell line showed a decrease of proliferation (0.64 fold) relative to Mock (Figure 2-7 A). In P30, both tRNA^{Ser}(S) and tRNA^{Ser}(A) cell lines proliferated more than the Mock cell line (1.08 fold in both). The tRNA^{Ser}(H) cell line, which showed decrease in proliferation from P1 to P15 (0.33 fold), restored its proliferation values in P30 (1.03 fold) (Figure 2-7 B). The tRNA^{Ser}(A) cell line, which showed an increase in proliferation from P1 to P15 (0.36 fold), in P30 recovered the proliferation values observed in P1 (Figure 2-7 B).

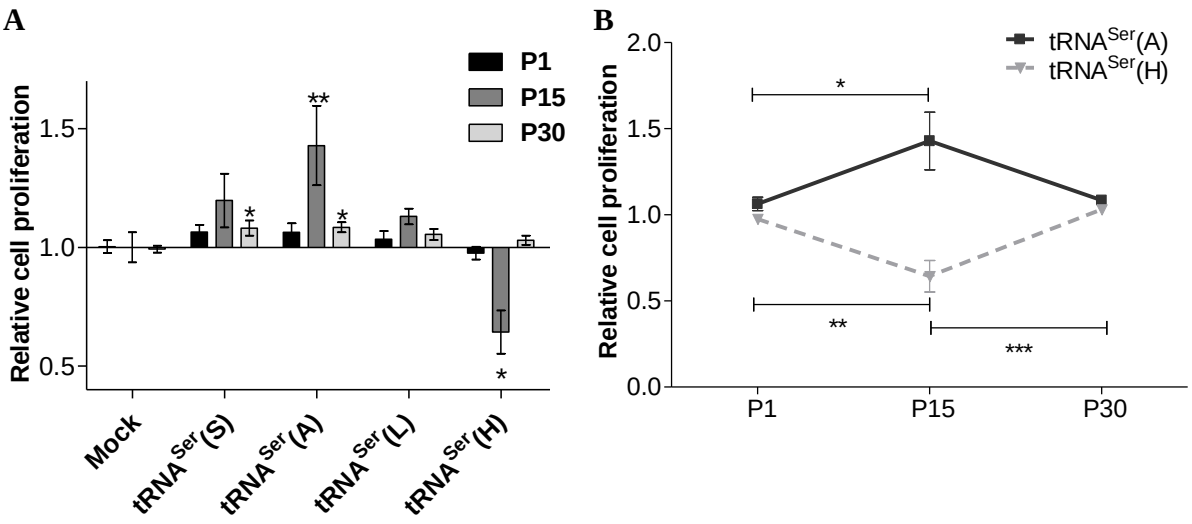


Figure 2-7. Relative cell proliferation determined using a BrdU ELISA Kit. A – Cell proliferation in comparison with Mock in each passage (one-way ANOVA, Dunnett's post-test, (**p*<0.05;***p*<0.01) **B –** Relative proliferation of tRNA^{Ser}(A) and tRNA^{Ser}(H) cell lines (one-way ANOVA, Bonferroni's post-test, **p*<0.05;***p*<0.01;****p*<0.001).Values were normalized to the Mock cell line of each passage and represent Average±SEM of three independent experiments in triplicate.

Cell proliferation and survival was also assessed using anchorage-dependent colony formation assay, which consists in the ability of a single cell to grow into a colony. Only tRNA^{Ser}(H) cell line had lower colony formation capacity relative to Mock cell line in P15, with an average 12.87 colonies for tRNA^{Ser}(H) vs. 18.44 for Mock (Figure 2-8). This was consistent with proliferation data, as tRNA^{Ser}(H) was the only cell line that displayed decreased proliferation, at this same passage.

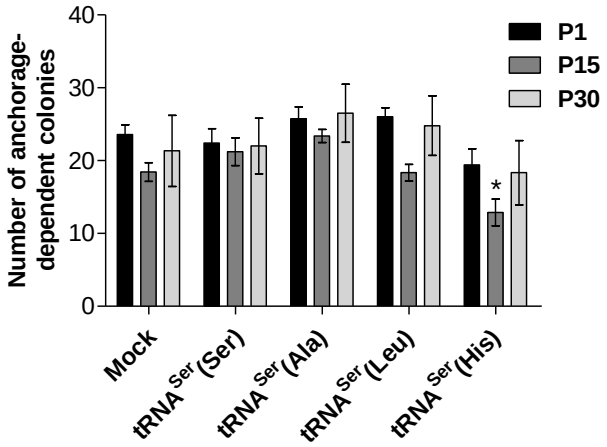


Figure 2-8. Number of colonies formed after 12 days in culture. Values were normalized to the Mock cell line of each passage and represent Average±SEM of three independent experiments in duplicate (one-way ANOVA, Dunnett's post-test, **p*<0.05)

Since these engineered cell lines are error prone at the level of protein synthesis, we questioned if protein synthesis rate could be affected during the evolution of these cell lines along 30 passages in culture. To measure protein synthesis rate, we took advantage of the SunSET method, based in the detection of puromycin incorporation into proteins, as described in the methods section (181). Protein synthesis rate increased in $tRNA^{Ser}$(A) cells in P1 (1.23 fold) and P15 (1.48 fold) relative to Mock, but decreased in the $tRNA^{Ser}$(L) cells in P30 (0.61 fold) relative to Mock (Figure 2-9 A). Despite showing increased protein synthesis rate relative to Mock, when we compare the three time points (P1, P15 and P30) of $tRNA^{Ser}$(A) cell line, there was a decrease from 1.23 fold in P1 to 0.76 fold in P30. The same occurred in $tRNA^{Ser}$(L) cells, in which protein synthesis rate decreased from 1.20 fold in P1 to 0.61 fold in P30 (Figure 2-9 B).

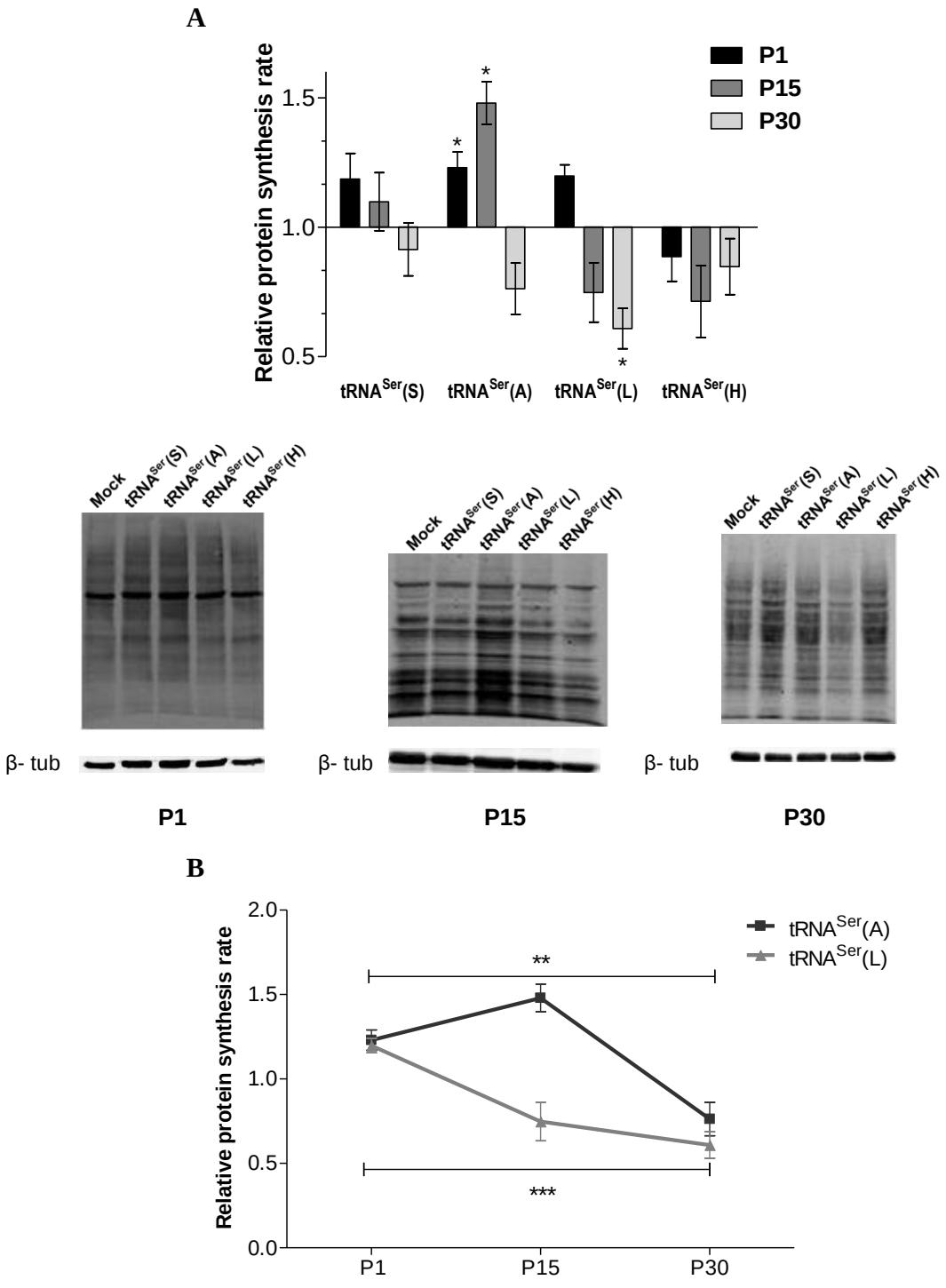


Figure 2-9. Protein synthesis rate determined by SunSET method adapted to immunoblot with anti-puromycin. A – Top panel: Relative protein synthesis rate in comparison with Mock in each passage (one-way ANOVA, Dunnett's post-test, **p*<0.05). Mock values for each passage were considered 1 and were not represented in the graph. **Lower panel:** Representative immunoblot images for each time point and cell line and β-tubulin. **B –** Relative protein synthesis rate during evolution of tRNA^{Ser}(A) and tRNA^{Ser}(L) cell lines (one-way ANOVA, Bonferroni's post-test, ***p*<0.01;****p*<0.001). Values were normalized to the Mock cell line of each passage and represent Average±SEM of at least three independent experiments in triplicate.

To test whether mutant and Wt $tRNA^{Ser}$ destabilized the proteome and lead to misfolded protein accumulation and aggregates formation, we have quantified the insoluble protein fraction in P1, P15 and P30. In P15, there was a tendency for increasing levels of insoluble proteins, but it was only statistically significant for the $tRNA^{Ser}$(L) cell line (2.66 fold) (Figure 2-10). This protein aggregation effect was also transient, as the level of insoluble proteins returned to control levels at P30. This result may be explained by the lower expression levels of the mutant tRNAs in P30, but it may also indicate that cells counteracted protein aggregation, through activation of PQC mechanisms. Interestingly, the level of insoluble proteins decreased from P15 (2.1 fold) to P30 (0.71 fold) in the $tRNA^{Ser}$(S) cell line that misexpresses the Wt $tRNA^{Ser}$ (Figure 2-10 B).

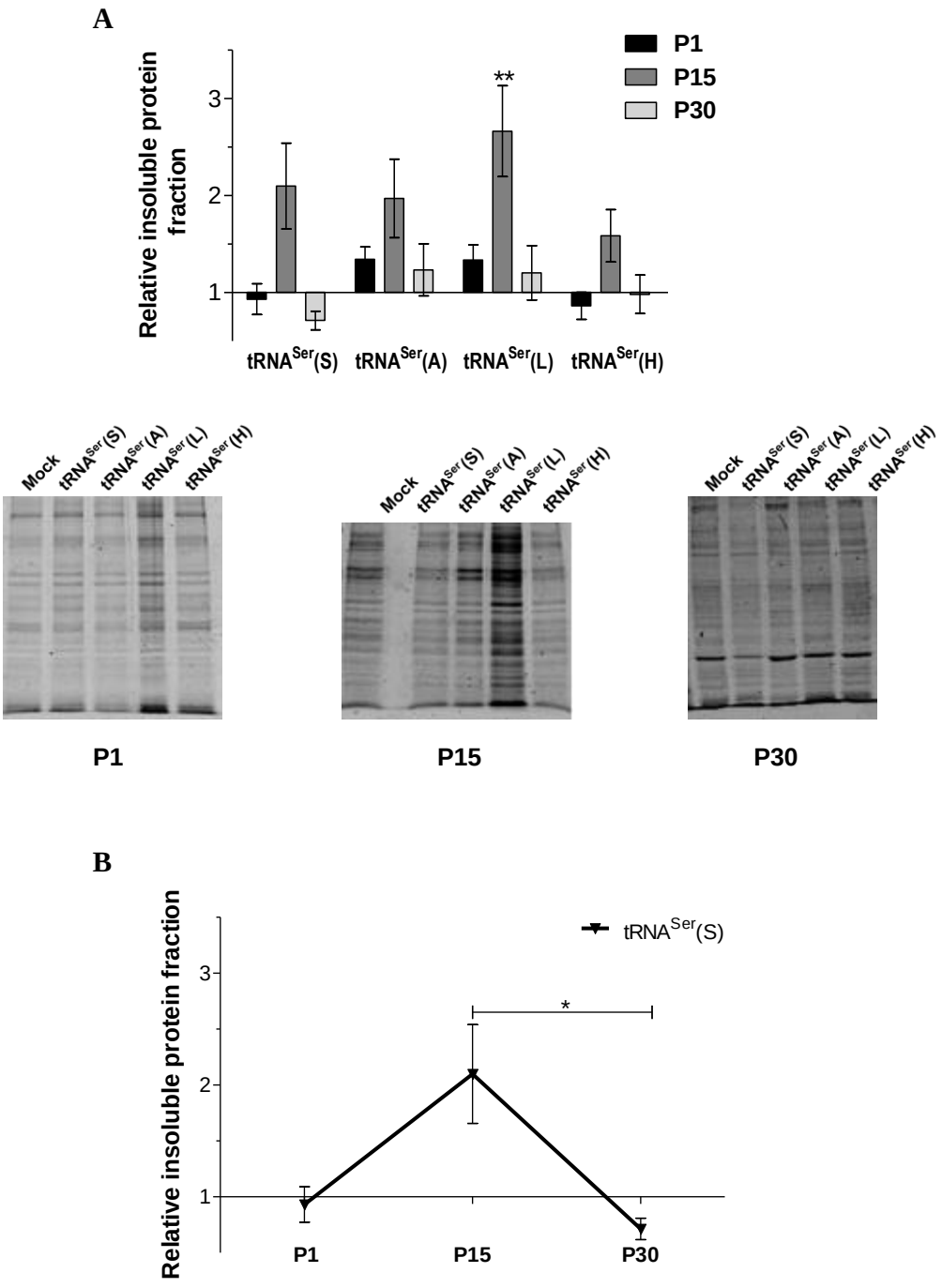


Figure 2-10. Relative insoluble protein fraction. **A – Top panel:** Relative insoluble fraction in comparison with Mock in each passage (one-way ANOVA, Dunnett's post-test, $**p<0.01$). Mock values for each passage were considered 1 and were not represented in the graph. **Lower panel:** Representative acrylamide gel after commassie staining of each time point and cell line. **B –** Protein insoluble fraction during evolution of the tRNA^{Ser}(S) cell line (one-way ANOVA, Bonferroni's post-test, $*p<0.05$). Values were normalized to the Mock cell line of each passage and represent Average±SEM of at least three independent experiments in triplicate.

2.3.4 The impact of PSE in the ubiquitin-proteasome system and molecular chaperones

Eukaryotic cells encompass several mechanisms of PQC to avoid protein aggregation and to eliminate aggregates, if they accumulate. These mechanisms include molecular chaperones, the ubiquitin-proteasome system (UPS), the unfolded protein response (UPR), the endoplasmic reticulum associated protein degradation (ERAD) and autophagy (82,88,182). Since we did not observe significant accumulation of insoluble proteins, particularly in P30, we wondered if erroneous proteins could have been marked by ubiquitin for degradation by the UPS (101,106). We observed increased levels of ubiquitinated proteins in the $tRNA^{Ser}$(H) cell line (1.43 fold) in P1 (Figure 2-11 A). In P15, all other cell lines, $tRNA^{Ser}$(S), $tRNA^{Ser}$(A) and $tRNA^{Ser}$(L), also showed higher levels of ubiquitinated proteins (1.36, 1.37 and 1.48 fold respectively) (Figure 2-11 A). In P30, the amount of ubiquitinated proteins in all cell lines was similar to Mock in the same passage. Importantly, from P15 to P30 the amount of ubiquitinated proteins decreased in all cell lines, being statistically significant in $tRNA^{Ser}$(S) (from 1.36 fold in P15 to 0.96 fold in P30) (Figure 2-11 B). The latter data indicates that these cells were able to somehow degrade all proteins targeted by ubiquitination.

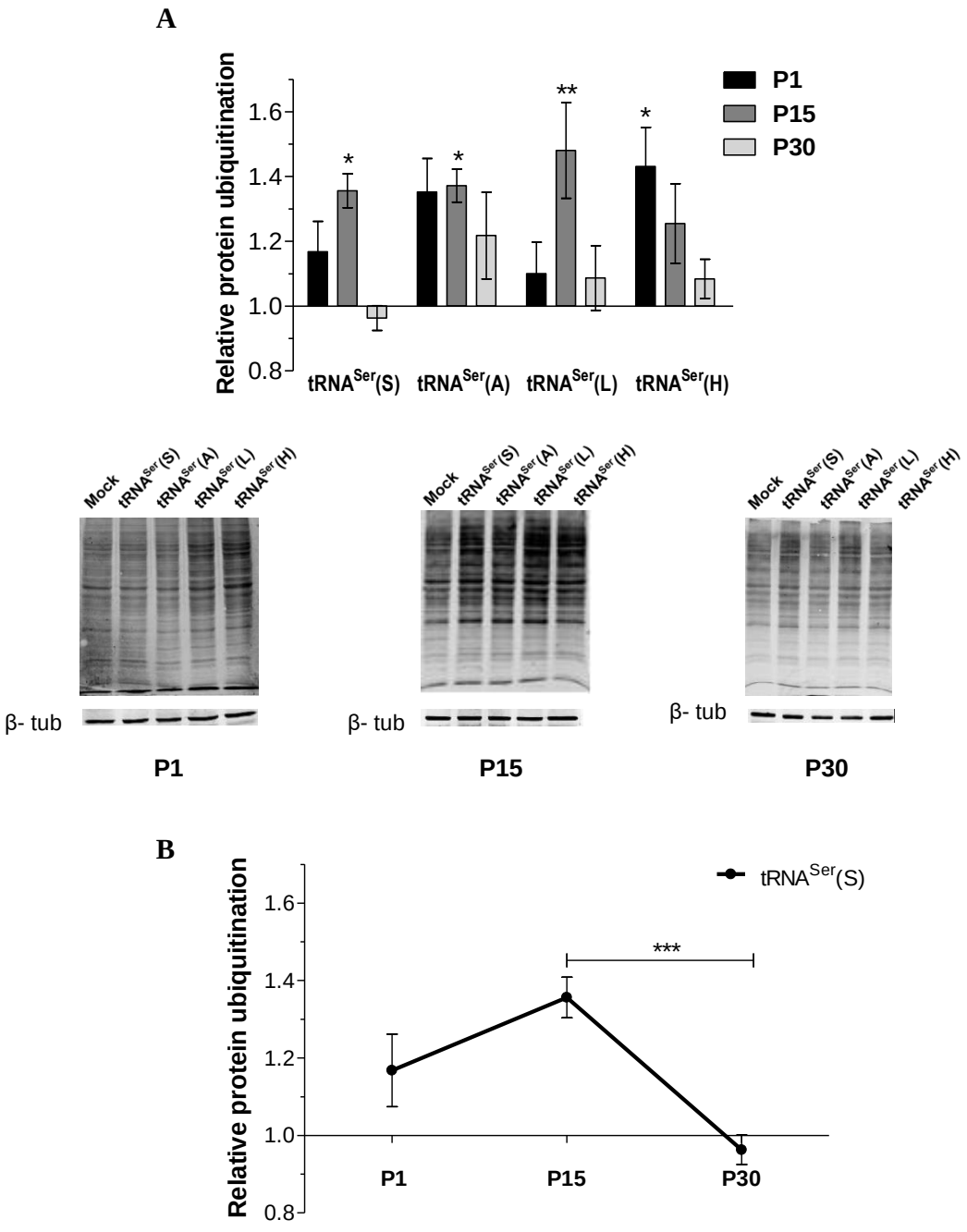


Figure 2-11. Relative protein ubiquitination determined by immunoblot. A – Top panel: Relative protein ubiquitination in comparison with Mock in each passage (one-way ANOVA, Dunnett's post-test, **p*<0.05;***p*<0.01). Mock values for each passage were considered 1 and were not represented in the graph. **Lower panel:** Representative immunoblot images for each time point and cell line plus β-tubulin are shown. **B –** Relative protein ubiquitination during evolution of the RNA^{Ser}(S) cell line (one-way ANOVA, Bonferroni's post-test, ****p*<0.001). Values were normalized to the Mock cell line of each passage and represent Average±SEM of at least three independent experiments in triplicate.

We then checked whether proteasome activity was altered in the same cell lines and time points, using a proteasome activity assay that measures chymotrypsin-like activity. The tRNA^{Ser}(L) cell line showed higher proteasome activity in P1(1.27 fold) (Figure 2-12 A), while tRNA^{Ser}(A) cell line had higher activity relative to Mock in P15 and P30 (2.17 fold and 1.66 fold, respectively) (Figure 2-12 A). However, proteasome activity decreased in the tRNA^{Ser}(L) cell line from 1.85 fold in P15 to 0.97 fold in P30 (Figure 2-12 B). The accumulation of ubiquitinated proteins and insoluble proteins in P15, particularly in this cell line, was likely overwhelming for the proteasome, affecting its degradative capacity. On the other hand, in the tRNA^{Ser}(A) cell line the accumulation of ubiquitinated proteins in P15 was concomitant with increased proteasome activity. tRNA^{Ser}(S) and tRNA^{Ser}(H) cell lines did now show significant alterations in proteasome activity.

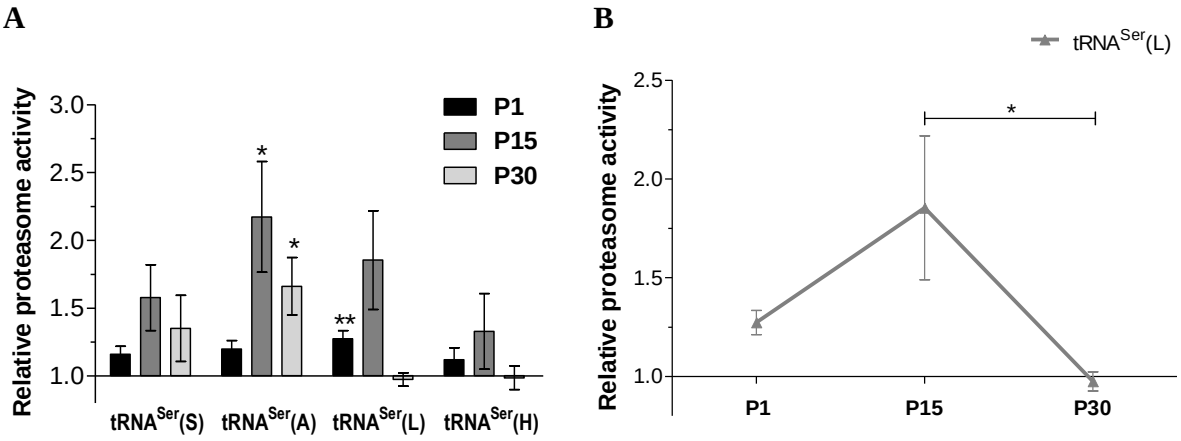


Figure 2-12. Relative proteasome activity. A – Relative proteasome activity assessed by fluorescent measurement of the labeled substrate Suc-LLVY-AMC (one-way ANOVA, Dunnett's post-test, *p<0.05;**p<0.01). Mock values for each passage were considered 1 and were not represented in the graph. **B –** Proteasome activity of the tRNA^{Ser}(L) cell line (one-way ANOVA, Bonferroni's post-test, *p<0.05). Values were normalized to the Mock cell line of each passage and represent Average±SEM of at least three independent experiments in triplicate.

Finally, we checked the molecular chaperones branch of the PQC system. We assessed the expression of those chaperones that are known to play a critical role in protein folding during stress, and whose expression is frequently altered in human diseases, namely Hsp70, Hsp27, Hsp60, Hsp90α and BiP (183,184). Hsp70 (heat shock protein 70) binds to a wide range of nascent polypeptides in stress conditions and by shielding hydrophobic regions, prevents aggregation and promotes proper folding. It also recruits ubiquitin ligases, such as

CHIP (carboxyl terminus of Hsp70-interacting proteins) to tag proteins for proteasomal degradation (108). The expression of Hsp70 did not change in P1 and P15, but decreased in P30 in tRNA^{Ser}(A), tRNA^{Ser}(L) and tRNA^{Ser}(H) cell lines: 0.67, 0.55 and 0.64 fold, respectively (Figure 2-13 A).

In general, the expression of this molecular chaperone decreased during evolution until P30, being this decrease statistically significant in tRNA^{Ser}(A) cells (from 1.02 fold in P1 to 0.67 fold in P30) (Figure 2-13 B).

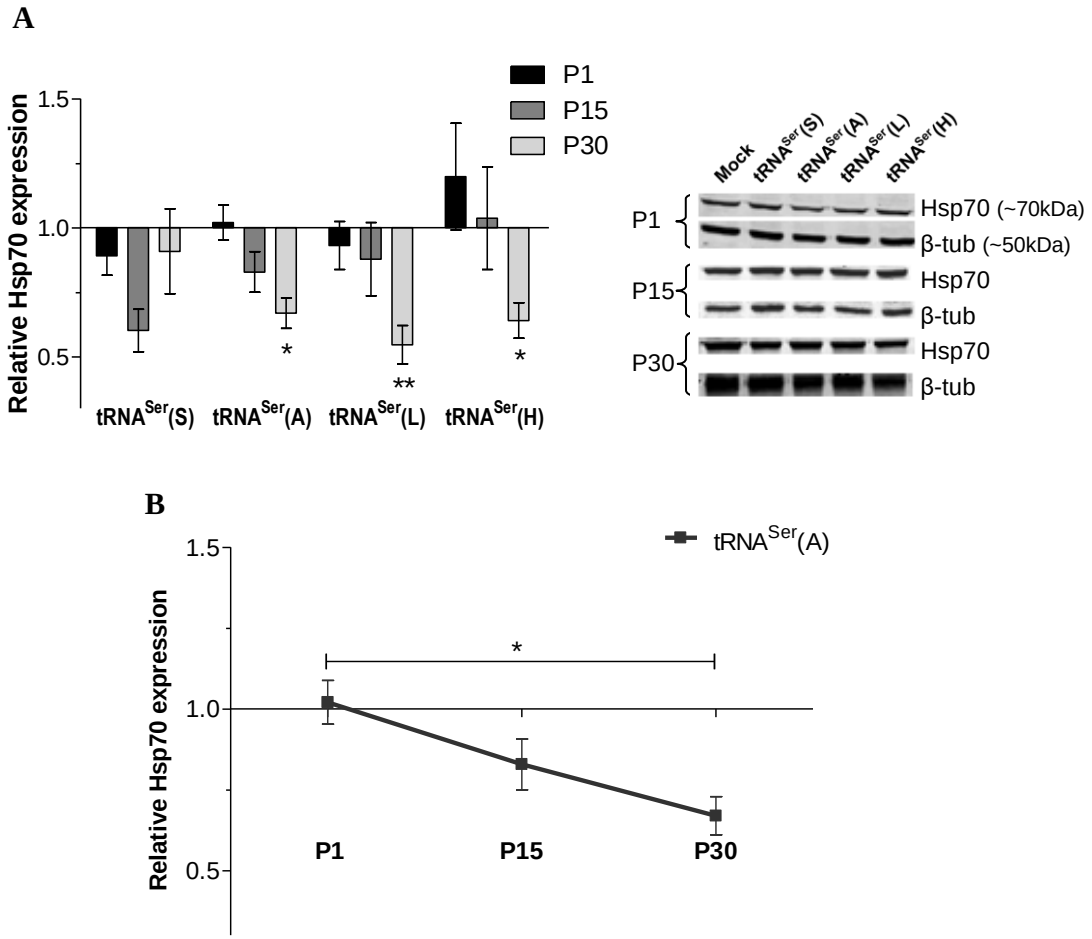


Figure 2-13. Relative HSP70 expression. A – Relative HSP70 expression and respective immunoblots (one-way ANOVA, Dunnett's post-test, **p*<0.05;***p*<0.01). Mock values for each passage were considered 1 and were not represented in the graph. **B –** Expression in the three different time points of HSP70 in tRNA^{Ser}(A) cell line (one-way ANOVA, Bonferroni's post-test, **p*<0.05). Values were normalized to the Mock cell line of each passage and represent Average±SEM of at least three independent experiments in triplicate.

Hsp27 minimizes protein aggregation, by destabilizing aggregates, bind to proteins and aids in the refolding processes, favors degradation of some ubiquitinated proteins by the

proteasome and is involved in apoptotic signaling pathways (185). Its expression increased in $tRNA^{Ser}$(L) cells in P15 (1.69 fold) and decreased in $tRNA^{Ser}$(L) and $tRNA^{Ser}$(H) cells in P30; 0.53 and 0.56 fold, respectively (Figure 2-14 A). In P15, the increase in Hsp27 levels in the $tRNA^{Ser}$(L) cell line was consistent with the higher levels of insoluble proteins, which may indicate that Hsp27 is being recruited to destabilize aggregated proteins. $tRNA^{Ser}$(S) and $tRNA^{Ser}$(A) cell lines showed increased Hsp27 during early evolution from P1 to P15 (from 0.75 to 1.48 fold and from 0.74 to 1.35 fold, in $tRNA^{Ser}$(S) and $tRNA^{Ser}$(A) cells, respectively). However, from P15 to P30 the level of this chaperone decreased (from 1.49 to 0.96 fold and from 1.35 to 0.69 fold, in $tRNA^{Ser}$(S) and in $tRNA^{Ser}$(A) cells, respectively) (Figure 2-14 B). From P15 to P30, the decrease in the expression of Hsp27 was also statistically significant in $tRNA^{Ser}$(L) (from 1.69 to 0.53 fold) and in $tRNA^{Ser}$(H) cells (from 1.31 to 0.56 fold) (Figure 2-14 C). These alterations in Hsp27 expression are consistent with the dynamics of increased protein ubiquitination in these cell lines in P15.

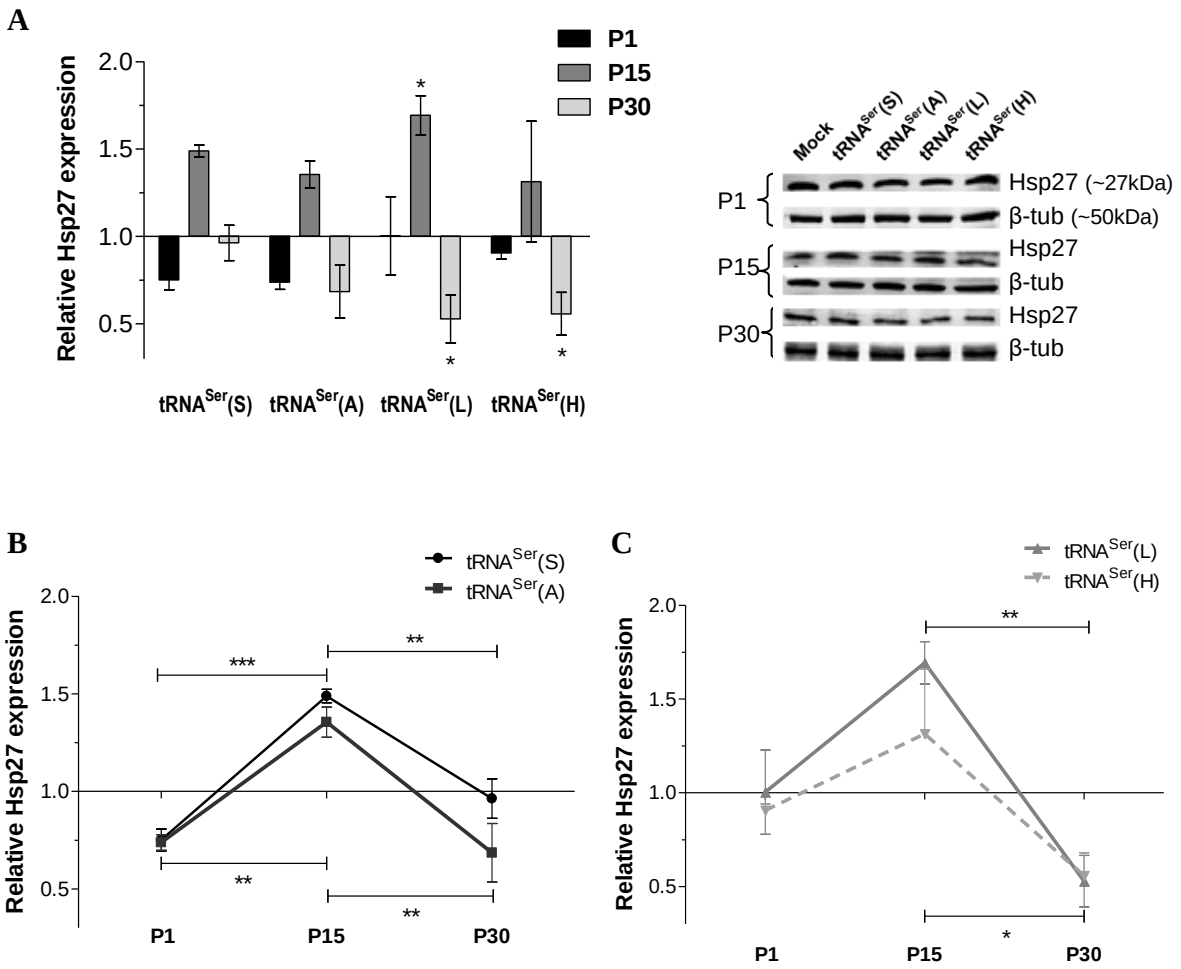


Figure 2-14. Relative Hsp27 expression. A - Relative Hsp27 expression the respective immunoblot (one-way ANOVA, Dunnett's post-test, **p*<0.05). Mock values for each passage were considered 1 and were not represented in the graph. **B and C –** Hsp27 expression during evolution line (one-way ANOVA, Bonferroni's post-test, (**p*<0.05;***p*<0.01; ****p*<0.001).Values were normalized to the Mock cell line of each passage and represent Average±SEM of at least three independent experiments in triplicate.

Hsp60 (heat shock protein 60) is a mitochondrial chaperonin involved in protein refolding in the mitochondrial matrix under stress conditions. In our cell lines, the expression of Hsp60 increased in P15 in tRNA^{Ser}(H) cell line (1.48 fold) and decreased in P30 in tRNA^{Ser}(L) (0.69 fold) and tRNA^{Ser}(H) cell lines (0.76 fold) relative to Mock (Figure 2-15 A). The expression of Hsp60 in these two cell lines, increased from P1 to P15 (from 0.83 to 1.4 fold in tRNA^{Ser}(L) cells and from 0.78 to 1.48 fold in tRNA^{Ser}(H) cells), but decreased from P15 to P30 (from 1.4 to 0.69 fold in tRNA^{Ser}(L) cells and from 1.48 to 0.76 fold in tRNA^{Ser}(H) cells) (Figure 2-15 B), indicating that PSE may also affect protein folding in the mitochondria.

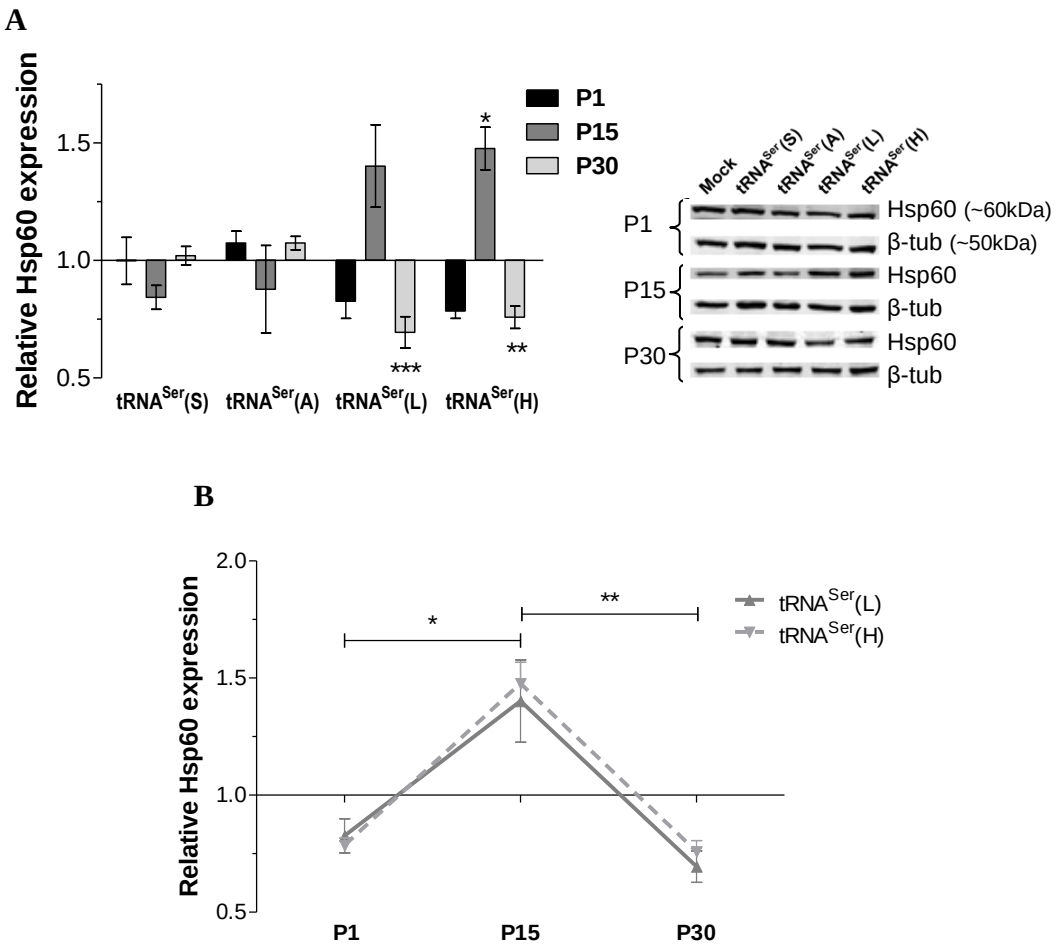


Figure 2-15. Relative Hsp60 expression. A - Relative Hsp60 expression the respective immunoblot (one-way ANOVA, Dunnett's post-test,*p<0.05;**p<0.01; ***p<0.001). Mock values for each passage were considered 1 and were not represented in the graph. **B –** Hsp60 expression during evolution (one-way ANOVA, Bonferroni's post-test, *p<0.05;**p<0.01). Values were normalized to the Mock cell line of each passage and represent Average±SEM of at least three independent experiments in triplicate.

Hsp90α is a molecular chaperone involved in the refolding of specific proteins (normally proteins involved in signal transduction, some of which are ER transmembrane kinases that participate in the UPR) (186). This protein displayed lower levels in tRNA^{Ser}(L) (0.71 fold) and tRNA^{Ser}(H) (0.69 fold) cell lines in P1, relative to Mock in the same passage (Figure 2-16). In tRNA^{Ser}(S) cell line, Hsp90α increased in P30 (1.31 fold) (Figure 2-16).

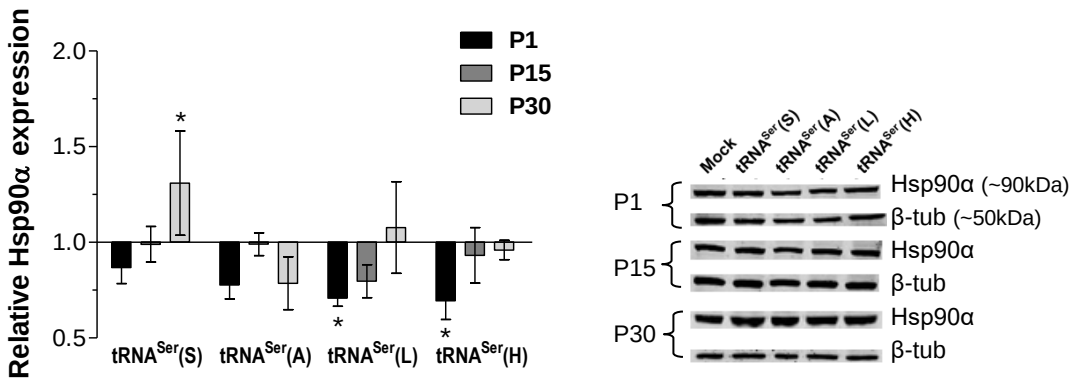


Figure 2-16. Relative Hsp90α expression. Graphic and immunoblot image of Hsp90α expression. Values were normalized to the Mock cell line of each passage and represent Average±SEM of at least three independent experiments in triplicate (one-way ANOVA, Dunnett's post-test, **p*<0.05). Mock values for each passage were considered 1 and were not represented in the graph.

Finally, we accessed the expression of BiP, a molecular chaperone of 70 KDa located in the lumen of the ER that senses ER stress and activates UPR signaling (187). Its expression was low compared to Mock in tRNA^{Ser}(H) cell line (0.61 fold) in P1 and higher compared to Mock in the tRNA^{Ser}(L) cell line in P30 (1.34 fold) (Figure 2-17 A). In tRNA^{Ser}(L) there was upregulation of this chaperone (from 0.77 to 1.34 fold) from P1 to P30 (Figure 2-17 B), suggesting the presence of misfolded proteins in the ER.

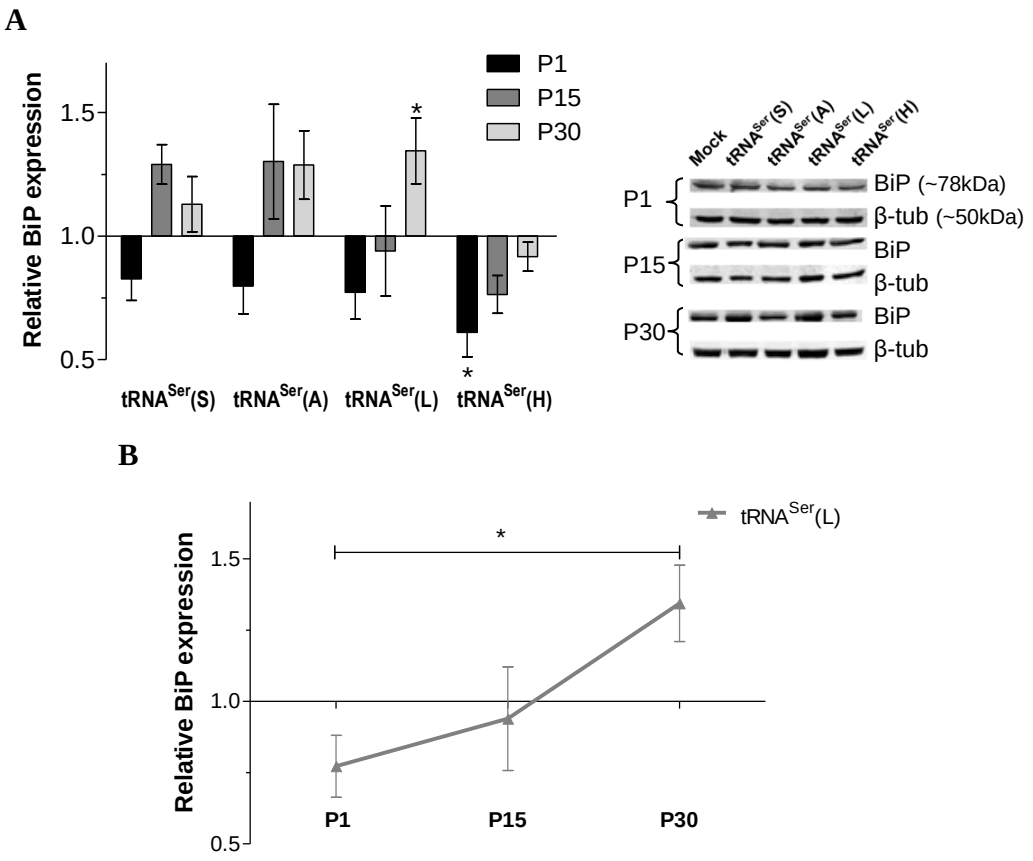


Figure 2-17. Relative BiP expression. A – BiP expression relative to Mock and the corresponding immunoblots (one-way ANOVA, Dunnett's post-test,**p*<0.05). Mock values for each passage were considered 1 and were not represented in the graph. **B –** BiP expression during evolution in tRNA^{Ser}(L) cell line (one-way ANOVA, Bonferroni's post-test, **p*<0.05). Values were normalized to the Mock cell line of each passage and represent Average±SEM of at least three independent experiments in triplicate.

In this part of the study, we have observed accumulation of ubiquitinated proteins in all cell lines, but proteasome activity was only altered in tRNA^{Ser}(A) and tRNA^{Ser}(L) cells. While tRNA^{Ser}(A) cells maintained high proteasome activity, tRNA^{Ser}(L) cells did not. The expression of molecular chaperones changed depending on the cell line and throughout evolution. Hsp70 expression tended to decrease during evolution, while expression of Hsp27 had a peak in P15, which could be correlated with higher levels of ubiquitinated proteins and insoluble proteins in some of the cell lines analyzed. Hsp60 expression was also increased in tRNA^{Ser}(H) cells in P15, but its levels decreased in P30. Hsp90α and BiP presented a different pattern of expression as their expression increased from P1 to P30, in some of our cell lines.

2.3.5 Effects of PSE in the UPR

Accumulation of misfolded proteins in the ER leads to activation of the UPR, which consists in transcriptional activation of genes required for protein folding, ER expansion and ER-associated protein degradation (ERAD). Activation of UPR reduces ER stress, but prolonged activation leads to apoptosis and also the accumulation of reactive oxygen species (ROS) via UPR-regulated oxidative protein folding machinery in the ER, contributing in this way to cell death (188). To clarify whether constitutive PSE could lead to activation of the UPR, we studied some molecular markers of the UPR branches. Activation transcription factor 6 (ATF6) is a transmembrane protein embedded in the ER. Following ER stress-induced proteolysis, it functions as a nuclear transcription factor (176). Although in P1 there were no alterations in the ratio of fragmented ATF6/total ATF6 among cell lines, compared to the Mock, $tRNA^{Ser}$(S) cell line displayed an increase in P15 and $tRNA^{Ser}$(A) an increase in P30 (1.69 fold in both cases) (Figure 2-18 A). In $tRNA^{Ser}$(S) cell line the ratio of fragmented ATF6/total ATF6 increased from P1 to P15 (from 0.72 to 1.69 fold) and decreased from P15 to P30 (from 1.69 fold to 0.67 fold) (Figure 2-18 B).

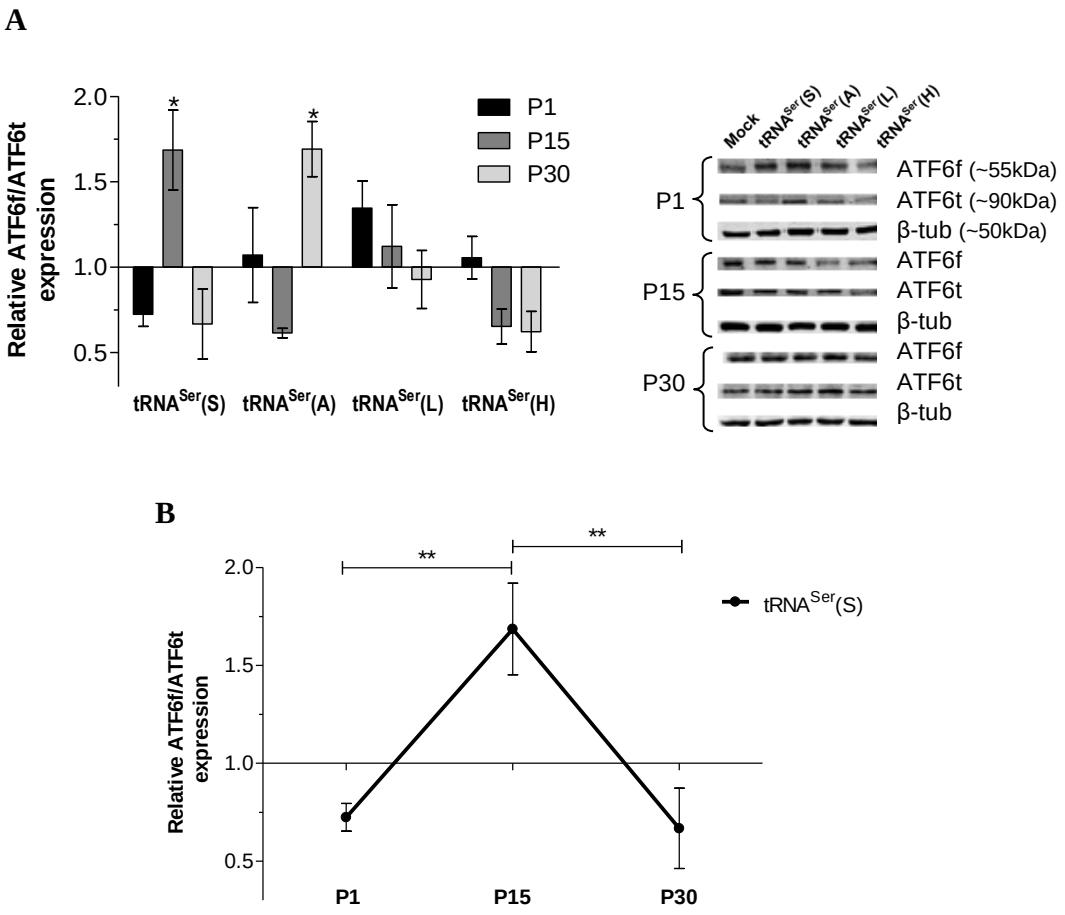


Figure 2-18. Relative ATF6f/ATF6t expression. **A –** ATF6f/ATF6t expression relative to Mock and the corresponding immunoblot (one-way ANOVA, Dunnett's post-test,*p<0.05). Mock values for each passage were considered 1 and were not represented in the graph. **B –** ATF6f/ATF6t expression during evolution in tRNA^{Ser}(S) (one-way ANOVA, Bonferroni's post-test, **p<0.01). Values were normalized to the Mock cell line of each passage and represent Average±SEM of at least three independent experiments in triplicate.

Another marker of UPR activation is the phosphorylation of the eukaryotic initiation factor 2α (eIF2α-P). In cases of ER stress, PERK phosphorylates and inactivates eIF2α, shutting down mRNA translation to reduce protein load in the ER (189). Only in P30 were observed diffrences in the ratio eIF2α-P/total eIF2α in our cell lines. The ratio eIF2α-P/total eIF2α decreased in tRNA^{Ser}(A) (0.53 fold) and increased in tRNA^{Ser}(L) (1.6 fold), comparing to Mock (Figure 2-19 A). In the latter, this happened gradually throughout evolution, from P1 to P30 (from 0.8 to 1.6 fold) (Figure 2-19 B) and this decrease is consistent with the decreased protein synthesis rate in P30 (Figure 2-9). The decrease of eIF2α-P in the tRNA^{Ser}(A) cell line made us wonder whether it was caused by downregulation of PERK, or upregulation of the eIF2α phosphatase PP1's regulatory domain, GADD34. Our data showed an increase in

GADD34 expression only in tRNA^{Ser}(A), in P30 (1.71 fold) (Figure 2-20 A) and from P1 to P30 (from 0.98 to 1.71 fold) (Figure 2-20 B). These results confirm that GADD34 plays an important role in the maintenance of protein translational rate in this particular cell line.

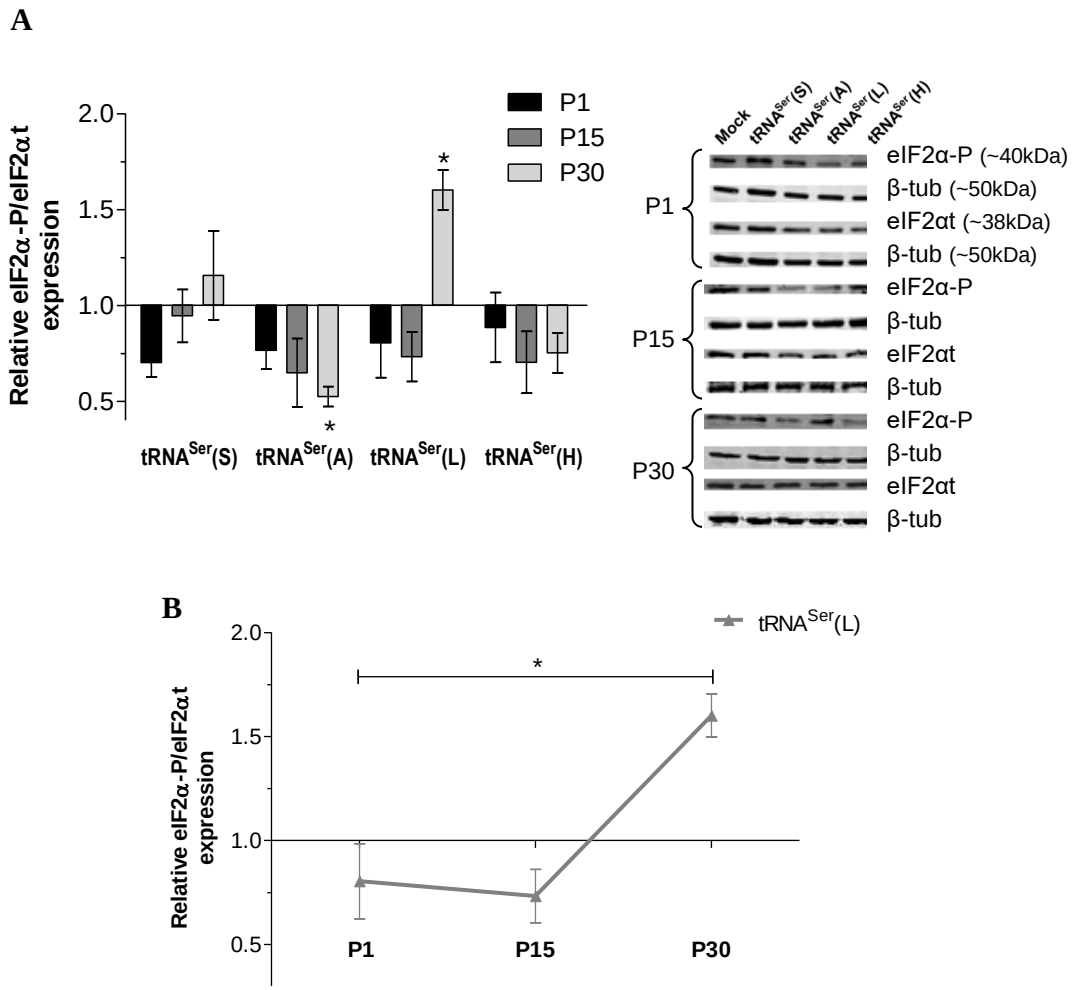


Figure 2-19. Relative eIF2αP/eIF2αt expression. A – eIF2αP/ eIF2αt expression relative to Mock and the corresponding immunoblots (one-way ANOVA, Dunnett's post-test,*p<0.05). Mock values for each passage were considered 1 and were not represented in the graph. **B –** eIF2αP/ eIF2αt expression during evoltution in tRNA^{Ser}(L) cells (one-way ANOVA, Bonferroni's post-test,*p<0.05). Values were normalized to the Mock cell line of each passage and represent Average±SEM of at least three independent experiments in triplicate.

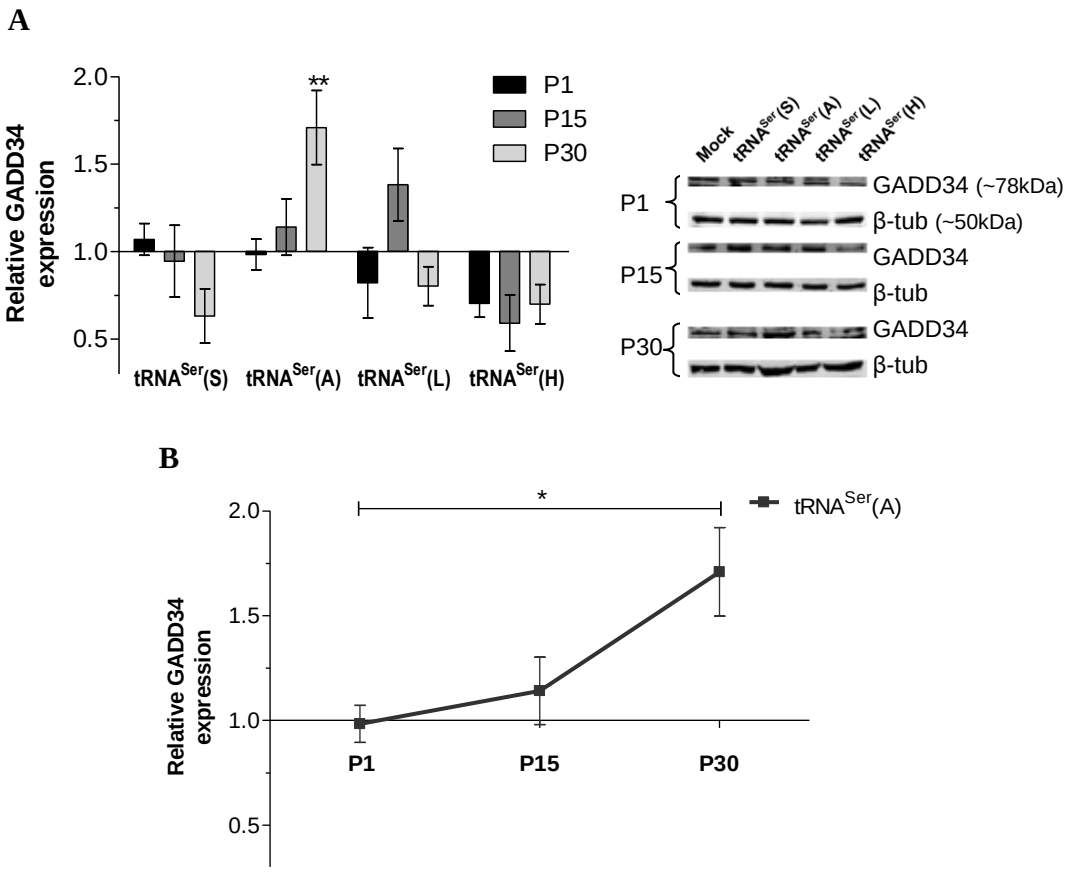


Figure 2-20. Relative GADD34 expression. A – GADD34 expression relative to Mock and the corresponding immunoblots (one-way ANOVA, Dunnett's post-test, **p<0.01). Mock values for each passage were considered 1 and were not represented in the graph. **B –** GADD34 expression during evoltution in tRNA^{Ser}(A) cells (one-way ANOVA, Bonferroni's post-test,*p<0.05). Values were normalized to the Mock cell line of each passage and represent Average±SEM of at least three independent experiments in triplicate.

Therefore, different types of PSE activated different UPR pathways, probably due to differences in ER stress intensity. Indeed, in the tRNA^{Ser}(A) cell line, where protein misfolding levels and ER stress are probably milder (Ser and Ala are chemical similar amino acids), there was an increase in fragmented ATF6, concomitant with increased accumulation of ubiquitinated proteins and proteasome activity. This is physiologically relevant since ATF6 is also responsible for protection against ER stress-induced apoptosis and cell survival (190). However, in the tRNA^{Ser}(L) cell line, where the level of misfolded proteins and ER stress should be more intense due to bigger chemical differences between Ser and Leu, phosphorylation of eIF2α repressed protein synthesis. This did not happen in the tRNA^{Ser}(A) cell line due to eIF2α dephosphorylation by GADD34, that maintained the levels of protein synthesis rate and allowed a balance between protein degradation and protein synthesis.

2.3.6 Transcriptional deregulation induced by PSE

To obtain a better picture of PQC activation in our model, we have characterized our cell lines using cDNA microarrays. We focused our gene expression data analysis on UPR, UPS, autophagy, translational factors and ribosomal protein genes (Table 2-1). Regarding the UPR genes, *ERN1*(endoplasmic reticulum to nucleus signaling 1, IRE1), *XBP1*(X-box binding protein 1) and *EIF2AK3* (eukaryotic translation factor 2 alpha kinase 3, PERK) were deregulated in our cells. In P1, *ERN1* was downregulated 2.4 fold in tRNA^{Ser}(S), 2.0 fold in tRNA^{Ser}(A) and 1.8 fold in tRNA^{Ser}(L) cells. On the contrary, *XBP1* (wich correspondes to the unspliced transcript) was upregulated 1.5 fold in P1 and 1.8 fold in P15 in tRNA^{Ser}(S) cells. *ERN1* catalyzes the splicing of XBP1 mRNA. Probably, since *ERN1* is downregulated in P1 in tRNA^{Ser}(S) cells, there is an accumulation of *XBP1u*. The upregulation of *XBP1* mRNA in P15, in tRNA^{Ser}(S) cells, is consistent with increased fragmented ATF6. ATF6 can induce not only the expression of ERAD genes, but also *XBP1* gene (119). During evolution, *EIF2AK3* expression increased from -1.2 fold in P1 to 1.1 fold in P30, relative to Mock in tRNA^{Ser}(L) and decreased in tRNA^{Ser}(H) cells from 1.3 in P1 to -1.2 in P30. *EIF2AK3* encodes the kinase responsible for eIF2α phosphorylation (PERK) and its upregulation is consistent with increased levels of eIF2α-P, in P30, in the tRNA^{Ser}(L) cell line.

The cDNA microarray data also confirmed activation of the UPS, as alterations in the expression of some ubiquitin ligases (*UBE2Z*, *UBE2I* and *SMURF2*) were observed over time. Overall, there was a tendency to increased expression in ubiquitin ligases in P15 in tRNA^{Ser}(S), tRNA^{Ser}(L) and tRNA^{Ser}(H) cells, despite the values were below 1.5 fold. In tRNA^{Ser}(S), *UBE2Z* expression decreased from P15 (1.3 fold) to P30 (-1.3 fold), while *SMURF2* expression increased from P1 (-1.2 fold) to P15 (1.3 fold). For the last one, in tRNA^{Ser}(L), there was also an increase in expression from P1 (0.8 fold) to P15 (1.1 fold). In tRNA^{Ser}(H), the expression of *UBE2I* decreased from P15 (1.1 fold) to P30 (-1.1 fold). Deubiquitinating enzymes (*USP48* and *USP36*) were deregulated in tRNA^{Ser}(S), tRNA^{Ser}(A) and tRNA^{Ser}(L) cell lines. *USP48* was upregulated in tRNA^{Ser}(L) in P30 (1.6 fold). *USP36* was upregulated in tRNA^{Ser}(S) in P15 (1.5 fold) and its expression decreased in tRNA^{Ser}(A) cells from P1 (1.2 fold) to P15 (-1.0 fold). Also, the expression of *PSMC1* gene (proteasome 26S subunit, ATPase 1) suffered alterations during evolution, its expression decreased from P15 to P30 in tRNA^{Ser}(S) (1.3 to -1.0 fold), tRNA^{Ser}(A) (1.1 to -1.1 fold) and tRNA^{Ser}(H) cells (1.2 to -1.1 fold). tRNA^{Ser}(S) cell line presented more alterations in UPS genes comparatively to the other cell lines. We could also observe that UPS related genes tended to be upregulated in P15,

concomitantly with the accumulation of ubiquitinated proteins in most cell lines, with exception for *UPS48* in $tRNA^{Ser}$(S) (upregulated in P30, 1.6 fold).

Regarding autophagy, the microarray data showed upregulation of *ATG16L1* during evolution in $tRNA^{Ser}$(S) (from -1.0 fold in P15 to 1.3 fold in P30), $tRNA^{Ser}$(A) (-1.1 fold in P15 to 2.2 fold in P30) and $tRNA^{Ser}$(H) cells (-2.2 fold in P1 to 1.6 fold in P30). *ATG12* and *ATG5* expression also increased in the $tRNA^{Ser}$(L) cell line, from -1.0 fold in P1 to 1.3 fold in P15 and -1.2 fold in P1 to 1.1 fold in P15, respectively. These two genes encode proteins that form a complex involved in the formation of the autophagosomes and may be involved in the degradation of protein aggregates (191). These results indicate that autophagy was probably activated in our cell lines. The increase in *ATG12* and *ATG5* is consistent with the higher levels of insoluble proteins in the $tRNA^{Ser}$(L) cell line.

Beyond the above mentioned deregulations, a set of translational factors, namely *EIF4EBP1* (eukaryotic translation initiation factor 4E binding protein 1), *EEF1A1* (eukaryotic translation elongation factor 1 alpha 1) and *EEF2* (eukaryotic translation elongation factor 2), were also deregulated during evolution. *EIF4EBP1* expression decreased in $tRNA^{Ser}$(A) celss from P15 (1.3 fold) to P30 (-1.1 fold). The eIF4EBP1 interacts with eIF4E and inhibits initiation complex assembly with consequent translation repression (192). The decreased expression of *EIF4EBP1* is in accordance with the translation derepression observed in this cell line. *EEF1A1* expression decreased in $tRNA^{Ser}$(S) and $tRNA^{Ser}$(H) cells from 1.4 fold in P15 to -1.0 fold in P30 and from 1.0 fold in P1 to -1.2 fold in P15, respectively. Also, in $tRNA^{Ser}$(S) cell line, *EEF2* expression was downregulated from 1.0 fold in P15 to -1.5 fold in P30. These data suggest that translation elongation rate may also been affected during evolution, but our data was not able to reveal it. Also, these results may indicate that translation was remodeled for efficient translation of stress genes rather than for global translational repression. Indeed, the *RPS6KL1* (ribosomal protein S6 kinase like 1) that encodes a member of the ribosomal S6 kinase family was upregulated during evolution from -1.3 fold in P15 to 1.0 fold in P30 in $tRNA^{Ser}$(A) cell line. S6 kinases are part of the mammalian target of rapamycin (mTOR) pathway, which is a key regulator of cell growth via the regulation of protein synthesis. S6 kinases are activated by serine/threonine phosphorylation and phosphorylate ribosomal protein 6, increasing translation of a set of proteins, including ribosomal proteins (193).

Altogether, the expression microarrays data are in accordance with our previous results and supported the idea that PQC mechanisms are being activated in a time dependent manner, and in response to mutant tRNAs or tRNA pool deregulation alone. UPR related genes were

upregulated (tRNA^{Ser}(S) cell line), or showed increased fold change during evolution (tRNA^{Ser}(L) cell line). UPS related genes were also upregulated in the same cell lines. Variation of expression through time in UPS genes, was observed in tRNA^{Ser}(S), tRNA^{Ser}(A) and tRNA^{Ser}(H) cells. These gene expression alterations are different for each cell line, probably depending on the need for ubiquitinating or deubiquitinating enzymes during each passage. Autophagy seems to play a role in the response against misfolded proteins in our cell lines, since there was an increase in the expression of autophagy related genes through their evolution from P1 to P30. Alterations in translational factors were also seen, as well as a tendency to decreased expression though time in tRNA^{Ser}(S) and tRNA^{Ser}(A) cells, with no implications in the overall protein synthesis rate. Finally, also the expression of a ribosomal protein kinase (*RPS6KL1*) was altered in tRNA^{Ser}(A) and tRNA^{Ser}(H) cell lines.

Table 2-1. Deregulation of PQC genes induced by protein synthesis errors.
Red color represents upregulated genes, fold change above 1.5. Green color represents dowregulated genes, variation below -1.5 fold. Numbers in blue highlight changes in gene expression that occurred during evolution.

		tRNA^{Ser}(S)			tRNA^{Ser}(A)			tRNA^{Ser}(L)			tRNA^{Ser}(H)		
	Time Point / Gene symbol	P1	P15	P30	P1	P15	P30	P1	P15	P30	P1	P15	P30
UPR related genes	*ERN1*	-2.4			-2.0			-1.8					
	XBP1	1.5	1.8										
	EIF2AK3							-1.2	-1.1	1.1	1.3	-1.1	-1.2
UPS related genes	*UBE2Z*		1.3	-1.3									
	UBE2I											1.1	-1.1
	SMURF2	-1.2	1.3					0.8	1.1				
	USP48									1.6			
	USP36		1.5		1.2	-1.0							
	PSMC1		1.3	-1.0		1.1	-1.1					1.2	-1.1
Autophagy related genes	*ATG16L1*		-1.0	1.3		-1.1	2.2				-2.2	1.4	1.6
	ATG12							-1.0	1.3				
	ATG5							-1.2	1.1				
Translational factors	*EIF4EBP1*					1.3	-1.1						
	EEF1A1		1.4	-1.0							1.0	-1.2	
	EEF2		1.0	-1.5		1.0	-1.3						
Ribosomal proteins	*RPS6KL1*					-1.3	1.0					1.1	-1.1

2.4 Discussion

Our study demonstrates that HEK293 cells expressing mutant tRNAs or overexpressing the Wt $tRNA^{Ser}$ are highly tolerant to PSE, despite the accumulation of misfolded proteins over time. PQC mechanisms were activated in a time and stress dependent manner, allowing these cell lines to thrive, after several generations in culture.

PSE destabilize protein structure and cause disease by overloading chaperones and the proteasome and inducing autophagy, increasing protein cleaning up energetic costs, altering cell signaling and metabolism, producing toxic protein aggregates, repressing protein synthesis and inducing major genomic alterations (75,167,194,195). Kalapis and Bezerra have demonstrated recently that a yeast strain misincorporating Ser at Leu sites could adapt to PSE by up-regulating protein synthesis, protein degradation and glucose up-take (196). Remarkably, clones that evolved for approximately 250 generations were able to reduce protein aggregates and recovered fitness to almost wild type levels, but at a high metabolic cost (196). Our data are in line with those data. Indeed, protein synthesis and degradation rates increased during evolution in $tRNA^{Ser}$(A) cell line. In contrast, while mistranslation had a major negative initial impact on yeast growth rate and viability, no consequence at all were seen in HEK293 cells doubling time and an increase was even observed in the viability of $tRNA^{Ser}$(A) and $tRNA^{Ser}$(L) cell lines and proliferation of $tRNA^{Ser}$(S) and $tRNA^{Ser}$(A) cell lines (Figure 2-5, Figure 2-6 and Figure 2-7).

Also, tolerance to mutant tRNAs increased in yeast during evolution, while in human cells the mutant tRNAs expression was strongly repressed throughout evolution. These data suggest that yeast adapted to mistranslation using error mitigation mechanisms, while human cells preferred error prevention.

The decrease in protein aggregation levels observed during evolution in yeast and HEK293 cells (namely in $tRNA^{Ser}$(L) cell line) has implications for understanding the biology of protein misfolding diseases. Protein aggregation studies use cell models where expression of aggregation prone proteins is induced and adaptation is not evaluated (197–199). Even in cases where these proteins are expressed constitutively the norm is to maintain cell passages as low as possible to avoid genomic instability (200). Considering our data on adaptation during evolution, human cell models of Alzheimer's, Parkinson's and other protein misfolding diseases should be studied to capture the full spectrum of metabolic and physiological changes induced by protein aggregation. Aggregates associated with neurological disorders are known to block proteasome activity and activate mechanisms that lead to the repression of protein

synthesis (41,201), suggesting that these cells are unable to tolerate and adapt to them (202,203). However, different types of human cells cope differently with protein aggregation, raising the question of whether adaptation to aggregation may follow different routes in different cell types. Recent studies showing that suppression of eIF2α kinases alleviates Alzheimer's symptoms in mice (129) support this hypothesis.

Different amino acid misincorporations in the proteome activated different cellular responses and led to differential adaption routes during the evolution of our cell lines. Interestingly, increased expression of the Wt $tRNA_{AGA}^{Ser}$ had phenotypic consequences that were clearly distinct from the Mock (control), suggesting that human cells are highly sensitive to tRNA gene copy number alterations and expression levels. It is known that increased expression of Wt tRNAs (which is common in cancer) alters translation rate, enhances expression of oncogenes and may also increase the level of protein errors, leading to accumulation of misfolded proteins (158,204). In P15, we observed accumulation of ubiquitinated proteins, UPS and UPR activation (Figure 2-21) confirming the hypothesis that misexpression of Wt tRNAs may have major physiological consequences due to activation of the stress response and translational deregulation of gene expression (159).

In the $tRNA^{Ser}$(A) cell line there was an initial (P1 to P15) increase in protein synthesis rate, followed by increased proteasome activity (P15 to P30), in other words increased protein synthesis and turnover, as occurred in yeast (Figure 2-21) (196). Previous studies show that short-living proteins have on average higher aggregation propensity and fewer chaperone interactions than long-living proteins and high protein turnover seems to be sufficient to prevent aggregation (205). Therefore, the Ser-to-Ala misincorporation model may be relevant to address the biology of PSE for instance in cancer, where PQC mechanisms are highly activated (206), but generalized protein aggregation is not commonly observed. Indeed, PSE have not yet been quantified in a systematic manner in cancer, and it is unclear whether it plays a relevant role, however chaperones are often uperegulated in tumors (207), suggesting high demand for protein folding/refolding, which is a hallmark of PSE.

The $tRNA^{Ser}$(L) cell line had higher proteasome activity (P1), but accumulation of aggregated proteins was also visible (P15) (Figure 2-21). Adaptation of this cell line to mutant tRNA expression and protein aggregation was more dependent on UPR activation, namely phosphorylation of eIF2α, and consequent inhibition of protein synthesis rate in order to alleviate ER stress, than other cell lines. At the protein synthesis level, phosphorylation of eIF2α promotes polysome disassembly resulting in the accumulation of untranslated messenger ribonucleoprotein particles (mRPs) that can form stress granules, and are responsible for

reprogramming mRNA metabolism and contribute to cell survival (208,209). Together with the decrease in protein synthesis rate, also dilution by cell division and autophagy may have contributed to the decrease in protein aggregates observed in P30, in this cell line.

tRNA^{Ser}(H) activated mainly PQC mechanisms in P1 and P15, while in P30, a decrease in molecular chaperones occurred (Figure 2-21). With exception of Hsp90α and BiP, there was a general decrease in molecular chaperones expression in our cell lines (Figure 2-21). Decrease in chaperones with aging have been already reported in several studies (210–213). Molecular chaperones are, usually, the first line of defense against misfolded proteins, and probably, other mechanisms can be more efficient to remove these proteins after prolonged proteotoxic stress.

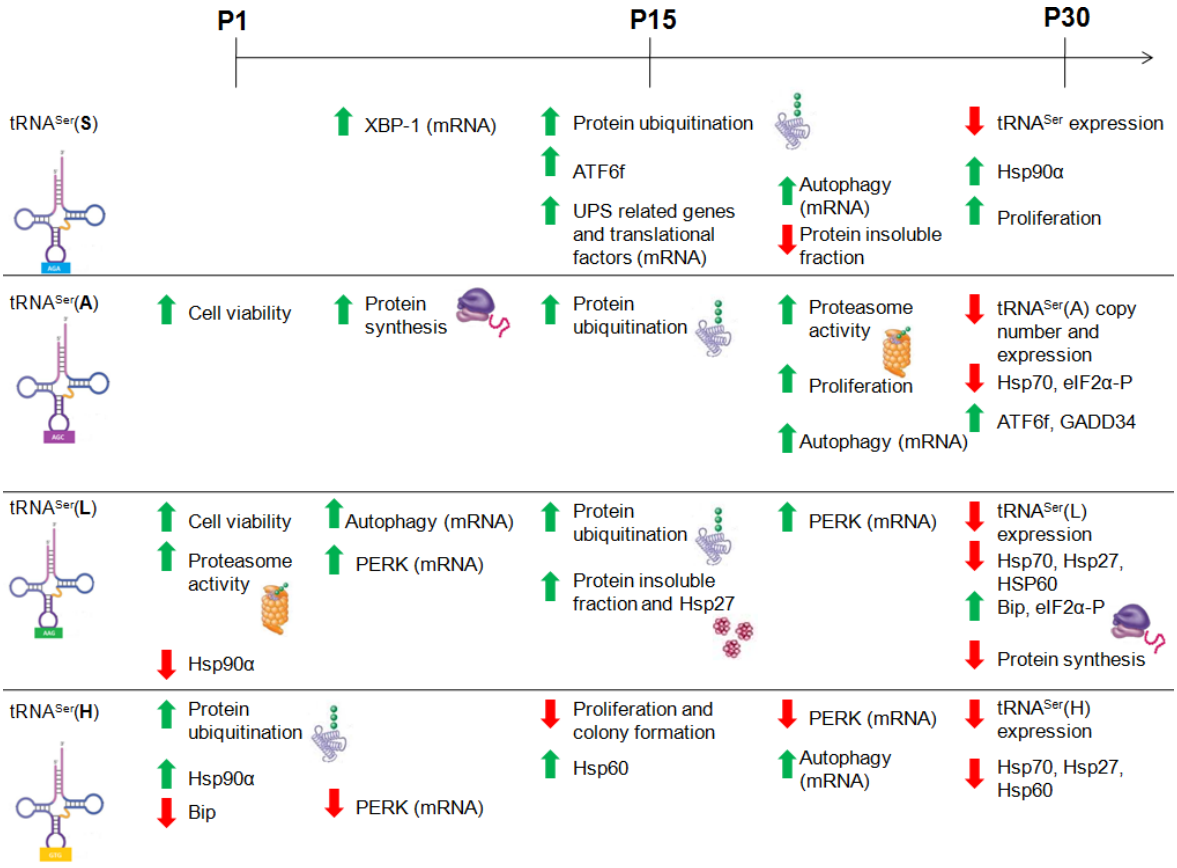


Figure 2-21. Summary of the PQC alterations identified in the different cell lines. Mutant tRNAs and misexpression of tRNA^{Ser} led to accumulation of ubiquitinated proteins, suggesting increased levels of misfolded proteins. PQC mechanisms were recruited in an error type and time dependent manner to counteract proteotoxic stress. Increase in protein turnover and decrease in protein synthesis seem to be two important mechanisms that cells used to thrive in culture after several generations accumulating errors in proteins.

2.5 Conclusion

PSE have been extensively studied in *E. coli* and yeast, but little is known about their biology in human cells. We show here that experimental evolution provides an important tool to study such errors in human cells, although they have only been applied to microorganisms.

HEK293 cells activated PQC mechanisms in order to respond to the accumulation of misfolded proteins caused by the introduction of mutant tRNAs or tRNA pool deregulation. As we expected, this activation depends on the cell passage and the type of proteome destabilization that we are creating. With evolution of the cell lines, we observed that there was a decrease in protein ubiquitination and in some cases protein aggregates with concomitant increase in UPR activation or protein degradation.

Clearly, and in contrast to models of protein misfolding diseases, we did not observe significant effects on cell viability or proteasome inhibition suggesting that human cells cope better with PSE than with protein aggregation associated with Alzheimer's, Parkinson's and other neuropathies. Adaptation to PSE and protein aggregation may also suggest that protein synthesis and degradation rates are more relevant to mitigate and erase aggregates than chaperones or autophagy. In fact, expression of molecular chaperones was unchanged or decreased during evolution, an effect also observed in yeast, where chaperones expression increases initially but decreases gradually or is even repressed in some cases.

2.6 Materials and Methods

2.6.1 Cell culture

Human embryonic kidney 293 (HEK293) cells were purchased from American Type Culture Collection (ATCC®CRL-1573). Cells were grown in Minimum Essential Medium (Gibco, Cat.41090-028) supplemented with 10% fetal bovine serum (FBS) (Sigma, Cat.F1051), 1% of Pen/Strep (Gibco, Cat.15070-063) and 1% of non-essential amino acids (Gibco, Cat.11140-050) in a humidified atmosphere at 37ºC in the presence of 5% CO_2.

2.6.2 Construction of mutant tRNA plasmids

A DNA fragment of 248kb, corresponding to part of the gene encoding human wild type $tRNA_{AGA}^{Ser}$ (Chr6 tRNA#5) and its flanking region, was amplified by PCR. The primers

used were the following: forward 5'-GCCGAATTCAGCTATTATTAAATCCCTAATAAAAGG-3' and reverse 5'-CGCCTCGAGTTGTAAAAGATCGAAAGCCTTTTTA-3'. The region amplified has the following sequence: 5'-AGCTATTATTAAATCCCTAATAAAAGGAGTACAATATGTGATGTATGGAAACATG TAAGACATTTAATAAGGTTTTTGGTATCTGTAGTCGTGGCCGAGTGGTTAAGGCGA TGGACTAGAAATCCATTGGGGTCTCCCCGCGCAGGTTCGAATCCTGCCGACTACGG CAGTGGGTTTTTGCATCTTCAAGCAGGTTTCATCCGACCGATCGATATTTCACGTG TAAAAAGGCTTTCGATCTTTTACAA-3'. This region was cloned into the modified vector pIRES2-DsRed with new multiple cloning sites, using the enzymes EcoRI (Thermo Scientific, Cat.ER0271), XhoI (Thermo Scientific, Cat.ER0691) and T4 DNA ligase (Thermo Scientific, Cat.EL0011). To change the anticodon of the $tRNA_{AGA}^{Ser}$, to other anticodons, we performed site-directed mutagenesis. The primers used were the following: forward to $tRNA_{AGC}^{Ser}$ (A) 5'-GTTAAGGCGATGGACTAGCAATCCATTGGGGTCTCCC-3'; reverse to $tRNA_{AGC}^{Ser}$ (A) 5'-GGGAGACCCCAATGGATTGCTAGTCCATCGCCTTAAC-3';forward to $tRNA_{AAG}^{Ser}$ (L) 5'-GGTTAAGGCGATGGACTAAGAATCCATTGGGGTCTCCC-3'; reverse to $tRNA_{AAG}^{Ser}$ (L) 5'-GGGAGACCCCAATGGATTCTTAGTCCATCGCCTTAACC-3';forward to $tRNA_{GTG}^{Ser}$ (H) 5'-GGTTAAGGCGATGGACTGTGATCCATTGGGGTCTCC-3'; reverse to $tRNA_{GTG}^{Ser}$ (H) 5'-GGAGACCCCAATGGATTCACAGTCCATCGCCTTAACC-3'.

Contruction of mutant tRNA plasmids was done by Patrícia Pereira in the RNA Biology Laboratory, Aveiro, Portugal, in 2007.

2.6.3 Generation of mistranslating cell lines

HEK293 cells with approximately 60% of confluency were transfected with 1µg of plasmid DNA using Lipofectamine2000 (Invitrogen, Cat.11668019), following manufacturer's instructions. Cells were transfected with an empty vector (Mock), and with the plasmid carrying $tRNA_{AGA}^{Ser}$(S), $tRNA_{AGC}^{Ser}$(A), $tRNA_{AAG}^{Ser}$(L) or $tRNA_{GTG}^{Ser}$(H) genes. To establish stable cell lines, 72h after transfection, geneticin (Formedium, Cat.G4185) was added to the medium at a concentration of 800µg/ml and selection lasted for 1 month. Cells were kept in low concentration of geneticin (100µg/ml) after selection and during evolution in culture. Geneticin was not added to the medium when cells were plated for the experiments.

2.6.4 Evolution of cells in culture

After transfection with the plasmids and selection, cells were kept in culture dishes (60mm) and subcultured ever 3 days using the same dilution (1/6) until passage 30. Cell culture conditions were the same during evolution. In P1, P15 and P30 cells were plated in 100mm culture dishes, to have enough cells to perform the experiments and extract DNA, RNA and protein.

2.6.5 Total RNA extraction

RNA was extracted using Trizol®Reagent (Thermo Fisher Scientific, Cat.15596026). The content of one well from a 6 well plate, with around $5x10^5$ cells, was collected for each experimental condition. Purification of RNA was done using DNaseI, Amplification Grade kit (Invitrogen, Cat.18068015), following manufacturer's instructions. RNA was then precipitated with a standard Phenol/Chlorophorm/Isoamylalchohol (25:24:1) (Acros Organics, Cat.327111000) extraction protocol and conserved at -80ºC. RNA concentration was determined using NanoDrop1000 (Thermo Scientific). RNA quality was verified using Agilent 2100 Bioanalyser.

2.6.6 Quantification of tRNA expression and tDNA copy number

The expression of the tRNAs was quantified by extracting total RNA from the transfected cell lines. 200ng of total RNA were used for cDNA convertion using NCode™ VILO™ miRNA cDNA Synthesis Kit (Invitrogen, Cat.A11193050), following manufacturer's instructions. To determine the copy number of the Wt $tDNA^{Ser}$ and the mutant tDNA genes, genomic DNA was extracted using Wizard® Genomic DNA Purification Kit (Promega, Cat.TM050), following the manufacturer's instructions. Amplification of the Wt and mutant tRNAs from cDNA (2µL) or DNA (200ng) was done by PCR using the following primers: forward 5'-CGTAGTCGGCAGGATTCGAA-3' and reverse 5'-GTA GTC GTG GCC GAG TGG TT-3'. As an internal control, GAPDH was also amplified in the same PCR reaction, using the following primers: forward 5´-CTC CTG TTC GAC AGT CAG CC -3' and reverse 5'-CCC ACT TGA TTT TGG AGG GA-3'.

PCR conditions for DNA amplification were: 95ºC for 15min, (95ºC for 30 sec, 62ºC for 1 min and 30 sec, 72ºC for 1 min and 30 sec, 3 cycles); (95ºC for 30 sec, 60ºC for 1 min and 30 sec, 72ºC for 1 min and 30 sec, 3 cycles); (95ºC for 30 sec, 58ºC for 1 min and 30 sec, 72ºC for 1 min and 30 sec, 30 cycles) and a final step of extension at 72°C for 10 min. 8µL of the PCR product was run on a 2% agarose gel to confirm the amplification of the two bands (Figure 2-23 A).

PCR conditions for cDNA amplification were: 95ºC for 15min, (95ºC for 30 sec, 62ºC for 1 min and 30 sec, 72ºC for 1 min and 30 sec, 3 cycles); (95ºC for 30 sec, 60ºC for 1 min and 30 sec, 72ºC for 1 min and 30 sec, 3 cycles); (95ºC for 30 sec, 58ºC for 1 min and 30 sec, 72ºC for 1 min and 30 sec, 30 cycles) and a final step of extension at 72°C for 10 min. 5µL of PCR product were purified with 1µL of ExoI (Thermo Fisher Scientific, Cat.EN0581) and 1µL of FastAp (Thermo Fisher Scientific, Cat.EF0654) for 60min at 37ºC followed by 15min at 85ºC. 2µL of purified PCR product were reamplified. PCR conditions were the same, with the exception of the number of cycles in the last amplification step (cycle 25). 8µL of the PCR product were run on a 2% agarose gel to confirm the amplification of the two DNA fragments (Figure 2-23 B).

5µL of PCR products (from DNA or cDNA) were purified using 1µL of ExoI (Thermo Fisher Scientific, Cat.EN0581) and 1µL of FastAp (Thermo Fisher Scientific, Cat.EF0654) for 60min at 37ºC followed by 15min at 85ºC.

SNaPshot reaction was performed using the following primers: 5'-GGGAGACCCCAATGGATT-3' for tRNAs and 5'-CCC ACT TGA TTT TGG AGG GA-3' for GAPDH and a SNaPshot Multiplex Ready Reaction Mix (Applied Biosystems, Cat.4323163). Reaction cycles for tRNAs were: 96ºC for 10 sec, 54ºC for 5 seconds and 60ºC for 30 sec, 25 cycles. Reaction cycles for GAPDH were: 96ºC for 10 sec, 64ºC for 5 seconds and 60ºC for 30 sec, 10cycles. SNaPshot products were purified with 1µL of FastAp (Thermo Fisher Scientific, Cat.EF0654) for 60min at 37ºC, followed by 15min at 85ºC. Samples were then sequenced and analyzed with Peak Scanner software (Applied Biosystems). The peak area corresponding to each mutant tRNA and Wt $tRNA^{Ser}$ was determined and a ratio calculated. *GAPDH* was quantified and used as a internal control to normalize tRNA expression levels (214,215).

2.6.7 Cell fitness assessment

To measure cells doubling time, $3x10^4$cells/well were plated in 6-well plates. After 72h, cells were detached and counted in a Neubauer chamber with Tripan blue 0.4% (Lonza, Cat.17-942E). Population doubling time was calculated using the formula: Doubling time=duration*log(2)/(log(final concentration)-log(initial concentration)) (216). For viability assays, the number of viable cells in culture was determined with Tripan blue exclusion assay. $3x10^4$cells/well were plated in 6-well plates. After 72h, cells were counted in a Neubauer chamber using Tripan Blue 0.4% (Lonza, Cat.17-942E). For the quantification of cell proliferation, we used a colorimetric immunoassay ELISA, based on the measurement of BrdU incorporation during DNA synthesis (Roche, Cat.11647229001), following manufacturer's instructions. $1x10^5$cells/well were plated in a 96-well and analysis was performed after 48h. To access the ability of a single cell to grow into a colony we performed an anchorage-dependent colony formation assay. 100cells/well were plated in 6-well plates. The medium was renewed every 3 days. After 12 days, cells were fixed in ice-cold methanol and stained with 1% cristal violet (Sigma Aldrich, Cat.C6158) in 20% methanol to count the foci.

2.6.8 Protein synthesis determination

In order to determine protein synthesis rate, we took advantage of a non-reactive fluorescence-activated cell sorting-based assay, called SUnSET (181) with few modifications. $2x10^5$cells/well were plated in 6-well plates and after 48h, puromycin (Sigma Aldrich, Cat. 07635) was added to each well in a final concentration of 10%. Cells were kept in the incubator for 10 min, washed twice with 1%PBS and returned to the incubator for 50 min. After protein extraction with Lysis Buffer (0.5% Triton X-100, 50mM HEPES, 250mM NaCl, 1mM DTT, 1mM NaF, 2mM EDTA, 1mM EGTA, 1mM PMSF, 1mM Na_3VO_4 supplemented with a cocktail of protease inhibitors (Complete, EDTA-free, Roche, Cat. 11873580001); as recommended by the manufacturer) and denaturation, 100 µg of protein were resolved in 10% SDS-PAGE and blotted onto nitrocellulose membranes (0.2µm) (GE Healthcare Life Sciences). Anti-puromycin, clone 12D10 (kindly given by Philippe Pierre) was used (1:5000 dilution) to detect the incorporation of puromycin into proteins. IRDye800 goat anti-mouse secondary antibody (Li-cor Biosciences, Cat.400-33) was used (1:10000 dilution) and detected

in an Odyssey Infrared Imaging System (Licor Biosciences). Membranes were also probed with Anti-β-tubulin (Invitrogen, Cat.32-2600) (1:1000 dilution) as a loading control.

2.6.9 Quantification of the insoluble protein fraction

To quantify the insoluble protein fraction, $2x10^5$cells/well were plated in 6-well plates. After 48h cells were detached and 100µL of Lysis Buffer (0.5% Triton X-100, 50mM HEPES, 250mM NaCl, 1mM DTT, 1mM NaF, 2mM EDTA, 1mM EGTA, 1mM PMSF, 1mM Na_3VO_4 supplemented with a cocktail of protease inhibitors (Complete, EDTA-free, Roche, Cat. 11873580001) as recommended by the manufacturer) was added to the cells' pellet. Cells were sonicated with a probe sonicator in 5 pulses of 5 seconds, incubated on ice for 30min and centrifuged at 5000rpm for 15min at 4ºC. 10µL of the supernatant (total protein fraction) were stored to measure protein concentration with bicinchoninic acid (BCA) assay (Thermo Fisher Scientific, Cat. 23225). 80µL of supernatant were centrifuged again at 12000rpm for 20min at 4ºC. The pellet (insoluble fraction) was then washed with 160µL of LB and 40µL of 10%Triton X-100 (Sigma-Aldrich, Cat.X100) and centrifuged at 15000g for 20min at 4ºC. The pellet was solubilized in 50µL of LB. 15µL of samples were denaturated with loading buffer (6x) (0.375M Tris pH 6.8, 12% SDS, 60% glycerol, 0.6M DTT, 0.06% bromophenol blue) at 95ºC for 5 min and resolved in a 10% SDS-PAGE. The gel was stained with 0.1% Comassie Brilliant Blue G solution (Sigma-Aldrich, Cat.B0770) for at least 2h. After destaining with a solution of 10% ethanol and 7.5% acetic acid, gels were scanned using Odyssey Infrared Imaging System (Licor Biosciences). Lane signals corresponding to each sample were quantified and normalized to the total amount of protein determined with BCA assay.

2.6.10 Quantification of proteasome activity

In order to test the activity of the proteasome, $2x10^5$cells/well were plated in 6-well plates. After 48h cells were washed with 1%PBS and resuspended in 100µL of Lysis Buffer (1mM EDTA, 10mM Tris-HCl pH 7.5; 20% glycerol; 4mM DTT; 2mM ATP). Cells were sonicated with a probe sonicator in 5 pulses of 5 seconds, and centrifuged at 13000rpm for 10min at 4ºC. The supernatant was diluted (1:20) and protein content was quantified using Bradford method (Sigma-Aldrich, Cat.B6916). 20µg of protein in Lysis Buffer were incubated with the substrate suc-LLVY-MCA (50µM) (Sigma-Aldrich, Cat.S4939) in the presence or in

the absence of the proteasome inhibitor MG132 (10µM) (Sigma-Aldrich, Cat.SML1135) in a medium containing 1mM EDTA, 10mM Tris-HCl pH 7.5; 2mM ATP (final volume 100µL). Substrate degradation was monitored every 5min during 1h at 37ºC in a fluorescence-luminescence detector Synergy™ HT Multi-Mode Microplate Reader (Biotek), set to 380 and 460nm, excitation and emission wavelengths, respectively. Specific proteasome activity was determined by subtracting the values for each sample without MG132 to the values with MG132. Final activity was calculated as fluorescence emission at 0 min subtracted from fluorescence after 1h relative to control (Mock).

2.6.11 Immunoblots

$2x10^5$ cells were plated in 6-well plates. After 48h, cells were washed with 1%PBS and then lysed with protein Lysis Buffer (0.5% Triton X-100, 50mM HEPES, 250mM NaCl, 1mM DTT, 1mM NaF, 2mM EDTA, 1mM EGTA, 1mM PMSF, 1mM Na_3VO_4 supplemented with a cocktail of protease inhibitors (Complete, EDTA-free, Roche, Cat. 11873580001) as recommended by the manufacturer). Cells were sonicated with a probe sonicator in 5 pulses of 5 seconds. After centrifugation, 16000g for 30min, protein in the supernatants was quantified using the BCA assay (Thermo Fisher Scientific, Cat. 23225). Samples were denaturated with loading buffer (6x) at 95ºC for 5 min.

Protein samples were separated by SDS-PAGE in 10% polyacrylamide gels (or 8% for molecular chaperones and ATF6), transferred to nitrocellulose membranes (0.2µm) and immunobloted. Membranes were incubated with primary antibodies overnight (4ºC), washed and incubated with secondary antibodies (1:10000 dilution, 2h at room temperature), IRDye800 goat anti-mouse (Li-cor Biosciences, Cat.400-33) or IRDye680 goat anti-rabbit (Li-cor Biosciences, Cat.925-68070). Secondary antibodies were detected using an Odyssey Infrared Imaging System (Licor Biosciences). The following primary antibodies were used: anti-ubiquitin (1:1000 dilution) (Sigma-Aldrich, Cat.U0508), anti-Hsp70 (1:1000 dilution) (Stress Marq Biosciences, Cat.SMC-100B), anti-Hsp27 (0.5µg/ml dilution) (Stress Marq Biosciences, Cat.SMC-161A), anti-Hsp60 (1:1000 dilution) (Stress Marq Biosciences, Cat.SPC-105), anti-Hsp90α (1:1000 dilution) (Stress Marq Biosciences, Cat.SMC-147), anti-BiP (1:1000 dilution) (Stress Marq Biosciences, Cat.SPC-180), anti-ATF6 (1:400 dilution) (Stressgen, Cat.70B1413.1), anti-eIF2α (1:1000 dilution) (Cell Signalling, Cat.9722), anti-phosphorylated eIF2α (1:400 dilution) (Abcam, Cat.ab4837), anti-GADD34 (1:500 dilution)

(Thermo Fisher Scientific, Cat.PA1-12376), anti-β-tubulin (Invitrogen, Cat.32-2600). β-tubulin was used in all the immunoblots as a loading control.

2.6.12 Gene expression microarrays

Gene expression microarrays profiling was performed using the Agilent protocol for One-Color Microarray Based Gene Expression Analysis Low Input Quick Amp Labeling v6.9 (Agilent Technologies). RNA quality determination was performed using 2100 Bioanalyser (Agilent Technologies) and the kit Agilent RNA 6000 Nano kit (Agilent Technologies, Cat.5067-1512). 100ng of total RNA were used to synthesize labeled cDNA (with Cyanine 3-CTP) using Agilent T7 Promoter Primer and T7 RNA polymerase Blend (Agilent Technologies, Cat.5190-2305). cDNA was purified with RNAeasy mini spin columns (Quiagen, Cat.74104). Dye incorporation was quantified using Nanodrop 1000 Spectophotometer. 600ng of labeled cDNA were hybridized in Sure Print G3 Human Gene Expression 8x60k v2 microarrays (Agilent Technologies, Cat.G4851B). Hybridizations were carried out using Agilent gasket slides in a rotating oven for 17h at 65ºC. Slides were then washed following manufacturer's instructions and scanned in an Agilent G2565AA microarrays scanner.

Probes signal values were extracted using Agilent Feature Extraction Software. Data were normalized using median centering of signal distribution with Biometric Research Branch BRB-Array tools v3.4.o software (217,218). Microarrays statistical analysis was carried out using Mev software (TM4 Microarray Software Suite) (219,220). T-test was performed to identify genes that showed differences in expression between control (Mock) and samples. Significant genes that present a fold change above 1.5 or bellow -1.5 were considered for downstream analysis.

The microarray raw data was submitted to the GEO database and has been given the following accession number: GSE93854.

2.6.13 Statistical analysis

For all assays, our data represent at least 3 independent experiments and 3 replicates. Statistical analysis was performed using One-way ANOVA analysis of variance followed by

the Dunnett's or Bonferroni's post-tests, as indicated in the figures. In all cases, *p*-values<0.05 were considered statistically significant.

2.7 Supplementary Figures

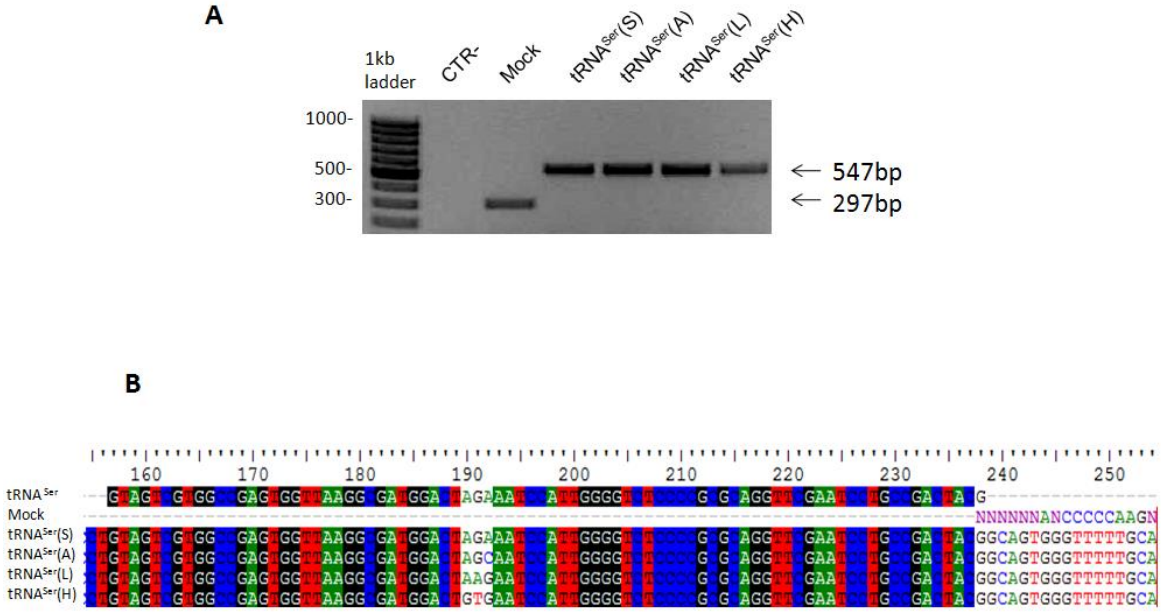


Figure 2-22. tRNA amplification with primers for the plasmid and Sanger sequencing.

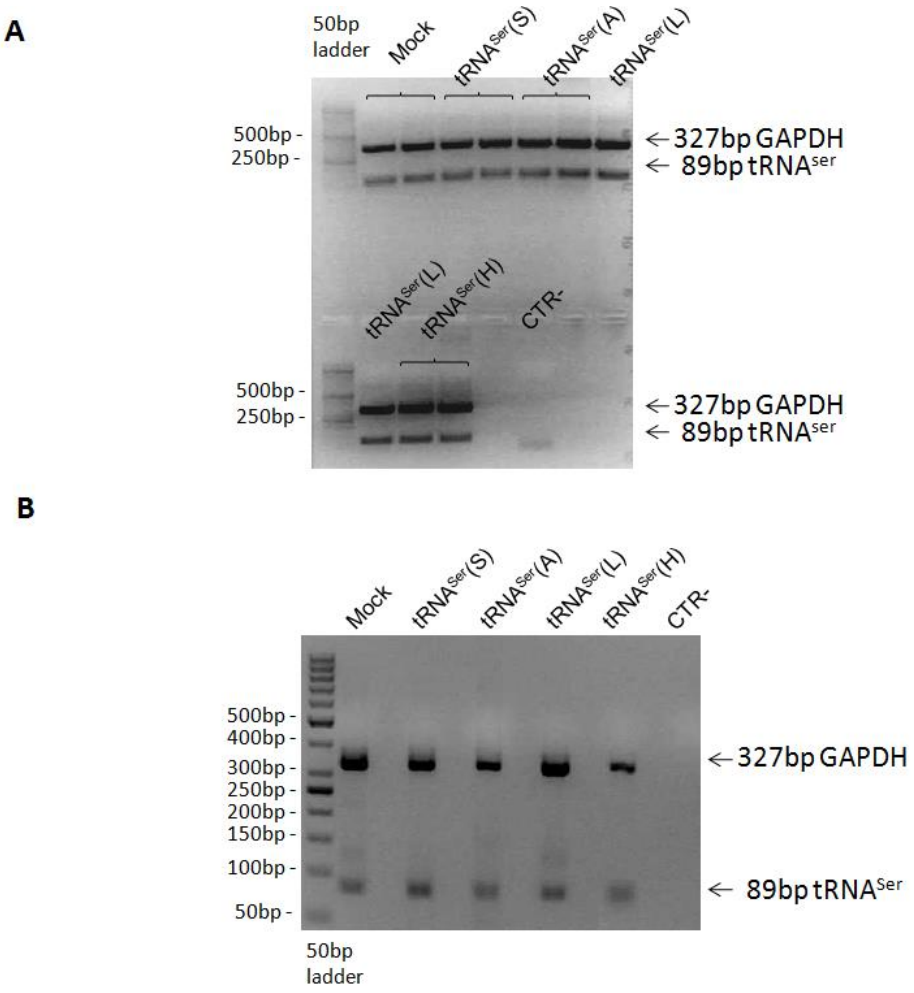


Figure 2-23. SNaPshot analysis – Agarose gels representative of the amplification of tRNA and GAPDH. A – Amplification from DNA of mistranslating cells; **B –** Amplification from cDNA of mistranslating cells.

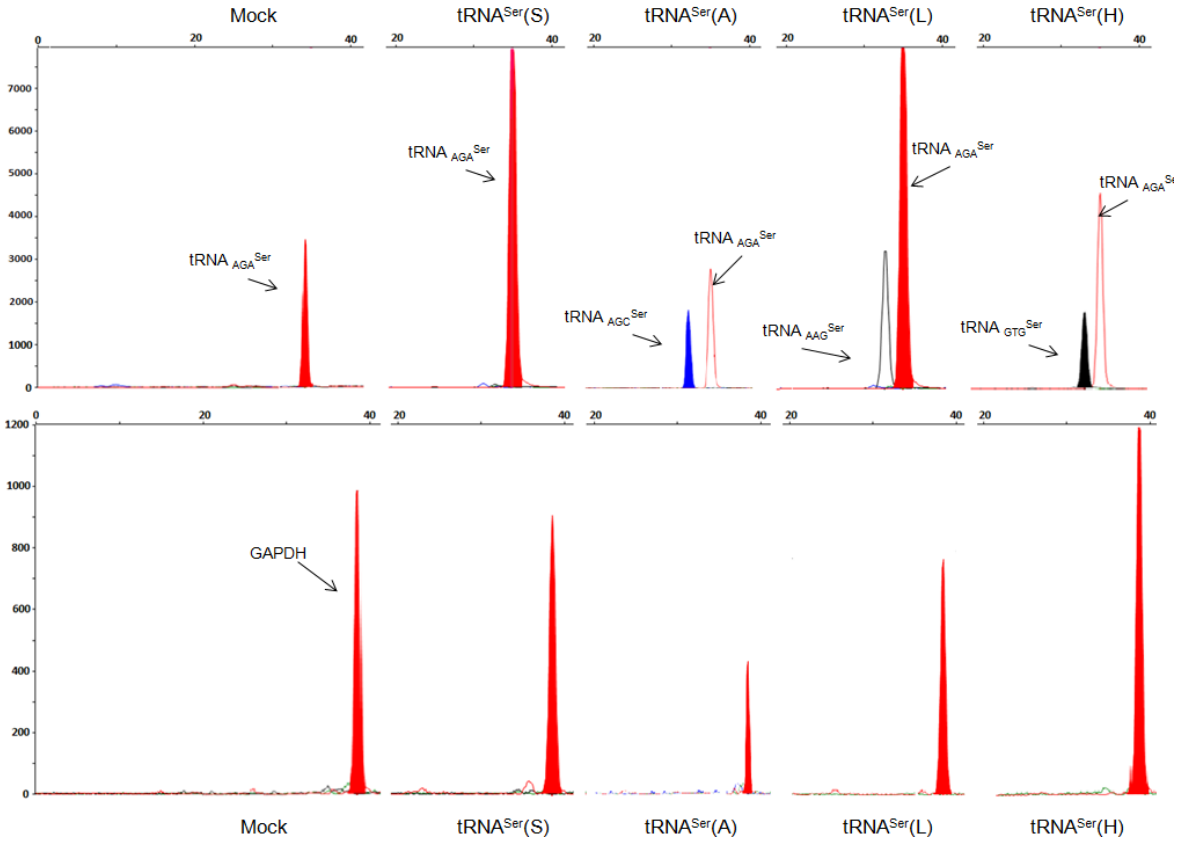


Figure 2-24. SNaPshot peaks of the endogenous tRNA^{Ser}, mutant tRNAs and the respective control, GAPDH. Sequenced samples correspond to cDNA from P1. The primer extension (SNaPshot) reaction utilizes a multiplex kit that contains a reaction mix of four differentially fluorescently labeled ddNTPs, allowing the detection of the incorporated base correspondent to the last base of anticodons. In the case of tRNA^{Ser}(S) cell line, expressing tRNA$_{AGA}^{Ser}$, the peak corresponds to the incorporation of T (red peak). For tRNA^{Ser}(A) cells (expressing tRNA$_{AGC}^{Ser}$), the peak corresponds to the incorporation of G (blue peak). For tRNA^{Ser}(L) (expressing tRNA$_{AAG}^{Ser}$) and tRNA^{Ser}(H) (expressing tRNA$_{GTG}^{Ser}$) cells, the peak corresponds to the incorporation of C (black peak). In tRNA^{Ser}(A), tRNA^{Ser}(L) and tRNA^{Ser}(H) cell lines, the Wt tRNA^{Ser} and mutant tRNAs were detected (upper panel). The peak area corresponding to each mutant tRNA and Wt tRNA^{Ser} was determined for each cell line and a ratio was calculated. Amplification of GAPDH was used as an internal control to normalize the values (lower panel).

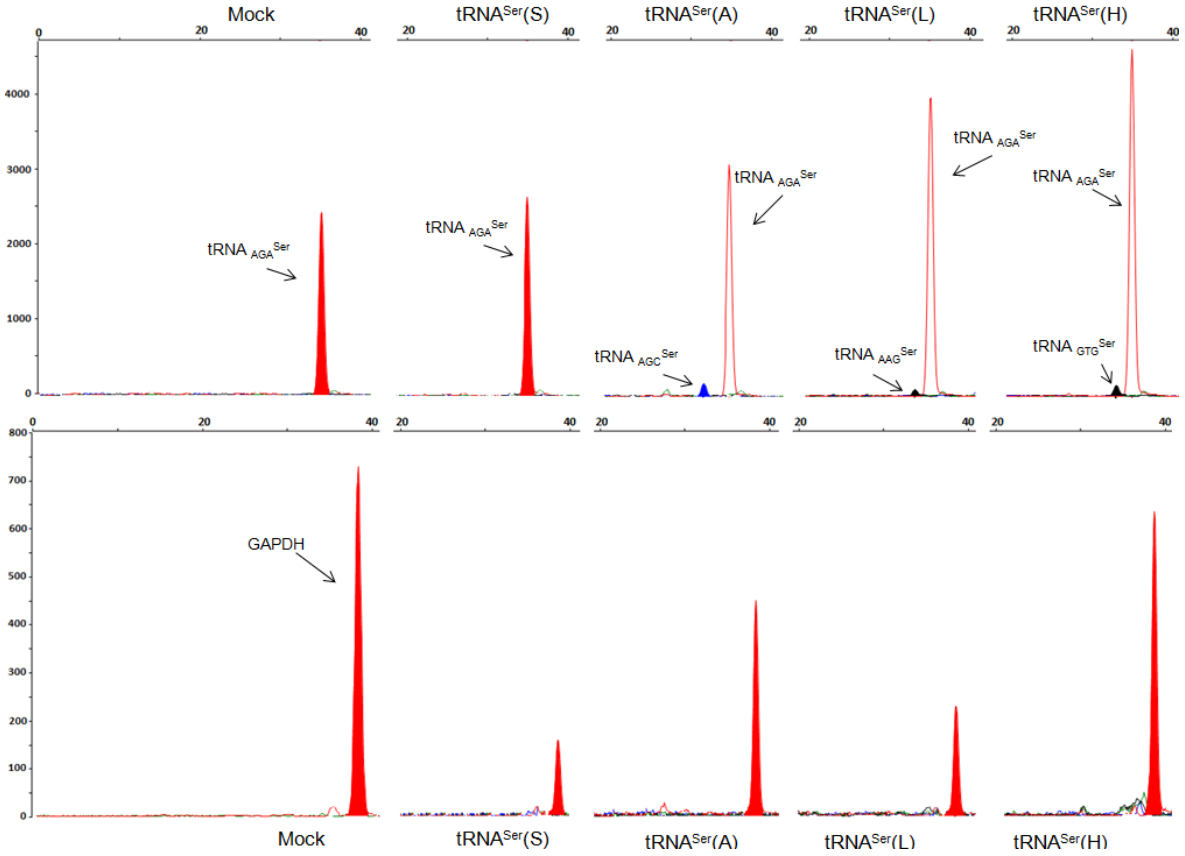


Figure 2-25. SNaPshot peaks of the endogenous tRNA^{Ser}, mutant tRNAs and the respective control, GAPDH. Sequenced samples correspond to cDNA from P15. For tRNA^{Ser}(S) cell line (expressing tRNA$_{AGA}^{Ser}$), the peak corresponds to the incorporation of T (red peak). For tRNA^{Ser}(A) cells (expressing tRNA$_{AGC}^{Ser}$), the peak corresponds to the incorporation of G (blue peak). For tRNA^{Ser}(L) (expressing tRNA$_{AAG}^{Ser}$) and tRNA^{Ser}(H) (expressing tRNA$_{GTG}^{Ser}$) cells, the peak corresponds to the incorporation of C (black peak). In tRNA^{Ser}(A), tRNA^{Ser}(L) and tRNA^{Ser}(H) cell lines, the Wt tRNA^{Ser} and mutant tRNAs were detected (upper panel). The peak area corresponding to each mutant tRNA and Wt tRNA^{Ser} was determined for each cell line and a ratio was calculated. Amplification of GAPDH was used as an internal control to normalize the values (lower panel).

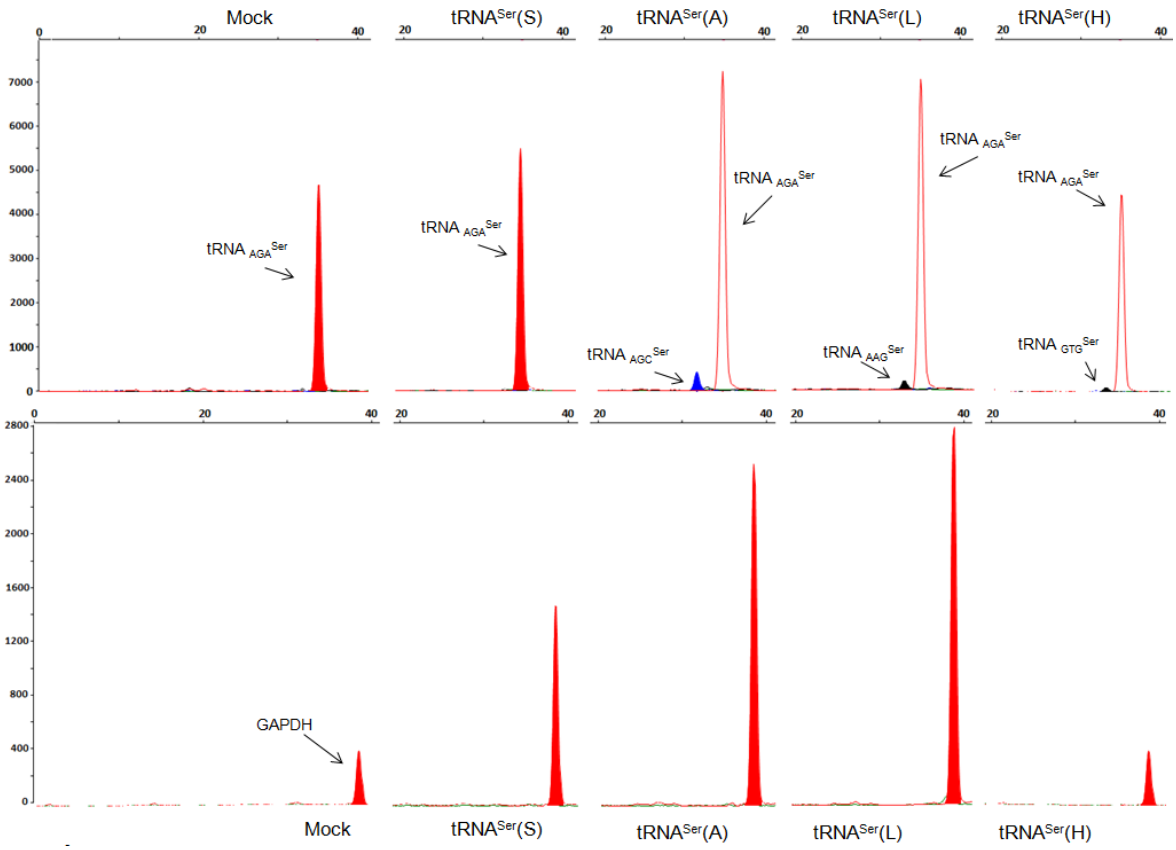


Figure 2-26. SNaPshot peak of the endogenous tRNA^{Ser}, mutant tRNAs and the respective control, GAPDH. Sequenced samples correspond to cDNA from P30. For tRNA^{Ser}(S) cell line (expressing tRNA$_{AGA}^{Ser}$), the peak corresponds to the incorporation of T (red peak). For tRNA^{Ser}(A) cells (expressing tRNA$_{AGC}^{Ser}$), the peak corresponds to the incorporation of G (blue peak). For tRNA^{Ser}(L) (expressing tRNA$_{AAG}^{Ser}$) and tRNA^{Ser}(H) (expressing tRNA$_{GTG}^{Ser}$) cells, the peak corresponds to the incorporation of C (black peak). In tRNA^{Ser}(A), tRNA^{Ser}(L) and tRNA^{Ser}(H) cell lines, the Wt tRNA^{Ser} and mutant tRNAs were detected (upper panel). The peak area corresponding to each mutant tRNA and Wt tRNA^{Ser} was determined for each cell line and a ratio was calculated. Amplification of GAPDH was used as an internal control to normalize the values (lower panel).

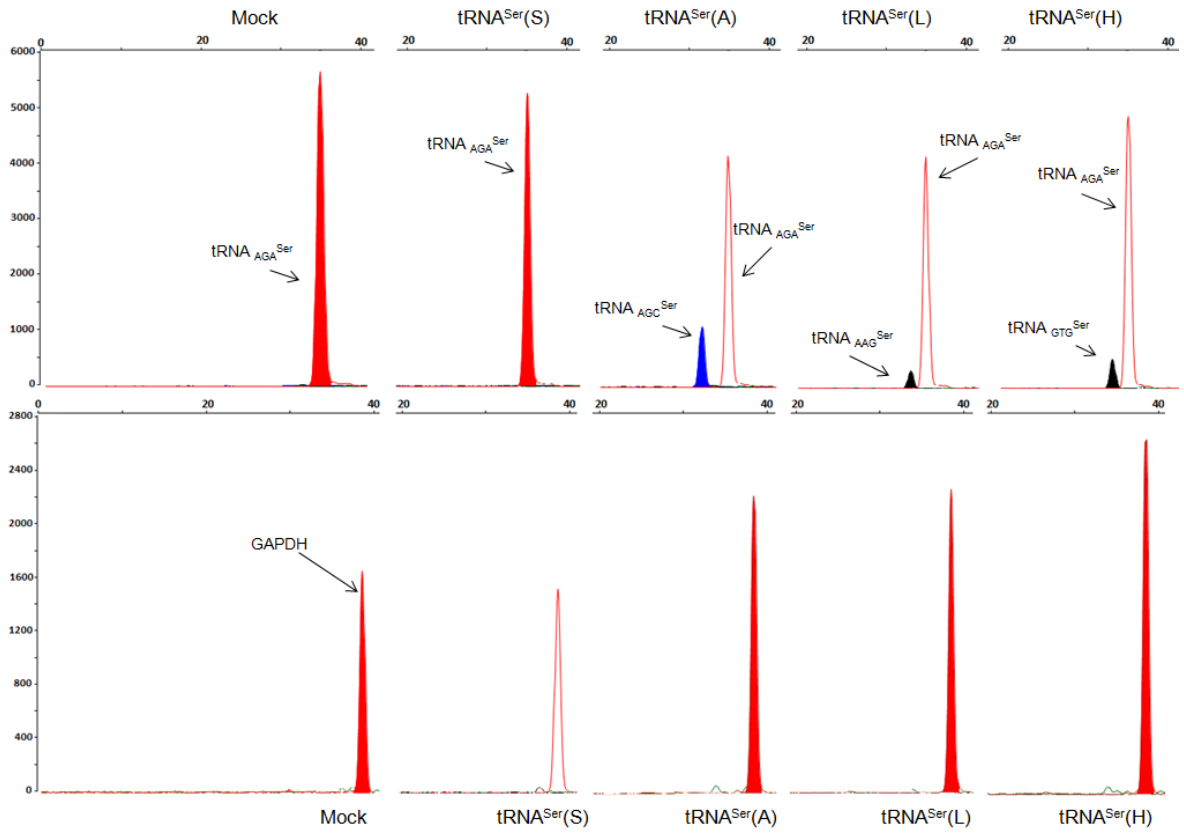


Figure 2-27. SNaPshot peaks of the endogenous tRNA^{Ser}, mutant tRNAs and the respective control, GAPDH. Sequenced samples correspond to DNA from P1. For tRNA^{Ser}(S) cell line (expressing tRNA$_{AGA}^{Ser}$), the peak corresponds to the incorporation of T (red peak). For tRNA^{Ser}(A) cells (expressing tRNA$_{AGC}^{Ser}$), the peak corresponds to the incorporation of G (blue peak). For tRNA^{Ser}(L) (expressing tRNA$_{AAG}^{Ser}$) and tRNA^{Ser}(H) (expressing tRNA$_{GTG}^{Ser}$) cells, the peak corresponds to the incorporation of C (black peak). In tRNA^{Ser}(A), tRNA^{Ser}(L) and tRNA^{Ser}(H) cell lines, the Wt tRNA^{Ser} and mutant tRNAs were detected (upper panel). The peak area corresponding to each mutant tRNA and Wt tRNA^{Ser} was determined for each cell line and a ratio was calculated. Amplification of *GAPDH* was used as an internal control to normalize the values (lower panel).

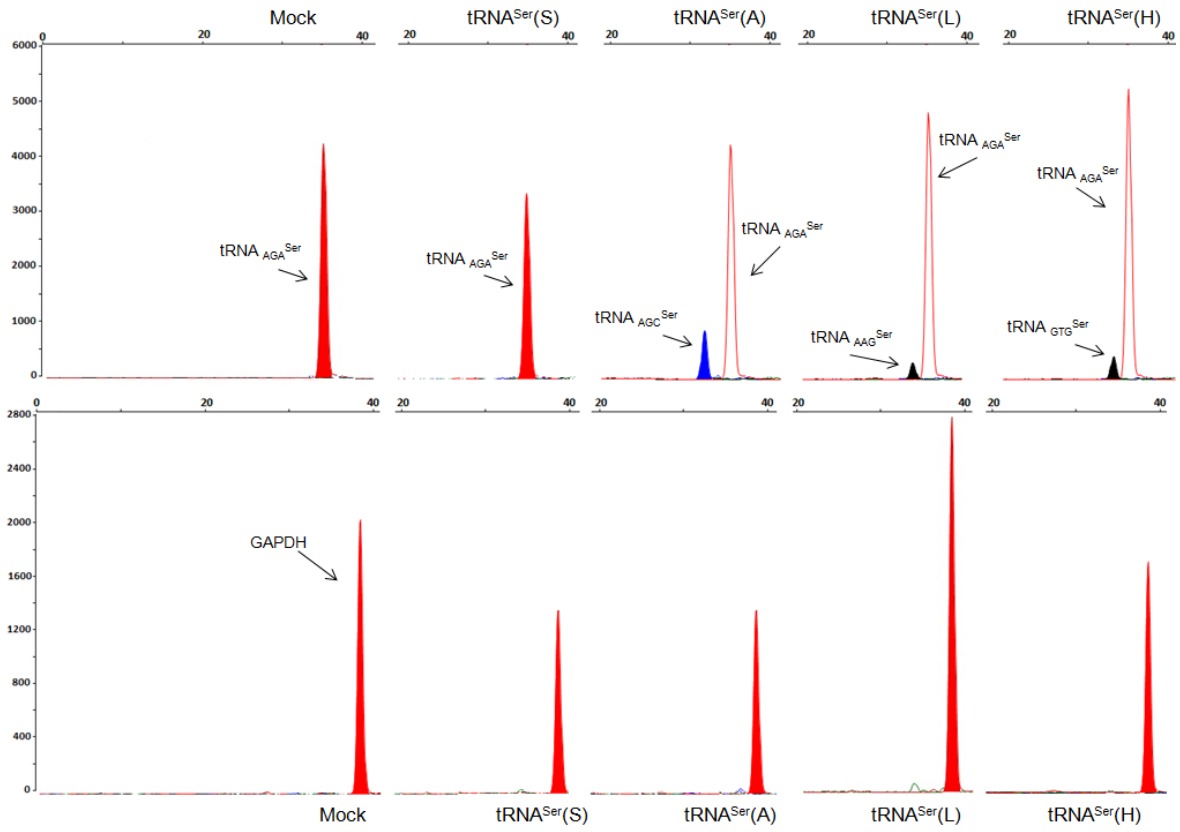


Figure 2-28. SNaPshot peaks of the endogenous tRNA^{Ser}, mutant tRNAs and the respective control, GAPDH. Sequenced samples correspond to DNA from P15. For tRNA^{Ser}(S) cell line (expressing tRNA$_{AGA}^{Ser}$), the peak corresponds to the incorporation of T (red peak). For tRNA^{Ser}(A) cells (expressing tRNA$_{AGC}^{Ser}$), the peak corresponds to the incorporation of G (blue peak). For tRNA^{Ser}(L) (expressing tRNA$_{AAG}^{Ser}$) and tRNA^{Ser}(H) (expressing tRNA$_{GTG}^{Ser}$) cells, the peak corresponds to the incorporation of C (black peak). In tRNA^{Ser}(A), tRNA^{Ser}(L) and tRNA^{Ser}(H) cell lines, the Wt tRNA^{Ser} and mutant tRNAs were detected (upper panel). The peak area corresponding to each mutant tRNA and Wt tRNA^{Ser} was determined for each cell line and a ratio was calculated. Amplification of *GAPDH* was used as an internal control to normalize the values (lower panel).

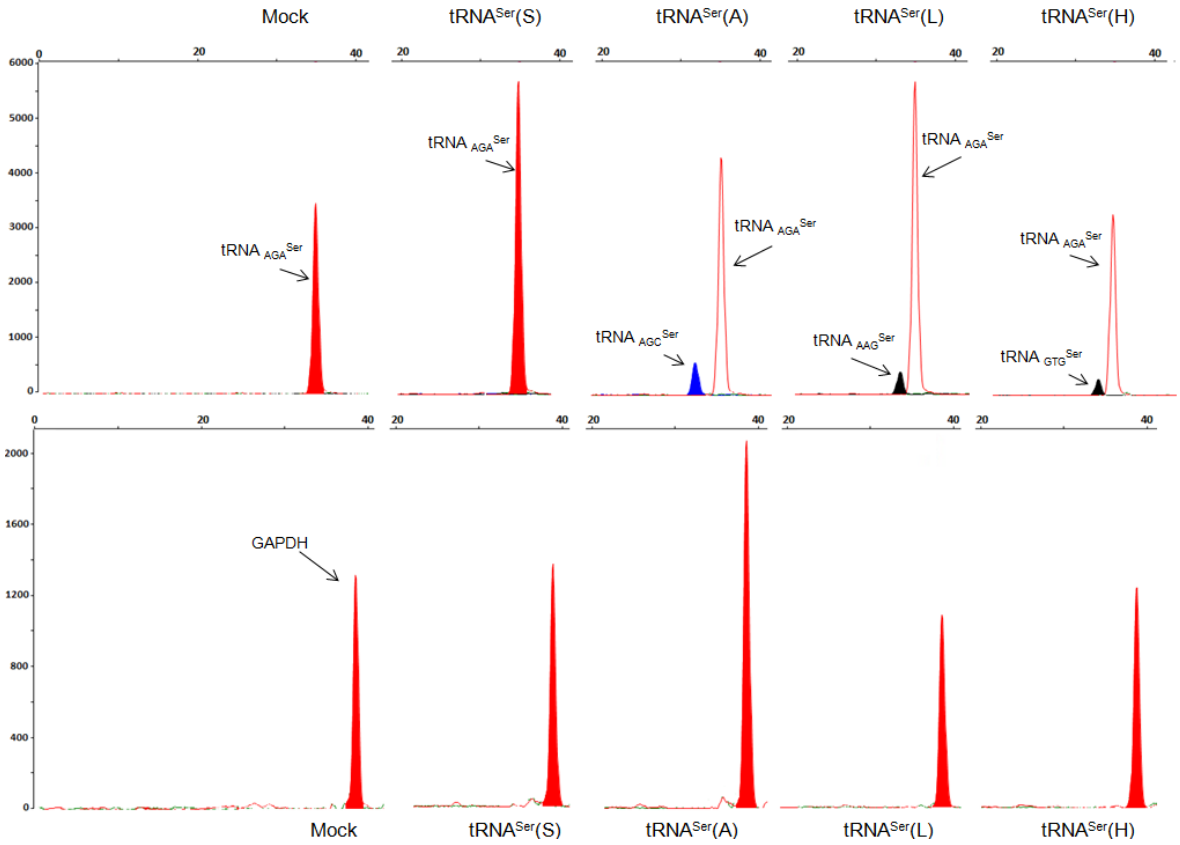


Figure 2-29. SNaPshot peaks of the endogenous tRNA^{Ser}, mutant tRNAs and the respective control, GAPDH. Sequenced samples correspond to DNA from P30. For tRNA^{Ser}(S) cell line (expressing tRNA$_{AGA}^{Ser}$), the peak corresponds to the incorporation of T (red peak). For tRNA^{Ser}(A) cells (expressing tRNA$_{AGC}^{Ser}$), the peak corresponds to the incorporation of G (blue peak). For tRNA^{Ser}(L) (expressing tRNA$_{AAG}^{Ser}$) and tRNA^{Ser}(H) (expressing tRNA$_{GTG}^{Ser}$) cells, the peak corresponds to the incorporation of C (black peak). In tRNA^{Ser}(A), tRNA^{Ser}(L) and tRNA^{Ser}(H) cell lines, the Wt tRNA^{Ser} and mutant tRNAs were detected (upper panel). The peak area corresponding to each mutant tRNA and Wt tRNA^{Ser} was determined for each cell line and a ratio was calculated. Amplification of *GAPDH* was used as an internal control to normalize the values (lower panel).

3. Transcriptional alterations in mRNA mistranslating HEK239 cells

3.1 Abstract

In human cells, protein synthesis errors (PSE) activate several protein quality control (PQC) mechanisms that are highly dynamic and respond to type and duration of stress stimuli (Chapter 2). Transcriptome analysis revealed that PSE also induce mechanisms that protect cells against stress. For example, mistranslating yeast strains upregulate genes involved in protein folding, response to oxidative stress, carbohydrate and energy reserve metabolism (195).

Many studies have characterized the cellular responses to PSE, but a global picture of how human cells respond and adapt in culture to PSE is still missing. To clarify this issue we have used the same cell lines created in the previous Chapter (Mock, $tRNA^{Ser}$(S); $tRNA^{Ser}$(A); $tRNA^{Ser}$(L); $tRNA^{Ser}$(H)) and profiled them using DNA microarrays. Our objective was to identify the transcriptional deregulation induced by PSE of cells in culture at different evolution time points (P1, P15 and P30). We also wanted to identify the transcriptome alterations that occur in each cell line throughout time, with an increase in the number of passages.

As in the previous Chapter, transcriptional alterations were dependent on error type and cells passage number. Several genes that encode endoplasmic reticulum (ER) and ER stress proteins were upregulated. In addition, metabolic alterations (glycolysis/gluconeogenesis) and activation of signaling pathways (p53 and MAPK) were deregulated in our cell lines. These transcriptional alterations provide new insights to understand the pathways involved in the response to PSE.

3.2 Introduction

Eukaryotic cells have evolved sophisticated sensing mechanisms and signal transduction systems that can produce accurate dynamic outcomes in response to stress. This ultimately allows cellular stress protection, homeostasis and survival (221). The stress responses involve gene expression alterations that are essential for the slower long term adaptation and recovery of cells. Control of gene expression is tightly regulated and has fast response kinetics and controlled reversibility, which allow cells to change their transcriptional capacity within minutes in the presence of the stress (221). Thus, the transcriptome (set of all RNA molecules in a cell) is highly dynamic and reflects the homeostatic imbalances of cells (221,222).

Microarray-based transcriptome comparisons have been widely used to identify potentially adaptive expression changes to several stressors (223) or characterize disease states, such as cancer (224,225).

PSE and accumulation of misfolded proteins in cells lead to the activation of PQC mechanisms, already described previouly. In Chapter 2, we have highlighted that increase or decrease in transcription and translation of PQC components fluctuate over time. Transcriptional responses of cells induced by PSE have already been characterized. For example, protein misfolding, caused by misincorporation of the proline analogue azetidine-2-carboxilic acid (AZC) in *S. cerevisiae* leads to induction of heat shock factor-regulated genes and selective repression of ribosomal protein genes (226). In human HEK293FT cells, the misincorporation of AZC leads to the upregulation of specific genes related to the response to unfolded proteins, namely *DDIT3* (a transcription factor activated by ER stress), *DNAJB1* (a molecular chaperone that stimulates ATPase activity of Hsp70 heat shock proteins) and *HSPH1* (a heat shock protein that prevents the aggregation of denatured proteins in cells under severe stress) (227). In yeast, the expression of a Ser tRNA that misincorporates Ser at Leu CUG codons sites, leads to the accumulation of misfolded proteins and upregulation of molecular chaperones, but also oxidative and general stress, carbohydrate and energy reserve metabolism genes. Genes encoding translational factors, ribosome biogenesis and assembly were downregulated (195). The deregulation of carbohydrate and energy metabolism genes shows that adaptation to PSE also prompts alterations in metabolism. These data show that upregulation of translation and metabolic processes occurs as an initial response to mistranslation, but such deregulation changes from positive to negative over time. Molecular chaperones are the first line of defense against misfolded proteins while downregulation of protein synthesis is a downstream adaptive event (195). Activation of different stress genes (so-called environmental stress response) plus those belonging to PQC, also confer stress cross-protection and increased thermotolerance, tolerance to oxidative and osmotic stress and also resistance to heavy metals and drugs, in yeast (73,228).

In zebrafish embryos, the expression of tRNAs that misincorporate Ser at non-cognate protein sites, also leads to upregulation of molecular chaperones (such as *hsp70l*, *hsp90a.2*, and *hspa5/bip*), ER stress response related genes (such as *eif2ak3/perk* and *xbp1*), apoptosis and autophagy related genes. It was also reported downregulation of several mitochondrial genes that confer protection against ROS (167).

Transcriptional tools have been essential to better characterize and understand the different effects (positive and negative) of PSE in cells. Yeast cells exposed for approximately

250 generations to PSE were characterized at the transcriptional level and alterations were identified in the evolved cell lines, relative to the ancestor. Genes belonging to ribosome biogenesis, rRNA processing and ribosomal large subunit assembly were upregulated while genes involved in glycogen biosynthetic processes were downregulated in the evolved cell lines (196).

Despite all these studies, there is little data about transcriptional responses of human cells exposed to PSE. To address this issue, we have used the cell lines described previously, which express mutant tRNAs and overexpress the Wt $tRNA_{AGA}^{Ser}$, and analyzed their transcriptional profile. Transcriptional alterations were identified in each cell line comparatively with the Mock in each passage and a comparison between the alterations in each cell line through time was also carried out. This, allowed us to complement our previous transcriptional analysis, from Chapter 2, which was focused on specific genes belonging to PQC processes.

Although, there were a small number of deregulated genes in these mistranslating cell lines, we could correlate some of the deregulated Gene Ontology (GO) categories with the phenotypes that we have described previously. Transcriptional alterations were dependent on the type of the error introduced into proteins and cells passage. In all of our mistranslating cells there was upregulation of ER and ER stress response genes. In addition, in some cell lines ($tRNA^{Ser}$(S) and $tRNA^{Ser}$(A)) genes involved in metabolic pathways such as glycolysis and gluconeogenesis were deregulated. Signaling pathways, for example p53 and MAPK, were also deregulated. This study provides a comphreensive overview of the transcriptional alterations and permited the identification of putative adaptive pathways.

3.3 Results

3.3.1 The impact of PSE in gene expression deregulation

In the previous Chapter, we have described phenotypic and molecular alterations of HEK293 cells expressing tRNAs that misincorporate Ser at non-cognate codon sites, namely, $tRNA_{AGC}^{Ala}$, $tRNA_{AAG}^{Leu}$ and $tRNA_{GTG}^{His}$, or overexpressing the Wt $tRNA_{AGA}^{Ser}$. These cells could cope relatively well with those tRNAs despite the accumulation of misfolded proteins or accumulation of protein aggregates ($tRNA^{Ser}$(L) cell line). Interstingly, $tRNA^{Ser}$(L) cell line in particular was able to erase protein aggregates after several generations in culture.

Depending on the type of the misincorporated amino acid introduced into proteins and also cells passage number, there was differential activation of PQC mechanisms. The activation of PQC mechanisms was essential for cell survival in culture along with decreased expression of the mutant tRNAs. The phenotypes observed suggest that these cells may activate others stress pathways and adaptation mechanisms. In order to understand and identify those putative mechanisms involved in adaptation to PSE, cDNA microarrays were performed using 3 biological replicates for each of our created cell lines (Mock, $tRNA^{Ser}$(S), $tRNA^{Ser}$(A), $tRNA^{Ser}$(L) and $tRNA^{Ser}$(H)) at passages number, 1, 15 and 30. These studies allowed us to analyze the expression level of 26.083 genes encoded by the human genome. Only significantly deregulated genes were considered in our analysis ($p<0.05$) with a fold change above 1.5 or below -1.5.

For this first analysis, each cell line ($tRNA^{Ser}$(S), $tRNA^{Ser}$(A), $tRNA^{Ser}$(L) and $tRNA^{Ser}$(H)) was compared with the control in the respective passage, thus the fold change was calculated relatively to Mock. Figure 3-1, shows the percentage of genes that were significantly deregulated with a fold change above 1.5 (upregulated genes) or below -1.5 (downregulated genes), in each cell line in P1, P15 and P30.

In P1, the $tRNA^{Ser}$(S) cell line had the highest number of deregulated genes (7.4%), followed by $tRNA^{Ser}$(A) cell line (5.5%). $tRNA^{Ser}$(L) and $tRNA^{Ser}$(H) cells, had 2.1% and 2.9% of deregulated genes, respectively (Figure 3-1 A). All of the cell lines in P1 had a higher number of downregulated genes ($tRNA^{Ser}$(S)-5.1%; $tRNA^{Ser}$(L)-1.8% and $tRNA^{Ser}$(H)-1.8%) comparatively to the number of upregulated genes ($tRNA^{Ser}$(S)-2.3%; $tRNA^{Ser}$(L)-0.4% and $tRNA^{Ser}$(H)-1.1%) with the exception of $tRNA^{Ser}$(A) cell line, that had fewer downregulated (1.9%) than upregulated genes (3.6%) (Figure 3-1 A).

In P15, the cell line with higher number of deregulated genes was also $tRNA^{Ser}$(S) (4.9%) (Figure 3-1 B). The remaining cell lines had similar levels of deregulation (($tRNA^{Ser}$(A)-2.8%; $tRNA^{Ser}$(L)-2.8% and $tRNA^{Ser}$(H)-2.7%). For each cell line in P15, the percentage of upregulated and downregulated genes was similar (Figure 3-1 B).

In P30, the $tRNA^{Ser}$(S) cell line had the lowest percentage of deregulated genes, only 1.3%, followed by $tRNA^{Ser}$(A) with 2.3% (Figure 3-1 C). $tRNA^{Ser}$(L) and $tRNA^{Ser}$(H) cells showed the highest deregulation, 5.9% and 5%, respectively. In $tRNA^{Ser}$(S) and $tRNA^{Ser}$(L) cells there were more genes downregulated (0.8% and 4.6%, respectively) than upregulated (0.4% and 1.3%, respectively), while in $tRNA^{Ser}$(A) cell line the levels of upregulated and downregulated genes was similar. $tRNA^{Ser}$(H) cell line showed fewer downregulated (1.2%) than upregulated (3.8%) genes (Figure 1-3 C).

In all the three time points (P1, P15 and P30), the percentage of deregulated genes was low (always below 8%), showing that only a small set of genes is being deregulated in response to PSE in our cells.

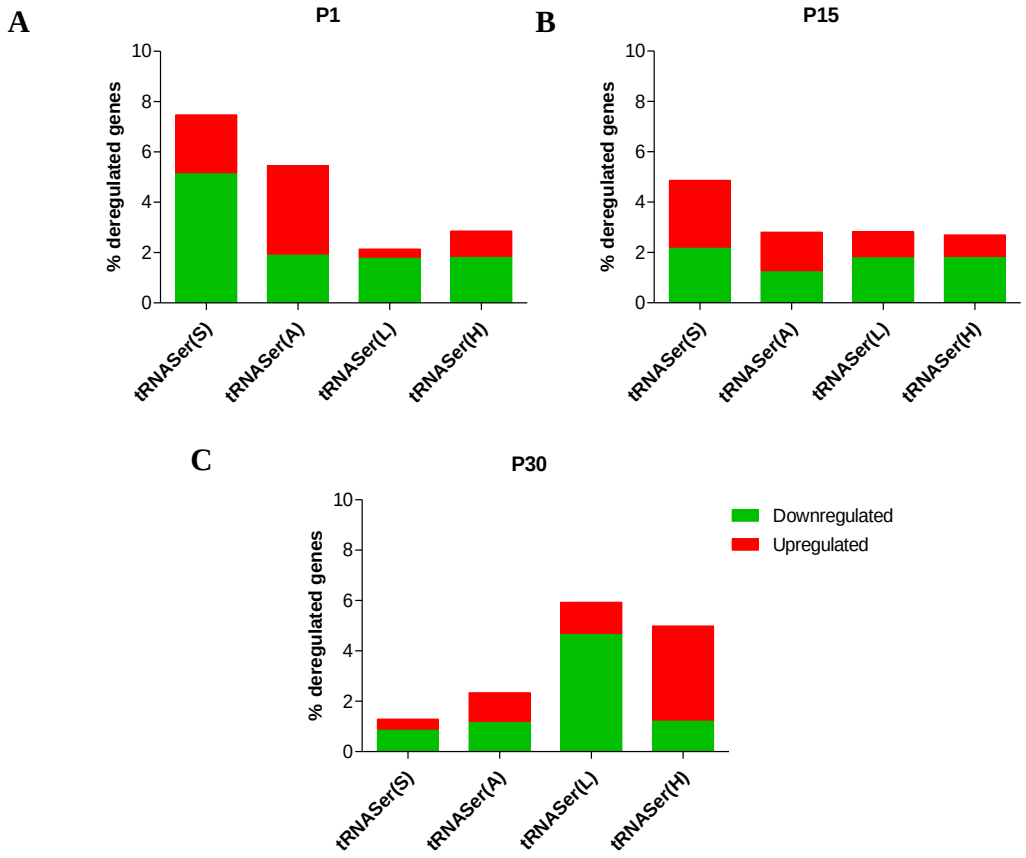


Figure 3-1. Percentage of deregulated genes for each cell line comparatively with the Mock in each time point. **A** – P1; **B** – P15; **C** – P30

3.3.2 Ontology of the genes deregulated in response to PSE

In order to understand in more detail the relevance of the above deregulation for adaptation to PSE, we annotated deregulated genes in GO Categories using DAVID Bioinformatics Resources 6.7. Even though the percentage of deregulation is low in these cells, many GO categories were significantly deregulated (p-value<0.05). Here, we only present GO categories that are theoretically more relevant for adaptation to PSE. The Categories chosen in the DAVID annotation tool were the following: OMIM_DISEASE, GOTERM_BP_FAT (Biological Process), GOTERM_CC_FAT (Cellular Compartment), GOTERM_MF_FAT (Molecular Function) and KEGG_PATHWAY and will be indicated in the following figures as well as the number of annotated genes and p-values for each category.

In P1, $tRNA^{Ser}$(S) and $tRNA^{Ser}$(A) cell lines had higher number of deregulated genes compared with the other cell lines (see above) and also higher number of deregulated GO categories (Figure 3-2 A,B). In $tRNA^{Ser}$(S) cell line there was upregulation of categories related to the ER and the nuclear envelope-endoplasmic reticulum network. In these categories were annotated genes such as molecular chaperones (for example *DNAJB11*, which encodes an Hsp40), several ER enzymes and transporters, and elements of response to ER stress, namely *HERPUD1*, which encodes an inducible ER protein that is expressed under ER stress and plays a role in ERAD. Some genes involved in homeostatic processes (for instance, *TEP1*, a telomerase associated protein) and carbohydrate transport and biosynthesis (for example *PCK2*, which encodes an enzyme that participates in gluconeogenesis) were also upregulated. Categories related to cell adhesion processes and extracellular region were downregulated (Figure 3-2 A). On the contrary, in $tRNA^{Ser}$(A) cells, also in P1, cell adhesion, migration and proliferation were upregulated (Figure 3-2 B). This result was in agreement with our previous data (Chapter 2), where we could see an increase in cell proliferation from P1 to P15 (Figure 2-7 B). Genes involved in homeostatic process, several transcription factors and ER genes, namely the molecular chaperones encoded by *DNAJB11* and *HSP90B1* and a protein responsible for degradation of misfolded proteins, encoded by *DERL3*, were also upregulated. Downregulated categories in this cell line were involved in fatty acid metabolism and lipid binding (Figure 3-2 B).

$tRNA^{Ser}$(L) and $tRNA^{Ser}$(H) cell lines deregulated genes that were annotated in fewer categories than $tRNA^{Ser}$(S) and $tRNA^{Ser}$(A) cell lines (Figure 3-2 C, D). Some genes upregulated in $tRNA^{Ser}$(L) cells were membrane receptors and transporters annonated in intrinsic to plasma membrane GO category. Categories related to lipid biosynthesis and cell adhesion were downregulated in this cell line (Figure 3-2 C). In the $tRNA^{Ser}$(H) cell line categories related to ER (where was included, once again the molecular chaperone encoded by *DNAJB11* gene) and proteolysis were upregulated. Genes involved in transcription, macromolecule biosynthetic processes and constituents of the cytoskeleton were downregulated (Figure 3-2 D).

In P1, upregulation of ER components occured in all of our cell lines, with the exception for $tRNA^{Ser}$(L) cell line. This can be correlated with increased levels in misfolded proteins in these cell lines and consequent increase in ER capacity to fold and counteract protein aggregation. In fact, in these categories we could find some genes involved in the ER stress response. Cell adhesion and processes related to lipid biosynthesis seem to be common downregulated processes in our mistranslating cells. The $tRNA^{Ser}$(A) cell line presented a

different transcriptional profile, which can probably explain its proliferating phenotype (Table 3-1).

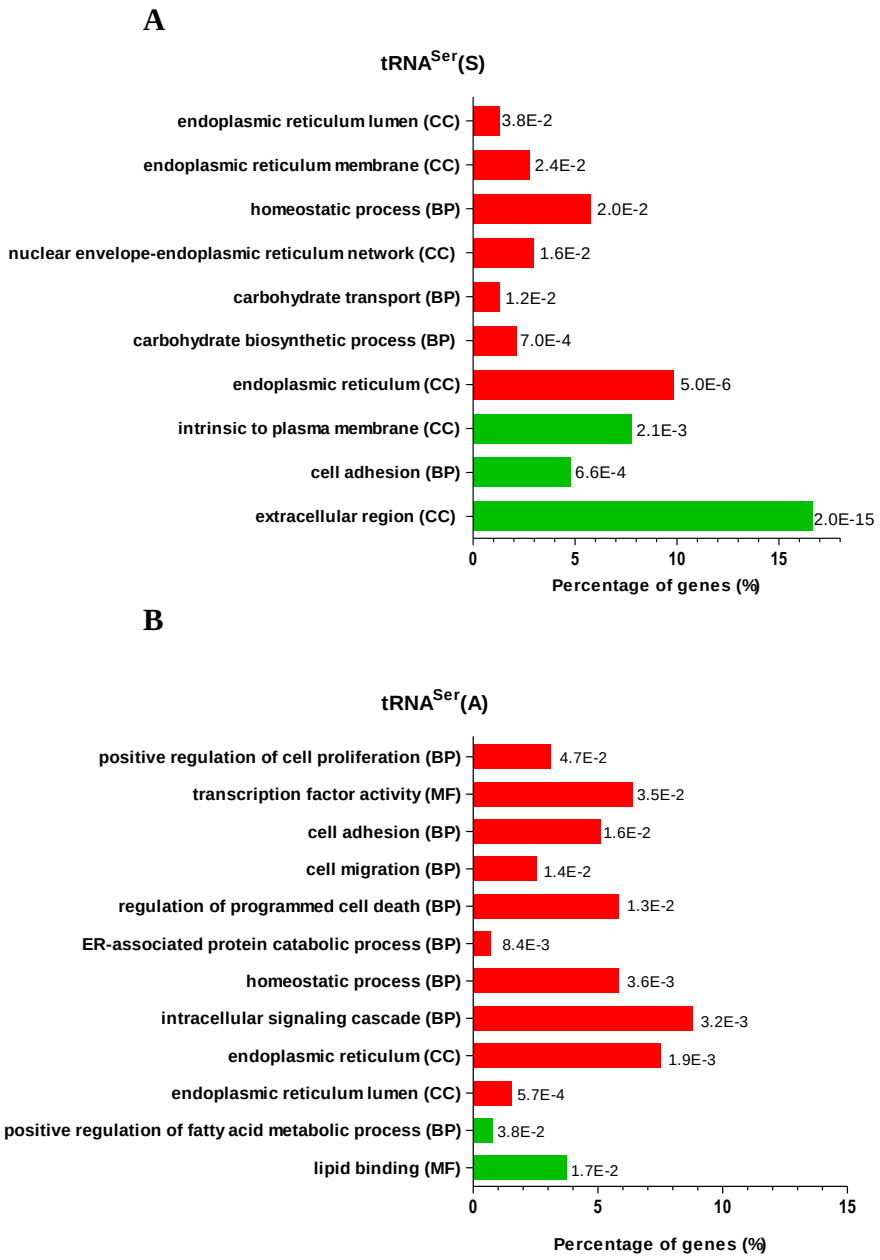
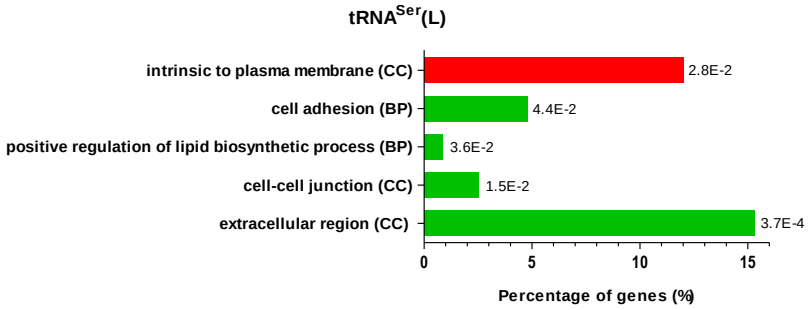


Figure 3-2.►

►C



D

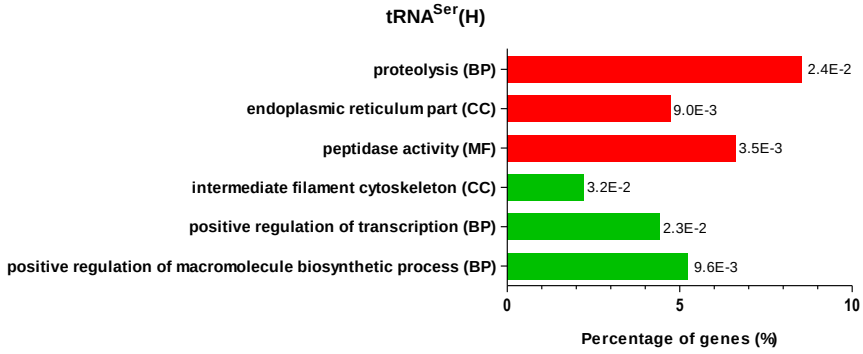


Figure 3-2. Functional class enrichment analysis (GO categories) of genes upregulated (red bars) and downregulated (green bars) in mistranslating cell lines in P1. A - tRNA^{Ser}(S); **B -** tRNA^{Ser}(A); **C -** tRNA^{Ser}(L); **D –** tRNA^{Ser}(H). *p*-values of each GO category are displayed in the graphs next to the respective bar.

In our intermediary time point, P15, the tRNA^{Ser}(S) cell line had the highest number of deregulated categories (Figure 3-3 A). Most of them are related to the response to misfolded proteins; UPS and ER stress: response to unfolded proteins, response to ER stress, proteasomal ubiquitin-dependent protein catabolic process, ER-nuclear signaling pathway, protein folding, unfolded protein binding, ER-Golgi intermediate compartment. The deregulation of these stress related categories is in line with our previous results of Chapter 2. The expression of the tRNA^{Ser} was higher in P15, in this cell line, and consequently there was also an increase in protein ubiquitination and fragmented ATF6. This protein is known to activate the transcription of ER genes to alleviate ER stress, which can explain the upregulation of ER related categories. Interestingly, the deregulation of the endogenous tRNA^{Ser} (in tRNA^{Ser}(S)) also lead to upregulation of several aaRSs genes, namely *AARS* - AlaRS, *NARS*-AspRS, *GARS*-GlyRS, *IARS*-IleRS, *MARS*-MetRS, *WARS*-TrpRS and *YARS*-TyrRS. Other upregulated

categories were mitochondrial lumen and pyruvate metabolism. The last one may indicate higher energetic requirement by these cells. Cell adhesion and peptidase inhibitor activity were downregulated (Figure 3-3 A).

The tRNA^{Ser}(A) cell line upregulated a GO category related to cell proliferation, as well as positive regulation of transcription factor activity and response to oxygen levels. Categories related to chromatin assembly, cytoskeleton regulation and cell adhesion were downregulated in P15 in this cell line (Figure 3-3 B). In tRNA^{Ser}(L) cell line the ER GO category was upregulated. Also, genes annotated in the GO category chromosome were upregulated, namely genes that encode histones (*HDAC8* and *HIST1H3A*) and centromere (*CENPE*) proteins and the telomerase associated protein-1(*TEP1*). Cell adhesion and components of extracellular matrix were downregulated in tRNA^{Ser}(L) cells, in P15 (Figure 3-3 C). The tRNA^{Ser}(H) cell line upregulated genes belonging to the plasma membrane part, namely several G-proteins and ion channels. Cell differentiation and adhesion processes were dowregulated in this cell line (Figure 3-3 D).

In P15, tRNA^{Ser}(S) and tRNA^{Ser}(L) cell lines upregulated GO categories related to ER stress and ER, respectively. In tRNA^{Ser}(S) cells, a higher number of categories related to misfolding protein accumulation, proteotoxic stress and cellular metabolism and energy production were upregulated. This may indicate higher consumption of energy in these cells. Despite some specific alterations, cell adhesion was clearly a dowregulated category in P15 (Table 3-1).

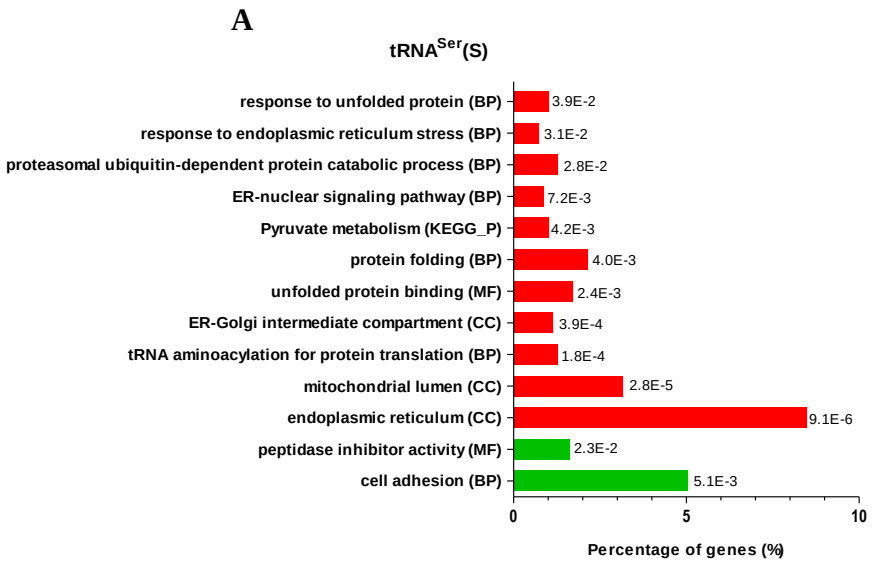


Figure 3-3.►

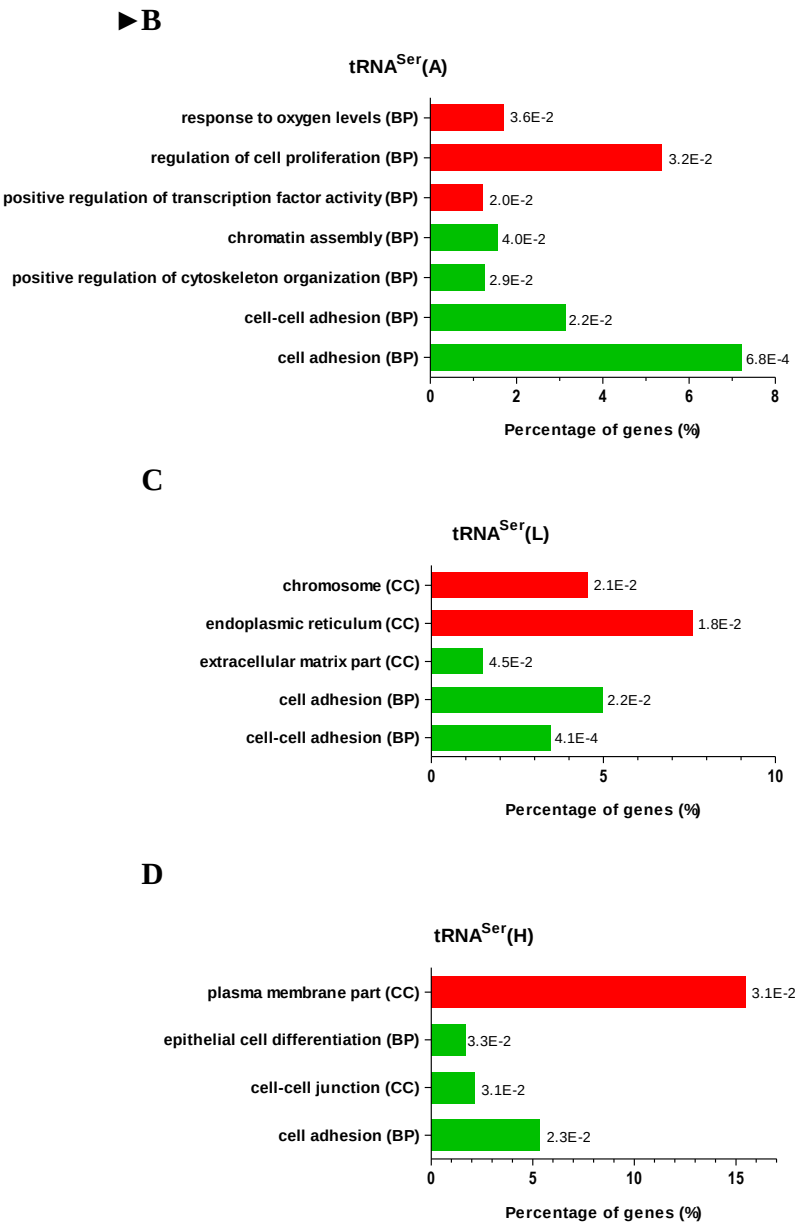


Figure 3-3. Functional class enrichment analysis (GO categories) of genes upregulated (red bars) and downregulated (green bars) in mistranslating cell lines in P15. A - $tRNA^{Ser}$(S); **B -** $tRNA^{Ser}$(A); **C -** $tRNA^{Ser}$(L); **D –** $tRNA^{Ser}$(H). *p*-values of each GO category are displayed in the graphs next to the respective bar.

Although the percentage of deregulated genes was similar for the three time points, few processes were deregulated in P30. At this passage number, $tRNA^{Ser}$(S) cells showed upregulation of genes involved in signal transduction and cell proliferation (Figure 3-4 A) and in fact, in Chapter 2 there was an increase in cell proliferation relative to Mock in this passage (Figure 2-7 A). Categories related to fatty acids and lipids metabolism were downderegulated

in $tRNA^{Ser}$(S) cells (Figure 3-4 A). In $tRNA^{Ser}$(A) cells, there was upregulation of genes involved in AMP catabolic process (Figure 3-4 B), which suggests metabolic alterations in these cells since AMP can function as an energy sensor (229). Genes involved in adhesion to the extracellular matrix belonging to the GO categories: focal adhesion, extracellular region and ECM-receptor interaction, were downregulated in $tRNA^{Ser}$(A) (Figure 3-4 B). In $tRNA^{Ser}$(L) cells, genes involved in transcription from RNA polymerase II promoter, negative regulation of cell proliferation and cell cycle arrest were up-regulated, while categories related to cell adhesion were downregulated (Figure 3-4 C). Finally, in $tRNA^{Ser}$(H) cells, genes encoding extracellular proteins, namely collagen and laminin were upregulated. Genes involved in positive regulation of cell differentiation, extracellular matrix (ECM-receptor interaction) and positive regulation of organelle organization were downregulated in this cell line (Figure 3-4 D).

Altogether these results show that genes involved in ER function and ER stress were upregulated in P1 and/or P15 in our cell lines, indicating that transcriptional responses to proteotoxic stress are very important during early adaptation to PSE, probably because mutant tRNA expression was higher (Table 3-1). The $tRNA^{Ser}$(S) cell line is interesting because it activated carbohydrate biosynthesis and pyruvate metabolism suggesting that adaptation to PSE requires high leves of energy. The proliferating phenotype of the $tRNA^{Ser}$(A) cell line, was confirmed since these cells upregulated genes related to proliferation and migration. Also, this cell line upregulated genes that respond to oxygen levels and others involved in AMP catabolic process and responders to metabolic alterations. Surprisingly no transcriptional responses to proteotoxic stress were observed in $tRNA^{Ser}$(L) cells in P1, and only the ER category was upregulated in P15. Despite this result, some individual stress genes may have been deregulated. Regarding dowregulated genes, they were mainly involved in cell adhesion or lipids metabolism. In $tRNA^{Ser}$(H) cell line, in addition to cell adhesion, also genes annotated in the categories of cytoskeleton, cell differentiation and organelle organization were downregulated. Table 3-1 summarizes the main processes altered in each passage in our mistranslating cell lines.

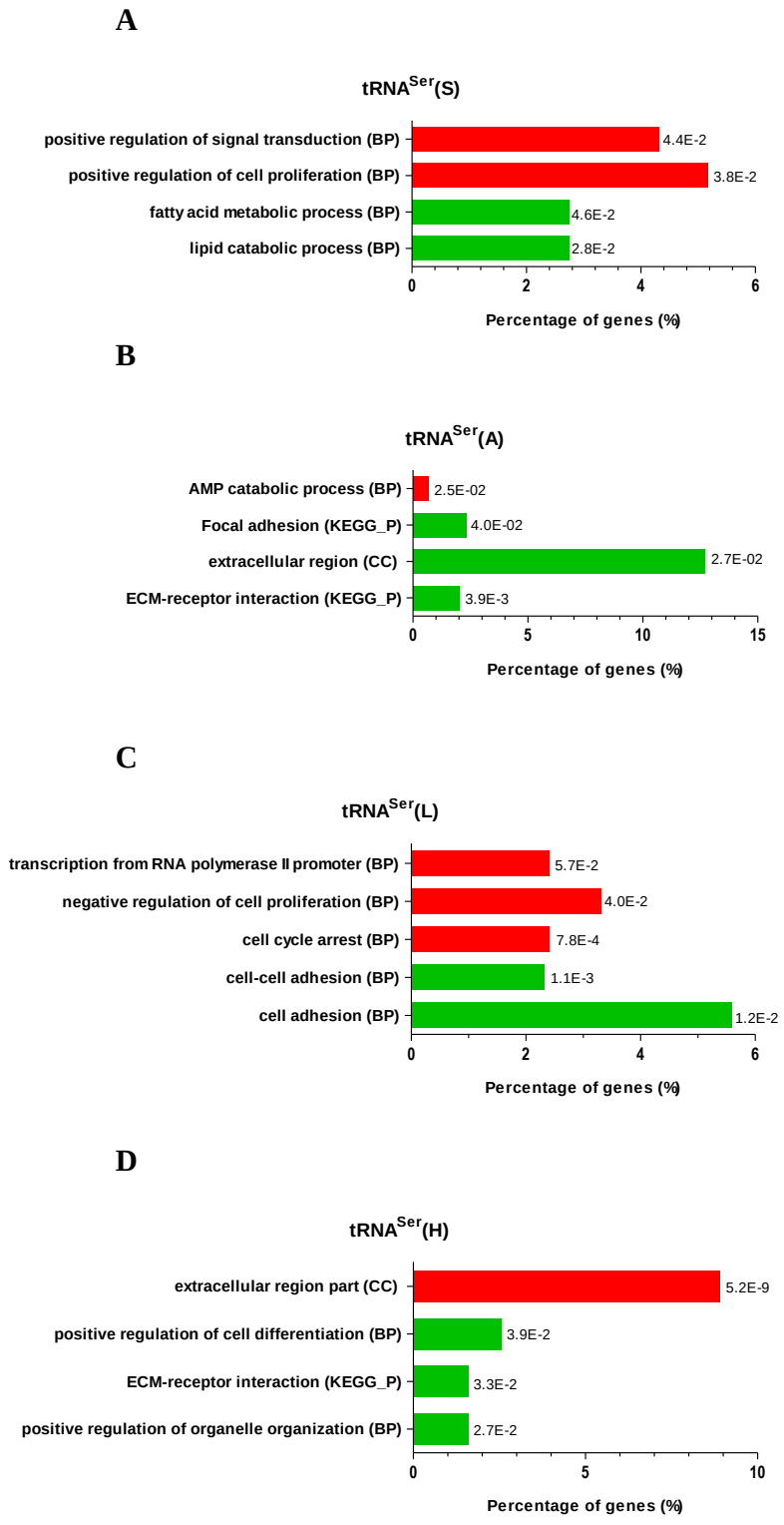


Figure 3-4. Functional class enrichment analysis (GO categories) of genes upregulated (red bars) and downregulated (green bars) in mistranslating cell lines in P30. A - $tRNA^{Ser}$(S); **B -** $tRNA^{Ser}$(A); **C -** $tRNA^{Ser}$(L); **D –** $tRNA^{Ser}$(H). *p*-values of each GO category are displayed in the graphs next to the respective bar.

Table 3-1. Summary of the cellular processes upregulated (red arrow) and downregulated (green arrow) in the mistranslating cell lines in the three passages.

	P1	P15	P30
tRNA^{Ser}(S)	ER Carbohydrate biosynthesis ↑ Cell adhesion Extracellular region ↓	ER and ER stress UPS Pyruvate metabolism tRNA aminoacylation ↑ Cell adhesion ↓	Signal transduction Proliferation ↑ Lipids metabolism ↓
tRNA^{Ser}(A)	ER Proliferation Cell adhesion Migration ↑ Lipids metabolism and binding ↓	Response to oxygen levels Proliferation ↑ Chromatin assembly Cell adhesion Cytoskeleton ↓	AMP catabolic process ↑ Focal adhesion Extracellular region ↓
tRNA^{Ser}(L)	Intrinsic to plasma membrane ↑ Cell adhesion Lipids biosynthesis Extracellular region ↓	ER Chromosome ↑ Cell adhesion Extracellular matrix ↓	Cell cycle arrest Transcription poll II Negative regulation of proliferation ↑ Cell adhesion ↓
tRNA^{Ser}(H)	ER Proteolysis ↑ Cytoskeleton Transcription ↓	Plasma membrane part ↑ Cell differentiation Cell adhesion ↓	Extracellular region ↑ Cell differentiation Organelle organization ↓

3.3.3 Long term gene expression deregulation in mistranslating cells

To identify gene expression alterations during the successive passages of our cell lines the genes that were up and downregulated between each time point for each cell line were identified. For instance, to identify genes up and downregulated from P1 to P15 in tRNA^{Ser}(S) cells we calculated the fold change of tRNA^{Ser}(S) in P15 relative to tRNA^{Ser}(S) in P1. Significantly altered genes were considered in our analysis (p<0.05) if fold change was above 1.5 or below -1.5. The same approach was used for our control cell line (Mock). We then removed the deregulated genes that appeared in our control cell (Mock P1-P15) from the list tRNA^{Ser}(S) P1-P15 (Figure 3-5). This was done for our four cell lines (tRNA^{Ser}(S), tRNA^{Ser}(A), tRNA^{Ser}(L), tRNA^{Ser}(H)) between the three time points, P1-P15, P15-P30 and P1-

P30. With this strategy we could identify genes that were specifically deregulated in each cell line, in response to PSE and excluded from the analysis genes whose deregulation was due to cell evolution in culture.

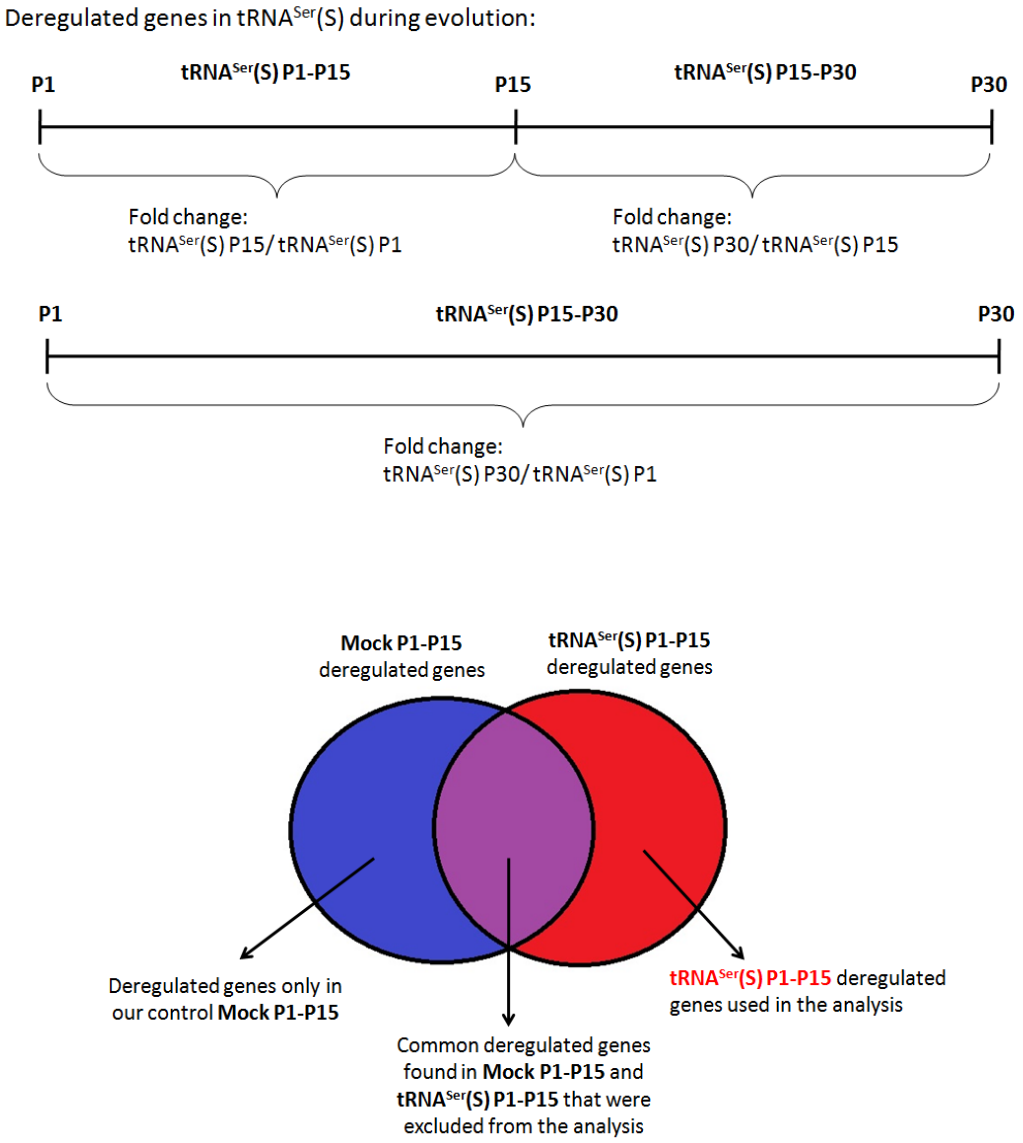


Figure 3-5. Diagram showing the experimental design used to obtain the list of tRNA^{Ser}(S) genes used in the transcriptional analysis of cell evolution.

From P1 to P15, all cell lines deregulated approximately 4% of their genes, except for tRNA^{Ser}(L) which deregulated 2.5% (Figure 3-6 A). The percentage of upregulated genes was higher than that of downregulated genes in tRNA^{Ser}(S) (2.3% up; 1.5% down), tRNA^{Ser}(L) (1.7% up; 0.8% down) and tRNA^{Ser}(H) (1.8% up;1.5% down) cells. tRNA^{Ser}(A) cells had more

genes downregulated (2.5%) than upregulated (1.2%) (Figure 3-6 A). From P15-P30, tRNA^{Ser}(H) cells showed higher percentage of deregulated genes (6.9%) followed by tRNA^{Ser}(L) (5.8%), tRNA^{Ser}(A) (4.1%) and tRNA^{Ser}(S) cells (2.5%) (Figure 3-6 B). tRNA^{Ser}(L) and tRNA^{Ser}(H) cells had higher number of upregulated genes (2% and 5.5%, respectively). During the evolution in culture, from P1 to P30, the cell lines that had more genes deregulated were tRNA^{Ser}(L) and tRNA^{Ser}(H) (5.6% and 4.9%) (Figure 3-6 C). tRNA^{Ser}(S) and tRNA^{Ser}(A) cell lines showed lower percentage of deregulated genes (2.8% and 3.8%, respectively) (Figure 3-6 C). Regarding the percentage of deregulated genes, the pattern is very similar in P15-P30 and P1-P30.

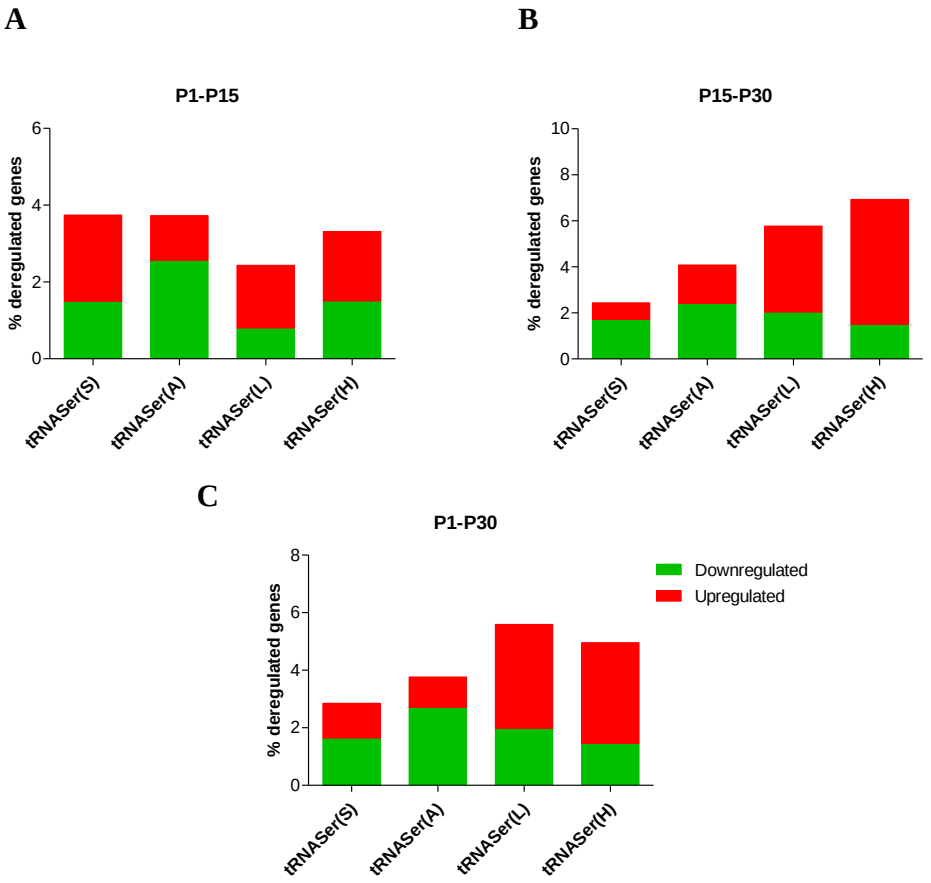


Figure 3-6. The percentage of deregulated genes in each cell line between passages. A - P1-P15; **B** – P15-P30; **C**- P1-P30

3.3.4 Gene Ontology categories deregulated during evolution

To complete gene expression deregulation analysis in each time point, we also studied each cell line between the time points P1-P15, P15-P30 and P1-P30. Deregulated genes were

also annotated in GO categories using DAVID Bioinformatics Resources 6.5. As before, we discuss the GO categories that were potentially relevant to understand the adaptation to PSE during evolution. The Categories chosen in DAVID annotation tool were the following: OMIM_DISEASE, GOTERM_BP_FAT (Biological Process), GOTERM_CC_FAT (Cellular Compartment), GOTERM_MF_FAT (Molecular Function) and KEGG_PATHWAY.

Table 2-1 shows the GO categories upregulated from P1 to P15 and from P15 to P30. During the initial part of the evolution, genes involved in carbohydrate catabolic process and glycolysis were upregulated in $tRNA^{Ser}$(S) and $tRNA^{Ser}$(A) cell lines, respectively. This may indicate that these two cell lines require glucose to boost ATP production during adaptation to PSE. The $tRNA^{Ser}$(A) cell line also showed upregulation of mitosis from P1 to P15, which correlates with an increase in cell proliferation capacity of this cell line (Chapter 2, Figure 2-7). Also, between P1 and P15, $tRNA^{Ser}$(L) and $tRNA^{Ser}$(H) cell lines showed upregulation of genes related to cell adhesion, while $tRNA^{Ser}$(L) cell line, upregulated ER genes. In particular stress related genes such as *DNAJC3*, which encodes a homolog of Hsp40, involved in the UPR. p53 signaling pathway was another category upregulated in this cell line that contained genes involved in regulation of the cell cycle and apoptosis (such as *CASP8*, which encodes caspase 8). Between P15 and P30, genes involved in transcription process were up-regulated in $tRNA^{Ser}$(S), $tRNA^{Ser}$(A) and $tRNA^{Ser}$(L) cells. Also, signaling pathways were upregulated in $tRNA^{Ser}$(A) and $tRNA^{Ser}$(L) cells. Specifically, in $tRNA^{Ser}$(L) cell line, there was upregulation of genes related to MAPK signaling, namely *MAP3K4*, which encodes a kinase that modulates the UPR.

Table 3-2. GO categories upregulated in mistranslating cell lines during evolution in culture.

	P1 → P15			P15 → P30		
	GO category	**%genes**	***p*-value**	**GO category**	**%genes**	***p*-value**
tRNA^{Ser}(S)	cellular carbohydrate catabolic process (BP)	1.0	5.0E-2	superoxide anion generation (BP)	1.5	5.5E-3
				regulation of specific transcription from RNA polymerase II promoter (BP)	2.0	4.9E-2
tRNA^{Ser}(A)	glycolysis (BP)	1.3	2.6E-2	transcription activator activity (MF)	3.6	1.9E-2
	positive regulation of mitosis (BP)	1.0	4.5E-2	intracellular signaling cascade (BP)	8.1	4.1E-2
	Glycolysis / Gluconeogenesis (KEGG_P)	1.3	4.8E-2			
tRNA^{Ser}(L)	positive regulation of cell adhesion (BP)	1.4	5.6E-3	RNA polymerase II transcription factor activity (MF)	2.6	3.9E-4
	p53 signaling pathway (KEGG_P)	1.2	3.9E-2	MAPK signaling pathway (KEGG_P)	2.7	6.1-4
	endoplasmic reticulum (CC)	6.3	4.2E-2	transcription (BP)	12.2	9.5E-3
				ubiquitin cycle (BP)	0.5	2.2E.2
tRNA^{Ser}(H)	cell adhesion (BP)	6.2	5.2E-4	cell adhesion (BP)	4.7	2.0E-4
				cellular ion homeostasis (BP)	2.4	1.9E-2

Regarding downregulated categories, from P1 to P15 (Table 3-3), in tRNA^{Ser}(S) cells the p53 signalling pathway was visible, where annotated genes are involved in DNA damage repair and cell cycle arrest. From P1 to P15, tRNA^{Ser}(H) cells dowregulated ER genes, namely *HSP90B1* and *DNAJB11*, which are two molecular chaperones, and cell motility and migration genes. As shown before, this cell line upregulated genes related to ER stress in P1 (Figure 3-2 D), but between P1 and P15 there was downregulation of ER genes, such as molecular chaperones. Results described in Chapter 2, showed that the levels of some chaperones were decreased in the end of evolution in tRNA^{Ser}(H) cell line. Also, Chapter 2 data showed that tRNA^{Ser}(H) cells had lower capacity to form colonies in culture compared with Mock in P15. This was probably due to downregulation of several genes involved in cell motility and migration.

Between P15 and P30 (Table 3-3), tRNA^{Ser}(S) cells dowregulated genes involved in cAMP biosynthesis, which plays important roles in signaling processes. In tRNA^{Ser}(A) cells regulation of growth and a KEGG pathway related to the extracellular matrix were downregulated. tRNA^{Ser}(L) cells showed downregulation of categories related to cell adhesion and plasma membrane, but also chromatin assembly and disassembly. In this last category,

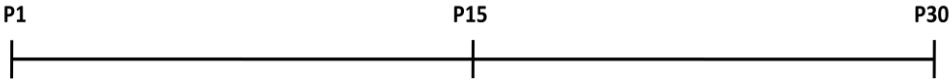
several histone coding genes, which play a central role in transcription regulation, DNA repair, DNA replication and chromosomal stability, were annotated. Finally, in $tRNA^{Ser}$(H) cells processes involved in regulation of cell division and differentiation, protein polyubiquitination and regulation of the actin cytoskeleton were downregulated from P15 to P30.

Table 3-3. GO Categories downregulated in mistranslating cell lines during evolution

	P1 → P15			P15 → P30		
	GO category	**%genes**	***p*-value**	**GO category**	**%genes**	***p*-value**
$tRNA^{Ser}$(S)	p53 signaling pathway (KEGG_P)	1.3	2.8E-2	regulation of cAMP biosynthetic process (BP)	1.4	4.0E-2
$tRNA^{Ser}$(A)	amino acid transmembrane transporter activity (MF)	1.1	1.1E-2	regulation of cell growth (BP)	2.1	7.6E-3
				ECM-receptor interaction (KEGG_P)	1.4	3.0E-2
$tRNA^{Ser}$(L)	extracellular region part (CC)	8.6	2.5E-2	intrinsic to plasma membrane (CC)	8.3	1.0E-2
				chromatin assembly or disassembly (BP)	1.6	1.9E-2
				cell adhesion (BP)	4.7	2.7E-2
$tRNA^{Ser}$(H)	homeostatic process (BP)	7.8	2.7E-4	positive regulation of cell division (BP)	1.3	4.5E-3
	endoplasmic reticulum part (CC)	3.6	1.3E-2			
				protein polyubiquitination (BP)	1.0	1.0E-2
	cell motility (BP)	3.1	3.6E-2			
				Regulation of actin cytoskeleton (KEGG_P)	2.4	3.2E-2
	cell migration (BP)	2.9	4.1E-2			
				epithelial cell differentiation (BP)	1.9	3.2E-2

The Figure 3-4 shows a summary of the main processes altered between P1 and P15 and P15 and P30. Each cell line shows a specific transcriptional response although some processes are common to some cells, such as transcription and cell adhesion. In $tRNA^{Ser}$(H) cells, it is interesting to note that some categories related to the ER and the response to proteotoxic stress (polyubiquitination of proteins) were downregulated. This may indicate that adaptation and survival of $tRNA^{Ser}$(H) cells in the last part of evolution does not rely on activation of PQC mechanisms.

Table 3-4. GO categories that are deregulated in mistranslating cell lines between P1 and P15 and at later stages (between P15 and P30).



tRNA^{Ser}(S)	Catabolism of carbohydrates ↑ p53 signaling pathway ↓	Regulation of specific transcription from pol II ↑ Regulation of cAMP biosynthesis ↓
tRNA^{Ser}(A)	Glycolysis Mitosis ↑ Amino acid transport ↓	Transcription activator activity Intracellular signaling cascades ↑ Regulation of cell growth Extracellular matrix ↓
tRNA^{Ser}(L)	Cell adhesion ER ↑ Extracellular region ↓	Transcription Ubiquitin cycle MAPK signalling ↑ Chromatin Cell adhesion ↓
tRNA^{Ser}(H)	Cell adhesion ↑ ER Cell migration ↓	Cell adhesion ↑ Cell division and differentiation Polyubiquitination ↓

From P1 to P30 (Table 3-5), we could observe upregulation of genes related to glycolysis and gluconeogenesis in tRNA^{Ser}(S) cells, similarly tRNA^{Ser}(A) cells upregulated categories involved in metabolism and energy, such as regulation of glycogen biosynthetic process. In tRNA^{Ser}(L) cells several GO categories were upregulated, in particular, cell cycle, transcription and carbohydrate catabolic process. Gluconeogenesis and fatty acid metabolism were upregulated in tRNA^{Ser}(H) cells.

A common category was downregulated between P1 and P30, namely extracellular region (Table 3-6). In tRNA^{Ser}(H) cells genes involved in migration, cell division and the cytoskeleton were also downregulated. An interesting category that was downregulated in tRNA^{Ser}(H) cell line was L-serine biosynthetic process. Since in these cells Ser is misincorporated at non-cognate codon sites it is likely that decrease in Ser synthesis decreases PSE.

Table 3-5. GO categories upregulated in mistranslating cell lines during evolution in culture (P1-P30).

P1 → P30

	GO category	%genes	*p*-value
$tRNA^{Ser}(S)$	gluconeogenesis (BP)	1.2	5.0E-3
	Glycolysis / Gluconeogenesis (KEGG_P)	1.6	1.8E-2
$tRNA^{Ser}(A)$	regulation of generation of precursor metabolites and energy (BP)	1.4	6.2E-3
	regulation of glycogen biosynthetic process (BP)	1.1	1.3E-2
$tRNA^{Ser}(L)$	cell cycle process (BP)	4.0	5.7E-3
	condensed chromosome (CC)	1.4	1.3E-2
	negative regulation of transcription, DNA-dependent (BP)	2.5	2.9E-2
	transcription from RNA polymerase II promoter (BP)	1.8	4.1E-2
	rRNA processing (BP)	1.0	4.2E-2
	carbohydrate catabolic process (BP)	1.1	4.2E-2
$tRNA^{Ser}(H)$	extracellular region (CC)	15.9	1.2E-11
	gluconeogenesis (BP)	0.6	1.3E-2
	positive regulation of fatty acid metabolic process (BP)	0.5	4.1E-2

Table 3-6. GO categories downregulated in mistranslating cell lines during evolution in culture (P1-P30).

P1 → P30

	GO category	%genes	*p*-value
$tRNA^{Ser}(S)$	extracellular region (CC)	12.1	3.2E-2
$tRNA^{Ser}(A)$	extracellular region part (CC)	8.7	1.5E-6
$tRNA^{Ser}(L)$	extracellular region part (CC)	8.0	1.3E-3
$tRNA^{Ser}(H)$	cell migration (BP)	4.1	2.6E-4
	positive regulation of cell division (BP)	1.6	4.8E-4
	extracellular region part (CC)	8.4	8.3E-4
	L-serine biosynthetic process (BP)	0.8	8.5E-4
	actin cytoskeleton (CC)	2.7	4.0E-2

3.4 Discussion

To understand the molecular bases of the phenotypes induced by PSE in human cells we have analyzed the transcriptional profile of mistranslating cells described previously. We wanted to unveil possible pathways and mechanisms of gene expression reprogramming by PSE. Our results showed that transcriptional responses were different between cell lines and culture passage number. Many upregulated genes were annotated into categories related to ER and ER stress response, as we would expect. However, mistranslating cell lines deregulated several genes involved in energy metabolism, transcription, chromosome, signaling pathways (p53, MAPK) and many other processes, indicating that other mechanisms could be important to adaptation to PSE.

We have observed that the expression of our mutant tRNAs or overexpression of the Wt $tRNA^{Ser}$ in HEK293 cells leads to the degerulation of a small percentage of genes, always below 8%. In *C. albicans* mistranslating strains the number of deregulated genes in cDNA microarrays was correlated with the level of Leu misincorporation (75,230). While low-level mistranslation (0.64%) deregulated only 2% of genes, high levels of mistranslation (20% to 80%) deregulated approximately 16.9% to 26.2% of the *C. albicans* genes (75,230). Since the expression of our mutant tRNAs decreases abruptly from P1 to P15, possibly leading to a sharp decrease in Ser misincorporation levels, it was not surprising that a low percentage of deregulated genes was observed in our cells. Also, we have to take into consideration that before the first evolution time point (P1), HEK293 cells expressed the mutant tRNAs during 5 passages. This was necessary to select the cells that expressed the mutant tRNAs stably.

Due to the structural differences between the amino acids Ser and Leu and the phenotypes observed, specially the accumulation of protein aggregates in $tRNA^{Ser}$(L) cells in P15, we were expecting higher percentage of deregulated genes in this cell line. In fact, $tRNA^{Ser}$(L) cells in P1 had the lowest percentage of deregulated genes (2.1%), relative to the other cell lines (Figure 3-1 A). Only in P30 the percentage of deregulated genes increased (5.9%) (Figure 3-1 C). Also from P1 to P15, $tRNA^{Ser}$(L) cells had the lowest number of deregulated genes (2.5%), which increased from P15 to P30 (5.8%) (Figure 3-6). From P1 to P30, $tRNA^{Ser}$(L) cells had higher percentage of deregulated genes comparatively to the other cell lines, suggesting that the impact of mistranslation is higher in this cell line (Figure 3-6 C). The overexpression of the Wt $tRNA^{Ser}$ had higher impact in transcriptional deregulation comparatively to other cell lines, particularly in P1 and P15 (7.4% and 4.9% of deregulated genes, respectively) (Figure 3-1). This is not surprising since previous studies show that

overexpression of a specific tRNA in human cells can modulate protein expression (159,204). In any case, this cell line had the lowest percentage of deregulated genes (2.8%) between P1 and P30 (Figure 3-6 C).

In the previous chapter, we used cDNA microarrays and selected genes that are known to be involved in the response to misfolded proteins in human cells. From these, only a few genes were deregulated. The analysis of the GO categories, in this chapter, allowed us to identify other genes that also participate in response to misfolded proteins. In particular, ER stress response genes. The UPR is a very complex cellular process that includes changes in expression of many genes and different cell lines and individuals show variability in these genes expression patterns (231). Dombroski *et al.*, exposed B cells to tunicamycin or thapsigargin and observed that among the top 100 genes that showed the most significant variation in fold change in response to ER stress, only 40% were known "UPR genes" (231). In our data set we have identified several genes belonging to the classic UPR pathway, for example; *DNAJB11*, *HERPUD1*, *HSP90AB1* and *DERL3*. Other genes that do not participate in UPR, but are usually upregulated in cases of proteotoxic stress were also identified (231). This was the case of *PCK2*, which encodes phosphoenolpyruvate carboxykinase 2 (mitochondrial), which was upregulated in $tRNA^{Ser}$(S) cells in P1, and annotated in carbohydrate biosynthetic process category (Figure 3-2 A). These results complement our previous data on the PQC mechanisms activated in response to PSE. The activation and synchronization throughout time of these mechanisms in all cell lines supports the hypothesis that PQC mechanisms are essential for cells to maintain homeostasis and to survive (232–234).

Besides PQC mechanisms, several other genes were deregulated in our cells. Genes involved in glycosis process were upregulated in $tRNA^{Ser}$(S) cells from P1 to P30 (Table 3-5) and in $tRNA^{Ser}$(A) cells from P1 to P15 (Table 3-2). A previous study correlated ER stress in hematopoietic cells with a relative inability of these cells to use extracellular glucose, resulting in reduced glycolysis and Kreb's cycle activity (235), however in cancer cells, where PQC mechanisms are usually upregulated (including UPR), there is increased glucose uptake and glycolysis (236–239). Glycolysis is a source of ATP and plays a key role in cell survival and adaptation of proliferating cells (236,239). In cells producing misfolded proteins, maintenance of ATP levels is essential since many PQC systems are ATP dependent, such as some chaperones, the UPS, and mechanisms of folding and degradation in the ER (175,237,240–242). Increasing glycolysis could help $tRNA^{Ser}$(S) and $tRNA^{Ser}$(A) cells to fold and degrade misfolded proteins and, even increase proliferation rate (Chapter 2, Figure 2-7). In a previous evolutionary study, using mistranslating strains of *S. cerevisiae*, it was observed that evolved

cell lines had an accelerated protein turnover. In response to this high-energy demand, glucose uptake was increased due to the selective duplication of hexose transporter genes. These cells were only able to thrive in nutrient-rich environments because of its exceptionally high energy demands (196).

Usually the control of glycolysis and energy production in proliferating cells, for example cancer cells, primarily resides at the transport and phosphorylation steps. For example, in upregulation of glucose transporters and hexokinases (239). However, in mammals, some tissues such as liver, kidney and gastrointestinal tract have gluconeogenic enzyme activity (243). Since our HEK293 cells are human embryonic kidney cells, it is not surprising that we could see upregulation of genes related to gluconeogenesis in $tRNA^{Ser}$(S) and $tRNA^{Ser}$(H) cell lines from P1 to P30 (Table 3-5), suggesting that these cell lines increase glucose production.

Cell signaling pathways, such as p53 and MAPK, were also deregulated in our cell lines. p53 signaling pathway was downregulated in $tRNA^{Ser}$(S) cells from P1 to P15 (Table 3-3), but is was upregulated in $tRNA^{Ser}$(L) cells also in P1 to P15 (Table 3-2). This pathway is activated in cells in response to various forms of stress, but mainly genotoxic stress. p53 targeted genes are involved in cell-cycle arrest, DNA repair or apoptosis if the damage is irreparable (244). In the case of $tRNA^{Ser}$(S) cells, genes belonging to this pathway were downregulated, possibly to avoid cell-cylce arrest and apoptosis. On the other hand, in $tRNA^{Ser}$(L) cells, the p53 pathway was upregulated. The upregulation of genes involved in control of the cell cycle and apoptosis effectors (such as *CASP8*, which encodes caspase 8) can possibly explain the phenotype observed in this cell line, a decrease in the number of viable cells from P1 to P15 (Chapter 2, Figure 2-6). The connection between misfolded proteins accumulation and p53 signaling pathway activation is still poorly understood. It its known, however, that severe or prolonged proteotoxic stress induces activation of cell death through apoptosis and p53 signaling pathway may be involved (245,246).

In $tRNA^{Ser}$(L) cells there was also upregulation of MAPK signaling pathway, from P15 to P30 (Table 3-2). Recent evidences show that mitogen-activated protein kinase (MAKP) signaling pathways have a role in the response to ER stress (247). Indeed, some MAPK pathways are activated in response to ER stress (such as JNK and p38) (247). One of the genes upregulated in $tRNA^{Ser}$(L) cells encodes mitogen-activated protein kinase kinase 4 (MKK4), which can activate JNK, leading to apoptosis, and modulate UPR via p38-dependent phosphorylation of CHOP and ATF6 (247–249). MAKP signaling may be involved, in this

way, in the response to the accumulation of misfolded proteins and consequent ER stress in the $tRNA^{Ser}$(L) cell line.

Genes involved in lipid metabolism were downregulated in our mistranslating cell lines (Table 3-1). In addition to its activities in the processing of secreted and membrane proteins, the ER is a major site of synthesis of membrane lipids in eukaryotic cells, particularly phospholipids and sterols (250). It is not surprising though, that accumulation of misfolded proteins in the ER lead to alterations in lipid metabolism. In *C. elegans*, acute ER stress (caused by tunicamycin) cause adaptative changes in the expression of several genes involved in phospholipids metabolism and phosphatidylcholine, the most abundant ER membrane lipid. The induction of these genes may reflect demands for ER expansion (250–252). Indeed, elements of the UPR are involved in lipids biosynthesis, for example, XBP-1 promotes transcription of genes involved in fatty acid and cholesterol biosynthesis (253). Despite these results, we have observed downregulation of lipid metabolism categories. Possibly, the induction of genes involved in lipid metabolism could be relevant in severe and acute stress, but not in prolonged adaptation to PSE.

Our cell lines showed several transcriptional alterations that go beyond activation of PQC mechanisms. Specially, $tRNA^{Ser}$(S) and $tRNA^{Ser}$(A) cells showed transcriptional activation of genes that can explain adaptation to PSE and survival. Together, this and Chapter 2 data, raises the hypothesis that these advantageous features could be important in different microenvironments, for example if these cells transform in an *in vivo* situation. A previous study using NIH3T3 mistranslating cells, showed that expression of mutant tRNAs lead to increased cell transformation *in vivo*, UPR and cancer-related pathways activation (unpublished data, Annex E), which may indicate that low levels of PSE in mammalian cells could be advantageous in specific conditions.

3.5 Conclusions

The microarray analysis carried out in this chapter complemented our previous study on the activation of PQC mechanisms in cells that produced aberrant proteins. It was also possible to unveil new pathways (p53 and MAPK) and adaptation mechanisms (increase in gluconeogenesis/ glycolysis) involved in the response to PSE during human cell evolution in culture.

As we would expect, the different mistranslating cell lines created have different transcriptional profiles, and the majority of these results can be correlated with our previous

phenotypes observed in Chapter 2. The most interesting cell lines, regarding evolutionary adaptation, were the $tRNA^{Ser}$(S) and $tRNA^{Ser}$(A) cells. Despite activation of several PQC mechanisms, transcriptional analysis showed alterations in many other categories, for example cell proliferation, cell adhesion, glycolysis, suggesting that PSE may even have positive effects in human cells. Another interesting feature is that misincorporation of Ser at Ala codons ($tRNA^{Ser}$(A)) is an error that occurs at the physiological level due to chemical similarity between the two amino acids. It will be interesting to determine in future studies if human cells are better adapted to Ser-Ala misincorporations than to other non-phisiological errors.

3.6 Materials and Methods

3.6.1 Cell culture

Human embryonic kidney 293 (HEK293) were grown in Minimum Essential Medium (Gibco, Cat.41090-028) supplemented with 10% fetal bovine serum (FBS) (Sigma, Cat.F1051), 1% of Pen/Strep (Gibco, Cat.15070-063) and 1% of non-essential amino acids (Gibco, Cat.11140-050) in a humidified atmosphere at 37ºC in the presence of 5% CO_2.

3.6.2 Mistranslating cell lines used

HEK293 cells stably transfected with an empty vector (Mock), and with the plasmid carrying $tRNA_{AGA}^{Ser}$(S), $tRNA_{AGC}^{Ser}$(A), $tRNA_{AAG}^{Ser}$(L) or $tRNA_{GTG}^{Ser}$(H) genes were used in this study in the passages 1, 15 and 30 (see Methods from Chapter 2).

3.6.3 Total RNA extraction

RNA was extracted using Trizol®Reagent (Thermo Fisher Scientific, Cat.15596026). The content of one well from a 6 well plate, with around $5x10^5$cells, was collected for each experimental condition. Purification of RNA was done using DNaseI, Amplification Grade kit (Invitrogen, Cat.18068015), following manufacturer's instructions. RNA was then precipitated with a standard Phenol/Chlorophorm/Isoamylalchohol (25:24:1) (Acros Organics, Cat.327111000) extraction protocol and conserved at -80ºC. RNA concentration was

determined using NanoDrop1000 (Thermo Scientific). RNA quality was verified using Agilent 2100 Bioanalyser.

3.6.4 Gene expression microarrays

Gene expression microarrays profiling was performed using the Agilent protocol for One-Color Microarray Based Gene Expression Analysis Low Input Quick Amp Labeling v6.9 (Agilent Technologies). RNA quality determination was performed using 2100 Bioanalyser (Agilent Technologies) and the Agilent RNA 6000 Nano kit (Agilent Technologies, Cat.5067-1512). 100ng of total RNA were used to synthesize labeled cDNA (with Cyanine 3-CTP) using Agilent T7 Promoter Primer and T7 RNA polymerase Blend (Agilent Technologies, Cat.5190-2305). cDNA was purified with RNAeasy mini spin columns (Quiagen, Cat.74104). Dye incorporation was quantified using Nanodrop 1000 Spectophotometer. 600ng of labeled cDNA were hybridized in Sure Print G3 Human Gene Expression 8x60k v2 microarrays (Agilent Technologies, Cat.G4851B). Hybridizations were carried out using Agilent gasket slides in a rotating oven for 17h at 65ºC. Slides were then washed following manufacturer's instructions and scanned in an Agilent G2565AA microarrays scanner. Probes signal values were extracted using Agilent Feature Extraction Software.

The microarray raw data was submitted to the GEO database and has been given the following accession number: GSE93854.

3.6.5 Statistical analysis

Data were normalized using median centering of signal distribution with Biometric Research Branch BRB-Array tools v3.4.o software (217,218). Each gene's measured intensity was median normalized to correct for differences in the labeling efficiency between samples using the most variable probe (set) measured by IQR across arrays. Median normalized values were then divided by its value in the control sample in order to obtain the M values (Log_2 intensity ratios) and Fold Change. Microarrays statistical analysis was carried out using Mev software (TM4 Microarray Software Suite) (219,220). T-test (calculated based on Welch approximation; *p*-value based on permutations; alpha critical *p*-value=0.05) was performed to identify genes that showed differences in expression between our control (Mock) and our samples.

3.6.6 Gene expression deregulation analysis

Significant genes that present a fold change above 1.5 or bellow -1.5 (between each cell line and the control) were considered as differentially expressed and used for our analysis. GO term enrichment analysis were carried out using DAVID Bioinformatics Resources 6.7, an online database (254,255).

4. Genome alterations of HEK293 mistranslating cells

4.1 Abstract

In the previous chapters, we characterized mistranslating cell lines regarding their phenotypes, activation of protein quality control (PQC) mechanisms and transcriptional profiles. Some molecular mechanisms involved in adaptation of these cells to protein synthesis errors (PSE) were identified, but the genomic mutations that permit that adaptation were not identified. In this chapter, we tackle this question to better understand the biology of PSE. For this, we took advantage of next-generation sequencing tools, namely whole-exome sequencing.

The exomes of our cell lines Mock, $tRNA^{Ser}$(S), $tRNA^{Ser}$(A), $tRNA^{Ser}$(L) and $tRNA^{Ser}$(H) were sequenced at the end of the evolution experiment (P30), to identify variants such as single nucleotide polymorphisms (SNPs) and small Insertions and Deletions (INDELs) that accumulated in the coding genome.

We were able to identify mutations which were absent in the Mock cell line, in genes that may be important for the adaptation to PSE. Mutated genes were mainly involved in mRNA binding, transcription by RNA polymerase II, cytoskeleton and Endoplasmic Reticulum (ER).

4.2 Introduction

Next-generation sequencing (NGS) technologies are able to tackle the complexity of the human genome and provide a variety of tools to understand how human genome sequence variants underlie phenotypes and diseases (256).

Whole-genome sequencing provides the most comprehensive view of genomic information and associated biological implications (256). However, it remains expensive on a large scale and poses considerable challenges in the analysis and interpretation of the data. Whole-exome sequencing targets only a subset of the genome, namely the protein coding region, and also some non-protein coding elements (e.g. microRNA, long intergenic noncoding RNA, etc.) (257). Since the size of the genomic material used in genetic analysis (~1% of the human genome) is reduced, also the costs of genome analysis are lower comparatively to whole-genome sequencing. This permits sequencing higher number of samples, allowing for higher breadth and depth of genomic studies (256,258).

Therefore, whole-exome sequencing has become the method of choice for identification of causative mutations of human diseases, or to discover functional variation of traits and drug responses (257,259). Examples of diseases where exome sequencing has been

used to detect causative variants include Alzheimer's disease, amyotrophic lateral sclerosis, and a number of cancer predisposition mutations (147,257,260–264). Especially in cancer, exome sequencing proved to be an important tool to study clonal evolution and response to therapeutic drugs (265–268).

Although most mutations are deleterious, they also provide the raw material for adaptation and evolution. In fact, stressors can induce mutagenesis adaptive programs that fuel evolution of microbial pathogenesis and antibiotic-resistance, tumor progression and chemotherapy resistance (269,270).

Stress-induced genomic instability has been studied in a variety of organisms such as bacteria, yeast and human cancer cells. Alterations in the genome include small changes (of 1 nucleotide), deletions and insertions, gross chromosomal rearrangements and copy-number variations (269). For example, in yeast cells, proteotoxic stress caused by transient Hsp90 inhibition, potentiates adaptation through induction of aneuploidy. This form of stress-inducible mutation in eukaryotes was capable of fueling rapid phenotypic evolution and drug resistance (271). Also, in cancer cells, aneuploidy is a form of mutation that may contribute to oncogenesis, metastasis and the development of chemotherapy resistance (272,273).

Previous studies have reported alterations at the genomic level of cells exposed to PSE for several generations. In these experiments, PSE were produced by an engineered $tRNA_{CAG}^{Ser}$ that misincorporated serine (Ser) at leucine (Leu) sites encoded by the CUG codon. In yeast CUG codons are distributed over 89% of its protein coding genes and since Ser and Leu are chemically different from each other, these mistranslation events generated protein misfolding and proteotoxic stress (196). During evolution for ~250 generations, the fitness of mistranslating strains increased and large-scale chromosomal duplication and deletion events were observed. Indeed, 431 independent mutational events, namely SNPs and large genomic rearrangements were detected. Several functional units were repeatedly mutated in the 11 clones sequenced, such as rRNA maturation, transcription initiation and elongation and the ubiquitin-proteasome system. Large scale duplication events were enriched in genes involved in ribosomal biogenesis, rRNA processing, superoxide dismutase 1 (SOD1) gene and protease core complex assembly. Since these chromosomal rearrangements were observed at the same breakpoints in all of the sequenced clones, it is likely that such mutational events confer adaptive advantage to the carrying cells (196). Also, in *C. albicans* mistranslating strains evolved for ~100 generations, full genome resequencing identified various strain specific single nucleotide polymorphism (SNPs) and one SNP in the deneddylase (JAB1) gene in all

strains. This gene is part of a complex, which mediates degradation of a variety of cellular proteins and may play a role in tolerance and adaptation to PSE in *C. albicans* (274).

In the previous chapters, we saw that human cells adapt relatively well to the expression of mutant tRNAs and to overexpression of the Wt $tRNA^{Ser}$. We observed the accumulation of misfolded proteins and in some cases protein aggregates ($tRNA^{Ser}$(L)) that activated PQC mechanisms (Chapter 2). We were also able to identify transcriptional alterations in these cells that may explain survival and adaptation, beyond PQC mechanisms, such as alterations in metabolic pathways (glycolysis/gluconeogenesis) and signaling pathways (p53 and MAKP) (Chapter 3). In order to understand those phenotypes, we sequenced the whole-exome of our cell lines, Mock, $tRNA^{Ser}$(S), $tRNA^{Ser}$(A), $tRNA^{Ser}$(L) and $tRNA^{Ser}$(H) in the last point of the evolution experiment (P30). Our hypothesis was that mistranslating cells accumulated adaptive mutations to overcome the stress imposed by PSE.

With the application of high quality criteria, we were able to detect several alterations in mistranslating cell lines, but the majority was common to our Mock control cell line. A small number of alterations (SNPs and Insertions), detected in our cell lines, were different from the Mock cell line and occurred in genes that are mainly were involved in mRNA binding, transcription from RNA polymerase II, and in some cases, cytoskeleton and Endoplasmic Reticulum. Some of the alterations could be correlated to our previous phenotypes and probably confer advantages to these mistranslating cells in culture.

4.3 Results

4.3.1 Whole-exome sequencing of evolved mistranslating HEK293 cells

Our previous data suggested that our cell lines adapt relatively well to PSE probably by accumulating mutations in the genome. In order to clarify this hypothesis, we have sequenced the whole-exome of the Mock, $tRNA^{Ser}$(S), $tRNA^{Ser}$(A), $tRNA^{Ser}$(L) and $tRNA^{Ser}$(H) cell lines at the end of the evolution (P30).

For exome sequencing, genomic DNA was extracted from those cell lines, and the samples were sequenced using a HiSeq2000 instrument (Illumina) with 75bp paired-end reads. After sequencing, images were acquired using Illumina Genome Analyzer Sequencing Control Software and analyzed. FASTQ files with base calls, quality metrics and read calls were generated. Sequences were then aligned to the Human Genome Reference and single

nucleotide polymorphisms (SNPs) and small insertions or deletions (INDELs) were detected. Several quality criteria were applied during the analysis and will be referred in the text, but the final criteria to identify variants and annotated genes was: unique variants (defined by chromosome, coordinate and reference), variant quality (Q) ≥100, genotype quality (GQ) ≥90 and number of reads≥40. This was done to ensure identification of high quality variants.

4.3.2 Variants Lost, Shared and Gained

Table 4-1 shows the number of variants detected in each cell line, relative to the Human Reference Genome (HRG). The cell lines that had the highest number of mutations relative to the HRG were tRNA^{Ser}(A) and tRNA^{Ser}(L), with 2.030.408 and 2.066.886 variants, respectively. tRNA^{Ser}(H) had the lowest number of variants (1.307.461). After the application of the first quality criteria: unique variants, defined by chromosome, coordinate and reference, variant quality (Q) ≥20, genotype quality (GQ) >20, the number of variants decreased. The tRNA^{Ser}(A) and tRNA^{Ser}(L) cell lines had higher number of variants, 306.719 and 298.472, respectively, while tRNA^{Ser}(S) cell line had the lower value, 173.980 variants.

Table 4-1. Number of variants in each cell line (vs. *Homo sapiens*) with no criteria and with the application of the first set of quality criteria.

Cell line	Total number of variants	**Number of unique variants** (chromosome, coordinate, reference) **Q ≥20, GQ>20**
Mock	1.897.546	273.515
tRNA^{Ser}(S)	1.797.217	173.980
tRNA^{Ser}(A)	2.030.408	306.719
tRNA^{Ser}(L)	2.066.886	298.472
tRNA^{Ser}(H)	1.307.461	199.064

The number of variants (unique variants, Q≥20, GQ>20) was normalized to the sum of mapped reads of each condition, because differences in variants number can be due to sequencing efficiency. In fact, when the values of variants were normalized (Figure 4-1), we could see that Mock and tRNA^{Ser}(S) cells had the same number of variants, 2,09. tRNA^{Ser}(A) and tRNA^{Ser}(L) cells had also the same number of variants, 2,14. tRNA^{Ser}(H) had the lower value, 1,64 variants.

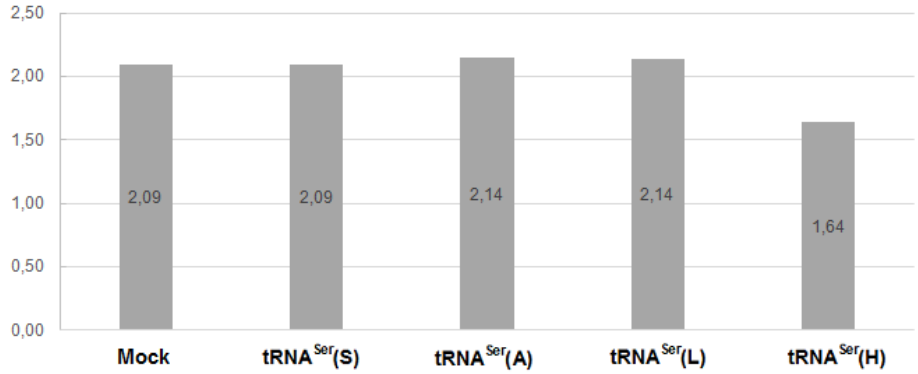


Figure 4-1. Number of variants (unique variants, Q≥20, GQ>20) in each cell line (vs. *Homo sapiens*) normalized to the total number of mapped reads in each cell line (x1000).

We then selected the variants (unique variants, Q≥20, GQ>20) with a coverage of reads equal or above 20 and compared each cell line with our control, the Mock cell line. Variants identified by this comparison were categorized in:

Lost variants (Lost) – are present in Mock, but absent in tRNA^{Ser}(S) or tRNA^{Ser}(A) or tRNA^{Ser}(L) or tRNA^{Ser}(H);

Gained variants (Gained) – are absent in Mock, but present in tRNA^{Ser}(S) or tRNA^{Ser}(A) or tRNA^{Ser}(L) or tRNA^{Ser}(H);

Shared variants (Shared) – present in Mock and tRNA^{Ser}(S) or tRNA^{Ser}(A) or tRNA^{Ser}(L) or tRNA^{Ser}(H).

The following patterns were observed: in tRNA^{Ser}(S) and tRNA^{Ser}(H) cells there was higher number of Lost variants, 164.410 and 170.420 variants, respectively; while in tRNA^{Ser}(A) and tRNA^{Ser}(L) cells the highest number of variants were categorized as Gained, 188.428 and 181.700 variants, respectively (Figure 4-2 A). Normalization of these values resulted in the same pattern (Figure 4-2 B). Shared variants were similar in the four cell lines (Figure 4-2 A). Normalization of the values showed that the tRNA^{Ser}(S) cell line had higher number of Shared variants, 1,31 (Figure 4-2 B).

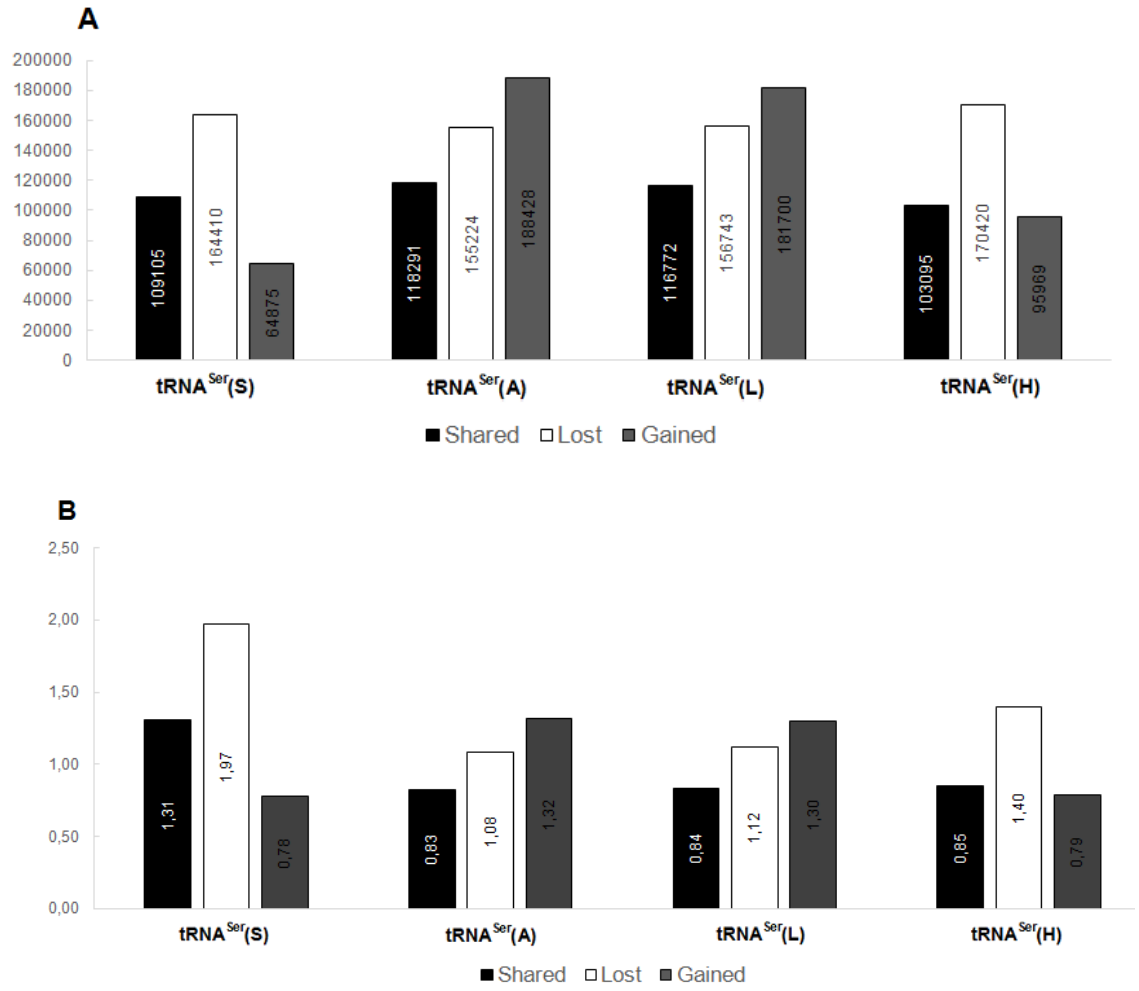


Figure 4-2. Number of variants (unique variants, Q≥20, GQ>20, number of reads≥20) in each cell line vs. Mock, categorized as "Shared", "Lost" and "Gained". A – Total number of variants; **B –** Number of variants normalized to the total number of mapped reads for each cell line (x1000).

In order to increase the sensitivity of variants detection, we analyzed the variants Shared, Lost and Gained with new criteria. According to previous results of Meynert *et al.*, 40x mean on target depth is required to obtain 95% overall estimated sensitivity for heterozygous SNPs (275). Using this value as reference (Table 4-2), we compared the number of variants Shared, Lost and Gained (relative to Mock), using a coverage of at least 20 reads, and using a coverage of at least 40 reads (non-normalized values). The number of variants decreased when we analyzed variants with coverage of a minimum of 40 reads, relatively to the analysis carried out with at least 20 reads. The numbers of Lost and Gained variants were similar in each cell line. The number of Shared variants in all cell lines was higher than Lost or Gained variants.

Table 4-2. Number of variants Shared, Lost or Gained relative to Mock, with a minimum of 20 or 40 reads (non-normalized values).

Mock vs.	**Shared**		**Lost**		**Gained**	
	#unique variants **Q ≥20, GQ>20, reads≥20**	#unique variants, **Q ≥20, GQ>20, reads≥40**	#unique variants **Q ≥20, GQ>20, reads≥20**	#unique variants, **Q ≥20, GQ>20, reads≥40**	#unique variants **Q ≥20, GQ>20, reads≥20**	#unique variants, **Q ≥20, GQ>20, reads≥40**
tRNA^{Ser}(S)	109.105	31.884	164.410	2.219	64.875	1.977
tRNA^{Ser}(A)	118.291	50.504	155.224	2.097	188.428	2.317
tRNA^{Ser}(L)	116.772	49.187	156.743	2.133	181.700	2.261
tRNA^{Ser}(H)	103.095	43.649	170.420	3.462	95.969	2.124

The values of variants obtained with a minimum of 40 reads were normalized to the total number of mapped reads for each cell line. The number of variants Shared was much higher than the number of Lost or Gained variants in all cells lines (Figure 4-3), suggesting that the majority of the alterations accumulated during evolution are not related to PSE, since these alterations were also present in our control cell line (Mock). The number of Lost and Gained variants decreased with the application of this new criterion, meaning that Lost and Gained variants previously identified had lower coverage and higher level of associated error. In tRNA^{Ser}(S) and tRNA^{Ser}(H) cells there was higher number of Lost variants (0,27 and 0,28, respectively) relatively to the number of Gained variants (tRNA^{Ser}(S): 0,24 and tRNA^{Ser}(H): 0,17). In tRNA^{Ser}(A) and tRNA^{Ser}(L) cells the pattern observed was the opposite, the number of Gained variants (0,16 for both cases) was higher than Lost variants (0,15 for both cases). tRNA^{Ser}(S) and tRNA^{Ser}(H) cells showed higher values of Lost or Gained variants relatively to tRNA^{Ser}(A) and tRNA^{Ser}(L) cells (Figure 4-3 B).

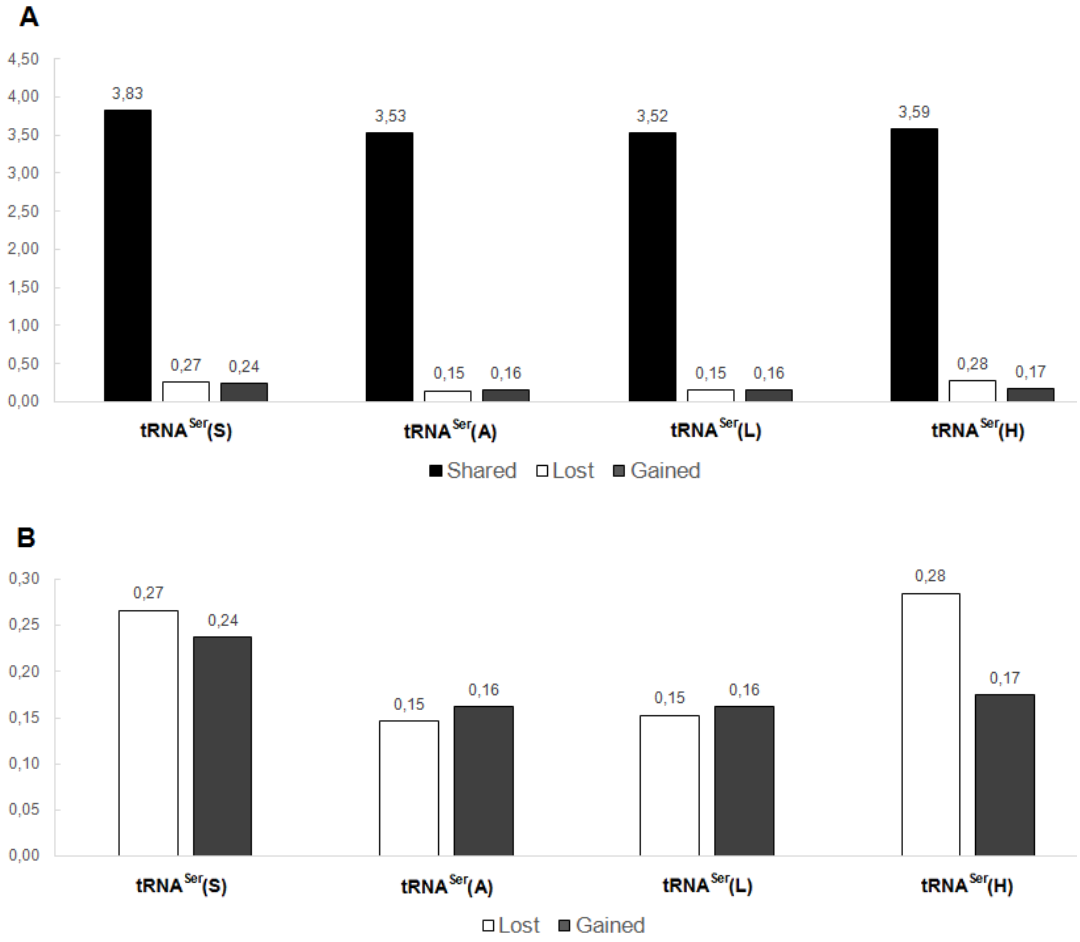


Figure 4-3. Number of variants (unique variants, Q≥20, GQ>20, number of reads≥40) in each cell line relative to Mock, normalized to the total number of mapped reads for each cell line (x10000). A – Variants categorized as Shared, Lost or Gained for each cell line. **B –** Highlight of Lost and Gained variant categories.

Since after the application of the second criterion we still had high numbers of variants we decided to apply a final set of criteria to select variants and reduce scoring error levels. We used high quality variants, unique variants, Q≥100, GQ≥90, with a number of reads≥40. As expected the application of the new criteria reduced the number of variants. The number of Lost variants in tRNA^{Ser}(A) and tRNA^{Ser}(L) cells (559 and 558 variants, respectively) was very similar to the number of Gained variants (601 and 557 variants, respectively) (Table 4-3). tRNA^{Ser}(S) and tRNA^{Ser}(H) cells had a different pattern since the number of Lost variants (1494 and 1360 variants, respectively) was higher than the number of Gained variants (557 and 602 variants, respectively) (Table 4-3).

Table 4-3. Number of high quality (HQ) variants Lost or Gained in comparison with Mock, with a minimum of 40 reads or with the minimum of 40 reads with the application of the new quality criteria (Q≥100, GQ≥90) (non-normalized values).

Mock vs.	**Lost**		**Gained**	
	#unique variants, **Q ≥20, GQ>20, reads≥40**	#unique variants, **Q ≥100, GQ>90, reads≥40**	#unique variants, **Q ≥20, GQ>20, reads≥40**	#unique variants, **Q ≥100, GQ>90, reads≥40**
tRNA^{Ser}(S)	2.219	1.494	1.977	557
tRNA^{Ser}(A)	2.097	559	2.317	601
tRNA^{Ser}(L)	2.133	558	2.261	557
tRNA^{Ser}(H)	3.462	1.360	2.124	602

Table 4-3 values of high quality variants (unique variants, Q≥100, GQ≥90, with number of reads≥40) were normalized to the total number of mapped reads for each cell line. Figure 4-4 shows the final distribution of Lost and Gained variants in our cell lines. The patterns already identified were similar for our cell lines. In tRNA^{Ser}(A) and tRNA^{Ser}(L) cells the number of Lost variants (0,039 and 0,040 variants, respectively) was similar to the number of Gained variants (0,042 and 0,040 variants, respectively). In tRNA^{Ser}(S) and tRNA^{Ser}(H) cells the number of Lost variants (0,179 and 0,112 variants, respectively) was higher than the number of Gained variants (0,067 and 0,049 variants, respectively).

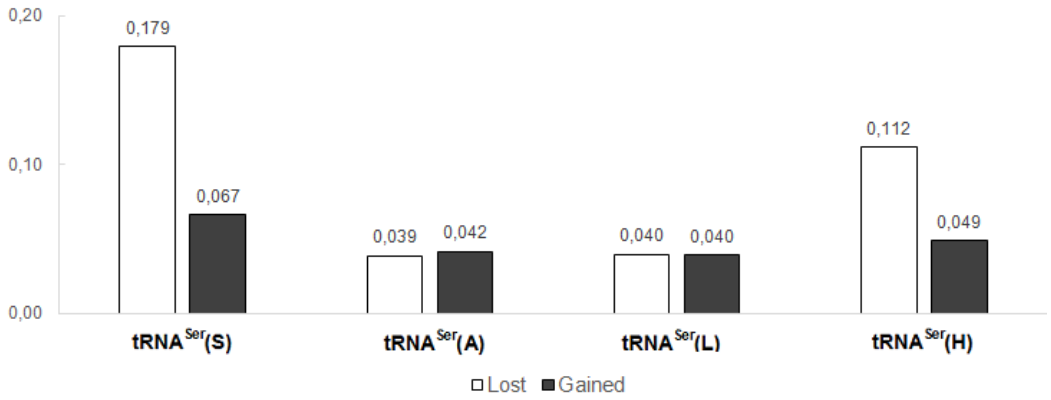


Figure 4-4. Number of Lost and Gained HQ variants (unique variants, Q≥100, GQ≥90, with a number of reads≥40) in each cell line comparatively to Mock, normalized to the total number of mapped reads for each cell line (x10000).

To better understand the relevance of the previous patterns we have determined the ratio of Gained/Lost variants relative to Mock in each cell line (Figure 4-5). We can observe that tRNA^{Ser}(S) had the lowest value (0,37), followed by tRNA^{Ser}(H) cells (0,44), meaning that these two cell lines, have a tendency to lose variants that were present in Mock. tRNA^{Ser}(L) ratio of Gained/Lost variants was 1, while tRNA^{Ser}(A) ratio was 1,08 indicating that in these cell lines the number of Lost and Gained variants was identical or very similar. Until the introduction of these new criteria **(**Q≥100, GQ≥90) we had higher number of Gained variants relative to Lost variants in tRNA^{Ser}(A) and tRNA^{Ser}(L) cell lines.

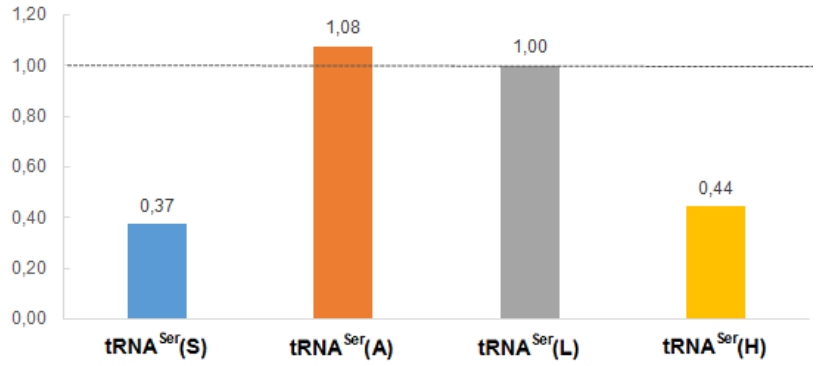


Figure 4-5. Ratio of Gained/Lost HQ variants (unique variants, Q≥100, GQ≥90, with a number of reads≥40) in each cell line comparatively to Mock.

We then analyzed the distribution of Lost and Gained variants per chromosome. The ratio of Ganeid/Lost variants was in most of the cases lower or equal to 1, but there were some exceptions (Figure 4-6). In tRNA^{Ser}(S) and tRNA^{Ser}(H) cell lines there was a variant loss trend, relative to the Mock cell line, in all chromosomes. In the tRNA^{Ser}(S) cell line in chromosomes 1, 12, 13 and 19 the ratio of Gained/Lost variants was higher than 1 and in the chromosome 19, the ratio Gained/Lost was around 2. This increase in gained variants was observed throughout chromosome 19. The genes that accumulated variants in this chromosome did not belong to any relevant functional Gene Ontology (GO) category. However, we can highlight *USE1* gene that contained an upstream variant. This gene encodes a SNARE protein that may be involved in targeting and fusion of Golgi-derived retrograde transport vesicles with the endoplasmic reticulum (ER). In chromosome 19, *ZNF480*, *ZNF805*, *ZNF766* genes that encode zinc finger proteins and are involved in transcriptional regulation, also accumulated mutations.

In tRNA^{Ser}(A) cell line there was accumulation of Gained variants in chromosome 6. The ratio Gained/Lost variants was around 4. The genes in chromosome 6 that accumulated

mutations were annotated in the GO category of "extracellular matrix structural constituent" (*p*-value=2.6E-2). Two genes that encode collagen molecules were annotated in this category, *COL12A1* and *COL19A1*. Other genes also accumulated alterations, but were not annotated in any GO category, namely:

- *TBP*, which encodes a transcription factor that functions at the core of the DNA-binding multiprotein factor TFIID, playing a role in the activation of eukaryotic genes transcribed by RNA polymerase II. This gene accumulated a frameshit variant with predicted high impact in the protein structure;
- *HSF2*, which encodes a DNA binding protein that specifically binds to heat shock promoter elements and activates transcription. In this case an intronic mutation was detected;
- *POLH*, which encodes a DNA polymerase involved in DNA repair, accumulated two intron variants;
- *MRPL18*, which encodes a mitochondrial ribosomal protein, accumulated an intron variant;
- *ECT2L*, which encodes the protein epithelial cell transformin sequence 2 oncogene-like, accumulated an intron variant. This gene was also downregulated in our microarray data set, namely from P1 to P30 in tRNA^{Ser}(A) cell line.

Finally, in tRNA^{Ser}(L) cells, there was a strong decrease in Gained variants in chromosome 13, relative to the other cell lines.

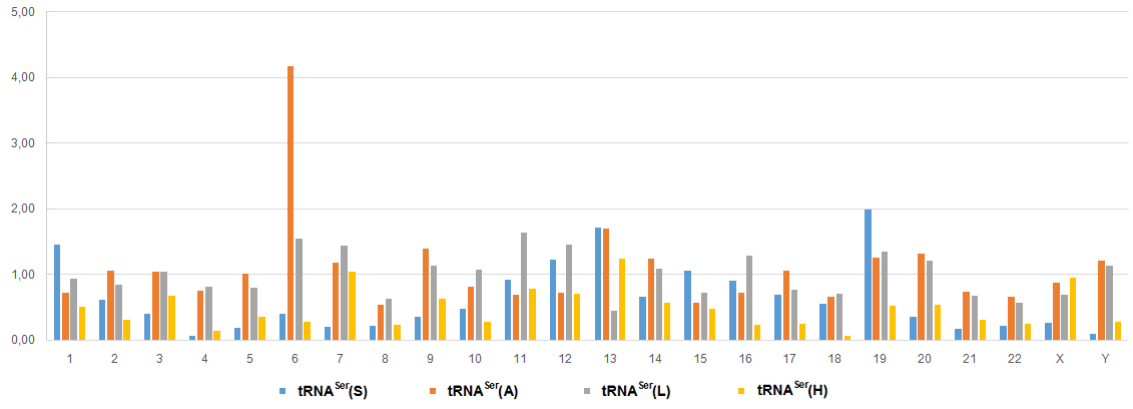


Figure 4-6. Ratio of Gained/Lost HQ variants (unique variants, Q≥100, GQ≥90, number of reads≥40) per chromosome, in each cell line, comparatively to Mock.

4.3.3 SNPs, INDELs and variants quality criteria

We identified next the type of alterations (SNPs or INDELs) that were present in the exome of our cell lines in P30. Whole-exome sequencing analysis allows identification of SNPs, namely Transitions and Transversions, and small Insertions and Deletions (INDELs). Transitions are interchanges of two-ring purines (A↔G) or of one-ring pyrimidines (C↔T) (Figure 4-7). Transversions are interchanges of purine for pyrimidine bases, which involve exchange of one-ring and two-ring structures (Figure 4-7).

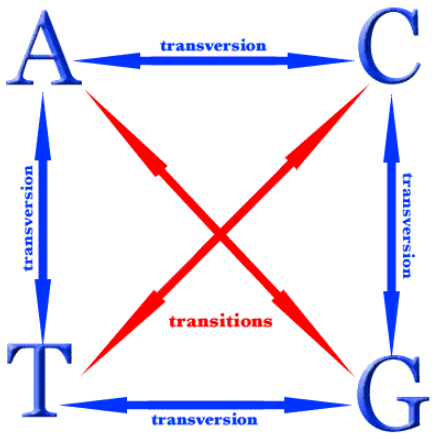


Figure 4-7. Representative scheme of Transitions and Transversions (adapted from (276)).

Table 4-4 shows the number of variants; Transitions or Transversions for each cell line, using the different quality criteria already described. The number of Transitions was higher than the number of Transversions, for all cell lines and independently of the quality criteria used. This is not surprising since, in humans transition mutations are at least twice as likely as transversion mutations, resulting in many SNPs that are A/G or C/T (277). $tRNA^{Ser}$(A) and $tRNA^{Ser}$(L) cell lines had higher number of Transitions and Transversions, relative to $tRNA^{Ser}$(S) and $tRNA^{Ser}$(H) cells. $tRNA^{Ser}$(S) cell line had the lowest values of Transitions and Transversions.

Table 4-4. Number of variants (Transitions or Transversions) for each cell line, using different quality criteria.

# Variants (not normalized)					
Type	**Cell Line**	**No Criteria**	**#unique variants Q≥20, GQ≥20, Number of reads≥20**	**#unique variants Q≥20, GQ≥20, Number of reads≥40**	**#unique variants Q≥100, GQ≥90, Number of reads≥40**
TRANSITIONs	**$tRNA^{Ser}$(S)**	173.438	44.010	23.293	20.245
	$tRNA^{Ser}$(A)	223.954	52.490	35.392	31.630
	$tRNA^{Ser}$(L)	218.596	51.396	34.259	30.537
	$tRNA^{Ser}$(H)	178.217	48.739	31.594	28.225
TRANSVERSIONs	**$tRNA^{Ser}$(S)**	145.552	21.000	10.377	8.211
	$tRNA^{Ser}$(A)	213.875	25.881	15.870	13.019
	$tRNA^{Ser}$(L)	212.901	25.589	15.626	12.885
	$tRNA^{Ser}$(H)	169.791	24.152	14.051	11.585

We then, normalized the values of Transitions and Transversions to the number of reads for each cell line, using the last quality criteria applied (unique variants, Q≥100, GC≥90 and number of reads≥40). With the normalization of the values we could observe that $tRNA^{Ser}$(S) had slightly more Transitions than the other cell lines (0,81) (Figure 4-8). This was probably due to a technical problem; since the sequencing of this cell line was less efficient than the other cell lines ($tRNA^{Ser}$(S) had around 83 million reads while the other cell lines had above 120 million reads). When we normalized $tRNA^{Ser}$(S) variants to the total number of reads there was an apparent higher number of variants. The number of Transversions was very similar in all mistranslating cells, the average of Transversions was around 0.3. Once again, it was clear that the number of Transitions was higher than the number of Transversions, in all cell lines (Figure 4-8).

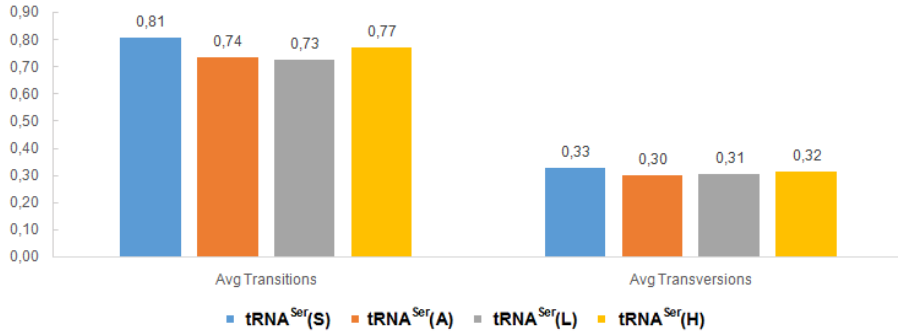


Figure 4-8. Number of Transitions and Transversions (unique variants, Q≥100, GQ≥90, number of reads≥40), for each cell line, relative to Mock, normalized to the total number of mapped reads for each cell line (x10000).

Next, the number of variants; Transitions and Transversions, and Insertions or Deletions was categorized in Gained, Shared or Lost, relative to our control, the Mock cell line (Figure 4-9). When we introduce the more stringent criteria the number of Gained and Lost variants was very low in all cell lines. This happened for all variant types, Transitions, Transversions, Insertions and Deletions. The decrease in Gained and Lost variants with the introduction of quality criteria was due to variants of low quality, the low genome quality and also the decreased number of reads (low coverage). In all cell lines, genomic alterations were mainly Shared with the control, and only a very low number were specific of each cell line.

A

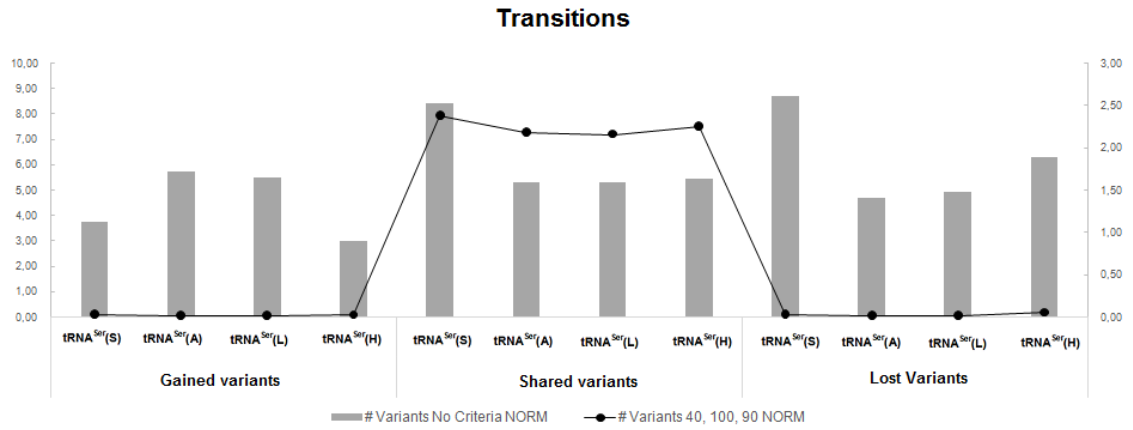
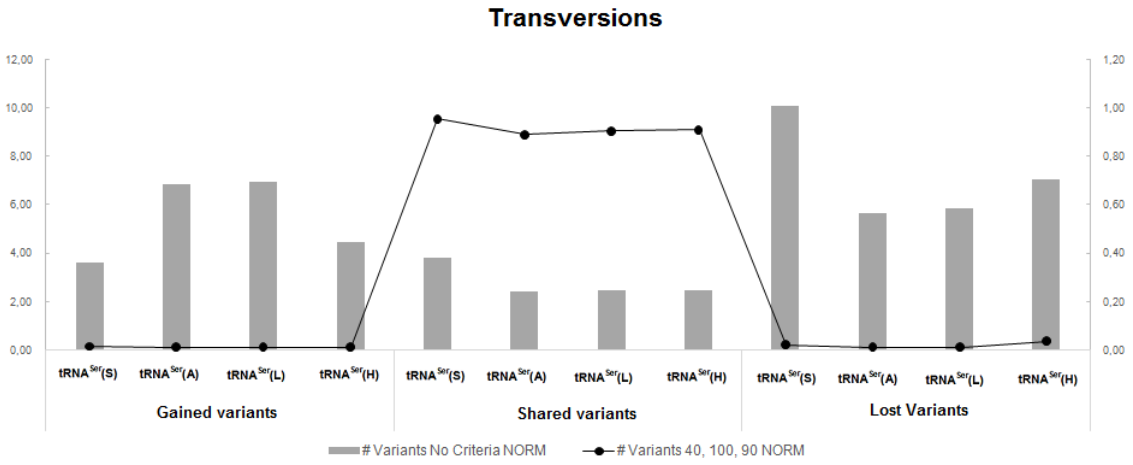
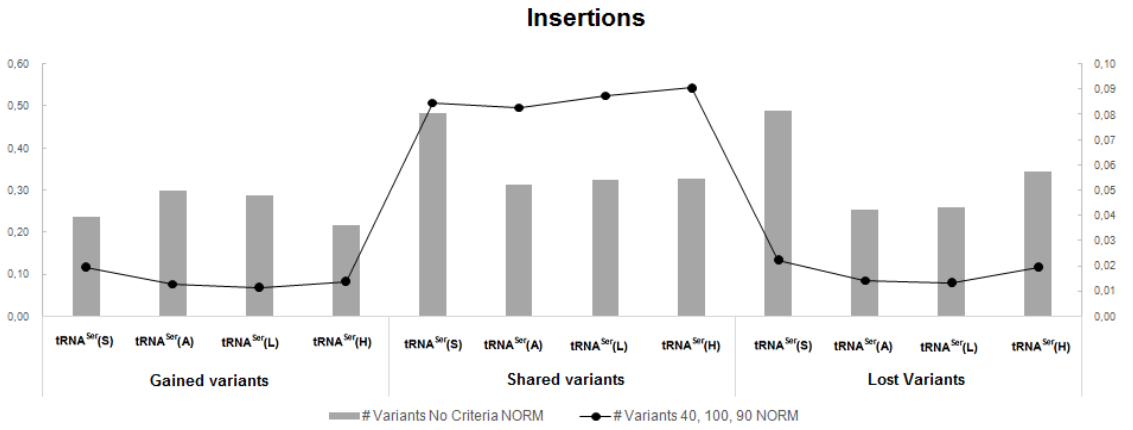


Figure 4-9.►

►**B**



C



D

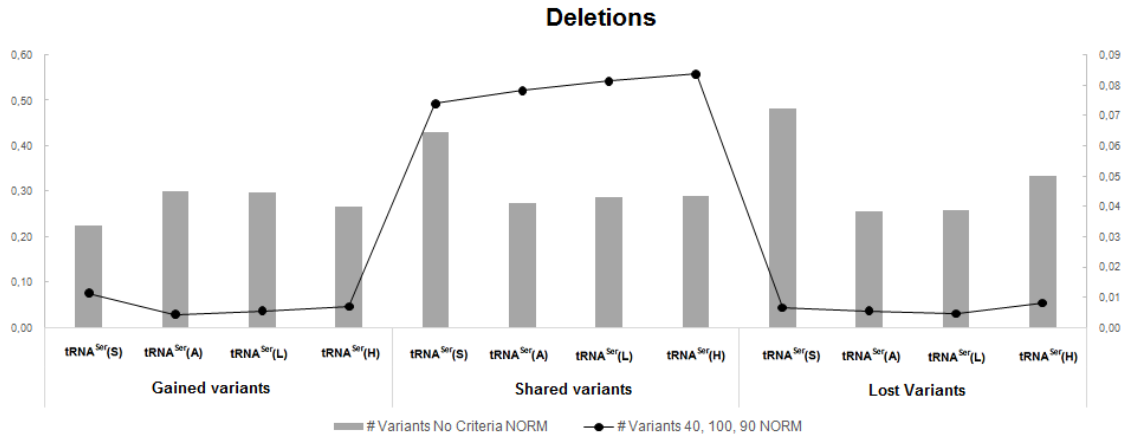


Figure 4-9. Number of variants with no criteria (columns) or with application of the maximum quality criteria (unique variants, Q≥100, GQ≥90, number of reads≥40) (dots) that were Gained, Shared or Lost for each cell line relative to Mock, normalized to the total number of mapped reads for each cell line (x10000). A – Transitions; **B –** Transversions; **C –** Insertions; **D –** Deletions

4.3.3 SNPs and Insertions annotated to genes

SNPs (Transitions and Transversions) and some Insertions with high quality (unique variants, Q≥100, GQ≥90, number of reads≥40) were annotated to genes in order to identify biological functions that can be affected and explain adaptation to PSE. Overall, 10.198 high quality INDELs were detected, however only 1.438 were annotated to genes. This is due to a technical problem, since paired-end reads and depth coverage approaches frequently miss small INDELs, and are unable to predict the breakpoints accurately (278). For this reason, we could not identify the genes with Deletions and the majority of Insertions. For each cell line, we identified genes that accumulated Transitions, Transversions and Insertions, relative to Mock, and also cross checked these data with our previous microarray data, whenever possible.

The number of annotated and identified (DAVID Bioinformatic Resources 6.7) genes in our cell lines that have Transitions, Transversions or Insertions in their sequence is presented in Table 4-5.

Table 4-5. Number of annotated and identified (DAVID Bioinformatic Resources 6.7) genes in our cell lines that have Transitions, Transversions or Insertions in their sequence.

Mistranslating cell lines	#Genes with Transitions	#Genes with Transversions	#Genes with Insertions
tRNA^{Ser}(S)	17	12	24
tRNA^{Ser}(A)	20	22	42
tRNA^{Ser}(L)	13	20	32
tRNA^{Ser}(H)	14	23	38

In tRNA^{Ser}(S) cell line, 17 genes with Transitions were annotated and identified, namely *ROS1* (containing an intron variant) which codes a cell growth or differentiation factor receptor and *CDK19* (containing an intron variant), which encodes a component of the Mediator co-activator complex, required for transcriptional activation by DNA binding transcription factors that regulate RNA polymerase II. Two genes that encode proteins that

bind to mRNA poly(A) tails, *PABPC1* (containing a missense variant) and *PABPC1P5* (containing an upstream gene variant) also had Transitions in tRNA^{Ser}(S) cell line (Table 4-6). One of the 20 genes that accumulated Transitions in tRNA^{Ser}(A) cells was *HSF2* (containing an intron variant). This gene encodes the heat shock transcription factor 2 that induces the transcription of genes in response to alterations in the protein-degradative machinery. In the tRNA^{Ser}(L) cell line, genes with Transitions were annotated to the GO category "transcription from RNA polymerase II promoter". One of them is *HSF2* (containing an intron variant). The tRNA^{Ser}(H) cell line had 14 genes with Transitions, namely *DNASE1L1* (containing a downstream gene variant), which encodes a deoxyribonuclease protein very similar to DNaseI (Table 4-6).

The genes that accumulated Transitions in our mistranslating cell lines were mainly involved in RNA polymerase II transcription, indicating that transcription plays an important role in the response to proteotoxic stress in the evolved cell lines. *HSF2* presented alterations in two of our cell lines, tRNA^{Ser}(A) and tRNA^{Ser}(L). HSF2 is ubiquitously expressed in mammanlian tissues and is activated by the downregulation of the ubiquitin-proteasome pathway (UPS). When activated, the response is similar to the activation of HSF1 and leads to upregulation of molecular chaperones, likely due to accumulation of misfolded proteins (279). Our previous data from Chapter 2 indicated that the levels of HSP27, HSP60 and HSP70 chaperones seem to be downregulated in our cells. This may indicate that sustained activation through time of molecular chaperones is not beneficial and alterations in *HSF2* gene could have an effect in this process.

Table 4-6. Examples of genes with Transitions in mistranslating cell lines.

Mistranslating cell lines	Examples of genes with Transitions
tRNA^{Ser}(S)	***ROS1*** (ROS proto-oncogene1, receptor tyrosine kinase) ***CDK19*** (cyclin dependent kinase 19) ***PABPC1*** (poly(A) binding protein cytoplasmic 1) ***PABPC1P5*** (poly(A) binding protein cytoplasmic 1 pseudogene 5)
tRNA^{Ser}(A)	***HSF2*** (heat shock transcription factor 2)
tRNA^{Ser}(L)	***HSF2*** (heat shock transcription factor 2)
tRNA^{Ser}(H)	***DNASE1L 1*** (desoxyribonuclease 1 like 1)

Regarding the genes with Tranversions, 12 were identified in tRNA^{Ser}(S) cell line, in genes such as *PABPC3* (containing a missense variant) and *ZNF75D* (Table 4-7). *PABPC3* gene encodes a protein that binds to the poly(A) tail of mRNA and may be involved in cytoplasmic regulatory processes of mRNA metabolism, while the *ZNF75D* gene encodes a zinc finger protein involved in transcriptional regulation. tRNA^{Ser}(A) cell line had 22 genes with Transversions. Within these there are zinc finger proteins (*ZNF676*, *ZNF75D* and *ZC3H12D*). The *ZC3H12D* gene product may function as an RNase, regulating levels of target RNAs. Also, the genes *TBX18* (containing an intron variant) and *ZYG11B* (containing a missense variant) had Transversions in tRNA^{Ser}(A) cell line. *TBX18* encodes a transcriptional regulator involved in developmental process, and *ZYG11B* encodes a protein that is part of an E3 ubiquitin ligase complex, probably acting as a target recruitment subunit (Table 4-7).

In tRNA^{Ser}(L) cell line, the genes that accumulated Transversions were annotated in GO categories such as "mRNA binding" and "transcription regulation". The "mRNA binding" category includes the *KHDRBS2* gene (containing an intron variant), which encodes an RNA-binding protein that plays a role in the regulation of alternative splicing and influences mRNA splice selection and exon inclusion. In the second category, "transcription regulation" we found two genes that encode zinc finger proteins, both involved in transcriptional regulation, *ZNF75D* and *ZNF441*. Also, *MED23* gene was found in this category. The protein encoded by *MED23* gene is required for transcriptional activation subsequent to the pre-initiation complex and is a component of the Mediator complex that functions as a coactivator involved in the regulated transcription of nearly all RNA polymerase II-dependent genes (Table 4-7). In tRNA^{Ser}(H) cell line 23 genes had Transversions and some of them were annotated in the GO category "mRNA binding", such as *KHDRBS2*. Two genes that encode zinc finger proteins involved in transcriptional regulation also had Transversions, *ZNF75D* and *ZNF774* (Table 4-7).

The accumulation of Transversions occurred in genes involved mainly in RNA binding and transcription. In tRNA^{Ser}(S) alterations in the potential RNase gene, *ZC3H12D*, could be related with the alterations of tRNA^{Ser} expression. The gene *ZYG11B*, encodes a substrate-recognition subunit of a CUL-2-based complex (E3 ubiquitin ligase complex) that regulates cell division and transcription (280). Alterations in this gene seem to be important in tRNA^{Ser}(A), since its expression is upregulated in our microarray data (P15-P30). Therefore, it may be important for cell division and proliferation of tRNA^{Ser}(A) cells. In tRNA^{Ser}(L) cells, alterations that occurred at the genomic level in *MED23* gene may explain the up-regulation in P30.

Table 4-7. Example of genes with Transversions in mistranslating cell lines.

Mistranslating cell lines	Examples of genes with Transversions
tRNA^{Ser}(S)	***PABPC3*** (poly(A) binding protein cytoplasmic 3) ***ZNF75D*** (zinc finger protein 75D)
tRNA^{Ser}(A)	***ZNF676*** (zinc finger protein 676) ***ZNF75D*** (zinc finger protein 75D) ***ZC3H12D*** (zinc finger CCCH-type containing 12D) ***TBX18*** (T-box 18) ***ZYG11B*** (zyg-11 family member B, cell cycle)
tRNA^{Ser}(L)	***KHDRBS2*** (KH RNA binding domain containing, signal transduction associated 2) ***ZNF75D*** (zinc finger protein 75D) ***ZNF441*** (zinc finger protein 441) ***MED23*** (mediator complex subunit 23)
tRNA^{Ser}(H)	***KHDRBS2*** (KH RNA binding domain containing, signal transduction associated 2) ***ZNF75D*** (zinc finger protein 75D) ***ZNF774*** (zinc finger protein 774)

Regarding genes that contained Insertions in tRNA^{Ser}(S) cells, 24 were identified and annotated in GO categories, including "cytoskeletal part" and "microtubule cytoskeleton". Other genes not annotated in these categories were for example, *EIF3A* (containing two downstream gene variants) (Table 4-8). This gene encodes the eukaryotic translation initiator factor 3 subunit A. In tRNA^{Ser}(A) cell line genes with Insertions were annotated in categories such as "transcription cofactor activity" and "alternative splicing". The first category included *TCERG1* (containing a downstream gene variant) and *MYSM1* (containing an intron variant) genes (Table 4-8). *TCERG1* gene encodes a transcription factor that binds to RNA polymerase II and inhibits elongation of transcripts from target promoters. The *MYSM1* gene encodes a metalloprotease that specifically deubiquitinates monoubiquitinated histone H2A, a specific tag for epigenetic transcriptional repression, thereby acting as a coactivator. In addition to these, *PSMD9* (containing a downstream gene variant), encodes a protein that acts as a regulatory subunit of the 26S proteasome. Also in tRNA^{Ser}(L) cell line, genes that accumulated

Insertions were mainly involved in alternative splicing, such as *MYSM1* and cytoskeletal protein binding. The genes *TRNT1, DNAJC14, POLR3G PSMD9* and *CIDEB* also had Insertions in tRNA^{Ser}(L) cells (Table 4-8). The *TRNT1* gene encodes a protein that adds and repairs the conserved 3´-CCa sequence of tRNAs; the *DNAJC14* gene encodes an Hsp70 co-chaperone that regulates the export of target proteins from the ER to the cell surface; *POLR3G* gene encodes a specific peripheric component of RNA polymerase III, which synthesizes small RNAs and tRNAs and *CIDEB* gene encodes a protein that activates apoptosis. Finally regarding tRNA^{Ser}(H) cell line, 38 genes had Insertions. Some were annotated in the GO category "endoplasmic reticulum", such as *DNAJC14* (Table 4-8).

The number of genes with Insertions was higher than the number of genes with Transitions or Transvertions. These genes were mainly involved in transcription processes, alternative splicing and cytoskeleton. It is interesting to note that tRNA^{Ser}(A) cells accumulated alterations in genes that belong to the ubiquitin-proteasome system, which may indicate that these mutations can be important to the activity of this system. A component of polymerase III was mutated in tRNA^{Ser}(L) cells, *POLR3G*, which may influence tRNA synthesis in this cell line. We wonder if this mutation is involved in the dowregulation of the mutant tRNA. Also, the mutation detected in *CIDEB* gene in tRNA^{Ser}(L) cells may explain its uperegulation in our microarray data set (P1-P30 and M30-L30). The respective protein is involved in apoptosis, but we did not observe significant alterations in cell death in tRNA^{Ser}(L) cells in P30. Genes that belong to the "endoplasmic reticulum" category accumulated insertions in tRNA^{Ser}(H) cells, namely *DNAJC14* which may prevent the the activation of Hsp70 and the export of misfolded proteins to cells surface.

Table 4-8. Example of genes with Insertions in mistranslating cell lines.

Mistranslating cell lines	Examples of genes with Insertions
tRNA^{Ser}(S)	***EIF3A*** (eukaryotic translation initiation factor 3 subunit A)
tRNA^{Ser}(A)	***TCERG1*** (transcription elongation regulator 1) ***MYSM1*** (Myb like, SWIRM and MPN domains 1) ***PSMD9*** (proteasome 26S subunit, non-ATPase 9)
tRNA^{Ser}(L)	***MYSM1*** (Myb like, SWIRM and MPN domains 1) ***TRNT1*** (tRNA nucleotidyl transferase 1) ***DNAJC14*** (DnaJ heat shock protein family (Hsp40) member C14) ***POLR3G*** (RNA polymerase III subunit G) ***PSMD9*** (proteasome 26S subunit, non-ATPase 9) ***CIDEB*** (cell death-inducing DFFA-like effector b)
tRNA^{Ser}(H)	***DNAJC14*** (DnaJ heat shock protein family (Hsp40) member C14)

4.4 Discussion

Whole-exome sequencing analysis was performed to understand the genomic alterations in our cell lines in P30. Our hypothesis was that these cells accumulated adaptive mutations to survive or counteract the negative effects of PSE. In addition to adaptive mutations, our results show alterations in some genes that could be related to the phenotypes that we have observed in previous chapters, for example mutations in *DNAJC14* gene among others. Genes involved in RNA binding and Transcription from RNA polymerase II showed higher level of alterations in our cell lines, indicating that genomic alterations were mainly related to the transcription process.

In *C. albicans*, amino acid misincorporations lead to alterations in genome structure, detected by whole-genome sequencing. *C. albicans* has the ability to incorporate Ser (~97%) and Leu (~3%) at CUG codons. Engineered *C. albicans* strains that misincorporate increasing levels of Leu at protein CUG sites showed high number of genotype changes at polymorphic sites, and 80% of Leu misincorporation results in complete loss of heterozygosity in a large region of chromosome V (75). In the affected region of chromosome V most of the ORFs had no known function, but the known genes were involved in the stress response, antifungal drug

resistance and pathogenesis, which is in line with the acquired resistance to antifungals (75). Whole-exome sequencing is a good tool to study alterations in the coding part of the genome, but unlike whole-genome sequencing, it is very limited to unveil genomic structural variants (281).

The identification of Indels represents a challenge in the analysis of NGS data. In our study, we could identify and annotate some Insertions, but no Deletions. Due to the short sequence reads produced by NGS, Insertions and Deletions involving large DNA segments are difficult to accurately align. For example, when the size of the Indel is greater than half of the sequence read length, the NGS reads containing the mutant sequence may fail to align properly (278,282).

The majority of variants detected in mistranslating cell lines were also present in Mock. This could be due to the low quality of Gained and Lost variants, as described in the results. One of the major complications with NGS data is the identification of high quality variants and for this reason during the bioinformatic analysis the quality criteria was increased (275,283).

In $tRNA^{Ser}$(A) and $tRNA^{Ser}$(L) cell lines the *HSF2* gene had the same single nucleotide variation (SNV) in its intron, suggesting that this gene may be relevant to adaptation to PSE. Heat shock factors (HSFs) are transcriptional regulators of genes encoding molecular chaperones and several other stress proteins and for this reason are essential for organisms survival (284). HSF2 depends on HSF1 for its stress-related functions, it is recruited to HSP gene promoters only in the presence of HSF1 and can positively or negatively modulate HSPs expression (284). Previous studies revealed that HSF2 is able to modulate HSP70, HSP27, HSP40 and HSP90 expression in stress situations, but is also important for normal levels of constitutive expression of HSE-containing genes (285–287). Protein chaperones levels decreased in P30 in our cell lines (Chapter 2), and in $tRNA^{Ser}$(A) and $tRNA^{Ser}$(L) this decrease may be correlated with the mutation found in *HSF2* gene. Although HSF1 is the major regulator of Hsps expression and no impact in the protein structure can be atributed to the variant found in *HSF2* gene, we can not exclude alteration in expression of HSF2 mediated by this mutation, but further studies are needed to clarify the importance of this alteration in these mistranslating cell lines. Particularly in the case of Hsp70, which is downregulated in $tRNA^{Ser}$(A) and $tRNA^{Ser}$(L) cells, it is well decribed that HSF2 is able to form heterotrimers with HSF1, bind to *hsp70.1* promoter and increase Hsp70 expression in response to heat shock and during proteasome inhibition in mouse embryonic fibroblasts (MEFs) (285). Also *DNAJC14* gene accumulated a 3' UTR variant in $tRNA^{Ser}$(L) and $tRNA^{Ser}$(H) cells. The J protein encoded by this gene is a co-chaperone of Hsp70 and has a J domain that interacts and

stimulates Hsp70 ATPase activity (91). In addition, DNAJC14 is bound to ER and modulates transport of client proteins from this organelle to the plasma membrane (288,289). During ER stress, DANJC14 can bind to client misfolded proteins and in cooperation with Hsc70 targeting them to unconventional secretion, a mechanism suggested to relieve the protein burden in the ER (289). Although this may not be an adaptive mutation and the impact of the downstream variant found in *DNAJC14* gene, is not known, in $tRNA^{Ser}$(L) cell line we do not exclude that this alteration may prevent ER stress alleviation leading to accumulation of protein aggregates (Chapter 2). Previous exome sequencing analyses have identified several mutions in *DNAJC* genes (*DNAJC13*, *DNAJC6* and *DNAJC5*) that may be implicated in protein aggregation diseases such as Parkinson's disease (290).

The $tRNA^{Ser}$(A) cell line presented a ratio of Gained mutations higher than the other cell lines, especially in chromosome 6. This cell line adapts rather well to PSE, through activation of the ubiquitin-proteasome system, increasing protein synthesis rate and also increasing the expression of genes involved in glycolysis (Chapters 2 and 3). One of the genes that accumulated mutations was *POLH* that encodes the DNA polymerase eta, involved in DNA repair. This polymerase has the ability of repairing damage from ultraviolet radiation and oxidative stress (291), suggesting that PSE may increase mutation rate. This is interesting because cancer cells speed their growth by acquiring mutations in genes that normally fix DNA damage, fueling the evolution of cancer (292,293).

The $tRNA^{Ser}$(S) cell line had mutations in some genes involved in translation, such as poly(A) binding proteins and the eukaryotic translation initiation factor; *EIF3A*. The eIF3a protein participates in translation of global mRNAs, but also regulates the translation of a subset of mRNAs that may be involved in cell proliferation, cell cycle control, cell differentiation and DNA repair (294,295). In many cancers, such as lung, breast and prostatic cancers, mutations in *EIF3A* gene are associated with susceptibility to tumor growth (296,297). Despite these connections with disease, previous studies using mistranslating cell lines do not report mutations in protein synthesis machinery components. The effect of these mutations should be elucidated in future studies.

Transcription factor genes accumulated a significant number of mutations in our mistranslating cell lines. This phenomenon was already observed in *S. cerevisiae* evolved mistranslating strains, where genes belonging to the category: transcription initiation and elongation also accumulated mutations (196).

One of the limitations of this work was the analysis of only one biological replicate. It would be important to complement this analysis with several biological replicates to

distinguish the variants that occur in all replicates and are specific for the adaptation of each cell line to PSE. One alternative could be the sequencing of single-cells (298). This approach is important to overcome key issues, such as population heterogeneity. In fact, this technique has been used to characterize some types of cancer (299,300). Also, our study would benefit from sequencing cells in P1 to identify the alterations that were present in the lines at the beginning of the evolution experiment (196,274).

4.5 Conclusions

Whole-exome sequencing, allowed us to detect genomic alterations that accumulated during evolution in the exome of mistranslating cell lines. The cell line that had more mutations was $tRNA^{Ser}$(A) and the majority of mutations in our cells occurred in genes involved in mRNA binding and RNA polymerase II transcription. Also, genes annotated in categories such as cytoskeleton and endoplasmic reticulum accumulated mutations. We were able to identify genes that could be related to the phenotypes described in Chapter 2; such as the mutations that occur in *HSF2* gene in $tRNA^{Ser}$(A) and in $tRNA^{Ser}$(L) cells and mutations in *DNAJC14* gene in $tRNA^{Ser}$(L) cells, or may explain adaptation; such as mutations in *POLH* and *EIF3A* genes in $tRNA^{Ser}$(A) and $tRNA^{Ser}$(S) cells, respectively. Nevertheless, how these mutations contribute to adaption to PSE remains to be clarified.

Future studies involving sequencing of evolved and non-evolved single-cells would be extremely useful to reduce population heterogeneity and identify mutations that are relevant to adaptation to PSE.

4.6 Materials and Methods

4.6.1 Cell Culture

Human embryonic kidney 293 (HEK293) were grown in Minimum Essential Medium (Gibco, Cat.41090-028) supplemented with 10% fetal bovine serum (FBS) (Sigma, Cat.F1051), 1% of Pen/Strep (Gibco, Cat.15070-063) and 1% of non-essential amino acids (Gibco, Cat.11140-050) in a humidified atmosphere at 37ºC in the presence of 5% CO_2.

4.6.2 Evolution of cells in culture

After transfection and plasmid selection, cells were kept in culture dishes (60mm) and subcultured ever 3 days using the same dilution (1/6) untill passage 30. Cell culture conditions were the same during evolution.

4.6.3 Mistranslating cell lines used

HEK293 cells stably transfected with an empty vector (Mock), and with the plasmid carrying $tRNA^{ser}_{AGA}$(S), $tRNA^{Ser}_{AGC}$(A), $tRNA^{Ser}_{AAG}$(L) or $tRNA^{Ser}_{GTG}$(H) genes were used in this study in passage 30 (see Methods from Chapter 2).

4.6.4 DNA Extraction

Genomic DNA extraction was carried out using the Genomic-tip 100/G kit (Qiagen) according to manufacturer's protocol. Quantification and quality assessment were performed using the Picogreen fluorescence bases quantification assay.

4.6.5 Exome Capture Nimblegen v3 and Illumina sequencing

Library preparation for exome sequencing was done using the capture kit "SeqCap EZ Human Exome Library v3" (Roche NimbleGen) that covers a total of 64 Mb. The insert size of the libraries was between 150 and 300 nucleotides. The initial DNA fragmentation was done using a Covaris S2 instrument. After library preparation, quality of libraries was checked using a capillary instrument, Bioanalyzer 2100(Agilent) to measure the size of the library and real time PCR to calculate its concentration. In order to load different libraries in the same lane of sequencing different adaptors were used (1-12).

Libraries were sequenced using a HiSeq2000 instrument (Illumina) with paired end reads of 2x75pb read length. For each sample was generated more than 50 fold average coverage. More than 80% of the filtered reads had a quality value higher than Q30 (*Phred score*).

Samples were sequenced in CNAG (Centre Nacional D'anàlisi Genòmica) in Barcelona in collaboration with Marta Gut, Mónica Bayés, Lídia Águeda and Sergi Beltran.

4.6.6 Alignment and analysis of sequences

After sequencing, images were acquired using Illumina Genome Analyzer Sequencing Control Software and analyzed. FASTQ files with base calls, quality metrics and read calls were generated. BWA (Burrows-Wheeler Aligner) software package was used for mapping sequences against the human reference genome (GRCh38) (301). SAMtools (Sequence Alignment/Map) was used for variant call and to generate VCF files (302). We also used PERL scripts developed in house and data stored in MySQL databases. For statistical analysis, we used the R software. Several quality criteria were applied during the analysis, but the final criteria to identify variants and annotated genes was: unique variants (defined by chromosome, coordinate and reference), variant quality (Q) ≥100, genotype quality (GQ) ≥90 and number of reads≥40. This was done to ensure the identification of high quality variants.

Alignment of sequences and analysis was done by Patrícia Oliveira from IPATIMUP/I3S from Porto.

4.6.7 Gene Ontology analysis

Genes that presented variants with high quality (unique variants; defined by chromosome, coordinate and reference, variant quality (Q) ≥100, genotype quality (GQ) ≥90 and number of reads≥40) were annotated using in GO categories using DAVID Bioinformatics Resources 6.7, an online database (254,255).

5. General Discussion and Future Perspectives

5.1 General Discussion

Protein synthesis is essential for life and maintenance of its fidelity is critical. Although cells possess several mechanisms to ensure fidelity and proteome homeostasis, errors are an intrinsic characteristic of protein synthesis and destabilize the proteome (51). Many different studies have shown that protein synthesis errors (PSE) are associated with degenerative phenotypes and disease (32,69,86). Surprisingly, recent studies suggest that PSE may be also advantageous depending on environmental conditions and physiology (67,75,196,228).

In our study, we wanted to understand the long-term effects of PSE in human cells in culture, namely the cellular stress responses pathways that are activated and the connections between them. For this, we used mutant tRNAs that introduce random errors in the proteome and created five stable cell lines: Mock (control), $tRNA^{Ser}$(S) (overexpressing the Wt $tRNA^{Ser}$), $tRNA^{Ser}$(A), $tRNA^{Ser}$(L) and $tRNA^{Ser}$(H) (Figure 2-1). This allowed us to create different proteome destabilization strategies to clarify if specific amino acid substitutions in proteins would produce different phenotypes. A previous study using similar Ser tRNAs transfected transiently into HEK293 cells showed severe alterations in cell death and activation of the UPS and UPR (165), suggesting that HEK293 cells or other types of human cells, could be important to understand the effects of PSE. Activation of the UPS and UPR were proportional to the severity of the errors being accumulated in the whole proteome, for example incorporation of Ser at isoleucine (Ile) codons sites strongly activated PQC, while incorporation of Ser at lysine (Lys) codon sites had a much lower impact on cells (165). In our study, we have observed that cells overexpressing the Wt $tRNA^{Ser}$ and mistranslating cell lines were also able to activate different PQC mechanisms, depending on the type of error introduced in the proteome (Figure 2-21).

Two of our cell lines, $tRNA^{Ser}$(S) and $tRNA^{Ser}$(A), activated the ATF6 branch of the UPR wich may explain their adaptation capacity to stress caused by PSE (Figure 2-18). When cells are under ER stress, ATF6 is cleaved and the cytosolic fragment (ATF6f) controls directly the expression of genes encoding ERAD components and XBP-1 (124). In addition, ATF6 has been associated with tolerance to chronic ER stress, since it leads to the expression of several genes that preserve ER function, including protein folding, protein degradation and maintenance of general ER homeostasis (303). In $tRNA^{Ser}$(S) cell line, in P15, ATF6f levels increased (Figure 2-18 A) and consequently the expression of genes involved in protein folding (*ERO1LB*, *PDIA4*, *DNAJC3*, *HYOU1*, *GRP94*), protein trafficking/ER homeostasis (*ATP2A2*,

NUCB2) and biosynthesis (*PDXK*), among others, was upregulated (Figure 3-3 A). Recent studies suggest that activation of the ATF6 pathway could correct the aberrant ER proteostasis of pathologic destabilized protein mutants involved in protein misfolding diseases (304,305). For example, in HEK293 cells expressing Wt and mutant forms of the protein rhodopsin, ATF6f was able to significantly reduce abnormal rhodopsin aggregates, while the levels of Wt proteins were not affected (306). Increasing ATF6 activation also attenuates aggregation of amyloidogenic proteins that can form extracellular aggregates (305). In $tRNA^{Ser}$(S) and $tRNA^{Ser}$(A) cell lines where ATF6f levels increased (Figure 2-18), there was accumulation of ubiquitinated proteins (Figure 2-11), but protein aggregates were not observed (Figure 2-10). Therefore, it is likely that activation of the ATF6 branch of the UPR, and the expression of targeted genes, was essential to maintain protein homeostasis and avoid the accumulation of aggregated proteins.

The UPR branches can interact with each other. In ER stress situations the ATF6 pathway can induce *XBP1* expression, which can be spliced by IRE1α to induce transcription of ER chaperones and ERAD components (119). In addition, the unspliced XBP-1 (XBP-1u) can negatively regulate the XBP-1 spliced form (XBP-1s), benefiting cell survival (307). In $tRNA^{Ser}$(S) cell line there was increased ATF6f (Figure 2-18) and *XBP1*u (Table 2-1). Whether *XBP1*u increased the level of the active transcription factor or induced a negative feedback is an important issue to address in future studies to clarify the importance of XBP-1 in adaptation of $tRNA^{Ser}$(S) cells to PSE.

Cells respond to various forms of stress by inhibiting or downregulating protein synthesis through phosphorylation of the translation initiation factor eIF2α. If the ER is exposed to stress, eIF2α is phosphorylated by PERK. The shutting down of protein synthesis, allows cells to recover from stress or result in cell death if the damage is beyond repair (308,309). Some mRNAs, such as ATF4 mRNA are efficiently translated when eIF2α is phosphorylated, through a mechanism that involves delayed translation re-initiation (310). Increased translation of ATF4 is required to increase the expression of genes that facilitate the adaptation of cells to stress (121,309). One of these genes encodes GADD34 that exerts a negative feedback loop by directly interacting with the catalytic subunit of type 1 protein serine/threonine phosphatase (PP1). Activation of PP1's results in dephosphorylation of eIF2α allowing for translational recovery (311). In our $tRNA^{Ser}$(A) cells, ER stress leads to downregulation of eIF2α-P (Figure 2-19 A), in P30, due to the upregulation of GADD34 (Figure 2-20 A), exerting a protective effect by preventing shutdown of protein synthesis. Studies with GADD34 knockout mice showed that GADD34 promotes the recovery from

protein synthesis inhibition induced by ER stress leading to increased cell survival (312). Also, in cells expressing mutant huntingtin protein, overexpression of GADD34 increased cell viability and induced cytoprotective autophagy (313). In line with this, in our $tRNA^{Ser}$(A) cells, *ATG16L1*, which encodes a protein component of a large complex essential for autophagy, was upregulated, from P15 to P30 (Table 2-1).

The activation of PQC mechanisms was concomitant with decreased expression of the mutant tRNAs (Figure 2-3). We were not able to elucidate how tRNA expression decreased, but it is likely that their transcription was downregulated or, for some unknown reason, they were targeted to degradation. Changes in endogenous tRNA abundance are regulated by a translational program where tRNA expression profiles vary depending on the differentiation or proliferation status of cells (314). One of the pathways that may be involved in tRNA expression regulation is the mammalian target of rapamycin (mTOR), since it controls tRNA and ribosome biogenesis (315,316) by interacting with Maf1, a known repressor of Pol III transcription. Other regulators of Pol III include the tumor suppressors retinoblastoma protein and p53 (317,318). The role of these factors in the expression of the mutant tRNAs should be studied in the future.

Since expression of our tRNAs imbalances the pool of serine tRNAs in the cell lines, it will be interesting to determine tRNAs levels during evolution. Indeed, expression of other serine tRNAs, that compete for the active site of SerRS, and other tRNAs that compete in the A-site of the ribosome with the mutant tRNAs carrying the correct amino acid ($tRNA_{AGC}^{Ala}$; $tRNA_{AAG}^{Leu}$ and $tRNA_{GTG}^{His}$) are likely to be affected. Upregulation of these tRNAs could be important to decrease the level of PSE and amino acids misincorporation. Breast cancer cells, not only overexpress tRNAs, but also perturb expression of isoacceptor tRNAs to optimize codon usage of tumor-promoting genes (158,316). Similarly, our mistranslating cells may be able to decrease mistranslation levels by adjusting tRNAs pool. In $tRNA^{Ser}$(S) cells, overexpression of $tRNA^{Ser}$ (Figure 2-3) also leads to overexpression of several genes that encode aaRS, such as AlaRS, AspRS, GlyRS, IleRS, MetRS, TrpRS and TyrRS (Figure 3-3 A), indicating that deregulation of a $tRNA^{Ser}$ also influences the expression of other components of the protein synthesis machinery. How alterations in aaRS expression can be involved in adaptation to PSE should be addressed in future studies, especially because aaRS can have several non-canonical functions, namely in signaling pathways (p53 and mTOR), immune response and translational control, and are associated with tumorigenesis (37,319).

Whole-exome sequencing allowed us to identify several mutations that were present in evolved cell lines in P30. Although most variants found are modifiers (there is no evidence of

impact in proteins structure), some of them can be correlated with the phenotypes observed. For example, mutations in *EIF3A* gene in tRNA(S) cells (Table 4-8) can be associated with up regulation of cell proliferation genes (Figure 3-4 A) and increased proliferation (Figure 2-7 A) in P30. This gene codes for the largest subunit of eIF3, which binds to the 40S ribosome and promotes the loading of the ternary complex of Met-tRNA/eiF2-GTP, forming the 43S preinitiation complex and can also assist eIF4 to recruit mRNAs to this complex. The eIF3a subunit regulates various genes involved in DNA repair, proliferation, differentiation and cell cycle, and is overexpressed in breast, cervical, esophageal and lung cancers (320–322). In addition to these, two SNPs (rs10787899 and rs3824830) have been identified as risk factores for the development of breast and pancreas adenocarcinomas (323). Several genes involved in transcription had mutations in our cell lines. Two examples were *MED23* (Table 4-7) and *HSF2* (Table 4-6) genes in tRNA(L) cells. *MED23* encodes a subunit of the Mediator complex that transduces regulatory information from transcription regulators to the RNA polymerase II apparatus, namely heat shock factors (HSF) mediated transcriptional activation (324,325). In stress conditions, trimerized HSF proteins bind to heat shock promoters and recruit MED23 to induce chaperones transcription by Pol II (325,326). Interestingly, another gene involved in induction of molecular chaperone transcription also had a mutation in this cell line, namely *HSF2* gene. Although further investigations will be needed to clarify the role of these genes, the mutations detected are likely responsible for the downregulation of chaperones observed in tRNA(L) cell line. Even though mRNA levels of chaperones were not altered in our study (Chapter 3), we observed downregulation of protein levels of HSP27, HSP60 and HSP70 (Chapter 2) in P30.

Depending on the type of error introduced in the proteome of our cell lines, some interesting phenotypes were observed. The $tRNA^{Ser}$(S) and $tRNA^{Ser}$(A) cells had increased proliferation (Figure 2-7) accompanied by upregulation of genes involved in glycolysis (Chapter 3), showing high capacity to deal with PSE, while $tRNA^{Ser}$(L) and $tRNA^{Ser}$(H) cells, were not able to increase proliferation rate at the end of evolution and showed accumulation of protein aggregates and decreased proliferation in P15, respectively (Figure 2-21). In the case of $tRNA^{Ser}$(S) cells, it is not surprinsing to observe these phenotypes since tRNA overexpression is frequently associated to proliferating phenotypes and serine tRNAs are the most overexpressed in breast tumors (158,314). tRNA imbalance can affect translation rate, fidelity and/or protein folding (327). In the case of cancer cells, overexpression of specific tRNAs can form "inducible" pathways with their direct target transcripts, enriched for their cognate codons, since these target transcripts can be more effectively translated (159). The tRNAs

carrying the amino acid Ser, might potentiate post-translational regulation of proteins involved in signal transduction, because Ser is a target for kinases and phosphatases, although the mechanism had been not identified (158). But, increase in the abundance of the tRNA pool may also accelerate elongation rate, compromising the folding of proteins, especially if there is an increase in low-abundance tRNAs. These tRNAs usually pair with slow-translating codons, thus facilitating co-translational folding (328).

In general, translational errors can occur during the discrimination of the correct amino acid by aaRS or during decoding in the ribosomes (329). In the case of aaRSs, errors are more likely to occur between amino acids with similar structure, for example between Ser and Ala. Such errors are usually reduced by the aaRS editing mechanisms, leading to the hydrolysis of misactivated or mischarged amino acids, before or after atachement to the tRNAs (31,32,330). If an editing error occurs, for example in AlaRS, this synthetase can aminoacylate $tRNA^{Ala}$ with Gly or Ser and these amino acids are incorporated into proteins in Ala sites (330,331). Therefore, in $tRNA^{Ser}$(A) cell line, the expression of the mutant $tRNA_{AGC}^{Ser}$, originates a type of error that is more likely to occur physiologically than Ser-for-Leu or Ser-for-His misincorporations, which are much more disruptive to the proteome and may influence cells phenotypes and adaptation.

Previous work from our lab showed that, NIH3T3 mouse embryo fibroblast cells stably transfected with the Wt $tRNA^{Ser}$ ($tRNA_{AGA}^{Ser}$) and a mutant tRNA that misincorporates Ser at Ala codons ($tRNA_{AGC}^{Ser}$) were angiogenic and tumorigenic, in the chorioallantoid membrane assays (CAM) and nude mice (unpublished data, Annex E, Figure 2). Tumor analyses showed strong selection of the mutant tRNA expression suggesting that mistranslation is advantageous for tumor growth (unpublished data, Annex E, Figure 1 and 2). In fact, these cell lines could activate the UPR and cancer-associated pathways. Moreover, tumors derived from cells expressing the mutant tRNA showed low levels of phosphorylated eIF2α due to GADD34 upregulation (unpublished data, Annex E, Figures 3-5). These results are consistent with our data in $tRNA^{Ser}$(A) cell line (Chapter 2) and show that PSE are likely relevant in cancer. Deregulation of the protein synthesis machinery has been associated to cancer, but the mechanism by which these alterations may lead to disease are still poorly understood (16,161,332,333). One clue unveiled in this thesis is the increase in protein turnover that allows $tRNA^{Ser}$(A) cells (Figures 2-9 and 2-12) to thrive and may be relevant in a cancer microenvironment. This was already reported in yeast cells with high PSE levels (196). But, this increase in protein turnover has an energetic cost that requires upregulation of glycolysis (Chapter 3).

The $tRNA^{Ser}$(L) cells accumulated protein aggregates in P15, but not in P30 (Figure 2-10). It was interesting that these cells could erase protein aggregates by activating PQC mechanisms. Another factor that may contribute to the dilution of protein aggregates may be cell division. A recent study showed that in cells that divide asymmetrically, such as *S. cerevisiae*, aggregated proteins are retained by the mother cell, through association with the mother-cell mitochondria. This asymmetric cell division results in an aging mother cell and a daughter cell free of aggregates and rejuvenated (334,335). Since dividing cells are able to erase protein aggregates, probably PSE and consequently accumulation of misfolded proteins can have different effects depending on the type of cells. Neurons, for instance, are probably not able to counteract PSE, in the same way dividing cells do, which will ultimately lead to cell death (69).

Previous work and our study show that PSE may have important consequences to the biology of eukaryotic cells, but we still do not yet understand the full impact of PSE in the proteome. PSE rates have been determined using specific proteins and biochemical or genetic reporters (75,76,167,195,228), but these methods have several disadvantages because they only measure one type of error at a time and do not provide a global picture of the amino acid misincorporations tolerated by cells. Mass spectrometry has also been used to search for specific types of amino acids misincorporations (336). Recently, our group was able to detect Ser misincorporations using mass spectrometry, in mice tumors originated from NIH3T3 cells that expressed the mutant $tRNA_{AGC}^{Ser}$, the same that we used in this work to create the $tRNA^{Ser}$(A) cell line. In comparison with Mock (empty plasmid), this cell line had more misincorporations in the two codons that are decoded by the $tRNA_{AGC}^{Ser}$ (GCT and GCC) (unpublished data, Annex E, Supplementary Figure 3). This technique is being improved to detect misincorporations and is very promising, since it may allow to obtain a global view of PSE in human cells in different pathologies or physiological conditions.

We have demonstrated that human cells can tolerate PSE in long term evolution experiments. Although the level of error was not quantified in our cell lines, we demonstrated that the mutant tRNAs were expressed and since previous studies showed that these tRNAs are fully functional in HEK293 cells (165,337), we are confident that the phenotypes described in this thesis represent cells responses to PSE.

The main alterations detected in the different cell lines used in this thesis are summarized in Figure 5-1. All cell lines accumulated ubiquitinated proteins in P15, but not in P30, and $tRNA^{Ser}$(L) was the only one that showed accumulation of protein aggregates. Two interesting mechanisms were identified in $tRNA^{Ser}$(A) and $tRNA^{Ser}$(L) cells that could explain

adaptation. In tRNA^{Ser}(A) cells (physiological error) there was increased protein turnover and activation of UPR, namely increase in fragmented ATF6 and decreased phosphorylation of eIF2α by GADD34. In tRNA^{Ser}(L) cells (non-physiological error) there was decreased protein synthesis rate due to the phosphorylation of eIF2α (Figure 5-1). At the transcriptional level, there was upregulation of genes involved in ER function and ER stress, and in some cases, pathways, such as glycolysis (Figure 5-1). Mutations in these cells occurred mainly in genes involved in RNA binding and transcription.

Further investigations are needed to obtain a global picture of the underlying adaptation to PSE, but our study provides the first evolutionary view of human cell adaptation to different types of PSE.

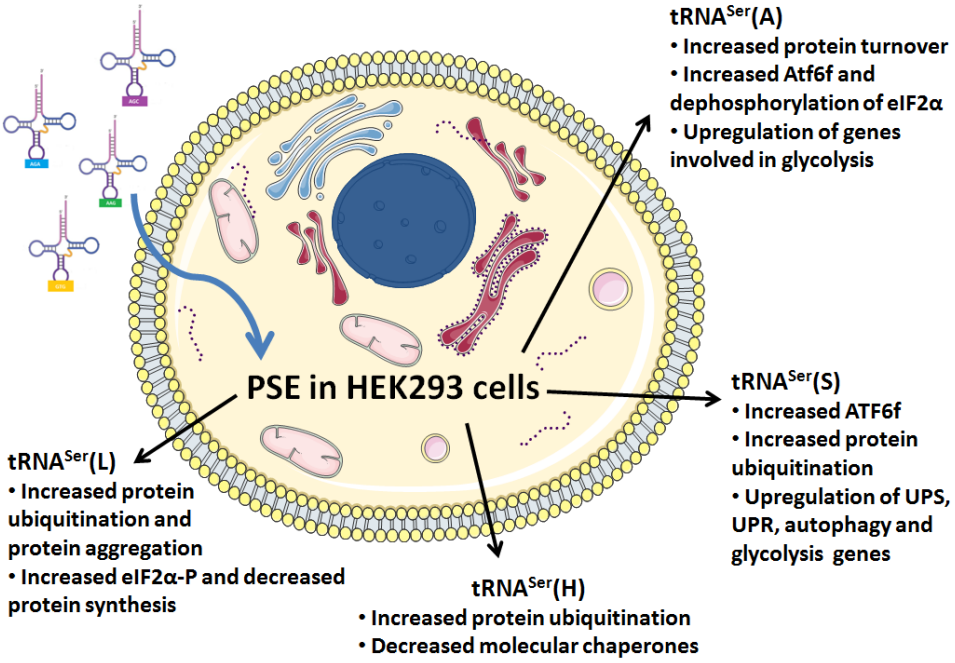


Figure 5-1. Summary of the main alterations detected during evolution of cells expressing the Wt tRNA^{Ser} and mistranslating cell lines [adapted from Servier Medical Art image bank (338)].

5.2 Future Perspectives

Our data unveiled several new issues that should be addressed and clarified in future studies.

1) Quantification of Ser misincorporations at non-cogante sites in proteins. This issue should be addressed for each cell line at each passage. Mass spectrometry identification of proteins containing higher level of PSE, would help better understand the phenotypes of the mistranslating cell lines.
2) Other cell or animal models should be used to identify common mechanisms of adaptation to PSE. Mutations in other tRNA isoacceptors and overexpression of other Wt tRNAs affecting different codons could help understand the diversity of stress responses to PSE. Overexpression of Wt tRNAs is very interesting and physiologically more relevant because alterations in the expression of protein synthesis components are usually associated with disease (159).
3) It would be interesting to inhibit the proteasome and/or protein synthesis rate of $tRNA^{Ser}$(A) cells and determine if such cells would accumulate protein aggregates. Also, understanding if protein turnover is important to overcome the effects of PSE would help understand adaptation to PSE. Since genes involved in glycolysis are upregulated, in this cell line, testing whether these cells have a higher glucose comsumption could reveal if adaptation to PSE requires higher ATP production.
4) Validation of the deregulation of MAPK and p53 pathways and the consequences of the mutations found in *EIF3A*, *MED23*, *HSF2* and other genes would help unveil the cellular networks that regulate proteotoxic stress induced by mistranslation.
5) Quantification of amino acids misincorporations in different pathologies would help understand the biomedical relevance of the PSE phenomenon.

6. References

1. Lodish HF, Berk A, Zipursky SL, Matsudaira P, Baltimore D, James D. Molecular Cell Biology. Vol. 5, Book. 2008. 801-846 p.
2. Lewin B. Genes VIII. Benjamin Cummings; United States ed edition; 2004. 1-1006 p.
3. Crick F. Central Dogma of Molecular Biology. Nature. 1970 Aug;227(5258):561–3.
4. Agris PF. Decoding the genome: A modified view. Vol. 32, Nucleic Acids Research. 2004. p. 223–38.
5. Crick FH. Codon--anticodon pairing: the wobble hypothesis. J Mol Biol. 1966;19(2):548–55.
6. Knight RD, Freeland SJ, Landweber LF. Selection, history and chemistry: The three faces of the genetic code. Vol. 24, Trends in Biochemical Sciences. 1999. p. 241–7.
7. Crick FHC. The origin of the genetic code. J Mol Biol. 1968;38(3):367–79.
8. Woese CR, Olsen GJ, Ibba M, Soll D. Aminoacyl-tRNA Synthetases, the Genetic Code, and the Evolutionary Process. Microbiol Mol Biol Rev. 2000;64(1):202–36.
9. Giegé R. Toward a more complete view of tRNA biology. TL - 15. Nat Struct Mol Biol. 2008;15 VN-r(10):1007–14.
10. Goodenbour JM, Pan T. Diversity of tRNA genes in eukaryotes. Nucleic Acids Res. 2006;34(21):6137–46.
11. Marquéz V, Nierhaus KH, Ribas de Pouplana L, Schimmel P. tRNA and Synthetases. In: Protein Synthesis and Ribosome Structure. Wiley-VCH Verlag GmbH & Co. KGaA; 2006. p. 145–84.
12. Tamura K. Origins and Early Evolution of the tRNA Molecule. Life. 2015;5(4):1687–99.
13. Lenhard B, Orellana O, Ibba M, Weygand-Durašević I. tRNA recognition and evolution of determinants in seryl-tRNA synthesis. Vol. 27, Nucleic Acids Research. 1999. p. 721–9.
14. Asahara H, Himeno H, Tamura K, Hasegawa T, Watanabe K, Shimizu M. Recognition Nucleotides of Escherichia coli tRNALeu and Its Elements Facilitating Discrimination from tRNASer and tRNATyr. J Mol Biol. 1993;231(2):219–29.
15. Towns WL, Begley TJ. Transfer RNA Methytransferases and Their Corresponding Modifications in Budding Yeast and Humans: Activities, Predications, and Potential Roles in Human Health. DNA Cell Biol. 2012;31(4):434–54.
16. Torres AG, Batlle E, Ribas de Pouplana L. Role of tRNA modifications in human diseases. Trends Mol Med. 2014 Jun;20(6):306–14.
17. Kirchner S, Ignatova Z. Emerging roles of tRNA in adaptive translation, signalling dynamics and disease. Nat Rev Genet. 2015 Feb;16(2):98–112.
18. Phizicky EM, Hopper AK. tRNA biology charges to the front. Vol. 24, Genes and Development. 2010. p. 1832–60.
19. El Yacoubi B, Bailly M, de Crécy-Lagard V. Biosynthesis and Function of Posttranscriptional Modifications of Transfer RNAs. Annu Rev Genet. 2011;46(1):120820103026000.
20. Björk GR, Durand JMB, Hagervall TG, Leipuvien R, Lundgren HK, Nilsson K, et al. Transfer RNA modification: influence on translational frameshifting and metabolism. FEBS Lett. 1999 Jun;452(1–2):47–51.
21. van Raam BJ, Salvesen GS. Transferring Death: A Role for tRNA in Apoptosis Regulation. Vol. 37, Molecular Cell. 2010. p. 591–2.
22. Raab JR, Chiu J, Zhu J, Katzman S, Kurukuti S, Wade PA, et al. Human tRNA genes function as chromatin insulators. EMBO J. 2012;31(2):330–50.
23. Park SG, Schimmel P, Kim S. Aminoacyl tRNA synthetases and their connections to disease. Proc Natl Acad Sci U S A. 2008 Aug;105(32):11043–9.
24. Reynolds NM, Lazazzera BA, Ibba M. Cellular mechanisms that control mistranslation. Vol. 8, Nature reviews Microbiology. 2010. p. 849–56.
25. Ribas de Pouplana L, Schimmel P. Aminoacyl-tRNA synthetases: Potential markers of genetic code development. Vol. 26, Trends in Biochemical Sciences. 2001. p. 591–6.
26. O'Donoghue P, Luthey-schulten Z. On the Evolution of Structure in Aminoacyl-tRNA Synthetases. Society. 2003;67(4):550–73.

27. Eriani G, Delarue M, Poch O, Gangloff J, Moras D. Partition of tRNA synthetases into two classes based on mutually exclusive sets of sequence motifs. Lett to Nat. 1990;347:203–6.
28. Rould MA, Perona JJ, Soll D, Steitz TA. Structure of E. coli glutaminyl-tRNA synthetase complexed with tRNA(Gln) and ATP at 2.8 A resolution. Science (80-). 1989 Dec;246(4934):1135–42.
29. Sherman JM, Soll D. Aminoacyl-tRNA synthetases optimize both cognate tRNA recognition and discrimination against noncognate tRNAs. Biochemistry. 1996;35(2):601–7.
30. Ibba M, Soll D. Aminoacyl-tRNA synthesis. Annu Rev Biochem. 2000 Jan;69:617–50.
31. Yadavalli SS, Musier-Forsyth K, Ibba M. The return of pretransfer editing in protein synthesis. Proc Natl Acad Sci. 2008 Dec;105(49):19031–2.
32. Schimmel P. Development of tRNA synthetases and connection to genetic code and disease. Protein Sci. 2008;17(10):1643–52.
33. Ling J, Reynolds N, Ibba M. Aminoacyl-tRNA synthesis and translational quality control. Annu Rev Microbiol. 2009;63:61–78.
34. Jakubowski H. Quality control in tRNA charging. Vol. 3, Wiley Interdisciplinary Reviews: RNA. 2012. p. 295–310.
35. Ryckelynck M, Giegé R, Frugier M. tRNAs and tRNA mimics as cornerstones of aminoacyl-tRNA synthetase regulations. In: Biochimie. 2005. p. 835–45.
36. Hausmann CD, Ibba M. Aminoacyl-tRNA synthetase complexes: Molecular multitasking revealed. Vol. 32, FEMS Microbiology Reviews. 2008. p. 705–21.
37. Guo M, Schimmel P. Essential nontranslational functions of tRNA synthetases. Nat Chem Biol. 2013;9(3):145–53.
38. Yannay-Cohen N, Carmi-Levy I, Kay G, Yang CM, Han JM, Kemeny DM, et al. LysRS Serves as a Key Signaling Molecule in the Immune Response by Regulating Gene Expression. Mol Cell. 2009;34(5):603–11.
39. Guo M, Yang X-L, Schimmel P. New functions of aminoacyl-tRNA synthetases beyond translation. Nat Rev Mol Cell Biol. 2010;11(9):668–74.
40. Khatter H, Myasnikov AG, Natchiar SK, Klaholz BP. Structure of the human 80S ribosome. Nature. 2015 Apr 22;520(7549):640–5.
41. Scheper GC, van der Knaap MS, Proud CG. Translation matters: protein synthesis defects in inherited disease. Nat Rev Genet. 2007;8(9):711–23.
42. Kapp LD, Lorsch JR. The molecular mechanics of eukaryotic translation. Annu Rev Biochem. 2004;73(1):657–704.
43. Jackson RJ, Hellen CUT, Pestova T V. The mechanism of eukaryotic translation initiation and principles of its regulation. Nat Rev Mol Cell Biol. 2010;11(2):113–27.
44. Hinnebusch AG, Lorsch JR. The mechanism of eukaryotic translation initiation: New insights and challenges. Cold Spring Harb Perspect Biol. 2012;4(10).
45. Klann E, Dever TE. Biochemical mechanisms for translational regulation in synaptic plasticity. Nat Rev Neurosci. 2004;5(12):931–42.
46. Dever TE, Green R. The elongation, termination, and recycling phases of translation in eukaryotes. Cold Spring Harb Perspect Biol. 2012;4(7):1–16.
47. Wohlgemuth I, Pohl C, Mittelstaet J, Konevega AL, Rodnina M V. Evolutionary optimization of speed and accuracy of decoding on the ribosome. Philos Trans R Soc B Biol Sci. 2011;366(1580):2979–86.
48. Manickam N, Joshi K, Bhatt MJ, Farabaugh PJ. Effects of tRNA modification on translational accuracy depend on intrinsic codon-anticodon strength. Nucleic Acids Res. 2015;44(4):1871–81.
49. Zaher HS, Green R. Quality control by the ribosome following peptide bond formation. Nature. 2009;457(7226):161–6.
50. Ibba M, Söll D. Aminoacyl-tRNAs: Setting the limits of the genetic code. Vol. 18, Genes and Development. 2004. p. 731–8.
51. Cochella L, Green R. Fidelity in protein synthesis. Curr Biol. 2016 Jul 12;15(14):R536–40.

52. Gromadski KB, Rodnina M V. Kinetic Determinants of High-Fidelity tRNA Discrimination on the Ribosome. Mol Cell. 2004;13(2):191–200.
53. Rodnina M V, Wintermeyer W. Fidelity of aminoacyl-tRNA selection on the ribosome: kinetic and structural mechanisms. Annu Rev Biochem. 2001;70:415–35.
54. Zaher HS, Green R. Fidelity at the molecular level: lessons from protein synthesis. Cell. 2009 Feb;136(4):746–62.
55. van Hoof A. Exosome-Mediated Recognition and Degradation of mRNAs Lacking a Termination Codon. Science (80-). 2002;295(5563):2262–4.
56. Brandman O, Hegde RS. Ribosome-associated protein quality control. Nat Struct Mol Biol. 2016;23(1):7–15.
57. Akimitsu N, Tanaka J, Pelletier J. Translation of nonSTOP mRNA is repressed post-initiation in mammalian cells. EMBO J. 2007;26(9):2327–38.
58. Simms CL, Thomas EN, Zaher HS. Ribosome-based quality control of mRNA and nascent peptides. Wiley Interdiscip Rev RNA. 2016;
59. Fredrick K, Ibba M. Errors rectified in retrospect Magnetic bond. Nature. 2009;457(January):157–8.
60. Moura GR, Carreto LC, Santos MA. Genetic code ambiguity: an unexpected source of proteome innovation and phenotypic diversity. Curr Opin Microbiol. 2009;12(6):631–7.
61. Farabaugh PJ, Björk GR. How translational accuracy influences reading frame maintenance. EMBO J. 1999;18(6):1427–34.
62. Ogle JM, Ramakrishnan V. Structural insights into translational fidelity. Annu Rev Biochem. 2005 Jan;74:129–77.
63. Loftfield RB, Vanderjagt D. The frequency of errors in protein biosynthesis. Biochem J. 1972;128(5):1353–6.
64. gallant J, Palmer L. Error propagation in viable cells. Mech Ageing Dev. 1979 Apr;10(1–2):27–38.
65. Moghal A, Mohler K, Ibba M. Mistranslation of the genetic code. Vol. 588, FEBS Letters. 2014. p. 4305–10.
66. Drummond DA, Wilke CO. The evolutionary consequences of erroneous protein synthesis. Nat Rev Genet. 2009;10(10):715–24.
67. Ribas de Pouplana L, Santos M, Zhu JH, Farabaugh PJ, Javid B. Protein mistranslation: Friend or foe? Trends Biochem Sci. 2014;39(8):355–62.
68. Dobson CM. The structural basis of protein folding and its links with human disease. Philos Trans R Soc Lond B Biol Sci. 2001;356(1406):133–45.
69. 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 Sep;443(7107):50–5.
70. Suzuki T, Ueda T, Watanabe K. The “polysemous” codon - A codon with multiple amino acid assignment caused by dual specificity of tRNA identity. EMBO J. 1997;16(5):1122–34.
71. Rocha R, Pereira PJB, Santos M a S, Macedo-Ribeiro S. Unveiling the structural basis for translational ambiguity tolerance in a human fungal pathogen. Proc Natl Acad Sci U S A. 2011;108(34):14091–6.
72. Gomes AC, Miranda I, Silva RM, Moura GR, Thomas B, Akoulitchev A, et al. A genetic code alteration generates a proteome of high diversity in the human pathogen Candida albicans. Genome Biol. 2007;8(10):R206.
73. Santos MAS, 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. Mol Microbiol. 1999 Feb;31(3):937–47.
74. Miranda I, Silva R, Santos MAS. Evolution of the genetic code in yeasts. Yeast. 2006;23(3):203–13.
75. Bezerra AR, Simões 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. Proc

Natl Acad Sci U S A. 2013;110:11079–84.
76. Netzer N, Goodenbour JM, David A, Dittmar K a, Jones RB, Schneider JR, et al. Innate immune and chemically triggered oxidative stress modifies translational fidelity. Nature. 2009;462(7272):522–6.
77. Pan T. Adaptive translation as a mechanism of stress response and adaptation. Annu Rev Genet. 2013;47:121–37.
78. Jones TE, Alexander RW, Pan T. Misacylation of specific nonmethionyl tRNAs by a bacterial methionyl-tRNA synthetase. Proc Natl Acad Sci U S A. 2011;108(17):6933–8.
79. Wiltrout E, Goodenbour JM, Fréchin M, Pan T. Misacylation of tRNA with methionine in Saccharomyces cerevisiae. Nucleic Acids Res. 2012;40(20):10494–506.
80. Miranda I, Silva-Dias A, Rocha R, Teixeira-Santos R, Coelho C, Gonçalves T, et al. Candida albicans CUG mistranslation is a mechanism to create cell surface variation. MBio. 2013;4(4).
81. Ecroyd H, Carver JA. Unraveling the mysteries of protein folding and misfolding. Vol. 60, IUBMB Life. 2008. p. 769–74.
82. Gregersen N, Bross P, Vang S, Christensen JH. Protein misfolding and human disease. Annu Rev Genomics Hum Genet. 2006;7:103–24.
83. Gidalevitz T, Prahlad V, Morimoto RI. The stress of protein misfolding: From single cells to multicellular organisms. Cold Spring Harb Perspect Biol. 2011;3(6):1–18.
84. Moreno-Gonzalez I, Soto C. Misfolded protein aggregates: Mechanisms, structures and potential for disease transmission. Vol. 22, Seminars in Cell and Developmental Biology. 2011. p. 482–7.
85. Herczenik E, Gebbink MFBG. Molecular and cellular aspects of protein misfolding and disease. FASEB J. 2008;22(7):2115–33.
86. Nangle LA, Motta CM, Schimmel P. Global Effects of Mistranslation from an Editing Defect in Mammalian Cells. Chem Biol. 2006;13(10):1091–100.
87. Amm I, Sommer T, Wolf DH. Protein quality control and elimination of protein waste: The role of the ubiquitin-proteasome system. Vol. 1843, Biochimica et Biophysica Acta - Molecular Cell Research. 2014. p. 182–96.
88. Chen B, Retzlaff M, Roos T, Frydman J. Cellular Strategies of Protein Quality Control. Cold Spring Harb Perspect Biol. 2011 Aug 1;3(8):a004374–a004374.
89. Beckerman M. Chaperones, Endoplasmic Reticulum Stress, and the Unfolded Protein Response. In: Cellular Signaling in Health and Disease. New York, NY: Springer US; 2009. p. 391–410.
90. Kim YE, Hipp MS, Bracher A, Hayer-Hartl M, Hartl FU. Molecular chaperone functions in protein folding and proteostasis. Vol. 82, Annual review of biochemistry. 2013. 323-55 p.
91. Kampinga HH, Craig EA. The HSP70 chaperone machinery: J proteins as drivers of functional specificity. Nat Rev Mol Cell Biol. 2010;11(8):579–92.
92. Taipale M, Jarosz DF, Lindquist S. HSP90 at the hub of protein homeostasis: emerging mechanistic insights. Nat Rev Mol Cell Biol. 2010;11(7):515–28.
93. Bryantsev AL, Kurchashova SY, Golyshev SA, Polyakov VY, Wunderink HF, Kanon B, et al. Regulation of stress-induced intracellular sorting and chaperone function of Hsp27 (HspB1) in mammalian cells. Biochem J. 2007;407(3):407–17.
94. Mayer MP, Bukau B. Hsp70 chaperones: cellular functions and molecular mechanism. Cell Mol Life Sci. 2005;62(6):670–84.
95. Hartl FU, Bracher A, Hayer-Hartl M. Molecular chaperones in protein folding and proteostasis. Nature. 2011;475(7356):324–32.
96. Auluck PK, Chan HYE, Trojanowski JQ, Lee VMY, Bonini NM. Chaperone suppression of alpha-synuclein toxicity in a Drosophila model for Parkinson's disease. Science. 2002;295(5556):865–8.
97. Bobkova N V., Garbuz DG, Nesterova I, Medvinskaya N, Samokhin A, Alexandrova I, et al. Therapeutic effect of exogenous Hsp70 in mouse models of Alzheimer's disease. J Alzheimer's Dis. 2014;38(2):425–35.

98. Lu R-C, Tan M-S, Wang H, Xie A-M, Yu J-T, Tan L. Heat shock protein 70 in Alzheimer's disease. Biomed Res Int. 2014;2014:435203.
99. Deocaris CC, Kaul SC, Wadhwa R. On the brotherhood of the mitochondrial chaperones mortalin and heat shock protein 60. Vol. 11, Cell stress & chaperones. 2006. p. 116–28.
100. Sun Y, MacRae TH. Small heat shock proteins: Molecular structure and chaperone function. Vol. 62, Cellular and Molecular Life Sciences. 2005. p. 2460–76.
101. Vilchez D, Saez I, Dillin A. The role of protein clearance mechanisms in organismal ageing and age-related diseases. Nat Commun. 2014;5:5659.
102. Navon A, Ciechanover A. The 26 S proteasome: From basic mechanisms to drug targeting. Vol. 284, Journal of Biological Chemistry. 2009. p. 33713–8.
103. Corn PG. Role of the ubiquitin proteasome system in renal cell carcinoma. BMC Biochem. 2007;8 Suppl 1(4):S4.
104. Tanaka K. The proteasome: overview of structure and functions. Proc Jpn Acad Ser B Phys Biol Sci. 2009;85(1):12–36.
105. Tanaka K, Mizushima T, Saeki Y. The proteasome: Molecular machinery and pathophysiological roles. In: Biological Chemistry. 2012. p. 217–34.
106. Voges D, Zwickl P, Baumeister W. The 26S proteasome: a molecular machine designed for controlled proteolysis. Annu Rev Biochem. 1999;68(1):1015–68.
107. Dantuma NP, Lindsten K. Stressing the ubiquitin-proteasome system. Cardiovasc Res. 2010 Jan 15;85(2):263–71.
108. Esser C, Alberti S, Höhfeld J. Cooperation of molecular chaperones with the ubiquitin/proteasome system. Biochim Biophys Acta - Mol Cell Res. 2004 Nov;1695(1–3):171–88.
109. Shimura H, Schwartz D, Gygi SP, Kosik KS. CHIP-Hsc70 Complex Ubiquitinates Phosphorylated Tau and Enhances Cell Survival. J Biol Chem. 2004;279(6):4869–76.
110. Imai Y, Soda M, Hatakeyama S, Akagi T, Hashikawa T, Nakayama KI, et al. CHIP is associated with Parkin, a gene responsible for familial Parkinson's Disease, and enhances its ubiquitin ligase activity. Mol Cell. 2002;10(1):55–67.
111. Jäger R, Bertrand MJM, Gorman AM, Vandenabeele P, Samali A. The unfolded protein response at the crossroads of cellular life and death during endoplasmic reticulum stress. Biol Cell. 2012;104(5):259–70.
112. Ruggiano A, Foresti O, Carvalho P. ER-associated degradation: Protein quality control and beyond. Vol. 204, Journal of Cell Biology. 2014. p. 869–79.
113. Janssens S, Pulendran B, Lambrecht BN. Emerging functions of the unfolded protein response in immunity. Nat Immunol. 2014;15(10):910–9.
114. Lin JH, Li H, Yasumura D, Cohen HR, Zhang C, Panning B, et al. IRE1 signaling affects cell fate during the unfolded protein response. Science. 2007;318(5852):944–9.
115. Hollien J, Lin JH, Li H, Stevens N, Walter P, Weissman JS. Regulated Ire1-dependent decay of messenger RNAs in mammalian cells. J Cell Biol. 2009;186(3):323–31.
116. Harding HP1, Novoa I, Zhang Y, Zeng H, Wek R, Schapira M RD. Regulated translation initiation controls stress-induced gene expression in mammalian cells. - PubMed - NCBI [Internet]. Mol. Cell. 2000. p. 1099–108. Available from: http://www.ncbi.nlm.nih.gov/pubmed/11106749
117. Kaufman RJ. Regulation of mRNA translation by protein folding in the endoplasmic reticulum. Vol. 29, Trends in Biochemical Sciences. 2004. p. 152–8.
118. Brush MH, Weiser DC, Shenolikar S. Growth arrest and DNA damage-inducible protein GADD34 targets protein phosphatase 1 alpha to the endoplasmic reticulum and promotes dephosphorylation of the alpha subunit of eukaryotic translation initiation factor 2. Mol Cell Biol. 2003;23(4):1292–303.
119. Yoshida H, Matsui T, Yamamoto A, Okada T, Mori K. XBP1 mRNA Is Induced by ATF6 and Spliced by IRE1 in Response to ER Stress to Produce a Highly Active Transcription Factor. Cell. 2001 Jul 28;107(7):881–91.
120. Puthalakath H, O'Reilly LA, Gunn P, Lee L, Kelly PN, Huntington ND, et al. ER Stress

Triggers Apoptosis by Activating BH3-Only Protein Bim. Cell. 2007;129(7):1337–49.
121. Urano F, Wang X, Bertolotti A, Zhang Y, Chung P, Harding HP, et al. Coupling of Stress in the ER to Activation of JNK Protein Kinases by Transmembrane Protein Kinase IRE1. Science (80-). 2000 Jan 28;287(5453):664–6.
122. Nishitoh H, Matsuzawa A, Tobiume K, Saegusa K, Takeda K, Inoue K, et al. ASK1 is essential for endoplasmic reticulum stress-induced neuronal cell death triggered by expanded polyglutamine repeats. Genes Dev. 2002;16(11):1345–55.
123. Hetz C, Chevet E, Harding HP. Targeting the unfolded protein response in disease. Nat Rev Drug Discov. 2013;12(9):703–19.
124. Hetz C. The unfolded protein response: controlling cell fate decisions under ER stress and beyond. Nat Rev Mol Cell Biol. 2012;13(2):89–102.
125. Lin JH, Li H, Zhang Y, Ron D, Walter P. Divergent effects of PERK and IRE1 signaling on cell viability. PLoS One. 2009;4(1).
126. Mahdi AA, Rizvi SHM, Parveen A. Role of Endoplasmic Reticulum Stress and Unfolded Protein Responses in Health and Diseases. Indian J Clin Biochem. 2016;31(2):127–37.
127. Kondo T, Asai M, Tsukita K, Kutoku Y, Ohsawa Y, Sunada Y, et al. Modeling Alzheimer's Disease with iPSCs Reveals Stress Phenotypes Associated with Intracellular Aβ and Differential Drug Responsiveness. Cell Stem Cell. 2013 Apr;12(4):487–96.
128. Baleriola J, Walker CA, Jean YY, Crary JF, Troy CM, Nagy PL, et al. Axonally Synthesized ATF4 Transmits a Neurodegenerative Signal across Brain Regions. Cell. 2014 Aug;158(5):1159–72.
129. Ma T, Trinh M a, Wexler AJ, Bourbon C, Gatti E, Pierre P, et al. Suppression of eIF2α kinases alleviates Alzheimer's disease-related plasticity and memory deficits. Nat Neurosci. 2013;16(9):1299–305.
130. Wang M, Kaufman RJ. Protein misfolding in the endoplasmic reticulum as a conduit to human disease. Nature. 2016;529(7586):326–35.
131. Wang M, Kaufman RJ. The impact of the endoplasmic reticulum protein-folding environment on cancer development. Nat Rev Cancer. 2014;14(9):581–97.
132. Vandewynckel Y-P, Laukens D, Geerts A, Bogaerts E, Paridaens A, Verhelst X, et al. The Paradox of the Unfolded Protein Response in Cancer. Anticancer Res. 2013;33(11):4683–94.
133. Bobrovnikova-Marjon E, Grigoriadou C, Pytel D, Zhang F, Ye J, Koumenis C, et al. PERK promotes cancer cell proliferation and tumor growth by limiting oxidative DNA damage. Oncogene. 2010;29(27):3881–95.
134. B'chir W, Maurin A-C, Carraro V, Averous J, Jousse C, Muranishi Y, et al. The eIF2α/ATF4 pathway is essential for stress-induced autophagy gene expression. Nucleic Acids Res. 2013 Sep 25;41(16):7683–99.
135. Nagy P, Varga Á, Pircs K, Hegedus K, Juhász G. Myc-Driven Overgrowth Requires Unfolded Protein Response-Mediated Induction of Autophagy and Antioxidant Responses in Drosophila melanogaster. PLoS Genet. 2013;9(8).
136. Hong SY, Hagen T. Multiple myeloma Leu167Ile (c.499C>A) mutation prevents XBP1 mRNA splicing. Br J Haematol. 2013;161(6):898–901.
137. Leung-Hagesteijn C, Erdmann N, Cheung G, Keats JJ, Stewart AK, Reece DE, et al. Xbp1s-Negative Tumor B Cells and Pre-Plasmablasts Mediate Therapeutic Proteasome Inhibitor Resistance in Multiple Myeloma. Cancer Cell. 2013;24(3):289–304.
138. Niederreiter L, Fritz TMJ, Adolph TE, Krismer A-M, Offner FA, Tschurtschenthaler M, et al. ER stress transcription factor Xbp1 suppresses intestinal tumorigenesis and directs intestinal stem cells. J Exp Med. 2013;210(10):2041–56.
139. Chen X, Iliopoulos D, Zhang Q, Tang Q, Greenblatt MB, Hatziapostolou M, et al. XBP1 promotes triple-negative breast cancer by controlling the HIF1α pathway. Nature. 2014;advance on(7494):103–7.
140. Nixon RA. The role of autophagy in neurodegenerative disease. Nat Med. 2013;19(8):983–97.

141. Wong E, Cuervo AM. Integration of clearance mechanisms: the proteasome and autophagy. Vol. 2, Cold Spring Harbor perspectives in biology. 2010.
142. McEwan DG, Dikic I. The Three Musketeers of Autophagy: Phosphorylation, ubiquitylation and acetylation. Vol. 21, Trends in Cell Biology. 2011. p. 195–201.
143. Glick D, Barth S, Macleod KF. Autophagy: cellular and molecular mechanisms. J Pathol. 2010 May;221(1):3–12.
144. Kim J, Kundu M, Viollet B, Guan K-L. AMPK and mTOR regulate autophagy through direct phosphorylation of Ulk1. Nat Cell Biol. 2011;13(2):132–41.
145. Narla A, Ebert BL. Ribosomopathies: Human disorders of ribosome dysfunction. Vol. 115, Blood. 2010. p. 3196–205.
146. Sulima SO, Patchett S, Advani VM, De Keersmaecker K, Johnson AW, Dinman JD. Bypass of the pre-60S ribosomal quality control as a pathway to oncogenesis. Proc Natl Acad Sci U S A. 2014;111(15):5640–5.
147. Yan X, Xu J, Gu Z-H, Pan C, Lu G, Shen Y, et al. Exome sequencing identifies somatic mutations of DNA methyltransferase gene DNMT3A in acute monocytic leukemia. Nat Genet. 2011;43(4):309–15.
148. Wills NM, Atkins JF. The potential role of ribosomal frameshifting in generating aberrant proteins implicated in neurodegenerative diseases. RNA. 2006;12(7):1149–53.
149. Antonellis A, Ellsworth RE, Sambuughin N, Puls I, Abel A, Lee-Lin S-Q, et al. Glycyl tRNA synthetase mutations in Charcot-Marie-Tooth disease type 2D and distal spinal muscular atrophy type V. Am J Hum Genet. 2003;72:1293–9.
150. Jordanova A, Irobi J, Thomas FP, Van Dijck P, Meerschaert K, Dewil M, et al. Disrupted function and axonal distribution of mutant tyrosyl-tRNA synthetase in dominant intermediate Charcot-Marie-Tooth neuropathy. Nat Genet. 2006;38(2):197–202.
151. Nangle L a, Zhang W, Xie W, Yang X-L, Schimmel P. Charcot-Marie-Tooth disease-associated mutant tRNA synthetases linked to altered dimer interface and neurite distribution defect. Proc Natl Acad Sci U S A. 2007;104(27):11239–44.
152. Abbott J a, Francklyn CS, Robey-Bond SM. Transfer RNA and human disease. Front Genet. 2014;5(June):158.
153. Kobayashi Y, Momoi MY, Tominaga K, Momoi T, Nihei K, Yanagisawa M, et al. A point mutation in the mitochondrial tRNALeu(UUR) gene in melas (mitochondrial myopathy, encephalopathy, lactic acidosis and stroke-like episodes). Biochem Biophys Res Commun. 1990;173(3):816–22.
154. Tsutomu S, Asuteka N, Takeo S. Human mitochondrial diseases caused by lack of taurine modification in mitochondrial tRNAs. Wiley Interdiscip Rev RNA. 2011 May;2(3):376–86.
155. Shoffner JM, Lott MT, Lezza a M, Seibel P, Ballinger SW, Wallace DC. Myoclonic epilepsy and ragged-red fiber disease (MERRF) is associated with a mitochondrial DNA tRNA(Lys) mutation. Cell. 1990;61(6):931–7.
156. White RJ. RNA polymerase III transcription and cancer. Oncogene. 2004 Apr;23(18):3208–16.
157. White RJ. RNA polymerases I and III, growth control and cancer. Nat Rev Mol Cell Biol. 2005;6(1):69–78.
158. Pavon-Eternod M, Gomes S, Geslain R, Dai Q, Rosner MR, Pan T. tRNA over-expression in breast cancer and functional consequences. Nucleic Acids Res. 2009 Nov;37(21):7268–80.
159. Goodarzi H, Nguyen HCB, Zhang S, Dill BD, Molina H, Tavazoie SF. Modulated Expression of Specific tRNAs Drives Gene Expression and Cancer Progression. Cell. 2016 Aug 11;165(6):1416–27.
160. Najmabadi H, Hu H, Garshasbi M, Zemojtel T, Abedini SS, Chen W, et al. Deep sequencing reveals 50 novel genes for recessive cognitive disorders. Nature. 2011;478(7367):57–63.
161. Rodriguez V, Chen Y, Elkahloun A, Dutra A, Pak E, Chandrasekharappa S. Chromosome 8 BAC array comparative genomic hybridization and expression analysis identify

amplification and overexpression of TRMT12 in breast cancer. Genes Chromosom Cancer. 2007;46(7):694–707.
162. Wei FY, Suzuki T, Watanabe S, Kimura S, Kaitsuka T, Fujimura A, et al. Deficit of tRNALys modification by Cdkal1 causes the development of type 2 diabetes in mice. J Clin Invest. 2011;121(9):3598–608.
163. Wei F-Y, Tomizawa K. Functional loss of Cdkal1, a novel tRNA modification enzyme, causes the development of type 2 diabetes [Review]. Endocr J. 2011;58(10):819–25.
164. Knowles SE, Gunn JM, Hanson RW, Ballard FJ. Increased degradation rates of protein synthesized in hepatoma cells in the presence of amino acid analogues. Biochem J. 1975;146(3):595–600.
165. Geslain R, Cubells L, Bori-Sanz T, Alvarez-Medina R, Rossell D, Martí E, et al. Chimeric tRNAs as tools to induce proteome damage and identify components of stress responses. Nucleic Acids Res. 2010;38:e30.
166. Brehme M, Voisine C, Rolland T, Wachi S, Soper JH, Zhu Y, et al. A chaperome subnetwork safeguards proteostasis in aging and neurodegenerative disease. Cell Rep. 2014;9(3):1135–50.
167. Reverendo M, Soares AR, Pereira PM, Carreto L, Ferreira V, Gatti E, et al. TRNA mutations that affect decoding fidelity deregulate development and the proteostasis network in zebrafish. RNA Biol. 2014;11(9):1199–213.
168. Gingold H, Pilpel Y. Determinants of translation efficiency and accuracy. Mol Syst Biol. 2011;7(481):481.
169. Ruusala T, Andersson D, Ehrenberg M, Kurland CG. Hyper-accurate ribosomes inhibit growth. EMBO J. 1984;3(11):2575–80.
170. Zhou Y, Goodenbour JM, Godley LA, Wickrema A, Pan T. High levels of tRNA abundance and alteration of tRNA charging by bortezomib in multiple myeloma. Biochem Biophys Res Commun. 2009;385(2):160–4.
171. Frye M, Watt FM. The RNA Methyltransferase Misu (NSun2) Mediates Myc-Induced Proliferation and Is Upregulated in Tumors. Curr Biol. 2006;16(10):971–81.
172. Begley U, Sosa MS, Avivar-Valderas A, Patil A, Endres L, Estrada Y, et al. A human tRNA methyltransferase 9-like protein prevents tumour growth by regulating LIN9 and HIF1-α. EMBO Mol Med. 2013;5(3):366–83.
173. Liu Y, Satz JS, Vo M-NN, Nangle LA, Schimmel P, Ackerman SL. Deficiencies in tRNA synthetase editing activity cause cardioproteinopathy. TL - 111. Proc Natl Acad Sci U S A. 2014;111 VN-(49):17570–5.
174. Bayat V, Thiffault I, Jaiswal M, T??treault M, Donti T, Sasarman F, et al. Mutations in the mitochondrial methionyl-tRNA synthetase cause a neurodegenerative phenotype in flies and a recessive ataxia (ARSAL) in humans. PLoS Biol. 2012;10(3).
175. Bukau B, Weissman J, Horwich A. Molecular chaperones and protein quality control. Cell. 2006;125(3):443–51.
176. Wang S, Kaufman RJ. The impact of the unfolded protein response on human disease. J Cell Biol. 2012;197:857–67.
177. Tuller T. The Effect of Dysregulation of tRNA Genes and Translation Efficiency Mutations in Cancer and Neurodegeneration. Front Genet. 2012 Oct;3:201.
178. Wang G, Yang Z-Q, Zhang K. Endoplasmic reticulum stress response in cancer: molecular mechanism and therapeutic potential. Am J Transl Res. 2010;2:65–74.
179. Normanly J, Ollick T, Abelson J. Eight base changes are sufficient to convert a leucine-inserting tRNA into a serine-inserting tRNA. Proc Natl Acad Sci U S A. 1992;89(12):5680–4.
180. Gray IC, Barnes MR. Amino Acid Properties and Consequences of Substitutions. Bioinforma Genet. 2003;4:289–304.
181. Schmidt EK, Clavarino G, Ceppi M, Pierre P. SUnSET, a nonradioactive method to monitor protein synthesis. Nat Meth. 2009 Apr;6(4):275–7.
182. Li X, Zhang K, Li Z. Unfolded protein response in cancer: the physician's perspective. J

Hematol {&} Oncol. 2011;4:8.
183. Macario AJL, Conway de Macario E. Stress and molecular chaperones in disease. Vol. 30, International Journal of Clinical and Laboratory Research. 2000. p. 49–66.
184. Broadley SA, Hartl FU. The role of molecular chaperones in human misfolding diseases. Vol. 583, FEBS Letters. 2009. p. 2647–53.
185. Parcellier A, Schmitt E, Gurbuxani S, Seigneurin-Berny D, Pance A, Chantôme A, et al. HSP27 is a ubiquitin-binding protein involved in I-kappaBalpha proteasomal degradation. Mol Cell Biol. 2003;23(16):5790–802.
186. Marcu MG, Doyle M, Bertolotti A, Ron D, Hendershot L, Neckers L. Heat shock protein 90 modulates the unfolded protein response by stabilizing IRE1alpha. Mol Cell Biol. 2002;22(24):8506–13.
187. Wang M, Wey S, Zhang Y, Ye R, Lee AS. Role of the unfolded protein response regulator GRP78/BiP in development, cancer, and neurological disorders. Antioxidants {&} redox Signal. 2009;11(9):2307–16.
188. Xu C, Bailly-Maitre B, Reed J. Endoplasmic reticulum stress: cell life and death decisions. J Clin Invest. 2005;115(10):2656–64.
189. DuRose JB, Scheuner D, Kaufman RJ, Rothblum LI, Niwa M. Phosphorylation of eukaryotic translation initiation factor 2alpha coordinates rRNA transcription and translation inhibition during endoplasmic reticulum stress. Mol Cell Biol. 2009;29(15):4295–307.
190. Tay KH, Luan Q, Croft A, Jiang CC, Jin L, Zhang XD, et al. Sustained IRE1 and ATF6 signaling is important for survival of melanoma cells undergoing ER stress. Cell Signal. 2014 Feb;26(2):287–94.
191. Otomo C, Metlagel Z, Takaesu G, Otomo T. Structure of the human ATG12~ATG5 conjugate required for LC3 lipidation in autophagy. Nat Struct Mol Biol. 2013 Jan 2;20(1):59–66.
192. Sonenberg N, Hinnebusch AG. Regulation of Translation Initiation in Eukaryotes: Mechanisms and Biological Targets. Vol. 136, Cell. 2009. p. 731–45.
193. Dufner a, Thomas G. Ribosomal S6 kinase signaling and the control of translation. Exp Cell Res. 1999;253:100–9.
194. Ruan B, Palioura S, Sabina J, Marvin-Guy L, Kochhar S, Larossa RA, et al. Quality control despite mistranslation caused by an ambiguous genetic code. Proc Natl Acad Sci U S A. 2008;105(43):16502–7.
195. Paredes JA, Carreto L, Simões J, Bezerra AR, Gomes AC, Santamaria R, et al. Low level genome mistranslations deregulate the transcriptome and translatome and generate proteotoxic stress in yeast. BMC Biol. 2012 Jan;10:55.
196. Kalapis D, Bezerra AR, Farkas Z, Horvath P, Bódi Z, Daraba A, et al. Evolution of Robustness to Protein Mistranslation by Accelerated Protein Turnover. PLoS Biol. 2015;13(11).
197. Khlistunova I, Biernat J, Wang Y, Pickhardt M, von Bergen M, Gazova Z, et al. Inducible Expression of Tau Repeat Domain in Cell Models of Tauopathy: AGGREGATION IS TOXIC TO CELLS BUT CAN BE REVERSED BY INHIBITOR DRUGS . J Biol Chem . 2006 Jan 13;281(2):1205–14.
198. Lim S, Haque MM, Kim D, Kim DJ, Kim YK. Cell-based models to investigate Tau aggregation. Vol. 12, Computational and Structural Biotechnology Journal. 2014. p. 7–13.
199. Falkenburger BH, Saridaki T, Dinter E. Cellular models for Parkinson's disease. Journal of Neurochemistry. 2016;
200. Stansley B, Post J, Hensley K. A comparative review of cell culture systems for the study of microglial biology in Alzheimer's disease. J Neuroinflammation. 2012;9(1):115.
201. Verhoef LGGC, Lindsten K, Masucci MG, Dantuma NP. Aggregate formation inhibits proteasomal degradation of polyglutamine proteins. Hum Mol Genet. 2002;11(22):2689–700.
202. Soto C. Unfolding the role of protein misfolding in neurodegenerative diseases. Nat Rev

Neurosci. 2003;4(1):49–60.
203. Ross C, Poirier M. Protein aggregation and neurodegenerative disease. Nat Med. 2004;10 Suppl(July):S10-7.
204. Pavon-Eternod M, Gomes S, Rosner MR, Pan T. Overexpression of initiator methionine tRNA leads to global reprogramming of tRNA expression and increased proliferation in human epithelial cells. RNA. 2013;19:461–6.
205. de Baets G, Reumers J, Blanco JD, Dopazo J, Schymkowitz J, Rousseau F. An evolutionary trade-off between protein turnover rate and protein aggregation favors a higher aggregation propensity in fast degrading proteins. PLoS Comput Biol. 2011;7(6).
206. Trcka F, Vojtesek B, Muller P. Protein quality control and cancerogenesis. Klin Onkol Cas Ces a Slov Onkol Spol. 2012;25 Suppl 2:2S38-44.
207. Calderwood SK, Khaleque MA, Sawyer DB, Ciocca DR. Heat shock proteins in cancer: Chaperones of tumorigenesis. Vol. 31, Trends in Biochemical Sciences. 2006. p. 164–72.
208. Kedersha N, Stoecklin G, Ayodele M, Yacono P, Lykke-Andersen J, Fitzler MJ, et al. Stress granules and processing bodies are dynamically linked sites of mRNP remodeling. J Cell Biol. 2005;169(6):871–84.
209. McEwen E, Kedersha N, Song B, Scheuner D, Gilks N, Han A, et al. Heme-regulated inhibitor kinase-mediated phosphorylation of eukaryotic translation initiation factor 2 inhibits translation, induces stress granule formation, and mediates survival upon arsenite exposure. J Biol Chem. 2005;280(17):16925–33.
210. Hay DG, Sathasivam K, Tobaben S, Stahl B, Marber M. Progressive decrease in chaperone protein levels in a mouse model of Huntingt ... Hum Mol Genet. 2004;13:1389–405.
211. Arumugam T V, Phillips TM, Cheng A, Morrell CH, Mattson MP, Wan R. Age and energy intake interact to modify cell stress pathways and stroke outcome. Ann Neurol. 2010;67(1):41–52.
212. Njemini R, Lambert M, Demanet C, Mets T. The effect of aging and inflammation on heat shock protein 27 in human monocytes and lymphocytes. Exp Gerontol. 2006;41(3):312–9.
213. Ben-Zvi A, Miller E a, Morimoto RI. Collapse of proteostasis represents an early molecular event in Caenorhabditis elegans aging. Proc Natl Acad Sci. 2009 Sep;106(35):14914–9.
214. Carvalho J, Van Grieken NC, Pereira PM, Sousa S, Tijssen M, Buffart TE, et al. Lack of microRNA-101 causes E-cadherin functional deregulation through EZH2 up-regulation in intestinal gastric cancer. J Pathol. 2012;228(1):31–44.
215. Hurst CD, Zuiverloon TCM, Hafner C, Zwarthoff EC, Knowles M a. A SNaPshot assay for the rapid and simple detection of four common hotspot codon mutations in the PIK3CA gene. BMC Res Notes. 2009;2:66.
216. Roth V. Doubling Time Computing. 2006.
217. Simon R, BRB-ArrayTools Development Team. BRB-ArrayTools.
218. Simon R, Lam A, Li M-C, Ngan M, Menenzes S, Zhao Y. Analysis of Gene Expression Data Using BRB-Array Tools. Cancer Inform. 2007 Feb;3:11–7.
219. Saeed AI, Sharov V, White J, Li J, Liang W, Bhagabati N, et al. TM4: A free, open-source system for microarray data management and analysis. Biotechniques. 2003;34(2):374–8.
220. Saeed A, Bhagabati N, Braisted J, Liang W, Sharov V, Howe E. TM4 microarray software suit. Methods Enzymol. 2006;411:134–93.
221. de Nadal E, Ammerer G, Posas F. Controlling gene expression in response to stress. Nat Rev Genet. 2011;12(12):833–45.
222. Brown JB, Boley N, Eisman R, May GE, Stoiber MH, Duff MO, et al. Diversity and dynamics of the Drosophila transcriptome. Nature. 2014;512(7515):1–7.
223. Gasch a P, Spellman PT, Kao CM, Carmel-Harel O, Eisen MB, Storz G, et al. Genomic expression programs in the response of yeast cells to environmental changes. Mol Biol Cell. 2000;11(12):4241–57.
224. DeRisi J, Penland L, Brown PO, Bittner ML, Meltzer PS, Ray M, et al. Use of a cDNA microarray to analyse gene expression patterns in human cancer. Nat Genet. 1996;14(4):457–60.

225. Ross DT, Scherf U, Eisen MB, Perou CM, Rees C, Spellman P, et al. Systematic variation in gene expression patterns in human cancer cell lines. Nat Genet. 2000;24(3):227–35.
226. Trotter EW. Misfolded Proteins Are Competent to Mediate a Subset of the Responses to Heat Shock in Saccharomyces cerevisiae. J Biol Chem. 2002 Nov 15;277(47):44817–25.
227. Varanda AS. Master Thesis - Aberrant protein synthesis in human HEK293FT cells. University of Aveiro; 2011.
228. Silva RM, Paredes JA, Moura GR, Manadas B, Costa TL, Rocha R, et al. Critical roles for a genetic code alteration in the evolution of the genus Candida. EMBO J. 2007;26(21):4555–65.
229. Hardie DG. AMP-activated protein kinase-an energy sensor that regulates all aspects of cell function. Vol. 25, Genes and Development. 2011. p. 1895–908.
230. Bezerra ARM. Doctoral Thesis - Molecular genomics of a genetic code alteration. University of Aveiro; 2013.
231. Dombroski BA, Nayak RR, Ewens KG, Ankener W, Cheung VG, Spielman RS. Gene Expression and Genetic Variation in Response to Endoplasmic Reticulum Stress in Human Cells. Am J Hum Genet. 2010;86(5):719–29.
232. Rutkowski DT, Kaufman RJ. That which does not kill me makes me stronger: adapting to chronic ER stress. Trends Biochem Sci. 2016 Sep 22;32(10):469–76.
233. Walter P, Ron D. The unfolded protein response: from stress pathway to homeostatic regulation. Science. 2011;334(6059):1081–6.
234. Rutkowski DT, Hegde RS. Regulation of basal cellular physiology by the homeostatic unfolded protein response. Vol. 189, Journal of Cell Biology. 2010. p. 783–94.
235. Wang X, Eno CO, Altman BJ, Zhu Y, Zhao G, Olberding KE, et al. ER stress modulates cellular metabolism. Biochem J. 2011;435(1):285–96.
236. DeBerardinis RJ, Lum JJ, Hatzivassiliou G, Thompson CB. The Biology of Cancer: Metabolic Reprogramming Fuels Cell Growth and Proliferation. Vol. 7, Cell Metabolism. 2008. p. 11–20.
237. Bravo R, Parra V, Gatica D, Rodriguez AE, Torrealba N, Paredes F, et al. Endoplasmic Reticulum and the Unfolded Protein Response. Dynamics and Metabolic Integration. Int Rev Cell Mol Biol. 2013;301:215–90.
238. Bi M, Naczki C, Koritzinsky M, Fels D, Blais J, Hu N, et al. ER stress-regulated translation increases tolerance to extreme hypoxia and promotes tumor growth. EMBO J. 2005;24:3470–81.
239. Gatenby RA, Gillies RJ. Why do cancers have high aerobic glycolysis? Nat Rev Cancer. 2004;4(11):891–9.
240. Peth A, Nathan JA, Goldberg AL. The ATP costs and time required to degrade ubiquitinated proteins by the 26 S proteasome. J Biol Chem. 2013;288(40):29215–22.
241. Saibil H. Chaperone machines for protein folding, unfolding and disaggregation. Nat Rev Mol Cell Biol. 2013;14(10):630–42.
242. Braakman I, Bulleid NJ. Protein folding and modification in the Mammalian endoplasmic reticulum. Annu Rev Biochem. 2011;80:71–99.
243. Gerich JE, Woerle HJ, Meyer C, Stumvoll M. Renal gluconeogenesis: Its importance in human glucose homeostasis. Vol. 24, Diabetes Care. 2001. p. 382–91.
244. Joerger AC, Fersht AR. The p53 Pathway: Origins, Inactivation in Cancer, and Emerging Therapeutic Approaches. Annu Rev Biochem. 2016;85(1):annurev-biochem-060815-014710.
245. Li J, Lee B, Lee AS. Endoplasmic reticulum stress-induced apoptosis: Multiple pathways and activation of p53-UP-regulated modulator of apoptosis (PUMA) and NOXA by p53. J Biol Chem. 2006;281(11):7260–70.
246. Lin W-C, Chuang Y-C, Chang Y-S, Lai M-D, Teng Y-N, Su I-J, et al. Endoplasmic reticulum stress stimulates p53 expression through NF-κB activation. PLoS One. 2012;7(7):e39120.
247. Darling NJ, Cook SJ. The role of MAPK signalling pathways in the response to

endoplasmic reticulum stress. Vol. 1843, Biochimica et Biophysica Acta - Molecular Cell Research. 2014. p. 2150–63.
248. Szegezdi E, Logue SE, Gorman AM, Samali A. Mediators of endoplasmic reticulum stress-induced apoptosis. EMBO Rep. 2006;7(9):880–5.
249. Zarubin T, Han J. Activation and signaling of the p38 MAP kinase pathway. Cell Res. 2005;15(1):11–8.
250. Hou NS, Taubert S. Membrane lipids and the endoplasmic reticulum unfolded protein response: An interesting relationship. Worm. 2014;3(3):e962405.
251. Shen X, Ellis RE, Sakaki K, Kaufman RJ. Genetic interactions due to constitutive and inducible gene regulation mediated by the unfolded protein response in C. elegans. PLoS Genet. 2005;1(3).
252. Bernales S, McDonald KL, Walter P. Autophagy counterbalances endoplasmic reticulum expansion during the unfolded protein response. PLoS Biol. 2006;4(12):2311–24.
253. Basseri S, Austin RC. Endoplasmic reticulum stress and lipid metabolism: Mechanisms and therapeutic potential. Biochemistry Research International. 2012.
254. Huang DW, Lempicki R a, Sherman BT. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc. 2009;4(1):44–57.
255. Huang DW, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: Paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res. 2009;37(1):1–13.
256. Goodwin S, McPherson JD, McCombie WR. Coming of age: ten years of next-generation sequencing technologies. Nat Rev Genet. 2016;17(6):333–51.
257. Warr A, Robert C, Hume D, Archibald A, Deeb N, Watson M. Exome Sequencing: Current and Future Perspectives. G3 Genes|Genomes|Genetics. 2015;5(8):1543–50.
258. Pabinger S, Dander A, Fischer M, Snajder R, Sperk M, Efremova M, et al. A survey of tools for variant analysis of next-generation genome sequencing data. Brief Bioinform. 2014;15(2):256–78.
259. Yang Y, Muzny DM, Reid JG, Bainbridge MN, Willis A, Ward P a, et al. Clinical Whole-Exome Sequencing for the Diagnosis of Mendelian Disorders. N Engl J Med. 2013;369(16):1502–11.
260. Sassi C, Guerreiro R, Gibbs R, Ding J, Lupton MK, Troakes C, et al. Exome sequencing identifies 2 novel presenilin 1 mutations (p.L166V and p.S230R) in British early-onset Alzheimer's disease. Neurobiol Aging. 2014;35(10).
261. Johnson JO, Mandrioli J, Benatar M, Abramzon Y, Van Deerlin VM, Trojanowski JQ, et al. Exome Sequencing Reveals VCP Mutations as a Cause of Familial ALS. Neuron. 2010;68(5):857–64.
262. Kiiski JI, Pelttari LM, Khan S, Freysteinsdottir ES, Reynisdottir I, Hart SN, et al. Exome sequencing identifies FANCM as a susceptibility gene for triple-negative breast cancer. Proc Natl Acad Sci U S A. 2014;111(42):15172–7.
263. Cai J, Li L, Ye L, Jiang X, Shen L, Gao Z, et al. Exome sequencing reveals mutant genes with low penetrance involved in MEN2A-associated tumorigenesis. Endocr Relat Cancer. 2015;22(1):23–33.
264. Cirulli ET, Lasseigne BN, Petrovski S, Sapp PC, Dion PA, Leblond CS, et al. Exome sequencing in amyotrophic lateral sclerosis identifies risk genes and pathways. Science. 2015;347(6229):1436–41.
265. Hansen SN, Ehlers NS, Zhu S, Thomsen MBH, Nielsen RL, Liu D, et al. The stepwise evolution of the exome during acquisition of docetaxel resistance in breast cancer cells. BMC Genomics. 2016 Jun 9;17:442.
266. Wang Y, Waters J, Leung ML, Unruh A, Roh W, Shi X, et al. Clonal evolution in breast cancer revealed by single nucleus genome sequencing. Nature. 2014 Aug 14;512(7513):155–60.
267. Weston-Bell N, Gibson J, John M, Ennis S, Pfeifer S, Cezard T, et al. Exome sequencing in tracking clonal evolution in multiple myeloma following therapy. Leukemia. 2013 Apr;27(5):1188–91.

268. Kovac M, Blattmann C, Ribi S, Smida J, Mueller NS, Engert F, et al. Exome sequencing of osteosarcoma reveals mutation signatures reminiscent of BRCA deficiency. Nat Commun. 2015;6:8940.
269. Galhardo RS, Hastings PJ, Rosenberg SM. Mutation as a Stress Response and the Regulation of Evolvability. Crit Rev Biochem Mol Biol. 2007;42(5):399–435.
270. Rosenberg SM. Evolving responsively: adaptive mutation. Nat Rev Genet. 2001;2(7):504–15.
271. Chen G, Bradford WD, Seidel CW, Li R. Hsp90 stress potentiates rapid cellular adaptation through induction of aneuploidy. Nature. 2012;482(7384):246–50.
272. Pavelka N, Rancati G, Li R. Dr Jekyll and Mr Hyde: Role of aneuploidy in cellular adaptation and cancer. Vol. 22, Current Opinion in Cell Biology. 2010. p. 809–15.
273. Giam M, Rancati G. Aneuploidy and chromosomal instability in cancer: a jackpot to chaos. Cell Div. 2015;10:3.
274. Simões J, Bezerra AR, Moura GR, Araújo H, Gut I, Bayes M, et al. The fungus Candida albicans tolerates ambiguity at multiple codons. Front Microbiol. 2016;7(MAR).
275. Meynert AM, Ansari M, FitzPatrick DR, Taylor MS. Variant detection sensitivity and biases in whole genome and exome sequencing. BMC Bioinformatics. 2014;15(1):247.
276. Columbia U of B. Biology 334 Classical Genetics [Internet]. Available from: http://www.zoology.ubc.ca/genetics/chapter14.htm
277. Cristianini N, Hahn MW. Introduction to Computational Genomics A Case Studies Approach. In: Annals of Physics. 2006. p. 182.
278. Ratan A, Olson TL, Loughran TP, Miller W. Identification of indels in next-generation sequencing data. BMC Bioinformatics. 2015;16(1):1–8.
279. Mathew A, Mathur SK, Morimoto RI. Heat shock response and protein degradation: regulation of HSF2 by the ubiquitin-proteasome pathway. Mol Cell Biol. 1998;18(9):5091–8.
280. Vasudevan S, Starostina NG, Kipreos ET. The Caenorhabditis elegans cell-cycle regulator ZYG-11 defines a conserved family of CUL-2 complex components. EMBO Rep. 2007;8(3):279–86.
281. Biesecker LG, Shianna K V, Mullikin JC. Exome sequencing: the expert view. Genome Biol. 2011;12(9):128.
282. Wang J, Yu H, Zhang VW, Tian X, Feng Y, Wang G, et al. Capture-based high-coverage NGS: a powerful tool to uncover a wide spectrum of mutation types. Genet Med. 2015;(July):1–9.
283. Sims D, Sudbery I, Ilott NE, Heger A, Ponting CP. Sequencing depth and coverage: key considerations in genomic analyses. Nat Rev Genet. 2014;15(2):121–32.
284. Akerfelt M, Morimoto RI, Sistonen L. Heat shock factors: integrators of cell stress, development and lifespan. Nat Rev Mol Cell Biol. 2010;11(8):545–55.
285. Östling P, Björk JK, Roos-Mattjus P, Mezger V, Sistonen L. Heat Shock Factor 2 (HSF2) contributes to inducible expression of hsp genes through interplay with HSF1. J Biol Chem. 2007;282(10):7077–86.
286. Wilkerson DC, Skaggs HS, Sarge KD. HSF2 binds to the Hsp90, Hsp27, and c-Fos promoters constitutively and modulates their expression. Cell Stress Chaperones. 2007 Aug 9;12(3):283–90.
287. Stephanou A, Latchman DS. Transcriptional modulation of heat-shock protein gene expression. Biochemistry Research International. 2011.
288. Bermak JC, Li M, Bullock C, Zhou Q-Y. Regulation of transport of the dopamine D1 receptor by a new membrane-associated ER protein. Nat Cell Biol. 2001 May;3(5):492–8.
289. Jung J, Kim J, Roh SH, Jun I, Sampson RD, Gee HY, et al. The HSP70 co-chaperone DNAJC14 targets misfolded pendrin for unconventional protein secretion. Nat Commun. 2016;7(May 2015):11386.
290. Bonifati V. Genetics of Parkinson’s disease – state of the art, 2013. Parkinsonism Relat Disord. 2016 Dec 12;20:S23–8.

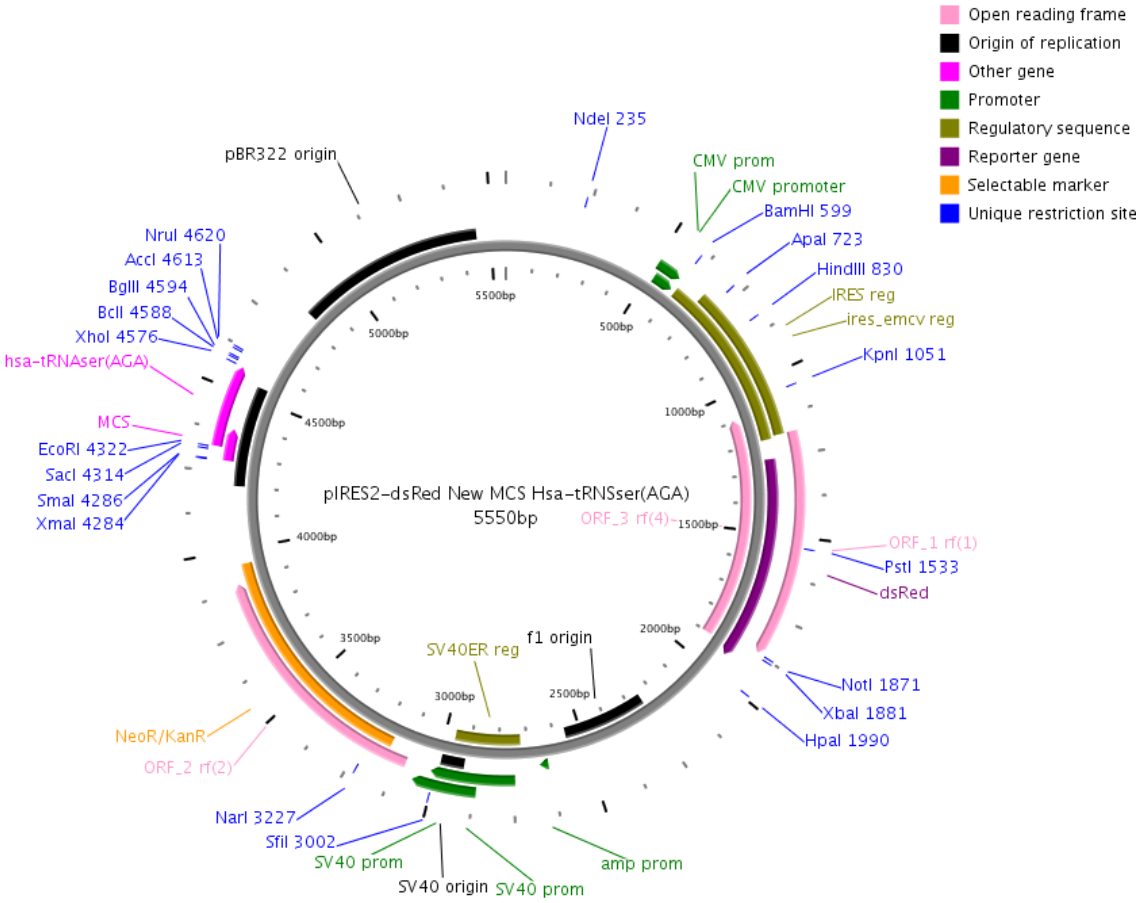
291. Lee DH, Pfeifer GP. Translesion synthesis of 7,8-dihydro-8-oxo-2′-deoxyguanosine by DNA polymerase eta in vivo. Mutat Res - Fundam Mol Mech Mutagen. 2008;641(1–2):19–26.
292. Leongamornlert D, Saunders E, Dadaev T, Tymrakiewicz M, Goh C, Jugurnauth-Little S, et al. Frequent germline deleterious mutations in DNA repair genes in familial prostate cancer cases are associated with advanced disease. Br J Cancer. 2014;110(6):1663–72.
293. Clancy S. DNA Damage & Repair: Mechanisms for Maintaining DNA Integrity. Nat Educ. 2008;1(1):103.
294. Yin JY, Shen J, Dong ZZ, Huang Q, Zhong MZ, Feng DY, et al. Effect of eIF3a on response of lung cancer patients to platinum-based chemotherapy by regulating DNA repair. Clin Cancer Res. 2011;17(13):4600–9.
295. Lee ASY, Kranzusch PJ, Cate JHD. eIF3 targets cell-proliferation messenger RNAs for translational activation or repression. Nature. 2015;522:111–4.
296. Liu K, Lei Z, Yao H, Lei S, Zhao H. Impact of a Eukaryotic Translation Initiation Factor 3a Polymorphism on Susceptibility to Gastric Cancer. Med Princ Pract. 2016;25(5):461–5.
297. Yin J-Y, Dong Z, Liu Z-Q, Zhang J-T. Translational control gone awry: a new mechanism of tumorigenesis and novel targets of cancer treatments. Biosci Rep. 2010;31(1):1–15.
298. Navin NE. Cancer genomics: one cell at a time. Genome Biol. 2014;15(8):452.
299. Hou Y, Song L, Zhu P, Zhang B, Tao Y, Xu X, et al. Single-Cell Exome Sequencing and Monoclonal Evolution of a JAK2-Negative Myeloproliferative Neoplasm. Cell. 2012 Mar 2;148(5):873–85.
300. Xu X, Hou Y, Yin X, Bao L, Tang A, Song L, et al. Single-cell exome sequencing reveals single-nucleotide mutation characteristics of a kidney tumor. Cell. 2012;148(5):886–95.
301. Li H, Durbin R. Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics. 2010;26(5):589–95.
302. Li H. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. Bioinformatics. 2011;27(21):2987–93.
303. Wu J, Rutkowski DT, Dubois M, Swathirajan J, Saunders T, Wang J, et al. ATF6?? Optimizes Long-Term Endoplasmic Reticulum Function to Protect Cells from Chronic Stress. Dev Cell. 2007;13(3):351–64.
304. Shoulders MD, Ryno LM, Genereux JC, Moresco JJ, Tu PG, Wu C, et al. Stress-Independent Activation of XBP1s and/or ATF6 Reveals Three Functionally Diverse ER Proteostasis Environments. Cell Rep. 2013;3(4):1279–92.
305. Chen JJ, Genereux JC, Qu S, Hulleman JD, Shoulders MD, Wiseman RL. ATF6 activation reduces the secretion and extracellular aggregation of destabilized variants of an amyloidogenic protein. Chem Biol. 2014;21(11):1564–74.
306. Chiang WC, Hiramatsu N, Messah C, Kroeger H, Lin JH. Selective activation of ATF6 and PERK endoplasmic reticulum stress signaling pathways prevent mutant rhodopsin accumulation. Investig Ophthalmol Vis Sci. 2012;53(11):7159–66.
307. Yoshida H, Oku M, Suzuki M, Mori K. pXBP1(U) encoded in XBP1 pre-mRNA negatively regulates unfolded protein response activator pXBP1(S) in mammalian ER stress response. J Cell Biol. 2006;172(4):565–75.
308. Rajesh K, Krishnamoorthy J, Kazimierczak U, Tenkerian C, Papadakis a I, Wang S, et al. Phosphorylation of the translation initiation factor eIF2α at serine 51 determines the cell fate decisions of Akt in response to oxidative stress. Cell Death Dis. 2015;6(1):e1591.
309. Koromilas AE. Roles of the translation initiation factor eIF2α serine 51 phosphorylation in cancer formation and treatment. Biochim Biophys Acta - Gene Regul Mech. 2014;1849(7):871–80.
310. Vattem KM, Wek RC. Reinitiation involving upstream ORFs regulates ATF4 mRNA translation in mammalian cells. Proc Natl Acad Sci U S A. 2004;101(31):11269–74.
311. Ma Y, Hendershot LM. Delineation of a negative feedback regulatory loop that controls protein translation during endoplasmic reticulum stress. J Biol Chem. 2003;278(37):34864–

73.
312. Kojima E, Takeuchi A, Haneda M, Yagi A, Hasegawa T, Yamaki K ichi, et al. The function of GADD34 is a recovery from a shutoff of protein synthesis induced by ER stress: elucidation by GADD34-deficient mice. FASEB J. 2003;17(11):1573–5.
313. Hyrskyluoto A, Reijonen S, Kivinen J, Lindholm D, Korhonen L. GADD34 mediates cytoprotective autophagy in mutant huntingtin expressing cells via the mTOR pathway. Exp Cell Res. 2012;318(1):33–42.
314. Gingold H, Tehler D, Christoffersen NR, Nielsen MM, Asmar F, Kooistra SM, et al. A Dual Program for Translation Regulation in Cellular Proliferation and Differentiation. Cell. 2015 Jun;158(6):1281–92.
315. Roux PP, Topisirovic I. Regulation of mRNA translation by signaling pathways. Cold Spring Harb Perspect Biol. 2012;4(11).
316. Topisirovic I, Sonenberg N. Distinctive tRNA Repertoires in Proliferating versus Differentiating Cells. Cell. 2014;158(6):1238–9.
317. White RJ. Transcription by RNA polymerase III: more complex than we thought. Nat Rev Genet. 2011;12(7):459–63.
318. Shor B, Wu J, Shakey Q, Toral-Barza L, Shi C, Follettie M, et al. Requirement of the mTOR kinase for the regulation of Maf1 phosphorylation and control of RNA polymerase III-dependent transcription in cancer cells. J Biol Chem. 2010;285(20):15380–92.
319. Kim S, You S, Hwang D. Aminoacyl-tRNA synthetases and tumorigenesis: more than housekeeping. Nat Rev Cancer. 2011 Oct;11(10):708–18.
320. Liu R-Y, Dong Z, Liu J, Yin J-Y, Zhou L, Wu X, et al. Role of eIF3a in regulating cisplatin sensitivity and in translational control of nucleotide excision repair of nasopharyngeal carcinoma. Oncogene. 2011;30(48):4814–23.
321. Dong Z, Liu LH, Han B, Pincheira R, Zhang J-T. Role of eIF3 p170 in controlling synthesis of ribonucleotide reductase M2 and cell growth. Oncogene. 2004;23(December 2003):3790–801.
322. Dong Z, Liu Z, Cui P, Pincheira R, Yang Y, Liu J, et al. Role of eIF3a in regulating cell cycle progression. Exp Cell Res. 2009;315(11):1889–94.
323. Saletta F, Rahmanto YS, Richardson DR. The translational regulator eIF3a: The tricky eIF3 subunit! Vol. 1806, Biochimica et Biophysica Acta - Reviews on Cancer. 2010. p. 275–86.
324. Wang W, Yao X, Huang Y, Hu X, Liu R, Hou D, et al. Mediator MED23 regulates basal transcription in vivo via an interaction with P-TEFb. Transcription. 2013;4(1):39–51.
325. Kim TW, Kwon Y-J, Kim JM, Song Y-H, Kim SN, Kim Y-J. MED16 and MED23 of Mediator are coactivators of lipopolysaccharide- and heat-shock-induced transcriptional activators. Proc Natl Acad Sci U S A. 2004;101(33):12153–8.
326. Park JM, Werner J, Kim JM, Lis JT, Kim YJ. Mediator, not holoenzyme, is directly recruited to the heat shock promoter by HSF upon heat shock. Mol Cell. 2001;8(1):9–19.
327. Iben JR, Maraia RJ. tRNA gene copy number variation in humans. Gene. 2014;536(2):376–84.
328. Fedyunin I, Lehnhardt L, Böhmer N, Kaufmann P, Zhang G, Ignatova Z. TRNA concentration fine tunes protein solubility. FEBS Lett. 2012;586(19):3336–40.
329. Jakubowski H, Goldman E. Editing of errors in selection of amino acids for protein synthesis. Microbiol Rev. 1992;56:412–29.
330. Schimmel P. Mistranslation and its control by tRNA synthetases. Philos Trans R Soc Lond B Biol Sci. 2011 Oct;366(1580):2965–71.
331. Pasman Z, Robey-Bond S, Mirando AC, Smith GJ, Lague A, Francklyn CS. Substrate specificity and catalysis by the editing active site of alanyl-tRNA synthetase from Escherichia coli. Biochemistry. 2011 Mar 8;50(9):1474–82.
332. Grzmil M, Hemmings BA. Translation regulation as a therapeutic target in cancer. Cancer Res. 2012 Aug;72(16):3891–900.
333. Spinola M, Galvan A, Pignatiello C, Conti B, Pastorino U, Nicander B, et al. Identification and functional characterization of the candidate tumor suppressor gene TRIT1 in human

lung cancer. Oncogene. 2005 Aug;24(35):5502–9.
334. Coelho M, Lade SJ, Alberti S, Gross T, Tolić IM. Fusion of Protein Aggregates Facilitates Asymmetric Damage Segregation. PLoS Biol. 2014;12(6):1–11.
335. Coelho M, Tolić IM. Asymmetric damage segregation at cell division via protein aggregate fusion and attachment to organelles. Bioessays. 2015;37(7):740–7.
336. 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. Proc Natl Acad Sci U S A. 2011;108(23):9378–83.
337. Geslain R, Pan T. Functional Analysis of Human tRNA Isodecoders. J Mol Biol. 2010;396(3):821–31.
338. Servier Medical Art [Internet]. Available from: http://www.servier.com/Powerpoint-image-bank

7. Annexes

Annex A - Map of the plasmid



Annex B – List of genes with Transitions in mistranslating cell lines

tRNA^{Ser}(S) - *PABPC1P5, ROS1, RP1-179P9.3, GPR6, CGB8, SASH1, RP11-236F9.2, CGB5, SLC25A14, PABPC1, CTB-60B18.17, CLUHP10, TRBV6-5, TRBV7-4, KRTAP9-9, RNU6-55P, MBNL2, RNA5SP37, SNRPEP6, NDUFV2, RP11-21J18.1, NDUFV2-AS1, HDDC2, GJE1, RNU6-1239P, POTEM, CTD-2311B13.5, USP17L18, USP17L17, CDK19, DAB2IP, LCN1P2*

tRNA^{Ser}(A) - *CGB8, SASH1, CGB5, CTB-60B18.17, RP11-301A5.2, NKAIN2, UPK3BP1, DCBLD1, POTEE, ARID4A, RNU6-55P, C17orf98, SNRPEP6, TUBB8P10, FAM166AP5, RP11-96F8.1, RSU1P2, CUBNP3, RP11-285G1.15, IGHVII-1-1, RP11-747H12.5, RPL7AP34, PTP4A1, PRELID1P1, IL22RA2, ADAT2, MTRF1L, RP1-101K10.6, NDUFV2, RP11-21J18.1, NDUFV2-AS1, HDDC2, STAC, RWDD1, HSF2, RPL23AP50, RNU6-437P*

tRNA^{Ser}(L) - *ANKRD20A9P, RP1-163M9.7, POTEF, CDC27P1, ANKRD30BL, ARFGEF3, PBOV1, CLUHP10, TRBV6-5, TRBV7-4, UPK3BP1, ARID4A, TMEM183A, KRT16P1, RNU6-55P, C17orf98, STAC, RWDD1, HSF2, LRRC10B, MIR4488, COLEC12, RP1-142L7.9, GJE1*

tRNA^{Ser}(H) - *PRB1, ADGRE5, PRB2, TNPO1P2, DYTN, ANKRD20A9P, GRK7, KRT31, CTB-60B18.17, CLUHP10, RP11-301A5.2, TRBV6-5, TRBV7-4, RP11-364B14.1, NECAP1, RNU6-537P, RP11-347H15.1, RP11-69M1.6, MAPK10, RP11-288H12.3, RP11-104G3.2, GLIDR, ABCD1, BCAP31, DNASE1L1, RPL10, XX-FW83563B9.5, TMEM183A, KRT16P1, RP11-23F23.3, RNU6-55P, MBNL2, RNA5SP37, C17orf98, SNRPEP6*

Annex C – List of genes with Transversions in mistranslating cell lines

tRNA^{Ser}(S) - *OR8U1, VWF, FRG1GP, FANCD2, NBPF20, CTD-3157E16.2, PABPC3, FOXE1, MAN1A1, ZNF75D, RNU7-195P, NIFKP8, RBMX, SNAP91, PFN1P3, RP11-1281K21.2, CBWD5, PPP1R14C, RP13-210D15.8, CSPG4P12, CHEK2*

tRNA^{Ser}(A) - *SIM1, PDXDC2P, KCNJ12, ZYG11B, FRG1GP, PSPC1P1, NBPF20, CTD-3157E16.2, PCSK6-AS1, ZNF890P, RP11-419C5.2, CTB-109A12.1, ZC3H18, KLF3, SAMD5, MAN1A1, ADGB, ZNF75D, ZNRF2P2, RP11-419C5.3, ZNF676, SEPP1, C17orf53, BCRP2, KB-1592A4.15, TBX18, FRK, SHANK3, RPS13P2, PFN1P3, RP11-1281K21.2, CBWD5, IGKV1-8, RPL12P4, CTD-2366F13.2, RP3-335E1.1, STX7, ZC3H12D, PPP1R14C, RP13-210D15.8, OR10AH1P, LSM1, RNU6-323P, NPM1P8*

tRNA^{Ser}(L) - *PDXDC2P, OR8U1, MZT1, ZNF441, HSPA8P8, CTD-2060L22.1, ITM2A, ETFA, TYRO3P, RP11-419C5.2, CTB-109A12.1, KHDRBS2, MAN1A1, ZNF75D, MED23, RNU7-195P, RP11-419C5.3, CTD-2561J22.2, CTD-2561J22.1, SEPP1, OPHN1, C17orf53, TTC39B, SUMO2P4, RPS13P2, PPP1R14C, RP13-210D15.8, NPM1P8, CDK17, RNU6-80P, MIR2113, RP11-436D23.1, HEY2, RP11-624M8.1, FBXO5*

tRNA^{Ser}(H) - *ZNF774, CTBP2P4, KCNJ12, OR8U1, GABARAPL3, B3GNTL1, MAP4K4, TTLL12, PLXNA4, PARP12, ARR3, RAB41, PDZD11, PCSK6-AS1, ZNF890P, KHDRBS2, ZNF75D, RNU7-195P, PABPC1P5, SLC9B1P4, BPI, CTD-2308N23.2, SEPP1, PABPC1P9, BRAFP1, RNA5SP283, LINC01410, RP11-23F23.3, RPS13P2, PPP1R14C, RP13-210D15.8, OR10AH1P, LSM1, RNU6-323P, TSPO, RARRES2P1, AGGF1P3, RP11-489M13.2, CCNJL, RP11-826N14.1, ZNRF2P1*

Annex D – List of genes with Insertions in mistranslating cell lines

tRNA^{Ser}(S) - *OLFM1, APP, DISP3, DENND2D, CHI3L2, RIIAD1, NANOS1, EIF3A, MRPL23, VEZT, LRRK1, MYOM1, ZNF805, CTC-444N24.8, LINC00461, WRN, PTPRD, MYSM1, MSH4, IFI44, RP4-641G12.3, PPP1R13B, ZFYVE21, MUC16, GNA12, AMZ1, STAU2, RP11-463D19.1, DNHD1, C12orf56, RPS11P6, ABCC13*

tRNA^{Ser}(A) - *SNRPC, CTBP2P1, MYSM1, DNHD1, C12orf56, RPS11P6, ABCC13, PADI6, ST6GALNAC3, STX6, RP11-46A10.5, RP11-46A10.8, FAM13C, STYK1, SNORD114-7, SNORD114-10, SNORD114-11, SNORD114-9, SNORD114-13, SNORD114-12, SNHG23, ALOX15, DNAH2, SPACA3, MPND, AC007292.3, AC007292.4, EIF1P6, ALK, CLIP4, PDCL3, CST1, SH3BGR, TMPRSS2, TMEM110, TMEM110-MUSTN1, MIR8064, RP11-894J14.2, CRYBG3, NUP54, TCERG1, GPR151, OFCC1, RP11-23J9.4, CCDC180, PSMD9, WDR66, RP11-87C12.2, FN3KRP, RP11-388C12.1, AC024361.1, LRP5L, CTA-246H3.11, L3MBTL2, RP4-756G23.5, ZNF717, MIR4273, RTN4, EML6*

tRNA^{Ser}(L) - *MYSM1, ABCC13, PSMD9, WDR66, RP11-87C12.2, FN3KRP, RP11-388C12.1, AC024361.1, LRP5L, CTA-246H3.11, L3MBTL2, RP4-756G23.5, ZNF717, MIR4273, A2ML1, RP11-259O18.5, CIDEB, NOP9, DHRS1, SAMD4A, MYH10, NDEL1, NEB, CXADR, AP000345.1, TRIOBP, TRNT1, IL5RA, PDHB, PXK, RP11-802O23.3, NKD2, LYSMD3, POLR3G, TSPYL1, DSE, RP1-93H18.1, ADAM9, NOL6, MIR6851, ORMDL2, DNAJC14, SARNP, RP11-762I7.5, CST2*

tRNA^{Ser}(H) - *CPT1A, OLFM1, MSH4, IFI44, RP4-641G12.3, PPP1R13B, ZFYVE21, MUC16, GNA12, AMZ1, STAU2, RP11-463D19.1, ABCC13, RTN4, EML6, ORMDL2, DNAJC14, SARNP, RP11-762I7.5, CST2, NVL, ART1, ART5, TRPC2, OR9G4, OR9G3P, RP11-100N3.2, MIR6128, KCNA6, GALNT8, RP11-234B24.4, OR7E47P, RP1-288H2.2, CIT, TDRD3, KHNYN, SDR39U1, AQR, RP11-83J16.1, ARIH1, RP11-1006G14.4, BLM, SLC5A11, RTTN, KIAA2012, DGKD, RNF114, ISY1, ISY1-RAB43, PRKAA1, STEAP2*

Annex E – Paper (Manuscript): Mutant tRNAs are selected in tumors and accelerate tumor growth in mice

Disclosure of Interests: The author of this thesis has contributed to the following experiments of this work in collaboration with the first author of the paper, Mafalda Santos:

1) Development of the technique that allowed for the detection of tRNAs (SNaPshot) in both cell lines and tumors;

2) Preparation of peptides from NIH3T3 derived tumors for Mass Spectrometry;

3) TNF alpha assay;

4) The author has also contributed to the writing and revision of the presented manuscript.

Patricia M. Pereira[a,b,1], Mafalda Santos[a,b,c,1], A. Sofia Varanda[a,b,c], Joana Carvalho[b,c], Mafalda Azevedo[b], Denisa D. Mateus[b], Nuno Mendes[b,c], Patricia Oliveira[b,c], Fábio Trindade[a,d] Marta Teixeira Pinto[b,c], Renata Bordeira-Carriço[b,c], Fátima Carneiro[b,c,e], Rui Vitorino[a], Carla Oliveira[b,c,e,2]* and Manuel A. S. Santos[a,2,]*

[a]Department of Medical Sciences and Institute of Biomedicine – iBiMED, University of Aveiro, 3810-193 Aveiro, Portugal. [b]Expression Regulation in Cancer, Institute of Molecular Pathology and Immunology, University of Porto (IPATIMUP), 4200-465 Porto, Portugal. [c]Instituto de Investigação e Inovação em Saúde, University of Porto, Portugal. [d]Department of Physiology and Cardiothoracic Surgery, Faculty of Medicine, University of Porto, Portugal. [e]Dept. Pathology and Oncology, Faculty of Medicine, University of Porto, 4200-465 Porto, Portugal
[1]Co-first authors;[2]Co-senior authors

Keywords: Cancer, tumor evolution, mRNA mistranslation, tRNAs, UPR, protein biosynthesis errors.

Abstract

The upregulation of protein synthesis, deregulation of tRNA expression and amino acid starvation are common features of human cancer. The occurrence of any of these mechanisms, or their concomitance, raises the hypothesis that translational fidelity is compromised in tumors. However, little is known about the relevance of protein synthesis errors (mistranslation) in cancer initiation and development. To clarify this issue, we expressed mutant tRNAs that misincorporate amino acids in non-cognate codons in mouse NIH3T3 cells and assessed their tumorigenic capacity *in vitro* and *in vivo*. These tRNAs were well tolerated, but in transformed

cells, stimulated angiogenesis and produced fast growing tumors in mice. Their expression increased sharply during tumor development, highlighting unanticipated advantages of mutant tRNAs in cancer growth. Since mistranslation diversifies the proteome, produces high phenotypic variation and drug resistance in various eukaryotic cell models, we postulate that tRNA decoding adds an important layer of complexity to tumor characterization.

Significance Statement: Increased rate of protein synthesis is normally associated with cell proliferation and cancer. In this study, we go beyond classical translational regulation studies to show important mRNA mistranslation effects on tumor growth, activation of the unfolded protein response and cancer signaling pathways in mice. Since protein synthesis errors diversify the proteome, produce cell population heterogeneity and genome instability, our work highlights unanticipated roles of tRNA misreading in tumor heterogeneity and evolution.

Introduction

Cancer is a multifactorial disease driven by accumulation of DNA mutations, chromosomal aberrations and epigenetic alterations (1). Transcriptional, post-transcriptional and translational deregulation is also well established; however little is known about the contribution of protein synthesis errors to tumor initiation and growth. Eukaryotic cells translate mRNAs (messenger RNAs) with basal error levels of 10^{-3} to 10^{-4} amino acid misincorporations per codon, after protein quality control (PQC) (2,3). Cells cope relatively well with this level of aberrant protein synthesis and it is unclear (controversial) whether such errors play any role in cell degeneration, aging or disease (4–7). Recent reports show, however, that small increase in translational errors leads to accumulation of misfolded proteins (4,8–10), saturation of PQC, proteotoxic stress, re-wiring of chaperone-clients interaction networks, and wide deregulation of cellular functions (9). Since translational fidelity depends mainly on tight regulation of tRNAs, aminoacyl-tRNA synthetases (aaRSs), RNA modifying enzymes, translation elongation factors, ribosome function and amino acid supply (11), we hypothesize that the deregulation of these translational factors in tumors (12–19) has profound effects on cancer biology and etiopathology. Indeed, aaRS are normally upregulated to sustain the high rate of protein synthesis and cell proliferation in tumors (20). Although the available methods for tRNA prediction assign the tRNA identity solely based on the anticodon, making it difficult to detect mutations in the anticodon, there are already some reports associating mitochondrial tRNA mutations with cancer (21–23).

Activation of the unfolded protein response (UPR) is consistent with deregulation of tRNA expression, and processing, amino acid starvation and aberrant protein synthesis (24–27).

Mistranslated proteins sequester BiP, activate the UPR sensors PERK, IRE-1 and ATF6 and upregulate molecular chaperones and the ubiquitin-proteasome system (24). Tumor cells highjack these endoplasmic reticulum (ER) adaptive measures to thrive (28), explaining, at least in part, the UPR association with malignancy and aggressiveness of several types of cancer (29–31).

With this work, we aimed at disentangling the role of mRNA mistranslation from other protein synthesis and protein quality control confounding effects. Our main objective was to clarify if mistranslation alone is sufficient to sustain cell transformation, tumor growth and UPR activation. Geslain *et al*. developed a method for inducing protein synthesis errors using mutant tRNAs that introduce serine (Ser) in non-cognate codons in transiently transfected human cells (32). Taking advantage of this model, we constructed mutant tRNAs that misincorporate Ser at alanine (Ala) and leucine (Leu) codons, and stably expressed them in the near-normal NIH3T3 cell line. We also created a cell line that expresses the normal Ser tRNA to see alterations caused by tRNA pool deregulation. In our model, there were no visible effects on cell viability, but cell transformation, angiogenesis, tumor growth, activation of the UPR and other cancer-related pathways, were clearly visible. Remarkably, expression of the mutant tRNAs was sharply selected in tumors, indicating that some types of mistranslation might be advantageous during tumor growth.

Results

Mistranslation in mammalian cell lines

The construction of mutant tRNAs that misincorporate Ser at Ala-GCU,GCC ($tRNA^{Ser}$(Ala)), and Leu-CUU,CUC ($tRNA^{Ser}$Leu) codons (Fig.1A, left panel) was possible due to the inability of the seryl-tRNA synthetase (SerRS) to recognize the anticodon of $tRNA^{Ser}_{AGA}$, i.e., introduction of mutations in this structural domain do not interfere with serylation (32,33). Those anticodons were chosen to test putative phenotypic diversity arising from Ser misincorporation at chemically distinct protein sites. Mutant tRNA genes and the Wt $tRNA^{Ser}$ ($tRNA^{Ser}_{AGA}$) were cloned into the pIRES2-DsRed and transfected into NIH3T3 cells. Cell lines stably expressing these tRNAs were then selected for phenotypic characterization (Fig.1A, right panel). Cells transfected with the empty vector (Mock) was used as a control. Transfection efficiency was determined by Real-Time PCR, using the pIRES2 *DsRed* gene as a readout probe. The data showed 100% efficiency for all cell lines but the $tRNA^{Ser}$(Leu) cell line, where 72% efficiency was observed.

The incorporation of the mutant tRNA genes in the transfected cells was confirmed by Polymerase Chain Reaction (PCR) and Sanger DNA sequencing of genomic DNA samples extracted from the selected cell lines (Fig. S1A). tRNA expression was determined using a primer extension assay (SNaPshot analysis) that detected each mutant tRNA and the endogenous copies of the $tRNA^{Ser}$ gene (n=12 genes). The cellular level of the endogenous $tRNA^{Ser}$ was 19.4-fold higher than the mutant $tRNA_{AGC}^{Ser}$(Ala) and 49.5-fold higher than the $tRNA_{AAG}^{Ser}$(Leu) (Fig.1B). This low expression level was expected to influence cell behavior, as previous reports showed that modest deregulation of endogenous tRNAs is sufficient to produce phenotypic effects in mammalian cells (17).

Cellular consequences of mRNA mistranslation

We used cell viability, proliferation and apoptosis assays to evaluate the phenotypic consequences of the mutant tRNAs and misexpression of Wt $tRNA^{Ser}$. Trypan Blue staining showed no impact on viability (Fig.S1B) and the Annexin V Apoptosis assay showed a basal necrosis level (≤1% of cells) and low percentage of cells in late (ca. 5% of cells) and early apoptosis (7-8% of cells) (Fig.1C). Cell proliferation was also not significantly affected (Fig.S1C) and cell cycle progression demonstrated a similar pattern in all cell lines (Fig.S1D). These data suggested that NIH3T3 cells are highly tolerant to the mutant tRNAs and misexpression of Wt $tRNA^{Ser}$, however there seems to be an effect on cell transformation, as measured by the foci formation assay. Mutant tRNA expressing cell lines produced more foci than the negative control cells (Mock) (Fig.1D), suggesting that mutant tRNAs were able to transform NIH3T3 cells.

Effect of mutant tRNAs and misexpression of Wt $tRNA^{Ser}$ in tumor formation in vivo

To clarify whether mutant tRNAs or the misexpression of Wt $tRNA^{Ser}$ could induce angiogenesis and tumorigenesis *in vivo,* we used the chick chorioallantoic membrane assay (CAM). Only $tRNA^{Ser}$(Ala) cells, produced larger tumors and had stronger angiogenic response in the CAM assay (Fig.2A). These results were further confirmed by inoculating cells expressing K-ras^{V12}, the mutant tRNAs ($tRNA^{Ser}$(Ala) or $tRNA^{Ser}$(Leu)) or misexpressing the Wt $tRNA^{Ser}$ ($tRNA^{Ser}$(WT)), on the left dorsal flank of at least five mice (for each cell line) and control cells (Mock) on the corresponding right dorsal flank of every mice. Within 14 to 21 days post-inoculation (p.i.), tumors were produced by cells expressing K-ras^{V12} (5/5 mice), $tRNA^{Ser}$(Ala) (6/6 mice) and $tRNA^{Ser}$(Leu) (5/5 mice) (Fig.2B, upper panel). At day 27 p.i., 12/21 (57.1%)

mice inoculated with Mock cells and 5/5 mice inoculated with $tRNA^{Ser}$(WT) cells developed small similar size tumors. At this stage, tumors produced by cells expressing $tRNA^{Ser}$(Ala) were the largest. At day 31 p.i., the experiment was terminated. $tRNA^{Ser}$(Ala) and K-ras^{V12} tumors were similar in size distribution and were different from Mock tumors (p<0.01) (Fig.2B, upper and middle panel).

Histological characterization of resected tumors unveiled high grade sarcomas with high proliferative index, as determined by Ki67 labeling (Fig.2B, lower panel, Fig.S2C). Histopathological analysis of murine organs (ganglion, lung, kidney, liver, bladder, pleura and stomach), collected at day 31 p.i., revealed the presence of lung metastases in K-ras^{V12} expressing tumors, and no metastases in all other mice. Therefore, the expression of mutant tRNA in $tRNA^{Ser}$(Ala) cells accelerated tumor growth, in agreement with the CAM assay results, but Ser misincorpoartion at Leu codons and misexpression of $tRNA^{Ser}$ also induced formation of smaller tumors.

DNA extracted from tumors, from both CAM and mice experiments, were sequenced and genomic incorporation of all plasmids was validated (Fig.S2A,B). tRNA expression in mice tumors was determined using the primer extension assay described above. Surprisingly, expression levels of mutant tRNAs were much higher in tumors than in the corresponding cell lines, i.e., 8 and 8.4 fold higher for the $tRNA^{Ser}$(Ala) and $tRNA^{Ser}$(Leu), respectively (Fig.2C). In other words, the mutant tRNAs were advantageous for tumor development (Fig.1B,2C).

We analysed the Soluble (SF) and Insoluble Fraction (IF) of proteins from tumors derived from our cell lines to confirm misincorporations by Mass Spectrometry. Overall, we detected more proteins with misincorporations (Ala→Ser and Leu→Ser) in the SF of tumor proteins. Both in SF and IF, $tRNA^{Ser}$(Ala) derived tumor presented more distinct proteins with misincorporations than the rest of the tumors (Fig. S3A). We further investigated whether the misincorporations detected corresponded to the anticodons of our tRNAs. In the IF, the misincorporations detected did not correspond to our tRNAs, occurring randomly in Ala or Leu codons. However, in the SF, there was increased incorporation of Ser at Ala sites in the $tRNA^{Ser}$(Ala) cells (Fig. S3B). This proves that our mutant tRNA is being recruited by the ribosome and that it is able to induce incorporation of Ser at the correspondent Ala sites (GCC and GCU). Using this approach, we could not detect an increase of Leu→Ser misincorporations. Since this misincorporation disrupts protein conformation, probably proteins that are being formed are also being degraded by the proteasome or autophagy, being this misincorporation difficult to detect.

Impact of mutant tRNAs and misexpression of Wt tRNA^{Ser} in the UPR

Since the UPR is frequently activated, highly relevant in cancer and an end point of protein mistranslation (34), we have tested whether it was also activated in our model. For this, we have monitored biomarkers of the three UPR branches, namely IRE-1, PERK and ATF6. Activation of the IRE-1 pathway was determined by measuring splicing levels of the XBP-1 transcription factor, using RT-PCR. Comparing to Mock tumors, its activation was increased 7% and 14% in tRNA^{Ser}(Ala) and tRNA^{Ser}(Leu) tumors, respectively (Fig.S4). tRNA^{Ser}(Ala) and tRNA^{Ser} tumors showed higher levels of ATF6 activation than Mock tumors (3.1- and 2.15-fold, respectively) (Fig.3A,C). We next assessed the phosphorylation status of eIF2α, the downstream target of PERK, to clarify whether these tRNAs affected translation initiation rate. The levels of eIF2α-P (the inactive form of eIF2α) were 77% lower in tRNA^{Ser}(Ala) tumors, comparing to Mock, and did not change in other tumors (Fig.3B,C). This raised the hypothesis that PERK could be downregulated or, alternatively, the regulatory subunit of the PP1 phosphatase GADD34 could be upregulated. Western blot analysis showed 6-fold upregulation of GADD34 (Fig.3B,C), confirming that the fast growth rate of tRNA^{Ser}(Ala) tumors was likely due to upregulation of protein synthesis through dephosphorylation eIF2α by the GADD34-PP1 complex.

The effect of misreading tRNAs in cancer-associated signaling pathways

Serine, threonine and tyrosine tRNAs are among the most overexpressed tRNAs in breast cancer (16). Since these amino acids can be phosphorylated, their misincorporation at non-cognate sites may cause aberrant phosphorylation and alteration of signaling transduction pathways (16). This lead us to hypothesize that Ras/Raf/MEK/ERK and Ras/PI3K/PTEN/Akt signaling pathways could be affected in our model system, promoting unrestrained cellular growth, proliferation and tumor formation (35). Indeed, global serine phosphorylation was increased in tRNA^{Ser}(Ala) and tRNA^{Ser}(Leu) tumors (1.53 and 1.71 fold, respectively) (Fig.4A), confirming that cell signaling could be deregulated. We then analyzed the activation of Akt, ERK1/2 and p38 in the same tumors and observed activation of the Akt pathway in all of our tumors (Fig.4B), and downregulation of the ERK1 (64%) and ERK2 (54%) pathways in tRNA^{Ser}(Leu) tumors (Fig.S5A,C). p38 was downregulated 82% in tRNA^{Ser}(WT) relative to Mock, but was unchanged in the other tumors (Fig.S4B,C). Therefore, tumorigenesis induced by mutant tRNAs and misexpression of tRNA^{Ser} is likely associated with activation of the Akt pathway, while growth rate differences between tumors could be linked to differential activation

of ERK1/2.

Microenvironment effects in mistranslation outcomes

Since the expression of mutant tRNAs and misexpression of $tRNA^{Ser}$ were highly tolerated and did not activate the UPR or cancer-associated pathways *in vitro* (data not shown), we reasoned that external stimuli would be necessary to expose the phenotypes observed *in vivo*. To clarify this issue, we treated cell lines expressing Mock, $tRNA^{Ser}$ and the mutant tRNAs with the pro-inflammatory tumor necrosis factor alpha (TNF-α) for 30 min and 4 hours, and used Akt and p38 phosphorylation as phenotypic readouts (Akt-P/Akt and p38-P/p38) (Fig.5 and S6). The p38 pathway was activated in $tRNA^{Ser}$(Leu) cells after 30 minutes and persisted up to 4 hours of exposure (3.43 and 2.1-fold change, respectively). A slight activation was also observed in $tRNA^{Ser}$(Ala) cells at 30 minutes, but was lost after 4h (Fig.5). The Akt pathway was only significantly activated in $tRNA^{Ser}$(Ala) cells after 30 minutes (2.6-fold change) (Fig.5A). Despite the lack of complete data overlap between tumors and cell lines, these responses indicated that the microenvironment is crucial for tRNA mediated activation of cancer-related pathways and tumorigenesis.

Discussion

Imbalance of tRNA pools promotes the formation of non-cognate tRNA-aaRS pairs leading to tRNA mischarging (36). Pavon *et al*. reported an increase of certain tRNAs associated with malignant phenotypes and Gingold *et al.* reported enrichment of tRNAs required for the fast and accurate translation of proliferation associated genes in cancer (16,17,37). These studies support the role of protein synthesis regulation and translational fidelity effects in tumors. Moreover, mistranslation has clear impact on proteome and phenotypic diversification, drug tolerance and resistance in other biological models (7,38), suggesting that it may interfere with tumor growth, heterogeneity and response to therapy. To disentangle protein synthesis rate and regulation from mRNA mistranslation effects on cancer, we have expressed several mutant tRNAs and misexpress the Wt $tRNA^{Ser}$ in near-normal NIH3T3 cells. Such tRNAs were tolerated *in vitro* and did not produce visible effects on cell viability, apoptosis, proliferation and cell cycle progression, contrary to previous reports using transient models, but promoted angiogenesis and tumor growth *in vivo* (32). In particular, $tRNA^{Ser}$(Ala) cells, produced tumors that grew as fast as K-ras^{V12} tumors. Previous studies have shown that Ser misincorporation at Ala codon sites, arising from a single mutation that prevents Ser vs Ala discrimination by the editing site of the

AlaRS, induces rapid loss of purkinge cells, ataxia and premature death in mice (39). These apparently contradictory phenotypic outcomes suggest that mistranslation effects are cell type dependent, i.e., mistranslation in purkinge cells leads to apoptosis while in other cells leads to transformation and neoplasia. Indeed, the higher expression of mutant tRNAs in our mice tumors (Fig.2C), compared to the levels detected in cell lines (Fig.1B) indicates that mistranslation is adaptive in tumor contexts. Our group has previously shown that mistranslation in yeast increases tolerance to stress and allows for growth in presence of lethal doses of drugs and chemicals (40,41). These yeast cells also adapt very fast to mistranslation by altering genomic architecture, increasing protein synthesis, protein degradation and glucose uptake (10). Thus, the deleterious effects of mutant tRNA expression are rapidly mitigated through genomic, metabolic and proteomic changes, allowing mistranslating cells to thrive in specific microenvironments.

The $tRNA^{Ser}$(Ala) expressing tumors showed low levels of eIF2α-P, due to GADD34 upregulation and activation of cancer-associated pathways. This is highly relevant since decreased levels of eIF2α-P are sufficient to trigger NIH3T3 transformation (42). ATF6 and IRE-1 activation have also been documented in tumors (30,43), where the former is associated with cellular protection and growth (44). Furthermore, UPR coupled with induced tumor dormancy protects neoplasic cells from apoptosis and permits recurrence once favorable growth conditions are restored (45). Microenvironment signaling may also be responsible for accentuating this phenotype when cells are inoculated, allowing them to adapt and thrive. It is well established that tumor development needs genetic and epigenetic changes as well as cooperation of microenvironment components to promote adaptation and growth (46). Common adaptive responses include enhanced plasticity, cell motility, resistance to apoptosis and survival in hostile environments characterized by hypoxia, acidity, amino acid deprivation, inflammatory cytokines delivery, and induction of the UPR (47,48). Importantly, PERK activity and eIF2α-P levels are reduced in mouse breast tumors, where Akt is activated, and in cells exposed to ER or oxidative stress (49). In line with these results, tumors expressing mutant tRNAs showed concomitant activation of the Akt pathway, low levels of eIF2α-P and UPR induction. We postulate that this conjugation of factors may drive apoptosis evasion, cell survival and potentiate tumor growth (Fig.6).

Although remarkable progress is being made on the elucidation of the molecular basis of cancer, the etiology of most cancers is still unknown. In the past few years, new molecular links between cancer and translation deregulation have been unraveled, highlighting this setting as etiopathogenic (50). Our model supports this link by disclosing that mutant tRNAs are upregulated and selected for in tumors and are sufficient to promote fast tumor growth. Since mistranslation leads to random amino acid misincorporation into proteins, producing highly

heterogeneous and unstable proteomes, the tumor mistranslation landscape, and the effects of proteome alterations on tumor heterogeneity and resistance to anti-cancer drugs will be fascinating future avenues of research.

Methods

Experimental procedures are briefly described here and detailed description is provided in SI text. A DNA fragment of the gene encoding human $tRNA_{AGA}^{Ser}$ (Chr6 tRNA#5) and its flanking region was amplified by PCR from genomic DNA and cloned into the vector pIRES2-DsRed. Site-directed mutagenesis was performed to change the anticodon of $tRNA_{AGA}^{Ser}$ to other anticodons. A mouse Embryo Fibroblast cell line (NIH3T3) (American Type Culture Collection) was transfected with 1µg DNA plasmid using Lipofectamine 2000 (Invitrogen), following manufacturer's instructions. Stably transfected cell lines were established with 1000µg/ml G418 for 1 month. Cells were transfected with the empty vector (Mock), the plasmid with the wild-type $tRNA^{Ser}$(WT) and the mutant $tRNA_{AGC}^{Ser}$(Ala) and $tRNA_{AAG}^{Ser}$(Leu). To confirm the expression of mutant tRNAs in cell lines and tumors we amplified the cDNA of interest by PCR. PCR products were analyzed on 2% agarose gels and bands were excised and purified using Ilustra™ GFX™ PCR DNA and Gel Band purification Kit (GE). SNaPshot reactions were performed using SNaPshot Multiplex Ready Reaction Mix (Applied Biosystems). Samples were then sequenced and analyzed using Peak Scanner software (Applied Biosystems). To access the tumorigenic ability of Mock, $tRNA^{Ser}$(WT), $tRNA^{Ser}$(Ala) and $tRNA^{Ser}$(Leu) cells, the chicken embryo chorioallantoic membrane (CAM) model and tumor induction assay in mice were preformed. Two rings were placed in each CAM, one was filed with Mock cell suspension and the other with tRNA transfected cells ($1x10^6$ cells). After 3 days, the CAM was excised, photographed *ex ovo* in a stereoscope at 20x magnification (Olympus SZX16, DP71 camera). The number of new vessels was counted and tumor area was determined using Cell^A Olympus program. Regarding the mice, 6-8 weeks-old male nude mice were used for *in vivo* experiments, in accordance with the Guidelines for the Care and Use of Laboratory Animals, directive 2010/63/EU. Each mouse was injected in the right flank with Mock cells and in the left flank with $tRNA^{Ser}$(WT), $tRNA^{Ser}$(Ala) and $tRNA^{Ser}$(Leu) cells. Mice were weighted and tumor width and length were measured. Tumor volumes were calculated assuming ellipsoid growth patterns. Regarding the Mass Spectrometry approach, complete lanes were cut of SDS-PAGE gels, destained with 25 mM ammonium bicarbonate/50% acetonitrile and dried under vacuum (SpeedVac®, Thermo Savant, USA). The dried gel pieces were rehydrated with 10 µg/mL trypsin (Promega V5111) and digested overnight at 37 °C. Tryptic peptides were extracted from

the gel with 10% formic acid/ 50% acetonitrile and were then dried under vacuum. Separation of tryptic peptides by nano-HPLC was performed on the module separation Proexeon EASY-nLC 1000 from Thermo equipped with a 50-cm EASY C18 column with particle size 2-µm. Each sample was separated over a gradient of 5-32 % ACN in 90 at 250 nl/min. Peptide cations were converted to gas-phase ions by electrospray ionization and analyzed on a Thermo Orbitrap Fusion Lumos. Precursor scans were performed from 300 to 1,500 m/z at 120K resolution (at 445 m/z) using a 1 x 10^5 AGC target. Precursors selected for tandem MS were isolated at 1 Th with the quadrupole, fragmented by HCD with normalized collision energy of 30, and analyzed using rapid scan in the ion trap. The maximum injection time for MS2 analysis was 50 ms, with an AGC target of 1 x 10^4. Precursors with a charge state of 2-5 were sampled for MS2. Dynamic exclusion time was set at 60 seconds, with a 5ppm tolerance around the selected precursor. The raw files were searched directly against the *Mus musculus* database using PatternLab (Version 4.0.0.48). Search criteria included a static modification of +57.0214 Da on cysteine residues, variable modification of +15.9949 Da on oxidized methionine, +15.994915 on Ala and -26.052036 on Ile/Leu.

Acknowledgements

The authors wish to thank: Paulo Matos and Peter Jordan for the plasmid pEGFP containing K-ras^{V12} used in this study; Tiago Pedrosa for helping with flow cytometry analysis; Hugo Osório for technical support with MS/MS tentative approach; Ana Maria Magalhães for help with TNF-α assay; IPATIMUP group (ERIC) and RNA Biology group, especially Ana Soares.

Grant Support

Ipatimup, i3S and iBiMED are partially supported by the Portuguese Foundation for Science and Technology (FCT). This work was funded by: 1) FEDER funds through the Operational Programme for Competitiveness Factors–COMPETE; National Funds through the FCT, under the projects "PEst-C/SAU/LA0003/2013"; 2) NORTE-07-0162-FEDER-00018-"Contributos para o reforço da capacidade do Ipatimup enquanto actor do sistema regional de inovação" and NORTE-07-0162-FEDER-000067-Reforço e consolidação da capacidade infraestrutural do IPATIMUP para o sistema regional de inovação", supported by Programa Operacional Regional do Norte (ON.2–O Novo Norte), through FEDER funds under the Quadro de Referência Estratégico Nacional (QREN); 3)FCT/FEDER project FCT-ANR/IMI-

MIC/0041/2012, FCT UID/BIM/04501/2013 and FCT Fellowships [SFRH/BPD/26611/2006-PMP; SFRH/BPD/89764/2012-PO; SFRH/BPD/86543/2012-JC; SFRH/BD/91020/2012-MS and SFRH/BD/76417/2011-ASV).

References

1. You JS, Jones PA. Cancer genetics and epigenetics: two sides of the same coin? Cancer Cell. 2012 Jul;22(1):9–20.
2. Jakubowski H, Goldman E. Editing of errors in selection of amino acids for protein synthesis. Microbiol Rev. 1992;56:412–29.
3. Loftfield RB, Vanderjagt D. The frequency of errors in protein biosynthesis. Biochem J. 1972;128(5):1353–6.
4. Drummond DA, Wilke CO. The evolutionary consequences of erroneous protein synthesis. Nat Rev Genet. 2009;10(10):715–24.
5. Silva RM, Duarte ICN, Paredes JA, Lima-Costa T, Perrot M, Boucherie H, et al. The Yeast PNC1 longevity gene is up-regulated by mRNA mistranslation. PLoS One. 2009;4(4):e5212.
6. Pan T. Adaptive translation as a mechanism of stress response and adaptation. Annu Rev Genet. 2013;47:121–37.
7. Ribas de Pouplana L, Santos M, Zhu JH, Farabaugh PJ, Javid B. Protein mistranslation: Friend or foe? Trends Biochem Sci. 2014;39(8):355–62.
8. Bloom JD, Labthavikul ST, Otey CR, Arnold FH. Protein stability promotes evolvability. Proc Natl Acad Sci U S A. 2006;103:5869–74.
9. Paredes JA, Carreto L, Simões J, Bezerra AR, Gomes AC, Santamaria R, et al. Low level genome mistranslations deregulate the transcriptome and translatome and generate proteotoxic stress in yeast. BMC Biol. 2012 Jan;10:55.
10. Kalapis D, Bezerra AR, Farkas Z, Horvath P, Bódi Z, Daraba A, et al. Evolution of Robustness to Protein Mistranslation by Accelerated Protein Turnover. PLoS Biol. 2015;13(11).
11. Zaher HS, Green R. Fidelity at the molecular level: lessons from protein synthesis. Cell. 2009 Feb;136(4):746–62.
12. Bjornsti MA, Houghton PJ. Lost in translation: Dysregulation of cap-dependent translation and cancer. Vol. 5, Cancer Cell. 2004. p. 519–23.
13. Mamane Y, Petroulakis E, Rong L, Yoshida K, Ler LW, Sonenberg N. eIF4E - from translation to transformation. Oncogene. 2004 Apr;23(18):3172–9.
14. Pandolfi PP. Aberrant mRNA translation in cancer pathogenesis: an old concept revisited comes finally of age. Oncogene. 2004;23:3134–7.
15. Park SG, Schimmel P, Kim S. Aminoacyl tRNA synthetases and their connections to disease. Proc Natl Acad Sci U S A. 2008 Aug;105(32):11043–9.
16. Pavon-Eternod M, Gomes S, Geslain R, Dai Q, Rosner MR, Pan T. tRNA over-expression in breast cancer and functional consequences. Nucleic Acids Res. 2009 Nov;37(21):7268–80.
17. Pavon-Eternod M, Gomes S, Rosner MR, Pan T. Overexpression of initiator methionine tRNA leads to global reprogramming of tRNA expression and increased proliferation in human epithelial cells. RNA. 2013;19:461–6.
18. Ruggero D, Pandolfi PP. Does the ribosome translate cancer? Nat Rev Cancer. 2003;3:179–92.
19. White RJ. RNA polymerase III transcription and cancer. Oncogene.

2004;23(18):3208–16.
20. Grzmil M, Hemmings BA. Translation regulation as a therapeutic target in cancer. Cancer Res. 2012 Aug;72(16):3891–900.
21. Sahyoun AH, Holzer M, Juhling F, Höner zu Siederdissen C, Al-Arab M, Tout K, et al. Towards a comprehensive picture of alloacceptor tRNA remolding in metazoan mitochondrial genomes. Nucleic Acids Res. 2015 Sep 18;43(16):8044–56.
22. Wang L, Chen Z-J, Zhang Y-K, Le H-B. The role of mitochondrial tRNA mutations in lung cancer. Int J Clin Exp Med. 2015;8(8):13341–6.
23. Suzuki Y, Nishimaki K, Taniyama M, Muramatsu T, Atsumi Y, Matsuoka K, et al. Lipoma and opthalmoplegia in mitochondrial diabetes associated with small heteroplasmy level of 3243 tRNALeu(UUR) mutation. Diabetes Res Clin Pract. 2004;63(3):225–9.
24. Clarke HJ, Chambers JE, Liniker E, Marciniak SJ. Endoplasmic Reticulum Stress in Malignancy. Cancer Cell. 2014;25:563–73.
25. Torres AG, Batlle E, Ribas de Pouplana L. Role of tRNA modifications in human diseases. Trends Mol Med. 2014 Jun;20(6):306–14.
26. Rodriguez V, Chen Y, Elkahloun A, Dutra A, Pak E, Chandrasekharappa S. Chromosome 8 BAC array comparative genomic hybridization and expression analysis identify amplification and overexpression of TRMT12 in breast cancer. Genes Chromosom Cancer. 2007;46(7):694–707.
27. Spinola M, Galvan A, Pignatiello C, Conti B, Pastorino U, Nicander B, et al. Identification and functional characterization of the candidate tumor suppressor gene TRIT1 in human lung cancer. Oncogene. 2005 Aug;24(35):5502–9.
28. Maas NL, Diehl JA. The PERKs and Pitfalls of Targeting the Unfolded Protein Response in Cancer. Clin Cancer Res. 2015;21(4):675–9.
29. Papandreou I, Denko NC, Olson M, Van Melckebeke H, Lust S, Tam A, et al. Identification of an Ire1alpha endonuclease specific inhibitor with cytotoxic activity against human multiple myeloma. Blood. 2011 Jan;117(4):1311–4.
30. Schewe DM, Aguirre-Ghiso JA. ATF6alpha-Rheb-mTOR signaling promotes survival of dormant tumor cells in vivo. Proc Natl Acad Sci U S A. 2008;105:10519–24.
31. Bi M, Naczki C, Koritzinsky M, Fels D, Blais J, Hu N, et al. ER stress-regulated translation increases tolerance to extreme hypoxia and promotes tumor growth. EMBO J. 2005;24:3470–81.
32. Geslain R, Cubells L, Bori-Sanz T, Alvarez-Medina R, Rossell D, Martí E, et al. Chimeric tRNAs as tools to induce proteome damage and identify components of stress responses. Nucleic Acids Res. 2010;38:e30.
33. Cusack S. Aminoacyl-tRNA synthetases. Curr Opin Struct Biol. 1997;7:881–9.
34. Wang S, Kaufman RJ. The impact of the unfolded protein response on human disease. J Cell Biol. 2012;197:857–67.
35. Chappell WH, Steelman LS, Long JM, Kempf RC, Abrams SL, Franklin RA, et al. Ras/Raf/MEK/ERK and PI3K/PTEN/Akt/mTOR inhibitors: rationale and importance to inhibiting these pathways in human health. Oncotarget. 2011;2:135–64.
36. Dittmar KA, Sørensen MA, Elf J, Ehrenberg M, Pan T. Selective charging of tRNA isoacceptors induced by amino-acid starvation. EMBO Rep. 2005;6(2):151–7.
37. Gingold H, Tehler D, Christoffersen NR, Nielsen MM, Asmar F, Kooistra SM, et al. A Dual Program for Translation Regulation in Cellular Proliferation and Differentiation. Cell. 2015 Jun;158(6):1281–92.
38. Bezerra AR, Simões 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. Proc Natl Acad Sci U S A. 2013;110:11079–84.
39. 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 Sep;443(7107):50–5.
40. Silva RM, Paredes JA, Moura GR, Manadas B, Costa TL, Rocha R, et al. Critical roles for a genetic code alteration in the evolution of the genus Candida. EMBO J. 2007;26(21):4555–65.
41. Santos MAS, 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. Mol Microbiol. 1999 Feb;31(3):937–47.
42. Perkins D, Barber G. Defects in Translational Regulation Mediated by the Subunit of Eukaryotic Initiation Factor 2 Inhibit Antiviral Activity and Facilitate the Malignant Transformation of Human Fibroblasts. Mol Cell Biol. 2004;24(5):2025–40.
43. Romero-Ramirez L, Cao H, Nelson D, Hammond E, Lee A-H, Yoshida H, et al. XBP1 is essential for survival under hypoxic conditions and is required for tumor growth. Cancer Res. 2004 Sep;64(17):5943–7.
44. Belmont PJ, Tadimalla A, Chen WJ, Martindale JJ, Thuerauf DJ, Marcinko M, et al. Coordination of growth and endoplasmic reticulum stress signaling by regulator of calcineurin 1 (RCAN1), a novel ATF6-inducible gene. J Biol Chem. 2008;283:14012–21.
45. Dong D, Stapleton C, Luo B, Xiong S, Ye W, Zhang Y, et al. A critical role for GRP78/BiP in the tumor microenvironment for neovascularization during tumor growth and metastasis. Cancer Res. 2011 Apr;71(8):2848–57.
46. Taddei ML, Giannoni E, Comito G, Chiarugi P. Microenvironment and tumor cell plasticity: An easy way out. Vol. 341, Cancer Letters. 2013. p. 80–96.
47. Clarke R, Cook KL, Hu R, Facey COB, Tavassoly I, Schwartz JL, et al. Endoplasmic reticulum stress, the unfolded protein response, autophagy, and the integrated regulation of breast cancer cell fate. Vol. 72, Cancer Research. 2012. p. 1321–31.
48. Fels DR, Koumenis C. The PERK/eIF2alpha/ATF4 module of the UPR in hypoxia resistance and tumor growth. Vol. 5, Cancer Biology and Therapy. 2006. p. 723–8.
49. Mounir Z, Krishnamoorthy JL, Wang S, Papadopoulou B, Campbell S, Muller WJ, et al. Akt determines cell fate through inhibition of the PERK-eIF2α phosphorylation pathway. Sci Signal. 2011 Sep;4(192):ra62.
50. Sang Won L, Young Sun K, Kim S. Multifunctional Proteins in Tumorigenesis: Aminoacyl-tRNA Synthetases and Translational Components. Curr Proteomics. 2006;3(4):233–47.

Figures:

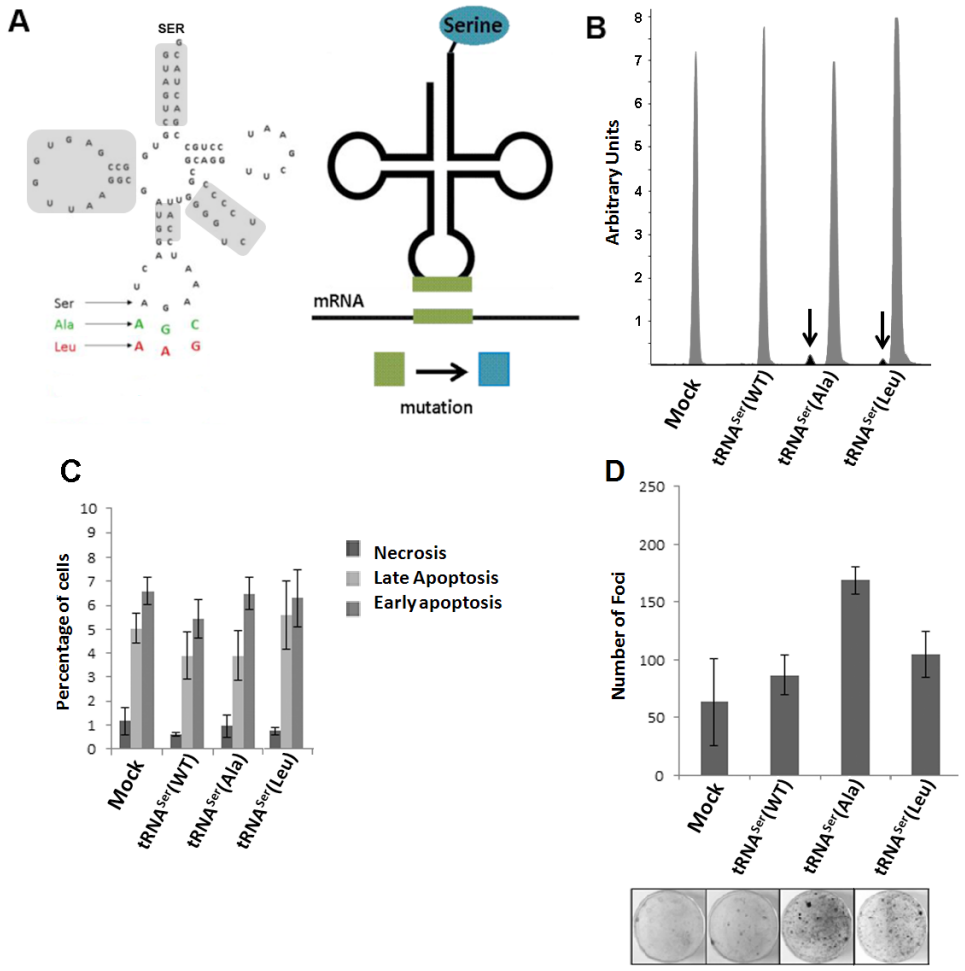


Figure 1. Schematic model and *in vitro* phenotypic effects of mutant tRNAs. **A)** Schematic model. Domains highlighted in grey are important for $tRNA^{Ser}$ recognition by SerRS. **Left panel)** The human $tRNA_{AGA}^{Ser}$ gene (Chr6 tRNA#5), was cloned into pIRES2-DsRed plasmid and mutant constructs were generated by site-directed mutagenesis (green: Ala(AGC) and red: Leu(AAG)). **Right panel)** Serylated mutant tRNAs misincorporate Ser at the non-cognate codons indicated. **B)** Expression of mutant tRNAs on stably expressing cells was confirmed using SNaPshot. Samples were sequenced and analyzed using Peak Scanner software. The endogenous copies of $tRNA^{Ser}$ were 32 and 49.5-fold more expressed then the mutant tRNAs in $tRNA^{Ser}$(Ala) and $tRNA^{Ser}$(Leu) cells, respectively. **Grey:** Non-mutated $tRNA^{Ser}$; **Black:** mutant tRNAs. **C)** Percentage of cells in necrosis, late and early apoptosis were determined by flow cytometry using AnnexinV-FITC (1:100) and Propidium iodide (2.5µg/ml) staining. **D)** The number of foci arising from NIH3T3 cells was counted after 13-21 days after transfection. Data represents average ± SEM (n=2-3) and was analyzed using One-way ANOVA followed by Dunnett's using Mock cell line as control. There are no significant differences among cell lines (p>0.05).

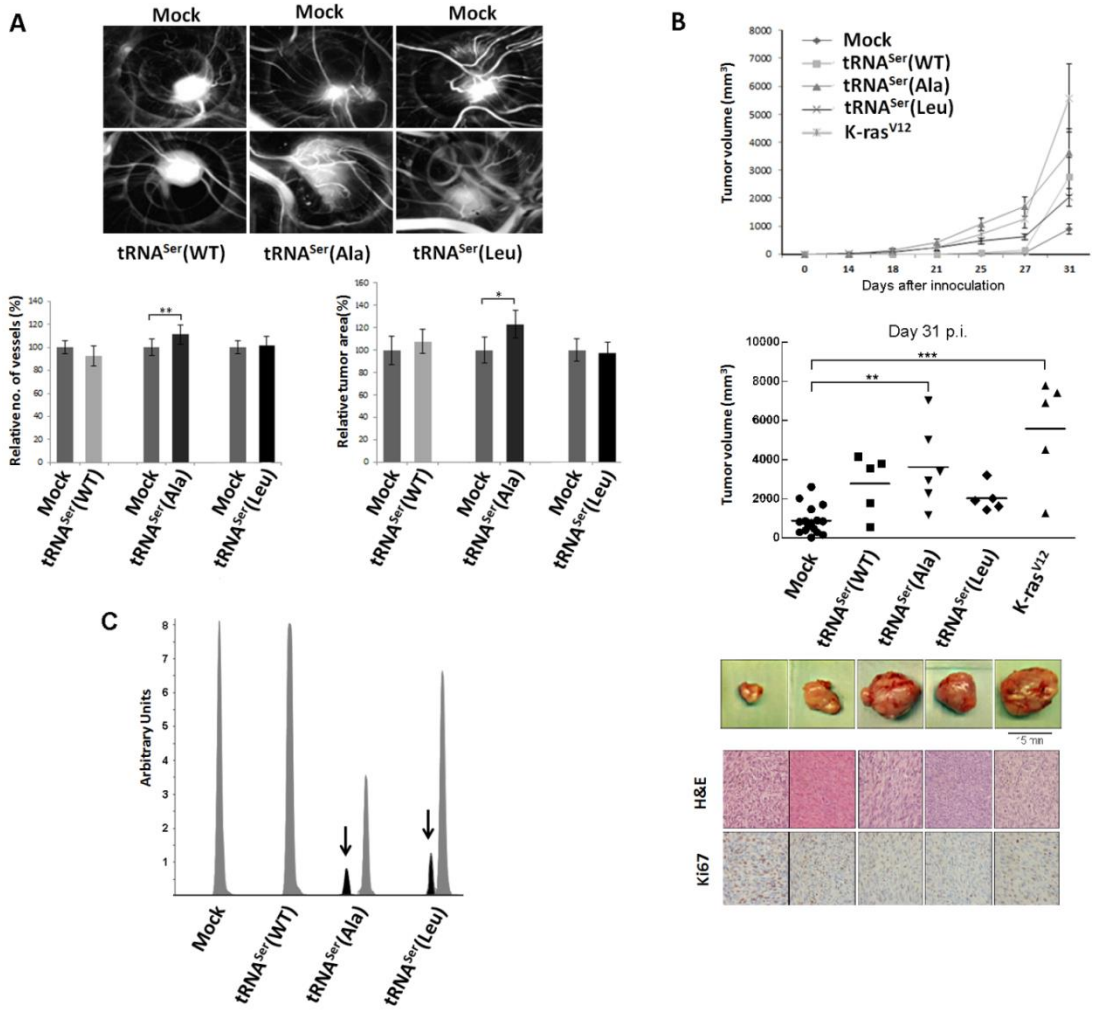


Figure 2. Impact of mistranslation on angiogenesis and tumor formation *in vivo*. **A)** CAM assay. **Upper panel)** Representative images of tumors and vessels produced by Mock, tRNA^{Ser}(WT), tRNA^{Ser}(Ala) and tRNA^{Ser}(Leu) cell lines. **Lower panel, left)** Quantitative evaluation of new vessels' formation. **Lower panel, right)** Relative tumor area. Data represents percentage relative to Mock. Graphics depict average ± SEM (n=12-14). Data was analyzed using two-tailed paired Student's *t* test (**p*<0.05; **p<0.01). **B)** Tumorigenic capacity of mistranslating cells in mice. **Upper panel)** Kinetics of tumor growth determined after inoculation of Mock, tRNA^{Ser}(WT), tRNA^{Ser}(Ala) and tRNA^{Ser}(Leu) cells and K-ras^{V12} expressing cells (positive control). **Middle panel)** Quantitative evaluation of tumor area at 31 days p.i.. Graphics depict the average ± SEM (n=5-10). Data was analyzed using One-way ANOVA followed by Dunnett's (***p*<0.01; ****p*<0.001). **Lower panel)** Photographs of representative tumors, H&E and Ki67 staining (40x amplification) from each condition. **C)** Expression of tRNA^{Ser} and mutant tRNAs in mice tumors measured by SNaPshot. Samples were sequenced and analyzed using Peak Scanner software. Expression of the mutant tRNAs in tRNA^{Ser}(Ala) and tRNA^{Ser}(Leu) cells were 4 and 5.9-fold lower than the endogenous tRNA^{Ser}, respectively. **Grey:** Non-mutated tRNA^{Ser}; **Black:** Mutant tRNA^{Ser}.

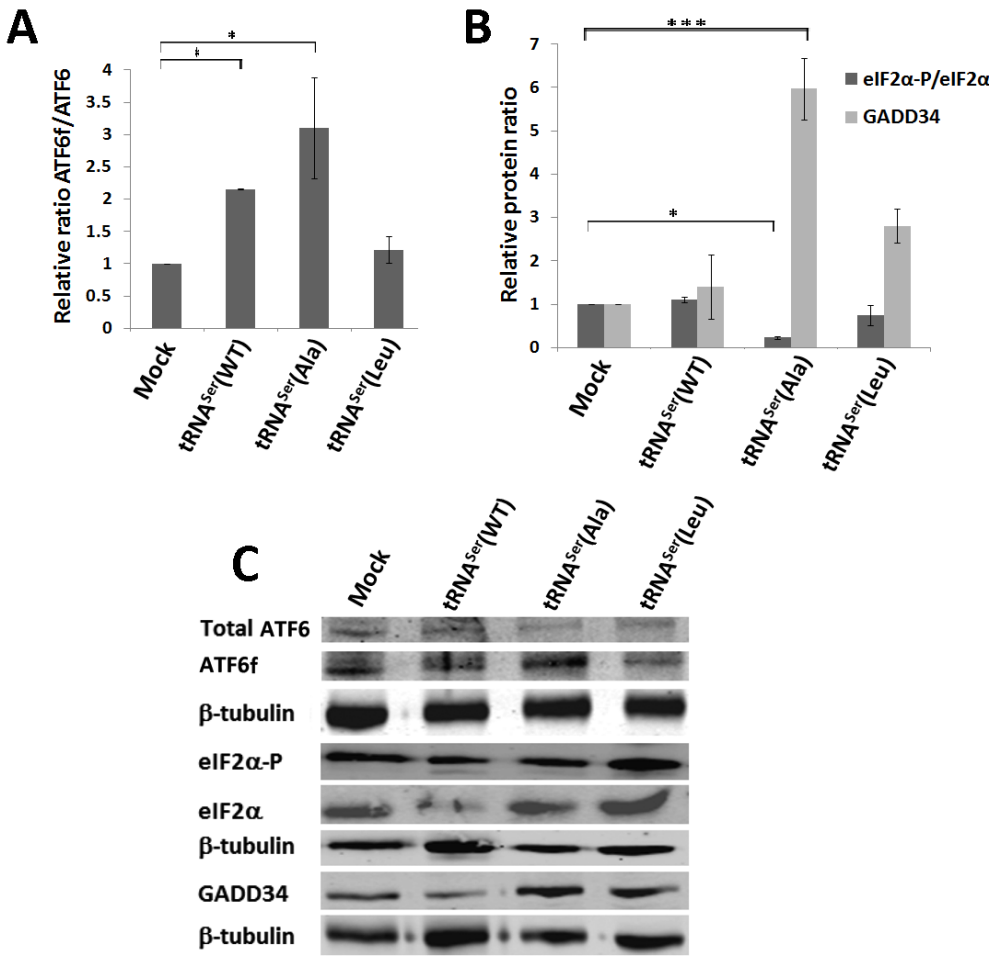


Figure 3. Activation of the UPR *in vivo*. **A)** Activation of ATF6 in tumors harboring the wild-type and mutant tRNAs. Total ATF6 and ATF6 fragment were detected by immunoblotting. **B)** eIF2α-P and GADD34 levels in each tumor lysate were analyzed by immunoblotting and relative expression values are shown. β-tubulin was used as loading control. **C)** Representative immunoblots for total ATF6, ATF6 fragment, total eiF2α, eIF2α-P, GADD34 and β-tubulin for each membrane. Graphics depict average ± SEM (n=3). Data was analyzed using One-way ANOVA followed by Dunnett's and significant *p-values* are shown (**p*<0.05; ****p*<0.001).

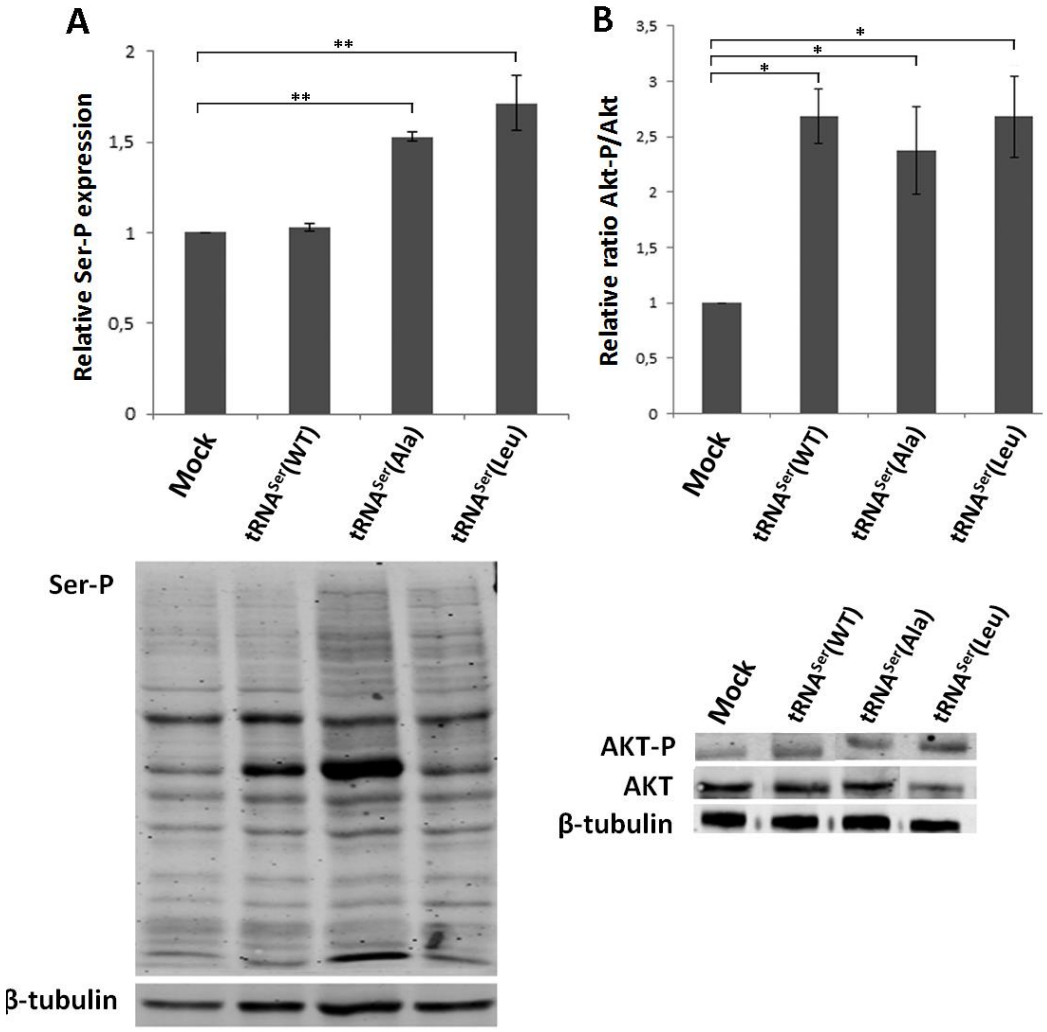


Figure 4. Classical cancer-associated pathways activated in mice tumors. **A)** Evaluation of total phosphoserine levels in tissue lysates from mice tumors. **B)** Relative activation ratio of Akt in tumor lysates compared to the Mock and representative immunoblots of Akt-P, total Akt and β-tubulin (loading control) from tumor lysates. Graphics depict average ± SEM (n=3). Data was analyzed using One-way ANOVA followed by Dunnett's and significant *p-values* are shown (**p*<0.05; ***p*<0.01).

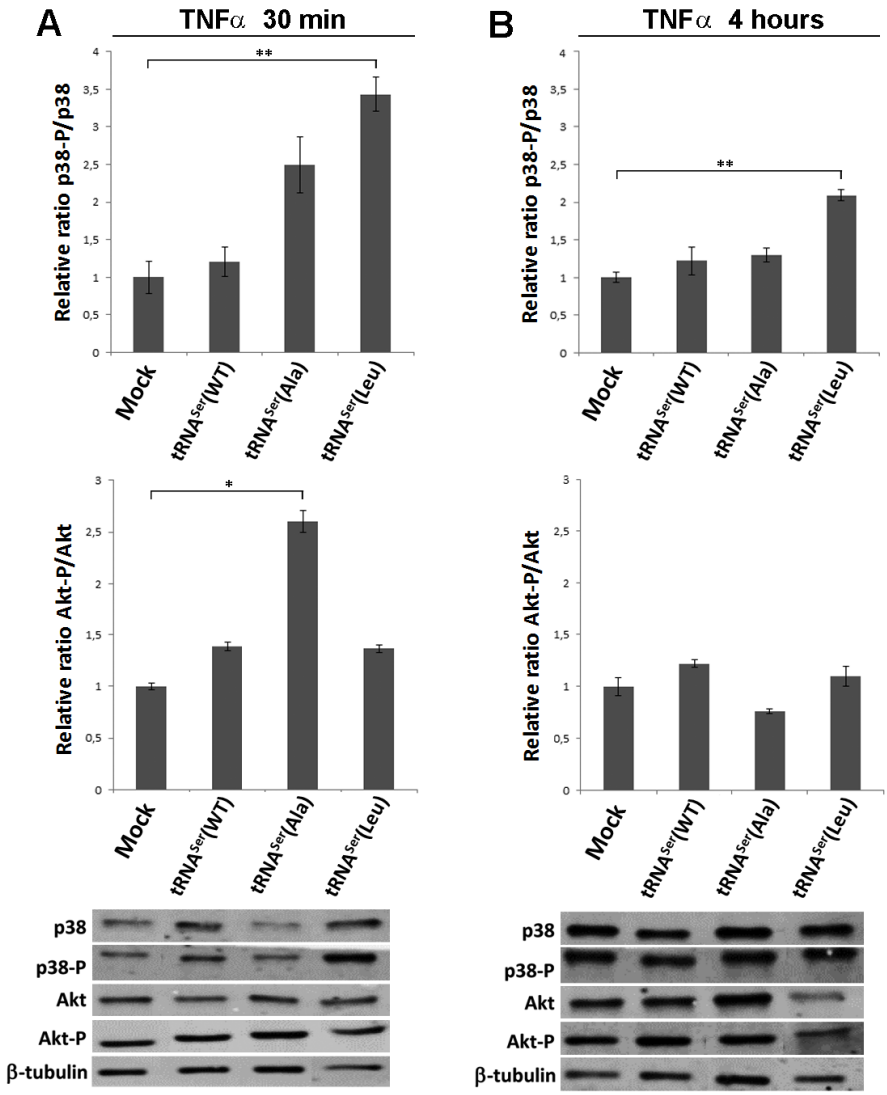


Figure 5. Pathways activated by TNFα induction. **A)** Treatment of cells misexpressing the Wt $tRNA^{Ser}$ and expressing mutant tRNAs with TNFα (30ng/ml) for 30 minutes. **Upper panel)** Relative activation ratios of p38 in cell lines exposed to TNFα. **Middle pannel)** Relative activation ratios of Akt in cell lines exposed to TNFα. **Lower panel)** Representative immunoblots of p38-P, total p38, Akt-P and total Akt in cell lines. β-tubulin was used as a loading control. **B)** Treatment of cells misexpressing the Wt $tRNA^{Ser}$ and expressing mutant tRNAs with TNFα (30ng/ml) for 4 hours. **Upper panel)** Relative activation ratios of p38 in cell lines exposed to TNFα. **Middle panel)** Relative activation ratios of Akt in cell lines exposed to TNFα. **Lower panel)** Representative Immunoblots of p38-P, total p38, Akt-P and total Akt in cells lines. β-tubulin was used as a loading control. Data represents average ± SEM (n=3), was analyzed using One-way ANOVA followed by Dunnett's and relevant *p-values* are displayed (**p*<0.05; ***p*<0.01).

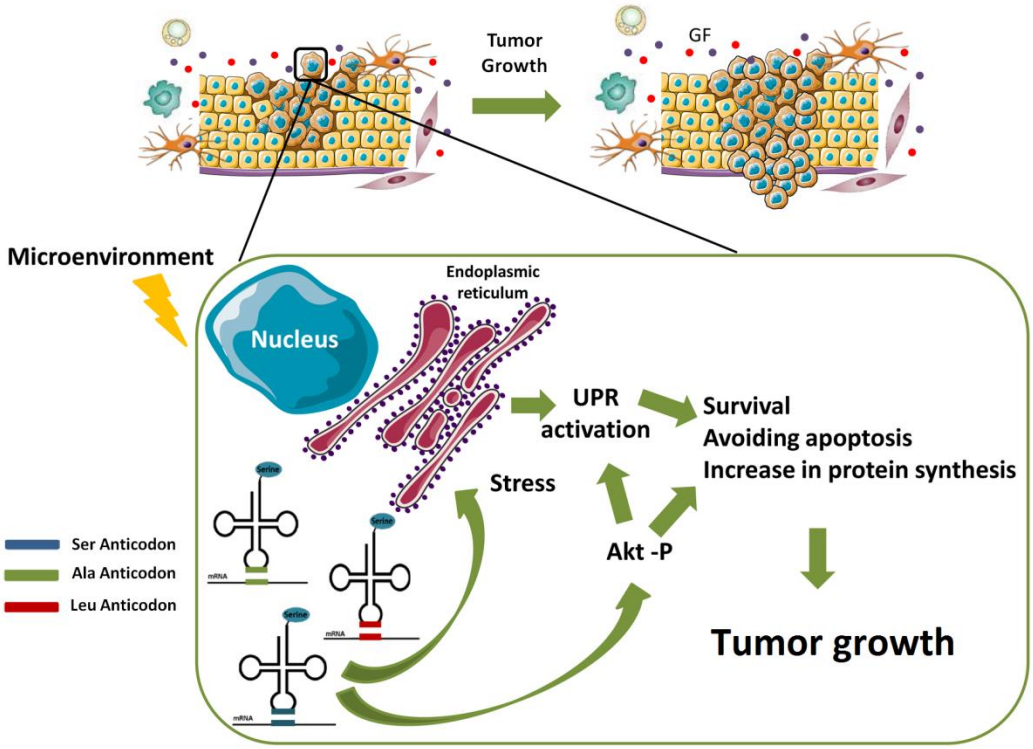
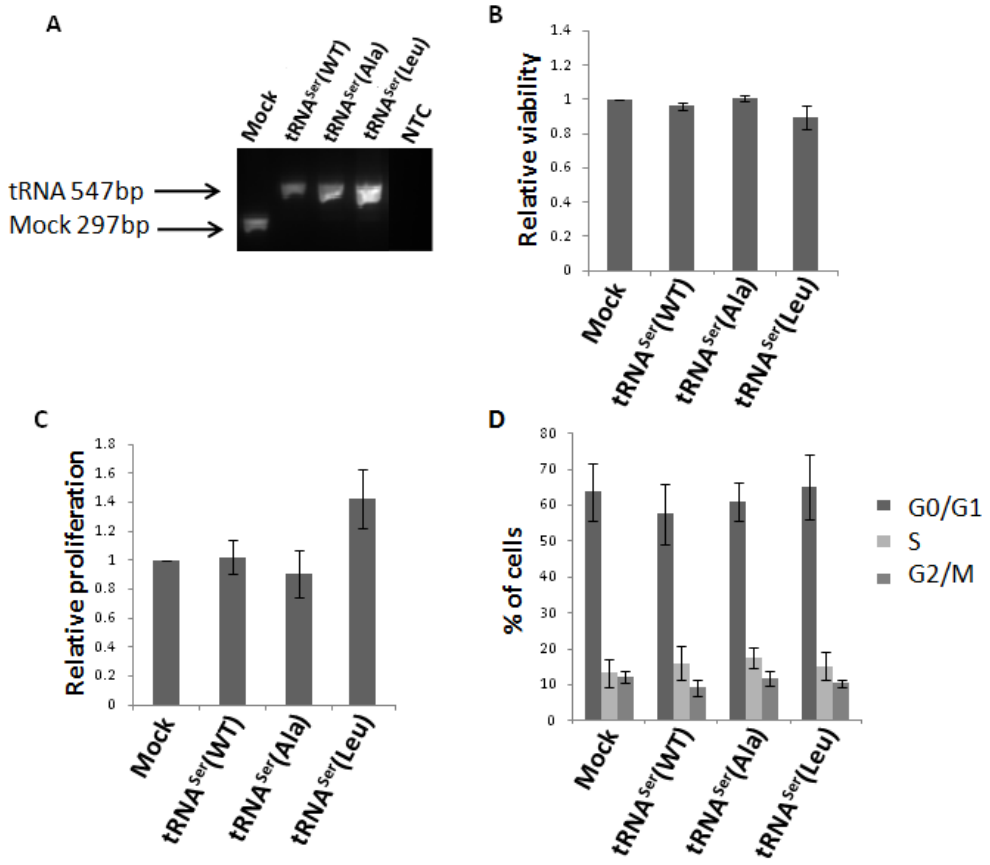
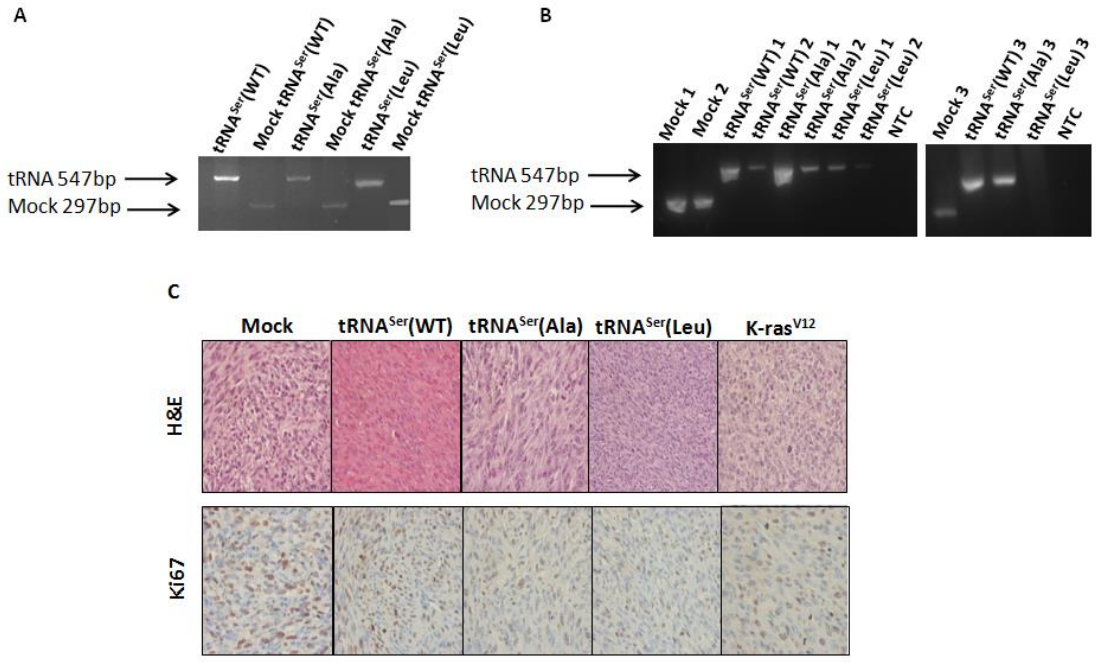


Figure 6. Representation of the stress response induced by mutant tRNAs and $tRNA^{Ser}$ misexpression. The expression of these tRNAs in NIH3T3 cells in combination with microenvironment stimuli *in vivo* induces ER stress and activation of the Akt pathway. These events lead to UPR activation, increasing cell capacity to survive, evade apoptosis and increased protein synthesis rate, especially in $tRNA^{Ser}$(Ala) cells where eIF2α-P is downregulated due to GADD34 overexpression. Overall these molecular mechanisms culminate in tumor growth. Adapted from Servier Medical Art collection (http://www.servier.com).

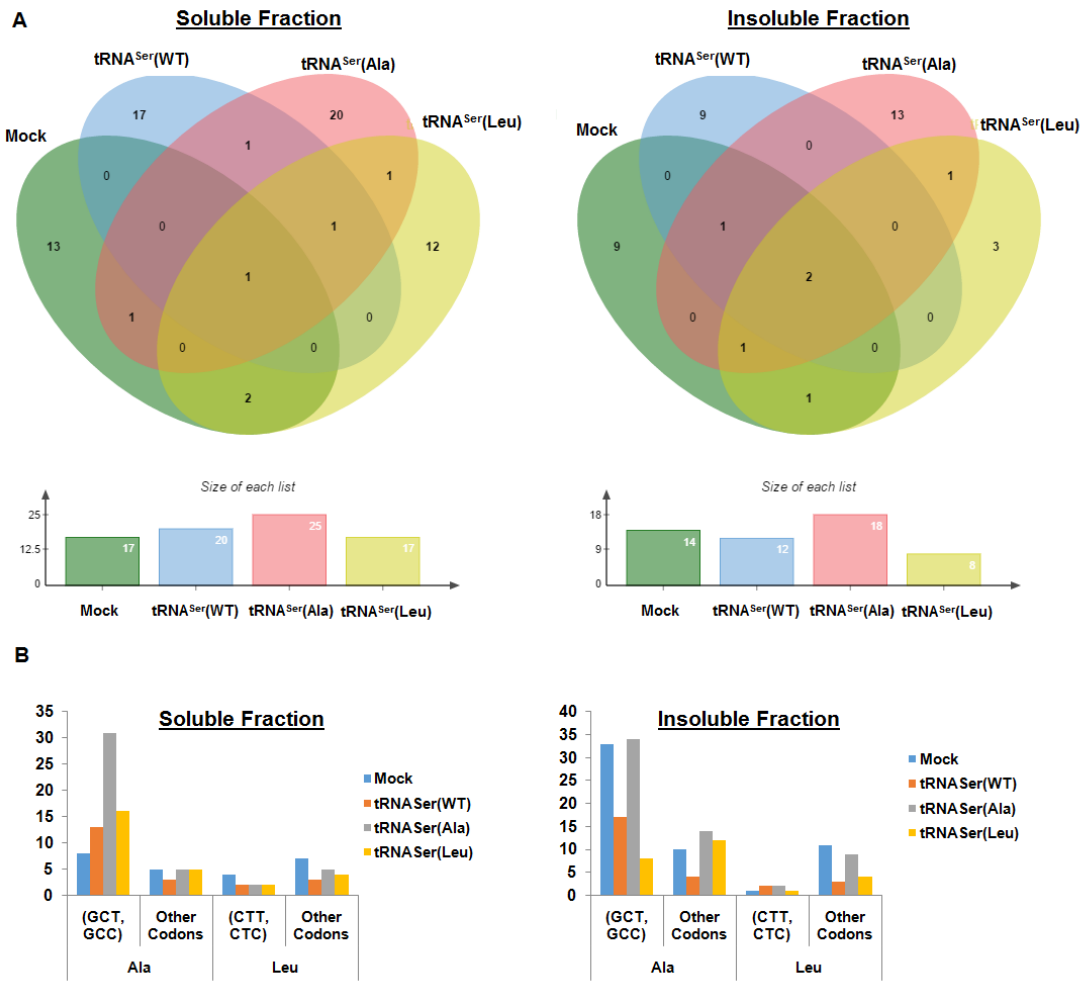
Supplementary Figures:



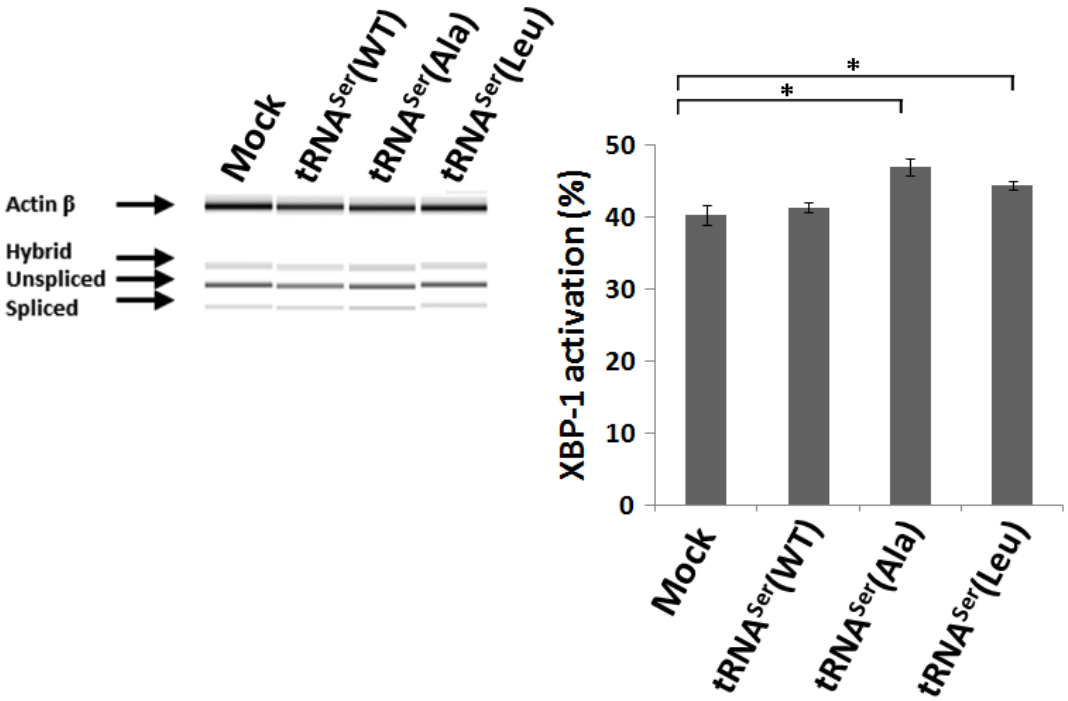
Supplementary Figure 1. Integration of pIRES2-DsRed and exogenous tRNAs in NIH3T3 cells and its phenotypic effects. **A)** Agarose gel (1%) of the PCR product obtained by amplification of genomic DNA from cell lines across the fragment of pIRES2-DsRed plasmid containing the tRNA insert. NTC represents the negative control. This is a cropped figure. NTC was at the end of the gel and there were three more samples which were excluded from the paper. **(B)** To access the cell viability, cells were grown for 48h and then stained with Trypan blue (0.4%). Number of viable cells was then registered and normalized to the Mock cell line. **(C)** To check if there were differences in proliferation, stable cell lines were grown in 10 mm cover slips for 48h and incubated with BrdU for 1h. Cover slips were incubated with1:10 mouse anti-BrdU and 1:500 goat anti-mouse Ig Alexa Fluor® 488. Total number of cells was counted using DAPI staining. Number of proliferating cells was determined and normalized to the Mock cell line. **(D)** For cell-cycle analysis we performed flow cytometry (Propidium Iodide (1mg/ml) staining). The percentage of cells detected in each phase of the cell cycle (G0/G1; S; G2/M) is shown in the graph. Data represents average ± SEM (n=3). Statistical significance was determined using ONE-way ANOVA with Dunnett's post-test. Results were not statistically significant ($p > 0.05$).



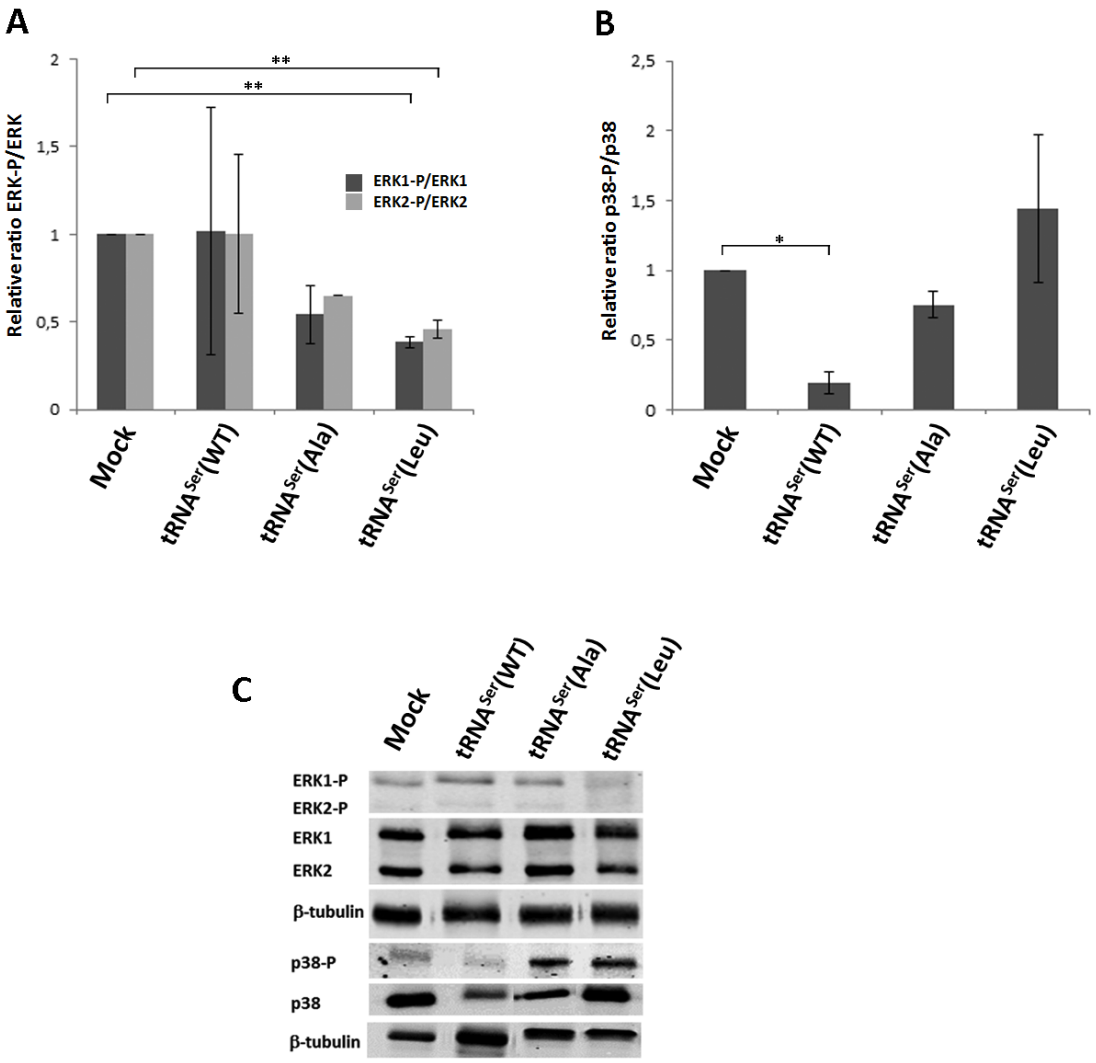
Supplementary Figure 2. Presence of the plasmid pIRES2-DsRed and exogenous tRNAs in the tumors (CAM and mice) and histological analysis of representative tumors. (A) Agarose gel of the PCR product obtained by amplification of genomic DNA from tumors, extracted from chicken embryo chorioallantoic membrane (CAM) model, across the fragment of pIRES2-DsRed plasmid containing the tRNA insert. NTC represents the negative control. **(B)** Agarose gel of the PCR product obtained by amplification of genomic DNA from tumors, extracted from mice model, across the fragment of pIRES2-DsRed plasmid containing the tRNA insert. NTC represents the negative control. **(C)** Histological analysis. **(Upper panel)** Hematoxylin and eosin staining staining (H&E) on paraffin-embedded tumor tissues showing high grade sarcomas (40x amplification). **(Lower panel)** Immunohistochemical analysis with anti-Ki67 antibody (1:400) on paraffin-embedded tumor tissues presenting high level of proliferation in all tumors (40x amplification).



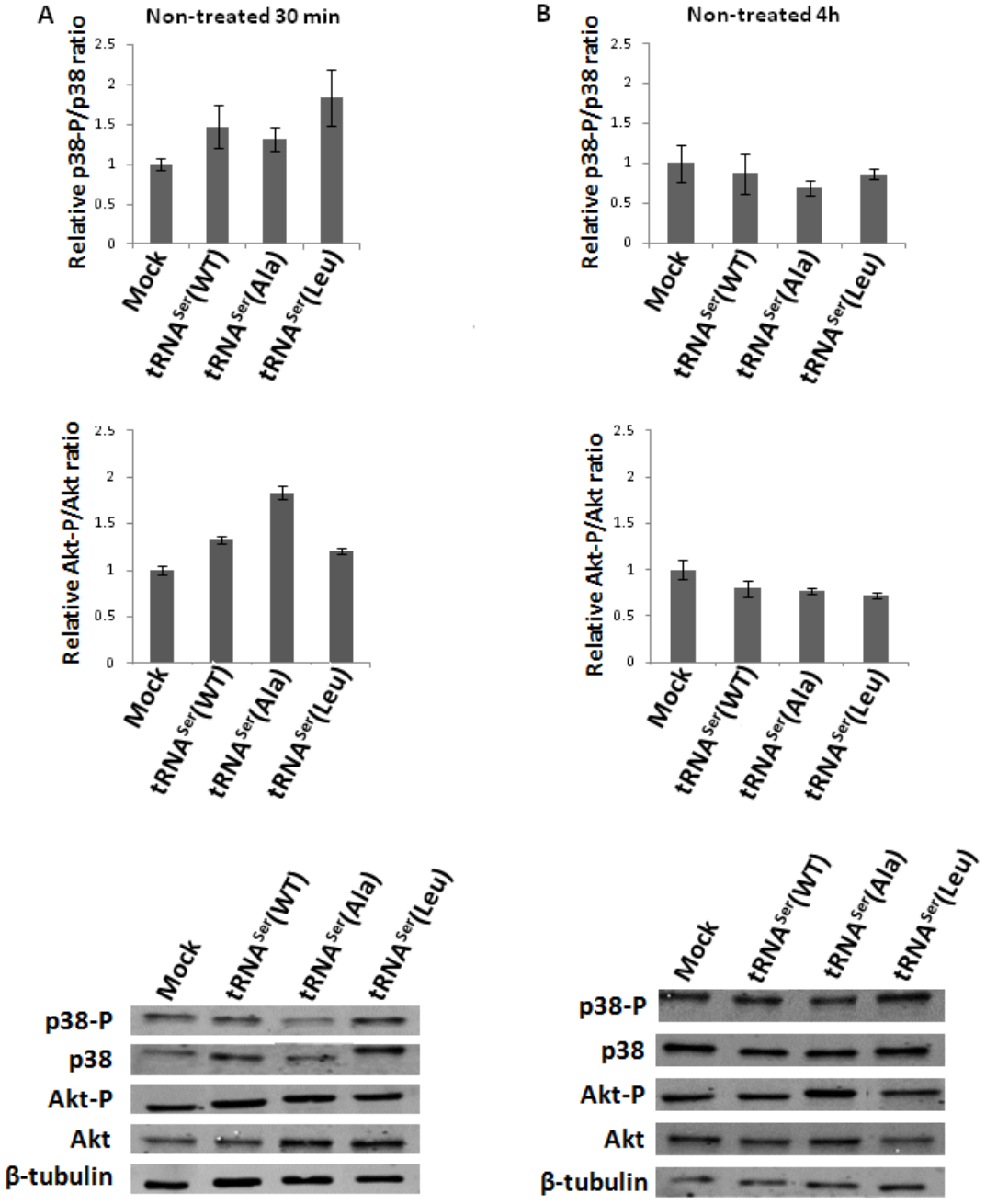
Supplementary Figure 3. Mass Spectrometry Analysis of protein fractions isolated from tumors derived from our cell lines. (A) Number of proteins that have at least one amino acid misincorporation Ala→Ser or Leu→Ser. **(Left panel)** Venn diagram showing the number of proteins with amino acid substitutions (Ala→Ser and Leu→Ser) in the soluble protein fraction. **(Right panel)** Venn diagram showing the number of proteins with amino acid substitutions (Ala→Ser and Leu→Ser) in the insoluble protein fraction. **(B)** Analysis of codons that correspond to the misincorporated amino acids **(Left panel)** Number of Ser misincorporations in Ala and Leu codons in the protein soluble fraction. The Ala codons, GCT and GCC, and the Leu codons, CTT and CTC correspond to the anticodons of our mutant tRNAs ($tRNA_{AGC}^{Ser}$(Ala) and $tRNA_{AAG}^{Ser}$(Leu)) **(Right panel)** Number of Ser misincorporations in Ala and Leu codons in the protein insoluble fraction.



Supplementary Figure 4. Activation of *XBP-1* in tumors expressing mutant tRNAs and the Wt serine tRNA. **Left panel)** The presence of unspliced (un), hybrid (H) and spliced (s) *XBP-1* forms was checked by RT-PCR. *Actin β* was used as a loading control. **Right panel)** Activation of *XBP-1* was determined using the formula: 100x[XBP-1s+0.5 XBP-1H]/[XBP-1s+XBP-1H+XBP-1un]. Graphics depict average ± SEM (n=3). Data was analyzed using One-way ANOVA with Dunnett's post-test and significant *p-values* are shown (*p<0.05).



Supplementary Figure 5. Classical cancer-associated pathways modulated in mice tumors. (A) ERK1/2 relative activation ratio among tumor lysates. **(B)** Relative activation ratios of p38 in tumor lysates. **(C)** Representative immunoblots of ERK1/2-P, ERK1/2, p38-P, total p38 and the respective β-tubulin (loading control) from tumor lysates. Graphics depict average ± SEM (n=3). Data was analyzed using One-way ANOVA with Dunnett's post-test and significant *p-values* are shown (**p*<0.05; ***p*<0.01).



Supplementary Figure 6. TNFα induction assay – results from non-treated cells (negative controls). (A) Negative controls of cells treated with with TNFα for 30 minutes**. (Upper panel)** Relative activation ratios of p38 in the cell lines not exposed to TNFα. **(Middle panel)** Relative activation ratios of Akt in the cell lines not exposed to TNFα. **(Lower panel)** Representative immunoblots of phosphorylated p38, p38, phosphorylated Akt and total Akt in cells lines that were not treated with TNFα. β-tubulin was used as a loading control. **(B)** Negative controls of cells treated with TNFα for 4 hours. **(Upper panel)** Relative activation ratios of p38 in the cell lines not exposed to TNFα. **(Middle panel)** Relative activation ratios of Akt in the cell lines not exposed to TNFα. **(Lower panel)** Representative immunoblots of phosphorylated p38, p38, phosphorylated Akt and total Akt in cells lines that were not treated with TNFα. β-tubulin was used as a loading control. Data

represents average ± SEM (n=3). Statistical significance was determined using unpaired two-tailored Student's t-test. Results are not statistically significant ($p > 0.05$), showing that there are no differences among cell lines when they are not exposed to any stimuli *in vitro*.