

Université de Montréal

**Caractérisation fonctionnelle du nouveau gène  
mitochondrial *mtalnd4* chez l'humain**

Par

Thierry Choquette

Département de sciences biologiques, Faculté des Arts et des Sciences

Mémoire présenté en vue de l'obtention du grade de maîtrise (M. Sc.) en sciences  
biologiques, option recherche en biologie

Décembre 2022

© Thierry Choquette, 2022

Université de Montréal  
Département de sciences biologiques, Faculté des Arts et des Sciences

*Ce mémoire intitulé*

**Caractérisation fonctionnelle du nouveau gène mitochondrial *mtatnd4* chez l'humain**

*Présenté par*

**Thierry Choquette**

*A été évalué par un jury composé des personnes suivantes*

**Matthew Regan**

Président-rapporteur

**Sophie Breton**

Directrice de recherche

**Annie Angers**

Codirectrice de recherche

**Marc Germain**

Membre du jury

## Résumé

Chez les cellules eucaryotes, la mitochondrie est une organelle impliquée dans plusieurs fonctions cellulaires (production d'énergie, apoptose, production de ROS, prolifération, signalisation cellulaire, vieillissement, immunité et plus encore) et possédant son propre génome, soit l'ADN mitochondrial (ADNmt). Chez l'humain, on croyait que l'ADNmt ne codait que pour 37 gènes impliqués dans la production d'énergie et la traduction mitochondriale. Cependant, le potentiel codant du génome mitochondrial aurait été sous-estimé. Il a récemment été démontré qu'à l'intérieur de ces principaux gènes connus pouvaient se cacher plusieurs petits gènes alternatifs. Ceux-ci se retrouvent au sein de régions non codantes ou possèdent des séquences d'initiation ou de terminaison de la traduction distinctes de celles du gène de référence dans lequel ils se retrouvent. Ils codent pour des micropeptides dérivés des mitochondries qui possèdent un large éventail de fonctions, s'ajoutant à la longue liste de fonctions dans lesquelles la mitochondrie est déjà impliquée. Parmi ces peptides, on retrouve l'Humanine, MOTS-c, SHLP1-6, Gau et SHMOOSE. Nous avons précédemment découvert un nouveau gène alternatif situé dans le gène *nd4*, nommé *mtalnd4*. Dans cette étude, nous visons à clarifier les fonctions du peptide alternatif correspondant MTALTND4 en étudiant son patron d'expression dans les tissus humains, l'impact de plusieurs stress sur son expression, l'impact du peptide sur la transcription des gènes, et ses partenaires d'interaction. Nous avons découvert que MTALTND4 pourrait être une molécule de signalisation sécrétée par les cellules en réponse au stress et affectant la physiologie pour induire un état de dépression bioénergétique en réduisant les processus de production et de demande en ATP. Plusieurs autres indices révélés par nos expériences suggèrent que MTALTND4 pourrait être une protéine multifonctionnelle impliquée dans de nombreuses voies de régulation.

**Mots-clé :** Mitochondrie, Protéome alternatif, Humain, altORF, sORF, Humanine, MOTS-c, SHLPs, SHMOOSE, MTALTND4

## Abstract

In eukaryotic cells, mitochondria are organelles involved in many cellular functions (energy production, apoptosis, ROS production, proliferation, cell signaling aging, immunity and more) and that possess their own genome, namely mitochondrial DNA (mtDNA). In humans, mtDNA was believed to encode only 37 genes involved in energy production and mitochondrial translation. However, the coding potential of the mitochondrial genome has been underestimated. It has recently been shown that within these main genes could hide several small alternative genes (*i.e.*, genes with non-coding regions or with translation initiation/termination sequences that are distinct from the reference gene sequences in which they are found). They code for mitochondrial-derived micropeptides (MDPs) that have a broad spectrum of functions, adding to the extensive list of functions in which mitochondria are already involved. These peptides include Humanin, MOTS-c, SHLP1-6, Gau and SHMOOSE. We have previously discovered a new alternative gene located in the *nd4* gene, termed *mtaltd4*. In this study, we aim to clarify the functions of the corresponding alternative peptide MTALTND4 by studying its expression pattern in human tissues, the impact of several stresses on its expression, the impact of the peptide on gene transcription, and its interaction partners. We have found that MTALTND4 could be a signaling molecule secreted by cells in response to stress and would affect physiology to induce a state of bioenergetic depression by reducing ATP-producing and ATP-demanding processes. Several other clues revealed by our experiments suggest that MTALTND4 could be a multifunctional protein involved in numerous regulatory pathways.

**Keywords:** Mitochondria, Alternative Proteome, Human, altORF, sORF, Humanin, MOTS-c, SHLPs, SHMOOSE, MTALTND4

## Table des matières

Résumé.....	3
Abstract.....	4
Table des matières.....	5
Liste des tableaux.....	7
Liste des figures.....	8
Liste des sigles et abréviations.....	12
Remerciements.....	16
1. Introduction.....	17
1.1 Les mitochondries et leur génome.....	17
1.1.1 Le génome mitochondrial humain.....	17
1.1.2 Le génome bactérien.....	18
1.2 Initiation de la traduction.....	19
1.2.1 Initiation de la traduction chez les procaryotes.....	20
1.2.2 Initiation de la traduction des gènes nucléaires chez les eucaryotes.....	21
1.2.3 Initiation de la traduction mitochondriale.....	22
1.3 L'existence d'un protéome alternatif.....	23
1.3.1 Le protéome alternatif procaryote.....	23
1.3.2 La découverte récente d'un protéome alternatif eucaryote.....	25
1.4 Les mécanismes de la traduction alternative chez les eucaryotes.....	28
1.4.1 Le <i>leaky scanning</i> .....	28
1.4.2 La réinitialisation ribosomale.....	29
1.4.3 Le <i>frameshifting</i> .....	29
1.4.4 Les IRES.....	30
1.5 Le protéome alternatif mitochondrial.....	31
1.6 La découverte de MTALTND4.....	33
1.7 Hypothèses et objectifs.....	34
2. ARTICLE : Functional characterization of the new human mitochondrial gene <i>mtaltn4</i> .....	35
ABSTRACT.....	36
INTRODUCTION.....	36
METHODS.....	38
Cell culture & sample preparation.....	38
Western blotting.....	39

Physiological cell treatments .....	39
Transcriptomics.....	40
Pull-down assay .....	42
RESULTS .....	44
MTALTND4 expression characterized in various tissues and physiological conditions .....	44
MTALTND4 has an exogenous effect on gene expression .....	45
MTALTND4 interacts with several proteins involved in mitochondrial functions.....	47
DISCUSSION.....	49
MTALTND4 exhibits a tissue-specific expression and reacts to different treatments known to affect mitochondrial functions .....	49
MTALTND4 depresses metabolism and biosynthetic processes .....	50
Interacting partners of MTALND4 also point to a possible role in regulation of gene expression .....	51
CONCLUSION.....	53
3. Discussion.....	54
3.1 Expression tissulaire de MTALTND4 .....	54
3.2 Expression de MTALTND4 en conditions stressantes.....	58
3.3 Influence de MTALTND4 sur le transcriptome .....	61
3.4 Partenaires d'interaction de MTALTND4 .....	63
3.5 Voies de signalisation affectées par MTALTND4 .....	65
4. Conclusion et perspectives.....	67
Références.....	68
Annexe .....	83
Annexe 1 : Article de Kienzle <i>et al.</i> (2022) en prépublication .....	83
Annexe 1.1 : L'intégrale de Kienzle <i>et al.</i> (2022) .....	83
Annexe 1.2 Figures qui seront incorporées à la version finale de l'article de Kienzle <i>et al.</i> (2022).....	103
Annexe 2: Supplementary Material .....	104
Annexe 2.1 Supplementary Figures .....	104
Annexe 2.2 Supplementary Tables .....	105

## Liste des tableaux

<b>Table A1.</b> Recurring signaling pathways found between common genes from microarrays and interacting partners from pull-down assays for DAVID Bioinformatics Resources and Reactome Knowledgebase analyses. ....	48
<b>Table S1.</b> Upregulated and downregulated genes between control and treated HeLa cells identified from microarray analysis. ....	105
<b>Table S2.</b> Upregulated and downregulated genes between control and treated HEK-293T cells identified from microarray analysis. ....	105
<b>Table S3.</b> Common (A) upregulated and (B) downregulated genes between HeLa and HEK-293T cells identified from microarray analysis. ....	106
<b>Table S4, related to Figure A3D.</b> Common Gene Ontology results from microarray data on HeLa and HEK-293T cells treated with exogenous MTALTND4 (10 $\mu$ M) for 4 hrs. ....	119
<b>Table S5.</b> Pathways associated with common downregulated and upregulated genes between HeLa and HEK-293T cells from microarrays. ....	120
<b>Table S6.</b> List of common proteins found by mass spectrometry following pull-down assay. Only results with 95% probability and two or more unique spectrum count are presented. ....	122
<b>Table S7.</b> Pathways associated with interaction partners from pull-down assays on HeLa and HEK-293T cells. ....	125
<b>Table S8.</b> Signaling pathways associated with C1QBP and YBX1 interactions. ....	130

## Liste des figures

- Figure 1.** Carte du génome mitochondrial humain comprenant 37 gènes essentiels à l'assemblage et la fonction de la chaîne respiratoire et l'ATP synthase. Les 13 gènes associés à des protéines dans la chaîne respiratoire et l'ATP synthase (vert, mauve, bleu et rouge) sont présentés, ainsi que les 2 gènes associés aux ARNr (jaune), les 22 gènes associés aux ARNt (gris) et les régions non-codantes (blanc). Figure tirée et modifiée de Picard *et al.* (2016)..... 18
- Figure 2.** Structure typique de molécules d'ARNm procaryote polycistronique et d'ARNm eucaryote monocistronique. Les ARNm bactériens peuvent posséder plusieurs sites d'initiation de la traduction et l'information nécessaire pour coder plusieurs protéines, tandis que les ARNm eucaryotes ne contiennent que l'information nécessaire pour coder une seule protéine. Figure tirée et modifiée d'Alberts *et al.* (2008)..... 19
- Figure 3.** Implication de la séquence de liaison au ribosome Shine-Dalgarno (SD) dans la synthèse protéique chez les procaryotes. La liaison de la séquence SD dans la région 5'UTR de l'ARNm avec la séquence anti-SD sur la queue 3' de la petite sous-unité ribosomale facilite la reconnaissance en aval des codons initiateurs de la traduction. Figure tirée et modifiée de Yang *et al.* (2016)..... 21
- Figure 4.** Vision globale de l'initiation de la traduction coiffe-dépendante chez les eucaryotes. Les facteurs eIF4E et eIF4G liés à la coiffe méthylée à l'extrémité 5' interagissent avec la protéine PABP, permettant la circularisation de l'ARNm par sa liaison à la queue polyA de l'extrémité 3'. La petite sous-unité ribosomale 40S est recrutée et balaye le transcrit jusqu'à reconnaître un premier codon initiateur. Figure tirée et modifiée de Blin & Ricci (2016). ..... 22
- Figure 5.** Modèle d'initiation de la traduction dans les mitochondries des mammifères. Dans cette voie, le facteur MtIF3 stabilise la petite sous-unité mitoribosomale 28S pour l'hébergement de MtIF2. La jonction de la grande sous-unité 39S entraîne un changement de conformation chez MtIF2, permettant le recrutement de l'ARNt et l'ARNm pour former le complexe d'initiation de la traduction. Figure tirée et modifiée de Khawaja *et al.* (2020). ..... 23



**Figure 6.** (A) Schématisation du principe des gènes chevauchants antiparallèles. Organisation de gènes sur une région d'ADN avec organisation antiparallèle entre les gènes 2 et 3 (figure tirée et modifiée de Tunca *et al.*, 2009). (B) Schématisation de différentes stratégies utilisées pour le principe de gènes à l'intérieur d'autres gènes chez les bactéries. Traduction de deux protéines dont (a) le codon stop diffère, (b) le codon start alternatif est présent dans la séquence primaire ou (c) les codons start et stop alternatifs sont présents dans la séquence primaire. Figure tirée et modifiée de Meydan *et al.* (2018)..... 24

**Figure 7.** ARNm eucaryote typique et ses possibles ORFs comprenant l'ORF de référence (refORF) annoté dans les bases de données nucléotidiques actuelles ainsi que les ORFs alternatifs (altORF) compris entre des codons start et stop distincts de ceux du refORF et codant pour une protéine alternative qui n'est pas annotée dans les bases de données. Ceux-ci peuvent se retrouver dans les ARNm ou dans les ARN dits « non codants ». Figure tirée et modifiée de Landry *et al.* (2015)..... 27

**Figure 8.** Deux méthodes les plus connues de la traduction alternative : A) Lors du *leaky scanning*, la sous-unité ribosomale 40S peut soit reconnaître un codon start en amont (altORF) et conduire à la traduction d'une protéine, soit balayer au-delà et initier la traduction au niveau d'un codon start en aval (refORF). B) La réinitialisation ribosomale se produit lorsque la traduction est initiée au codon start d'une séquence d'ADN codante en aval (refORF) après avoir terminé la traduction d'un ORF en amont (altORF) en raison de la non-dissociation de la sous-unité 40S de l'ARNm. La traduction de l'altORF conduit à la production d'un court peptide. Figure tirée et modifiée de Andrews & Rothnagel (2014)..... 30

**Figure 9.** Carte du génome mitochondrial humain montrant l'emplacement des nouveaux altORFs caractérisés (Humanine, MOTS-c, SHLPs et Gau en bleu), incluant le nouveau gène *mtaltnd4* (en rouge) et tous les cadres de lecture ouverts potentiels sur les brins sens (flèches orangées) et antisens (flèches vertes). Figure tirée et modifiée d'Angers *et al.* (2019)..... 33

**Figure A1.** Expression of nuclear-encoded mitochondrial ATP5 (loading control), mtDNA-encoded ND4 and MTALTND4 in various human tissues (kidney, heart, testis, brain, liver, skeletal muscle, spleen) and HeLa cells by immunoblotting. .... 44

<b>Figure A2.</b> Impact of (A) dose-dependent hydrogen peroxide treatment (0, 0.5, 1, 5, 10 mM) for 30 minutes (n=4-9); (B) starvation treatment (6, 24 hrs) (n=3-4); and (C) hypoxia treatment for 4 hrs (n=3) on normalized MTALTND4 expression within HeLa cells. Data shown as mean ± SEM, Student's t-test, *p < 0.05. ....	45
<b>Figure A3. Effect of MTALTND4 on gene expression.</b> Microarray analysis performed on HeLa and HEK-293T cells treated with MTALTND4 (10mM) for 4h (n=2). (A) Principal component analysis (PCA) of all samples passing quality control review. (B) Hierarchical clustering of differentially expressed genes depicting a high-level (red, fold change > 1.5) and low-level (green, fold change < 1.5) expression relative to control cells in both cell types. (C) Venn diagram depicting upregulated (red) and downregulated (blue) genes in both cell types. (D) Common Gene Ontology results of differentially expressed genes in both cell types. ....	46
<b>Figure A4.</b> Diagram showing the cellular components of the identified MTALTND4 interacting proteins.....	48
<b>Figure 10.</b> Expression globale de la protéine MT-ND4 codée par le gène mitochondrial <i>nd4</i> dans différents tissus humains. Figure tirée et modifiée de The Human Protein Atlas project. ....	56
<b>Figure 11.</b> Expression de SHLP4 et MTALTND4 pour divers tissus humains. Figure tirée et modifiée de Cobb <i>et al.</i> (2016). ....	58
<b>Figure 12.</b> Expression normalisée de MTALTND4 dans des cellules HeLa suite à un traitement hypoxique issu de lysats commerciaux (n=3) ou induit en laboratoire (n=11). Présenté par la moyenne ± SEM, Student's t-test, *p < 0.05. ....	61
<b>Kienzle <i>et al.</i>, 2022: Figure S8.</b> Detection of MTALTND4, mitochondrial-encoded ND4 and nuclear-encoded mitochondrial ATP5 by western blotting in human plasma. ....	103

**Kienzle *et al.*, 2022: Figure S9. GST pull-down assay indicating that MTALTND4 and C1QBP interact.** Western blot probed with anti-C1QBP antibodies. Lane 1: pull-down products from GST bound glutathione beads (HeLa cells). Lane 2: from GST-MTALTND4 bound glutathione beads (HeLa cells). Lane 3: HeLa cells lysate. Lane 4: pull-down products from GST bound glutathione beads (HEK-293T cells). Lane 5: from GST-MTALTND4 bound glutathione beads (HEK-293T cells). Lane 6: HEK-293T cells lysate..... 103

**Figure S1, related to Figure A1 .** Two immunoblots of the same samples showing MTALTND4 expression in various human tissues (K: kidney, H: heart, T: testis, B: brain, L: liver, M: skeletal muscle, S: spleen) and HeLa cells with presence of secondary bands (blue squares)..... 104

## Liste des sigles et abréviations

ADN : Acide désoxyribonucléique

ADNmt : Acide désoxyribonucléique mitochondrial

ADP : Adénosine Diphosphate

altORF : *Alternative Open Reading Frame* / Cadre de lecture ouvert alternatif

AMPK : *Adenosine Monophosphate-Activated Protein Kinase*

ANP32B : *Acidic Nuclear Phosphoprotein 32 Family Member B*

ARE : *Antioxidant response element* / Élément de réponse antioxydante

ARN : Acide ribonucléique

ARNm : Acide ribonucléique messenger

ARNr : Acide ribonucléique ribosomique

ARNt : Acide ribonucléique de transfert

ATCC : *American Type Culture Collection*

ATP : Adenosine triphosphate

ATP5/ATP5F1A : *ATP synthase subunit alpha*

ATP5F1C : *ATP synthase subunit gamma*

ATP6 : *Mitochondrially Encoded ATP Synthase Membrane Subunit 6*

ATP8 : *Mitochondrially Encoded ATP Synthase Membrane Subunit 8*

BSA : *Bovine serum albumin*

CNTFR : *Ciliary neurotrophic factor receptor*

COX1 : *Mitochondrially Encoded Cytochrome C Oxidase I*

C1QBP : *Complement component 1q subcomponent binding protein*

DAVID : *Database for Annotation, Visualization and Integrated Discovery*

DMEM : *Dulbecco's Modified Eagle Medium*

*E. coli* : *Escherichia coli*

EDTA : Éthylènediaminetétraacétique

eIF2 : *Eukaryotic initiation factor 2*

eIF2B : *Eukaryotic translation initiation factor 2B*

eIF3 : *Eukaryotic initiation factor 3*

eIF4E : *Eukaryotic initiation factor 4E*

eIF4G : *Eukaryotic initiation factor 4G*  
ERK1/2 : *Extracellular signal-regulated kinases 1 and 2*  
fMet-ARNt<sup>Met</sup> : *N-formylméthionine-acide ribonucléique de transfert méthionine*  
FPRL1 : *Formyl peptide receptor-like 1*  
GCN2 : *General control nonderepressible 2*  
GO : *Gene Ontology*  
GP130 : *Glycoprotéine 130*  
GST : *Glutathion S-transférase*  
HEK-293T : *Human Embryonic Kidney 293 SV40 large T antigen*  
HEPES : *Acide 4-(2-hydroxyéthyl)-1-pipérazine éthane sulfonique*  
HiF-1 $\alpha$  : *Hypoxia-inducible factor 1-alpha*  
HRI : *Heme-regulated inhibitor*  
IF1 : *Initiation factor 1*  
IF2 : *Initiation factor 2*  
IF3 : *Initiation factor 3*  
IgG : *Immunoglobuline G*  
IPTG : *Isopropyl  $\beta$ -D-1-thiogalactopyranoside*  
IRES : *Internal ribosome entry site*  
kDa : *Kilodalton*  
LB : *Lysogeny broth*  
LSB : *Laemmli Sample Buffer*  
MDP : *Mitochondrial-derived peptides / peptides dérivés de la mitochondrie*  
Met-ARNt<sup>Met</sup> : *Méthionine-acide ribonucléique de transfert méthionine*  
MOTS-c : *Mitochondrial open reading frame of the 12S rRNA-c*  
mS37 : *28S Ribosomal Protein S37, Mitochondrial*  
MTALTND4 : *Mitochondrial alternative ND4 protein*  
mtIF2 : *Mitochondrial Initiation Factor 2*  
mtIF3 : *Mitochondrial Initiation Factor 3*  
MUHC : *McGill University Health Center / Centre Universitaire de Santé McGill*  
NAP1L1 : *Nucleosome assembly protein 1-like 1*  
ND4 : *Mitochondrially Encoded NADH:Ubiquinone Oxidoreductase Core Subunit 4*

ND4L : *Mitochondrially Encoded NADH:Ubiquinone Oxidoreductase Core Subunit 4L*  
ND5 : *Mitochondrially Encoded NADH:Ubiquinone Oxidoreductase Core Subunit 5*  
NPM1 : *Nucleophosmin 1*  
NRF2 : *Nuclear factor erythroid 2-related factor 2*  
ORF : *Open Reading Frame / Cadre de lecture ouvert*  
PABP : *PolyA-binding protein*  
pAkt : *Phosphorylated protein kinase B*  
PANTHER : *Protein analysis through evolutionary relationships*  
pb : *Paires de bases*  
PBS : *Phosphate buffered saline*  
PCA : *Principal component analysis*  
PCNA : *Proliferating cell nuclear antigen*  
PERK : *Protein kinase R (PKR)-like endoplasmic reticulum kinase*  
PHB1 : *Prohibitine*  
PKR : *Protein kinase RNA-activated*  
PMSF : *Phenylmethylsulfonyl Fluoride*  
polyA : *Polyadénosine*  
PPM1G : *Protein Phosphatase Magnesium-Dependent 1 Gamma*  
PVDF : *Polyvinylidene Difluoride*  
refORF : *Reference Open Reading Frame / Cadre de lecture ouvert de référence*  
RNA-Seq : *Ribonucleic acid sequencing*  
ROS : *Reactive oxygen species / Espèces réactives de l'oxygène*  
SD : *Shine-Dalgarno*  
SDS-PAGE : *Sodium dodecyl sulfate-polyacrylamide gel electrophoresis*  
SEM : *Standard error of mean*  
SHLP1-6 : *Small Humanin-like peptides 1-6*  
SHMOOSE : *Small Human Mitochondrial ORF Over SERine tRNA*  
Site P : *Site peptidique*  
SLC25A6 : *Solute Carrier Family 25 Member 6*  
sORF : *Short Open Reading Frame / Cadre de lecture ouvert court*  
STAT3 : *Signal Transducer and Activator of Transcription 3*

TAC : *Transcriptome Analysis Console*

TBS-T : *Tris-buffered saline-Polysorbate 20*

TEMED : *Tétraméthyléthylènediamine*

TGF $\beta$  : *Transforming growth factor beta*

TRAIL : *Tumour necrosis factor-related apoptosis-inducing ligand*

UPLC-MS/MS : *Ultra performance liquid chromatography - tandem mass spectrometry*

UPR : *Unfolded protein response*

UTR : *Untranslated region*

WSX1 : *Interleukin-27 receptor subunit alpha*

YBX1 : *Y box binding protein-1*

## Remerciements

J'aimerais tout d'abord remercier ma directrice et ma codirectrice de recherche, Sophie Breton et Annie Angers, pour leur travail formidable, leur temps, leur appui et leurs nombreuses corrections. Merci Sophie pour ton entrain, ton encadrement et de toujours être là pour tes étudiants, jour et nuit. Je n'aurais pas pu demander pour une meilleure directrice. Merci Annie d'avoir partagé tes connaissances des techniques de laboratoire et tes conseils qui m'ont permis d'accomplir mon projet. Merci pour cette incroyable opportunité de maîtrise et, malgré avoir débuté au milieu d'une pandémie qui fut plus longue que prévu, vous avez su me garder motivé et nous avons pu nous adapter malgré quelques complications résultantes. Au final, nous en ressortons avec un beau projet dont je suis fier et dont j'en ressors grandi. Ce fut un plaisir de travailler avec vous et ce n'est pas la fin. Je tiens aussi à remercier mes collègues Laura (ma mentore qui m'a tout montré), Roza, Stefano, Julie, Thierry, Andréanne, Mélanie, Georges, Yousef, Émmélie, Alizée, Vincent, Ludovic, Francis, Rudy, Ariane, Sofia, Marianne, Aziz, et Kiara pour tous les beaux moments passés ensemble dans le laboratoire, les sorties, les travaux pratiques, les conférences, au Québec, comme à l'international. Vous avez rendu ma maîtrise divertissante avec vos discussions et débats captivants. Merci aux membres de ma famille et à mes ami(e)s qui m'ont toujours encouragé dans mon projet, même s'ils ne savent pas encore de quoi il traite, ni ce qu'est une mitochondrie, mais qui continuent de démontrer de l'intérêt envers mes études. Merci à mon papi que nous avons perdu en cours de route, mais qui serait fier de me voir ici. Un dernier gros merci à ma copine Nayla pour son soutien inconditionnel, sa capacité à me faire sourire et de toujours me pousser à me dépasser.



# 1. Introduction

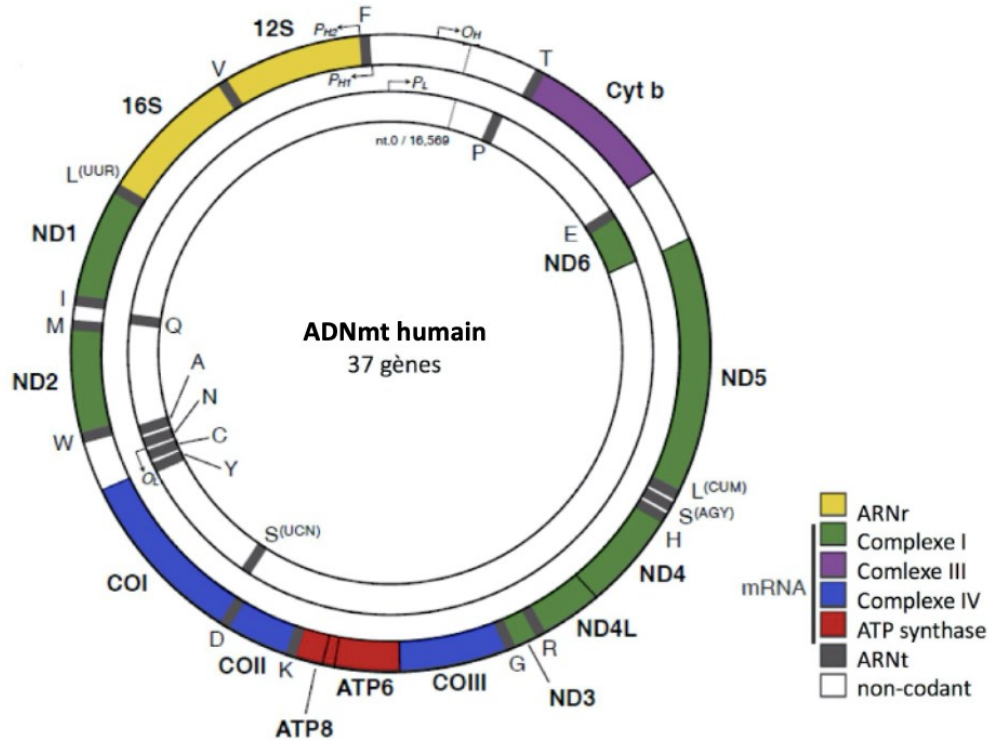
## 1.1 Les mitochondries et leur génome

Les premières cellules eucaryotes seraient originaires d'une endosymbiose survenue il y a environ 2,5 milliards d'années entre un groupe de microorganismes unicellulaires nommé les archées et une lignée de protéobactéries qui deviendra la mitochondrie (Roger *et al.*, 2017; Martijn *et al.*, 2018; Wallace, 2018). Étant un organisme à part entière, la protéobactérie possédait son propre génome indépendant de celui de la cellule primitive. Cette caractéristique est conservée encore à ce jour sous forme d'ADN mitochondrial. Au fil du temps, le génome nucléaire s'est complexifié, tandis qu'il s'est plutôt simplifié pour celui de la mitochondrie. En effet, le génome mitochondrial en symbiose a perdu plusieurs gènes associés au mode de vie libre de l'ancêtre protéobactérien ou devenus redondants avec la cellule hôte primitive. Une grande partie du matériel génétique a également été transférée vers le génome nucléaire de cette nouvelle cellule eucaryote (Roger *et al.*, 2017). Grâce à cette division des gènes, le noyau s'est consacré au codage des éléments anatomiques de la cellule, tandis que la mitochondrie s'est spécialisée dans le codage des éléments nécessaires à la production d'énergie (Wallace, 2018). Chez la cellule eucaryote, la mitochondrie est donc une organelle jouant un rôle central dans la phosphorylation oxydative pour produire l'ATP, la molécule clé de l'énergétique cellulaire. La mitochondrie est aussi impliquée dans de nombreuses autres fonctions. Parmi celles-ci, on retrouve la régulation des processus apoptotiques, la production d'espèces réactives de l'oxygène (les ROS qui sont à la fois éléments clés de la communication avec le noyau, mais aussi source de stress oxydatif), la prolifération cellulaire, la signalisation cellulaire, la régulation calcique, l'immunité, le développement de maladies, le vieillissement et plus encore (Kotrys & Szczesny, 2019).

### 1.1.1 Le génome mitochondrial humain

La publication de la première séquence complète du génome mitochondrial humain faisait mention d'un petit ADN double brin circulaire d'environ 16 000 paires de bases (pb) ne possédant que 37 gènes (Anderson *et al.*, 1981) : 13 gènes codant pour des protéines impliquées dans la phosphorylation oxydative (sous-unités des complexes I, III, IV et V), permettant la production d'ATP, ainsi que deux gènes codant pour des ARN ribosomiaux (ARNr — 12S et 16S) et 22 gènes codant pour des ARN de transfert (ARNt) qui sont des éléments responsables de la traduction

mitochondriale (Taanman, 1999) (Figure 1). La majorité de ces gènes se trouvent sur un seul des brins qui est principalement codant, tandis que le brin complémentaire est presque entièrement non codant. Le génome est organisé de manière efficace avec certains gènes se chevauchant de quelques nucléotides et peu de régions non codantes (Anderson *et al.*, 1981).

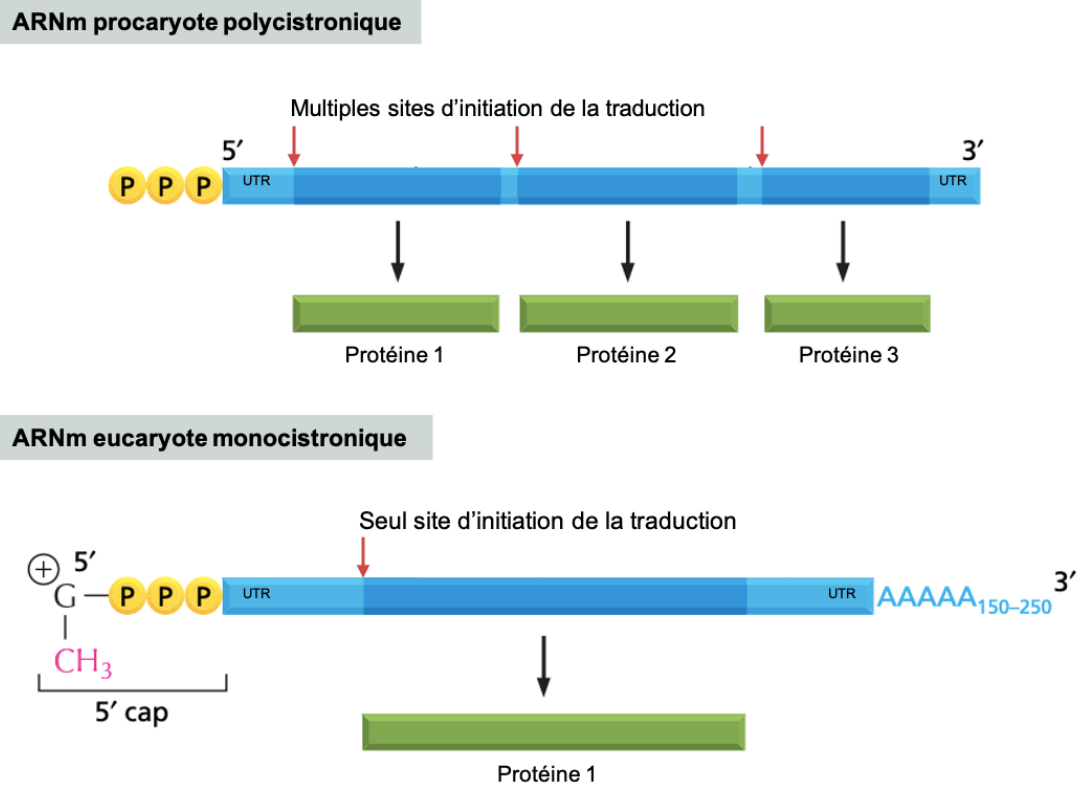


**Figure 1.** Carte du génome mitochondrial humain comprenant 37 gènes essentiels à l’assemblage et la fonction de la chaîne respiratoire et l’ATP synthase. Les 13 gènes associés à des protéines dans la chaîne respiratoire et l’ATP synthase (vert, mauve, bleu et rouge) sont présentés, ainsi que les 2 gènes associés aux ARNr (jaune), les 22 gènes associés aux ARNt (gris) et les régions non codantes (blanc). Figure tirée et modifiée de Picard *et al.* (2016).

### 1.1.2 Le génome bactérien

Le génome protéobactérien consiste, lui aussi, en un ADN double brin circulaire. L’ADN mitochondrial (ADNmt) humain aurait donc conservé la forme de son génome bactérien ancestral, quoique leur taille diffère grandement. Alors que l’ADNmt humain contient 16 569 pb, le génome bactérien possède en moyenne ~3 870 000 pb (Cobb *et al.*, 2016; diCenzo & Finan, 2017). La taille génomique n’est cependant pas la seule caractéristique divergente entre les deux groupes. Les bactéries possèdent des mécanismes d’initiation de la traduction particuliers permettant à un seul ARN messager (ARNm) de produire plusieurs protéines. En effet, alors que les ARNm

transcrits chez les eucaryotes sont habituellement monocistroniques, c'est-à-dire qu'ils ne peuvent coder que pour une seule protéine, plusieurs ARNm procaryotes sont polycistroniques, contenant les séquences codant plusieurs protéines (Orr *et al.*, 2020) (Figure 2).



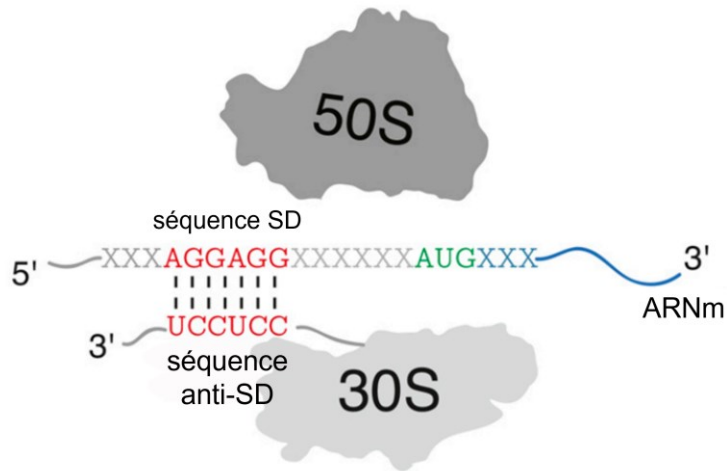
**Figure 2.** Structure typique de molécules d'ARNm procaryote polycistronique et d'ARNm eucaryote monocistronique. Les ARNm bactériens peuvent posséder plusieurs sites d'initiation de la traduction et l'information nécessaire pour coder plusieurs protéines, tandis que les ARNm eucaryotes ne contiennent que l'information nécessaire pour coder une seule protéine. Figure tirée et modifiée d'Alberts *et al.* (2008).

## 1.2 Initiation de la traduction

La machinerie de transcription et de traduction a évolué en développant des mécanismes propres à chaque type de génome. Ici, les différents mécanismes de l'initiation de la traduction seront présentés.

### 1.2.1 Initiation de la traduction chez les procaryotes

Chez les procaryotes, la traduction conventionnelle débute par la dissociation du ribosome en sa petite sous-unité (30S) et sa grande sous-unité (50S) (Wang *et al.*, 2021). Les facteurs d'initiation de la traduction IF1 et IF3 (*initiation factor 1 et 3*) ainsi que l'ARNm se lient à la petite sous-unité ribosomale. Ceci assure une conformation permettant le recrutement du facteur IF2 qui s'associe avec l'ARNt initiateur possédant une méthionine formylée (fMet-ARNt<sup>Met</sup>). Par la suite, la grande sous-unité ribosomale joint la petite pour former le complexe d'initiation pouvant initier la synthèse protéique (Khawaja *et al.*, 2020). Une des particularités de la traduction chez les procaryotes est la méthode de reconnaissance du codon initiateur par complémentarité nucléotidique. L'ARNm bactérien possède une séquence de liaison au ribosome (*ribosome binding sequence*) qui se retrouve quelques nucléotides en amont du codon initiateur de la traduction. Par exemple, plusieurs bactéries, dont *E. coli*, possèdent une séquence nommée Shine-Dalgarno qui se situe à environ sept nucléotides du codon d'initiation et est représentée par 5'-AGGAGG-3' (Amin *et al.*, 2018) (Figure 3). La sous-unité ribosomale, quant à elle, possède sur sa queue 3' la séquence complémentaire anti-Shine-Dalgarno (5'-CCUCCU-3') permettant l'interaction avec l'ARNm à ces sites (Shine & Dalgarno, 1974; Yang *et al.*, 2016). Cette interaction est importante afin de faciliter la reconnaissance des codons de départ et d'initier efficacement la traduction, surtout dans le génome bactérien abondant en ARNm polycistronique (Orr *et al.*, 2020). Par cette méthode, les ribosomes peuvent commencer la traduction à partir de n'importe quelle séquence de liaison, permettant de synthétiser les multiples protéines d'un même ARNm.

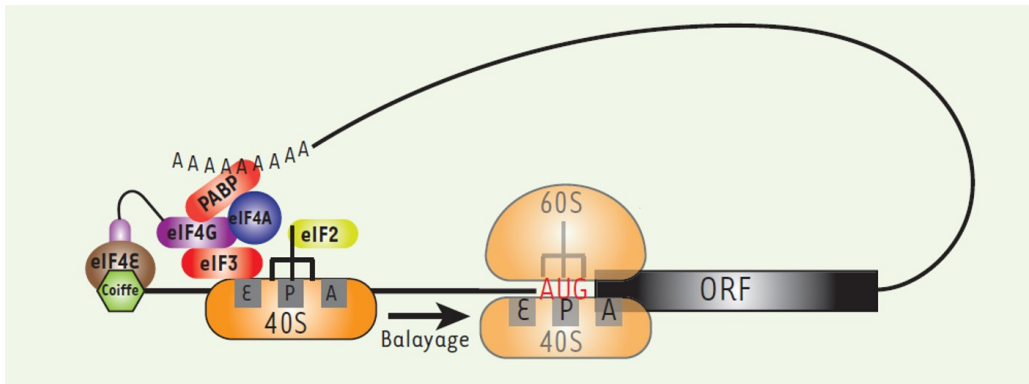


**Figure 3.** Implication de la séquence de liaison au ribosome Shine-Dalgarno (SD) dans la synthèse protéique chez les procaryotes. La liaison de la séquence SD dans la région 5'UTR de l'ARNm avec la séquence anti-SD sur la queue 3' de la petite sous-unité ribosomale facilite la reconnaissance en aval des codons initiateurs de la traduction. Figure tirée et modifiée de Yang *et al.* (2016).

### 1.2.2 Initiation de la traduction des gènes nucléaires chez les eucaryotes

La traduction eucaryote conventionnelle au niveau du noyau est dite coiffe-dépendante, car les ribosomes nécessitent la présence de structures conservées au niveau de l'ARNm afin de pouvoir le reconnaître et le traduire. On y retrouve la coiffe méthylée à l'extrémité 5' et la queue polyadénosine, ou queue polyA, à l'extrémité 3'. L'étape de l'initiation implique plusieurs facteurs d'initiation et protéines dont eIF2 (*eukaryotic initiation factor 2*), eIF4E, eIF4G, eIF3 et PABP (*polyA-binding protein*) (Figure 4). Un groupe de facteurs nommé *cap-binding-complex* se lie d'abord à la coiffe en 5' (Ramanathan *et al.*, 2016). Parmi ceux-ci, le facteur eIF4E interagit avec eIF4G, une protéine d'échafaudage qui se lie avec de multiples facteurs, dont PABP et eIF3. La protéine PABP va permettre de circulariser l'ARNm en se liant à la queue polyA de l'extrémité 3'. Le facteur eIF3, quant à lui, va venir recruter la petite sous-unité ribosomale (40S) associée à des facteurs d'initiation et à l'ARNt initiateur Met-ARNt<sup>Met</sup>. Le ribosome va ensuite balayer le long du transcrit de 5' vers 3' et il sera aidé de l'hélicase eIF4A qui déroule certaines structures secondaires de l'ARN pouvant bloquer sa progression. Le balayage s'arrête à la reconnaissance du premier codon d'initiation AUG possédant un contexte nucléotidique optimal nommé la séquence de Kozak (Blin & Ricci, 2016). Chez l'humain, cette séquence entourant le codon start est représentée par : 5'-GCCGCC(A/G)(C/A)CAUGGCG-3'. La séquence de Kozak est considérée

comme forte si le codon AUG est entouré d'une purine en position -3 ainsi qu'une guanine en position +4 par rapport à l'adénine du codon de départ (Kozak, 1987). Une fois la petite sous-unité ribosomale placée au codon initiateur, la grosse sous-unité (60S) et des facteurs supplémentaires s'y associent pour former un ribosome fonctionnel (80S) permettant l'élongation et la traduction de la protéine (Andrews & Rothnagel, 2014; Blin & Ricci, 2016).

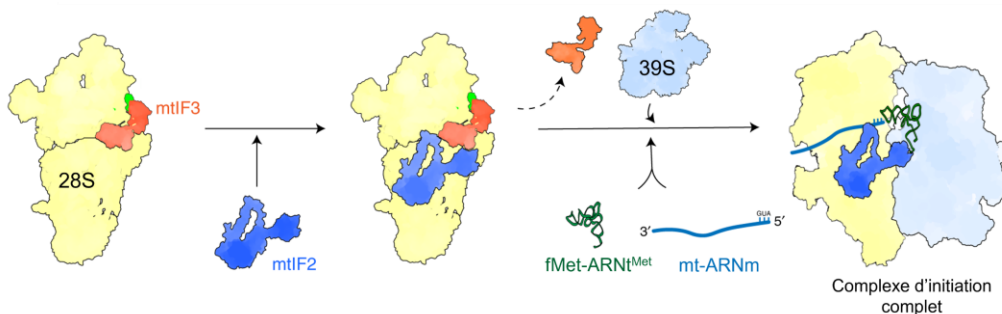


**Figure 4.** Vision globale de l'initiation de la traduction coiffe-dépendante chez les eucaryotes. Les facteurs eIF4E et eIF4G liés à la coiffe méthylée à l'extrémité 5' interagissent avec la protéine PABP, permettant la circularisation de l'ARNm par sa liaison à la queue polyA de l'extrémité 3'. La petite sous-unité ribosomale 40S est recrutée et balaye le transcrit jusqu'à reconnaître un premier codon initiateur. Figure tirée et modifiée de Blin & Ricci (2016).

### 1.2.3 Initiation de la traduction mitochondriale

Bien que la mitochondrie ait une origine bactérienne et possède encore une machinerie de traduction similaire, l'initiation de la traduction mitochondriale a grandement divergé de son homologue procaryote. Notamment, on constate l'absence de la séquence de liaison au ribosome et l'existence de mécanismes spécifiques aux mitochondries qui sont encore en train d'être élucidés (Khawaja *et al.*, 2020; Wang *et al.*, 2021). Tout comme chez les bactéries, il y a d'abord la dissociation du mitoribosome en sa petite (28S) et grande (39S) sous-unité. Ceci mène à la formation éventuelle du complexe d'initiation comprenant la sous-unité 28S, les facteurs d'initiation MtIF2 et MtIF3 (*mitochondrial initiation factor 2 et 3*), l'ARNm et l'ARNt-initiateur avec une méthionine formylée (fMet-ARNt<sup>Met</sup>) (Figure 5). En premier lieu, une liaison entre MtIF3 et la protéine mitoribosomale mS37 (*28S Ribosomal Protein S37, Mitochondrial*) dans la petite sous-unité assure le maintien d'une bonne conformation du ribosome pour le recrutement du facteur MtIF2. Celui-ci assure la liaison de l'ARNt au site peptidique (site P) chez le mitoribosome. Ensuite, la grande sous-unité ribosomale se lie à la petite sous-unité et ceci entraîne

le recrutement de l'ARNt et de l'ARNm. La structure de l'ARNm mitochondrial typique varie de ses homologues nucléaires et procaryotes, car il est dépourvu d'une séquence Shine-Delgarno, d'une coiffe à son extrémité 5' et d'introns. Ses régions UTR sont également absentes ou très courtes, mais il possède une queue polyA en 3' d'environ cinquante nucléotides (Temperley *et al.*, 2010). Le facteur MtIF3 va aider la liaison de l'ARNm à la petite sous-unité et ajuster la position de l'ARNm afin que le codon d'initiation soit enligné correctement au site P. Enfin, l'ARNt vient remplacer MtIF3 au site P de la sous-unité ribosomale pour venir former un complexe d'initiation fonctionnel permettant la traduction (Khawaja *et al.*, 2020; Wang *et al.*, 2021). Chez les mitochondries des mammifères, un codon AUG, AUA ou même AUU pourrait initier la traduction grâce à un code génétique qui varie légèrement de celui du noyau, ainsi qu'une discrimination moins stricte par l'anticodon de l'ARNt (Ayyub & Varshney, 2019). Il a aussi été démontré que les ribosomes mitochondriaux forment préférentiellement des complexes d'initiation au niveau d'un codon initiateur 5'-terminal comparativement à un codon initiateur interne (Christian & Spremulli, 2010).



**Figure 5.** Modèle d'initiation de la traduction dans les mitochondries des mammifères. Dans cette voie, le facteur MtIF3 stabilise la petite sous-unité mitoribosomale 28S pour l'hébergement de MtIF2. La jonction de la grande sous-unité 39S entraîne un changement de conformation chez MtIF2, permettant le recrutement de l'ARNt et l'ARNm pour former le complexe d'initiation de la traduction. Figure tirée et modifiée de Khawaja *et al.* (2020).

## 1.3 L'existence d'un protéome alternatif

### 1.3.1 Le protéome alternatif procaryote

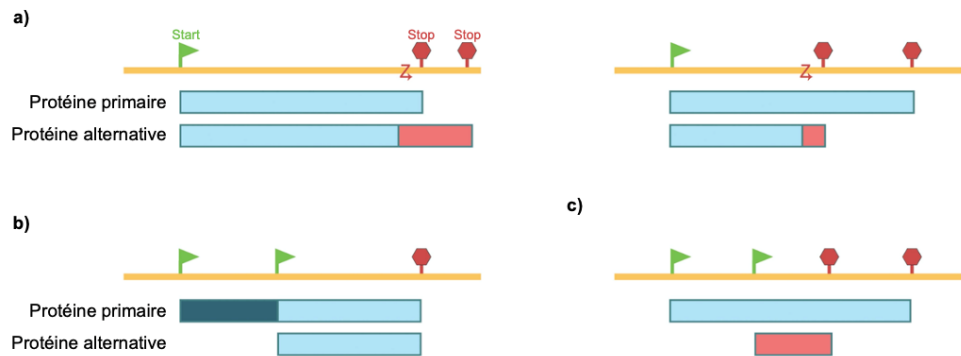
En plus des ARNm polycistroniques, une autre caractéristique remarquable des procaryotes est leur capacité à produire des protéines qui pourraient être décrites comme « alternatives » (par opposition aux protéines dites « de référence », voir plus bas). Le terme alternatif, ici, désigne le résultat des mécanismes d'optimisation du potentiel codant d'un génome par lesquels un gène est situé, au moins en partie, sur une séquence déjà occupée par un autre gène.

Ce gène alternatif peut ensuite coder pour un produit fonctionnel unique tel qu'une protéine alternative ou un ARN non codant. La production de protéines alternatives peut être due à la présence rare de gènes chevauchants antiparallèles (Figure 6A). Pour une même région génomique, la séquence complémentaire d'un gène est normalement non codante, mais le brin opposé d'un gène peut parfois être occupé par un second gène pouvant transcrire un ARNm et coder pour une protéine distincte (Tunca *et al.*, 2009). Un autre mécanisme possible est la présence de gènes à l'intérieur d'autres gènes (Figure 6B). Dans ce cas, lorsqu'un ARNm est produit à partir d'une région de l'ADN contenant des gènes alternatifs, il peut contenir les séquences nécessaires pour coder des protéines issues de plusieurs gènes différents. Ces séquences peuvent se chevaucher dans des cadres de lecture différents, avec des codons start et/ou stop distincts, ce qui leur permet d'être traduits simultanément (Meydan *et al.*, 2018). L'existence de ces mécanismes d'optimisation du potentiel codant chez les procaryotes pousse à se demander si les mitochondries auraient perdu cette capacité à transcrire ou à traduire des séquences alternatives.

#### A. Gènes chevauchants antiparallèles



#### B. Gènes à l'intérieur d'autres gènes



**Figure 6. (A)** Schématisation du principe des gènes chevauchants antiparallèles. Organisation de gènes sur une région d'ADN avec organisation antiparallèle entre les gènes 2 et 3 (figure tirée et modifiée de Tunca *et al.*, 2009). **(B)** Schématisation de différentes stratégies utilisées pour le principe de gènes à l'intérieur d'autres gènes chez les bactéries. Traduction de deux protéines dont (a) le codon stop diffère, (b) le codon start alternatif est présent dans la séquence primaire ou (c) les codons start et stop alternatifs sont présents dans la séquence primaire. Figure tirée et modifiée de Meydan *et al.* (2018).



### 1.3.2 La découverte récente d'un protéome alternatif eucaryote

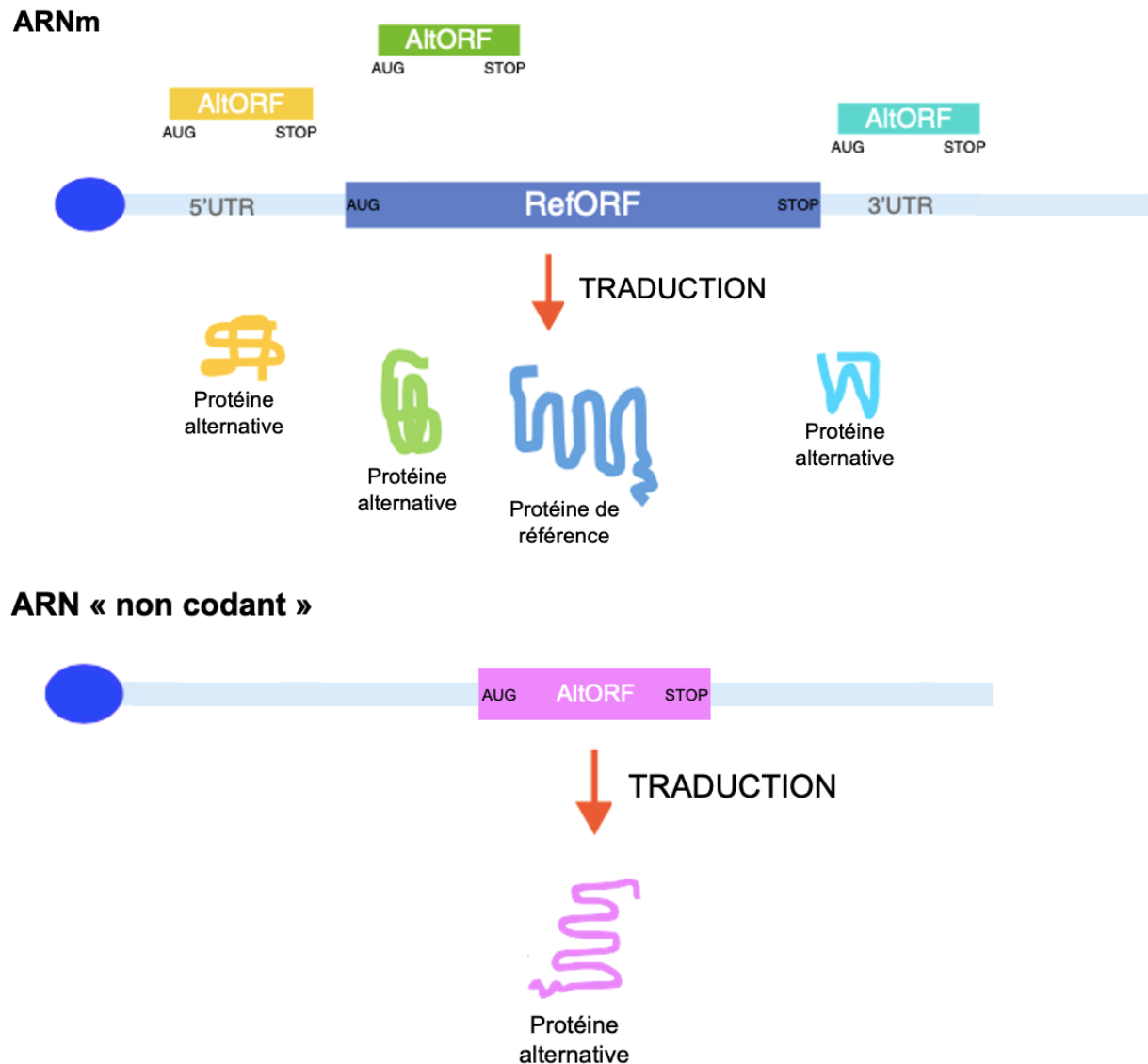
Le décryptage d'un génome consiste, entre autres, à identifier les cadres de lecture ouverts (Open Reading Frame – ORF) qui sont situés entre un codon start et un codon stop et codent pour une protéine fonctionnelle. Jusqu'à tout récemment, en raison des limitations de techniques liées aux méthodes d'analyses computationnelles, un seuil minimal de 100 codons était imposé afin qu'un ORF soit reconnu comme un gène, sans quoi il était plutôt considéré comme le fruit du hasard ou non fonctionnel. Chez les eucaryotes, pour chaque ARNm transcrit, une seule protéine y était associée, soit celle correspondant au plus grand ORF de l'ARNm et considérée comme étant la plus probable d'être produite par cet ARNm. La séquence codant cette protéine devenait alors l'ORF de référence (refORF) annoté dans les bases de données nucléotidiques (Landry *et al.*, 2015; Martinez *et al.*, 2020). Or, la découverte récente de cadres de lecture ouverts fonctionnels inférieurs à 100 codons (short ORF – sORF), avec des rôles clés chez les eucaryotes, notamment par l'encodage de protéines bioactives et la régulation des refORFs (Kotrys & Szczesny, 2019), a mis en lumière une toute nouvelle facette du protéome auparavant négligée.

Avec l'utilisation de techniques génomiques et analytiques plus performantes, ces sORFs ignorés ont pu être détectés en plus grand nombre chez diverses espèces. Par exemple, la drosophile possède le gène *tal* qui code pour un peptide à 11 acides aminés, la plus petite protéine décrite jusqu'à maintenant, et qui assure le bon développement de l'organisme (Pi *et al.*, 2011; Martinez *et al.*, 2020, Su *et al.*, 2013). Chez les mammifères, on retrouve aussi plusieurs petits peptides assurant des rôles essentiels tels que la réparation de l'ADN, la régulation de l'ARN, le développement des muscles et plus encore (Martinez *et al.*, 2020). Même chez l'humain, des analyses bio-informatiques ont démontré que le transcriptome contiendrait plus de 80 000 sORFs potentiels ayant moins de 40 codons alors que le protéome contient, à ce jour, environ 52 000 protéines annotées (Mouilleron *et al.*, 2015).

De plus, ces analyses ont mis en lumière le fait que les eucaryotes possèdent également un protéome pouvant être décrit comme alternatif (Pi *et al.*, 2011; Mouilleron *et al.*, 2015). Pour clarifier, pour un ARNm eucaryote typique, la protéine qui y est conventionnellement attribuée et qui se retrouve dans les bases de données est codée par une séquence qui se nomme l'ORF de

référence (refORF). À ce refORF peuvent s'ajouter plusieurs autres ORFs alternatifs (altORF) qui se définissent comme des séquences situées entre un codon start et stop différents de la séquence de référence et ayant la possibilité de coder pour une protéine dite « alternative » (Figure 7). Dans l'ARNm, ces altORFs peuvent se situer sur des segments dans la région 5'UTR (*untranslated region*) et 3'UTR, soit en amont et en aval du refORF, mais aussi à l'intérieur du refORF ou chevauchant deux régions. Des altORFs peuvent également se trouver sur des segments d'ARN considérés auparavant comme non codants (ARNt et ARNr) (Landry *et al.*, 2015; Orr *et al.*, 2020). Malgré la petite quantité d'altORFs étudiée de manière approfondie, ceux-ci possèdent d'importants rôles. Les ORFs situés en amont d'un refORF peuvent généralement influencer la traduction d'un ORF en aval et peuvent même l'atténuer en réponse à des signaux environnementaux via des interactions avec les ribosomes ou les facteurs de transcription (Andrews & Rothnagel, 2014; Orr *et al.*, 2020). Cependant, les fonctions des ORFs en aval ou chevauchants sont moins connues, mais les produits semblent être davantage ciblés vers diverses organelles (Andrews & Rothnagel, 2014).

Parmi les premiers altORFs découverts chez l'humain, on retrouve *altmid51*, un altORF du gène *miefl* dont la protéine alternative serait intégrée dans les mitoribosomes et servirait de lien entre ceux-ci et certaines protéines de liaison à l'ARN. Un autre exemple est le gène *altslc35a4*, un altORF du gène *slc35a4*, dont la protéine alternative interagit avec plusieurs protéines membranaires et serait surexprimée dans des conditions de stress (Brunet *et al.*, 2020). Plusieurs cas ont été observés chez l'humain où des conditions de stress entraînent une augmentation de la traduction d'altORFs, en amont et en aval des refORFs, comme c'est le cas pour *altslc35a4*. En présence de stress oxydatif, d'une privation d'oxygène, d'une privation de glucose ou autre perturbation, les altORFs sont régulés grâce à des modifications au complexe d'initiation (Andreev *et al.*, 2015b; Brunet *et al.*, 2020). Toutes ces découvertes indiquent que d'importants détails concernant le génome ont trop longtemps été ignorés et méritent d'être étudiés, car le protéome impacte toutes les facettes du bon fonctionnement des individus.



**Figure 7.** ARNm eucaryote typique et ses possibles ORFs comprenant l'ORF de référence (refORF) annoté dans les bases de données nucléotidiques actuelles, ainsi que les ORFs alternatifs (altORF) compris entre des codons start et stop distincts de ceux du refORF et codant pour une protéine alternative qui n'est pas annotée dans les bases de données. Ceux-ci peuvent se retrouver dans les ARNm ou dans les ARN dits « non codants ». Figure tirée et modifiée de Landry *et al.* (2015).

## 1.4 Les mécanismes de la traduction alternative chez les eucaryotes

La cellule eucaryote possède divers mécanismes permettant de réguler la traduction qui peuvent potentiellement être utilisés pour coder des protéines alternatives, comme le *leaky scanning*, la réinitialisation ribosomale, le *frameshifting* et les IRES. Ces mécanismes sont fréquemment utilisés par les virus afin que leurs hôtes eucaryotes traduisent leurs nombreux ARNm polycistroniques et peuvent être utilisés par les eucaryotes eux-mêmes (Blin & Ricci, 2016; Mouilleron *et al.*, 2015).

### 1.4.1 Le *leaky scanning*

Le *leaky scanning* est un mécanisme permettant aux ribosomes d'ignorer des codons start lors du balayage de l'ARNm, résultant en la traduction des altORFs qui se retrouvent en aval (Figure 8A). Spécifiquement, lors du balayage de la région 5', le complexe d'initiation peut soit s'arrêter au niveau du premier codon initiateur AUG et débiter l'élongation, ou bien la petite sous-unité ribosomale peut plutôt continuer de balayer l'ARNm pour atteindre le prochain site d'initiation. Ce phénomène peut même être répété plus d'une fois en ignorant plusieurs codons initiateurs d'affilée afin d'atteindre des altORFs à l'intérieur du refORF ou dans la région 3'UTR. Le nombre de ribosomes s'arrêtant au « premier codon start » est déterminé par les séquences entourant celui-ci, soit la présence d'une forte séquence de Kozak qui peut grandement réduire le *leaky scanning* (Blin & Ricci, 2016; Kozak, 1987; Mouilleron *et al.*, 2015). Le *leaky scanning* est notamment utilisé pour réguler l'expression protéique en réponse au stress pendant la répression du facteur eIF2 (Brunet *et al.*, 2020). En effet, divers stressseurs peuvent déclencher une réponse intégrée au stress chez la cellule par l'activation de kinases spécifiques, dont PKR (présence d'ARN double brin lors d'infection virale), PERK (mauvais repliement de protéines dans le réticulum endoplasmique), HRI (privation d'hème) et GCN2 (privation d'acides aminés) (Andreev *et al.*, 2015b; Sonenberg & Hinnebusch, 2009). Ces kinases mènent à l'inactivation du facteur clé d'initiation de la traduction eIF2 par la phosphorylation de sa sous-unité alpha et la formation d'un complexe avec le facteur d'échange de nucléotides de guanine eIF2B. Ceci cause une inhibition générale de la synthèse protéique, permettant la conservation d'énergie ainsi que de réduire la charge sur la machinerie cellulaire (Sonenberg & Hinnebusch, 2009). Cependant, ce n'est pas l'entièreté de la traduction qui est interrompue. Des mécanismes de lutte au stress, comme le *leaky scanning* qui est favorisé lors d'un chargement réduit des ribosomes, font en sorte que certains

facteurs évitent non seulement cette inhibition de la traduction, mais sont synthétisés davantage en tant qu'éléments intégraux pour répondre adéquatement à l'état de stress, malgré les faibles niveaux d'eIF2 disponibles (Andreev *et al.*, 2015a). Il est intéressant de noter que le *leaky scanning*, lorsqu'utilisé pour la traduction d'ARNm viraux polycistroniques, peut se réaliser avec des sites d'initiation non AUG grâce à la dégénérescence du code génétique (Blin & Ricci, 2016).

#### **1.4.2 La réinitialisation ribosomale**

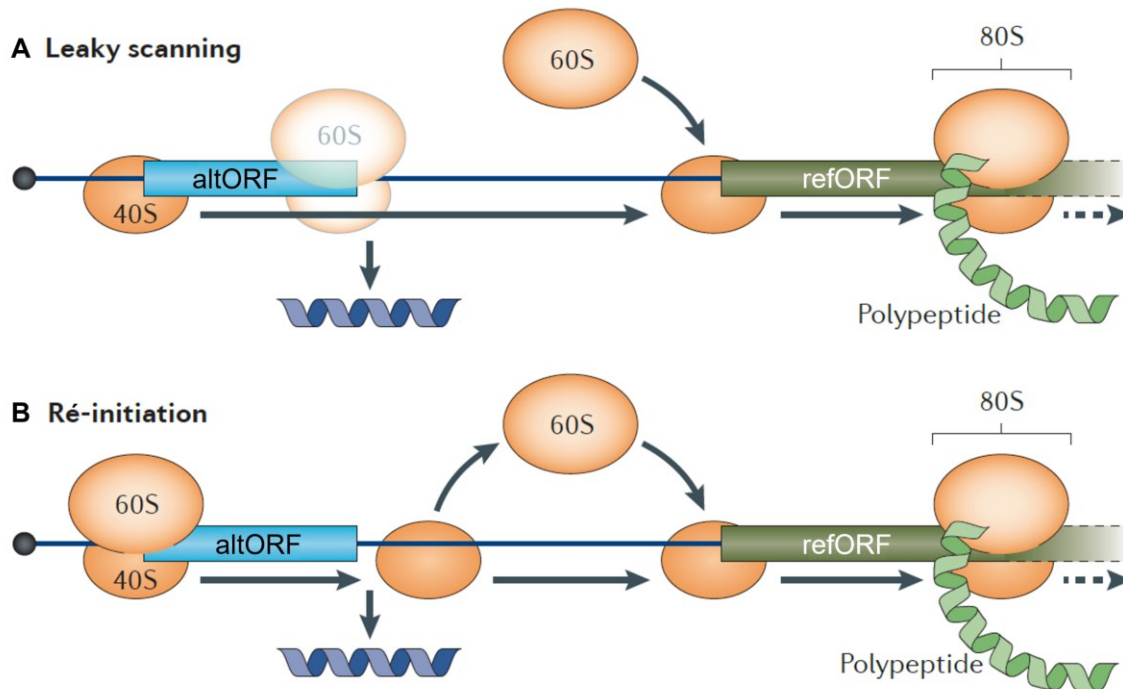
La réinitialisation ribosomale accorde aux ribosomes la capacité d'initier la traduction d'un second ORF malgré la traduction complète d'un ORF au préalable. Après le relargage de la protéine synthétisée à la fin de la traduction, il y a normalement dissociation des sous-unités ribosomales de l'ARNm. Ici, cependant, la petite sous-unité demeure plutôt accrochée et continue de balayer l'ARNm à la recherche du prochain codon start pour initier de nouveau la traduction (Andrews & Rothnagel, 2014; Blin & Ricci, 2016) (Figure 8B). Ce phénomène semble être le plus fréquemment utilisé pour la traduction d'un ou plusieurs altORF dans la région 5'UTR avant la traduction du refORF. La réinitialisation nécessite la présence d'une forte séquence de Kozak chez l'altORF et dépend de facteurs tels que la longueur des ORFs (< 55 codons), la distance les séparant ainsi que des éléments structuraux (Andrews & Rothnagel, 2014; Mouilleron *et al.*, 2015). Dans des conditions de stress, la faible disponibilité en eIF2 signifie une réinitialisation plus lente et donc une plus grande distance parcourue par la sous-unité avant de pouvoir reprendre la traduction, ignorant donc parfois certains altORFs ou refORFs (Andreev *et al.*, 2015a).

#### **1.4.3 Le *frameshifting***

Le *frameshifting* ou glissement de phase de lecture cause un décalage du ribosome pour permettre la traduction dans un autre cadre de lecture. Traditionnellement, le ribosome fait la lecture de l'ARNm par groupes de trois nucléotides, soit un codon. Lors du *frameshifting*, le ribosome est ralenti par une structure secondaire de l'ARNm, ce qui cause un décalage d'un ou deux nucléotides. Le ribosome se retrouve alors dans un cadre de lecture différent qui traduit pour une protéine entièrement distincte de celle du refORF, et ce, malgré une séquence nucléotidique identique (Andrews & Rothnagel, 2014; Blin & Ricci, 2016).

#### 1.4.4 Les IRES

Les IRES (*internal ribosome entry site*) sont des structures de l'ARN capables de recruter la machinerie de traduction dans des conditions où la traduction coiffe-dépendante est perturbée. Les séquences IRES sont hautement variables et possèdent plutôt des régions structurées pouvant directement interagir avec le ribosome ou certains facteurs d'initiation. Cette variabilité crée une diversité de mécanismes pouvant initier la traduction grâce à l'utilisation de facteurs variables ou même en l'absence complète de facteurs d'initiation. Ces séquences semblent assez limitées dans les transcrits eucaryotes, mais elles assureraient la synthèse de protéines importantes en conditions de stress (Blin & Ricci, 2016; Mouilleron *et al.*, 2015). Toutefois, leur rôle dans la cellule et la traduction des altORFs est encore sujet de débats (Yang & Wang, 2019).



**Figure 8.** Deux méthodes les plus connues de la traduction alternative : **A)** Lors du *leaky scanning*, la sous-unité ribosomale 40S peut soit reconnaître un codon start en amont (altORF) et conduire à la traduction d'une protéine, soit balayer au-delà et initier la traduction au niveau d'un codon start en aval (refORF). **B)** La réinitialisation ribosomale se produit lorsque la traduction est initiée au codon start d'une séquence d'ADN codante en aval (refORF) après avoir terminé la traduction d'un ORF en amont (altORF) en raison de la non-dissociation de la sous-unité 40S de l'ARNm. La traduction de l'altORF conduit à la production d'un court peptide. Figure tirée et modifiée de Andrews & Rothnagel (2014).

## 1.5 Le protéome alternatif mitochondrial

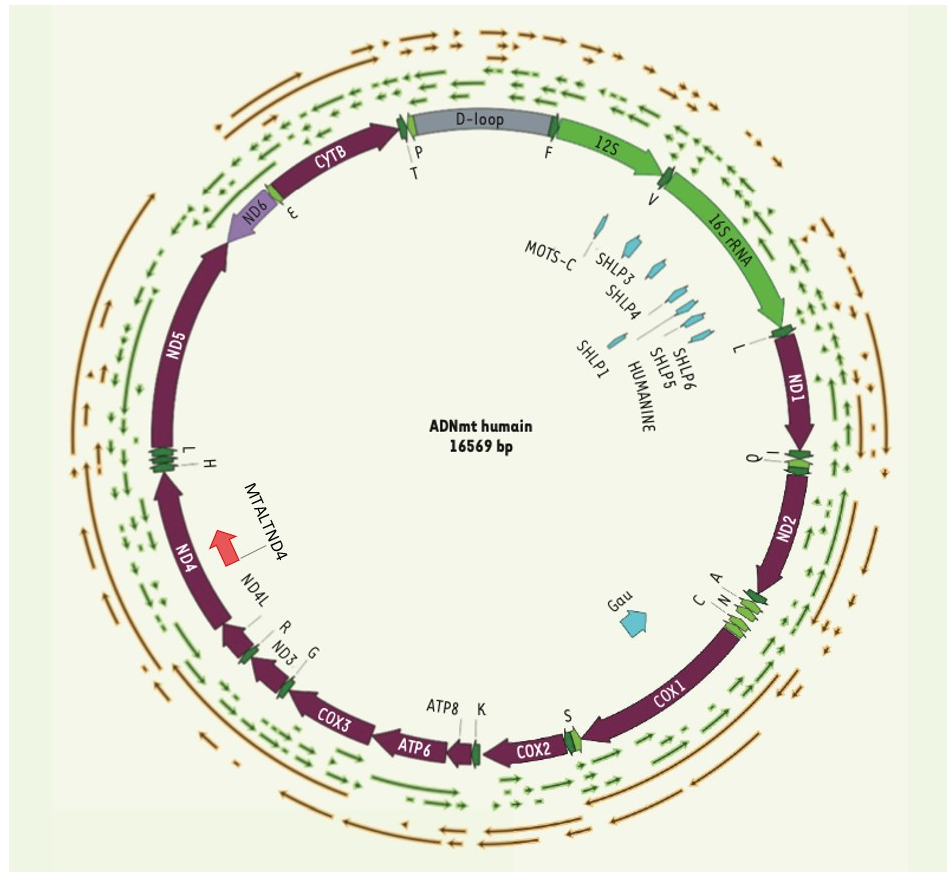
L'existence d'un protéome alternatif n'est donc pas exclusive au génome procaryote, mais il est également présent dans le génome nucléaire des eucaryotes. Mais qu'en est-il du génome mitochondrial, descendant de l'ADN bactérien? Nous savons déjà que les mitochondries peuvent posséder plus d'un gène par ARNm. En effet, la mitochondrie humaine possède deux ARNm dicistroniques : un codant pour les protéines ATP8/ATP6 et un autre pour ND4L/ND4 (Figure 1). Dans ces cas, les deux gènes se retrouvent dans des ORFs qui se chevauchent de quelques nucléotides (46 nucléotides pour *atp8/atp6* et 7 nucléotides pour *nd4l/nd4*) (Christian & Spremulli, 2012). Il a été découvert que la synthèse des deux protéines se produit grâce à une seule interaction entre l'ARNm et le ribosome. Puis, la traduction de la protéine en aval (ATP6 ou ND4) dépend de la traduction de celle en amont (ATP8 ou ND4L). De plus, une hypothèse proposée est que, suivant la synthèse de la première protéine, le ribosome reculerait pour ensuite « réinitialiser » la traduction de la seconde protéine (Cruz-Zaragoza *et al.*, 2021). C'est donc une forte possibilité que des mécanismes similaires à ceux décrits plus haut puissent être impliqués pour traduire des protéines alternatives chez la mitochondrie.

De précédentes recherches ont permis d'identifier de petits peptides alternatifs dérivés de l'ADNmt humain (*mitochondrial-derived peptides* – MDP) ayant un large spectre de fonctions. Le premier MDP découvert est l'Humanine (24 acides aminés), dont le gène est situé à l'intérieur de la séquence de l'ARNr 16S (Figure 9) (Hashimoto *et al.*, 2001). Celui-ci assure, par exemple, des fonctions neuroprotectrices contre la maladie d'Alzheimer et offre une protection contre le stress oxydatif et l'apoptose (Angers *et al.*, 2019). Ensuite, MOTS-c (*mitochondrial open reading frame of the 12S rRNA-c* – 16 acides aminés), dont le gène est situé dans l'ARNr 12S (Figure 9), joue un rôle dans la régulation du métabolisme mitochondrial, de l'insuline et de l'obésité et il contribue à la longévité (Angers *et al.*, 2019; Lee *et al.*, 2015). Puis, les « small Humanin-like proteins » 1 à 6 (SHLPs – entre 20 et 38 acides aminés) sont des peptides également encodés dans la région 16S (Figure 9) qui assurent la régulation du métabolisme mitochondrial, de l'apoptose et de la prolifération cellulaire. Certains SHLPs offrent aussi une protection contre la dégénérescence maculaire et possèdent une sensibilité à l'insuline (Angers *et al.*, 2019; Cobb *et al.*, 2016). D'ailleurs, l'expression de MOTS-c et des SHLPs peut varier entre les tissus (Cobb *et al.*, 2016; Lee *et al.*, 2015). De plus, SHMOOSE (*Small Human Mitochondrial ORF Over SERine tRNA* – 58

acides aminés) est le MDP le plus récent qui se situe majoritairement dans une séquence codant pour un ARNt (sérine) et chevauche le gène ND5 (Miller *et al.*, 2022). Il agit dans le cerveau en affectant l'expression génique mitochondriale et augmente la consommation d'oxygène. Enfin, le gène mitochondrial *Gau* se situe sur le brin complémentaire du gène *cox1* (Figure 9), mais ses fonctions demeurent inconnues (Angers *et al.*, 2019; Faure *et al.*, 2011).

La régulation des MDPs lors d'un stress a déjà été étudiée. Certains MDPs dans les muscles humains, par exemple, sont surexprimés lors de stress perturbant les mitochondries, tels que l'exercice physique (Merry *et al.*, 2020). Les MDPs ont d'étonnants mécanismes d'action leur permettant d'interagir avec des facteurs de transcription afin d'influencer l'expression génique. MOTS-c, par exemple, a la capacité de transloquer au noyau en coordination avec la protéine AMPK (*Adenosine Monophosphate-Activated Protein Kinase*) en réponse au stress métabolique. Il peut alors réguler l'expression des gènes de réponse antioxydante (*Antioxidant response element* – ARE) dans le noyau en interagissant avec le facteur de transcription NRF2 (*Nuclear factor erythroid 2-related factor 2*) sensible aux ROS issus du stress oxydatif (Kim *et al.*, 2018). Humanine, quant à lui, peut se lier à un récepteur trimérique à la surface cellulaire composé des protéines WSX1 (*Interleukin-27 receptor subunit alpha*), GP130 (*Glycoprotein 130*) et CNTFR (*Ciliary neurotrophic factor receptor*) ou bien au récepteur FPRL1 (*Formyl peptide receptor-like 1*). Une fois activé, plusieurs cascades de signalisation intracellulaires sont initiées impliquant des facteurs de transcription comme STAT3 (*Signal Transducer and Activator of Transcription 3*) et des kinases comme ERK1/2 (*Extracellular signal-regulated kinases 1 and 2*) et pAkt (*phosphorylated protein kinase B*) qui affectent la survie cellulaire, la prolifération et la biologie mitochondriale (Miller *et al.*, 2020). Le stress métabolique affecte donc non seulement le patron d'expression des MDPs, mais ces peptides influencent, eux aussi, l'expression des gènes. Sachant que plusieurs dysfonctions mitochondriales sont reliées à des maladies telles que l'Alzheimer, le Parkinson, la fibrose kystique, les cancers, les maladies cardiovasculaires ainsi que le diabète, l'obésité et l'autisme (Wallace, 2018), davantage de recherches sur les MDPs et leurs effets protecteurs et régulateurs mèneront certainement à des progrès importants sur le plan médical ainsi qu'au développement d'outils thérapeutiques pour lutter contre ces dysfonctions.





**Figure 9.** Carte du génome mitochondrial humain montrant l'emplacement des nouveaux altORFs caractérisés (Humanine, MOTS-c, SHLPs et Gau en bleu), incluant le nouveau gène *mtaltnd4* (en rouge) et tous les cadres de lecture ouverts potentiels sur les brins sens (flèches orangées) et antisens (flèches vertes). Figure tirée et modifiée d'Angers *et al.* (2019).

## 1.6 La découverte de MTALTND4

L'existence de petits altORFs fonctionnels, cachés dans les séquences ARNr 12S et 16S, et codant pour Humanine, MOTS-c et les SHLPs, a mené à l'hypothèse que des altORFs pourraient aussi se retrouver à l'intérieur des autres gènes mitochondriaux ou sur le brin complémentaire. Pour vérifier ceci, une réanalyse du génome mitochondrial humain a été réalisée et a révélé l'existence de plus de 200 altORFs d'au moins 20 codons pouvant potentiellement coder pour des protéines fonctionnelles (Angers *et al.*, 2019) (Figure 9). Une récente étude (à laquelle j'ai activement participé avec ce projet de maîtrise) de neuf de ces altORFs a démontré qu'au moins un d'entre eux code pour une protéine fonctionnelle (Kienzle *et al.*, 2022 – Annexe 1).

Kienzle *et al.* (2022) rapportent l'existence d'un altORF situé dans le gène mitochondrial *nd4* (Figure 9) et nommé *mtaltnd4*. Cet altORF code pour une protéine de 99 acides aminés avec une séquence relativement bien conservée chez certains mammifères dont les primates. Le peptide MTALTND4 est localisé dans le cytoplasme et dans les mitochondries. Sa présence a aussi été détectée dans le plasma (Annexe 1.1 : Figure 2G; Annexe 1.2 : Figure S8), suggérant une potentielle fonction de signalisation. Cette protéine aurait notamment un impact sur la physiologie mitochondriale. En effet, les résultats *in vitro* suggèrent que ce peptide impacte négativement le taux de respiration cellulaire, tout en augmentant l'afflux de ROS, induisant une possible dépression bioénergétique. De plus, la protéine multifonctionnelle C1QBP (*complement component 1q subcomponent binding protein*) serait un partenaire d'interaction potentiel, identifié par la technique de pull-down (Annexe 1.1 : Table S5 – objectif 4) et confirmé par western blot (Annexe 1.2 : Figure S9 – objectif 4). Ces découvertes ouvrent la porte à un large potentiel de fonctions à explorer.

## 1.7 Hypothèses et objectifs

Ce projet de recherche a pour but de préciser les fonctions du nouveau peptide alternatif MTALTND4. Nous posons l'hypothèse que, tout comme les autres gènes alternatifs mitochondriaux découverts jusqu'à présent, le gène *mtlnd4* joue un rôle dans les grandes fonctions associées à la mitochondrie (*e.g.* métabolisme, régulation du stress oxydatif, prolifération, *etc.* [Kotrys & Szczesny, 2019]). On s'attend aussi à ce que le peptide se comporte de manière similaire aux autres MDPs étudiés. Autrement dit, MTALTND4 possédera une expression tissu-spécifique, son expression endogène sera influencée par le stress et le peptide affectera l'expression génique et le métabolisme cellulaire (Cobb *et al.*, 2016; Hashimoto *et al.*, 2001; Lee *et al.*, 2015). Ce projet est composé de quatre objectifs. Le premier vise à analyser les patrons d'expression de la protéine MTALTND4 dans différents tissus humains. Le second objectif vise à exposer une lignée cellulaire humaine à différentes conditions de stress (milieu riche en ROS, milieu de jeûne, milieu hypoxique) ayant un impact sur les fonctions mitochondriales afin de mesurer l'effet sur le patron d'expression de *mtaltnd4*. Le troisième objectif consiste à étudier le transcriptome de cellules exposées au peptide de manière exogène pour déterminer son effet sur les autres gènes cellulaires. Le quatrième objectif consiste en la confirmation des partenaires d'interaction de MTALTND4 par la technique de pull-down pour deux lignées cellulaires.

## **2. ARTICLE : Functional characterization of the new human mitochondrial gene *mtatnd4***

Thierry Choquette, Laura Kienzle, Hajar Hosseini Khorami, Yousef Alaessa, Stefano Bettinazzi, Annie Angers, Sophie Breton\*

Département de Sciences Biologiques, Université de Montréal, Montréal, Québec, Canada

**\*Corresponding author:** Sophie Breton, Département de Sciences Biologiques, Université de Montréal, Montréal, QC H3C 3J7, Canada

Email: [s.breton@umontreal.ca](mailto:s.breton@umontreal.ca)

Article en préparation

## ABSTRACT

Mitochondria are involved in numerous cellular functions, and its mitogenome's coding potential has been underestimated. Recent studies have shown that its genes could hide several small alternative genes (genes with translation initiation or termination sequences distinct from the reference gene sequences in which they are found) coding for micropeptides with a broad spectrum of functions. Following the discovery of the mitochondrial alternative gene *mtaltn4*, we try to elucidate the functions of the gene and its peptide (MTALTND4) by studying its expression in human tissues and in response to stress, its effects on gene expression and its interacting partners. Our research has revealed that MTALTND4 exhibits tissue-specific expression and shows potential as a stress-responsive protein inducing a state of bioenergetic depression. This bioactive peptide could act as a signaling molecule to regulate various aspects of cell physiology, including metabolic processes, protein activity, mitochondrial transport, transcription, and more, through putative interaction with C1QBP and YBX1. Further studies on alternative peptides, such as MTALTND4, could provide valuable insights into the functions of mitochondria in cells and present novel therapeutic targets with significant potential.

## INTRODUCTION

The mitochondrion is an organelle of bacterial origin that is best known for its role in metabolism and energy production (Kotrys & Szczesny, 2019; Roger *et al.*, 2017). However, mitochondria are also involved in a multitude of other cellular functions, including reactive oxygen species (ROS) production, cell proliferation, apoptosis, cellular signaling, aging, calcic regulation, immunity, diseases and more (Kotrys & Szczesny, 2019). Mitochondria are also known to possess their own genome, a vestige of their bacterial ancestors (Roger *et al.*, 2017). In humans, the mitochondrial DNA (mtDNA) is traditionally described as having a functional repertoire limited to 37 described genes involved in energy production and mitochondrial translation (Anderson *et al.*, 1981; Taanman, 1999). However, recent studies demonstrated that the coding potential of the mitogenome has been underestimated, by revealing the existence of micropeptides derived from previously ignored short open reading frames (sORF) and alternative ORFs (altORFs with a frameshift relative to their reference gene's reading frame) present within previously identified reference genes (Angers *et al.*, 2019; Capt *et al.*, 2016; Kienzle *et al.*, 2022; Miller *et al.*, 2020; Miller *et al.*, 2022). Such micropeptides can have vastly different functions compared to their

reference protein and their expression can even be independently regulated in response to stress or metabolic changes (Brunet *et al.*, 2020; Merry *et al.*, 2020; Miller *et al.*, 2020).

The mitochondrial-derived micropeptides (MDP) Humanin, MOTS-c, SHLP1-6 (Small Humanin Like Peptides) and SHMOOSE appear to modulate cell metabolism and global physiology (Cobb *et al.*, 2016; Hashimoto *et al.*, 2001; Lee *et al.*, 2015; Miller *et al.*, 2020). They possess a broad spectrum of functions such as neuroprotection and insulin regulation, and could protect against various diseases including Alzheimer's disease, macular degeneration, diabetes, cardiovascular disease, and prostate cancer (Angers *et al.*, 2019; Miller *et al.*, 2020). MDPs play important roles in homeostasis and exert functions through binding to both intracellular molecules and cell membrane receptors (Cobb *et al.*, 2016; Kim *et al.*, 2018; Lee *et al.*, 2015; Miller *et al.*, 2020). For example, Humanin increases mitochondrial respiration, cell proliferation and cell survival through cell membrane receptors (Miller *et al.*, 2020). MOTS-c translocates to the nucleus and interacts with transcription factors to regulate gene expression in response to metabolic stress (Kim *et al.*, 2018). It is worth noting that MOTS-c was first identified as a gene expression regulator following microarrays performed on HEK-293T cells exogenously treated with synthetic MOTS-c, which indicated a marked effect after a 4-hour treatment on expression of genes associated with cellular metabolism and inflammation (Lee *et al.*, 2015). Globally, these recent findings highlight the need for systematic analyses of the unresolved protein-coding potential of human mtDNA.

The most recently discovered human MDP is MTALTND4, an altORF found within the mitochondrial *nd4* gene that is relatively well-conserved in some mammals and in many primates (Kienzle *et al.*, 2022). *Mtaltnd4* is translated in the mitochondria into a 99 amino acid peptide that can be found in mitochondria, in the cytoplasm as well as in plasma, possibly acting as a signaling molecule that helps regulate cellular physiology through a putative interaction with the complement C1QBP, a multifunctional and multicompartamental protein that regulates cellular energy metabolism (Kienzle *et al.*, 2022; Wang *et al.*, 2022). To date, MTALTND4 has been studied only in HeLa or HEK-293T human cell lines, and nothing is known about its expression in human tissues or in different conditions known to affect mitochondrial physiology. Moreover, it is not yet known whether this plasmatic peptide affects gene expression and its putative interacting partners have not been thoroughly studied. The present study aims to address these knowledge gaps.

## **METHODS**

### **Cell culture & sample preparation**

#### *Cell culture and collection*

HeLa and HEK-293T cell lines were cultured following ATCC's animal cell culture guide (2022). Cells were maintained at 37°C with 5% CO<sub>2</sub> in complete culture media that was prepared by supplementing low glucose (1 mg/ml) Dulbecco's Modified Eagle's Medium (DMEM) with 10% calf bovine serum, 1% penicillin and streptomycin antibiotics and 1% fungizone antifungal. Cells were kept in 10 cm Petri dishes and passaged before reaching confluency. In summary, semi-confluent cells were washed with PBS and 2 ml of Trypsin-EDTA solution (0.05%) was added. Cells were then incubated for 2 minutes at 37°C and, once detached, 10 ml of complete growth medium was added to the cell suspension to inactivate the trypsin. Detached cells were transferred into a 15 ml tube and were centrifuged for 5 minutes at 1000 rpm. The supernatant was discarded, and the cell pellet was resuspended in 2 ml for cell counting using a hemacytometer. A new Petri dish was prepared with 10 ml of fresh culture medium with the desired number of cells for splitting, and it was placed back into the incubator. For cell collection, cultivated control or treated samples were washed with PBS and ice-cold lysis buffer was added comprised of 10 mM HEPES, 150 mM NaCl and a cocktail of protease inhibitors (benzamidine, PMSF, aprotinin and leupeptin). Collected cells were kept on ice.

#### *Sample preparation*

The various tissue samples (kidney, heart, testis, brain, liver, skeletal muscle, spleen) were obtained from Novus Biologicals (NB820-59231; NB820-59217; NB820-59266; NB820-59177; NB820-59232; NB820-59253; NB820-59259). Blood-derived human plasma containing K2 EDTA anticoagulant was obtained from Innovative Research (IPLASK2E2ML). Tissue and cell samples were lysed with a sonic dismembrator sonicator for 3 X 3 seconds. Triton X-100 10% was added to a final concentration of 1% and samples were incubated for 20 minutes at 4°C. They were centrifuged at 15 000 rpm for 15 minutes at 4°C and the supernatant was collected and kept on ice. Protein concentration was then assessed by a Bradford protein assay. Samples were suspended in Laemmli Sample Buffer (LSB) 1X and heated for 5 minutes at 95°C.

## Western blotting

A Tris-Tricine SDS-PAGE was performed using a 3-part polyacrylamide gel (Separation gel: 16,5% T, 6% C; Spacer gel: 10% T, 3% C; Stacking gel: 5% T, 3% C), ideal for protein separation ranging from 1 to 100 kDa (Schägger & von Jagow, 1987). The separation gel was first poured into an electrophoresis montage and polymerized using APS 10% and TEMED. Following the same process, it was then overlaid by the spacer gel (2-3 cm) and then again by the stacking gel (1-2 cm). Samples were then loaded onto the gel at equal protein concentrations (~100-200 µg). Tricine anode and cathode buffers were added as lower and upper electrode buffers respectively. Samples were migrated at low voltage (~43 V) overnight at room temperature. The gel was then transferred onto a previously hydrated PVDF membrane with a 0.2 µm pore size, allowing for better retention of smaller proteins. The transfer was executed at 1000 mA for an hour at 4°C in electrophoresis transfer buffer (Towbin buffer). The membrane was then incubated for an hour at room temperature in a PBS blocking buffer with 5% non-fat dry milk and 0.05% tween-20. Membranes were then incubated with primary antibodies for 2 hours at room temperature. The primary antibodies used were diluted in a PBS solution with 0.05% tween-20 as follows: rabbit anti-MTALTND4 (1:1000; MediMabs), mouse anti-ATP5 (1:4000; Abcam, ab14748), rabbit anti-MTND4 (1:1000; Boster, A04180-1), mouse anti-C1QBP (1:1000; Abcam, ab29889). Following primary incubation, membranes were washed 3 X 10 minutes in TBS-T. They were then incubated with goat anti-rabbit IgG (1:2000) or goat anti-mouse IgG (1:2000) secondary antibodies coupled to horseradish peroxidase (HRP). These were diluted in PBS-T and membranes were incubated for 30 minutes at room temperature. They were washed 3 X 10 minutes in TBS-T and 1 X 5 minutes in TBS. Equal parts of chemiluminescent substrate and enhancer solutions (Montreal Biotech Inc.; western blot detection kit) were added to the membranes and signals were captured using a FUSION FX chemiluminescence imaging system.

## Physiological cell treatments

### *Hydrogen peroxide treatment*

For ROS treatment, semi-confluent (70-80%) HeLa cells were washed with PBS and media was replaced with DMEM containing 0, 0.5, 1, 5 or 10 mM of hydrogen peroxide (Abcam, 2021; Murray *et al.*, 2004). Treated cells were incubated for 30 minutes at 37°C before being collected, lysed and frozen.

### *Starvation treatment*

Serum-free media was prepared in the same way as complete growth media, though lacking the usual 10% calf bovine serum supplement, and therefore removing growth factors and other small molecules essential for normal cell growth (Johnson, 2012). For the starvation treatment, HeLa cells were first washed with PBS followed by either serum-free media or complete media for controls. Serum-free DMEM was added to treated Petri dishes and they were incubated at 37°C for 6 or 24 hours (Aghababazadeh & Kerachian, 2014). They were then collected, lysed and frozen.

### *Hypoxia treatment*

Commercial hypoxic and normoxic HeLa cell lysates were obtained from Novus Biologicals (NBP2-36452). HeLa cells were cultured until semi-confluent (70-80%). Normoxic cells were grown in an environment containing 21% pure oxygen, while hypoxic cells were grown under true hypoxic conditions where pure oxygen levels were reduced from 21% to 2% for 4 hours. Success of hypoxic treatment was confirmed by detection of HiF-1a levels.

### *Statistical analysis*

Western blot band density of MTALTND4 from cell treatments was quantified using ImageJ version 1.53k (Schneider *et al.*, 2012). MTALTND4 density was adjusted to ATP5 loading control. Data was analyzed in GraphPad Prism (version 9.3.1, GraphPad Software, San Diego, California USA, [www.graphpad.com](http://www.graphpad.com)) using independent parametric unpaired *t*-tests with Welch's correction to compare the differences between control and experimental groups for analysis, such as ROS, starvation and hypoxia. Values shown are presented as mean  $\pm$  SEM (standard error of mean).  $P < 0.05$  was considered statistically significant.

## **Transcriptomics**

### *Exogen peptide treatment*

Semi-confluent HeLa and HEK-293T cells were suspended in MiR05 mitochondrial respiratory buffer (110 mM D-sucrose, 60 mM lactobionic acid, 20 mM taurine, 20 mM HEPES, 10 mM  $\text{KH}_2\text{PO}_4$ , 3 mM  $\text{MgCl}_2$ , 0.5 mM EGTA, BSA 1  $\text{g}\cdot\text{L}^{-1}$ ) supplemented with 5 mM pyruvate as an external energy source to allow respiration of intact cell (Doerrier *et al.*, 2018; Kienzle *et al.*, 2022). Treated cells were exposed exogenously to 10  $\mu\text{M}$  of synthetic MTALND4 peptide



(LifeTein) for 4 hours, while water was added to control cells. Cells were then pelleted by centrifugation. The supernatant was then removed and samples were frozen. In total, samples were collected for treated HeLa cells, control HeLa cells, treated HEK-293T cells and control HEK-293T cells. RNA of treated and control samples was extracted using Quick-RNA™ Miniprep Kit as described by the manufacturer (Zymo Research, 2021; R1054). In brief, pelleted cells were resuspended in provided RNA lysis buffer. RNA was purified by filtration through centrifuge columns to remove genomic DNA and washed multiple times using ethanol, prep buffer and wash buffer. RNA was collected and DNase/RNase-Free Water was added. Quality and quantity of the RNA were assessed by electrophoresis on a 1% agarose gel as well as with a BioDrop  $\mu$ LITE spectrophotometer. Quality samples were then sent to Génome Québec for a Human Clariom S Assay via ThermoFisher Scientific (Affymetrix).

### *Statistical analysis*

Processed data from the microarray analysis was interpreted using ThermoFisher's Transcriptome Analysis Console (TAC) Software version 4.0.2.15 (Affymetrix, Inc., Santa Clara, CA, USA). Principal component analysis (PCA) was performed and the analysis was executed using the Gene Level - SST-RMA summarization method and an ANOVA method with eBayes approach as determined by the TAC guide (ThermoFisher Scientific, 2019, Transcriptome Analysis Console 4.0.2.). Significant genes were selected by a false-discovery rate  $< 0.30$ , fold-change  $> 1.5$  in both directions and an ANOVA  $p$ -value  $< 0.05$ , as used previously for similar analyses (Lee *et al.*, 2015). Over and under-expressed genes were compared between HeLa and HEK-293T and significant genes common in both cell types were further analyzed.

Gene ontology was performed on all significantly upregulated or downregulated genes using a PANTHER Overrepresentation Test (GO Ontology database, released 2022-03-22, doi:10.5281/zenodo.6399963; Ashburner *et al.*, 2000; The Gene Ontology Consortium, 2020; Mi *et al.*, 2018). This allowed to identify impacted biological processes and molecular functions in response to exogenous MTALTND4 treatment. Up or down-regulated gene lists were compared to a custom *Homo sapiens* reference list of all genes detected in the microarray analysis. For results, Fischer's Exact Test, with FDR correction was used and significant GO were determined with  $p < 0.05$ .

Enrichment analysis of bio-pathways was performed on either upregulated or downregulated genes common in both cell types using DAVID Bioinformatics Resources against KEGG, Biocarta and Reactome pathway databases (Huang *et al.*, 2009; Sherman *et al.*, 2022). Fisher's Exact test was used to measure the gene enrichment in annotation terms and Fisher's Exact *p*-values were computed by summing probabilities *p* over defined sets of tables. Functional Annotation Clustering was then used to group similar significant pathways ( $p < 0.05$ ) together containing similar gene members.

### **Pull-down assay**

#### *Competent cells transformation and protein expression*

Competent BL21(DE3) cells from *E. coli* B strain (ThermoFisher, EC0114) were transformed following the manufacturer's instructions (ThermoFisher Scientific, 2019, BL21(DE3) Competent Cells). To perform protein expression, transformant colonies were picked and inoculated in LB medium with ampicillin, growing to saturation overnight at 37°C. The overnight culture was added to fresh LB medium and was incubated for 2 hours. A concentration of 0.5 mM of IPTG was then added, and the culture was incubated for an additional 2 hours. The induced culture was centrifuged at 2500 rpm for 10 minutes and the pellet was resuspended in PBS 1X with 4M urea. The bacterial pellet was lysed in PBS 1% triton X-100 and a cocktail of protease inhibitors.

#### *Protein extraction & pull-down assay*

For GST and GST-MTALTND4 fusion protein extraction, the bacterial lysates were incubated overnight with 100  $\mu$ l of a 50% slurry of Glutathione Sepharose<sup>®</sup> 4B (BioWorld, 20181088) beads at 4°C, allowing for pre-coupling. The beads were then washed 3 times with PBS 1X. Different quantities of the beads (1 and 10  $\mu$ l) were loaded onto an SDS-PAGE gel along with a bovine serum albumin gradient (BSA; 0.5 to 10  $\mu$ g) and stained with Coomassie blue to estimate protein concentration and confirm expected band sizes. Pre-clearing of HeLa cell lysates was performed by adding 50  $\mu$ l of Glutathione Sepharose beads and 25  $\mu$ g of synthetic GST protein (Sino Biological, 11213-HNAE) and incubating for 2 hours at 4°C with end-over-end mixing. The lysates were then centrifuged for 2 minutes at 15 000 rpm at 4°C and the supernatant was collected.

For pull-down assays (Cold Spring Harbor Laboratory, 2004), ~10 µg of pre-coupled GST or GST-MTALTND4 protein was added to 1 ml of the pre-cleared HeLa lysates and incubated overnight at 4°C with end-over-end mixing. Samples were washed 6 times with lysis buffer and suspended in LSB 1X. Samples were then sent to the proteomics platform of the McGill University Health Center (MUHC) Research Institute for mass spectrometry analysis (UPLC-MS/MS – Ultra performance liquid chromatography-tandem mass spectrometry) and analyzed with Mascot (Matrix Science) against a human database. The final report was generated using Scaffold5 (Proteome Software, Inc., Portland, OR 97219, USA).

### *Statistical analysis*

Processed data was analyzed based on total spectrum count and using a protein and peptide threshold of > 95% significance. Only proteins presenting a spectrum count  $\geq 2$  compared to their control counterpart were conserved. The remaining proteins were then compared between the three pull-down assays (two on HeLa and one on HEK-293T) and proteins present in all three were conserved for further functional analyses.

Enrichment analysis of bio-pathways was performed on common interaction partners using DAVID Bioinformatics Resources against KEGG, Biocarta and Reactome pathway databases (Huang *et al.*, 2009; Sherman *et al.*, 2022). Fisher's Exact test was used to measure the gene enrichment in annotation terms and Fisher's Exact *p*-values were computed by summing probabilities *p* over defined sets of tables. Functional Annotation Clustering was then used to group pathways together containing similar gene members.

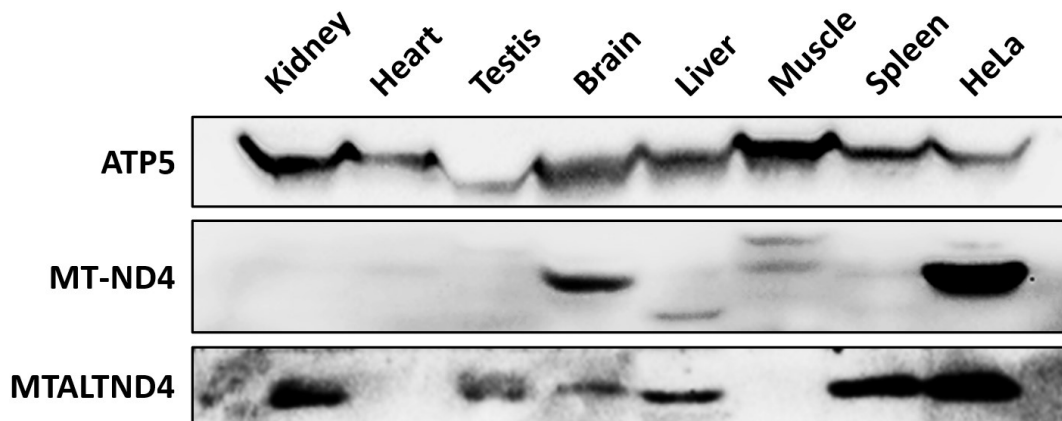
To further compare relevant pathways between transcriptome and pull-down experiments, results from the two DAVID Bioinformatics analyses performed on commonly regulated genes in both cell types from transcriptomics and on potential interacting partners from pull-down assays were compared and common resulting pathways were identified.

Furthermore, a pathway overrepresentation analysis using the Reactome Knowledgebase (Gillespie *et al.*, 2022) was performed on both data sets: the commonly regulated genes in both cell types from transcriptomics and potential interaction partners from pull-down assays. A hypergeometric distribution was used, corrected for false discovery rate using the Benjamini-Hochberg method with *p*-value < 0.05. Both results were compared and common significant pathways were identified.

## RESULTS

### MTALTND4 expression characterized in various tissues and physiological conditions

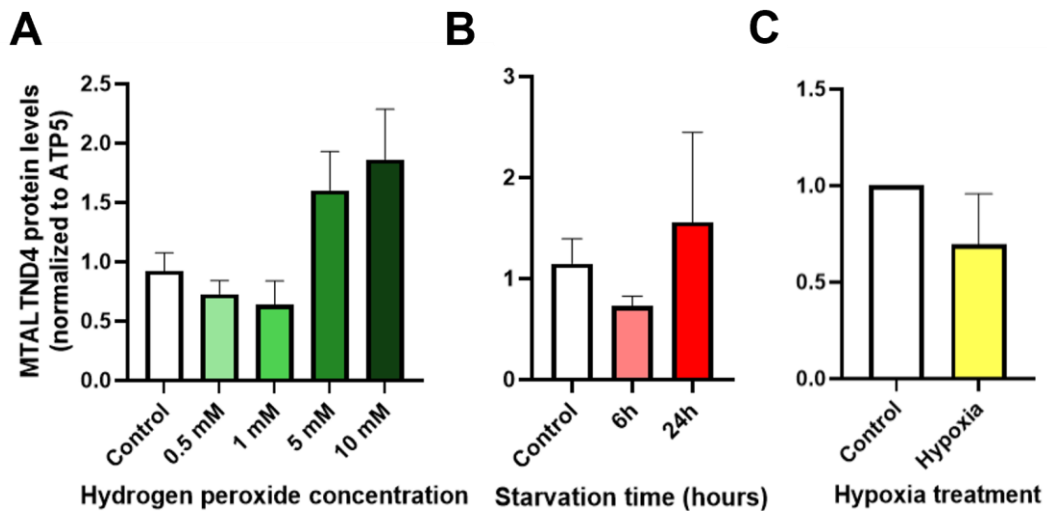
Endogenous expression of MTALTND4 was assessed by immunoblotting in various human tissues tested in other MDP studies including kidney, heart, testis, brain, liver, skeletal muscle and spleen (Cobb *et al.*, 2016; Lee *et al.*, 2015). Tissular expression of ND4 was also tested alongside of MTALTND4 to identify a converging or diverging expression pattern between the reference protein and its alternative protein. MTALTND4 is expressed at varying levels across human tissues (Figures A1 and S1). It was detected in the kidney, liver and spleen; low signals were also seen in the brain and testis, whereas no signal was found in the heart and skeletal muscle (Figure A1). Possible degradation or MTALTND4 monomer may also be present in the brain, liver and spleen (Figure S1). It seems that MTALTND4 does not follow the expression pattern of its ND4 reference protein, which is for example not found in kidney and testis, and is lowly detected in spleen, liver and muscle.



**Figure A1.** Expression of nuclear-encoded mitochondrial ATP5 (loading control), mtDNA-encoded ND4 and MTALTND4 in various human tissues (kidney, heart, testis, brain, liver, skeletal muscle, spleen) and HeLa cells by immunoblotting.

To explore the physiological roles of MTALTND4, HeLa cells underwent hydrogen peroxide treatment, starvation treatment and hypoxia treatment (Figures A2), known to affect mitochondrial functions such as ROS production, energy production, cell proliferation and metabolism (Aghababazadeh & Kerachian, 2014; Zorov *et al.*, 2006). Variation in expression was assessed by immunoblotting. While none of these treatments yielded statistically significant

results, extended starvation and exposure to high concentrations of hydrogen peroxide, a major cytotoxic ROS, seemed to increase expression of MTALTND4 in cells (Figures A2A and A2B). By contrast, it appears that hypoxia may have a negative impact on MTALTND4 expression (Figure A2C).

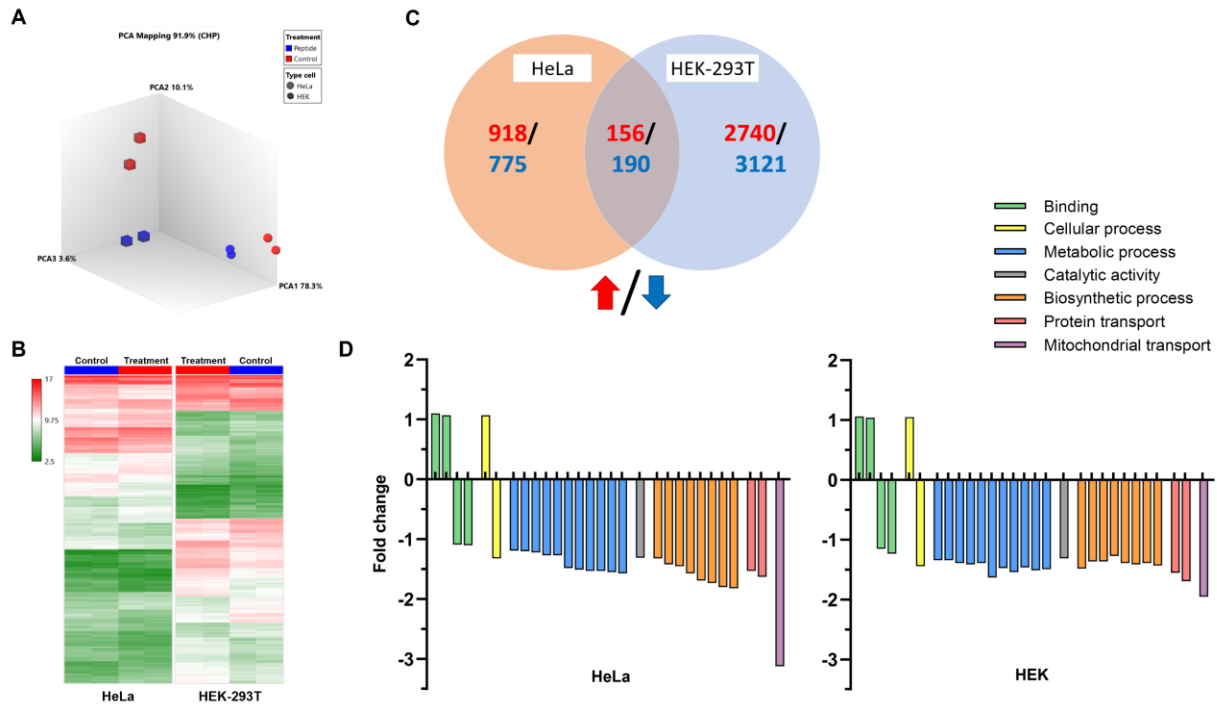


**Figure A2.** Impact of (A) dose-dependent hydrogen peroxide treatment (0, 0.5, 1, 5, 10 mM) for 30 minutes (n=4-9); (B) starvation treatment (6, 24 hrs) (n=3-4); and (C) hypoxia treatment for 4 hrs (n=3) on normalized MTALTND4 expression within HeLa cells. Data shown as mean ± SEM, Student's t-test, \*p < 0.05.

### MTALTND4 has an exogenous effect on gene expression

Following MTALTND4's detection in plasma by Kienzle *et al.* (2022) and to confirm its previously identified exogenous impact on mitochondrial physiology, microarray analyses were performed on HeLa and HEK-293T cells treated exogenously with MTALTND4 to clarify impacted genes and processes as a signaling peptide (Figures A3). Principal component analysis (PCA) revealed a global gene expression profile shift between control and treated cell types, especially for HEK-293T cells (Figure A3A). Hierarchical clustering was employed to further highlight gene expression differences in treated cells when compared to controls (Figure A3B). A notable number of genes were differentially expressed after treatment, but more genes were differentially expressed in HEK-293T cells (Figure A3C, Tables S1 and S2). A moderate overlap of 346 common genes were similarly affected in both cell types (156 upregulated and 190 downregulated genes, Table S3) and were further inspected in gene ontology and pathway analyses

to identify functions impacted by the exogenous treatment. Out of all differentially regulated genes, it was found that MTALTND4 had an overall reductive effect on metabolism, biosynthetic processes, catalytic activity and protein transport, as well as the biggest impact on mitochondrial transport, especially in HeLa cells (Figure A3D, Table S4). Pathways associated with commonly expressed genes also revealed that MTALTND4 downregulates mitosis, cellular senescence, autophagy and protein localization, and upregulates pathways related to cancer, apoptosis, metabolism, signaling and more (Table S5).



**Figure A3. Effect of MTALTND4 on gene expression.** Microarray analysis performed on HeLa and HEK-293T cells treated with MTALTND4 (10mM) for 4h (n=2).

(A) Principal component analysis (PCA) of all samples passing quality control review.

(B) Hierarchical clustering of differentially expressed genes depicting a high-level (red, fold change > 1.5) and low-level (green, fold change < 1.5) expression relative to control cells in both cell types.

(C) Venn diagram depicting upregulated (red) and downregulated (blue) genes in both cell types.

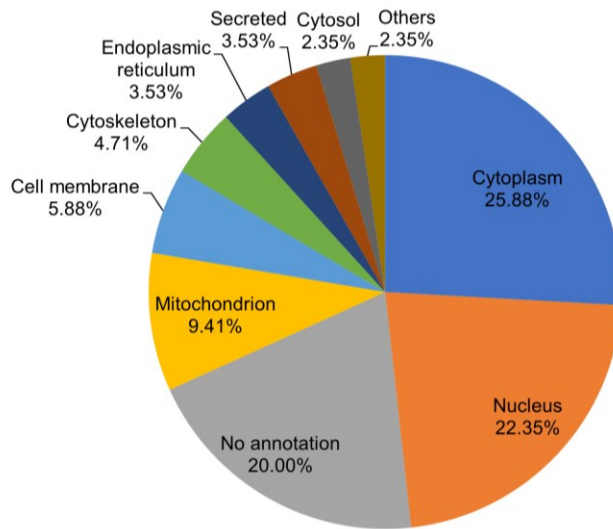
(D) Common Gene Ontology results of differentially expressed genes in both cell types.

## **MTALTND4 interacts with several proteins involved in mitochondrial functions**

MTALTND4's interacting partners were previously identified in HeLa cells (Kienzle *et al.*, 2022), though common partners in other cell types could help reveal potential functions associated with these proteins. To assess this, pull-down assays were performed with MTALTND4 on HEK-293T cells and were compared to the previous analysis by Kienzle *et al.* (2022), revealing 48 significant potential interaction partners present in all three pull-down assays (Table S6). Notably, the complement component 1q subcomponent binding protein (C1QBP) was identified as previously described and validated by western blot (Kienzle *et al.*, 2022), further solidifying it as an interacting partner to MTALTND4. The 48 interacting partners were widespread throughout the cell and found in the cytoplasm, the nucleus and the mitochondrion. A few proteins were also localized in the cell membrane, cytoskeleton, endoplasmic reticulum, cytosol and secreted proteins (The UniProt Consortium, 2021) (Figure A4).

In addition to C1QBP, MTALTND4 seemed to interact with a handful of other mitochondrial proteins such as two ATP synthase subunits (ATPF1A et ATP5F1C – *ATP synthase subunit alpha et gamma*) and an ADP/ATP translocase (SLC25A6 – *Solute Carrier Family 25 Member 6*) (Table S6). We also detected a high number of ribosomal proteins (26) and, though their presence may solely be due to an active synthesis of MTALTND4 during pull-down assays, we have still found a few non-ribosomal partners involved in protein synthesis (*e.g.*, NPM1 – *Nucleophosmin 1* and PPM1G – *Protein Phosphatase Magnesium-Dependent 1 Gamma*) and this function is therefore not to be completely ruled out. Several interacting proteins also seem to be involved in some other major functions associated with mitochondria such as cell proliferation (*e.g.*, PCNA – *Proliferating cell nuclear antigen* and NAP1L1 – *Nucleosome assembly protein 1-like 1*), apoptosis (*e.g.*, SLC25A6 and NPM1), immunity (*e.g.*, ANP32B – *Acidic Nuclear Phosphoprotein 32 Family Member B*) and respiration (PHB1 – *Prohibitin*) (Table S6). Furthermore, signaling pathways related to the interacting partners, which were identified using DAVID, indicated that MTALTND4 could be involved in regulation of gene expression, metabolism, mitochondrial biogenesis and maintenance (Table S7). Finally, we compared overrepresented pathways for both the interaction partners and the commonly expressed genes using Reactome Knowledgebase and DAVID (Gillespie *et al.*, 2022; Huang *et al.*, 2009; Sherman *et al.*, 2022). Recurring significant pathways in both experimentations were shown to affect metabolism, mRNA splicing, protein localization and stress response (Table A1).

**Figure A4.** Diagram showing the cellular components of the identified MTALTND4 interacting proteins



**Table A1.** Recurring signaling pathways found between common genes from microarrays and interacting partners from pull-down assays for DAVID Bioinformatics Resources and Reactome Knowledgebase analyses.

DAVID Bioinformatics Resources					
Related pathways	Interacting partners		Common genes		
	Gene count	P-value	# of genes	P-value	Expression
Metabolism of RNA	28	2.2E-23	12	3.2E-02	-
Metabolism	27	8.5E-10	24	1.6E-02	+

Reactome Knowledgebase					
Related pathways	Interacting partners		Common genes		
	Gene count	P-value	# of genes	P-value	Expression
Response of EIF2AK4 (GCN2) to amino acid deficiency	29	1.11E-16	5	1.23E-02	+
Cellular response to starvation	29	1.11E-16	6	1.90E-02	+
ATF6 (ATF6-alpha) activates chaperone genes	2	1.45E-03	2	1.39E-02	+
ATF6 (ATF6-alpha) activates chaperones	2	1.85E-03	2	1.76E-02	+
Protein localization	3	2.51E-02	7	6.03E-03	-
mRNA Splicing - Major Pathway	3	3.24E-02	6	3.23E-02	-
mRNA Splicing	3	3.73E-02	6	4.07E-02	-



## DISCUSSION

We have previously suggested that, as an adaptive response to certain stress, MTALTND4 could act as a signaling molecule secreted from cells to regulate physiology by inducing a state of bioenergetic depression (Kienzle *et al.*, 2022). Indeed, MTALTND4 is secreted in the plasma, it decreases mitochondrial respiration in a dose-dependent manner and increases ROS efflux, with a lack of potential compensatory mechanisms such as increased antioxidant activity and upregulation of anaerobic glycolysis (Kienzle *et al.*, 2022). Together with the little to no impact on cell viability and an increase in ATP content, this suggests that the effects of MTALTND4 do not appear to be the result of a deleterious impairment of mitochondrial functions but rather reflect a global bioenergetic depression (Kienzle *et al.*, 2022). The present study aimed to better understand the functions of MTALTND4 by assessing its tissue-specific expression, its expression in response to stress, its impact on gene expression and its potential interacting partners.

### **MTALTND4 exhibits a tissue-specific expression and reacts to different treatments known to affect mitochondrial functions**

Tissular expression analysis of MTALTND4 has shown an independent regulation from its reference gene-encoded protein ND4, suggesting independent functional roles for the two proteins, at least to some degree (Mouilleron *et al.*, 2015). This may be the result of alternative translation mechanisms such as leaky scanning and reinitiation (Andrews & Rothnagel, 2014), stress-induced changes in transcription (Himanen & Sistonen, 2019), or earlier degradation of mitochondrial proteins by mitophagy (Kubli & Gustafsson, 2012). The observed tissue-specific expression pattern of MTALTND4 is interesting in that it is low or even absent in certain tissues normally requiring substantial energy metabolism such as the brain, heart, skeletal muscles and testes (Figure A1) (Ghebrehiwet *et al.*, 2019). Knowing MTALTND4's effect on mitochondrial respiration, a high expression in these vital high energy-demanding tissues may need to be avoided to maintain proper functions even in periods of stress. This could be a way to reduce global energy consumption in situations where ATP supplies are being limited, such as during hypoxia for example (Marchetti *et al.*, 2020; Withers & Cooper, 2008).

Various stresses affecting mitochondrial functions like exposure to ROS and starvation seemed to increase MTALTND4 expression (Figures A2A and A2B). This could be a way for cells to conserve nutrients by reducing mitochondrial respiration and cell metabolism (Aghababazadeh

& Kerachian, 2014; Kienzle *et al.*, 2022). Exposure to hypoxia treatments seemed to counterintuitively decrease peptide expression (Figure A2C). Knowing the observed respiratory effects of MTALTND4 (Kienzle *et al.*, 2022), an increase in peptide expression during hypoxia would have seemed like the logical step to preserve precious oxygen resources. Perhaps MTALTND4's mechanisms of action and interaction partners differ according to stress conditions in ways that could explain this decrease. However, bigger sample sizes and a wider range of parameters for the treatments, as well as different treatments will be necessary to determine any definite physiological roles based on this approach.

### **MTALTND4 depresses metabolism and biosynthetic processes**

Mitochondrial MDPs such as MOTS-c are known to regulate nuclear gene expression, this has been demonstrated following microarrays performed on HEK-293T cells exogenously treated with the synthetic peptide, which indicated a marked effect after a 4-hour treatment on expression of genes associated with cellular metabolism and inflammation (Lee *et al.*, 2015). Similarly, we investigated the effect of MTALTND4 on gene expression by performing microarrays on HeLa and HEK-293T cells treated with this synthetic peptide for 4 hours.

The data reported here shows that gene expression was significantly altered after treatment, but more genes were differentially expressed in HEK-293T cells (Figure A3C, Tables S1 and S2). This is in line with the previous observation that HeLa and HEK-293T had dissimilar sensitivities to MTALTND4 at the phenotypic level (Kienzle *et al.*, 2022). For all differentially expressed genes, gene ontology terms indicated that metabolism, biosynthesis, protein transport, mitochondrial transport and more were significantly downregulated in both cell lines (Figure A3D, Table S4). This is also in line with their phenotypic response to MTALTND4 treatment inducing a global bioenergetic depression that would involve a coordinated downregulation of ATP-producing and ATP-demanding processes (Kienzle *et al.*, 2022). Hence, as other MDPs such as MOTS-c and SHLPs (Cobb *et al.*, 2016; Fuku *et al.*, 2015; Klaus & Ost, 2020), one of the roles of MTALTND4 could be to act as a mitokine, a mitochondrial-derived molecule released by the cell in response to mitochondrial stress and acting on other cells or tissues in an endocrine-like manner. During stress conditions the integrated stress response is triggered by the cell, causing a general inhibition of protein synthesis (Andreev *et al.*, 2015a), observed here through downregulation of biosynthetic, protein transport and catalytic activity pathways. A cellular respiration decrease by

MTALTND4 (Kienzle *et al.*, 2022) could also be explained by the downregulation of mitochondrial transport (*i.e.*, transport of substances such as molecules and ions into, out of or within a mitochondrion). Through inhibition of substance transport, respiration becomes less efficient, therefore lowering oxygen demands (Zimmerman *et al.*, 2011).

### **Interacting partners of MTALND4 also point to a possible role in regulation of gene expression**

To better identify the molecular pathways affected by MTALTND4, we used complementary approaches to identify its interacting partners. Kienzle *et al.* (2022) detected the sequence of MTALTND4 by mass spectrometry after immunoprecipitation from HeLa cell extracts. Other proteins were also found that could be potential interacting partners. One of these proteins, the complement component 1q subcomponent binding protein (C1QBP/p32/HABP1), was also identified as an interacting partner using pull-down assays on HeLa cells, an interaction that was validated by western blot (Kienzle *et al.*, 2022). Here, we have shown that this is also the case in HEK-293T cells (Table S6), further solidifying the possibility of it being an interaction partner. C1QBP is a multifunctional and multicompartamental protein primarily localized in the mitochondrial matrix, but also present in the cytosol, nucleus and cell surface of different cells (Zhang *et al.*, 2013). C1QBP has been reported to affect a variety of cellular processes. Notably, it has a major role in cellular energy metabolism by modulating the mitochondrial respiratory chain complexes' translation and their enzymatic activity, protein-protein interactions and  $\text{Ca}^{2+}$  concentrations (Chen *et al.*, 2016; Wang *et al.*, 2022). C1QBP is also involved in ROS production and oxidative stress protection, cancer metastasis and some neuronal functions such as neuroplasticity (Barna *et al.*, 2019; McGee and Baines, 2011; Zhang *et al.*, 2013). It also has an important role in the immune response through inflammation processes and its ability to interact with a variety of proteins from the plasma, cell surface and from pathogenic microorganisms (Wang *et al.*, 2022). This interaction with C1QBP could potentially explain the observed MTALTND4-mediated changes in mitochondrial respiration, ROS production (Kienzle *et al.*, 2022) and gene expression (present study). However, the cellular compartments in which this interaction may occur have yet to be determined and could offer insight to clarify potential functions.

Another protein of interest identified with our pull-down assay, the nuclease-sensitive element-binding protein 1 or Y box binding protein-1 (YBX1), has also been identified as a potential interacting partner for C1QBP (Zhang *et al.*, 2013). YBX1 is a multifunctional RNA/DNA binding protein involved in regulation of transcription, translation, DNA repair, pre-mRNA splicing and more (Prabhu *et al.*, 2015; Yue *et al.*, 2017). Following a pathway analysis using the Reactome Knowledgebase (Gillespie *et al.*, 2022), C1QBP and YBX1 seem to interact together in several pathways related to mRNA splicing, generic transcription, disease and cell signaling (Table S8). Cell signaling is involved through the Notch signaling pathway, a highly conserved pathway for communication between cells involved in cellular differentiation, proliferation, and specification (Jassal, 2006), and through the TGF $\beta$  signaling pathway, involved in gene regulation for cell differentiation, proliferation and apoptosis (Khang *et al.*, 2009). Along with MTALTND4, all three proteins could possibly interact together directly or indirectly in one or more of these pathways to accomplish their proper functions, but this remains to be demonstrated. Among the other putative interacting partners identified in our pull-down assay, a few are known as cell surface receptors (Figure A4). This includes a form of C1QBP with the ability to bind plasma proteins and, potentially, act as a receptor for circulating MTALTND4, suggesting that the peptide, like other MDPs, could act as a signaling molecule (Wang *et al.*, 2022; Zhang *et al.*, 2013; Miller *et al.*, 2020).

## CONCLUSION

Through our functional analyses of MTALTND4 by the means of various methods, we found that MTALTND4 seems to possess general characteristics shared by previously published MDPs. It is a bioactive peptide that impacts biological processes in treated cells such as metabolic process, transcription, protein activity, mitochondrial transport and more. It is found in plasma, is hinted to be a signaling molecule and its expression is tissue-specific, suggesting precise mechanisms of action (Cobb *et al.*, 2016; Hashimoto *et al.*, 2001; Lee *et al.*, 2015; Voigt & Jelinek, 2016). Complementary approaches using different cell lines point to C1QBP as a main interacting partner of MTALTND4, a protein that is involved in various intracellular processes and extracellular signaling pathways (Wang *et al.*, 2022). As an adaptive response to stress, MTALTND4 could be secreted, regulating cell physiology and reducing both ATP producing and consuming processes, in means of conserving limited energy stocks by shutting down non-essential functions (Kienzle *et al.*, 2022). However, further investigations into the precise roles of MTALTND4 are required.

### 3. Discussion

Ce projet de recherche avait pour objectif d'identifier les fonctions potentielles du nouveau peptide alternatif d'origine mitochondriale MTALTND4. Les MDPs découverts jusqu'à présent sont des peptides multifonctionnels impliqués dans une large gamme de fonctions, dont la neuroprotection, la régulation de l'insuline, la régulation du métabolisme mitochondrial, la dégénérescence maculaire et plus encore (Cobb *et al.*, 2016; Hashimoto *et al.*, 2001; Lee *et al.*, 2015; Miller *et al.*, 2022). Il serait fort probable qu'une multifonctionnalité soit également le cas pour MTALTND4.

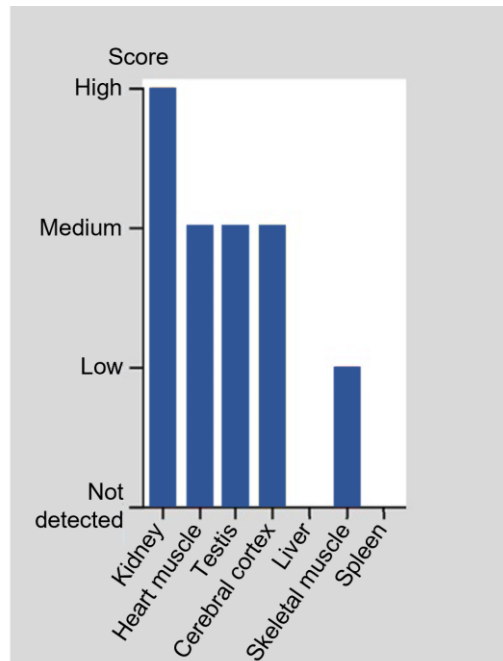
Kienzle *et al.* (2022) ont réalisé une étude préliminaire sur les fonctions potentielles de MTALTND4. Les premiers résultats ont révélé que MTALTND4, lorsqu'appliqué de manière exogène sur des cultures cellulaires, provoque une diminution du taux de respiration cellulaire ainsi qu'une augmentation de l'afflux de ROS. Cependant, ceci n'est pas accompagné de mécanismes compensatoires comme une augmentation de la glycolyse anaérobie ou de l'activité antioxydante. Une augmentation de la quantité d'ATP cellulaire a toutefois été observée, accompagnée d'une diminution importante de la prolifération cellulaire, avec une viabilité stable. Pour ces raisons, Kienzle *et al.* (2022) ont posé l'hypothèse que MTALTND4 pourrait induire une dépression bioénergétique chez la cellule en réponse à un stress. Afin de creuser cette hypothèse et de mieux cerner les fonctions potentielles de MTALTND4, des expériences ont été réalisées pour déterminer l'expression tissulaire, les effets du stress sur l'expression, l'impact de MTALTND4 sur le transcriptome et ses partenaires d'interaction.

#### 3.1 Expression tissulaire de MTALTND4

L'étude de l'expression tissulaire de MTALTND4 a révélé que le peptide semble être exprimé indépendamment de ND4 et présente un niveau d'expression tissu-spécifique (Figure A1). Autrement dit, pour un même tissu, les bandes pour ND4 et MTALTND4 présentent des niveaux d'intensité divergents, et ce, pour la majorité des tissus étudiés. Certaines protéines alternatives ont la capacité d'interagir physiquement avec leur protéine de référence. Dans ces cas, les deux ORFs peuvent alors être coexprimés et présentent un degré d'expression similaire. Les interactions fonctionnelles et réversibles entre les protéines codées par différents ORFs dans un même ARNm pourraient être un mécanisme répandu pour réguler l'activité protéique et, par le fait même,

certaines voies biologiques (Landry *et al.*, 2015; Mouilleron *et al.*, 2015). Ce n'est cependant pas ce qui est observé dans le cas de MTALTND4, signifiant que cet altORF possède au moins un certain niveau d'expression indépendant. Le peptide est donc sans doute soumis à des mécanismes de traduction alternatifs permettant sa production sans nécessairement induire celle de sa protéine de référence (Blin & Ricci, 2016). Peut-être même qu'il est soumis à des changements dans la transcription induits par le stress, pouvant modifier l'expression de gènes cytoprotecteurs et de gènes liés à la régulation du cycle cellulaire, de la transcription et du métabolisme (Himanen & Sistonen, 2019). On constate aussi que le patron d'expression de ND4 ne concorde pas tout à fait avec la littérature (Figure 10). Pour plusieurs tissus, le niveau protéique est largement inférieur à celui attendu (rein, cœur, testicule), voire quasi nul, ou présente des bandes plus élevées ou plus basses à la taille attendue de ~37 kilodaltons (kDa) (foie, muscle squelettique) (Figures A1, S1). Les bandes de faible intensité ou de taille inférieures peuvent s'expliquer par une dégradation *post mortem*. La dégradation précoce des protéines mitochondriales dans nos échantillons aurait pu être déclenchée par un niveau surélevé de stress oxydatif et de protéases (Kubli & Gustafsson, 2012). En effet, lors d'une nécrose ischémique, les protéines mitochondriales sont plus susceptibles de se dégrader en raison d'une réduction abrupte du débit sanguin aux tissus. Ceci déclenche des processus délétères entraînant des dommages importants au niveau des mitochondries, qui entraînent la mort éventuelle des cellules. Les protéases agissent comme des ciseaux qui clivent une protéine à un endroit spécifique dans sa séquence. Ce clivage peut être réalisé de manière plus généraliste (trypsine : ...K\... ou ...R\...) ou plus spécifique (*Human rhinovirus 3C protease* : ...LEVLFQ\GP...) en fonction de la protéase qui agit (López-Otín & Bond, 2008; Rodriguez *et al.*, 2008; Wagner & Weber, 2019). Ceci expliquerait notamment chez MTALND4 la présence de bandes inférieures similaires dans le cerveau et la rate, mais distinctes du foie présentant une bande encore plus basse (Figure S1). Lorsqu'on compare le niveau d'expression entre ND4 et ATP5, deux sous-unités de la chaîne de transport des électrons, on constate qu'ATP5 semble avoir subi une dégradation moindre et est observable dans tous les tissus. Ce contraste pourrait s'expliquer par les taux de dégradation hautement variables chez les protéines de la membrane interne, même au sein d'un même complexe respiratoire. Aussi, la protéine ATP5F1 $\alpha$  étant une des plus abondantes dans la membrane interne a sans doute aidé à sa bonne détection, malgré la dégradation (Hare & Hodges, 1982; King & Plun-Favreau, 2017). La présence d'une bande supérieure à la taille attendue chez ND4 (muscle squelettique) pourrait être le résultat de modifications post-

traductionnelles, des événements qui modifient la structure et les propriétés d'une protéine, par exemple, par l'addition covalente d'un groupe modificateur à un ou plusieurs acides aminés, ce qui a pour effet d'augmenter la taille de la protéine et impacter sa migration (Ramazi & Zahiri, 2021).



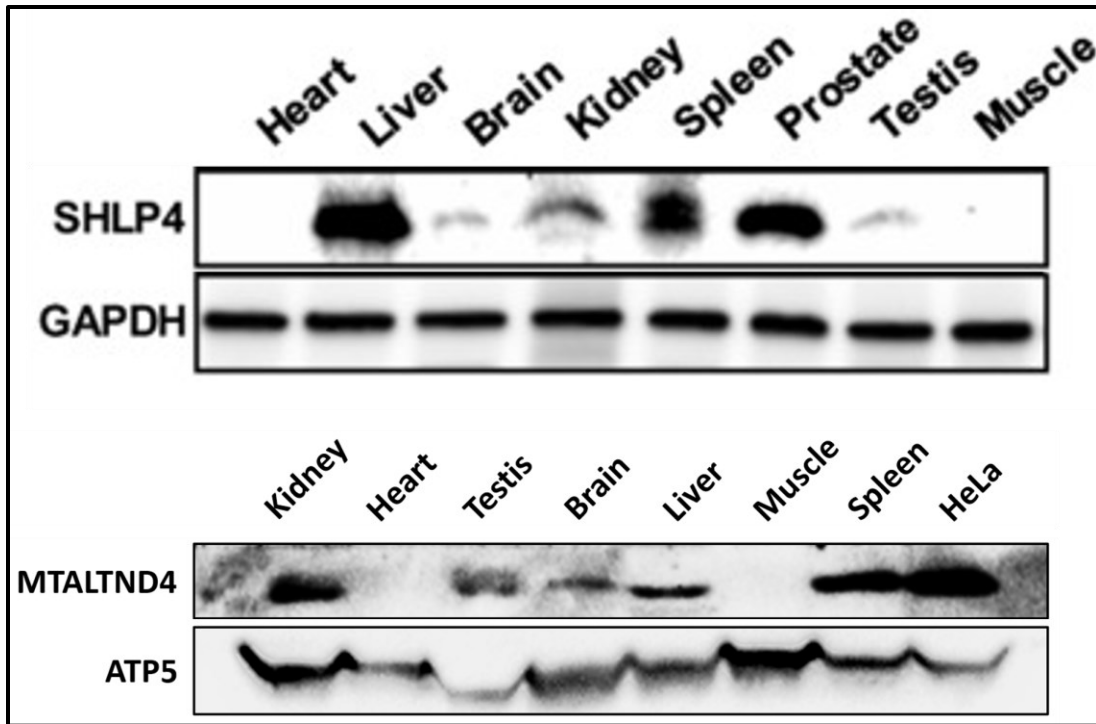
**Figure 10.** Expression globale de la protéine MT-ND4 codée par le gène mitochondrial *nd4* dans différents tissus humains. Figure tirée et modifiée de The Human Protein Atlas project.

Pour ce qui est de l'expression tissu-spécifique de MTALTND4, on constate que l'expression n'est pas constante entre les tissus, tout comme pour d'autres gènes alternatifs mitochondriaux (MOTS-c, SHLPs) (Cobb *et al.*, 2016; Lee *et al.*, 2015). Ce patron variable suggère donc des fonctions spécifiques à MTALTND4. Si le peptide induit une dépression bioénergétique chez les cellules par la diminution de la respiration (Kienzle *et al.*, 2022), il serait logique qu'il soit exprimé en quantité inférieure dans des tissus métaboliquement actifs et énergivores, comme observé dans le cas du cœur, du muscle squelettique, du cerveau et du testicule (Figure A1). En effet, ces tissus requièrent un important métabolisme énergétique et effectuent des fonctions immédiates importantes pour la survie et le fitness (Ghebrehiwet *et al.*, 2019; Matos *et al.*, 2021; Rink & Khanna, 2011). Il est aussi intéressant de noter que MTALTND4 semble absent des deux échantillons correspondant à des tissus musculaires, soit le cœur et les muscles



squelettiques, suggérant qu'il pourrait effectuer des fonctions qui n'ont pas lieu chez ce type de tissu. Il serait intéressant de tester d'autres tissus musculaires, comme les muscles lisses du petit et du gros intestin, afin de voir si le patron est similaire.

Enfin, un autre élément intéressant ressortant de cette expérience est la similarité du patron d'expression de MTALTND4 à celui du MDP SHLP4 découvert par Cobb *et al.* (2016). Dans les deux cas, le peptide ne semble pas s'exprimer dans le cœur et le muscle squelettique, il présente une expression faible dans le cerveau et le testicule ainsi qu'une expression intense dans le foie et la rate (Figure 11). Le peu que l'on connaît du MDP SHLP4 est qu'il semble notamment favoriser la prolifération des cellules NIT-1  $\beta$ , une lignée de cellules pancréatiques provenant de souris transgéniques portant un gène de rat promoteur d'insuline (Cobb *et al.*, 2016). Il serait intéressant de tester si MTALTND4 pourrait avoir un rôle similaire chez l'humain. Il serait aussi intéressant d'effectuer d'autres expériences sur des tissus frais, évitant la dégradation *post mortem*, ou sur des lignées cellulaires d'un organisme où la séquence de MTALTND4 est bien conservée, comme c'est le cas chez plusieurs primates (Kienzle *et al.*, 2022). Il serait aussi pertinent de vérifier l'expression dans les tissus d'individus d'âges différents. Pour d'autres MDPs connus, les niveaux en circulation semblent être intimement liés à l'âge, suggérant une implication dans le vieillissement et la progression de maladies liées à l'âge (Cobb *et al.*, 2016; Hashimoto *et al.*, 2001; Lee *et al.*, 2015).



**Figure 11.** Expression dans divers tissus humains pour SHLP4 et son contrôle de chargement (figure tirée et modifiée de Cobb *et al.*, 2016), puis pour MTALTND4 et son contrôle de chargement.

### 3.2 Expression de MTALTND4 en conditions stressantes

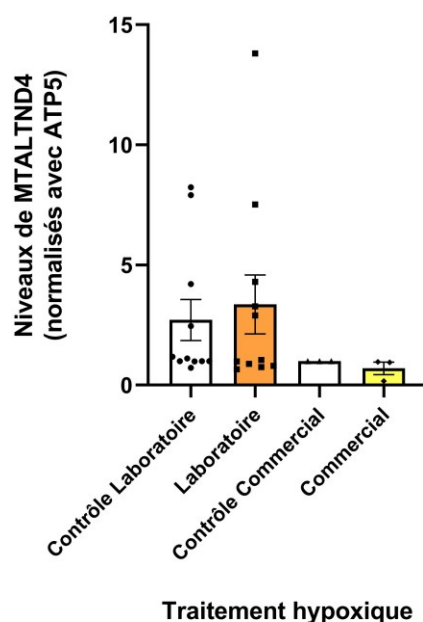
Afin de mieux cerner les rôles physiologiques de MTALTND4 dans les cellules, nous avons exposé une lignée cellulaire à différents stress affectant les fonctions mitochondriales, et étudié les changements dans l'expression de MTALTND4. Un changement dans l'expression pourrait indiquer une implication dans les fonctions associées, agissant comme mécanisme compensatoire ou promoteur. Des cellules HeLa en culture ont été traitées avec du peroxyde d'hydrogène, simulant les ROS, pour induire un stress oxydatif influençant les fonctions productrices de ROS (Zorov *et al.*, 2006). Un second traitement simule un jeûne et cible les fonctions liées à la production d'énergie et la prolifération cellulaire (Aghababazadeh & Kerachian, 2014). Un troisième traitement expose les cellules à une hypoxie, permettant de cibler les fonctions mitochondriales impliquées dans la respiration cellulaire (Turcotte *et al.*, 2002). Les résultats obtenus pour les différents traitements (Figures A2), suite à l'analyse des échantillons par western blot, n'ont montré aucun changement significatif dans l'expression de MTALTND4, mais des tendances peuvent être observées.

Pour ce qui est du traitement aux ROS (Figures A2A), on constate une augmentation de l'expression de MTALTND4 en fonction de la concentration en peroxyde d'hydrogène. D'autres MDPs étudiés (SHLP2, SHLP3, SHMOOSE) ont un effet d'atténuation sur la production de ROS et sur les niveaux de superoxydes dans la mitochondrie (Cobb *et al.*, 2016; Miller *et al.*, 2022). Cependant, il a été démontré que MTALTND4 augmente l'afflux de H<sub>2</sub>O<sub>2</sub>, sans toutefois augmenter sa capacité de piégeage des ROS (Kienzle *et al.*, 2022). Couplé avec la tendance positive entre la concentration de peroxyde d'hydrogène et la quantité de MTALTND4, ce pourrait être un mécanisme en place afin de favoriser la libération de ROS. Les ROS ne sont pas uniquement des sous-produits néfastes de la respiration mitochondriale. Ils peuvent aussi réguler divers processus cellulaires comme la prolifération et la différenciation, la biogenèse mitochondriale, la réponse à l'hypoxie, l'inflammation, l'autophagie et l'apoptose, le vieillissement et plus encore (Sena *et al.*, 2012; Sies & Jones, 2020). En réponse à des perturbations dans la cellule, les mitochondries peuvent influencer la transcription des gènes en communiquant avec le noyau comme adaptation au stress. Cette rétrocommunication se fait par la « réponse rétrograde mitochondriale » où les ROS sont une des molécules de signalisation clé agissant comme intermédiaire entre les deux organelles (Desai *et al.*, 2020). L'augmentation du peptide MTALTND4 en fonction de la concentration en H<sub>2</sub>O<sub>2</sub> et, par le fait même, l'augmentation de l'afflux de ROS pourrait être une mesure médiant la physiologie cellulaire pour répondre au stress présent. Cependant, cela reste à être démontré.

Pour ce qui est du traitement de jeûne (Figure A2B), on constate aussi une augmentation de l'expression de MTALTND4 suite à un jeûne prolongé (avec des résultats variant beaucoup d'une expérience à l'autre). On pourrait penser que, lorsque les cellules se retrouvent limitées en ressources, le peptide agirait afin de préserver celles-ci en diminuant la respiration mitochondriale, en limitant l'utilisation d'ATP et en ralentissant la prolifération, sans toutefois affecter la viabilité des cellules (Aghababazadeh & Kerachian, 2014; Kienzle *et al.*, 2022). Cependant, ces mesures n'ont pas été réalisées sur les cellules en traitement de jeûne, il sera donc intéressant de tester cette hypothèse.

Pour ce qui est du traitement hypoxique (Figure A2C), on constate une diminution de l'expression de MTALTND4 face à un manque d'oxygène. Kienzle *et al.* (2022) ont révélé qu'à

de faibles concentrations, la présence de MTALTND4 exogène a un effet négatif et immédiat sur la consommation basale d'oxygène chez les deux lignées cellulaires étudiées. Ces résultats laissent penser que, en conditions hypoxiques, l'expression du peptide serait favorisée pour protéger la cellule en diminuant son taux de respiration. Cependant, le contraire a plutôt été observé. Il est possible que MTALTND4 ne soit simplement pas impliqué lors de l'hypoxie et ses effets seraient plus avantageux lors de stress chroniques requérant une réponse de survie moins immédiate. Il est aussi possible que le nombre de lysats commerciaux utilisés ne soit pas suffisant (n=3). Des traitements hypoxiques ont également été réalisés en laboratoire et suggèrent plutôt une augmentation de l'expression de MTALTND4 après le traitement (Figure 12). Cependant, ces traitements n'ont pas été retenus, car les conditions expérimentales n'étaient pas uniformes et les résultats présentaient des fluctuations importantes. Brièvement, pour chaque échantillon, un Pétri de cellules HeLa a été divisé entre deux chambres respiratoires d'un respiromètre (Oroboros Instruments, Innsbruck, Austria). Une des chambres avait un accès à l'air, agissant comme contrôle, tandis que l'autre chambre était scellée et la consommation d'oxygène a été surveillée. En fonction de la quantité de cellules dans la chambre et de leur taux de respiration, les échantillons étaient exposés à une durée variable en hypoxie jusqu'à l'épuisement complet de l'oxygène pendant cinq minutes avant la réoxygénation de la chambre. Il serait intéressant de refaire les expériences hypoxiques en laboratoire telles que celles effectuées par Pastukh *et al.* (2016) en utilisant deux incubateurs ajustés avec des niveaux différents d'oxygène afin d'obtenir des traitements plus stables. Aussi, vu que MTALTND4 a préalablement été retrouvé dans le plasma (Kienzle *et al.*, 2022), une quantité importante du peptide aurait pu être exportée hors des cellules pour se retrouver dans le milieu de culture où elle n'aurait pas pu être détectée. Dans tous les cas, il serait intéressant de confirmer les résultats avec davantage d'échantillons. Puis, il serait bien de tester différentes durées et concentrations pour les traitements réalisés dans le but d'obtenir des données significatives et approfondir le rôle de MTALTND4 face au stress. Des traitements complémentaires pourraient également être réalisés pour cibler d'autres fonctions mitochondriales. Par exemple, un milieu incitant le déclenchement de l'apoptose grâce à TRAIL (*tumour necrosis factor-related apoptosis-inducing ligand*), un ligand qui engendre le relâchement du cytochrome *c* par la mitochondrie, devrait avoir un effet sur le rôle majeur de la mitochondrie dans la régulation des processus apoptotiques (Azakir *et al.*, 2010). Ce genre de traitement pourrait nous donner d'autres indices fonctionnels sur MTALTND4.



**Figure 12.** Expression normalisée de MTALTND4 dans des cellules HeLa suite à un traitement hypoxique issu de lysats commerciaux (n=3) ou induit en laboratoire (n=11). Présenté par la moyenne  $\pm$  SEM, Student's t-test, \*p < 0.05.

### 3.3 Influence de MTALTND4 sur le transcriptome

Notre étude précédente sur MTALTND4 a révélé un impact exogène sur la respiration mitochondriale et l'afflux de ROS (Kienzle *et al.*, 2022). Puisque le peptide a aussi été détecté dans le plasma, une fonction de signalisation régulant la physiologie cellulaire a été suggérée. Afin de mieux comprendre les fonctions et processus biologiques affectés, nous avons étudié l'effet exogène de MTALTND4 sur l'expression génique à l'aide de biopuces. Pour chacune des lignées cellulaires, l'entièreté des gènes différentiellement régulés et détectés par transcriptomique (excluant les ARN non codants) a pu être analysée par Gene Ontology, permettant l'interprétation fonctionnelle des données expérimentales. Nos résultats montrent une sous-régulation importante de plusieurs processus biologiques et fonctions cellulaires (Figure A3D, Tableau S4). Les processus métaboliques et les processus biosynthétiques représentaient les sous-groupes communs les plus touchés. L'activité catalytique et le transport de certaines protéines étaient aussi régulés négativement, tandis que le transport mitochondrial était la fonction la plus touchée de toutes, particulièrement dans les cellules HeLa. La liaison protéique ainsi que des processus cellulaires, quant à eux, étaient à la fois sur et sous-régulés.

Afin de bien comprendre chaque processus affecté, ils seront définis ici à l'aide de l'application AmiGO (Carbon *et al.*, 2009), permettant de parcourir la base de données en Gene Ontology. Les processus métaboliques sont les réactions par lesquelles les organismes transforment les substances chimiques, telles que l'anabolisme et le catabolisme. Ils comprennent les transformations de petites molécules ainsi que des processus macromoléculaires comme la réparation et la réplication de l'ADN, la synthèse et la dégradation des protéines. Les processus biosynthétiques sont un sous-groupe des processus métaboliques axés sur les fonctions anaboliques où de l'énergie est utilisée pour transformer les molécules simples en molécules complexes. L'activité catalytique représente l'augmentation du taux des réactions chimiques à l'aide d'enzymes qui sont impliquées dans la majorité des réactions biologiques (Cooper, 2000). Le transport protéique concerne plus précisément le transport de composés azotés dans la cellule, essentiel pour la croissance et la prolifération (Kurmi & Haigis, 2020). Le transport mitochondrial comprend le déplacement de molécules ou d'ions à l'intérieur, à l'extérieur ou au sein de la mitochondrie. Enfin, la liaison protéique concerne simplement l'interaction avec une protéine et les processus cellulaires, notamment l'organisation d'organelles assure l'assemblage ou le désassemblage d'une organelle au sein d'une cellule.

Nos résultats semblent donc en accord avec l'hypothèse de la dépression bioénergétique, observée précédemment quant à l'effet exogène de MTALTND4 sur la physiologie cellulaire (Kienzle *et al.*, 2022), par une sous-régulation coordonnée des processus producteurs et consommateurs d'ATP. Au niveau du transcriptome, il y a notamment une diminution globale des processus métaboliques, soit l'ensemble des réactions cataboliques productrices d'énergie et anaboliques demandant de l'énergie. De plus, la synthèse protéique est l'un des processus biosynthétiques qui demande le plus d'énergie et, face à un stress, sera strictement régulée en diminuant sa capacité de production afin de conserver ses ressources (Andreev *et al.*, 2015a). Puis, la catalyse de l'ATP est une des réactions enzymatiques les plus fréquentes chez les organismes permettant de transférer l'énergie chimique nécessaire pour accomplir les diverses tâches vitales chez la cellule (Kiani & Fischer, 2016). Il serait alors normal d'avoir une baisse au niveau des activités cataboliques lorsque les demandes en ATP sont aussi à la baisse. La sous-régulation du transport des composés azotés pourrait aussi expliquer la diminution de la prolifération cellulaire

observée, tandis que la sous-régulation importante du transport mitochondrial pourrait expliquer la chute de la respiration cellulaire (Kienzle *et al.*, 2022). En effet, la phosphorylation oxydative requiert le transport de nombreuses substances, par exemple le passage du pyruvate issu de la glycolyse dans la matrice mitochondriale. Si ce transport est perturbé et que la respiration est moins efficace, les demandes en oxygène seront moindres (Zimmerman *et al.*, 2011). MTALTND4 affecte donc des fonctions similaires chez HeLa et HEK-293T de manière indirecte en modulant la transcription de différents groupes et quantités de gènes (Figure A3C, Tableaux S1 et S2). Ceci se reflète dans les réponses variables entre les deux lignées cellulaires lors des expériences physiologiques effectuées à ce jour (Kienzle *et al.*, 2022).

### 3.4 Partenaires d'interaction de MTALTND4

Afin de préciser les fonctions et partenaires d'interaction de MTALTND4, la chromatographie d'affinité avec une fusion GST-MTALTND4 (ou la technique *pull-down*) a été utilisée sur des cellules HEK-293T afin de compléter les résultats obtenus précédemment sur des cellules HeLa par Kienzle *et al.* (2022). Les échantillons ont ensuite été analysés par spectrométrie de masse (UPLC-MS/MS – *Ultra performance liquid chromatography - tandem mass spectrometry*) par la plateforme de protéomique du centre universitaire de santé McGill. Les résultats sont compilés dans le tableau S6.

Un des partenaires d'interaction les plus prometteurs, à la fois révélé par Kienzle *et al.* (2022) dans les cellules HeLa (par co-immunoprécipitation et *pull-down*) et dans la présente étude dans les cellules HEK-293T (par *pull-down*), est la protéine C1QBP (*complement component 1q subcomponent binding protein/p32/HABP1*). C1QBP est une protéine multifonctionnelle présente dans plusieurs compartiments cellulaires, dont la matrice mitochondriale et la surface cellulaire (Zhang *et al.*, 2013). En particulier, elle est impliquée dans la modulation des complexes respiratoires et leur activité enzymatique, la production de ROS, la protection contre le stress oxydatif et la transcription mitochondriale (Rahman *et al.*, 2020; Wang *et al.*, 2022). C1QBP peut interagir avec une variété de protéines plasmatiques et servir de récepteur sur la membrane plasmique (Barna *et al.*, 2019; Wang *et al.*, 2022). L'activation du récepteur C1QBP par MTALTND4 pourrait à la fois expliquer les résultats physiologiques obtenus par Kienzle *et al.* (2022), ainsi que les résultats transcriptomiques discutés plus haut suite à une exposition exogène

de cellules en culture au peptide MTALTND4. Ceci est en outre appuyé par la présence d'un autre partenaire d'intérêt, soit YBX1 (*nuclease-sensitive element-binding protein 1 or Y box binding protein-1*) associée à la régulation de la transcription (Yue *et al.*, 2017). En effet, cette protéine interagirait non seulement avec MTALTND4, mais c'est aussi un partenaire de C1QBP et tous les trois pourraient potentiellement former un complexe protéique (Zhang *et al.*, 2013). Ce mécanisme d'action pourrait se rapprocher de celui utilisé par l'Humanine qui peut se lier au récepteur trimérique de la membrane plasmique WSX1-GP130-CNTFR ou au récepteur FPRL1 pour déclencher des cascades de signalisation impliquant plusieurs facteurs de transcription et kinases affectant physiologie cellulaire (Miller *et al.*, 2020). MTALTND4 pourrait agir de manière similaire où la fraction extracellulaire activerait un récepteur sur la membrane plasmique, ce qui déclencherait des voies de signalisations pour engendrer des réponses cellulaires telles que des modifications dans la régulation de la transcription. Une analyse avec la base de données Réactome (Reactome Knowledgebase, Gillespie *et al.*, 2022) a révélé que C1QBP et YBX1 interagissent dans certaines des voies de signalisation, dont la transduction de signal, le métabolisme de l'ARN, l'expression génique et certaines maladies (Tableau S8). Les voies de transduction de signal permettent aux signaux extracellulaires de provoquer des changements dans l'état et l'activité des cellules réceptrices. Le complexe C1QBP-YBX1 est impliqué dans la signalisation cellulaire par la voie de signalisation Notch, une voie hautement conservée pour la communication entre cellules qui régule la différenciation, la prolifération et la spécification cellulaire. Elle assure, par exemple, le renouvellement de tissus adultes comme le sang, la peau et l'épithélium intestinal et maintient les cellules souches dans un état de prolifération (Jassal, 2006). La signalisation et l'expression génique sont aussi médiées par la voie de signalisation TGF $\beta$  (*Transforming growth factor beta*). Celle-ci régule l'expression de la prolifération, du développement de tissus et de l'immunité (Kang *et al.*, 2009; Morikawa *et al.*, 2016).

L'implication de C1QBP et YBX1 dans des fonctions de signalisation, de régulation génique et de prolifération, en plus du fait que MTALTND4 semble interagir avec ces protéines, appuie l'hypothèse comme quoi MTALTND4 serait une molécule signalétique qui induirait une dépression bioénergétique (Kienzle *et al.*, 2022). Il serait intéressant de réaliser des expériences en immunofluorescence ciblant les protéines MTALTND4 et C1QBP afin d'identifier dans quels compartiments cellulaires leur interaction a lieu pour préciser les fonctions et voies de signalisation



potentielles de MTALTND4. Cela dit, quelques autres partenaires potentiels retrouvés à la fois dans les cellules HeLa et HEK-293T, en grand nombre (*spectrum counts*), indiquent que MTALTND4 pourrait, tout comme C1QBP, se retrouver dans différents compartiments cellulaires, incluant les mitochondries (*e.g.*, ATP5F1A/ATP5F1C – *ATP synthase subunit alpha* et *gamma* – et SLC25A6 – *Solute Carrier Family 25 Member 6* : des protéines mitochondriales identifiées comme partenaires potentiels), le cytoplasme, le noyau (*e.g.*, NPM1 – *Nucleophosmin 1* – et PCNA – *Proliferating cell nuclear antigen* : des protéines nucléaires) et le plasma (*e.g.*, C1QBP et YBX1 se retrouvent dans le plasma, comme mentionné plus haut) (Figure A4, Tableau S6). Il est donc tout à fait probable que MTALTND4 soit, tout comme C1QBP, une protéine multifonctionnelle.

Parmi ses multiples rôles, MTALTND4 pourrait participer aux fonctions de la mitochondrie par l'interaction avec des éléments mitochondriaux. Parmi ces interacteurs, on retrouve notamment des sous-unités de l'ATP synthase (ATP5F1A et ATP5F1C) qui régulent son activité et convertissent l'énergie (Mnatsakanyan *et al.*, 2009; Walker *et al.*, 1982), ainsi qu'une translocase ATP/ADP (SLC25A6), qui assure l'importation d'ADP dans la matrice mitochondriale ainsi que l'exportation d'ATP pour alimenter la cellule et des produits induisant l'apoptose (Stelzer *et al.*, 2016) (Tableau S6). De plus, plusieurs partenaires participent aux principales fonctions mitochondriales telles que la prolifération cellulaire (*e.g.*, PCNA et NAP1L1 – *Nucleosome assembly protein 1-like 1*), l'apoptose (*e.g.*, SLC25A6 et NPM1), l'immunité (*e.g.*, ANP32B – *Acidic Nuclear Phosphoprotein 32 Family Member B*), ainsi que la respiration et le vieillissement (PHB1 – *Prohibitin*) (Tableau S6). MTALTND4 pourrait aussi réguler la synthèse protéique par la présence de partenaires impliqués dans cette fonction (*e.g.*, 26 protéines ribosomales, NPM1 et PPM1G – *Protein Phosphatase Magnesium-Dependent 1 Gamma*) (Tableau S6).

### **3.5 Voies de signalisation affectées par MTALTND4**

L'ensemble des partenaires d'interaction de MTALTND4 (Tableau S6), ainsi que les gènes régulés similairement entre HeLa et HEK-293T suivant une exposition au peptide MTALTND4 (Tableau S3), ont été analysés via la base de données Réactome (Gillespie *et al.*, 2022) et les ressources bio-informatiques DAVID (*DAVID Bioinformatics Resources*, Huang *et al.*, 2009; Sherman *et al.*, 2022) pour révéler les voies de signalisation associées. La base de données DAVID

offre le regroupement fonctionnel de voies similaires possédant des éléments communs, permettant une vision plus globale des fonctions associées. Généralement, les partenaires d'interaction issus des essais pull-down semblent être impliqués au niveau de la biogenèse mitochondriale, l'importation de protéines mitochondriales, l'expression génique, le métabolisme et dans certaines maladies par la signalisation dans les cancers, le Parkinson et la maladie de Huntington (Tableau S7). Les gènes communs issus de l'analyse transcriptomique semblent réguler négativement les voies liées à la mitose, la sénescence cellulaire, l'autophagie et le métabolisme de l'ARN, et ils indiquent une régulation positive des voies reliées au métabolisme des lipides, à l'apoptose, au système immunitaire et aux cancers (Tableau S5). En comparant les résultats de ces analyses pour les deux types d'expérimentations, il a été possible de révéler les voies de signalisation communes (Tableau A1). Le métabolisme de l'ARN et la localisation des protéines seraient sous-régulés, tandis que la réponse aux protéines mal repliées (UPR – *unfolded protein response*) et la réponse à la famine, deux voies de réponse au stress utilisées afin de maintenir l'homéostasie cellulaire (Kültz, 2005), ainsi que le métabolisme seraient sur-régulés. Tous les éléments énumérés ici permettent de faire des liens avec les résultats observés précédemment sur les effets de MTALTND4 sur la respiration mitochondriale, la prolifération cellulaire, la viabilité cellulaire, et la possibilité que MTALTND4 agisse comme « molécule signal » dans le plasma.

## 4. Conclusion et perspectives

L'ensemble de nos résultats concernant la caractérisation fonctionnelle du nouveau peptide mitochondrial MTALTND4 semble confirmer les hypothèses proposées par Kienzle *et al.* (2022) comme quoi MTALTND4 serait une molécule de signalisation sécrétée par les cellules qui viendrait affecter la physiologie pour induire un état de dépression bioénergétique en réponse au stress. Ceci est appuyé par une présence plus faible de peptide dans les tissus énergivores, son implication dans la production de ROS, son effet inhibiteur sur de nombreux gènes et processus biologiques, son interaction avec C1QBP ainsi que son implication dans certaines voies de signalisation. Plusieurs autres indices ont également été révélés par les expérimentations et suggèrent que MTALTND4 pourrait, tout comme C1QBP, être une protéine multifonctionnelle (ou être un élément d'un complexe multifonctionnel). MTALTND4 pourrait être une cible intéressante en tant que molécule thérapeutique comme l'ont été plusieurs autres MDPs grâce au rôle central des mitochondries dans les fonctions cellulaires, les maladies et le vieillissement (Hashimoto *et al.*, 2001; Lee *et al.*, 2015; Picard *et al.*, 2016).

Bien que ce projet de recherche ait offert plusieurs pistes exploratoires, d'autres études seront nécessaires pour épurer les fonctions potentielles obtenues et pouvoir cibler des rôles précis associés à MTALTND4. Au sein de l'équipe de laboratoire de Sophie Breton et Annie Angers, d'autres expériences seront réalisées prochainement afin de pousser les résultats d'analyses transcriptomique. Cette fois-ci, les lignées cellulaires HeLa et HEK seront traitées de manière exogène avec MTALTND4 pendant 4 h et 72 h, comme l'a fait Lee *et al.* (2015), et le transcriptome sera séquencé via RNA-Seq (*RNA sequencing*). Ceci permettra d'obtenir des résultats plus précis par le séquençage complet de l'ensemble du transcriptome, contrairement à l'analyse de transcrits prédéfinis avec les biopuces. D'autres techniques de surexpression génique de manière endogène sont en développement et pourront potentiellement être utilisées pour étudier plus concrètement l'effet de MTALTND4 sur la physiologie cellulaire.

## Références

- Abcam. (2021). *HeLa Oxidative Damage Lysate Set: Hydrogen Peroxide-Treated and Vehicle-Treated Control ab181414*. <https://www.abcam.com/hela-oxidative-damage-lysate-set-hydrogen-peroxide-treated-and-vehicle-treated-control-ab181414.html>
- Aghababazadeh, M., & Kerachian, M. A. (2014). Cell Fasting: Cellular Response and Application of Serum Starvation. *Journal of Fasting and Health*, 2(4), 147-150. doi:10.22038/JFH.2014.3928
- Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., & Walter, P. (2008). *Molecular biology of the cell* (5e éd.). Garland Science.
- American Type Culture Collection (ATCC). (2022). *ATCC Animal Cell Culture Guide*. <https://www.atcc.org/resources/culture-guides/animal-cell-culture-guide>
- Amin, M. R., Yurovsky, A., Chen, Y., Skiena, S., & Futcher, B. (2018). Re-annotation of 12,495 prokaryotic 16S rRNA 3' ends and analysis of Shine-Dalgarno and anti-Shine-Dalgarno sequences. *PLOS ONE*, 13(8), e0202767. <https://doi.org/10.1371/journal.pone.0202767>
- Anderson, S., Bankier, A. T., Barrell, B. G., de Bruijn, M. H. L., Coulson, A. R., Drouin, J., Eperon, I. C., Nierlich, D. P., Roe, B. A., Sanger, F., Schreier, P. H., Smith, A. J. H., Staden, R., & Young, I. G. (1981). Sequence and organization of the human mitochondrial genome. *Nature*, 290(5806), 457–465. doi:10.1038/290457a0
- Andreev, D. E., O'Connor, P. B., Fahey, C., Kenny, E. M., Terenin, I. M., Dmitriev, S. E., Cormican P., Morris, D. W., Shatsky, I. N., & Baranov, P. V. (2015). Translation of 5' leaders is pervasive in genes resistant to eIF2 repression, *eLife*, 4, e03971. doi:10.7554/eLife.03971.001

- Andreev, D. E., O'Connor, P. B., Zhdanov, A. V., Dmitriev, R. I., Shatsky, I. N., Papkovsky, D. B., & Baranov, P. V. (2015). Oxygen and glucose deprivation induces widespread alterations in mRNA translation within 20 minutes. *Genome Biology*, *16*(1). doi:10.1186/s13059-015-0651-z
- Andrews, S. J., & Rothnagel, J. A. (2014). Emerging evidence for functional peptides encoded by short open reading frames. *Nature Reviews Genetics*, *15*(3), 193–204. doi:10.1038/nrg3520
- Angers, A., Ouimet, P., Tsyvian-Dzyabko, A., Nock, T., & Breton, S. (2019). L'ADN mitochondrial, un potentiel codant mésestimé. *Médecine/sciences*, *35*(1), 46–54. doi:10.1051/medsci/2018308
- Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., Davis, A. P., Dolinski, K., Dwight, S. S., Eppig, J. T., Harris, M. A., Hill, D. P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J. C., Richardson, J. E., Ringwald, M., Rubin, G. M., & Sherlock, G. (2000). Gene Ontology: tool for the unification of biology. *Nature Genetics*, *25*(1), 25–29. doi:10.1038/75556
- Ayyub, S. A., & Varshney, U. (2019). Translation initiation in mammalian mitochondria- a prokaryotic perspective. *RNA Biology*, *17*(2), 165-175. doi:10.1080/15476286.2019.1690099
- Azakir, B. A., Desrochers, G., & Angers, A. (2010). The ubiquitin ligase Itch mediates the antiapoptotic activity of epidermal growth factor by promoting the ubiquitylation and degradation of the truncated C-terminal portion of Bid. *FEBS Journal*, *277*(5), 1319–1330. doi:10.1111/j.1742-4658.2010.07562.x
- Barna, J., Dimén, D., Puska, G., Kovács, D., Csikós, V., Oláh, S., Udvari, E. B., Pál, G., & Dobolyi, Á. (2019). Complement component 1q subcomponent binding protein in the brain of the rat. *Scientific Reports*, *9*(1). doi:10.1038/s41598-019-40788-z

- Blin, J., & Ricci, E. P. (2016). Le profilage ribosomique. *Médecine/sciences*, *32(10)*, 849–860. doi:10.1051/medsci/20163210018
- Brunet, M. A., Leblanc, S., & Roucou, X. (2020). Reconsidering proteomic diversity with functional investigation of small ORFs and alternative ORFs. *Experimental Cell Research*, *393(1)*, 112057. doi:10.1016/j.yexcr.2020.112057
- Capt, C., Passamonti, M., & Breton, S. (2016). The human mitochondrial genome may code for more than 13 proteins. *Mitochondrial DNA Part a*, *27(5)*, 3098-3101. doi:10.3109/19401736.2014.1003924
- Carbon, S., Ireland, A., Mungall, C. J., Shu, S., Marshall, B., Lewis, S., AmiGO Hub, & Web Presence Working Group (2009). AmiGO: online access to ontology and annotation data. *Bioinformatics*, *25(2)*, 288–289. doi:10.1093/bioinformatics/btn615
- Chen, R., Xiao, M., Gao, H., Chen, Y., Li, Y., Liu, Y., & Zhang, N. (2016). Identification of a novel mitochondrial interacting protein of C1QBP using subcellular fractionation coupled with CoIP-MS. *Analytical and Bioanalytical Chemistry*, *408(6)*, 1557–1564. doi:10.1007/s00216-015-9228-7
- Christian, B. E., & Spremulli, L. L. (2010). Preferential Selection of the 5'-Terminal Start Codon on Leaderless mRNAs by Mammalian Mitochondrial Ribosomes. *Journal of Biological Chemistry*, *285(36)*, 28379–28386. doi:10.1074/jbc.m110.149054
- Christian, B. E., & Spremulli, L. L. (2012). Mechanism of protein biosynthesis in mammalian mitochondria. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*, *1819(9-10)*, 1035–1054. doi:10.1016/j.bbagr.2011.11.009

- Cobb, L. J., Lee, C., Xiao, J., Yen, K., Wong, R. G., Nakamura, H. K., Mehta, H. H., Gao, Q., Ashur, C., Huffman, D. M., Wan, J., Muzumdar, R., Barzilai, N., & Cohen, P. (2016). Naturally occurring mitochondrial-derived peptides are age-dependent regulators of apoptosis, insulin sensitivity, and inflammatory markers. *Aging*, *8*(4), 796–808. doi:10.18632/aging.100943
- Cold Spring Harbor Laboratory. (2004). Detection of protein–protein interactions using the GST fusion protein pull-down technique. *Nat Methods*, *1*, 275-276. doi:10.1038/nmeth1204-275
- Cooper, G. M. (2000). The central role of enzymes as biological catalysts. *The Cell: A Molecular Approach* (2e éd.). Sinauer Associates.
- Cruz-Zaragoza, L. D., Dennerlein, S., Linden, A., Yousefi, R., Lavdovskaia, E., Aich, A., Falk, R. R., Gomkale, R., Schöndorf, Bohnsack, M. T., Richter-Dennerlein, R., Urlaub, H., & Rehling, P. (2021). An in vitro system to silence mitochondrial gene expression. *Cell*, *184*(23), 5824-5837. doi:10.1016/j.cell.2021.09.033
- Desai, R., East, D. A., Hardy, L., Faccenda, D., Rigon, M., Crosby, J., Alvarez, M. S., Singh, A., Mainenti, M., Hussey, L. K., Bentham, R., Szabadkai, G., Zappulli, V., Dhoot, G. K., Romano, L. E., Xia, D., Coppens, I., Hamacher-Brady, A., Chapple, J. P., ... & Campanella, M. (2020). Mitochondria form contact sites with the nucleus to couple prosurvival retrograde response. *Science Advances*, *6*(51). doi:10.1126/sciadv.abc9955
- diCenzo, G. C., & Finan, T. M. (2017). The Divided Bacterial Genome: Structure, Function, and Evolution. *Microbiology and Molecular Biology Reviews*, *81*(3). doi:10.1128/mnbr.00019-17

- Doerrier, C., Garcia-Souza, L. F., Krumschnabel, G., Wohlfarter, Y., Mészáros, A. T., & Gnaiger, E. (2018). High-Resolution FluoRespirometry and OXPHOS Protocols for Human Cells, Permeabilized Fibers from Small Biopsies of Muscle, and Isolated Mitochondria. *Methods in Molecular Biology*, 31–70. doi:10.1007/978-1-4939-7831-1\_3
- Faure, E., Delaye, L., Tribolo, S., Levasseur, A., Seligmann, H., & Barthélémy, R.-M. (2011). Probable presence of an ubiquitous cryptic mitochondrial gene on the antisense strand of the cytochrome oxidase I gene. *Biology Direct*, 6(1), 56. doi:10.1186/1745-6150-6-56
- Fuku, N., Pareja-Galeano, H., Zempo, H., Alis, R., Arai, Y., Lucia, A., & Hirose, N. (2015). The mitochondrial-derived peptide MOTS-c: a player in exceptional longevity? *Aging Cell*, 14(6), 921–923. doi:10.1111/accel.12389
- Ghebrehiwet, B., Geisbrecht, B. V., Xu, X., Savitt, A. G., & Peerschke, E. I. B. (2019). The C1q Receptors: Focus on gC1qR/p33 (C1qBP, p32, HABP-1)1. *Seminars in Immunology*, 45, 101338. doi:10.1016/j.smim.2019.101338
- Gillespie, M., Jassal, B., Stephan, R., Milacic, M., Rothfels, K., Senff-Ribeiro, A., Griss, J., Sevilla, C., Matthews, L., Gong, C., Deng, C., Varusai, T., Ragueneau, E., Haider, Y., May, B., Shamovsky, V., Weiser, J., Brunson, T., Sanati, N., ... & D'Eustachio, P. (2022). The reactome pathway knowledgebase 2022, *Nucleic Acids Research*, 50(D1), D687-D692. doi:10.1093/nar/gkab1028
- Hare, J. F., & Hodges, R. (1982). Turnover of mitochondrial inner membrane proteins in hepatoma monolayer cultures. *Journal of Biological Chemistry*, 257(7), 3575-3580. doi:10.1016/S0021-9258(18)34818-X



- Hashimoto, Y., Niikura, T., Tajima, H., Yasukawa, T., Sudo, H., Ito, Y., Kita, Y., Kawasumi, M., Kouyama, K., Doyu, M., Sobue, G., Koide, T., Tsuji, S., Lang, J., Kurokawa, K., & Nishimoto, I. (2001). A rescue factor abolishing neuronal cell death by a wide spectrum of familial Alzheimer's disease genes and A $\beta$ . *Proceedings of the National Academy of Sciences*, *98(11)*, 6336–6341. doi:10.1073/pnas.101133498
- Himanen, S. V., & Sistonen, L. (2019). New insights into transcriptional reprogramming during cellular stress. *Journal of Cell Science*, *132(21)*, jcs238402. doi:10.1242/jcs.238402
- Huang, D. W., Sherman, B. T., & Lempicki, R. A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature Protocols*, *4(1)*, 44–57. doi:10.1038/nprot.2008.211
- Jassal, B. (2006). Signaling by NOTCH. *Reactome - a curated knowledgebase of biological pathways*, *16*, doi:10.3180/REACT\_299.3 10.3180/REACT\_2105.2
- Johnson, M. (2012). Fetal bovine serum. *Mater methods*, *2(117)*. doi:10.13070/mm.en.2.117
- Khawaja, A., Itoh, Y., Remes, C., Spähr, H., Yukhnovets, O., Höfig, H., Amunts, A., & Rorbach, J. (2020). Distinct pre-initiation steps in human mitochondrial translation. *Nature Communications*, *11(1)*. doi:10.1038/s41467-020-16503-2
- Kiani, F. A., & Fischer, S. (2016). Comparing the catalytic strategy of ATP hydrolysis in biomolecular motors. *Physical Chemistry Chemical Physics*, *18(30)*, 20219–20233. doi:10.1039/c6cp01364c
- Kienzle, L., Bettinazzi, S., Brunet, M., Choquette, T., Khorami, H. H., Roucou, X., Landry, C. R., Angers, A., & Breton, S. (2022). MTALTND4, a second protein coded by *nd4* impacts mitochondrial bioenergetics. *bioRxiv*. doi:10.1101/2022.04.28.489924

- Kim, K. H., Son, J. M., Benayoun, B. A., & Lee, C. (2018). The Mitochondrial-Encoded Peptide MOTS-c Translocates to the Nucleus to Regulate Nuclear Gene Expression in Response to Metabolic Stress. *Cell Metabolism*, 28(3), 516-524. doi:10.1016/j.cmet.2018.06.008
- King, L., & Plun-Favreau, H. (2017). Mitophagy. *Parkinson's Disease*, 139–177. Academic Press. doi:10.1016/b978-0-12-803783-6.00005-5
- Klaus, S., & Ost, M. (2020). Mitochondrial uncoupling and longevity – A role for mitokines? *Experimental Gerontology*, 130, 110796. doi:10.1016/j.exger.2019.110796
- Kotrys, & Szczesny. (2019). Mitochondrial Gene Expression and Beyond—Novel Aspects of Cellular Physiology. *Cells*, 9(1), 17. doi:10.3390/cells9010017
- Kozak, M. (1987). At least six nucleotides preceding the AUG initiator codon enhance translation in mammalian cells. *Journal of Molecular Biology*, 196(4), 947–950. doi:10.1016/0022-2836(87)90418-9
- Kubli, D. A., & Gustafsson, A. B. (2012). Mitochondria and Mitophagy: The Yin and Yang of Cell Death Control. *Circulation Research*, 111(9), 1208–1221. doi:10.1161/circresaha.112.265819
- Kültz, D. (2005). Molecular and evolutionary basis of the cellular stress response. *Annual Review of Physiology*, 67(1), 225–257. doi:10.1146/annurev.physiol.67.040403.103635
- Kurmi, K., & Haigis, M. C. (2020). Nitrogen Metabolism in Cancer and Immunity. *Trends in Cell Biology*, 30(5), 408-424. doi:10.1016/j.tcb.2020.02.005
- Landry, C. R., Zhong, X., Nielly-Thibault, L., & Roucou, X. (2015). Found in translation: functions and evolution of a recently discovered alternative proteome. *Current Opinion in Structural Biology*, 32, 74–80. doi:10.1016/j.sbi.2015.02.017

- Lee, C., Zeng, J., Drew, B. G., Sallam, T., Martin-Montalvo, A., Wan, J., Kim, S.-J., Mehta, H., Hevener, A. L., de Cabo, R., & Cohen, P. (2015). The Mitochondrial-Derived Peptide MOTS-c Promotes Metabolic Homeostasis and Reduces Obesity and Insulin Resistance. *Cell Metabolism*, *21*(3), 443–454. doi:10.1016/j.cmet.2015.02.009
- López-Otín, C., & Bond, J. S. (2008). Proteases: Multifunctional Enzymes in Life and Disease. *Journal of Biological Chemistry*, *283*(45), 30433–30437. doi:10.1074/jbc.r800035200
- Marchetti, P., Fovez, Q., Germain, N., Khamari, R., & Kluza, J. (2020). Mitochondrial spare respiratory capacity: Mechanisms, regulation, and significance in non-transformed and cancer cells. *The FASEB Journal*, *34*(10), 13106–13124. doi:10.1096/fj.202000767r
- Martijn, J., Vosseberg, J., Guy, L., Offre, P., & Ettema, T. J. G. (2018). Deep mitochondrial origin outside the sampled alphaproteobacteria. *Nature*, *557*(7703), 101–105. doi:10.1038/s41586-018-0059-5
- Martinez, T. F., Chu, Q., Donaldson, C., Tan, D., Shokhirev, M. N., & Saghatelian, A. (2020). Accurate annotation of human protein-coding small open reading frames. *Nature chemical biology*, *16*(4), 458-468. doi:10.1038/s41589-019-0425-0
- Matos, B., Publicover, S. J., Castro, L. F. C., Esteves, P. J., & Fardilha, M. (2021). Brain and testis: more alike than previously thought? *Open Biology*, *11*(6), 200322. doi:10.1098/rsob.200322
- McGee, A. M., & Baines, C. P. (2011). Complement 1q-binding protein inhibits the mitochondrial permeability transition pore and protects against oxidative stress-induced death. *Biochemical Journal*, *433*(1), 119–125. doi:10.1042/bj20101431

- Merry, T. L., Chan, A., Woodhead, J. S. T., Reynolds, J. C., Kumaga, H., Kim, S.-J., & Lee, C. (2020). Mitochondrial-derived peptides in energy metabolism. *American Journal of Physiology-Endocrinology and Metabolism*, 319(4), E659-E666. doi:10.1152/ajpendo.00249.2020
- Meydan, S., Vázquez-Laslop, N., & Mankin, A. S. (2018). Genes within Genes in Bacterial Genomes. *Microbiology Spectrum*, 6(4). doi:10.1128/microbiolspec.rwr-0020-2018
- Mi, H., Muruganujan, A., Ebert, D., Huang, X., & Thomas, P. D. (2018). PANTHER version 14: more genomes, a new PANTHER GO-slim and improvements in enrichment analysis tools. *Nucleic Acids Research*, 47(D1), D419-D426. doi:10.1093/nar/gky1038
- Miller, B., Kim, S.-J., Kumagai, H., Mehta, H. H., Xiang, W., Liu, J., Yen, K., Cohen, P. (2020). Peptides derived from small mitochondrial open reading frames: Genomic, biological, and therapeutic implications. *Experimental Cell Research*, 393(2), 112056.. doi:10.1016/j.yexcr.2020.112056
- Miller, B., Kim, S.-J., Mehta, H. H., Cao, K., Kumagai, H., Thumaty, N., Leelaprachakul, N., Jiao, H., Vaughan, J., Diedrich, J., Saghatelian, A., Arpawong, T. E., Crimmins, E. M., Ertekin-Taner, N., Tubi, M. A., Hare, E. T., Braskie, M. N., Décarie-Spain, L., Kanoski, S. E., ... & Cohen, P. (2022). Mitochondrial DNA variation in Alzheimer's disease reveals a unique microprotein called SHMOOSE. *Molecular Psychiatry*, 1-14. doi:10.1038/s41380-022-01769-3
- Mnatsakanyan, N., Hook, J. A., Quisenberry, L., & Weber, J. (2009). ATP Synthase with Its  $\gamma$  Subunit Reduced to the N-terminal Helix Can Still Catalyze ATP Synthesis. *Journal of Biological Chemistry*, 284(39), 26519–26525. doi:10.1074/jbc.m109.030528
- Moulleron, H., Delcourt, V., & Roucou, X. (2015). Death of a dogma: eukaryotic mRNAs can code for more than one protein. *Nucleic Acids Research*, 44(1), 14–23. doi:10.1093/nar/gkv1218

- Murray, J. I., Whitfield, M. L., Trinklein, N. D., Myers, R. M., Brown, P. O., & Botstein, D. (2004). Diverse and Specific Gene Expression Responses to Stresses in Cultured Human Cells. *Molecular Biology of the Cell*, *15*(5), 2361–2374. doi:10.1091/mbc.e03-11-0799
- Orr, M. W., Mao, Y., Storz, G., & Qian, S. B. (2020). Alternative ORFs and small ORFs: shedding light on the dark proteome. *Nucleic Acids Research*, *48*(3), 1029-1042. doi:10.1093/nar/gkz734
- Pastukh, V. M., Gorodnya, O. M., Gillespie, M. N., & Ruchko, M. V. (2016). Regulation of mitochondrial genome replication by hypoxia: The role of DNA oxidation in D-loop region. *Free Radical Biology and Medicine*, *96*, 78–88. doi:10.1016/j.freeradbiomed.2016.04.011
- Pi, H., Huang, Y.-C., Chen, I.-C., Lin, C.-D., Yeh, H.-F., & Pai, L.-M. (2011). Identification of 11-amino acid peptides that disrupt Notch-mediated processes in *Drosophila*. *Journal of Biomedical Science*, *18*(1), 42. doi:10.1186/1423-0127-18-42
- Picard, M., Wallace, D. C., & Burrelle, Y. (2016). The rise of mitochondria in medicine. *Mitochondrion*, *30*, 105–116. doi:10.1016/j.mito.2016.07.003
- Prabhu, L., Hartley, A.-V., Martin, M., Warsame, F., Sun, E., & Lu, T. (2015). Role of post-translational modification of the Y box binding protein 1 in human cancers. *Genes & Diseases*, *2*(3), 240–246. doi:10.1016/j.gendis.2015.05.001
- Rahman, J., Singh, P., Merle, N. S., Niyonzima, N., & Kemper, C. (2020). Complement's favorite organelle – mitochondria? *British Journal of Pharmacology*, *178*(14), 2771-2785. doi:10.1111/bph.15238
- Ramanathan, A., Robb, G. B., & Chan, S.-H. (2016). mRNA capping: biological functions and applications. *Nucleic Acids Research*, *44*(16), 7511–7526. doi:10.1093/nar/gkw551

- Ramazi, S., & Zahiri, J. (2021). Post-translational modifications in proteins: resources, tools and prediction methods. *Database*, 2021. doi:10.1093/database/baab012
- Rink, C., & Khanna, S. (2011). Significance of Brain Tissue Oxygenation and the Arachidonic Acid Cascade in Stroke. *Antioxidants & Redox Signaling*, 14(10), 1889–1903. doi:10.1089/ars.2010.3474
- Rodriguez, J., Gupta, N., Smith, R. D., & Pevzner, P. A. (2008). Does Trypsin Cut Before Proline? *Journal of Proteome Research*, 7(1), 300–305. doi:10.1021/pr0705035
- Roger, A. J., Muñoz-Gómez, S. A., & Kamikawa, R. (2017). The Origin and Diversification of Mitochondria. *Current Biology*, 27(21), R1177–R1192. doi:10.1016/j.cub.2017.09.015
- Schägger, H., & von Jagow, G. (1987). Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Analytical Biochemistry*, 166(2), 368–379. doi:10.1016/0003-2697(87)90587-2
- Schneider, C. A., Rasband, W. S., & Eliceiri, K. W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nature Methods*, 9(7), 671–675. doi:10.1038/nmeth.2089
- Sena, L. A., & Chandel, N. S. (2012). Physiological Roles of Mitochondrial Reactive Oxygen Species. *Molecular Cell*, 48(2), 158–167. doi:10.1016/j.molcel.2012.09.025
- Sherman, B. T., Hao, M., Qiu, J., Jiao, X., Baseler, M. W., Lane, H. C., Imamichi, T., & Chang, W. (2022). DAVID: a web server for functional enrichment analysis and functional annotation of gene lists (2021 update). *Nucleic Acids Research*, 50(W1), W216–W221, doi: 10.1093/nar/gkac194

- Shine, J., & Dalgarno, L. (1974). The 3'-Terminal Sequence of Escherichia coli 16S Ribosomal RNA: Complementarity to Nonsense Triplets and Ribosome Binding Sites. *Proceedings of the National Academy of Sciences*, *71(4)*, 1342–1346. doi:10.1073/pnas.71.4.1342
- Sies, H., & Jones, D. P. (2020). Reactive oxygen species (ROS) as pleiotropic physiological signalling agents. *Nature Reviews Molecular Cell Biology*, *21(7)*, 363–383. doi:10.1038/s41580-020-0230-3
- Sonenberg, N., Hinnebusch, A. G. (2009). Regulation of translation initiation in eukaryotes: mechanisms and biological targets. *Cell*, *136(4)*, 731–745. doi: 10.1016/j.cell.2009.01.042.
- Stelzer, G., Rosen, N., Plaschkes, I., Zimmerman, S., Twik, M., Fishilevich, S., Iny Stein, T., Nudel, R., Lieder, I., Mazor, Y., Kaplan, S., Dahary, D., Warshawsky, D., Guan-Golan, Y., Kohn, A., Rappaport, N., Safran, M., & Lancet, D. (2016). The GeneCards suite: from gene data mining to disease genome sequence analyses. *Current protocols in bioinformatics*, *54(1)*, 1-30. doi:10.1002/cpbi.5
- Su, M., Ling, Y., Yu, J., Wu, J., & Xiao, J. (2013). Small proteins: untapped area of potential biological importance. *Frontiers in Genetics*, *4*, 286. doi:10.3389/fgene.2013.00286
- Temperley, R. J., Wydro, M., Lightowlers, R. N., & Chrzanowska-Lightowlers, Z. M. (2010). Human mitochondrial mRNAs—like members of all families, similar but different. *Biochimica et Biophysica Acta (BBA) - Bioenergetics*, *1797(6-7)*, 1081–1085. doi:10.1016/j.bbabi.2010.02.036
- Temperley, R. J., Wydro, M., Lightowlers, R. N., & Chrzanowska-Lightowlers, Z. M. (2010). Human mitochondrial mRNAs—like members of all families, similar but different. *Biochimica et Biophysica Acta (BBA) - Bioenergetics*, *1797(6-7)*, 1081–1085. doi:10.1016/j.bbabi.2010.02.036
- The Gene Ontology Consortium. (2020). The Gene Ontology resource: enriching a GOLD mine. *Nucleic Acids Research*, *49(D1)*, D325–D334. <https://doi.org/10.1093/nar/gkaa1113>

- The Human Protein Atlas. (s. d.). Tissue expression of *MT-ND4*.  
<https://www.proteinatlas.org/ENSG00000198886-MT-ND4/tissue>
- The UniProt Consortium. (2021). UniProt: the universal protein knowledgebase in 2021. *Nucleic Acids Research*, *49(D1)*, D480–D489. doi:10.1093/nar/gkaa1100
- ThermoFisher Scientific. (2019). *BL21(DE3) Competent Cells*. (Pub. No. MAN0018595).  
[https://assets.thermofisher.com/TFS-Assets/LSG/manuals/MAN0018595\\_BL21\\_DE3competent\\_cells\\_UG.pdf](https://assets.thermofisher.com/TFS-Assets/LSG/manuals/MAN0018595_BL21_DE3competent_cells_UG.pdf)
- ThermoFisher Scientific. (2019). *Transcriptome Analysis Console (TAC) 4.0.2*. (Pub. No. 703150). [https://tools.thermofisher.com/content/sfs/brochures/tac\\_software\\_datasheet.pdf](https://tools.thermofisher.com/content/sfs/brochures/tac_software_datasheet.pdf)
- Tunca, S., Barreiro, C., Coque, J.-J. R., & Martín, J. F. (2009). Two overlapping antiparallel genes encoding the iron regulator DmdR1 and the Adm proteins control siderophore and antibiotic biosynthesis in *Streptomyces coelicolor* A3(2). *FEBS Journal*, *276(17)*, 4814–4827. doi:10.1111/j.1742-4658.2009.07182.x
- Turcotte, M. L., Parliament, M., Franko, A., & Allalunis-Turner, J. (2002). Variation in mitochondrial function in hypoxia-sensitive and hypoxia-tolerant human glioma cells. *British Journal of Cancer*, *86(4)*, 619–624. doi:10.1038/sj.bjc.6600087
- Voigt, A., & Jelinek, H. F. (2016). Humanin: a mitochondrial signaling peptide as a biomarker for impaired fasting glucose-related oxidative stress. *Physiological Reports*, *4(9)*, e12796. doi:10.14814/phy2.12796
- Wagner, H. J., & Weber, W. (2019). Design of a Human Rhinovirus-14 3C Protease-Inducible Caspase-3. *Molecules*, *24(10)*, 1945. doi:10.3390/molecules24101945



- Walker, J. E., Saraste, M., Runswick, M. J., & Gay, N. J. (1982). Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *The EMBO Journal*, *1*(8), 945–951. doi:10.1002/j.1460-2075.1982.tb01276.x
- Wallace, D. C. (2018). Mitochondrial genetic medicine. *Nature Genetics*. *50*(12), 1642-1649. doi:10.1038/s41588-018-0264-z
- Wang, F., Zhang, D., Zhang, D., Li, P., & Gao, Y. (2021). Mitochondrial protein translation: emerging roles and clinical significance in disease. *Frontiers in Cell and Developmental Biology*, *9*. doi:10.3389/fcell.2021.675465
- Wang, J., Huang, C. L., & Zhang, Y. (2022). Complement C1q Binding Protein (C1QBP): Physiological Functions, Mutation-Associated Mitochondrial Cardiomyopathy and Current Disease Models. *Frontiers in Cardiovascular Medicine*, *9*. doi:10.3389/fcvm.2022.843853
- Withers, P. C., & Cooper, C. E. (2008). Dormancy. *Encyclopedia of Ecology*, 952–957. Elsevier. doi:10.1016/b978-008045405-4.00503-6
- Yang, C., Hockenberry, A. J., Jewett, M. C., & Amaral, L. A. (2016). Depletion of Shine-Dalgarno sequences within bacterial coding regions is expression dependent. *G3: Genes, Genomes, Genetics*, *6*(11), 3467–3474. doi:10.1534/g3.116.032227
- Yang, Y., & Wang, Z. (2019). IRES-mediated cap-independent translation, a path leading to hidden proteome. *Journal of Molecular Cell Biology*, *11*(10), 911–919. doi:10.1093/jmcb/mjz091
- Yue, D., Wang, Y., Sun, Y., Niu, Y., & Chang, C. (2017). C1QBP regulates YBX1 to suppress the androgen receptor (AR)-enhanced RCC cell invasion. *Neoplasia*, *19*(2), 135-144. doi : 10.1016/j.neo.2016.12.003

Zhang, X., Zhang, F., Guo, L., Wang, Y., Zhang, P., Wang, R., Zhang, N., & Chen, R. (2013). Interactome Analysis Reveals that C1QBP (complement component 1, q subcomponent binding protein) Is Associated with Cancer Cell Chemotaxis and Metastasis. *Molecular & Cellular Proteomics*, *12*(11), 3199–3209. doi:10.1074/mcp.m113.029413

Zimmerman, J. J., von Saint André-von Arnim, A., & McLaughlin, J. (2011). Cellular Respiration. *Pediatric Critical Care*, 1058–1072. doi:10.1016/b978-0-323-07307-3.10074-6

Zorov, D. B., Juhaszova, M., & Sollott, S. J. (2006). Mitochondrial ROS-induced ROS release: An update and review. *Biochimica et Biophysica Acta (BBA) - Bioenergetics*, *1757*(5-6), 509–517. doi:10.1016/j.bbabi.2006.04.029

Zymo Research. (2021). *Quick-RNA™ Miniprep Kit*.  
[https://files.zymoresearch.com/protocols/\\_r1054\\_r1055\\_quick-rna\\_miniprep\\_kit.pdf](https://files.zymoresearch.com/protocols/_r1054_r1055_quick-rna_miniprep_kit.pdf)

## Annexe

### Annexe 1 : Article de Kienzle *et al.* (2022) en prépublication

#### Annexe 1.1 : L'intégrale de Kienzle *et al.* (2022)

bioRxiv preprint doi: <https://doi.org/10.1101/2022.04.28.489924>; this version posted April 28, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

#### **MTALTND4, a second protein coded by *nd4* impacts mitochondrial bioenergetics**

Laura Kienzle<sup>1,\*</sup>, Stefano Bettinazzi<sup>1,\*</sup>, Marie Brunet<sup>2,3</sup>, Thierry Choquette<sup>1</sup>, Hajar Hosseini Khorami<sup>1</sup>, Xavier Roucou<sup>4</sup>, Christian R Landry<sup>5-9</sup>, Annie Angers<sup>1</sup>, Sophie Breton<sup>1,†</sup>

<sup>1</sup>Département de Sciences Biologiques, Université de Montréal, Montréal, Canada

<sup>2</sup>Service de Génétique Médicale, Département de Pédiatrie, Université de Sherbrooke, Sherbrooke, Canada

<sup>3</sup>Centre de Recherche du Centre hospitalier universitaire de Sherbrooke (CRCHUS), Sherbrooke, Canada

<sup>4</sup>Département de Biochimie, Université de Sherbrooke, Sherbrooke, Canada

<sup>5</sup>Département de Biochimie, de Microbiologie et de Bio-informatique, Faculté des Sciences et de Génie, Université Laval, Québec, Canada

<sup>6</sup>Institut de Biologie Intégrative et des Systèmes, Université Laval, Québec, Canada

<sup>7</sup>PROTEO, Le regroupement québécois de recherche sur la fonction, l'ingénierie et les applications des protéines, Université Laval, Québec, Canada

<sup>8</sup>Centre de Recherche sur les Données Massives, Université Laval, Québec, Canada

<sup>9</sup>Département de Biologie, Faculté des Sciences et de Génie, Université Laval, Québec, Canada

**\*First co-authors**

†**Corresponding author:** Sophie Breton, Département de Sciences Biologiques, Université de Montréal, Montréal, QC H3C 3J7, Canada  
Email: s.breton@umontreal.ca

## **ABSTRACT**

Recent evidence suggests that the coding potential of the mitogenome is underestimated. We found a downstream alternative ATG initiation codon in the +3 reading frame of the human mitochondrial *nd4* gene. This newly characterized alternative open reading frame (altORF) encodes a 99-amino acids long polypeptide, MTALTND4, which is conserved in primates. This small protein is localized in mitochondria and cytoplasm and is also found in the plasma, and it impacts mitochondrial physiology. Alternative mitochondrial peptides such as MTALTND4 may offer a new framework for the investigation of mitochondrial functions and diseases.

## **MAIN**

The recent identification in bacterial and eukaryotic genomes (including human) of thousands of unannotated yet translated small open reading frames (smORFs) and alternative open reading frames in reference genes (altORFs) has brought to light a whole new class of important genes (*1-4*). Although the elucidation of their function remains challenging, peptides derived from smORFs and altORFs reveal a new facet of eukaryote and bacterial proteomes, enabling a better understanding of key evolutionary processes such as the birth of novel genes and proteins, and providing novel fundamental biological knowledge as well as new avenues for therapeutic discovery. Mitochondria-derived proteomes do not escape this reality, given that the proteomes of their bacterial ancestors and ‘modern host cells’ are richer than previously anticipated (*1-4*). Up to now, 16-38 amino acid long micropeptides of functional importance have been shown to

be encoded in the human mitochondrial 16S and 12S rRNA genes (5-9). These small proteins, called Humanin, SHLPs (small humanin-like peptides 1-6) and MOTS-c appear to modulate mitochondrial and cellular biology (8), as well as global physiology, acting as signaling molecules secreted from cells and found in circulation (plasma) (5, 6). However, whether altORFs and smORFs exist in mitochondrial protein-coding genes or elsewhere in the human mitogenome has not been addressed yet.

A preliminary *in silico* examination of the unresolved protein-coding potential of the human mtDNA has revealed 227 unannotated ORFs of at least 60 nucleotides on either strand (Fig. 1). To identify potentially novel mtDNA-encoded proteins, unannotated ORFs (Fig. 1 and table S1) were selected for further investigation using three different approaches: (i) we searched among the 227 unannotated ORFs for those having a minimal conserved Kozak sequence known to be favorable for efficient translation initiation (5, 10) (table S2), (ii) we interrogated the OpenProt database (table S3), which enforces a polycistronic model of eukaryotic genome annotations (11), and (iii) we interrogated mitochondria-derived mass spectrometry datasets to see if peptides derived from our 227 unannotated ORFs could be detected (table S4 and fig. S1). These three approaches respectively returned 13, 14 and 36 candidates (two OpenProt candidates were found also in the MS dataset) and we respectively chose one, 4 and 4 of these candidates for antibody production (Fig. 1 and tables S1-S4). Immunoblotting was used to assess their expression in human HeLa and HEK-293T cells.

Of these nine candidates, one unannotated ORF-derived peptide, which was identified by mass spectrometry (approach 3; see tables S1 and S4), was detected by western blotting and its mitochondrial origin was validated with the loss of western blot signal in cells devoid of endogenous mtDNA (Rho0) and in cells treated with chloramphenicol that inhibits mitochondrial

protein synthesis, limiting the possibility of a nuclear origin due to a nuclear mtDNA transfer [NUMT; (14)] (fig. S2). This previously unannotated ORF-derived peptide is referred to as MTALTND4 (Mitochondrial Alternative ND4 protein) because the ORF sequence is found within the mitochondrial *nd4* gene in an alternative reading frame (Fig. 1 and Fig. 2A-E). The *mtalnd4* altORF is 300 nucleotides long and code for a putative 99 amino-acid peptide (with no predicted transmembrane helix) using the mitochondrial vertebrate genetic code and considering that human mitochondria use only UAA and UAG as stop codons as recently suggested (12, 13) (Fig. 2A and table S1; see supplementary text 1). Multiple peptide sequence alignments indicate that MTALTND4 is well conserved in primates in general (Fig. 2B and fig. S3; see supplementary text 2). Although somewhat higher than expected, a specific 25 kDa western blot signal was observed in both HeLa and HEK-293T cells for endogenous MTALTND4 (Fig. 2C). Our results showed that endogenous MTALTND4 is localized in the cytoplasm where it is co-localized to mitochondria (Fig. 2F), a pattern that is similar also for Humanin, MOTs-c and SHLPs (5-7), and the absence of signal in HeLa-chloramphenicol treated cells confirmed again the mitochondrial origin of the alternative peptide (fig. S4). Our custom antibody, but not the pre-immune serum, was able to immunoprecipitate MTALTND4 from HeLa cell lysates, confirming the existence of an endogenous MTALTND4 peptide (Table S5).

Similar to Humanin, MOTs-c and SHLPs (5-7), the presence of MTALTND4 in the plasma was confirmed by western blot (Fig. 2G), which was not the case for the mitochondrial protein ATP5 (encoded by the nuclear DNA and also devoid of predicted TMH). Following its detection in the plasma, we hypothesized that MTALTND4, as other mitochondria-derived micropeptides (6, 7), could act as a signaling molecule secreted from cells in the circulation and regulate cell physiology. We thus examined its exogenous effect on cellular metabolism using

HeLa and HEK-293T cells treated with the complete synthetic peptide at different concentrations. Overall, we observed a strong, immediate effect of MTALTND4 on mitochondrial respiration. Specifically, the presence of the peptide decreased oxygen consumption in both HeLa and HEK-293T cells. The effect appeared to be dose-dependent, although the concentration threshold differed between cell lines. For example, MTALTND4 at a concentration of 5  $\mu\text{M}$  was sufficient to impact the basal oxygen consumption rate (i.e. the ROUTINE respiration) of HeLa cells, whereas up to 30  $\mu\text{M}$  was necessary in the case of HEK-293T cells (Fig. 3A,B; supplementary table S6). Such effect was not observed using the shorter antigenic peptide as a control. It is worth mentioning that a concentration of 30  $\mu\text{M}$  did not negatively impact HeLa or HEK cell proliferation or viability after 24h treatment (see below).

Additional evidence of the exogenous impact of MTALTND4 comes from analyzing mitochondrial respiration in different respiratory states (Fig. 3C,D; supplementary table S7). In HeLa cells, 10  $\mu\text{M}$  MTALTND4 significantly decreased both the basal respiration (routine respiration, ce-pept\_R) and the maximal uncoupled respiration (ce-pept\_E), but not the residual leak respiration (ce-pept\_L) (Fig. 3C). In HEK-293T cells, both leak respiration and maximal uncoupled respiration, but not the routine respiration, were impacted in presence of 10  $\mu\text{M}$  MTALTND4 (Fig. 3D). This is in line with the dose-dependent effect of the peptide and the higher threshold revealed for HEK-293T compared to HeLa cells (Fig. 3A,B). The analysis of the ‘spare respiratory capacity’ (SRC) – i.e. the difference between the basal respiration and the maximal uncoupled respiration (E-R) [(17) and reference therein] – revealed a main effect of both factors ‘treatment’ (control vs peptide addition) and ‘cell type’ (HeLa vs HEK-293T), with no interaction between the two (Fig. 3E; supplementary table S7). In accordance with the general notion that cancer cells generally show lower SRC levels compared to their normal counterparts

(17), our results suggest that HeLa cells have a physiological respiration much closer to their maximum capacity, thus a lower mitochondrial reserve compared to HEK-293T cells. A smaller reserve capacity likely makes these cells more sensitive to MTALTND4 treatment, as a lower decrease in maximal respiration is needed to impact the physiological routine respiration. In other words, HeLa cells were less able to withstand the inhibitory effect of MTALTND4 because when E-R reaches zero, mitochondrial energy production fails to meet the minimal needs of the cell.

The impact of MTALTND4 upon mitochondrial respiratory properties does not require cells to be intact as we examined its effect in permeabilized HEK-293T cells and observed that 10  $\mu$ M MTALTND4 impacted both coupled (CI+II-pept\_P) and uncoupled maximal respiration (CI+II-pept\_E) sustained by CI+II-linked substrates. Although not significant, a sharp decreasing trend in O<sub>2</sub> consumption was also revealed when testing cytochrome *c* oxidase standalone capacity (CIV-pept\_E) (Fig. 3F; supplementary table S8). This result is of particular interest and deserves further investigations, as it may indicate that the site of action for MTALTND4 within the electron transport system (ETS) could reside in its final oxidase.

A change in reactive oxygen species (ROS – hydrogen peroxide) homeostasis was also observed. Intact HEK-293T cells in presence of 30  $\mu$ M MTALTND4 showed an immediate, steady increase in H<sub>2</sub>O<sub>2</sub> efflux but no increase in the total H<sub>2</sub>O<sub>2</sub> scavenging capacity measured after 4 hours of incubation (Fig. 3G,H; supplementary table S9). An increased antioxidant capacity would be expected to lower cell oxidative stress. However, an alternative role for the revealed increased ROS efflux is also possible. Rather than being just deleterious by-products of an impaired mitochondrial respiration, mitochondrial ROS are pivotal mediators for many physiological processes including mitochondrial biogenesis, cell proliferation and differentiation,



aging, apoptosis and response to hypoxia (17, 18). Given its impact on mitochondrial activity (potentially targeting CIV itself) and ROS metabolism, the potential role of MTALTND4 in cell adaptive response to stress such as hypoxia surely need further attention.

The reduced mitochondrial respiration was not accompanied by a parallel increase in anaerobic glycolysis to partially fulfill the cell energy demand. Four hours of exogenous treatment with MTALTND4 (30  $\mu$ M) had no impact on the activity of lactate dehydrogenase in HEK-293T cells (Fig. 3I; supplementary table S9). Interestingly, despite a reduced mitochondrial capacity and no-change in lactic fermentation, the total ATP-content of treated cells increased (Fig. 3L; supplementary table S9). These results are potentially in line with a general cell metabolism undergoing downregulation, i.e. involving a coordinate downregulation of ATP-producing and ATP-demanding processes. This is further supported by the analyses of the effect of MTALTND4 (0.1  $\mu$ M, 10  $\mu$ M and 30  $\mu$ M) upon HeLa or HEK-293T cell proliferation and viability at three different time points (24, 48 and 72 hours). A dose-dependent impact of the peptide upon cell proliferation was revealed at 48 hours, with a trend still perceivable at 72 hours although no more significant (Fig. 3M; supplementary table S10). The impact of the peptide upon cell viability was less pronounced, with no significant impact after 72 hours of treatment (blue trypan) (Fig. 3N; supplementary table S10), or with a small, but significant negative impact mainly for HeLa cells after 48h (alamarBlue) (fig. S5).

To summarize, we identified a novel human mitochondrial alternative peptide of 99 amino acids (i.e. longer than ATP8 [68 a.a.] and ND4L [98 a.a.]) whose sequence resides inside the *nd4* gene. MTALTND4, which is relatively well conserved in other primate species, is translated inside mitochondria and is also found in the plasma, and it appears to help regulate mitochondrial physiology. The sequence of MTALTND4 was detected by mass spectrometry

after immunoprecipitation from HeLa cell extracts. Other proteins were also found that could be potential interacting partners of MTALTND4 (table S5). One of these proteins, the complement component 1q subcomponent binding protein (C1qbp), was also identified as an interacting partner using pull down assays, an interaction that was validated by Western blot (table S5). C1qbp, also known as p32, gC1qR, or HABP1, is a multifunctional and multicompartmental protein mostly localized in the mitochondrial matrix but also present at the cell surface, cytosol, and nucleus (19, 20). C1qbp is an evolutionarily conserved protein that can interact with a diverse array of mitochondrial and cellular, plasma, and microbial proteins (20). C1qbp regulates cellular energy metabolism through modulation of mitochondrial translation or through protein-protein interactions (21, 22). It has also a regulatory role in ROS production and can protect cells against oxidative stress (20, 23). This interaction could potentially explain the observed MTALTND4-mediated changes in mitochondrial respiration and ROS production. The lack of potential compensatory mechanisms accompanying the MTALTND4-mediated increase in H<sub>2</sub>O<sub>2</sub> efflux and downregulation of mitochondrial respiration such as increased antioxidant activity and upregulation of anaerobic glycolysis, together with the no/little impact on cell viability, suggest that the effects of MTALTND4 do not appear to be the result of a deleterious impairment of mitochondrial functions. This bioenergetic depression does however link with a sharp decrease in cell proliferation and increased total ATP content, suggesting that the cells might enter a sort of quiescent state.

Our findings have several implications. First, they suggest that mitochondrial genes in humans might have gone unnoticed. This calls into question the evolution of the mitochondrial genome as well as the selection pressures exerted on the mitochondria and the mechanisms allowing the translation of these alternative proteins. Our findings also offer a new framework

for the investigation of mitochondrial diseases. Mitochondrial DNA mutations are responsible for a variety of human disorders due to defects in oxidative energy metabolism, and they are also involved in aging, aging-related diseases and multiple types of cancer (24, 25). The diversity of pathologies caused by mtDNA mutations is remarkable and the relationship between genotype and disease phenotype is not always straightforward. For example, mutations within genes sometimes cannot be directly linked to effects on oxidative energy metabolism or mitochondrial protein synthesis. Breton (26) proposed the mitochondrial Russian doll genes hypothesis to explain some of these discrepancies between theoretical expectation and experimental observation when studying mitochondrial diseases linked to mutations in the mtDNA. Indeed, molecular screening of mitochondrial disorders has been usually restricted to common harmful mutations (i.e. nonsynonymous or “non-silent” mutations) and deletions of mtDNA, i.e. without taking into account rare mutations or silent mutations in protein-coding genes (27). However, a silent mutation in a typical mitochondrial gene could be non-silent in an altORF. In other words, the existence of an alternative mitochondrial proteome could well explain some of the discrepancies observed previously, and therefore, this possibility clearly requires a systematic analysis. In sum, not only does this study enhance our understanding of mitochondrial biology, but it also provides a new mechanism through which mtDNA mutations might generally affect health. Collectively, we anticipate that more studies on the alternative mitochondrial proteome will expedite the discovery of new mitochondrial genes with key biological roles.

## REFERENCES

1. C. R. Landry, X. Zhong, L. Nielly-Thibault, Found in translation: functions and evolution of a recently discovered alternative proteome. *Current Opinion in Structural Biology* **32**, 74-80 (2015).

2. M. A. Brunet, S. Leblanc, X. Roucou, Reconsidering proteomic diversity with functional investigation of small ORFs and alternative ORFs. *Experimental Cell Research* **393**, 112057 (2020).
3. T. F. Martinez, Q. Chu, C. Donaldson, D. Tan, M. N. Shokhirev, A. Saghatelian, Accurate annotation of human protein-coding small open reading frames. *Nature Chemical Biology* **16**, 458-468 (2020).
4. M. W. Orr, Y. Mao, G. Storz, S. B. Qian, Alternative ORFs and small ORFs: shedding light on the dark proteome. *Nucleic Acids Research* **48**, 1029-1042 (2020).
5. C. Lee, J. Zeng, B. G. Drew, T. Sallam, A. Martin-Montalvo, J. Wan, S. J. Kim, H. Mehta, A. L. Hevener, R. de Cabo, P. Cohen, The mitochondrial-derived peptide MOTS-c promotes metabolic homeostasis and reduces obesity and insulin resistance. *Cell Metabolism* **21**, 443-454 (2015).
6. L. J. Cobb, C. Lee, J. Xiao, K. Yen, R. G. Wong, H. K. Nakamura, H. H. Mehta, Q. Gao, C. Ashur, D. M. Huffman, J. Wan, R. Muzumdar, N. Barzilai, P. Cohen, Naturally occurring mitochondrial-derived peptides are age-dependent regulators of apoptosis, insulin sensitivity, and inflammatory markers. *Aging* **8**, 796-808 (2016).
7. K. H. Kim, J. M. Son, B. A. Benayoun, C. Lee, The mitochondrial-encoded peptide MOTS-c translocates to the nucleus to regulate nuclear gene expression in response to metabolic stress. *Cell Metabolism* **28**, 516-524 (2018).
8. B. Miller, S. J. Kim, H. Kumagai, H. H. Mehta, W. Xiang, J. Liu, K. Yen, P. Cohen, Peptides derived from small mitochondrial open reading frames: genomic, biological, and therapeutic implications. *Experimental Cell Research* **393**, 112056 (2020).
9. S. Breton, F. Ghiselli, L. Milani, Mitochondrial Short-Term Plastic Responses and Long-Term Evolutionary Dynamics in Animal Species, *Genome Biology and Evolution* **13**, evab084 (2021).
10. G. P. Harhay, T. S. Sonstegard, J. W. Keele, M. P. Heaton, M. L. Clawson, W. M. Snelling, R. T. Wiedmann, C. P. Van Tassell, T. P. Smith, Characterization of 954 bovine full-CDS cDNA sequences. *BMC Genomics* **6**, 166 (2005).
11. M. A. Brunet, J. F. Lucier, M. Levesque, S. Leblanc, J. F. Jacques, H. R. R. Al-Saedi, L. Guillo, F. Grenier, M. Avino, I. Fournier, M. Salzet, A. Ouangraoua, M. S. Scott, F. M. Boisvert, X. Roucou, OpenProt 2021: deeper functional annotation of the coding potential of eukaryotic genomes. *Nucleic Acids Research* **49**, D380-D388 (2021).
12. R. Temperley, R. Richter, S. Dennerlein, R. N. Lightowers, Z. M. Chrzanowska-Lightowers, Hungry codons promote frameshifting in human mitochondrial ribosomes. *Science* **327**, 301 (2010).

13. S. A. Ayyub, F. Gao, R. N. Lightowlers, Z. M. Chrzanowska-Lightowlers, Rescuing stalled mammalian mitoribosomes – what can we learn from bacteria?. *Journal of Cell Science* **133**, jcs231811 (2020).
14. M. Ricchetti, F. Tekaia, B. Dujon, Continued colonization of the human genome by mitochondrial DNA. *PLoS Biology* **2**, e273 (2004).
15. M. A. Rubio, J. J. Rinehart, B. Krett, S. Duvezin-Caubet, A. S. Reichert, D. Söll, J. D. Alfonzo, Mammalian mitochondria have the innate ability to import tRNAs by a mechanism distinct from protein import. *Proceedings of The National Academy of Sciences of The United States of America* **105**, 9186–9191 (2008).
16. D. Jeandard, A. Smirnova, I. Tarassov, E. Barrey, A. Smirnov, N. Entelis, Import of non-coding RNAs into human mitochondria: a critical review and emerging approaches. *Cells* **8**, 286 (2019).
17. P. Marchetti, Q. Fovez, N. Germain, R. Khamari, J. Kluza, Mitochondrial spare respiratory capacity: mechanisms, regulation, and significance in non-transformed and cancer cells. *The FASEB Journal* **34**, 13106-13124 (2020).
18. H. Sies, D. P. Jones, Reactive oxygen species (ROS) as pleiotropic physiological signalling agents. *Nature Reviews Molecular Cell Biology* **21**, 363-383 (2020).
19. J. Barna, D. Dimén, G. Puska, D. Kovács, V. Csikós, S. Oláh, E. B. Udvari, G. Pál, Á. Dobolyi, Complement component 1q subcomponent binding protein in the brain of the rat. *Scientific reports* **9**, 1-19 (2019).
20. X. Zhai , K. Liu, H. Fang, Q. Zhang, X. Gao, F. Liu, S. Zhou, X. Wang, Y. Niu, Y. Hong, S. H. Lin, W. H. Liu, C. Xiao, Q. Li, N. Xiao, Mitochondrial C1qbp promotes differentiation of effector CD8<sup>+</sup> T cells via metabolic-epigenetic reprogramming. *Science advances* **7**, eabk0490 (2021).
21. V. Fogal, A. D. Richardson, P. P. Karmali, I. E. Scheffler, J. W. Smith, E. Ruoslahti, Mitochondrial p32 protein is a critical regulator of tumor metabolism via maintenance of oxidative phosphorylation. *Molecular and Cellular Biology* **30**, 1303-18 (2010).
22. R. Chen, M. Xiao, H. Gao, Y. Chen, Y. Li, Y. Liu, N. Zhang, Identification of a novel mitochondrial interacting protein of C1QBP using subcellular fractionation coupled with CoIP-MS. *Analytical and Bioanalytical Chemistry* **408**, 1557-64 (2016).
23. A. M. McGee, C. P. Baines, Complement 1q-binding protein inhibits the mitochondrial permeability transition pore and protects against oxidative stress-induced death. *The Biochemical Journal* **433**, 119-125 (2011).
24. E. A. Schon, S. DiMauro, M. Hirano, Human mitochondrial DNA: roles of inherited and somatic mutations. *Nature Reviews Genetics* **12**, 878-890 (2012).

25. G. S. Gorman, P. F. Chinnery, S. DiMauro, M. Hirano, Y. Koga, R. McFarland, A. Suomalainen, D. R. Thorburn, M. Zeviani, D. M. Turnbull, Mitochondrial diseases. *Nature Reviews Disease Primers* **2**, 16080 (2016).
26. S. Breton, Mitochondrial Russian doll genes may explain some discrepancies in links between mtDNA mutations and mitochondrial diseases. *BioEssays* **43**, e2100104 (2021).
27. S. Bannwarth, V. Procaccio, A. S. Lebre, C. Jardel, A. Chaussenot, C. Hoarau, H. Maoulida, N. Charrier, X. Gai, H. M. Xie, M. Ferre, K. Fragaki, G. Hardy, B. Mousson de Camaret, S. Marlin, C. M. Dhaenens, A. Slama, C. Rocher, J. P. Bonnefont, A. Rötig, N. Aoutil, M. Gilleron, V. Desquiret-Dumas, P. Reynier, J. Ceresuela, L. Jonard, A. Devos, C. Espil-Taris, D. Martinez, P. Gaignard, K. H. Le Quan Sang, P. Amati-Bonneau, M. J. Falk, C. Florentz, B. Chabrol, I. Durand-Zaleski, V. Paquis-Flucklinger, Prevalence of rare mitochondrial DNA mutations in mitochondrial disorders. *Journal of Medical Genetics* **50**, 704-714 (2013).
28. A. Angers, P. Ouimet, A. Tsyvian-Dzyabko, T. Nock, S. Breton, L'ADN mitochondrial, un potentiel codant mésestimé [The underestimated coding potential of mitochondrial DNA]. *Medecine Sciences* **35**, 46-54 (2019).
29. E. Faure, L. Delaye, S. Tribolo, A. Levasseur, H. Seligmann, R. M. Barthélémy, Probable presence of an ubiquitous cryptic mitochondrial gene on the antisense strand of the cytochrome oxidase I gene. *Biology Direct* **6**, (2011).
30. M. A. Brunet, M. Brunelle, J. F. Lucier, V. Delcourt, M. Levesque, F. Grenier, S. Samandi, S. Leblanc, J. D. Aguilar, P. Dufour, J. F. Jacques, I. Fournier, A. Ouangraoua, M. S. Scott, F. M. Boisvert, X. Roucou, OpenProt: a more comprehensive guide to explore eukaryotic coding potential and proteoms. *Nucleic Acids Research* **47**, D403-D410 (2019).
31. B. Wen, X. Wang, B. Zhang, PepQuery enables fast, accurate, and convenient proteomic validation of novel genomic alterations. *Genome Research* **29**, 485-493 (2019).
32. S. F. Altschul, T. L. Madden, A. A. Schäffer, J. Zhang, Z. Zhang, W. Miller, D. J. Lipman, Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research* **25**, 3389-3402 (1997).
33. J. D. Thompson, D. G. Higgins, T. J. Gibson, CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research* **22**, 4673-4680 (1994).
34. K. Hofmann, W. Stoffel, TMbase-a database of membrane spanning proteins segments. *Biological Chemistry Hoppe Seyler* **347**, 166 (1993).
35. M. Audagnotto, M. Dal Peraro, Protein post-translational modifications: *in silico* prediction tools and molecular modeling. *Computational and Structural Biotechnology Journal* **15**, 307-319 (2017).

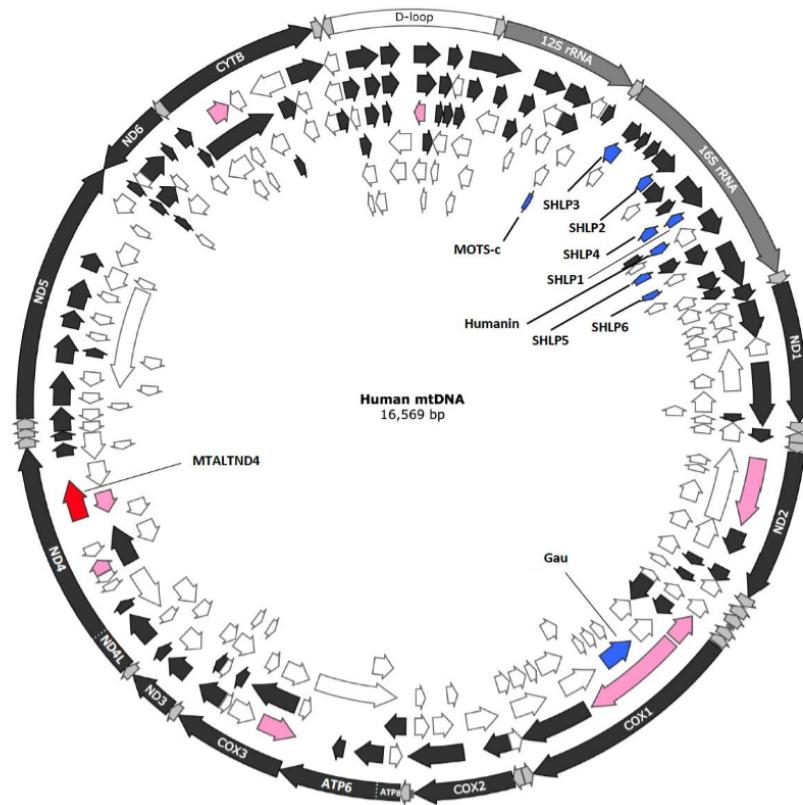


36. F. Tian, C. Wang, M. Tang, J. Li, X. Cheng, S. Zhang, D. Ji, Y. Huang, H. Li, The antibiotic chloramphenicol may be an effective new agent for inhibiting the growth of multiple myeloma. *Oncotarget* **7**, 51934-51942 (2016).
37. H. Schägger, G. von Jagow, Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Analytical Biochemistry* **166**, 368-379 (1987).
38. A. L. Grigorian, J. J. Bustamante, P. Hernandez, A. O Martinez, L. S. Haro, Extraordinarily stable disulfide-linked homodimer of human growth hormone. *Protein science: a publication of the Protein Society* **14**, 902-913 (2005).
39. C. Doerrier, L. F. Garcia-Souza, G. Krumschnabel, Y. Wohlfarter, A. T. Mészáros, E. Gnaiger, High-resolution fluorepirometry and OXPHOS protocols for human cells, permeabilized fibers from small biopsies of muscle, and isolated mitochondria. *Mitochondrial Bioenergetics* **1782**, 31-70 (2018).
40. E. Gnaiger, Mitochondrial pathways and respiratory control an introduction to OXPHOS analysis. *Bioenergetics Communications* **2020**, 122 (2020).
41. G. Oemer, K. Lackner, K. Muigg, G. Krumschnabel, K. Watschinger, S. Sailer, H. Lindner, E. Gnaiger, S. B. Wortmann, E. R. Werner, J. Zschocke, M. A. Keller, Molecular structural diversity of mitochondrial cardiolipins. *Proceedings of the National Academy of Sciences* **115**, 4158-4163 (2018).
42. D. Pesta, E. Gnaiger, High-resolution respirometry: OXPHOS protocols for human cells and permeabilized fibers from small biopsies of human muscle. *Methods in Molecular Biology* **810**, 25-58 (2012).
43. W. C. Orr, R. S. Sohal, The effects of catalase gene overexpression on life span and resistance to oxidative stress in transgenic *Drosophila melanogaster*. *Archives of Biochemistry and Biophysics* **297**, 35-41 (1992).
44. N. Pichaud, E. H. Chatelain, J. W. Ballard, R. Tanguay, G. Morrow, P.U. Blier, Thermal sensitivity of mitochondrial metabolism in two distinct mitotypes of *Drosophila simulans*: evaluation of mitochondrial plasticity. *The Journal of Experimental Biology* **213**, 1665-1675 (2010).
45. M. Thibault, P.U. Blier, H. Guderley, Seasonal variation of muscle metabolic organization in rainbow trout (*Oncorhynchus mykiss*). *Fish Physiology and Biochemistry* **16**, 139-155 (1997).
46. P. K. Smith, R. I. Krohn, G. T. Hermanson, A.K. Mallia, F. H. Gartner, M.D. Provenzano, E. K. Fujimoto, N. M. Goeke, B. J. Olson, D. C. Klenk, Measurement of protein using bicinchoninic acid. *Analytical Biochemistry* **150**, 76-85 (1985).

47. R Core Team, *R: A language and environment for statistical computing. R Foundation for Statistical Computing* (Vienna, Austria, 2021).
48. H. Wickham, *ggplot2: Elegant Graphics for Data Analysis* (Springer-Verlag, New York, 2016).
49. I. S. Logan, Pseudogenization of the *Humanin* gene is common in the mitochondrial DNA of many vertebrates. *Zoological Research* **38**, 198-202 (2017).
50. S. V. Emser, H. Schaschl, E. Millesi, R. Steinborn, Extension of Mitogenome Enrichment Based on Single Long-Range PCR: mtDNAs and Putative Mitochondrial-Derived Peptides of Five Rodent Hibernators. *Frontiers in Genetics* **12**, 685806 (2021).
51. D. Pierron, D. E. Wildman, M. Hüttemann, T. Letellier, L. I. Grossman, Evolution of the couple cytochrome c and cytochrome c oxidase in primates. *Advances in Experimental Medicine and Biology* **748**, 185-213 (2012).
52. B. Kadenbach, M. Hüttemann, The subunit composition and function of mammalian cytochrome c oxidase. *Mitochondrion* **24**, 64-76 (2015).
53. C. Rancurel, M. Khosravi, A. K. Dunker, P. R. Romero, D. Karlin, Overlapping genes produce proteins with unusual sequence properties and offer insight into de novo protein creation. *Journal of Virology* **83**, 10719-10736 (2009).
54. B. W. Wright, M. P. Molloy, P. R. Jaschke, Overlapping genes in natural and engineered genomes. *Nature Reviews Genetics* **23**, 154-168 (2022).

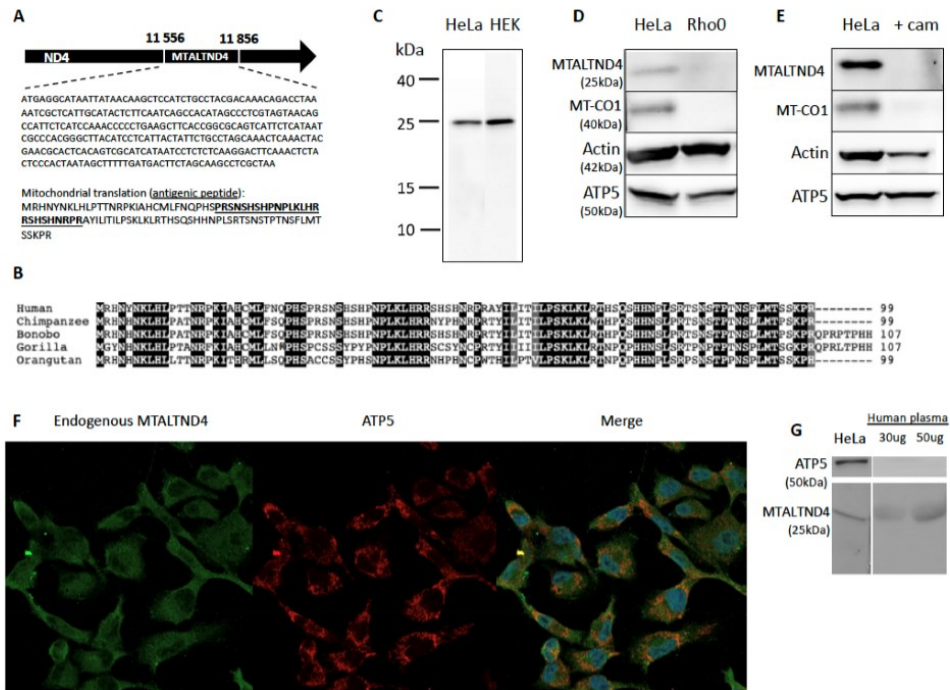


## FIGURES



**Figure 1. Small ORFs and alternative ORFs in the human mitochondrial genome.** In dark gray in the external circle: the typical 13 protein-coding genes. In gray in the external circle: the two rRNA genes. In white in the external circle: the d-loop. Arrows positioned inside the circle indicate smORFs and altORFs on the main coding 5'-3' (frames 1, 2 and 3: black arrows) and complementary 3'-5' (frames -1, -2 and -3: white arrows) strands. In blue: the ORFs coding for the micropeptides humanin, SHLP1-6 and MOTs-c found in the 16S and 12S rRNA genes. The

putative *gau* gene is also shown. In pink: the eight chosen candidates for antibody production. In red: MTALTND4.



**Figure 2. Identification of an alternative protein in the mtDNA-encoded *nd4* gene**

(A) The MTALTND4 ORF and its DNA and protein sequences (antigenic sequence underlined).

This ORF is frameshifted by 2 nucleotides relative to ND4 reading frame.

(B) Multiple peptide sequence alignment in five primate species. Black background for residues conserved in all species and grey for residues conserved in 4 out of 5 species.

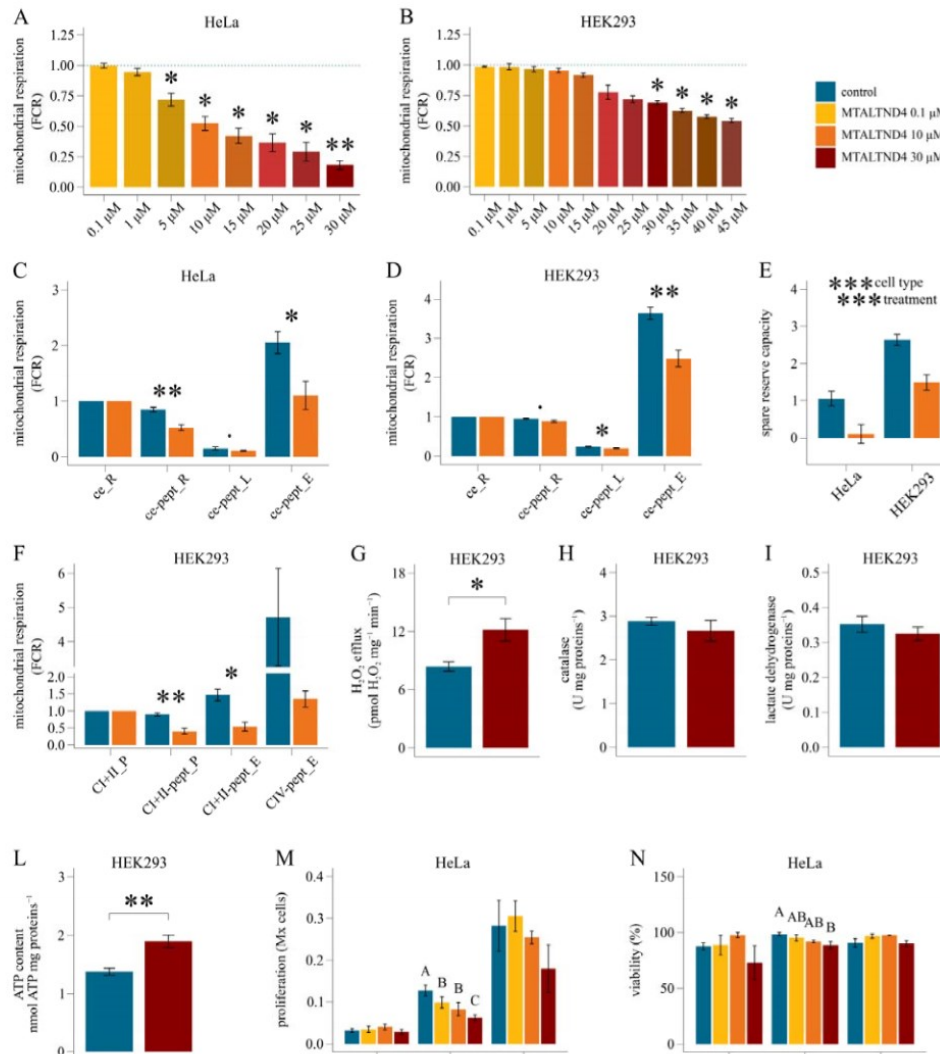
(C) MTALTND4 detected by western blotting in HeLa and HEK-293T cells.

(D) MTALTND4, mtDNA-encoded cytochrome c oxidase subunit I (CO1), nuclear-encoded Actin and nuclear-encoded mitochondrial ATP5 in HeLa (control) and HeLa Rho0 cells, which are devoid of mitochondrial DNA.

(E) MTALTND4, mtDNA-encoded cytochrome c oxidase subunit I (CO1), nuclear-encoded Actin and nuclear-encoded mitochondrial ATP5 in HeLa (control) and HeLa cells treated with chloramphenicol, which inhibits mitochondrial translation.

(F) Detection of MTALTND4 and nuclear-encoded mitochondrial ATP5 by immunofluorescence in HeLa cells.

(G) Detection of MTALTND4 and nuclear-encoded mitochondrial ATP5 by western blotting in human plasma.



**Figure 3. Impact of MTALTND4 upon mitochondrial and cell physiology.**

(A-B) Dose-dependent effect on intact (ce) HeLa and HEK-293T cells routine respiration ( $n=3$  and  $n=4$ , respectively). Respirometry data were normalized for the routine respiration in absence

of the peptide. For each titration point, data are represented as fraction of their own simultaneous control (blue dotted line).

(C-D) Mitochondrial respiration in intact (ce) HeLa and HEK-293T cells in presence of 10  $\mu\text{M}$  MTALTND4 ( $n=6-9$  and  $n=6$ , respectively). Respirometry data were expressed as flux control ratios (FCR), normalized for the routine respiration before treatment ('*ce\_R*'). '*ce-pept\_R*': routine respiration with MTALTND4; '*ce-pept\_L*': leak respiration with MTALTND4; '*ce-pept\_E*': maximal uncoupled respiration with MTALTND4.

(E) Mitochondrial spare reserve capacity (SRC) in HeLa and HEK-293T cells in presence of 10  $\mu\text{M}$  MTALTND4 ( $n=6$  and  $n=6$ , respectively).

(F) Mitochondrial respiration in permeabilized (pce) HEK-293T cells in presence of 10  $\mu\text{M}$  MTALTND4 ( $n=6$ ). Respirometry data were normalized for the CI+II-sustained coupled respiration before treatment ('*CI+II\_P*'). '*CI+II-pept\_P*': CI+II-linked coupled respiration with MTALTND4; '*CI+II-pept\_E*': CI+II-linked uncoupled respiration with MTALTND4; '*CIV-pept\_E*': Cytochrome *c* oxidase standalone capacity with MTALTND4.

(G) Hydrogen peroxide efflux rate ( $\text{pmol H}_2\text{O}_2 \cdot \text{mg proteins}^{-1} \cdot \text{minute}^{-1}$ ) in intact HEK-293T cells in presence of 30  $\mu\text{M}$  MTALTND4 ( $n=5$ ).

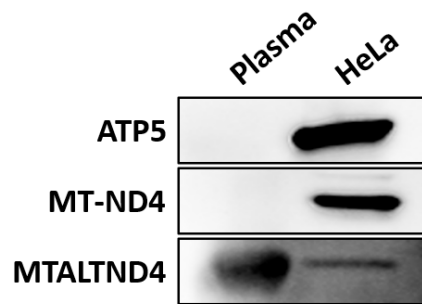
(H) Catalase activity ( $\text{U} \cdot \text{mg proteins}^{-1}$ ) in intact HEK-293T cells incubated 4 hours with 30  $\mu\text{M}$  MTALTND4 ( $n=5$ ).

(I) Lactate dehydrogenase activity ( $\text{U} \cdot \text{mg proteins}^{-1}$ ) in intact HEK-293T cells incubated 4 hours with 30  $\mu\text{M}$  MTALTND4 ( $n=5$ ).

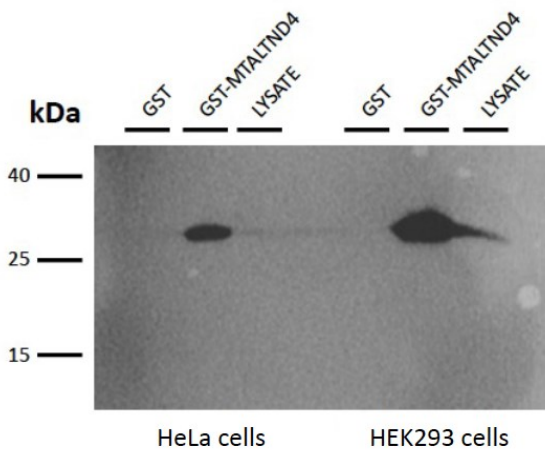
(L) ATP content ( $\text{nmol ATP} \cdot \text{mg proteins}^{-1}$ ) in intact HEK-293T cells incubated 4 hours with 30  $\mu\text{M}$  MTALTND4 ( $n=5$ ).

(*M-N*) Proliferation (Mx cells) and viability (%) of HeLa cells measured at different time points (24, 48 and 72 hours) and different MTALTND4 concentrations (0, 0.1, 10 and 30  $\mu$ M) ( $n=3-5$ ). Statistical analyses: (*A-B*) one sample *t* test; (*C, D, F, G, H, I, L*) paired *t* test; (*E*) linear mixed model. Factors '*cell type*' (2 levels) and '*treatment*' (2 levels) plus interaction. (*M, N*) linear mixed model. Factor '*treatment*' (4 levels), with letters indicating statistical difference following a *post hoc* multi comparison test. Data shown as mean  $\pm$  sem.  $0.05 > p \leq 0.09$ ;  $*p \leq 0.05$ ;  $**p \leq 0.01$ ;  $***p \leq 0.001$ . A detailed summary is reported in supplementary tables S6-S10.

Annexe 1.2 Figures qui seront incorporées à la version finale de l'article de Kienzle *et al.* (2022)



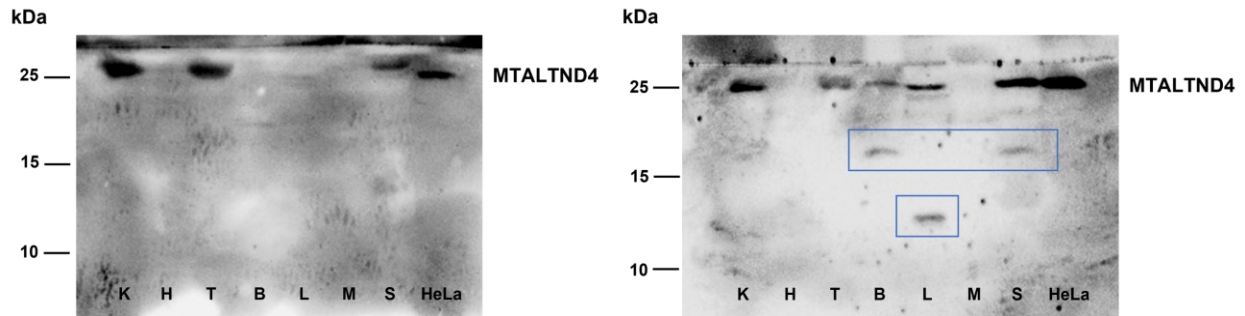
Kienzle *et al.*, 2022: Figure S8. Detection of MTALTND4, mitochondrial-encoded ND4 and nuclear-encoded mitochondrial ATP5 by western blotting in human plasma.



Kienzle *et al.*, 2022: Figure S9. GST pull-down assay indicating that MTALTND4 and C1QBP interact. Western blot probed with anti-C1QBP antibodies. Lane 1: pull-down products from GST bound glutathione beads (HeLa cells). Lane 2: from GST-MTALTND4 bound glutathione beads (HeLa cells). Lane 3: HeLa cells lysate. Lane 4: pull-down products from GST bound glutathione beads (HEK-293T cells). Lane 5: from GST-MTALTND4 bound glutathione beads (HEK-293T cells). Lane 6: HEK-293T cells lysate.

## Annexe 2: Supplementary Material

### Annexe 2.1 Supplementary Figures



**Figure S1, related to Figure A1** . Two immunoblots of the same samples showing MTALTND4 expression in various human tissues (K: kidney, H: heart, T: testis, B: brain, L: liver, M: skeletal muscle, S: spleen) and HeLa cells with presence of secondary bands (blue squares).



## **Annexe 2.2 Supplementary Tables**

**Table S1.** Upregulated and downregulated genes between control and treated HeLa cells identified from microarray analysis.

- Voir les fichiers complémentaires

**Table S2.** Upregulated and downregulated genes between control and treated HEK-293T cells identified from microarray analysis.

- Voir les fichiers complémentaires

**Table S3.** Common **(A)** upregulated and **(B)** downregulated genes between HeLa and HEK-293T cells identified from microarray analysis.

A. COMMON UPREGULATED GENES						
ID	Description	Gene Symbol	HeLa		HEK	
			Fold change	P-value	Fold change	P-value
TC1800008235.hg.1	Transcript Identified by aceview, Entrez Gene ID(s) 171586	ABHD3	1.88	9.10E-03	1.8	1.32E-02
TC0500011296.hg.1	Acyl-coa thioesterase 12	ACOT12	1.9	1.86E-02	2.47	2.70E-03
TC2000008950.hg.1	Memczak2013 ANTISENSE, CDS, coding, INTERNAL best transcript NM_001076552	ACSS2	1.92	1.25E-02	1.74	2.74E-02
TC2000007202.hg.1	Acyl-coa synthetase short-chain family member 2	ACSS2	1.79	9.20E-03	1.64	2.15E-02
TC1100006831.hg.1	Adrenomedullin	ADM	2.53	2.00E-04	1.91	3.20E-03
TC1100011083.hg.1	AHNAK nucleoprotein	AHNAK	1.56	4.21E-02	1.79	1.15E-02
TC0200015424.hg.1	Amyotrophic lateral sclerosis 2 chromosome region candidate 12	ALS2CR12	1.99	1.60E-03	1.55	2.19E-02
TC0100015350.hg.1	Adenosine monophosphate deaminase 1	AMPD1	1.93	2.40E-03	4.44	2.83E-06
TC1300008084.hg.1	Atpase, class VI, type 11A	ATP11A	1.84	1.80E-03	1.79	3.80E-02
TC1400010759.hg.1	Atpase, H+ transporting, lysosomal 34kda, V1 subunit D	ATP6V1D	1.81	1.58E-02	2.36	1.80E-03
TC0300013025.hg.1	Butyrylcholinesterase	BCHE	2	3.00E-03	2.11	1.80E-03
TC2000010009.hg.1	Bladder cancer associated protein	BLCAP	1.56	3.14E-02	2.08	2.00E-03
TC1500007304.hg.1	Chromosome 15 open reading frame 65	C15orf65	1.65	2.52E-02	3.31	7.37E-05
TC0500013298.hg.1	C1q and tumor necrosis factor related protein 3	C1QTNF3	2.25	1.50E-03	3.84	2.53E-05
TC0600013144.hg.1	Transcript Identified by aceview, Entrez Gene ID(s) 93663; 253582; Jeck2013 ALT_ACCEPTOR, ALT_DONOR, downstream_end, ncna, OVEXON, upstream_start best transcript TCONS_I2_00025470; Jeck2013 ALT_ACCEPTOR, ncna, OVEXON, upstream_start best transcript TCONS_I2_00024911; Salzman2013 ANTISENSE, CDS, coding, downstream_end, intronic, OVCODE, OVEXON, UTR3 best transcript NM_000426; Salzman2013 ANNOTATED, INTERNAL, ncna, OVEXON best transcript TCONS_I2_00025470; putative novel transcript; novel transcript	C6orf191andARHGAP18; RP1-69D17.4; RP1-69D17.3; TCONS_I2_00025470; TCONS_I2_00024911; LAMA2	1.55	2.38E-02	2.35	4.00E-04

TC1600011351.hg.1	Jeck2013 ANTISENSE, CDS, coding, INTERNAL, OVCODE, OVEXON best transcript NM_001042476	CARHSP1	1.94	2.90E-03	1.51	3.69E-02
TC1400010776.hg.1	Catsper channel auxiliary subunit beta	CATSPERB	1.74	3.80E-03	2.93	2.20E-05
TC1100010472.hg.1	Coiled-coil domain containing 73	CCDC73	2.83	2.00E-04	1.52	4.89E-02
TC1100008665.hg.1	Coiled-coil domain containing 81	CCDC81	2.68	9.36E-05	2.81	6.13E-05
TC0400012818.hg.1	Cyclin G2	CCNG2	1.68	1.30E-02	2.51	3.00E-04
TC0400011043.hg.1	Cyclin-dependent kinase-like 2 (CDC2-related kinase)	CDKL2	2.77	6.84E-05	1.74	6.10E-03
TC1700012361.hg.1	CMT1A duplicated region transcript 1	CDRT1	1.78	2.00E-02	2.66	8.00E-04
TC1600007985.hg.1	Cholesteryl ester transfer protein, plasma	CETP	2.19	7.00E-04	3.45	1.64E-05
TC0600014153.hg.1	Cilia and flagella associated protein 206	CFAP206	2.24	1.20E-03	1.83	7.50E-03
TC1200008686.hg.1	Transcript Identified by aceview, Entrez Gene ID(s) 50515	CHST11	1.68	3.84E-02	1.58	2.06E-02
TC0200015397.hg.1	CDC like kinase 1	CLK1	1.79	1.92E-02	2.49	1.40E-03
TC0500013020.hg.1	CDC like kinase 4	CLK4	1.92	2.80E-03	1.51	3.42E-02
TC1400006828.hg.1	Cochlin	COCH	1.75	4.96E-02	1.9	2.81E-02
TC1600008889.hg.1	Copine VII	CPNE7	1.83	3.95E-02	2.89	1.90E-03
TC0300010770.hg.1	Cysteine-serine-rich nuclear protein 1	CSRNP1	3.44	2.58E-05	1.63	1.81E-02
TC0800008481.hg.1	Collagen triple helix repeat containing 1	CTHRC1	1.76	2.32E-02	2.5	1.30E-03
TC0200014414.hg.1	Chemokine (C-X-C motif) receptor 4	CXCR4	2.42	3.00E-04	1.96	2.30E-03
TC1000012480.hg.1	Cytochrome P450, family 2, subfamily C, polypeptide 18	CYP2C18	1.65	2.09E-02	3.9	1.63E-05
TC1000011478.hg.1	Cytochrome P450, family 2, subfamily C, polypeptide 8	CYP2C8	1.65	3.27E-02	1.9	9.70E-03
TC0400007495.hg.1	DCN1, defective in cullin neddylation 1, domain containing 4	DCUN1D4	1.57	4.16E-02	3.57	4.89E-05
TC1200010968.hg.1	DNA-damage-inducible transcript 3	DDIT3	1.78	1.53E-02	2.89	3.00E-04
TC0800009094.hg.1	DENN/MADD domain containing 3	DENND3	2.29	6.00E-04	1.8	9.50E-03
TSUnmapped00000088.hg.1	Dual specificity phosphatase 16	DUSP16	1.54	4.25E-02	2.11	2.20E-03
TC1200009967.hg.1	Dual specificity phosphatase 16	DUSP16	1.63	2.44E-02	1.87	6.60E-03
TC1900008456.hg.1	Epididymal sperm binding protein 1	ELSPBP1	1.83	2.90E-03	2.34	3.00E-04
TC0800011621.hg.1	Ectonucleotide pyrophosphatase/phosphodiesterase 2	ENPP2	1.56	4.88E-02	1.56	4.80E-02
TC2100008143.hg.1	V-ets avian erythroblastosis virus E26 oncogene homolog	ERG	1.82	3.62E-02	1.91	2.57E-02
TC0100018403.hg.1	ERBB receptor feedback inhibitor 1	ERRFI1	1.65	1.36E-02	1.61	1.72E-02

TC0700011318.hg.1	Endogenous retrovirus group 3, member 1; zinc finger protein 117	ERV3-1; ZNF117	1.67	1.35E-02	1.73	9.60E-03
TC2000009945.hg.1	Family with sequence similarity 209, member A	FAM209A	1.87	9.50E-03	2.12	3.20E-03
TC0700009060.hg.1	Family with sequence similarity 71, member F1	FAM71F1	1.63	4.05E-02	2.36	1.80E-03
TC0300008532.hg.1	F-box protein 40	FBXO40	1.63	9.80E-03	1.81	3.10E-03
TC1500010853.hg.1	Fibrous sheath interacting protein 1	FSIP1	1.69	4.69E-02	1.78	3.11E-02
TC1900006588.hg.1	Growth arrest and DNA-damage-inducible, beta	GADD45B	2.19	8.00E-04	2.06	1.40E-03
TC0600012064.hg.1	Glial cells missing homolog 1 (Drosophila)	GCM1	1.93	7.60E-03	2.16	2.80E-03
TC0600012502.hg.1	Gap junction protein beta 7	GJB7	1.54	1.60E-02	1.52	1.95E-02
TC0500007465.hg.1	GC-rich promoter binding protein 1	GPBP1	1.69	9.40E-03	3.26	2.30E-05
TC1100012000.hg.1	G protein-coupled receptor 83	GPR83	2.05	1.90E-03	2.78	1.00E-04
TC1200009980.hg.1	G protein-coupled receptor, class C, group 5, member D	GPRC5D	2.85	1.00E-04	2.39	4.00E-04
TC0600013341.hg.1	GVQW motif containing 2	GVQW2	1.55	2.78E-02	1.64	1.57E-02
TC0100009608.hg.1	Hydroxyacid oxidase 2 (long chain)	HAO2	1.72	3.94E-02	2.64	1.60E-03
TC0700008747.hg.1	HMG-box transcription factor 1	HBP1	1.75	3.80E-03	2.52	9.22E-05
TC1200012248.hg.1	Hydroxycarboxylic acid receptor 3	HCAR3	2.03	7.50E-03	1.77	2.23E-02
TC0300012166.hg.1	Hematopoietic cell-specific Lyn substrate 1	HCLS1	1.75	2.47E-02	1.65	3.96E-02
TC1700008033.hg.1	Hexamethylene bis-acetamide inducible 1	HEXIM1	2.01	1.30E-03	2.67	9.18E-05
TC0600011125.hg.1	Histone cluster 1, h2ab	HIST1H2AB	1.76	3.40E-03	4.92	5.81E-07
TC0600011133.hg.1	Memczak2013 ANTISENSE, CDS, coding, upstream_start, UTR3, UTR5 best transcript NM_003523	HIST1H2BE	3.95	1.30E-05	1.87	5.50E-03
TC0100015698.hg.1	Histone cluster 2, h3d	HIST2H3D	1.6	2.48E-02	2.03	2.60E-03
TSUnmapped00000141.hg.1	Hydroxymethylbilane synthase	HMBS	2.28	1.10E-03	2.08	2.50E-03
TC0100008503.hg.1	Hook microtubule-tethering protein 1	HOOK1	2.15	1.03E-02	2.31	6.20E-03
TC2200007043.hg.1	HORMA domain containing 2	HORMAD2	1.72	1.73E-02	2.59	5.00E-04
TC1900006977.hg.1	Intercellular adhesion molecule 1	ICAM1	2.33	2.00E-03	2.13	4.20E-03
TC0100008653.hg.1	Interleukin 12 receptor, beta 2	IL12RB2	1.6	1.95E-02	3.56	1.52E-05
TSUnmapped00000135.hg.1	Inositol polyphosphate-5-phosphatase D	INPP5D	1.75	7.00E-03	1.77	2.20E-03
TC2100006784.hg.1	Junctional adhesion molecule 2	JAM2	1.63	3.04E-02	1.68	2.33E-02
TSUnmapped00000492.hg.1	K(lysine) acetyltransferase 6B	KAT6B	1.68	5.30E-03	1.61	9.00E-03
TSUnmapped00000282.hg.1	K(lysine) acetyltransferase 6B	KAT6B	1.66	1.64E-02	-1.65	1.70E-02
TC0300009673.hg.1	Kelch-like family member 24	KLHL24	1.8	6.10E-03	1.98	2.40E-03

TC1600011368.hg.1	Linker for activation of T-cells	LAT	1.56	4.35E-02	1.58	4.05E-02
TC1100007005.hg.1	Lactate dehydrogenase C	LDHC	1.84	1.52E-02	2.4	1.80E-03
TC1400010748.hg.1	Long intergenic non-protein coding RNA 1588	LINC01588	1.68	4.04E-02	2.51	1.80E-03
TC1200012778.hg.1	Lamin tail domain containing 1	LMNTD1	1.65	2.09E-02	1.63	2.45E-02
TC1200009198.hg.1	Leucine rich repeat containing 43	LRRC43	2.2	1.40E-03	2.53	4.00E-04
TC0700013442.hg.1	Leucine-rich single-pass membrane protein 1	LSMEM1	2.59	3.00E-04	7.65	3.28E-07
TC0100014769.hg.1	Mucolipin 3	MCOLN3	1.98	1.10E-03	1.85	2.30E-03
TC1700012310.hg.1	Major facilitator superfamily domain containing 11	MFSD11	1.79	2.22E-02	1.69	3.56E-02
TC1700008598.hg.1	Mast cell immunoglobulin-like receptor 1	MILR1	1.68	2.81E-02	1.81	1.48E-02
TC0200016770.hg.1	Microrna 6809; tensin 1	MIR6809; TNS1	1.87	4.75E-02	2.11	2.21E-02
TC1000006912.hg.1	Mannose receptor, C type 1	MRC1	2.01	3.50E-03	1.89	6.40E-03
TC1100010998.hg.1	Membrane-spanning 4-domains, subfamily A, member 4E	MS4A4E	1.61	3.06E-02	2.07	3.10E-03
TC0200012977.hg.1	Memczak2013 ANTISENSE, coding, INTERNAL, intronic best transcript NM_001202514	MXD1	3.43	1.45E-05	2.19	6.00E-04
TC0200009927.hg.1	Myosin IIIB	MYO3B	1.64	4.90E-02	2.7	1.00E-03
TC0200009927.hg.1	Myosin IIIB	MYO3B	1.64	4.90E-02	2.7	1.00E-03
TC1400008940.hg.1	Nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor, alpha	NFKBIA	5	2.06E-06	3.06	5.70E-05
TC0300013855.hg.1	Nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor, zeta	NFKBIZ	2.26	7.00E-04	1.74	8.80E-03
TC0200014672.hg.1	Nuclear receptor subfamily 4, group A, member 2	NR4A2	7.5	1.83E-06	1.86	1.54E-02
TC0900008219.hg.1	Nuclear receptor subfamily 4, group A, member 3	NR4A3	2.22	4.00E-04	1.94	1.40E-03
TC0900011199.hg.1	Olfactory receptor, family 2, subfamily K, member 2	OR2K2	1.79	1.23E-02	2.93	2.00E-04
TC1100007637.hg.1	Olfactory receptor, family 8, subfamily U, member 1; olfactory receptor, family 8, subfamily U, member 8	OR8U1; OR8U8	1.72	1.13E-02	1.98	3.00E-03
TC1100009833.hg.1	Memczak2013 ALT_ACCEPTOR, ALT_DONOR, coding, INTERNAL, intronic best transcript NM_020896	OSBPL5	1.69	7.50E-03	1.97	1.40E-03
TC1100011310.hg.1	Pyruvate carboxylase	PC	2.07	1.30E-03	1.75	3.14E-02
TC1600007007.hg.1	Pyridoxal-dependent decarboxylase domain containing 1	PDXDC1	1.55	4.25E-02	1.89	6.40E-03
TC0500007641.hg.1	Phosphoinositide-3-kinase, regulatory subunit 1 (alpha)	PIK3R1	1.99	2.10E-02	2.27	8.60E-03
TC0100014040.hg.1	Phosphoinositide-3-kinase, regulatory subunit 3 (gamma)	PIK3R3	1.65	1.52E-02	2.67	2.00E-04
TC0100018553.hg.1	Pleckstrin homology domain containing, family A member 6	PLEKHA6	2.18	3.50E-03	1.76	4.73E-02

TC0600008757.hg.1	Proline-rich nuclear receptor coactivator 1	PNRC1	2.15	2.00E-03	1.79	1.07E-02
TC1200012534.hg.1	Transcript Identified by aceview, Entrez Gene ID(s) 5426	POLE	1.77	1.95E-02	1.76	2.04E-02
TC0600011372.hg.1	Protein phosphatase 1, regulatory subunit 10	PPP1R10	3.56	1.96E-05	1.62	1.86E-02
TC0200010165.hg.1	Protein phosphatase 1, regulatory (inhibitor) subunit 1C	PPP1R1C	1.62	1.17E-02	2.03	1.10E-03
TC1400007444.hg.1	Protein phosphatase 1, regulatory subunit 36	PPP1R36	1.59	3.51E-02	2.68	4.00E-04
TC1100011182.hg.1	Phosphorylase, glycogen, muscle	PYGM	1.81	5.50E-03	2.13	1.10E-03
TC0200015476.hg.1	Ras association (ralgds/AF-6) and pleckstrin homology domains 1	RAPH1	1.94	1.70E-03	2.07	9.00E-04
TC0700012099.hg.1	Salzman2013 ALT_ACCEPTOR, ALT_DONOR, coding, INTERNAL, intronic best transcript NM_006989	RASA4	1.7	1.08E-02	1.66	1.39E-02
TC0100013636.hg.1	Ring finger protein 19B	RNF19B	1.6	1.91E-02	2.36	4.00E-04
TSUnmapped00000289.hg.1	Ribosomal protein l7a	RPL7A	1.57	1.27E-02	1.58	1.17E-02
TC0700010181.hg.1	Radial spoke head 10 homolog B2 (Chlamydomonas); radial spoke head 10 homolog B (Chlamydomonas)	RSPH10B2; RSPH10B	1.64	3.38E-02	2.06	4.80E-03
TC0200016673.hg.1	Reticulon 4	RTN4	1.52	2.32E-02	1.73	5.80E-03
TC1800007201.hg.1	Sialic acid binding Ig-like lectin 15	SIGLEC15	1.68	9.40E-03	1.73	7.00E-03
TSUnmapped00000474.hg.1	Solute carrier family 16 (monocarboxylate transporter), member 1	SLC16A1	1.62	2.28E-02	2.61	3.00E-04
TC0900010910.hg.1	Solute carrier family 35 (UDP-glcnac/UDP-glucose transporter), member D2	SLC35D2	1.93	3.70E-03	6.14	7.60E-07
TC1200009547.hg.1	Solute carrier family 6 (neurotransmitter transporter), member 13	SLC6A13	1.56	4.15E-02	3.28	7.63E-05
TC2000007670.hg.1	Snail family zinc finger 1	SNAI1	1.58	2.76E-02	1.52	3.87E-02
TC0400008624.hg.1	Transcript Identified by aceview, Entrez Gene ID(s) 166378	SPATA5	1.87	7.80E-03	1.69	9.20E-03
TC0200016423.hg.1	Speedy/RINGO cell cycle regulator family member A	SPDYA	1.53	2.15E-02	3.4	1.07E-05
TC0100010905.hg.1	SWT1 RNA endoribonuclease homolog	SWT1	1.83	5.30E-03	2.18	1.00E-03
TC0100015445.hg.1	T-box 15	TBX15	1.63	3.70E-02	2.08	4.50E-03
TC0100011621.hg.1	Transforming growth factor beta 2; TGFB2 overlapping transcript 1	TGFB2; TGFB2-OT1	1.69	6.00E-03	1.59	1.18E-02
TSUnmapped00000172.hg.1	Transglutaminase 4	TGM4	1.56	3.40E-02	1.72	2.17E-02
TC0900010388.hg.1	Transcript Identified by aceview, Entrez Gene ID(s) 23670	TMEM2	2.48	1.00E-03	2.95	1.87E-05
TC1000010744.hg.1	Transmembrane protein 26	TMEM26	1.52	2.63E-02	1.5	2.98E-02

TC0400010895.hg.1	Transmembrane protease, serine 11D	TMPRSS11D	1.83	1.30E-02	2.76	4.00E-04
TC0600009597.hg.1	Tumor necrosis factor, alpha-induced protein 3	TNFAIP3	4.71	5.73E-07	1.62	7.50E-03
TC1800007508.hg.1	Transcript Identified by aceview, Entrez Gene ID(s) 8792	TNFRSF11A	2.03	1.40E-03	2.03	1.40E-03
TC1900006804.hg.1	Tumor necrosis factor (ligand) superfamily, member 9	TNFSF9	1.86	5.70E-03	1.93	3.90E-03
TC1100013150.hg.1	Tryptophan hydroxylase 1	TPH1	1.65	2.03E-02	2.05	2.60E-03
TC1700012333.hg.1	Transient receptor potential cation channel, subfamily V, member 3	TRPV3	1.84	1.70E-03	1.56	1.15E-02
TC0200013602.hg.1	Testis specific 10	TSGA10	1.84	6.00E-03	1.59	2.44E-02
TC1800006655.hg.1	Thioredoxin domain containing 2 (spermatzoa)	TXNDC2	2.43	2.10E-03	2.05	7.80E-03
TC1900009984.hg.1	Unc-13 homolog A (C. Elegans)	UNC13A	1.62	3.34E-02	1.62	3.34E-02
TC0200007722.hg.1	Memczak2013 ANTISENSE, CDS, coding, INTERNAL best transcript NM_014709	USP34	2.18	2.50E-03	2.29	5.00E-04
TC1100008480.hg.1	UV radiation resistance associated	UVRAG	1.75	6.20E-03	1.92	4.19E-02
TC1900012055.hg.1	Vomeronal 1 receptor 1	VN1R1	1.75	8.00E-03	1.75	7.90E-03
TC0200016676.hg.1	WD repeat containing planar cell polarity effector	WDPCP	1.52	4.21E-02	1.99	3.00E-03
TC1000006466.hg.1	WD repeat domain 37	WDR37	1.54	3.79E-02	1.74	1.20E-02
TC0300014077.hg.1	WD repeat domain 49	WDR49	1.8	4.50E-03	1.76	5.90E-03
TC1200012719.hg.1	WD repeat domain 66	WDR66	1.73	4.74E-02	2.38	4.70E-03
TC0200007132.hg.1	Yippee like 5	YPEL5	2.04	4.40E-03	2.91	3.00E-04
TC0500013282.hg.1	Zinc finger, DHHC-type containing 11	ZDHHC11	1.83	3.27E-02	1.78	3.61E-02
TC1900012020.hg.1	Zinc finger protein 432	ZNF432	1.71	1.65E-02	1.58	3.41E-02
TC0500009673.hg.1	Zinc finger protein 454	ZNF454	2.16	1.20E-03	1.98	2.90E-03
TC1900011920.hg.1	Zinc finger protein 98	ZNF98	1.55	3.04E-02	1.53	3.33E-02

**B. COMMON DOWNREGULATED GENES**

ID	Description	Gene Symbol	HeLa		HEK	
			Fold change	P-value	Fold change	P-value
TC0700013065.hg.1	ATP binding cassette subfamily F member 2	ABCF2	-1.52	4.80E-02	-1.68	1.83E-02
TC1400007647.hg.1	Acyl-coa thioesterase 2	ACOT2	-1.57	4.44E-02	-1.89	2.80E-03
TC0200010577.hg.1	ADAM metalloproteinase domain 23	ADAM23	-2.01	5.40E-03	-1.98	6.00E-03
TC1400010715.hg.1	Ajuba LIM protein	AJUBA	-1.64	2.24E-02	-1.51	4.87E-02
TC1700007189.hg.1	Aldehyde dehydrogenase 3 family, member A2	ALDH3A2	-1.5	3.18E-02	-1.7	8.20E-03
TC1100007384.hg.1	Alkb homolog 3, alpha-ketoglutarate-dependent dioxygenase; SEC14-like 1 pseudogene 1	ALKBH3; SEC14L1P1	-1.61	2.96E-02	-1.66	2.21E-02
TC1000011050.hg.1	Adaptor-related protein complex 3, mu 1 subunit	AP3M1	-1.65	1.76E-02	-1.73	1.11E-02
TC0900009948.hg.1	Rho guanine nucleotide exchange factor 39	ARHGEF39	-2.04	1.70E-03	-3.03	5.25E-05
TC0200016115.hg.1	ADP-ribosylation factor like gtpase 4C	ARL4C	-1.69	1.65E-02	-1.9	5.20E-03
TC1200008213.hg.1	Ataxin 7-like 3B	ATXN7L3B	-1.62	2.40E-02	-3.23	5.88E-05
TC1200011320.hg.1	Bardet-Biedl syndrome 10	BBS10	-1.78	1.21E-02	-2.2	1.80E-03
TC0200010805.hg.1	BCS1 homolog, ubiquinol-cytochrome c reductase complex chaperone	BCS1L	-1.78	7.50E-03	-2.12	1.40E-03
TC0X00010409.hg.1	Brain expressed X-linked 2	BEX2	-1.68	2.47E-02	-1.79	1.43E-02
TC0700009996.hg.1	Chromosome 7 open reading frame 50	C7orf50	-1.59	2.14E-02	-1.55	2.81E-02
TC0900010586.hg.1	Chromosome 9 open reading frame 64	C9orf64	-2.67	2.50E-03	-2.8	1.80E-03
TC2100007032.hg.1	Carbonyl reductase 1	CBR1	-1.69	1.01E-02	-1.79	5.40E-03
TC0300011054.hg.1	Coiled-coil domain containing 71	CCDC71	-1.76	5.30E-03	-3.83	6.13E-06
TSUnmapped00000313.hg.1	Coiled-coil domain containing 84	CCDC84	-2.1	5.40E-03	-2.04	6.90E-03
TSUnmapped00000085.hg.1	Coiled-coil domain containing 84	CCDC84	-2.23	7.00E-04	-1.5	3.68E-02
TC1200006555.hg.1	Cyclin D2	CCND2	-1.71	9.60E-03	-1.99	2.10E-03
TC0600011809.hg.1	Cyclin D3	CCND3	-1.57	1.65E-02	-1.93	1.70E-03
TC0500012163.hg.1	Cell division cycle 23	CDC23	-2.12	1.85E-02	-1.92	3.59E-02
TC0700010435.hg.1	Cell division cycle associated 7-like	CDCA7L	-1.63	1.20E-02	-1.63	1.20E-02
TC0500012070.hg.1	CDKN2A interacting protein N-terminal like	CDKN2AIPNL	-1.66	6.90E-03	-1.71	4.90E-03
TC0300013719.hg.1	Centrosomal protein 19kda	CEP19	-1.79	4.50E-03	-1.67	9.60E-03
TC0100007447.hg.1	Centrosomal protein 85kda	CEP85	-1.52	2.97E-02	-2.83	7.82E-05
TC1200010653.hg.1	Ceramide synthase 5	CERS5	-2.09	1.40E-03	-1.95	2.70E-03



TC0300010357.hg.1	Coiled-coil-helix-coiled-coil-helix domain containing 4	CHCHD4	-1.56	2.37E-02	-1.91	3.00E-03
TC1900011635.hg.1	Charged multivesicular body protein 2A	CHMP2A	-1.53	2.03E-02	-3.01	2.45E-05
TC1700010747.hg.1	Cytochrome c oxidase assembly factor 3	COA3	-1.77	7.30E-03	-1.72	9.60E-03
TC1800007850.hg.1	Collectin sub-family member 12	COLEC12	-1.53	4.48E-02	-1.77	1.15E-02
TC1200007759.hg.1	Coatmer protein complex subunit zeta 1	COPZ1	-1.6	9.40E-03	-1.68	5.30E-03
TC1100011829.hg.1	CREB/ATF bzip transcription factor	CREBZF	-1.61	1.13E-02	-2.35	2.00E-04
TC0X00007923.hg.1	Cleavage stimulation factor, 3 pre-RNA, subunit 2	CSTF2	-1.67	8.80E-03	-1.93	1.90E-03
TC0100018442.hg.1	Cytochrome b5 reductase-like	CYB5RL	-1.71	1.23E-02	-1.58	2.68E-02
TC0100010661.hg.1	Aspartyl-trna synthetase 2, mitochondrial	DARS2	-1.6	1.02E-02	-1.88	1.70E-03
TC1200012629.hg.1	DAZ associated protein 2	DAZAP2	-1.68	6.90E-03	-1.81	3.10E-03
TC1600007234.hg.1	Dynactin 5 (p25)	DCTN5	-1.62	2.15E-02	-1.59	2.69E-02
TC1700012089.hg.1	Dicarbonyl/L-xylulose reductase	DCXR	-1.55	3.81E-02	-2.08	2.40E-03
TC0100006627.hg.1	DNA fragmentation factor, 40kda, beta polypeptide (caspase-activated dnase)	DFFB	-1.5	3.56E-02	-1.67	1.16E-02
TC1700010221.hg.1	Dehydrogenase/reductase (SDR family) member 13	DHRS13	-1.82	4.10E-02	-1.95	2.56E-02
TC0200012299.hg.1	DEAH (Asp-Glu-Ala-Asp/His) box polypeptide 57	DHX57	-1.65	3.44E-02	-1.98	7.20E-03
TC1100009043.hg.1	DIX domain containing 1	DIXDC1	-1.8	2.51E-02	-1.89	1.73E-02
TC1700010701.hg.1	Dnaj (Hsp40) homolog, subfamily C, member 7	DNAJC7	-1.74	2.48E-02	-3.29	2.00E-04
TC1700010700.hg.1	Zhang2013 ALT_ACCEPTOR, ALT_DONOR, coding, INTERNAL, intronic, OVERLAPTX, OVEXON best transcript NM_003315	DNAJC7	-2.08	2.40E-03	-2.39	7.00E-04
TC0200015779.hg.1	Aspartyl aminopeptidase	DNPEP	-1.51	3.28E-02	-1.9	3.10E-03
TC2000007502.hg.1	Deoxynucleotidyltransferase, terminal, interacting protein 1	DNTTIP1	-1.51	3.04E-02	-1.62	1.44E-02
TC1100009419.hg.1	Etoposide induced 2.4	EI24	-1.75	4.10E-03	-2.1	6.00E-04
TC1500006930.hg.1	Eukaryotic translation initiation factor 2 alpha kinase 4	EIF2AK4	-1.56	2.14E-02	-2.39	3.00E-04
TC1200010038.hg.1	Epidermal growth factor receptor pathway substrate 8	EPS8	-1.67	1.02E-02	-1.57	1.88E-02
TC0400009132.hg.1	Electron-transferring-flavoprotein dehydrogenase	ETFDH	-1.55	4.25E-02	-2.03	3.50E-03
TC0500009803.hg.1	Memczak2013 ANTISENSE, CDS, coding, INTERNAL best transcript NM_007277	EXOC3	-1.66	2.16E-02	-1.71	4.46E-02
TC0900008967.hg.1	Exosome component 2	EXOSC2	-1.66	4.90E-03	-1.57	9.70E-03
TC1600006493.hg.1	Family with sequence similarity 173, member A	FAM173A	-1.54	3.84E-02	-1.73	1.20E-02
TC0X00008833.hg.1	Family with sequence similarity 50, member A	FAM50A	-1.65	4.26E-02	-1.89	1.37E-02

TC0600009102.hg.1	FIG4 phosphoinositide 5-phosphatase	FIG4	-2.15	1.90E-03	-1.66	2.06E-02
TC0700010046.hg.1	Ftsj RNA methyltransferase homolog 2 (E. Coli)	FTSJ2	-1.88	9.40E-03	-2.19	2.60E-03
TC0800007985.hg.1	Ganglioside induced differentiation associated protein 1	GDAP1	-1.66	3.79E-02	-1.71	2.98E-02
TC0500012568.hg.1	Gem nuclear organelle associated protein 5	GEMIN5	-2.1	6.00E-04	-1.53	2.05E-02
TC2000007016.hg.1	GINS complex subunit 1 (Psf1 homolog)	GINS1	-1.9	1.10E-03	-1.57	1.06E-02
TC1400010632.hg.1	G-patch domain containing 2 like	GPATCH2L	-1.54	1.45E-02	-1.88	1.50E-03
TC0100012788.hg.1	G protein-coupled receptor 157	GPR157	-2.09	2.90E-03	-2.33	1.10E-03
TC1600007723.hg.1	Glutamic pyruvate transaminase (alanine aminotransferase) 2	GPT2	-1.61	3.20E-02	-1.9	6.80E-03
TC1700006592.hg.1	Germ cell associated 2 (haspin)	GSG2	-1.62	1.69E-02	-2.53	2.00E-04
TC0100013998.hg.1	HECT domain containing E3 ubiquitin protein ligase 3	HECTD3	-1.61	1.38E-02	-2.07	1.00E-03
TC0100009877.hg.1	Histone cluster 2, h3a	HIST2H3A	-3.84	9.82E-06	-2.76	1.00E-04
TC0100015701.hg.1	Histone cluster 2, h3a; histone cluster 2, h3c	HIST2H3A; HIST2H3C	-4.41	9.01E-06	-3.01	1.00E-04
TC0100015707.hg.1	Histone cluster 2, h4a; histone cluster 2, h4b	HIST2H4A; HIST2H4B	-3.38	7.44E-05	-2.77	3.00E-04
TC0100009870.hg.1	Histone cluster 2, h4b; histone cluster 2, h4a	HIST2H4B; HIST2H4A	-3.27	2.28E-05	-2.92	5.53E-05
TC0700010563.hg.1	Homeobox A7	HOXA7	-2.44	2.00E-04	-1.65	1.08E-02
TC1600011047.hg.1	Hydroxysteroid dehydrogenase like 1	HSDL1	-1.64	3.57E-02	-1.64	3.45E-02
TC1700008879.hg.1	Immature colon carcinoma transcript 1	ICT1	-1.67	2.08E-02	-1.75	1.35E-02
TC1300008713.hg.1	Kelch repeat and BTB (POZ) domain containing 6	KBTBD6	-2.01	8.00E-03	-3.26	2.00E-04
TC0600007836.hg.1	Potassium channel tetramerization domain containing 20	KCTD20	-1.71	4.20E-03	-2.08	5.00E-04
TC0100016476.hg.1	Kiaa0040	KIAA0040	-1.73	7.70E-03	-1.52	3.09E-02
TC1600009646.hg.1	Lysine-rich nucleolar protein 1	KNOP1	-1.7	1.16E-02	-1.53	3.47E-02
TC2100007362.hg.1	Leucine rich repeat containing 3	LRRC3	-1.51	4.66E-02	-1.76	1.07E-02
TC1900010011.hg.1	LSM4 homolog, U6 small nuclear RNA and mrna degradation associated	LSM4	-1.5	1.63E-02	-1.71	3.30E-03
TC0600011808.hg.1	Mediator complex subunit 20	MED20	-1.78	8.50E-03	-2.61	3.00E-04
TC1000011659.hg.1	Meningioma expressed antigen 5 (hyaluronidase)	MGEA5	-1.57	3.37E-02	-1.64	7.80E-03
TC0300006631.hg.1	Makorin ring finger protein 2	MKRN2	-1.68	6.70E-03	-1.8	3.20E-03
TC0100008124.hg.1	Methylmalonic aciduria (cobalamin deficiency) cblc type, with homocystinuria	MMACHC	-1.81	7.80E-03	-1.82	7.30E-03
TC1000011555.hg.1	MORN repeat containing 4	MORN4	-1.69	1.05E-02	-2.51	2.00E-04
TC1500007860.hg.1	Mannose phosphate isomerase	MPI	-1.96	1.30E-03	-2.02	1.00E-03

TC1900008036.hg.1	Mitochondrial ribosomal protein S12	MRPS12	-1.79	2.95E-02	-1.99	1.32E-02
TC1600009063.hg.1	Mitochondrial ribosomal protein S34	MRPS34	-1.66	9.20E-03	-1.68	7.80E-03
TC0500011255.hg.1	Metaxin 3	MTX3	-1.79	1.50E-02	-1.67	2.76E-02
TC0400012933.hg.1	N-acylethanolamine acid amidase	NAAA	-1.64	7.30E-03	-1.66	6.40E-03
TC0300013727.hg.1	Nuclear cap binding protein subunit 2	NCBP2	-1.62	2.73E-02	-1.62	2.68E-02
TC1000008961.hg.1	NHL repeat containing 2	NHLRC2	-1.97	6.00E-04	-1.52	1.38E-02
TC0500012937.hg.1	NOP16 nucleolar protein	NOP16	-1.88	3.10E-03	-2.24	5.00E-04
TC0600014283.hg.1	Nudix hydrolase 3	NUDT3	-1.59	1.17E-02	-1.74	4.30E-03
TC0X00009929.hg.1	Oligophrenin 1	OPHN1	-1.69	3.18E-02	-1.64	3.94E-02
TC0500010680.hg.1	Memczak2013 ANTISENSE, CDS, coding, INTERNAL, intronic best transcript NM_001178056	PARP8	-1.71	1.82E-02	-3.63	4.02E-05
TC1100009700.hg.1	Parkinson disease 7 domain containing 1	PDDC1	-1.72	2.33E-02	-1.85	1.25E-02
TC1000008574.hg.1	Phosphoglycerate mutase 1 (brain)	PGAM1	-1.51	3.00E-02	-2.07	1.10E-03
TC1700009623.hg.1	PHD finger protein 23	PHF23	-2.01	8.00E-04	-1.6	9.90E-03
TC1900006956.hg.1	Peptidylprolyl cis/trans isomerase, NIMA-interacting 1	PIN1	-1.56	4.91E-02	-2.59	7.00E-04
TC1100010216.hg.1	Pleckstrin homology domain containing, family A member 7	PLEKHA7	-1.72	4.70E-02	-3.23	6.00E-04
TC0100015931.hg.1	Phosphomevalonate kinase	PMVK	-1.6	2.80E-02	-2.21	1.40E-03
TC0600008095.hg.1	Polymerase (RNA) I polypeptide C	POLR1C	-1.84	4.40E-03	-1.6	1.89E-02
TC1000008193.hg.1	Peptidylprolyl isomerase F	PPIF	-1.51	2.89E-02	-2.97	4.02E-05
TC0900008902.hg.1	Protein phosphatase 2A activator, regulatory subunit 4	PPP2R4	-1.54	1.54E-02	-2.4	1.00E-04
TC0100008450.hg.1	Protein kinase, AMP-activated, alpha 2 catalytic subunit	PRKAA2	-1.65	2.98E-02	-1.73	1.95E-02
TC1100012717.hg.1	Pseudouridylate synthase 3	PUS3	-1.57	2.41E-02	-1.67	1.29E-02
TC0400012905.hg.1	Quinoid dihydropteridine reductase	QDPR	-1.64	1.74E-02	-2.17	1.10E-03
TC1500010756.hg.1	RAB11A, member RAS oncogene family	RAB11A	-1.57	2.73E-02	-1.63	1.79E-02
TC0800010138.hg.1	RAB11 family interacting protein 1 (class I)	RAB11FIP1	-1.5	4.02E-02	-1.81	6.10E-03
TC1400009422.hg.1	RAB15, member RAS oncogene family	RAB15	-1.62	1.94E-02	-1.75	9.00E-03
TC0X00008405.hg.1	RNA binding motif protein, X-linked 2	RBMX2	-1.55	1.89E-02	-1.76	4.90E-03
TC0200016618.hg.1	Ribonuclease H1	RNASEH1	-2.08	8.00E-04	-1.57	1.65E-02
TC1800009222.hg.1	Ring finger protein 125, E3 ubiquitin protein ligase	RNF125	-1.83	2.80E-02	-1.69	4.96E-02
TC1200010901.hg.1	Ring finger protein 41, E3 ubiquitin protein ligase	RNF41	-1.71	3.20E-03	-1.65	4.70E-03
TC0200008351.hg.1	Ribose 5-phosphate isomerase A	RPIA	-1.66	3.35E-02	-2.48	1.20E-03

TC0600014095.hg.1	Ribonuclease P/MRP 21kda subunit	RPP21	-1.93	1.40E-02	-1.85	1.95E-02
TC1700012287.hg.1	Ribosomal protein S6 kinase, 70kda, polypeptide 1	RPS6KB1	-1.57	2.86E-02	-1.54	3.36E-02
TC0600014130.hg.1	Ribosomal RNA processing 36	RRP36	-2.33	8.00E-04	-1.59	2.79E-02
TC1900009805.hg.1	Sterile alpha motif domain containing 1	SAMD1	-1.6	4.04E-02	-2.26	2.00E-03
TC0800012165.hg.1	Scribbled planar cell polarity protein; microrna 937	SCRIB; MIR937	-1.6	8.80E-03	-2.44	9.82E-05
TC1400008760.hg.1	Short chain dehydrogenase/reductase family 39U, member 1	SDR39U1	-1.59	1.18E-02	-1.79	3.00E-03
TC0100013760.hg.1	Splicing factor 3a subunit 3	SF3A3	-1.72	9.30E-03	-1.97	2.20E-03
TC1400009351.hg.1	SIX homeobox 1	SIX1	-1.54	2.15E-02	-1.74	5.40E-03
TC1700010826.hg.1	Solute carrier family 25, member 39	SLC25A39	-1.56	2.61E-02	-1.51	3.65E-02
TC0100007954.hg.1	Small arfgap2	SMAP2	-1.73	5.10E-03	-1.63	9.70E-03
TC1700012401.hg.1	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily e, member 1	SMARCE1	-1.83	4.30E-03	-1.57	2.11E-02
TC1700012286.hg.1	SMG8 nonsense mediated mrna decay factor	SMG8	-2.32	1.20E-03	-1.58	3.66E-02
TC0100018430.hg.1	Small integral membrane protein 12	SMIM12	-1.56	1.59E-02	-1.75	4.30E-03
TC0100011963.hg.1	Sprt-like N-terminal domain	SPRTN	-1.58	2.48E-02	-1.94	3.40E-03
TC1200007842.hg.1	SPRY domain containing 4	SPRYD4	-1.81	2.60E-03	-2.52	8.65E-05
TC1400006890.hg.1	Signal recognition particle 54kda	SRP54	-1.53	4.23E-02	-1.66	1.98E-02
TC1700012437.hg.1	Serine/arginine-rich splicing factor 1	SRSF1	-1.59	1.35E-02	-1.82	3.00E-03
TC0100013323.hg.1	Sperm-tail PG-rich repeat containing 1	STPG1	-1.83	1.80E-03	-1.61	8.10E-03
TC1600011428.hg.1	Telomeric repeat binding factor 2, interacting protein	TERF2IP	-1.59	8.60E-03	-1.96	8.00E-04
TC0300008477.hg.1	Translocase of inner mitochondrial membrane domain containing 1	TIMMDC1	-1.58	1.54E-02	-1.94	1.60E-03
TC1400010723.hg.1	Transmembrane 9 superfamily member 1	TM9SF1	-1.54	4.25E-02	-1.69	1.84E-02
TC0700010926.hg.1	Transmembrane p24 trafficking protein 4	TMED4	-1.76	1.27E-02	-1.7	1.76E-02
TC0200016516.hg.1	Transmembrane protein 177	TMEM177	-1.59	2.89E-02	-1.54	3.93E-02
TC1900011867.hg.1	Transmembrane protein 205	TMEM205	-1.5	3.74E-02	-2.22	7.00E-04
TC1700007360.hg.1	Transmembrane protein 97	TMEM97	-1.63	1.74E-02	-1.85	4.80E-03
TC0800006681.hg.1	Tankyrase, TRF1-interacting ankyrin-related ADP-ribose polymerase	TNKS	-1.66	2.17E-02	-2.57	4.00E-04
TC2200007361.hg.1	Translocase of outer mitochondrial membrane 22 homolog (yeast)	TOMM22	-1.78	3.20E-03	-1.56	1.38E-02

TC1700012352.hg.1	Trafficking protein particle complex 1	TRAPPC1	-1.51	1.61E-02	-2.16	3.00E-04
TSUnmapped00000493.hg.1	Trafficking protein particle complex 4	TRAPPC4	-1.51	2.13E-02	-1.6	2.87E-02
TC0900008574.hg.1	Tripartite motif containing 32	TRIM32	-1.75	3.77E-02	-3.12	6.00E-04
TC2200008018.hg.1	Trna methyltransferase 2 homolog A	TRMT2A	-1.57	1.43E-02	-1.58	1.29E-02
TC0600012958.hg.1	TSPY-like 4	TSPYL4	-2.23	6.00E-04	-2.55	2.00E-04
TC0900010946.hg.1	Thiosulfate sulfurtransferase (rhodanese)-like domain containing 2	TSTD2	-1.75	5.10E-03	-2.34	3.00E-04
TC2000009053.hg.1	TELO2 interacting protein 1	TTI1	-1.77	4.60E-03	-1.74	5.30E-03
TC1800006731.hg.1	Tubulin, beta 6 class V	TUBB6	-2.13	1.50E-03	-1.85	5.70E-03
TC1300008080.hg.1	Memczak2013 ANTISENSE, coding, INTERNAL, intronic best transcript NM_006322	TUBGCP3	-2.01	5.30E-03	-1.95	6.70E-03
TC1100012358.hg.1	Ubiquitin specific peptidase 28	USP28	-1.76	1.62E-02	-2.1	3.50E-03
TC0800008626.hg.1	UTP23, small subunit (SSU) processome component, homolog (yeast)	UTP23	-1.64	1.24E-02	-1.86	3.30E-03
TC1400009784.hg.1	VPS33B interacting protein, apical-basolateral polarity regulator, spe-39 homolog	VIPAS39	-1.52	4.00E-02	-1.58	2.74E-02
TC1400010764.hg.1	Vesicle transport through interaction with t-snares 1B	VTI1B	-1.59	2.69E-02	-2.03	2.50E-03
TC0100009580.hg.1	WD repeat domain 3	WDR3	-1.52	3.55E-02	-2.54	3.00E-04
TC0900011650.hg.1	WD repeat domain 34	WDR34	-1.96	1.30E-02	-1.68	4.44E-02
TC0200016681.hg.1	WD repeat domain 92	WDR92	-1.55	1.60E-02	-1.94	1.30E-03
TC0600006611.hg.1	Werner helicase interacting protein 1	WRNIP1	-1.7	3.52E-02	-1.75	2.84E-02
TC1200010944.hg.1	Zinc finger and BTB domain containing 39	ZBTB39	-1.9	2.20E-03	-3.15	2.22E-05
TC0600014112.hg.1	Zinc finger and BTB domain containing 9	ZBTB9	-1.98	8.60E-03	-1.73	2.59E-02
TC1900011935.hg.1	ZFP30 zinc finger protein	ZFP30	-1.72	1.84E-02	-1.6	3.51E-02
TC1900011814.hg.1	Zinc finger protein 17	ZNF17	-1.66	4.67E-02	-1.88	1.74E-02
TC1900008655.hg.1	Zinc finger protein 175	ZNF175	-2.19	5.00E-04	-2.66	8.46E-05
TC1900011752.hg.1	Zinc finger protein 223	ZNF223	-1.59	4.03E-02	-1.58	4.43E-02
TC1700009551.hg.1	Zinc finger protein 232	ZNF232	-2.16	8.20E-03	-2.57	2.30E-03
TC1200012737.hg.1	Zinc finger protein 268	ZNF268	-1.57	2.33E-02	-1.61	1.79E-02
TC2000009973.hg.1	Zinc finger protein 343	ZNF343	-1.57	3.27E-02	-2.79	2.00E-04
TSUnmapped00000461.hg.1	Zinc finger protein 35	ZNF35	-2.34	7.00E-04	-1.68	1.57E-02
TSUnmapped00000018.hg.1	Zinc finger protein 35	ZNF35	-2.58	1.00E-04	-1.53	2.63E-02

TC0600014086.hg.1	Zinc finger protein 391	ZNF391	-1.61	2.02E-02	-1.94	3.10E-03
TC0200016418.hg.1	Zinc finger protein 512	ZNF512	-1.9	2.40E-03	-1.72	7.00E-03
TC1900009002.hg.1	Zinc finger protein 530	ZNF530	-1.8	1.41E-02	-1.68	2.61E-02
TSUnmapped00000068.hg.1	Zinc finger protein 780A	ZNF780A	-2.23	9.50E-03	-2.01	1.60E-03
TSUnmapped00000038.hg.1	Zinc finger protein 780A	ZNF780A	-1.84	3.80E-03	-2.08	1.54E-02
TC2000007507.hg.1	Zinc finger, SWIM-type containing 1	ZSWIM1	-1.6	3.17E-02	-1.94	5.60E-03

**Table S4, related to Figure A3D.** Common Gene Ontology results from microarray data on HeLa and HEK-293T cells treated with exogenous MTALTND4 (10  $\mu$ M) for 4 hrs.

	Accession #	Gene Ontology Term	HeLa		HEK	
			Fold change	P-value	Fold change	P-value
Binding	GO:0005515	Protein binding	1.1	1.70E-06	1.06	1.47E-05
	GO:0005488	Binding	1.07	1.87E-05	1.04	1.42E-05
	GO:0005488	Binding	-1.09	5.10E-09	-1.15	1.97E-67
	GO:0005515	Protein binding	-1.1	9.44E-06	-1.23	6.03E-84
Cellular process	GO:0009987	Cellular process	1.07	1.63E-04	1.05	2.12E-05
	GO:0006996	Organelle organization	-1.32	5.23E-05	-1.44	2.32E-21
Metabolic process	GO:0008152	Metabolic process	-1.19	7.11E-06	-1.34	1.52E-43
	GO:0071704	Organic substance metabolic process	-1.2	9.31E-06	-1.34	2.39E-39
	GO:0044238	Primary metabolic process	-1.22	4.87E-06	-1.39	1.35E-44
	GO:0044237	Cellular metabolic process	-1.27	1.77E-08	-1.41	1.28E-50
	GO:0006807	Nitrogen compound metabolic process	-1.27	1.33E-07	-1.39	1.81E-40
	GO:0090304	Nucleic acid metabolic process	-1.48	8.29E-06	-1.63	7.62E-24
	GO:1901360	Organic cyclic compound metabolic process	-1.51	3.99E-09	-1.47	1.35E-21
	GO:0006139	Nucleobase-containing compound metabolic process	-1.53	2.27E-08	-1.54	5.33E-23
	GO:0034641	Cellular nitrogen compound metabolic process	-1.53	7.82E-11	-1.46	1.54E-22
	GO:0046483	Heterocycle metabolic process	-1.55	1.57E-09	-1.51	4.39E-22
	GO:0006725	Cellular aromatic compound metabolic process	-1.57	3.24E-10	-1.49	5.52E-21
Catalytic activity	GO:0003824	Catalytic activity	-1.31	6.83E-08	-1.31	9.34E-22
	GO:0044249	Cellular biosynthetic process	-1.32	3.45E-08	-1.48	3.25E-06
Biosynthetic process	GO:0009058	Biosynthetic process	-1.42	2.43E-05	-1.36	2.51E-10
	GO:1901576	Organic substance biosynthetic process	-1.45	9.34E-06	-1.36	4.94E-10
	GO:1901566	Organonitrogen compound biosynthetic process	-1.57	3.86E-05	-1.27	4.54E-04
	GO:1901362	Organic cyclic compound biosynthetic process	-1.69	2.01E-05	-1.39	1.43E-05
	GO:0044271	Cellular nitrogen compound biosynthetic process	-1.73	1.46E-07	-1.41	9.97E-08
	GO:0019438	Aromatic compound biosynthetic process	-1.8	5.34E-06	-1.39	5.95E-05
	GO:0018130	Heterocycle biosynthetic process	-1.82	3.32E-06	-1.43	1.31E-05
Protein transport	GO:0071705	Nitrogen compound transport	-1.53	3.34E-05	-1.55	5.76E-14
	GO:0015031	Protein transport	-1.63	3.03E-05	-1.69	6.27E-15
Mitochondrial transport	GO:0006839	Mitochondrial transport	-3.12	1.46E-05	-1.95	2.32E-04

**Table S5.** Pathways associated with common downregulated and upregulated genes between HeLa and HEK-293T cells from microarrays.

DOWNREGULATED COMMON GENES				
Database	Related pathways	Gene count	P-value	Gene symbols
REACTOME	Senescence-Associated Secretory Phenotype (SASP)	5	1.7E-02	H3C15, CDC23, H3C14, H4C14, RPS6KA1
KEGG	Herpes simplex virus 1 infection	10	2.1E-02	ZNF17, ZNF175, ZFP30, SRSF1, EIF2AK4, ZNF780A, ZNF268, ZNF223, ZNF530, ZNF343
REACTOME	Cell Cycle, Mitotic	11	2.4E-02	GIN51, CCND3, TUBB6, H3C15, CDC23, CCND2, H3C14, H4C14, CHMP2A, TUBGCP3, AJUBA
REACTOME	RNA Polymerase II Transcription	20	2.5E-02	ZNF17, SMARCE1, ZNF175, PRKAA2, NCBP2, CSTF2, SRSF1, CCND3, H3C15, CDC23, CCND2, ZFP30, H3C14, H4C14, MED20, PIN1, ZNF268, ZNF223, ZNF530, ZNF343
REACTOME	Cellular Senescence	6	2.9E-02	H3C15, CDC23, H3C14, H4C14, RPS6KA1, TERF2IP
REACTOME	RMTs methylate histone arginines	4	3.2E-02	SMARCE1, H3C15, H3C14, H4C14
REACTOME	Gene expression (Transcription)	21	3.3E-02	ZNF17, SMARCE1, ZNF175, PRKAA2, NCBP2, CSTF2, SRSF1, CCND3, H3C15, CDC23, CCND2, ZFP30, H3C14, H4C14, MED20, POLR1C, PIN1, ZNF268, ZNF223, ZNF530, ZNF343
REACTOME	Cell Cycle	12	3.8E-02	GIN51, CCND3, TUBB6, H3C15, CDC23, CCND2, H3C14, H4C14, CHMP2A, TERF2IP, TUBGCP3, AJUBA
REACTOME	Aberrant regulation of mitotic cell cycle due to RB1 defects	3	3.9E-02	CCND3, CDC23, CCND2
REACTOME	Diseases of mitotic cell cycle	3	4.3E-02	CCND3, CDC23, CCND2
REACTOME	B-WICH complex positively regulates rRNA expression	4	4.5E-02	H3C15, H3C14, H4C14, POLR1C
REACTOME	RNA Polymerase I Promoter Escape	4	4.5E-02	H3C15, H3C14, H4C14, POLR1C
REACTOME	Macroautophagy	5	3.1E-02	TUBB6, PRKAA2, ATG16L1, CHMP2A, TOMM22
REACTOME	Autophagy	5	4.3E-02	TUBB6, PRKAA2, ATG16L1, CHMP2A, TOMM22
REACTOME	Protein localization	6	1.4E-02	CHCHD4, ALDH3A2, ACOT2, BCS1L, GDAP1, TOMM22
REACTOME	Translocation of SLC2A4 (GLUT4) to the plasma membrane	4	2.5E-02	TUBB6, PRKAA2, EXOC3, RAB11A
REACTOME	Metabolism of RNA	12	3.2E-02	SF3A3, PUS3, NCBP2, WDR3, CSTF2, SRSF1, RPP21, RRP36, GEMIN5, EXOSC2, LSM4, SMG8
REACTOME	ER to Golgi Anterograde Transport	5	4.6E-02	DCTN5, TUBB6, TRAPPC1, TRAPPC4, COPZ1



**UPREGULATED COMMON GENES**

Database	Related pathways	Gene count	P-value	Gene symbols
KEGG	NF-kappa B signaling pathway	6	1.8E-03	NFKBIA, GADD45B, TNFAIP3, TNFRSF11A, LAT, ICAM1
KEGG	Chronic myeloid leukemia	5	3.8E-03	NFKBIA, TGFB2, GADD45B, PIK3R3, PIK3R1
KEGG	Small cell lung cancer	5	7.5E-03	NFKBIA, LAMA2, GADD45B, PIK3R3, PIK3R1
KEGG	TNF signaling pathway	5	1.5E-02	NFKBIA, PIK3R3, TNFAIP3, PIK3R1, ICAM1
KEGG	Leukocyte transendothelial migration	5	1.6E-02	PIK3R3, CXCR4, PIK3R1, JAM2, ICAM1
KEGG	Osteoclast differentiation	5	2.3E-02	NFKBIA, TGFB2, PIK3R3, TNFRSF11A, PIK3R1
KEGG	FoxO signaling pathway	5	2.5E-02	TGFB2, GADD45B, CCNG2, PIK3R3, PIK3R1
KEGG	Pancreatic cancer	4	2.6E-02	TGFB2, GADD45B, PIK3R3, PIK3R1
KEGG	Epstein-Barr virus infection	6	2.8E-02	NFKBIA, GADD45B, PIK3R3, TNFAIP3, PIK3R1, ICAM1
KEGG	Apoptosis	5	2.8E-02	NFKBIA, GADD45B, DDIT3, PIK3R3, PIK3R1
KEGG	B cell receptor signaling pathway	4	3.2E-02	NFKBIA, INPP5D, PIK3R3, PIK3R1
KEGG	Lipid and atherosclerosis	6	3.5E-02	NFKBIA, CYP2C8, DDIT3, PIK3R3, PIK3R1, ICAM1
KEGG	Colorectal cancer	4	3.6E-02	TGFB2, GADD45B, PIK3R3, PIK3R1
KEGG	PD-L1 expression and PD-1 checkpoint pathway in cancer	4	3.9E-02	NFKBIA, PIK3R3, PIK3R1, LAT
KEGG	Prostate cancer	4	4.9E-02	NFKBIA, PIK3R3, ERG, PIK3R1
REACTOME	Phospholipid metabolism	8	6.8E-04	BCHE, OSBPL5, ABHD3, INPP5D, CPNE7, PIK3R3, PLEKHA6, PIK3R1
REACTOME	Synthesis of PIPs at the plasma membrane	4	6.0E-03	INPP5D, PIK3R3, PLEKHA6, PIK3R1
KEGG	Pyruvate metabolism	4	7.2E-03	LDHC, PC, ACSS2, ACOT12
REACTOME	Interleukin receptor SHC signaling	3	1.5E-02	INPP5D, PIK3R3, PIK3R1
REACTOME	Metabolism	24	1.6E-02	BCHE, TPH1, SLC16A1, ACSS2, OSBPL5, ABHD3, CPNE7, SLC35D2, ACOT12, AMPD1, PIK3R3, PLEKHA6, PYGM, PIK3R1, CYP2C18, LDHC, RPL7A, CYP2C8, PC, CHST11, INPP5D, HMBS, ENPP2, HAO2
KEGG	Fc epsilon RI signaling pathway	4	2.0E-02	INPP5D, PIK3R3, PIK3R1, LAT
REACTOME	PI Metabolism	4	2.1E-02	INPP5D, PIK3R3, PLEKHA6, PIK3R1
REACTOME	GPVI-mediated activation cascade	3	2.5E-02	PIK3R3, PIK3R1, LAT
KEGG	B cell receptor signaling pathway	4	3.2E-02	NFKBIA, INPP5D, PIK3R3, PIK3R1
REACTOME	Metabolism of lipids	11	3.3E-02	BCHE, CYP2C8, OSBPL5, ABHD3, INPP5D, CPNE7, ACOT12, PIK3R3, PLEKHA6, PIK3R1, HAO2
KEGG	Lipid and atherosclerosis	6	3.5E-02	NFKBIA, CYP2C8, DDIT3, PIK3R3, PIK3R1, ICAM1
REACTOME	Interleukin-2 family signaling	3	3.8E-02	INPP5D, PIK3R3, PIK3R1
KEGG	PD-L1 expression and PD-1 checkpoint pathway in cancer	4	3.9E-02	NFKBIA, PIK3R3, PIK3R1, LAT
REACTOME	Interleukin-3, Interleukin-5 and GM-CSF signaling	3	4.4E-02	INPP5D, PIK3R3, PIK3R1
KEGG	Phosphatidylinositol signaling system	4	4.9E-02	DGKD, INPP5D, PIK3R3, PIK3R1
KEGG	Fc gamma R-mediated phagocytosis	4	4.9E-02	INPP5D, PIK3R3, PIK3R1, LAT
KEGG	Transcriptional misregulation in cancer	6	0.023	NR4A3, GADD45B, H3C13, DDIT3, NFKBIZ, ERG
REACTOME	Interleukin-7 signaling	3	0.026	H3C13, PIK3R3, PIK3R1
REACTOME	RUNX1 regulates genes involved in megakaryocyte differentiation and platelet function	4	0.03	NR4A3, H3C13, H2BC6, H2AC4
BIOCARTA	Selective expression of chemokine receptors during T-cell polarization	3	0.037	TGFB2, CXCR4, IL12RB2
REACTOME	Aspirin ADME	3	0.038	BCHE, CYP2C8, SLC16A1
REACTOME	DAP12 interactions	3	0.039	SIGLEC15, PIK3R1, LAT
KEGG	Rheumatoid arthritis	4	0.044	TGFB2, TNFRSF11A, ATP6V1D, ICAM1

**Table S6.** List of common proteins found by mass spectrometry following pull-down assay. Only results with 95% probability and two or more unique spectrum count are presented.

Protein Name	Accession #	Gene symbol	Molecular Weight	HeLa #1 Unique spectrum count	HeLa #2 Unique spectrum count	HEK Unique spectrum count
Cluster of Tubulin beta chain OS=Homo sapiens OX=9606 GN=TUBB PE=1 SV=2	P07437	TUBB	50 kDa	52	59	21
Glutathione S-transferase LANCL1 OS=Homo sapiens OX=9606 GN=LANCL1 PE=1 SV=1	O43813	LANCL1	45 kDa	34	13	12
Nucleophosmin OS=Homo sapiens OX=9606 GN=NPM1 PE=1 SV=2	P06748	NPM1	33 kDa	21	10	6
Nucleosome assembly protein 1-like 1 OS=Homo sapiens OX=9606 GN=NAP1L1 PE=1 SV=1	P55209	NAP1L1	45 kDa	20	10	11
Protein phosphatase 1G OS=Homo sapiens OX=9606 GN=PPM1G PE=1 SV=1	O15355	PPM1G	59 kDa	18	11	7
Complement component 1 Q subcomponent-binding protein, mitochondrial OS=Homo sapiens OX=9606 GN=C1QBP PE=1 SV=1	Q07021	C1QBP	31 kDa	17	37	16
Elongation factor 1-gamma OS=Homo sapiens OX=9606 GN=EEF1G PE=1 SV=3	P26641	EEF1G	50 kDa	17	15	8
Nucleosome assembly protein 1-like 4 OS=Homo sapiens OX=9606 GN=NAP1L4 PE=1 SV=1	Q99733	NAP1L4	43 kDa	14	5	3
Protein SET OS=Homo sapiens OX=9606 GN=SET PE=1 SV=3	Q01105	SET	33 kDa	14	4	3
Acidic leucine-rich nuclear phosphoprotein 32 family member B OS=Homo sapiens OX=9606 GN=ANP32B PE=1 SV=1	Q92688	ANP32B	29 kDa	13	27	6
Cluster of Tubulin alpha-1B chain OS=Homo sapiens OX=9606 GN=TUBA1B PE=1 SV=1	P68363	TUBA1B	50 kDa	12	37	11
40S ribosomal protein S18 OS=Homo sapiens OX=9606 GN=RPS18 PE=1 SV=3	P62269	RPS18	18 kDa	8	13	8
40S ribosomal protein S16 OS=Homo sapiens OX=9606 GN=RPS16 PE=1 SV=2	P62249	RPS16	16 kDa	8	11	7
ATP synthase subunit alpha, mitochondrial OS=Homo sapiens OX=9606 GN=ATP5F1A PE=1 SV=1	P25705	ATP5F1A	60 kDa	8	8	4

Proliferating cell nuclear antigen OS=Homo sapiens OX=9606 GN=PCNA PE=1 SV=1	P12004	PCNA	29 kDa	8	8	3
Y-box-binding protein 1 OS=Homo sapiens OX=9606 GN=YBX1 PE=1 SV=3	P67809	YBX1	36 kDa	7	10	12
Cluster of Heterogeneous nuclear ribonucleoprotein H OS=Homo sapiens OX=9606 GN=HNRNPH1 PE=1 SV=4	P31943	HNRNPH1	49 kDa	7	13	8
40S ribosomal protein S3 OS=Homo sapiens OX=9606 GN=RPS3 PE=1 SV=2	P23396	RPS3	27 kDa	7	15	6
Elongation factor 1-beta OS=Homo sapiens OX=9606 GN=EEF1B2 PE=1 SV=3	P24534	EEF1B2	25 kDa	7	6	4
40S ribosomal protein S5 OS=Homo sapiens OX=9606 GN=RPS5 PE=1 SV=4	P46782	RPS5	23 kDa	6	9	11
40S ribosomal protein S19 OS=Homo sapiens OX=9606 GN=RPS19 PE=1 SV=2	P39019	RPS19	16 kDa	6	8	2
60S ribosomal protein L27 OS=Homo sapiens OX=9606 GN=RPL27 PE=1 SV=2	P61353	RPL27	16 kDa	6	4	2
60S ribosomal protein L5 OS=Homo sapiens OX=9606 GN=RPL5 PE=1 SV=3	P46777	RPL5	34 kDa	5	12	6
40S ribosomal protein S13 OS=Homo sapiens OX=9606 GN=RPS13 PE=1 SV=2	P62277	RPS13	17 kDa	5	5	5
40S ribosomal protein S10 OS=Homo sapiens OX=9606 GN=RPS10 PE=1 SV=1	P46783	RPS10	19 kDa	5	9	3
60S ribosomal protein L23 OS=Homo sapiens OX=9606 GN=RPL23 PE=1 SV=1	P62829	RPL23	15 kDa	5	3	2
40S ribosomal protein S8 OS=Homo sapiens OX=9606 GN=RPS8 PE=1 SV=2	P62241	RPS8	24 kDa	4	10	11
Cluster of 40S ribosomal protein S4, X isoform OS=Homo sapiens OX=9606 GN=RPS4X PE=1 SV=2	P62701	RPS4X	30 kDa	4	10	10
40S ribosomal protein S3a OS=Homo sapiens OX=9606 GN=RPS3A PE=1 SV=2	P61247	RPS3A	30 kDa	4	15	7
40S ribosomal protein S11 OS=Homo sapiens OX=9606 GN=RPS11 PE=1 SV=3	P62280	RPS11	18 kDa	4	5	6
Alpha-S1-casein OS=Bos taurus OX=9913 GN=CSN1S1 PE=1 SV=2	P02662	CSN1S1	25 kDa	4	7	3

Elongation factor Tu, mitochondrial OS=Homo sapiens OX=9606 GN=TUFM PE=1 SV=2	P49411	TUFM	50 kDa	4	9	2
60S ribosomal protein L7 OS=Homo sapiens OX=9606 GN=RPL7 PE=1 SV=1	P18124	RPL7	29 kDa	3	12	6
40S ribosomal protein S2 OS=Homo sapiens OX=9606 GN=RPS2 PE=1 SV=2	P15880	RPS2	31 kDa	3	4	5
60S ribosomal protein L12 OS=Homo sapiens OX=9606 GN=RPL12 PE=1 SV=1	P30050	RPL12	18 kDa	3	11	3
Stress-70 protein, mitochondrial OS=Homo sapiens OX=9606 GN=HSPA9 PE=1 SV=2	P38646	HSPA9	74 kDa	3	22	2
60S ribosomal protein L3 OS=Homo sapiens OX=9606 GN=RPL3 PE=1 SV=2	P39023	RPL3	46 kDa	3	2	2
60S ribosomal protein L26 OS=Homo sapiens OX=9606 GN=RPL26 PE=1 SV=1	P61254	RPL26	17 kDa	3	2	2
40S ribosomal protein SA OS=Homo sapiens OX=9606 GN=RPSA PE=1 SV=4	P08865	RPSA	33 kDa	3	4	2
Endoplasmic reticulum chaperone BiP OS=Homo sapiens OX=9606 GN=HSPA5 PE=1 SV=2	P11021	HSPA5	72 kDa	2	19	15
40S ribosomal protein S17 OS=Homo sapiens OX=9606 GN=RPS17 PE=1 SV=2	P08708	RPS17	16 kDa	2	3	5
60S ribosomal protein L27a OS=Homo sapiens OX=9606 GN=RPL27A PE=1 SV=2	P46776	RPL27A	17 kDa	2	4	4
ADP/ATP translocase 3 OS=Homo sapiens OX=9606 GN=SLC25A6 PE=1 SV=4	P12236	SLC25A6	33 kDa	2	9	3
40S ribosomal protein S14 OS=Homo sapiens OX=9606 GN=RPS14 PE=1 SV=3	P62263	RPS14	16 kDa	2	8	3
40S ribosomal protein S6 OS=Homo sapiens OX=9606 GN=RPS6 PE=1 SV=1	P62753	RPS6	29 kDa	2	6	3
ATP synthase subunit gamma, mitochondrial OS=Homo sapiens OX=9606 GN=ATP5F1C PE=1 SV=1	P36542	ATP5F1C	33 kDa	2	3	3
60S ribosomal protein L10a OS=Homo sapiens OX=9606 GN=RPL10A PE=1 SV=2	P62906	RPL10A	25 kDa	2	9	4
Prohibitin OS=Homo sapiens OX=9606 GN=PHB PE=1 SV=1	P35232	PHB1	30 kDa	2	6	4

**Table S7.** Pathways associated with interaction partners from pull-down assays on HeLa and HEK-293T cells.

INTERACTION PARTNERS				
Database	Related pathways	Gene count	P-value	Gene symbols
REACTOME	Eukaryotic Translation Elongation	27	2.1E-45	RPL5, RPL3, RPL12, RPL10A, RPL7, RPS4X, RPS14, RPS17, RPS16, RPS19, RPS18, RPS3, RPS2, RPS11, RPS10, RPS13, RPS8, RPL23, RPS5, RPS6, RPSA, RPS3A, RPL27A, RPL27, RPL26
REACTOME	Peptide chain elongation	25	3.7E-41	RPL5, RPL3, RPL12, RPL10A, RPL7, RPS4X, RPS14, RPS17, RPS16, RPS19, RPS18, RPS3, RPS2, RPS11, RPS10, RPS13, RPS8, RPL23, RPS5, RPS6, RPSA, RPS3A, RPL27A, RPL27, RPL26
REACTOME	Viral mRNA Translation	25	3.7E-41	RPL5, RPL3, RPL12, RPL10A, RPL7, RPS4X, RPS14, RPS17, RPS16, RPS19, RPS18, RPS3, RPS2, RPS11, RPS10, RPS13, RPS8, RPL23, RPS5, RPS6, RPSA, RPS3A, RPL27A, RPL27, RPL26
REACTOME	Selenocysteine synthesis	25	1.2E-40	RPL5, RPL3, RPL12, RPL10A, RPL7, RPS4X, RPS14, RPS17, RPS16, RPS19, RPS18, RPS3, RPS2, RPS11, RPS10, RPS13, RPS8, RPL23, RPS5, RPS6, RPSA, RPS3A, RPL27A, RPL27, RPL26
REACTOME	Eukaryotic Translation Termination	25	1.2E-40	RPL5, RPL3, RPL12, RPL10A, RPL7, RPS4X, RPS14, RPS17, RPS16, RPS19, RPS18, RPS3, RPS2, RPS11, RPS10, RPS13, RPS8, RPL23, RPS5, RPS6, RPSA, RPS3A, RPL27A, RPL27, RPL26
REACTOME	Nonsense Mediated Decay (NMD) independent of the Exon Junction Complex (EJC)	25	2.2E-40	RPL5, RPL3, RPL12, RPL10A, RPL7, RPS4X, RPS14, RPS17, RPS16, RPS19, RPS18, RPS3, RPS2, RPS11, RPS10, RPS13, RPS8, RPL23, RPS5, RPS6, RPSA, RPS3A, RPL27A, RPL27, RPL26
REACTOME	Formation of a pool of free 40S subunits	25	1.2E-39	RPL5, RPL3, RPL12, RPL10A, RPL7, RPS4X, RPS14, RPS17, RPS16, RPS19, RPS18, RPS3, RPS2, RPS11, RPS10, RPS13, RPS8, RPL23, RPS5, RPS6, RPSA, RPS3A, RPL27A, RPL27, RPL26
REACTOME	Response of EIF2AK4 (GCN2) to amino acid deficiency	25	1.2E-39	RPL5, RPL3, RPL12, RPL10A, RPL7, RPS4X, RPS14, RPS17, RPS16, RPS19, RPS18, RPS3, RPS2, RPS11, RPS10, RPS13, RPS8, RPL23, RPS5, RPS6, RPSA, RPS3A, RPL27A, RPL27, RPL26
REACTOME	L13a-mediated translational silencing of Ceruloplasmin expression	25	1.5E-38	RPL5, RPL3, RPL12, RPL10A, RPL7, RPS4X, RPS14, RPS17, RPS16, RPS19, RPS18, RPS3, RPS2, RPS11, RPS10, RPS13, RPS8, RPL23, RPS5, RPS6, RPSA, RPS3A, RPL27A, RPL27, RPL26

REACTOME	SRP-dependent cotranslational protein targeting to membrane	25	1.9E-38	RPL5, RPL3, RPL12, RPL10A, RPL7, RPS4X, RPS14, RPS17, RPS16, RPS19, RPS18, RPS3, RPS2, RPS11, RPS10, RPS13, RPS8, RPL23, RPS5, RPS6, RPSA, RPS3A, RPL27A, RPL27, RPL26
REACTOME	GTP hydrolysis and joining of the 60S ribosomal subunit	25	1.9E-38	RPL5, RPL3, RPL12, RPL10A, RPL7, RPS4X, RPS14, RPS17, RPS16, RPS19, RPS18, RPS3, RPS2, RPS11, RPS10, RPS13, RPS8, RPL23, RPS5, RPS6, RPSA, RPS3A, RPL27A, RPL27, RPL26
REACTOME	Nonsense-Mediated Decay (NMD)	25	3.8E-38	RPL5, RPL3, RPL12, RPL10A, RPL7, RPS4X, RPS14, RPS17, RPS16, RPS19, RPS18, RPS3, RPS2, RPS11, RPS10, RPS13, RPS8, RPL23, RPS5, RPS6, RPSA, RPS3A, RPL27A, RPL27, RPL26
REACTOME	Nonsense Mediated Decay (NMD) enhanced by the Exon Junction Complex (EJC)	25	3.8E-38	RPL5, RPL3, RPL12, RPL10A, RPL7, RPS4X, RPS14, RPS17, RPS16, RPS19, RPS18, RPS3, RPS2, RPS11, RPS10, RPS13, RPS8, RPL23, RPS5, RPS6, RPSA, RPS3A, RPL27A, RPL27, RPL26
REACTOME	Selenoamino acid metabolism	25	7.6E-38	RPL5, RPL3, RPL12, RPL10A, RPL7, RPS4X, RPS14, RPS17, RPS16, RPS19, RPS18, RPS3, RPS2, RPS11, RPS10, RPS13, RPS8, RPL23, RPS5, RPS6, RPSA, RPS3A, RPL27A, RPL27, RPL26
REACTOME	Eukaryotic Translation Initiation	25	9.5E-38	RPL5, RPL3, RPL12, RPL10A, RPL7, RPS4X, RPS14, RPS17, RPS16, RPS19, RPS18, RPS3, RPS2, RPS11, RPS10, RPS13, RPS8, RPL23, RPS5, RPS6, RPSA, RPS3A, RPL27A, RPL27, RPL26
REACTOME	Cap-dependent Translation Initiation	25	9.5E-38	RPL5, RPL3, RPL12, RPL10A, RPL7, RPS4X, RPS14, RPS17, RPS16, RPS19, RPS18, RPS3, RPS2, RPS11, RPS10, RPS13, RPS8, RPL23, RPS5, RPS6, RPSA, RPS3A, RPL27A, RPL27, RPL26
REACTOME	Influenza Infection	26	9.3E-37	RPL5, RPL3, RPL12, RPL10A, RPL7, RPS4X, RPS14, RPS17, RPS16, RPS19, RPS18, RPS3, RPS2, RPS11, RPS10, RPS13, RPS8, RPL23, RPS5, RPS6, RPSA, RPS3A, RPL27A, RPL27, RPL26
REACTOME	Influenza Viral RNA Transcription and Replication	25	2.6E-36	RPL5, RPL3, RPL12, RPL10A, RPL7, RPS4X, RPS14, RPS17, RPS16, RPS19, RPS18, RPS3, RPS2, RPS11, RPS10, RPS13, RPS8, RPL23, RPS5, RPS6, RPSA, RPS3A, RPL27A, RPL27, RPL26
REACTOME	Cellular response to starvation	25	1.1E-34	RPL5, RPL3, RPL12, RPL10A, RPL7, RPS4X, RPS14, RPS17, RPS16, RPS19, RPS18, RPS3, RPS2, RPS11, RPS10, RPS13, RPS8, RPL23, RPS5, RPS6, RPSA, RPS3A, RPL27A, RPL27, RPL26
KEGG	Ribosome	25	1.4E-34	RPL5, RPL3, RPL12, RPL10A, RPL7, RPS4X, RPS14, RPS17, RPS16, RPS19, RPS18, RPS3, RPS2, RPS11, RPS10, RPS13, RPS8, RPL23, RPS5, RPS6, RPSA, RPS3A, RPL27A, RPL27, RPL26

REACTOME	Regulation of expression of SLITs and ROBOs	25	1.2E-33	RPL5, RPL3, RPL12, RPL10A, RPL7, RPS4X, RPS14, RPS17, RPS16, RPS19, RPS18, RPS3, RPS2, RPS11, RPS10, RPS13, RPS8, RPL23, RPS5, RPS6, RPSA, RPS3A, RPL27A, RPL27, RPL26
REACTOME	Translation	28	2.5E-33	RPL5, RPL3, RPL12, RPL10A, RPL7, EEF1B2, RPS4X, RPS14, RPS17, RPS16, RPS19, RPS18, RPS3, RPS2, RPS11, RPS10, RPS13, RPS8, RPL23, RPS5, RPS6, RPSA, RPS3A, TUFM, EEF1G, RPL27A, RPL27, RPL26
REACTOME	Major pathway of rRNA processing in the nucleolus and cytosol	25	7.5E-33	RPL5, RPL3, RPL12, RPL10A, RPL7, RPS4X, RPS14, RPS17, RPS16, RPS19, RPS18, RPS3, RPS2, RPS11, RPS10, RPS13, RPS8, RPL23, RPS5, RPS6, RPSA, RPS3A, RPL27A, RPL27, RPL26
REACTOME	rRNA processing in the nucleus and cytosol	25	2.9E-32	RPL5, RPL3, RPL12, RPL10A, RPL7, RPS4X, RPS14, RPS17, RPS16, RPS19, RPS18, RPS3, RPS2, RPS11, RPS10, RPS13, RPS8, RPL23, RPS5, RPS6, RPSA, RPS3A, RPL27A, RPL27, RPL26
REACTOME	rRNA processing	25	1E-31	RPL5, RPL3, RPL12, RPL10A, RPL7, RPS4X, RPS14, RPS17, RPS16, RPS19, RPS18, RPS3, RPS2, RPS11, RPS10, RPS13, RPS8, RPL23, RPS5, RPS6, RPSA, RPS3A, RPL27A, RPL27, RPL26
REACTOME	Signaling by ROBO receptors	25	5.4E-31	RPL5, RPL3, RPL12, RPL10A, RPL7, RPS4X, RPS14, RPS17, RPS16, RPS19, RPS18, RPS3, RPS2, RPS11, RPS10, RPS13, RPS8, RPL23, RPS5, RPS6, RPSA, RPS3A, RPL27A, RPL27, RPL26
KEGG	Coronavirus disease - COVID-19	25	2.4E-30	RPL5, RPL3, RPL12, RPL10A, RPL7, RPS4X, RPS14, RPS17, RPS16, RPS19, RPS18, RPS3, RPS2, RPS11, RPS10, RPS13, RPS8, RPL23, RPS5, RPS6, RPSA, RPS3A, RPL27A, RPL27, RPL26
REACTOME	SARS-CoV-2 modulates host translation machinery	16	4.9E-26	RPS8, RPS5, RPS6, RPS3A, RPSA, RPS4X, RPS14, RPS17, RPS16, RPS19, RPS18, RPS3, RPS2, RPS11, RPS10, RPS13
REACTOME	Formation of the ternary complex, and subsequently, the 43S complex	16	6.9E-26	RPS8, RPS5, RPS6, RPS3A, RPSA, RPS4X, RPS14, RPS17, RPS16, RPS19, RPS18, RPS3, RPS2, RPS11, RPS10, RPS13
REACTOME	Metabolism of amino acids and derivatives	25	3.1E-25	RPL5, RPL3, RPL12, RPL10A, RPL7, RPS4X, RPS14, RPS17, RPS16, RPS19, RPS18, RPS3, RPS2, RPS11, RPS10, RPS13, RPS8, RPL23, RPS5, RPS6, RPSA, RPS3A, RPL27A, RPL27, RPL26
REACTOME	Translation initiation complex formation	16	6.3E-25	RPS8, RPS5, RPS6, RPS3A, RPSA, RPS4X, RPS14, RPS17, RPS16, RPS19, RPS18, RPS3, RPS2, RPS11, RPS10, RPS13
REACTOME	Ribosomal scanning and start codon recognition	16	6.3E-25	RPS8, RPS5, RPS6, RPS3A, RPSA, RPS4X, RPS14, RPS17, RPS16, RPS19, RPS18, RPS3, RPS2, RPS11, RPS10, RPS13

REACTOME	Activation of the mRNA upon binding of the cap-binding complex and eIFs, and subsequent binding to 43S	16	8.5E-25	RPS8, RPS5, RPS6, RPS3A, RPSA, RPS4X, RPS14, RPS17, RPS16, RPS19, RPS18, RPS3, RPS2, RPS11, RPS10, RPS13
REACTOME	Metabolism of RNA	28	2.2E-23	RPL5, SET, RPL3, RPL12, YBX1, RPL10A, RPL7, RPS4X, RPS14, RPS17, RPS16, RPS19, RPS18, RPS3, RPS2, RPS11, RPS10, RPS13, RPS8, RPL23, RPS5, RPS6, RPSA, RPS3A, HNRNPH1, RPL27A, RPL27, RPL26
REACTOME	Axon guidance	26	1.2E-22	RPL5, RPL3, RPL12, RPL10A, RPL7, RPS4X, RPS14, TUBA1B, RPS17, RPS16, RPS19, RPS18, RPS3, RPS2, RPS11, RPS10, RPS13, RPS8, RPL23, RPS5, RPS6, RPSA, RPS3A, RPL27A, RPL27, RPL26
REACTOME	Nervous system development	26	3.5E-22	RPL5, RPL3, RPL12, RPL10A, RPL7, RPS4X, RPS14, TUBA1B, RPS17, RPS16, RPS19, RPS18, RPS3, RPS2, RPS11, RPS10, RPS13, RPS8, RPL23, RPS5, RPS6, RPSA, RPS3A, RPL27A, RPL27, RPL26
REACTOME	Cellular responses to stress	28	1.2E-21	RPL5, RPL3, RPL12, RPL10A, RPL7, RPS4X, RPS14, TUBA1B, RPS17, RPS16, RPS19, RPS18, RPS3, RPS2, RPS11, RPS10, RPS13, HSPA9, HSPA5, RPS8, RPL23, RPS5, RPS6, RPSA, RPS3A, RPL27A, RPL27, RPL26
REACTOME	Cellular responses to stimuli	28	1.9E-21	RPL5, RPL3, RPL12, RPL10A, RPL7, RPS4X, RPS14, TUBA1B, RPS17, RPS16, RPS19, RPS18, RPS3, RPS2, RPS11, RPS10, RPS13, HSPA9, HSPA5, RPS8, RPL23, RPS5, RPS6, RPSA, RPS3A, RPL27A, RPL27, RPL26
REACTOME	Infectious disease	30	2.2E-21	RPL5, RPL3, RPL12, RPL10A, RPL7, RPS4X, RPS14, TUBA1B, RPS17, RPS16, RPS19, RPS18, RPS3, RPS2, RPS11, RPS10, RPS13, NPM1, RPS8, RPL23, RPS5, TUBB, RPS6, RPSA, RPS3A, TUFM, RPL27A, RPL27, RPL26, SLC25A6
REACTOME	Disease	33	5.6E-18	RPL5, PHB1, RPL3, RPL12, RPL10A, RPL7, RPS4X, RPS14, TUBA1B, RPS17, RPS16, RPS19, RPS18, C1QBP, RPS3, RPS2, RPS11, RPS10, RPS13, NPM1, RPS8, RPL23, RPS5, TUBB, RPS6, RPSA, RPS3A, TUFM, EEF1G, RPL27A, RPL27, RPL26, SLC25A6
REACTOME	SARS-CoV-2-host interactions	17	7.3E-18	RPS8, RPS5, RPS6, RPS3A, RPSA, TUFM, RPS4X, RPS14, RPS17, RPS16, RPS19, RPS18, RPS3, RPS2, RPS11, RPS10, RPS13
REACTOME	SARS-CoV-2 Infection	17	1.9E-15	RPS8, RPS5, RPS6, RPS3A, RPSA, TUFM, RPS4X, RPS14, RPS17, RPS16, RPS19, RPS18, RPS3, RPS2, RPS11, RPS10, RPS13



REACTOME	SARS-CoV Infections	18	3.7E-15	NPM1, RPS8, RPS5, TUBB, RPS6, RPS3A, RPSA, TUFM, RPS4X, RPS14, RPS17, RPS16, RPS19, RPS18, RPS3, RPS2, RPS11, RPS10, RPS13
REACTOME	Developmental Biology	26	3.7E-15	RPL5, RPL3, RPL12, RPL10A, RPL7, RPS4X, RPS14, TUBA1B, RPS17, RPS16, RPS19, RPS18, RPS3, RPS2, RPS11, RPS10, RPS13, RPS8, RPL23, RPS5, RPS6, RPSA, RPS3A, RPL27A, RPL27, RPL26
REACTOME	Metabolism of proteins	31	1.7E-14	RPL5, RPL3, PCNA, RPL12, RPL10A, RPL7, EEF1B2, RPS4X, RPS14, TUBA1B, RPS17, RPS16, RPS19, RPS18, RPS3, RPS2, RPS11, RPS10, RPS13, NPM1, RPS8, RPL23, RPS5, RPS6, RPSA, RPS3A, TUFM, EEF1G, RPL27A, RPL27, RPL26
REACTOME	Metabolism	27	8.5E-10	RPL5, RPL3, RPL12, RPL10A, RPL7, RPS4X, RPS14, RPS17, ATP5F1A, RPS16, RPS19, RPS18, RPS3, RPS2, RPS11, RPS10, RPS13, RPS8, RPL23, RPS5, RPS6, RPSA, RPS3A, ATP5F1C, RPL27A, RPL27, RPL26
REACTOME	rRNA modification in the nucleus and cytosol	3	2.2E-02	RPS14, RPS6, RPS2
KEGG	Parkinson disease	6	4.6E-03	ATP5F1A, TUBA1B, HSPA5, TUBB, ATP5F1C, SLC25A6
KEGG	Prion disease	6	5.1E-03	ATP5F1A, TUBA1B, HSPA5, TUBB, ATP5F1C, SLC25A6
REACTOME	Organelle biogenesis and maintenance	5	2.5E-02	HSPA9, ATP5F1A, TUBA1B, TUBB, ATP5F1C
KEGG	Huntington disease	5	3.7E-02	ATP5F1A, TUBA1B, TUBB, ATP5F1C, SLC25A6
KEGG	Pathways of neurodegeneration - multiple diseases	6	4.6E-02	ATP5F1A, TUBA1B, HSPA5, TUBB, ATP5F1C, SLC25A6
REACTOME	Signaling by ALK in cancer	3	1.8E-02	EEF1G, NPM1, RPS6
REACTOME	Signaling by ALK fusions and activated point mutants	3	1.8E-02	EEF1G, NPM1, RPS6
REACTOME	Cristae formation	3	6.0E-03	HSPA9, ATP5F1A, ATP5F1C
REACTOME	Organelle biogenesis and maintenance	5	2.5E-02	HSPA9, ATP5F1A, ATP5F1C, TUBA1B, TUBB
REACTOME	Mitochondrial biogenesis	3	5.0E-02	HSPA9, ATP5F1A, ATP5F1C,
REACTOME	Mitochondrial protein import	3	2.4E-02	HSPA9, ATP5F1A, SLC25A6

**Table S8.** Signaling pathways associated with C1QBP and YBX1 interactions.

Pathway identifier	Related Pathways	Mapped entities	Interacts with	# of Reactions	Found reaction identifiers
R-HSA-162582	Signal Transduction	C1QBP; YBX1	C1QBP; YBX1	9	R-HSA-9017817;R-HSA-5687079;R-HSA-9733216;R-HSA-9013009;R-HSA-9733247;R-HSA-9017797;R-HSA-437192;R-HSA-374331;R-HSA-9013110
R-HSA-157118	Signaling by NOTCH	YBX1	C1QBP	2	R-HSA-9017817;R-HSA-9017797
R-HSA-9012852	Signaling by NOTCH3	YBX1	C1QBP	2	R-HSA-9017817;R-HSA-9017797
R-HSA-9013507	NOTCH3 Activation and Transmission of Signal to the Nucleus	YBX1	C1QBP	2	R-HSA-9017817;R-HSA-9017797
R-HSA-9017802	Noncanonical activation of NOTCH3	YBX1	C1QBP	2	R-HSA-9017817;R-HSA-9017797
R-HSA-162582	Signal Transduction	C1QBP; YBX1	C1QBP; YBX1	9	R-HSA-9017817;R-HSA-5687079;R-HSA-9733216;R-HSA-9013009;R-HSA-9733247;R-HSA-9017797;R-HSA-437192;R-HSA-374331;R-HSA-9013110
R-HSA-9006936	Signaling by TGFB family members	YBX1	C1QBP	2	R-HSA-9733216;R-HSA-9733247
R-HSA-170834	Signaling by TGF-beta Receptor Complex	YBX1	C1QBP	2	R-HSA-9733216;R-HSA-9733247
R-HSA-2173793	Transcriptional activity of SMAD2/SMAD3:SMAD4 heterotrimer	YBX1	C1QBP	2	R-HSA-9733216;R-HSA-9733247
R-HSA-2173796	SMAD2/SMAD3:SMAD4 heterotrimer regulates transcription	YBX1	C1QBP	2	R-HSA-9733216;R-HSA-9733247
R-HSA-8953854	Metabolism of RNA	YBX1	YBX1; C1QBP	31	R-HSA-72130;R-HSA-428293;R-HSA-429955;R-HSA-75079;R-HSA-159050;R-HSA-75081;R-HSA-75080;R-HSA-72139;R-HSA-75083;R-HSA-75082;R-HSA-111439;R-HSA-72143;R-HSA-428296;R-HSA-77587;R-HSA-158484;R-HSA-77591;R-HSA-77593;R-HSA-927830;R-HSA-77594;R-HSA-72160;R-HSA-72103;R-HSA-72231;R-HSA-77094;R-HSA-72107;R-HSA-158447;R-HSA-429992;R-HSA-450551;R-HSA-156661;R-HSA-72185;R-HSA-72124;R-HSA-72127

R-HSA-72203	Processing of Capped Intron-Containing Pre-mRNA	YBX1	YBX1; C1QBP	22	R-HSA-72160;R-HSA-72130;R-HSA-75079;R-HSA-72103;R-HSA-72231;R-HSA-159050;R-HSA-75081;R-HSA-75080;R-HSA-72107;R-HSA-72139;R-HSA-75083;R-HSA-75082;R-HSA-158447;R-HSA-111439;R-HSA-72143;R-HSA-77587;R-HSA-158484;R-HSA-156661;R-HSA-72185;R-HSA-77594;R-HSA-72124;R-HSA-72127
R-HSA-72172	mRNA Splicing	YBX1	YBX1; C1QBP	14	R-HSA-72160;R-HSA-72130;R-HSA-75079;R-HSA-72231;R-HSA-75081;R-HSA-75080;R-HSA-72107;R-HSA-72139;R-HSA-75083;R-HSA-75082;R-HSA-72143;R-HSA-156661;R-HSA-72124;R-HSA-72127
R-HSA-72163	mRNA Splicing - Major Pathway	YBX1	YBX1; C1QBP	9	R-HSA-72160;R-HSA-72130;R-HSA-72231;R-HSA-156661;R-HSA-72107;R-HSA-72139;R-HSA-72124;R-HSA-72143;R-HSA-72127
R-HSA-74160	Gene expression (Transcription)	YBX1	YBX1; C1QBP; C1QBP	8	R-HSA-6804998;R-HSA-5250947;R-HSA-9733216;R-HSA-8952382;R-HSA-9733247;R-HSA-163310;R-HSA-8864729;R-HSA-113454
R-HSA-73857	RNA Polymerase II Transcription	YBX1	C1QBP; YBX1	5	R-HSA-6804998;R-HSA-9733216;R-HSA-8952382;R-HSA-9733247;R-HSA-8864729
R-HSA-212436	Generic Transcription Pathway	YBX1	C1QBP; YBX1	5	R-HSA-6804998;R-HSA-9733216;R-HSA-8952382;R-HSA-9733247;R-HSA-8864729
R-HSA-2173793	Transcriptional activity of SMAD2/SMAD3:SMAD4 heterotrimer	YBX1	C1QBP	2	R-HSA-9733216;R-HSA-9733247
R-HSA-2173796	SMAD2/SMAD3:SMAD4 heterotrimer regulates transcription	YBX1	C1QBP	2	R-HSA-9733216;R-HSA-9733247
R-HSA-1643685	Disease	C1QBP	YBX1; C1QBP; C1QBP	22	R-HSA-9646295;R-HSA-9694452;R-HSA-9760510;R-HSA-8866857;R-HSA-5679000;R-HSA-9694780;R-HSA-9694364;R-HSA-9694467;R-HSA-168883;R-HSA-9694529;R-HSA-9694401;R-HSA-9694337;R-HSA-9716160;R-HSA-9696807;R-HSA-9645766;R-HSA-9694790;R-HSA-9694341;R-HSA-9755747;R-HSA-9694793;R-HSA-9694280;R-HSA-9694447;R-HSA-9700266