



HAL
open science

Maturation des ARN messagers mitochondriaux et système traductionnel chez l'algue verte *Chlamydomonas reinhardtii*

Herrade Meichel Claudel

► **To cite this version:**

Herrade Meichel Claudel. Maturation des ARN messagers mitochondriaux et système traductionnel chez l'algue verte *Chlamydomonas reinhardtii*. Biologie végétale. Université de Strasbourg, 2023. Français. NNT : 2023STRAJ067 . tel-04892970

HAL Id: tel-04892970

<https://theses.hal.science/tel-04892970v1>

Submitted on 17 Jan 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



École Doctorale
des Sciences de la Vie
et de la Santé
STRASBOURG

ibmp



Université de Strasbourg
École Doctorale 414 : Sciences de la Vie et de la Santé
UPR 2357 - Institut de Biologie Moléculaire des Plantes – CNRS

Thèse de Doctorat présentée par :
Herrade MEICHEL CLAUDEL

Soutenue le 15 septembre 2023
Pour obtenir le grade de Docteur de l'Université de Strasbourg
Discipline : Aspects moléculaires et cellulaires de la biologie

Maturation des ARN messagers mitochondriaux et
système traductionnel chez l'algue verte
Chlamydomonas reinhardtii

Thèse dirigée par :

Dr. Laurence DROUARD

Directrice de recherches, CNRS Strasbourg

Rapporteurs :

Dr. Claire REMACLE

Professeure, Université de Liège

Dr. Yves CHOQUET

Directeur de recherches, CNRS Paris

Examineurs :

Dr. José Manuel GUALBERTO

Directeur de recherches, CNRS Strasbourg

Invitée :

Dr Thalia SALINAS-GIEGE

Ingénieure de recherches, CNRS Strasbourg



« Bien sûr que ça va marcher ! Mais pas aujourd'hui.
Ça marchera peut-être l'année ou la saison prochaine. »

Théo Jansen

REMERCIEMENTS

Je souhaiterai tout d'abord remercier les membres de mon jury, Claire Remacle, Yves Choquet et José Gualberto, d'avoir accepté d'évaluer mes travaux de thèse.

Merci aux mécènes de cette thèse, le Labex Mitocross et la région Grand Est.

Je tiens également à remercier Laurence, ma directrice de thèse, pour son accueil dans l'équipe, ses conseils, son temps (ô combien précieux) et ses petites attentions gourmandes à chaque période de fête.

Merci à Thalia, de m'avoir donné ma chance et de m'avoir formée depuis le tout début. Merci pour votre patience et votre bonne humeur, vos conseils et votre écoute.

Merci également aux membres de l'équipe, présents et passés, pour l'ambiance de travail agréable dans le laboratoire, avec de l'entraide, des conseils et de la bonne humeur. Merci à Anne-Marie qui a toujours su être disponible quand j'avais « juste une petite question ». Une pensée particulière pour Jeanne, Lauriane, Mickaele et Romain qui ont grandement participé à ces bons moments au travail par le passé, et aux membres actuels de l'équipe, Alexandre, Christina, Grégory, Laureline et Valentina, dont la sympathie et la bonne humeur sont on ne peut plus agréables au quotidien !

Merci à l'équipe du 320, avec laquelle j'ai travaillé au quotidien également, et qui a vu mes premiers pas en laboratoire, lors de mon tout premier stage en 2015 avec Florent. Merci à Géraldine, Mathilde et Mathieu et, avec qui j'ai eu des interactions presque quotidiennes enrichissantes et agréables. Merci à Florent, mon mentor, d'avoir toujours été là, de m'avoir ouvert les portes du monde merveilleux des plantes, des mitochondries et du rhum, et de m'avoir laissé suivre ton chemin. Merci aussi à Pierre, pour nos petites discussions, nos raids et tes conseils toujours très justes.

Un grand merci surtout à Marjorie, nous avons commencé en même temps, mais ton départ un an avant le mien a laissé un grand vide dans mes journées au labo. Alors merci Jojo, pour ta joie de vivre et tes conseils, pour m'avoir écouté râler et pour avoir râler avec moi, pour ces soirées (nuits ?) à discuter devant chez moi, dans le froid et pour m'avoir initiée aux bons vins !

Speaking of bons vins, merci aussi à tout notre petit groupe des Assoiffés, Alex, Déborah, Émilie, Joël, Lucie, Mathilde, Nicolas, Valentin, pour nos IBeerMP et les soirées jeux, pour les encouragements et le réconfort (et le gout de la bonne bière).

Merci aux filles, ces copines là depuis presque toujours, pour leur amour et les week-end de déconnexions, sans lesquels il aurait été plus difficile de garder le moral à certains moments.

Évidemment, merci à ma famille, mes parents, ma sœur, pour leur soutien, leur amour, leurs tupperwares et pour m'avoir permis d'en arriver là. Merci à ma petite Terrine d'avoir été là au début, me donnant une bonne raison de rentrer du labo le soir. Je suis désolée pour tous ces moments câlins que j'ai remis à plus tard parce qu'il était l'heure d'aller travailler et qui n'arriveront finalement plus jamais. A mon Fitzouille, qui m'a permis de finir sans m'effondrer totalement, avec lequel je ne referai plus les mêmes erreurs.

Le plus grand de tous les **MERCIS** à Anne et Seb, mes collègues préférés, qui ont toujours été là. Vous êtes ce qui m'est arrivé de plus beau pendant ma thèse, de mes plus belles amitiés. Même si ma beaufitude et mon humour nul se sont gravement développés à votre contact, au jour d'aujourd'hui et en vrai de vrai, je n'imagine pas meilleur soutien au quotidien que le vôtre, vos câlins, vos messages et vos gratouilles. Je vous aime aussi fort que tâche le rouge ++.

Last but not least, merci à mon Titi, mon mari, mon âme sœur, qui me supporte depuis 13 ans dans la joie et l'alégresse. Merci pour ta patience et surtout ton aide, ton soutien et tes encouragements quotidiens. Merci pour ta confiance en moi et merci de m'avoir épaulée et guidée dans les moments les plus durs. Comme un vrai tournesol, tu es le soleil de ma vie.

TABLE DES MATIERES

| | |
|---|-----------|
| LISTE DES ABREVIATION | 1 |
| LISTE DES UNITES | 5 |
| LISTE DES FIGURES | 7 |
| INTRODUCTION | 13 |
| LES MITOCHONDRIES | 15 |
| <i>I. Origine</i> | 15 |
| <i>II. Morphologie</i> | 19 |
| <i>III. Structure</i> | 21 |
| 1. Membrane externe | 21 |
| 2. Espace intermembranaire | 21 |
| 3. Membrane interne | 23 |
| 4. Matrice | 25 |
| <i>IV. Importation des protéines mitochondriales</i> | 25 |
| <i>V. Fonctions mitochondriales</i> | 29 |
| 1. Production d'énergie et respiration | 29 |
| 2. Voie alternative à la respiration et tolérance aux stress | 31 |
| <i>VI. Génomes mitochondriaux</i> | 33 |
| 1. Taille et contenus des génomes mitochondriaux | 33 |
| 2. Structure et organisation des génomes mitochondriaux | 37 |
| CHLAMYDOMONAS REINHARDTII | 39 |
| <i>I. Organisation cellulaire</i> | 39 |
| <i>II. Cycle cellulaire</i> | 41 |
| <i>III. Génomes</i> | 43 |
| <i>IV. Organisme modèle</i> | 45 |
| EXPRESSION DU GENOME MITOCHONDRIAL DE <i>C. REINHARDTII</i> | 47 |
| <i>I. Génome mitochondrial et réplication</i> | 47 |
| 1. Génome mitochondrial | 47 |
| 2. Réplication | 47 |
| <i>II. Expression du génome mitochondrial</i> | 49 |
| 1. Transcription | 49 |
| 2. Maturation des transcrits | 53 |
| 3. Traduction | 59 |
| 4. Facteurs <i>trans</i> impliqués dans l'expression génétique mitochondriale | 73 |
| PROJET DE THESE | 81 |
| MATERIEL & METHODES | 83 |
| MATERIEL | 85 |

| | |
|--|-----|
| <i>I. Matériel végétal</i> | 85 |
| 1. Souches de <i>Chlamydomonas reinhardtii</i> | 85 |
| 2. <i>Nicotiana benthamiana</i> | 85 |
| <i>II. Souches bactériennes</i> | 87 |
| 1. <i>Escherichia coli</i> | 87 |
| 2. <i>Agrobacterium tumefaciens</i> | 89 |
| <i>III. Vecteurs</i> | 89 |
| 1. pChlamiRNA2 | 89 |
| 2. pGEM®-T Easy Vector System I (Promega) | 91 |
| 3. pDONR207 (Gateway – Invitrogen) | 91 |
| 4. pB7FWG2 (Gateway) | 91 |
| 5. pUPD2 (Golden Braid) | 91 |
| 6. pBACT α 1 (Golden Braid) | 93 |
| <i>IV. Oligonucléotides</i> | 93 |
| <i>V. Gène synthétique</i> | 93 |
| <i>VI. Anticorps</i> | 93 |
| 1. Goat@Mouse | 93 |
| 2. Goat@Rabbit | 93 |
| 3. Mouse@PDH E1 | 95 |
| 4. Rabbit@Cytf | 95 |
| 5. Rabbit@NAB1 | 95 |
| 6. Rabbit@PAP4 | 95 |
| 7. Rabbit@PAP6 | 95 |
| 8. Rabbit@VDAC1 | 95 |
| <i>VII. Outils informatiques</i> | 97 |
| 1. MacVector | 97 |
| 2. Blast (NCBI) | 97 |
| 3. Primer-BLAST | 97 |
| 4. Universal Probe Library (Roche) | 97 |
| 5. LightCycler480® Software (Roche) | 97 |
| 6. LinRegPCR | 99 |
| 7. WMD3 | 99 |
| 8. Agilent 2100 expert software | 99 |
| 9. Phytozome V13 | 99 |
| 10. PredAlgo | 99 |
| 11. TargetP | 99 |
| 12. ZEN lite (Zeiss) | 101 |
| 13. ImageJ | 101 |
| 14. Affinity Designer V1 | 101 |
| <i>VIII. Microscopie</i> | 101 |
| METHODES | 103 |

| | |
|---|-----|
| <i>I. Techniques relatives aux organismes photosynthétiques</i> | 103 |
| 1. Cultures de <i>Chlamydomonas reinhardtii</i> | 103 |
| 2. Quantification des <i>C. reinhardtii</i> | 103 |
| 3. Transformation nucléaire de <i>C. reinhardtii</i> | 105 |
| 4. Cryopréservation de <i>C. reinhardtii</i> | 107 |
| 5. Tests phénotypiques de <i>C. reinhardtii</i> | 109 |
| 6. Purification de mitochondries de <i>C. reinhardtii</i> | 109 |
| 7. Cultures de <i>N. benthamiana</i> | 111 |
| 8. Agro-infiltration de <i>N. benthamiana</i> | 111 |
| 9. Observation de feuilles agro-infiltrées de <i>N. benthamiana</i> | 111 |
| 10. Préparation de protoplastes de <i>N. benthamiana</i> | 113 |
| <i>II. Techniques de cultures et transformation des bactéries</i> | 113 |
| 1. Cultures d' <i>E. coli</i> | 113 |
| 2. Transformation d' <i>E. coli</i> | 113 |
| 3. Cultures d' <i>A. tumefaciens</i> | 113 |
| 4. Transformation d' <i>A. tumefaciens</i> | 115 |
| 5. Glycérol-stocks de bactéries..... | 115 |
| <i>III. Techniques liées à l'ADN</i> | 115 |
| 1. Extraction d'ADN brut de <i>C. reinhardtii</i> pour PCR de génotypage..... | 115 |
| 2. Amplifications par PCR | 115 |
| 3. Électrophorèse sur gel d'agarose ADN..... | 117 |
| 4. Purification des produits de PCR | 119 |
| 5. Purification de l'ADN au phénol-chloroforme..... | 119 |
| 6. Précipitation de l'ADN | 119 |
| 7. Quantification des acides nucléiques..... | 119 |
| 8. Clonage avec le système « pGEM T-Easy [®] vector »..... | 119 |
| 9. Clonage avec le système Gateway | 121 |
| 10. Clonage avec le système GoldenBraid | 121 |
| 11. Digestions enzymatiques..... | 123 |
| 12. Purification des plasmides..... | 125 |
| 13. Séquençage | 125 |
| 14. Constructions des micro RNA artificiels pour <i>C. reinhardtii</i> | 125 |
| <i>IV. Techniques liées à l'ARN</i> | 131 |
| 1. Extraction d'ARN total de <i>C. reinhardtii</i> | 131 |
| 2. Électrophorèse sur gel d'agarose ARN | 131 |
| 3. Traitement DNase | 131 |
| 4. Purification au phénol-chloroforme | 131 |
| 5. Précipitation | 131 |
| 6. Quantification..... | 133 |
| 7. Rétrotranscription | 133 |
| 8. RT-qPCR..... | 133 |
| 9. 3'RACE-seq | 135 |

| | |
|--|------------|
| V. Techniques liées aux protéines..... | 147 |
| 1. Quantification au Bradford..... | 147 |
| 2. Gel de protéines..... | 147 |
| 3. Western blot..... | 149 |
| 4. Production de protéines recombinantes..... | 151 |
| 5. Test d'activités..... | 153 |
| RESULTATS..... | 155 |
| CHAPITRE I : ÉTUDE DE DEUX NUCLEOTIDYL-TRANSFERASES : PAP4 ET PAP6 : ACTRICES DE LA POLYCYTIDYLATION ?..... | 157 |
| I. Mise en contexte..... | 157 |
| 1. Analyses des extrémités 3' des ARNm mitochondriaux..... | 157 |
| 2. Les Nucleotidyl-transférases : à l'origine de la polycytidylation ?..... | 159 |
| II. PAP4 et PAP6 sont adressées aux mitochondries..... | 165 |
| 1. Colocalisation de PAP4 et PAP6 avec les mitochondries..... | 165 |
| 2. Essais d'expression <i>in alga</i> | 169 |
| 3. Détection de PAP4 et PAP6 dans des mitochondries purifiées..... | 171 |
| 4. Conclusion..... | 175 |
| III. PAP4, impliquée dans la polycytidylation ?..... | 175 |
| 1. Architecture et classification..... | 175 |
| 2. Localisation uniquement mitochondriale de PAP4..... | 177 |
| 3. Obtention de lignées mutantes..... | 179 |
| 4. État des mitochondries dans les mutants de PAP4..... | 189 |
| 5. Il y a moins de queues polyC dans les mutants de PAP4..... | 193 |
| IV. PAP6, impliquée dans l'uridylation des ARNm mitochondriaux..... | 209 |
| 1. Architecture de PAP6..... | 209 |
| 2. Localisation uniquement mitochondriale de PAP6..... | 209 |
| 3. Obtention de lignées mutantes..... | 211 |
| 4. Caractérisation des mutants de PAP6..... | 213 |
| 5. Queues polyU-riche affectées dans les mutants de PAP6..... | 215 |
| 6. Activité <i>in vitro</i> de PAP6..... | 219 |
| CHAPITRE II : CARACTERISATION DU MITORIBOSOME DE <i>C. REINHARDTII</i> | 225 |
| I. Mise en contexte et résultats complémentaires à l'article..... | 225 |
| II. Article : <i>How to build a ribosome from RNA fragments in Chlamydomonas mitochondria</i> | 231 |
| DISCUSSION & PERSPECTIVES..... | 249 |
| VERS LA CARACTERISATION DES AJOUTS POST-TRANSCRIPTIONNELS DES TRANSCRITS MITOCHONDRIAUX DE <i>C. REINHARDTII</i> | 251 |
| I. PAP4 et la polycytidylation pour la stabilisation des transcrits ?..... | 253 |
| II. PAP6 et polyuridylation pour la dégradation des transcrits ?..... | 261 |
| III. Qu'en est-il de la polyadénylation des transcrits ?..... | 265 |
| IV. Les limites du 3'RACE-seq..... | 267 |
| CARACTERISATION DU MITORIBOSOME DE <i>C. REINHARDTII</i> | 273 |

| | |
|--|------------|
| BIBLIOGRAPHIE | 281 |
| ANNEXES..... | 307 |
| <i>I. Mutants CLiP</i> | <i>309</i> |
| <i>II. Oligonucléotides.....</i> | <i>310</i> |
| 1. PCR | 310 |
| 2. Séquençage | 312 |
| 3. Constructions pour les mutants <i>amiRNA</i> | 313 |
| 4. RT..... | 314 |
| 5. RT-qPCR..... | 314 |
| 6. 3'RACE-seq | 316 |
| <i>III. Anticorps</i> | <i>318</i> |
| <i>IV. Milieux et solutions</i> | <i>318</i> |
| 1. Milieux et solutions relatifs aux organismes | 318 |
| 2. Milieux et solutions relatifs aux ADN | 319 |
| 3. Milieux et solutions relatifs aux ARN..... | 319 |
| 4. Milieux et solutions relatifs aux protéines | 319 |
| <i>V. Modules Golden Braid</i> | <i>321</i> |
| <i>VI. Activités in vitro.....</i> | <i>321</i> |
| <i>VII. Données brutes des analyses de spectrométrie de masse.....</i> | <i>321</i> |
| <i>VIII. Données de 3'RACE-seq.....</i> | <i>325</i> |
| 1. Données des triplicats biologiques..... | 325 |
| 2. Dé-zoom des graphiques présentés dans les Résultats..... | 328 |
| <i>IX. Comparaison des données de 3'RACE et 3'RACE-seq</i> | <i>331</i> |
| <i>X. Pourcentages de différences – 3'RACE-seq</i> | <i>332</i> |
| <i>XI. Distribution des nt selon la position – 3'RACE-seq</i> | <i>333</i> |

LISTE DES ABREVIATION

| | | | |
|------------|--|----------------|--|
| @ | anti | CCAses | <i>CCA-adding enzymes</i> |
| 3'RACE-seq | <i>3' Rapid Amplification of cDNA Ends coupled with sequencing</i> | CCD | <i>core catalytic domain</i> |
| aa | acides aminés | CCM | <i>Carbon Concentrating Mechanism</i> |
| aaRS | <i>aminoacyl-tRNA synthetase</i> | CDS | <i>coding DNA sequence</i> |
| ADNc | ADN complémentaire | CLiP | <i>Chlamydomonas Library Project</i> |
| ADNdb | ADN double brin | cryo-EM | <i>cryo-electron microscopy</i> |
| ADNg | ADN génomique | C _t | <i>threshold cycle</i> |
| ADNk | ADN kinétoplaste | Cytf | <i>Cytochrome f</i> |
| ADNmt | ADN mitochondrial | dNTP | désoxyribonucléotide |
| ADNsb | ADN simple brin | <i>dum</i> | <i>Dark Uniparental transmission by the Minus parent</i> |
| ADP | Adénosine DiPhosphate | EGFP | <i>Enhanced Green Fluorescent Protein</i> |
| AEG | analyse de l'expression génique | EIM | espace intermembranaire |
| amiRNA | <i>artificial micro RNA</i> | FECA | <i>First Eukaryote Common Ancestor</i> |
| ANT | <i>Adenine Nucleotide Translocator</i> | IBMC | Institut de Biologie Moléculaire et Cellulaire |
| AOX | <i>Alternative Oxidase</i> | IBMP | Institut de Biologie Moléculaire des Plantes |
| APS | persulfate d'ammonium | IBPC | Institut de Biologie Physico-Chimique |
| ARNm | ARN messenger | IECB | Institut Européen de Chimie et Biologie |
| ARNr | ARN ribosomique | IF | <i>initiation factor</i> |
| ARNt | ARN de transfert | KD | <i>knock-down</i> |
| ATP | Adénosine TriPhosphate | KO | <i>knock-out</i> |
| BET | bromure d'éthidium | LB | <i>lysogeny broth</i> |
| BF | <i>flexible loop</i> | | |
| BLAST | <i>Basic Local Alignment Search Tool</i> | | |
| BSA | <i>bovin serum albumin</i> | | |
| CB | code barre | | |

| | | | |
|--------------|---|--------------|---|
| LECA | <i>Last Eukaryote Common Ancestor</i> | NRM | <i>nucleotide recognition motif</i> |
| LPCV | <i>Laboratoire Physiologie Cellulaire et Végétale</i> | NT | <i>Nucleotidyl Transferase</i> |
| LSU | <i>Large SubUnit</i> | NTD | <i>NT domain</i> |
| LTU | <i>Leftward Transcriptional Unit</i> | Oligos | <i>oligonucléotides</i> |
| MIA | <i>Mitochondrial Intermembrane space Assembly</i> | OPR | <i>Octatricopeptide Repeat</i> |
| miARN | <i>micro ARN</i> | ORF | <i>open reading frame</i> |
| mIF | <i>mitochondrial Initiation Factor</i> | P3P | <i>Purification et Production des Protéines</i> |
| Mim1 | <i>Mitochondrial importation 1</i> | PAM | <i>Presequence translocase-Associated Motor</i> |
| Mitoribosome | <i>ribosome mitochondrial</i> | PAP | <i>PolyA Polymerase</i> |
| MPP | <i>Mitochondrial Processing Peptidase</i> | PAPD | <i>PAP associated domain</i> |
| MS | <i>mass spectrometry</i> | PARN | <i>poly(A) specific ribonuclease</i> |
| Mt | <i>mating type</i> | PCR | <i>Polymerase Chain Reaction</i> |
| mtEFG | <i>mitochondrial Elongation Factor G</i> | PDH E1 | <i>pyruvate dehydrogenase E1</i> |
| mtEFTs | <i>mitochondrial Elongation Factor Ts</i> | PNPase | <i>polynucleotide phosphorylase</i> |
| mtEFTu | <i>mitochondrial Elongation Factor Tu</i> | <i>pol β</i> | <i>polymérase β</i> |
| mTERF | <i>mitochondrial Transcription Termination Factor</i> | PP | <i>purification protéines</i> |
| MTS | <i>Mitochondrial Targeting Sequence</i> | PPR | <i>pentatricopeptide repeat</i> |
| NAB1 | <i>Nucleic acid binding protein 1</i> | PRORP | <i>protein-only RNase P</i> |
| nanoLC-MS/MS | <i>nano-liquid chromatography coupled to tandem mass spectrometry</i> | RAP | <i>RNA binding abundant in Apicomplexa</i> |
| NCBI | <i>National Center for Biotechnology Information</i> | RBD | <i>RNA binding domain</i> |
| | | RBS | <i>ribosome binding site</i> |
| | | RFP | <i>Red Fluorescent Protein</i> |
| | | rNTP | <i>ribonucléotides</i> |
| | | ROS | <i>Reactive Oxygen Species</i> |
| | | RPI | <i>Read Primer Index</i> |

| | | | |
|---------|---|--------|--|
| RT | retrotranscription | Tm | <i>melting temperature</i> |
| RT-qPCR | <i>Reverse Transcription Quantitative PCR</i> | TOM | <i>Translocase of the Outer Membrane</i> |
| RTU | <i>Rightward Transcriptional Unit</i> | TUTase | <i>terminal uridylylate transferases</i> |
| SAM | <i>Sorting and Assembly Machinery</i> | UTR | <i>Untranslated Transcribed Region</i> |
| SSU | <i>Small SubUnit</i> | UV | ultraviolet |
| TAP | Tris Acétate Phosphate | VDAC | <i>Voltage Dependent Anion Channel</i> |
| TEMED | tétraméthyléthylène- diamine | Vmax | vitesse maximale |
| Ti | <i>Tumor induced</i> | WB | western blot |
| TIM | <i>Translocase of the Inner Membrane</i> | WT | <i>wild type</i> |

LISTE DES UNITES

| | | | |
|-----------------|--|--------------|--|
| °C | degrés Celsius | nt | nucléotides |
| Cfu/μg | <i>colony forming unit</i> par microgrammes d'ADN | p/v | poids à volume |
| cm ² | centimètre carré | pb | paires de bases |
| DO | densité optique | pH | potential hydrogène |
| g | gramme | pmol | picomoles |
| h | heure | rpm | rotation par minute |
| j | jour | s | seconde |
| kb | kilobase | U/μL | unités par microlitre |
| kDa | kiloDalton | V | Volt |
| mA | milliAmpère | v | volume |
| Mb | Megabase | v/v | volume à volume |
| Min | minute | μg | microgramme |
| mL | millilitre | μg/mL | microgramme par millilitre |
| mm | millimètre | μL | microlitre |
| mM | millimolaire | μm | micromètre |
| ng | nanogramme | μM | micromolaire |
| nm | nanomètre | μmol.m-2.s-1 | Densité de Flux Photonique Photosynthétique |
| nM | nanomolaire | | |

LISTE DES FIGURES

Introduction

Figure I-1 : La théorie endosymbiotique expliquant l'origine des organelles

Figure I-2 : Le modèle endosymbiotique E3

Figure I-3 : Structure de la mitochondrie

Figure I-4 : Deux types d'importation des protéines dans la mitochondrie

Figure I-5 : L'importation post-traductionnelle de protéines dans la mitochondrie de levure

Figure I-6 : La mitochondrie et la production d'énergie

Figure I-7 : La voie alternative à la respiration

Figure I-8 : Une diversité de génomes mitochondriaux

Figure I-9 : Arbre phylogénétique des Archaeplastidés

Figure I-10 : Structure d'une cellule végétative de *Chlamydomonas reinhardtii*

Figure I-11 : Structure en 3D d'une cellule végétative de *C. reinhardtii*

Figure I-12 : Cycle cellulaire de *C. reinhardtii*

Figure I-13 : Comparaison des génomes nucléaires de 3 espèces

Figure I-14 : Génome mitochondrial de *C. reinhardtii*

Figure I-15 : Modèle de réplication de l'ADN selon Georg Michaelis

Figure I-16 : Génome mitochondrial de *C. reinhardtii* et clivages des ARNm

Figure I-17 : Clivages post-transcriptionnels des transcrits de gènes de protéines

Figure I-18 : Plusieurs types de queues sont retrouvés dans les mitochondries

Figure I-19 : Diversité des ribosomes mitochondriaux

Figure I-20 : Structure des PPR en hélices α et code PPR

Figure I-21 : Différents types de domaines PPR et leur séquence consensus

Figure I-22 : Protéines OPR et leurs séquences consensus

Matériel & méthodes

Figure M-1 : Schémas récapitulatifs des différentes amorces utilisés pour le 3'RACE-seq

Figure M-2 : Détails du protocole et des résultats de la PCR1

Figure M-3 : Gels d'électrophorèse obtenus après les PCR2

Figure M-4 : Résultats d'analyses des pool/gènes au BioAnalyzer

Figure M-5 : Traitement bio-informatique des données de 3'RACE-seq

Figure M-6 : Vue d'ensemble des analyses de 3'RACE-seq

Résultats

Figure R-1 : Exemples de queues obtenues par des expériences de 3'RACE

Figure R-2 : Proportions des différentes queues ajoutées pour chaque transcrit

Figure R-3 : Classification des NT

Figure R-4 : La préférence des NT pour leur substrat dépend d'un acide aminé

Figure R-5 : Structure et spécificité des CCAsés

Figure R-6 : Arbre phylogénique des NT des Chlorophytes

Figure R-7 : Analyse phylogénique des NT de *C. reinhardtii*

Figure R-8 : Séquences des protéines candidates et régions clonées

Figure R-9 : Localisation subcellulaire de PAP2, PAP4, PAP6 et PAP10

Figure R-10 : Résultats des analyses de spectrométrie de masse

Figure R-11 : Résultats des analyses de protéomique des fractions matricielles et membranaires

Figure R-12 : PAP4 et PAP6 identifiées dans les extraits mitochondriaux

Figure R-13 : Structure de la protéine PAP4

Figure R-14 : PAP4, une CCAsé non canonique

Figure R-15 : Partie de PAP4 utilisée pour la production d'anticorps

Figure R-16 : PAP4 localisée dans les mitochondries en Western Blot

Figure R-17 : Cassette insérée dans la banque de mutants CLiP

Figure R-18 : Vérification de l'insertion de la cassette dans les mutants CLiP

Figure R-19 : Amplification des jonctions gène - cassette dans les mutants CLiP

Figure R-20 : Les mutants d'insertion PAP4 sont des mutants *knock-down*

Figure R-21 : Vérification de l'expression du transcrit du gène *PAP4* dans les mutants CLiP

Figure R-22 : Génération de mutants amiRNA de PAP4

Figure R-23 : Analyses des transcrits de la souche contrôle pHUN

Figure R-24 : Vérification de l'expression du transcrit du gène *PAP4* dans les mutants amiRNA

Figure R-25 : Le mutant *amipap4-37* est un mutant *knock-down*

Figure R-26 : Tests phénotypiques des mutants de PAP4

Figure R-27 : Analyses de l'expression des transcrits mitochondriaux des mutants de PAP4

Figure R-28 : Effets de la mutation sur les ARNm mitochondriaux selon leur position

Figure R-29 : Les résultats des triplicats sont similaires pour les mutants de PAP4

Figure R-30 : Comparaison des 3 catégories de queues ajoutées entre les mutants CLiP de PAP4 et la souche WT CC-5325

Figure R-31 : Le nombre de queues ajoutées selon leur longueur

Figure R-32 : Distribution des nt différentes dans les mutants, pour certains transcrits

Figure R-33 : Comparaison des 3 catégories de queues ajoutées entre le mutant amiRNA de PAP4 et la souche contrôle pHUN

Figure R-34 : Le nombre de queues ajoutées selon leur longueur

Figure R-35 : Distribution différente des nt dans le mutant amiRNA de PAP4, pour l'ARNm *nad5*

Figure R-36 : Effets de la mutation de PAP4 sur les ARNm mitochondriaux

Figure R-37 : Domaines de la protéine PAP6

Figure R-38 : Protéine PAP6 recombinante pour la production des anticorps polyclonaux

Figure R-39 : La protéine PAP6 est localisée uniquement dans les mitochondries

Figure R-40 : Expression relative du transcrit *PAP6* dans les mutants amiRNA

Figure R-41 : Les mutants amiRNA PAP6 sont des mutants *knock-down*

Figure R-42 : Tests phénotypiques des mutants de PAP6

Figure R-43 : Analyses de l'expression relative des transcrits mitochondriaux des mutants PAP6

Figure R-44 : Effets de la mutation de PAP6 sur les ARNm mitochondriaux selon leur position

Figure R-45 : Les résultats des triplicats sont similaires pour les mutants amiRNA de PAP6

Figure R-46 : Comparaison des 3 catégories de queues ajoutées entre les mutants amiRNA de PAP6 et la souche contrôle pHUN

Figure R-47 : La longueur des queues ajoutées varie dans les mutants

Figure R-48 : Distribution des nt différentes dans les mutants, pour certains transcrits

Figure R-49 : Effets de la mutation de PAP6 sur les ARNm mitochondriaux

Figure R-50 : Production des protéines recombinantes PAP6 et PAP6 mutée

Figure R-51 : Test d'activité *in vitro* de PAP6

Figure R-52 : Mutants d'insertion des protéines mitoribosomales

Figure R-53 : Mitoribosome de *Polytomella magna*

Discussion & perspectives

Figure D-1 : Modèle des ajouts de queues sur les ARNm dans les mitochondries de *C. reinhardtii*

INTRODUCTION

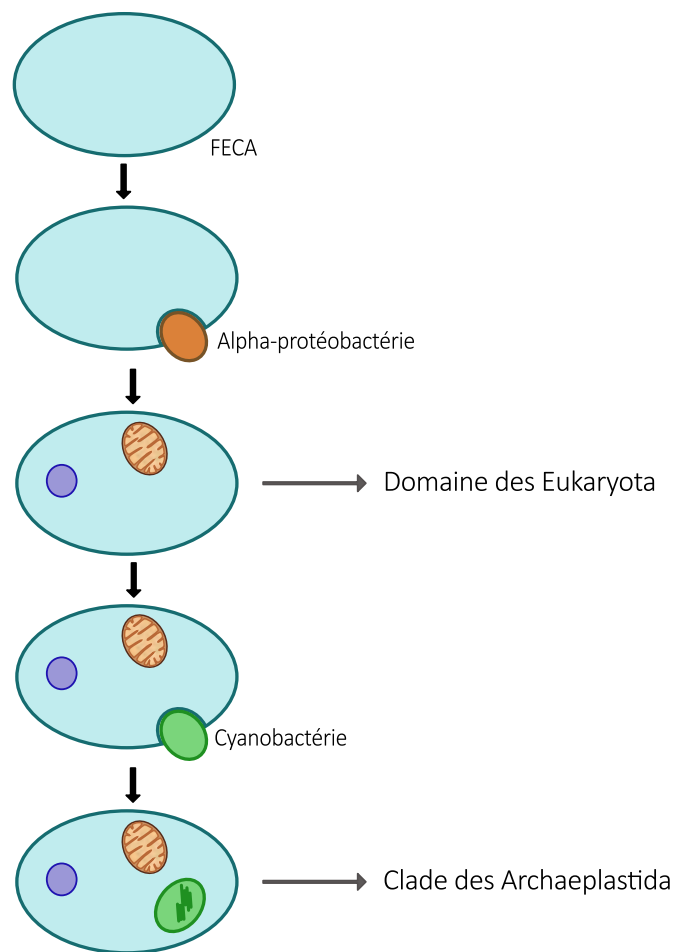


Figure I-1 : La théorie endosymbiotique expliquant l'origine des organelles

Schéma représentant l'acquisition des organites dans l'évolution des eucaryotes.

L'endosymbiose entre le FECA (*First Eukaryote Common Ancestor* - premier ancêtre commun aux eucaryotes) et une α -protéobactérie donnera les mitochondries et l'apparition du domaine des Eukaryota (eucaryotes). Plus tard, une cellule eucaryote endocytera une cyanobactérie, en résulteront les chloroplastes et le clade des Archaeplastida.

Dans les cellules : **bleu foncé** : noyau cellulaire ; **brun** : mitochondrie ; **vert** : chloroplaste.

LES MITOCHONDRIES

I. Origine

L'origine des organites (mitochondries et chloroplastes) est expliquée par la théorie endosymbiotique (Figure I-1). Il y a quelques 2 milliards d'années, une première endosymbiose se fait. Un organisme proto-eucaryote, considéré comme étant le premier ancêtre commun aux eucaryotes, ou le FECA (*First Eukaryote Common Ancestor*) aurait endocyté une α -protéobactérie, permettant l'acquisition des mitochondries, organites conservés dans presque tous les eucaryotes (Lang, Gray and Burger, 1999). L'organisme en résultant est considéré comme étant le dernier ancêtre commun des eucaryotes, le LECA (*Last Eukaryote Common Ancestor*) (Gray, 2015). Un demi-milliard d'années plus tard se fera le second évènement d'endosymbiose, lorsqu'une cellule eucaryote primitive endocyte une cyanobactérie, donnant les chloroplastes et amenant à l'apparition du clade des Archaeplastida, comprenant les Rhodophytes (les algues rouges), les Glaucophytes (les algues unicellulaire vert/bleu) et les Viridiplantae (les algues vertes et les plantes terrestres) (Dorrell and Howe, 2012).

L'évènement d'acquisition des mitochondries semble corrélér à l'augmentation du taux d'oxygène dans les océans et dans l'atmosphère (Lyons, Reinhard and Planavsky, 2014). La symbiose avec l' α -protéobactérie capable de respiration cellulaire aurait permis l'adaptation de la cellule hôte aux conditions aérobiques.

Concernant l'acquisition des mitochondries, la communauté scientifique s'accorde sur la nature de l'organisme endocyté. Des études phylogéniques indiquent que les mitochondries ont pour origine une α -protéobactérie de l'ordre des *Rickettsiales*, un groupe de bactéries dans lequel se trouvent de nombreuses espèces parasites intracellulaires obligatoires (Lang, Gray and Burger, 1999 ; Wang and Wu, 2015). Cela laisse à penser que la mitochondrie est issue d'un parasitisme de la cellule hôte par l' α -protéobactérie, qui a par la suite évolué en symbiose (Sassera *et al.*, 2011).

Par contre, des doutes subsistent concernant la nature exacte de la cellule hôte et à quel moment s'est faite l'acquisition de la mitochondrie durant l'eucaryogenèse. L'eucaryogenèse est le processus qui a donné lieu à l'apparition des eucaryotes. L'acquisition de la mitochondrie

se situe entre la formation du FECA et celle du LECA. Il existe 2 théories, selon si la mitochondrie a été acquise tôt ou tard lors de l'eucaryogenèse (Poole and Gribaldo, 2014 ; Gabaldón, 2018). La théorie « *mito-late* » considère que les bactéries ou les archées ne sont pas capables de phagocytose, donc la cellule hôte serait forcément une cellule ayant déjà acquis une certaine complexité, ce serait donc une cellule eucaryote primitive. Dans ce cas, l'apparition des mitochondries est la dernière étape de l'eucaryogenèse permettant la formation du LECA (Cavalier-Smith, 2002).

En opposition, la théorie « *mito-early* » considère que l'acquisition des mitochondries est un évènement survenu tôt dans l'eucaryogenèse, peut-être même comme la première étape, après quoi sont apparus les autres éléments et structures de la cellule eucaryote : membrane nucléaire, cytosquelette, etc. Dans ce cas, la cellule hôte serait une bactérie ou une archée. Cette théorie *mito-early* est la plus acceptée par la communauté scientifique aujourd'hui. Déjà dans les années 1980/1990, des analyses phylogéniques montrent que les archées et les eucaryotes sont des groupes proches (« *Archaeobacteria–Eukaryota sisterhood* » (Philippe and Forterre, 1999)). Ces résultats ont été renforcés par la découverte d'un groupe d'archées, les *Asgardarchaeota* (Zaremba-Niedzwiedzka *et al.*, 2017). Des études phylogéniques ont montré des points communs avec les eucaryotes, notamment les profilines. Ce sont des protéines régulatrices des filaments d'actine, suggérant les débuts du cytosquelette (Akil and Robinson, 2018). Elles sont également importantes pour le processus d'endocytose, processus au cœur de la théorie de l'endosymbiose. Toutefois, une étude suggère que les archées *Asgard* ne soient pas capables de phagocytose (Burns, Pittis and Kim, 2018). Dans cette étude des analyses *in silico* ont permis de construire un modèle permettant de discriminer les organismes selon leur capacité à faire de la phagocytose ou non. Les résultats montrent que les archées *Asgard* ne possèdent pas les protéines nécessaires au processus de phagocytose. Ces protéines seraient issues d'une combinaison de composants bactériens et archéens, supplémentées d'innovations eucaryotiques. Mais plus récemment une étude a proposé une nouvelle hypothèse : l'eucaryogenèse progressive (Imachi *et al.*, 2020). Pour la première fois, des archées *Asgard* appelées *Candidatus Prometheoarchaeum syntrophicum* ont pu être isolées à partir de sédiments marins et étudiées. Ce sont de petits organismes (environ 550 nm de diamètre), anaérobies et en relation syntrophique avec d'autres archées du genre *Methanogenium*. Les membranes de ces archées font des protubérances membranaires de

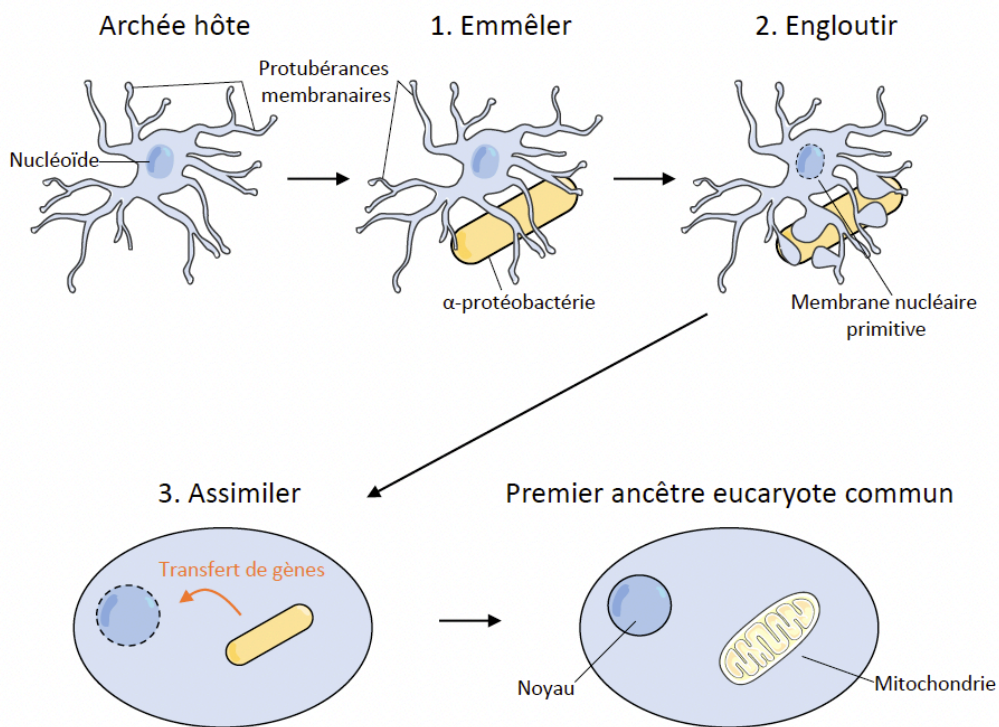


Figure I-2 : Le modèle endosymbiotique E3

De (Schatz-Daas, 2021).

Schéma de l'endosymbiose d'une α -protéobactérie par une archée de la famille *Asgard*, selon le modèle E3 (Imachi *et al.*, 2020). L'archée établit un lien avec la bactérie en l'emmêlant dans ses protubérances membranaires jusqu'à l'engloutir progressivement et l'assimiler. S'effectue alors un transfert de gènes de la bactérie vers le génome de l'archée et une apparition progressive du noyau, afin de donner le LECA.

80 à 100 nm. Ce type de structure aurait pu rentrer en contact avec l' α -protéobactérie, et l'endocytose aurait pu avoir eu lieu progressivement selon le modèle *E3*, en 3 étapes : *Entangle* – *Engulf* – *Endogenize*, ou Emmêler – Engloutir – Assimiler (Figure I-2). Dans ce modèle, la cellule proto-eucaryote aurait d'abord entouré l' α -protéobactérie grâce à ses protubérances membranaires et cette dernière aurait permis ainsi à la cellule hôte de faire la transition aux conditions aérobiques. La relation syntrophique entre les 2 cellules aurait abouti en une endocytose suivie d'une phase d'adaptation de la cellule hôte, par la perte de gènes redondants par exemple, ou le transfert de gènes du génome de la cellule endocytée à celui de la cellule hôte. Il en résulte donc des mitochondries dont les génomes ne codent plus que pour un nombre limité de gènes, et qui sont dépendantes du génome nucléaire pour l'expression de la majorité des protéines impliquées dans leurs fonctions et métabolismes.

Aujourd'hui, les mitochondries sont des organites présents chez la grande majorité des eucaryotes. *Monocercomonoides*, un genre de flagellés du règne des *Excavata*, est le premier genre des eucaryotes où l'absence de mitochondrie a été observée (Karnkowska *et al.*, 2016). Cette absence serait due à une perte durant l'évolution des mitochondries dans ces organismes

II. Morphologie

La forme des mitochondries varie selon les organismes. Par exemple, chez la levure *Saccharomyces cerevisiae* elles sont de forme tubulaires, tandis que chez *Arabidopsis thaliana* elles sont sphériques ou allongées et chez les mammifères, toutes sortes de morphologies sont observables (Logan, 2006). De fait, leur taille varie également. Par exemple chez les plantes, elles ont généralement une taille d'environ 1 μm , tandis que celle-ci varie de 2 à 8 μm chez les cellules animales.

Leur nombre dans chaque cellule varie selon les organismes, le type cellulaire et son besoin en énergie, mais aussi selon le cycle cellulaire, allant d'une mitochondrie à plus de 10^5 (D'Erchia *et al.*, 2015 ; Cole, 2016). En moyenne, les eucaryotes multicellulaires ont une à plusieurs centaines de mitochondries. Chez l'algue verte *Chlamydomonas reinhardtii*, il a été montré qu'il y a une grande variabilité de la forme et du nombre de mitochondries au cours des cycles cellulaires et circadiens, avec un passage d'une mitochondrie unique à un réseau de mitochondries (Osafune *et al.*, 1976 ; Ehara, Osafune and Hase, 1995).

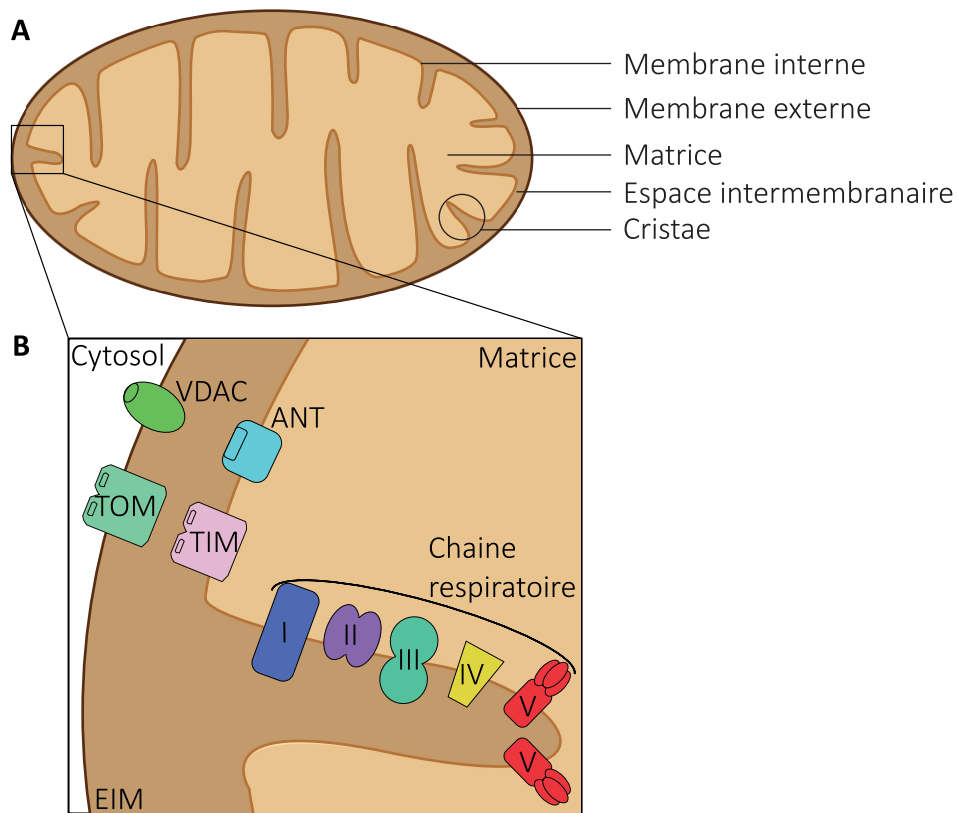


Figure I-3 : Structure de la mitochondrie

A : Schéma de la structure générale de l'organite, avec les différents composants légendés ;
B : Schéma agrandi des membranes, de l'espace intermembranaire et de la matrice avec certains complexes protéiques qui y sont ancrés.

III. Structure

La structure des mitochondries est similaire pour tous les organismes (Figure I-3A). Elles sont composées de 2 membranes, la membrane externe et la membrane mitochondriale interne. Celles-ci délimitent l'espace intermembranaire et la matrice.

1. Membrane externe

La membrane externe des mitochondries est composée de protéines et de phospholipides, avec un ratio de 1:1 en masse (Walther and Rapaport, 2009). Elle est poreuse et perméable aux ions et à des petits métabolites inférieurs à 5 kDa grâce notamment à la présence de porines, des protéines transmembranaires formant des canaux. Il existe des canaux ioniques spécifiques pour le passage de certains ions (Ca^{2+} , Na^+ , K^+ , etc.). La perméabilité de ces canaux ioniques peut être stimulée de manière chimio-dépendante, c'est-à-dire grâce à la fixation d'un ligand. Cette stimulation peut aussi se faire de façon voltage-dépendante, par une sensibilité au potentiel membranaire. Ce dernier type de canaux, aussi appelés VDAC (*Voltage Dependent Anion Channel*), régulent le flux métabolique et énergétique en permettant le passage d'ions, de métabolites, de nucléotides ou de petites molécules. Cependant, ils sont imperméables aux ions H^+ (Zizi *et al.*, 1995; Hodge and Colombini, 1997; Colombini, 2016). C'est notamment par ces VDAC que passe le pyruvate, molécule cruciale pour la respiration mitochondriale, afin d'entrer dans la mitochondrie. Chez les plantes, il a été montré que les VDAC permettent aussi le transport des ARNt (ARN de transfert) (Salinas *et al.*, 2006). Ces protéines sont les plus abondantes de la membrane externe, représentant 34% du total de la surface de la membrane externe (Fuchs *et al.*, 2020). L'importation de la majorité des protéines mitochondriales se fait par le complexe TOM (*Translocase of the Outer Membrane*) par un mécanisme qui sera développé plus loin.

2. Espace intermembranaire

L'EIM (espace intermembranaire) est délimité par les membranes externes et internes. La membrane externe étant imperméable aux ions H^+ , l'EIM est saturé en protons, produits de l'activité de la chaîne respiratoire, afin de produire de l'ATP (Adénosine TriPhosphate) par le complexe V de la chaîne respiratoire, l'ATP synthase. Cet EIM contient également des protéines, notamment impliquées dans la biogenèse et l'assemblage des différents complexes respiratoires ou dans la translocation des protéines passant par les complexes TOM/TIM

(*Translocase of the Inner Membrane*). La majorité des protéines de l'EIM sont petites (de 7 à 25 kDa chez l'Homme et la levure) et ont une structure en hélices α . Elles possèdent aussi des cystéines formant des ponts disulfures entre les 2 hélices α , ce qui leur confère une structure en hélice-coude-hélice (*helix-turn-helix*) (Backes and Herrmann, 2017; Edwards, Eaglesfield and Tokatlidis, 2021).

3. Membrane interne

La membrane interne entoure la matrice. Elle forme des replis membranaires, appelés crêtes ou *cristae*, grâce aux dimères d'ATP synthase (Miranda-Astudillo *et al.*, 2022). Ces invaginations sont dirigées vers la matrice. Elles permettent d'augmenter la surface de membrane interne et d'espace intermembranaire (Griparic and Van Der Blik, 2001). Ces 2 lieux sont essentiels pour la phosphorylation oxydative, qui est le rôle principal de la mitochondrie (O'Rourke, 2007). Ces structures caractéristiques sont facilement observables en microscopie électronique. La membrane interne est relativement imperméable : il n'y a que le dioxygène, le dioxyde de carbone et l'eau qui peuvent passer librement. Ceci est essentiel pour le bon fonctionnement de la chaîne de transport des électrons, dans l'espace intermembranaire, lorsque les protons H^+ doivent être pompés depuis la matrice vers l'espace intermembranaire.

Elle est composée à 75% de protéines et 25% de lipides. Parmi les 25% de lipides, 20% sont des cardiolipines, un phospholipide retrouvé spécifiquement dans des membranes couplant phosphorylation et transport d'électrons : membranes plasmiques bactériennes, chromatophores, mitochondries ou chloroplastes (Hoch, 1992). Ce cardiolipide contribue fortement à l'imperméabilité de la membrane interne.

Au niveau protéique, c'est dans la membrane interne que sont localisés tous les complexes protéiques nécessaires pour la phosphorylation oxydative, c'est-à-dire les 4 complexes de la chaîne respiratoire (Figure I-3B, complexes I, II, III et IV). Il y a aussi l'ATP synthase, aussi appelée complexe V (Figure I-3B), et de nombreux transporteurs permettant le passage d'acides gras, d'acides aminés, de l'ATP et de l'ADP (Adénosine DiPhosphate) (Ferramosca and Zara, 2021). On trouve notamment le transporteur ANT (*Adenine Nucleotide Translocator*), protéine majoritaire de la membrane interne qui permet le transport de l'ATP et l'ADP (Klingenberg, 1989). Il y a également les protéines du complexe TIM (*Translocase of the Inner Membrane*), qui, avec le complexe TOM, permettent l'importation de protéines du cytosol à la matrice.

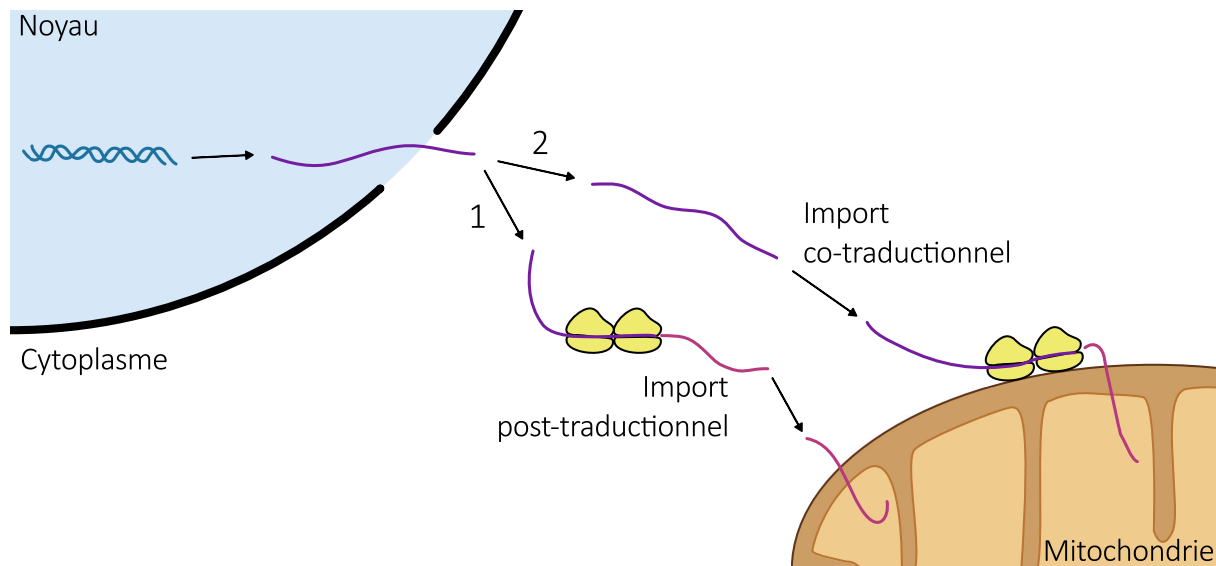


Figure I-4 : Deux types d'importation des protéines dans la mitochondrie

Schéma des 2 types d'imports des protéines dans la mitochondrie. L'ADN (**bleu**) est transcrit en ARNm (ARN messenger) (**violet**) dans le noyau. Une fois dans le cytosol, ces ARNm ont 2 voies possibles. **1** : ils peuvent être traduits en protéines par les ribosomes cytosoliques (**jaune**) puis importés dans la mitochondrie, on dit que l'import est post-traductionnel ; **2** : les ARNm peuvent aussi être adressé à la surface des mitochondries où se trouvent des ribosomes cytosoliques. Là, ils seront traduits en protéines (**rose**) et importés immédiatement dans l'organe. Ce processus s'appelle l'import co-traductionnel.

4. Matrice

La matrice contient le génome mitochondrial (aussi appelé ADNmt, pour ADN mitochondrial) ainsi que la machinerie nécessaire à son expression : protéines et enzymes impliquées entre autres dans la transcription, les étapes de maturation et de traduction. Chez la levure, l'Homme, *A. thaliana* ou *C. reinhardtii*, les mitoribosomes (ribosomes mitochondriaux), situés dans la matrice, sont en contact avec la membrane interne de la mitochondrie (Waltz *et al.*, 2021).

C'est également là que se font la conversion du pyruvate en acétyl-CoA, le Cycle de Krebs, la production d'ATP et d'autres métabolites utilisés pour la phosphorylation oxydative. Il a été montré que chez *A. thaliana* 80% des protéines matricielles sont impliquées dans le Cycle de Krebs (Fuchs *et al.*, 2020).

IV. Importation des protéines mitochondriales

Durant l'évolution, la grande majorité des gènes mitochondriaux ancestraux ont été perdus ou transférés au noyau. Par exemple, 33 protéines sont codées par le génome d'*A. thaliana*, tandis que 2934 protéines différentes ont été identifiées lors d'analyses protéomiques des mitochondries de cette plante (Fuchs *et al.*, 2020). La grande majorité des protéines est donc codée par le génome nucléaire, puis elles sont importées dans les mitochondries, selon 2 mécanismes différents : les importations post- et co-traductionnelles (Figure I-4).

L'importation post-traductionnelle est la plus étudiée à ce jour (Figure I-4-1). Dans ce cas, les gènes codés par le génome nucléaire sont transcrits dans le noyau et traduits en protéines dans le cytosol. Les protéines nouvellement synthétisées iront à la mitochondrie. Le mécanisme d'importation des protéines va dépendre de la destination finale de celle-ci. Le système le mieux étudié est celui de la levure (Schmidt, Pfanner and Meisinger, 2010; Zhao and Zou, 2021), qui est décrit dans les prochains paragraphes et en Figure I-5.

Les protéines de la membrane externe ont plusieurs possibilités d'importation. Les protéines en feuillet β sont reconnues par TOM, puis transloquées dans l'espace intermembranaire. Elles sont transférées par le complexe de chaperonnes Tim9-Tim10 au complexe SAM (*Sorting and Assembly Machinery*) pour être intégrées à la membrane externe mitochondriale. Les protéines en hélices α peuvent être intégrées directement du cytosol, ou en passant par Mim1 (*Mitochondrial importation 1*) puis SAM.

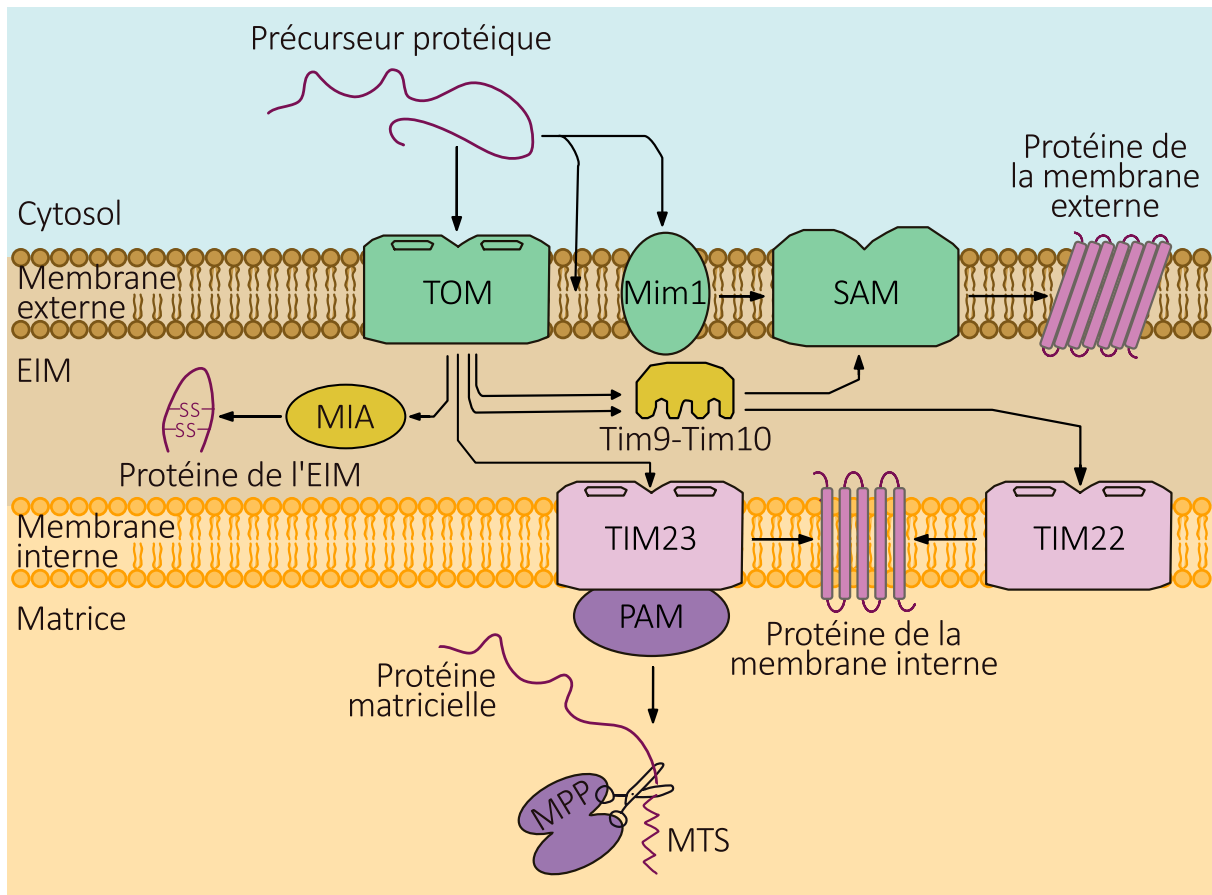


Figure I-5 : L'importation post-traductionnelle de protéines dans la mitochondrie de levure

D'après (Schmidt, Pfanner and Meisinger, 2010).

Schéma de l'importation post-traductionnelle de protéines dans la mitochondrie.

Jaune : protéines de l'EIM (espace intermembranaire) ; **Rose clair** : protéines dans la membrane interne ; **Rose foncé** : protéines importées ; **Vert** : protéines dans la membrane externe ; **Violet** : protéines matricielles. Les flèches **noires** indiquent le trajet suivi par la protéine à importer.

Les protéines de l'EIM sont riches en cystéines, un acide aminé soufré. Elles sont reconnues par le complexe TOM puis par le complexe MIA (*Mitochondrial Intermembrane space Assembly*) permettant l'oxydation des cystéines. Cela crée des ponts disulfures entre les résidus cystéines, permettant le bon repliement des protéines de l'EIM.

Les protéines de la membrane interne sont reconnues par des récepteurs du complexe TOM. Le précurseur est alors transloqué dans l'espace intermembranaire et peut être transféré par le complexe de chaperonnes Tim9-Tim10 au complexe TIM22 pour être intégré dans la membrane interne mitochondriale. Le précurseur peut également passer par le complexe TIM23 pour être intégré à la membrane interne.

Pour les protéines matricielles, des récepteurs du complexe TOM vont reconnaître la MTS (*Mitochondrial Targeting Sequence*). Celle-ci sert de séquence d'adressage à la mitochondrie. Le précurseur est alors transloqué dans l'espace intermembranaire et pris en charge par le complexe TIM23 avec l'aide de PAM (*Presequence translocase-Associated Motor*) pour arriver dans la matrice où la MTS est coupée par la MPP (*Mitochondrial Processing Peptidase*).

En comparaison des mammifères ou des levures, les MTS de plantes sont plus longues de 7 à 9 résidus, avec une moyenne de 42 acides aminés. Les prédictions des MTS se basent sur les critères suivants : elles comportent des résidus chargés positivement, hydrophobes et hydroxylés. Souvent, elles peuvent former des hélices- α amphipathiques, avec une surface chargée positivement et l'autre hydrophobe (Pujol, Maréchal-Drouard and Duchêne, 2007).

Les complexes TOM/TIM des plantes diffèrent un peu de ceux de la levure (Carrie, Murcha and Whelan, 2010). Par exemple, chez les plantes, TOM9, homologue de TOM22 de la levure, possède la même structure, sauf que TOM9 n'a pas le domaine récepteur cytosolique de TOM22. Chez *A. thaliana*, il y a en plus la protéine OM64 qui est impliquée dans l'importation de certains précurseurs protéiques (Lister *et al.*, 2007). Enfin, aucun orthologue de TOM70 n'a été trouvé chez les plantes. Concernant le complexe TIM, il y a moins de divergences entre les plantes et les levures. Chez les plantes, la protéine PAM17 n'est pas retrouvée, et TIM17 a une grande extension en C-terminal, qui est également retrouvée dans la membrane externe. Cette extension permet de faciliter les interactions avec le complexe TOM et avec les protéines à importer (Chaudhuri *et al.*, 2020).

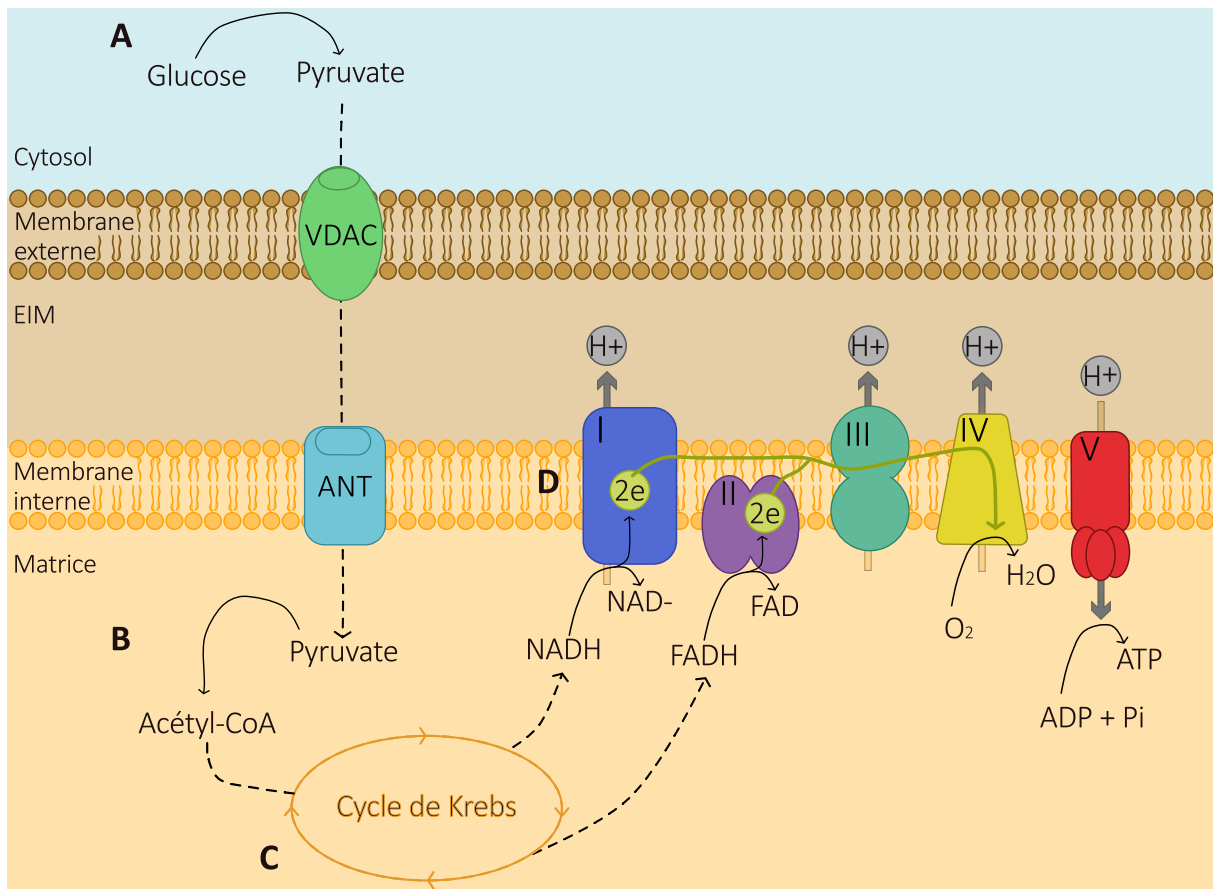


Figure I-6 : La mitochondrie et la production d'énergie

Schéma des différentes étapes de la production d'énergie par la mitochondrie. **A** : La glycolyse permet la dégradation d'une molécule de glucose en pyruvate. **B** : Le pyruvate diffuse dans la matrice mitochondriale par les canaux transporteur VDAC et ANT, où il sera transformé en Acétyl-CoA. **C** : l'Acétyl-CoA entre dans le cycle de Krebs (en orange) qui va produire du NADH et du FADH. **D** : Les molécules de NADH et de FADH sont métabolisées par la chaîne respiratoire, composée de 5 complexes (I à V), afin de produire de l'eau (H_2O) et de l'ATP. Cela se fait grâce à la chaîne de transport d'électrons (en vert) et des protons (H^+) contenus dans l'espace intermembranaire.

Flèches **noires** pleines : réactions chimiques ; flèches **noires** en pointillés : déplacement des molécules.

La deuxième voie possible est celle de l'importation co-translationnelle. Il a été étudié chez la levure, les plantes et chez les mammifères (Verner, 1993; Michaud et al., 2014; Avendaño-Monsalve, Ponce-Rojas and Funes, 2020). Les ARNm (ARN messagers) correspondant à des protéines mitochondriales codées par le génome nucléaire sont localisés à la surface des mitochondries, où ils seront traduits et immédiatement importés dans l'organelle (Figure I-4-2). Chez *S. cerevisiae*, environ la moitié des ARNm des protéines mitochondriales ont été trouvés à la surface des mitochondries (Williams, Jan and Weissman, 2014). Chez *A. thaliana*, cela concerne un tiers des ARNm correspondant à des protéines mitochondriales (Vincent et al., 2017).

V. Fonctions mitochondriales

Les mitochondries ont de nombreuses fonctions. Leurs rôles principaux sont la production d'énergie et la respiration cellulaire. Néanmoins, elles sont également impliquées dans la tolérance aux stress ainsi que dans le métabolisme des acides aminés et des nucléotides, mais aussi dans la biosynthèse des lipides et des stéroïdes ainsi que des clusters Fer/Soufre (McBride, Neuspiel and Wasiak, 2006 ; Spinelli and Haigis, 2018). De plus, les mitochondries permettent la régulation du processus de mort cellulaire programmée, l'apoptose, lors du développement ou lors d'une attaque par un pathogène, afin de limiter la propagation aux autres cellules (Scott and Logan, 2008).

1. Production d'énergie et respiration

La fonction principale des mitochondries dans les cellules eucaryotes est la production d'énergie. Celle-ci nécessite plusieurs étapes.

a) Cycle de Krebs

Les molécules de pyruvate, provenant de la glycolyse cytosolique (Figure I-6A), vont aller dans la matrice mitochondriale. Là, elles seront transformées en acétyl-CoA (Figure I-6B), libérant une molécule de CO₂ et une de NADH. Une molécule d'acétyl-CoA est métabolisée par le cycle de Krebs (Figure I-6C) : elle sera oxydée afin de récupérer 3 molécules de CO₂, 4 de NADH, 4 protons H⁺, une de FADH₂ (flavine adénine dinucléotide 2 hydrogènes, forme réduite du FAD) et un ATP. Pour une molécule de glucose de départ, ce rendement est donc doublé.

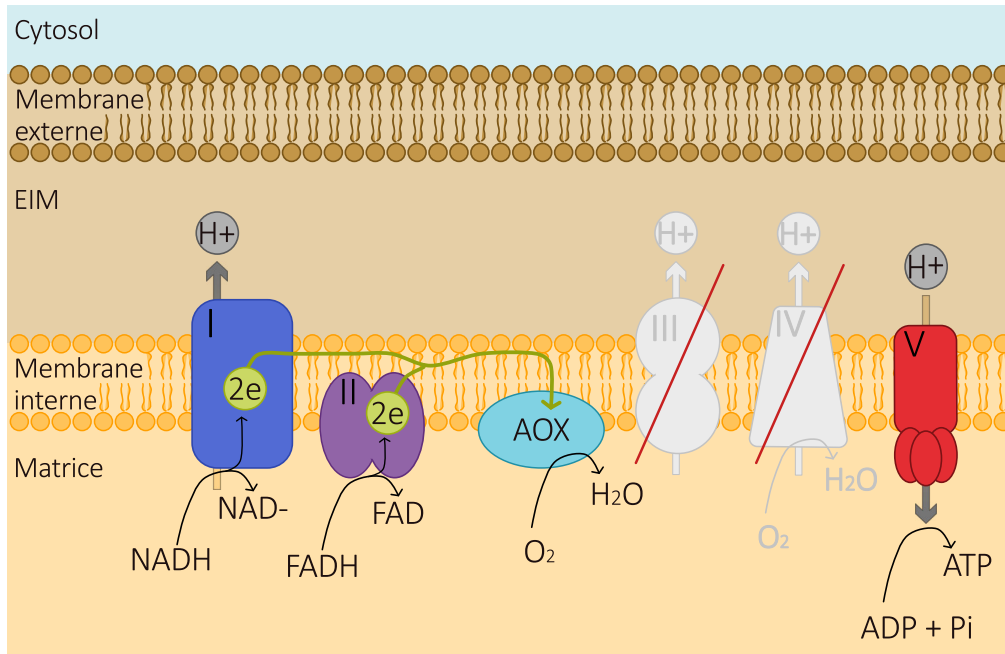


Figure I-7 : La voie alternative à la respiration

Schéma de la voie alternative à la respiration, la voie de l'AOX.

L'AOX court-circuite les complexes III et IV (grisés) de la chaîne respiratoire et catalyse la réduction de l'oxygène en H_2O grâce à la chaîne de transport d'électrons (en vert).

Flèches **noires** pleines : réactions chimiques.

b) Phosphorylation oxydative

Les réactions suivantes se passent au niveau de la chaîne respiratoire, composée de 5 complexes ancrés dans la membrane interne de la mitochondrie (Barros and McStay, 2020) (Figure I-6D). Les molécules de NADH et de FADH₂ seront métabolisées par les complexes I et II (respectivement) de la chaîne respiratoire, afin de créer un gradient de protons H⁺ dans l'espace intermembranaire. Ce gradient de protons est possible grâce à l'énergie des électrons provenant des molécules de NADH et de FADH₂. Ces électrons riches en énergie seront transportés dans les différents complexes de la chaîne respiratoire, c'est la chaîne de transport d'électrons. Le gradient de protons est renforcé par le complexe III qui permet le transport de 4 protons de la matrice à l'espace intermembranaire. De même, 2 protons sont transportés par le complexe IV qui est le lieu d'arrivée de la chaîne de transport d'électrons, réduisant l'oxygène en une molécule d'eau (H₂O). Une molécule de NADH permettra le transport de 10 H⁺ de la matrice à l'espace intermembranaire, et celle de FADH₂ de 6. Ceux-ci retourneront à la matrice par une pompe à protons, une ATP synthase, le complexe V, produisant ainsi de l'ATP.

Au total, à partir d'une molécule de glucose, sont produits 32 ATP.

2. Voie alternative à la respiration et tolérance aux stress

Une voie alternative à la respiration décrite ci-dessus existe chez les plantes, mais aussi chez les protistes et les champignons (Selinski *et al.*, 2018). Elle fait intervenir l'AOX (*Alternative Oxidase*) (Figure I-7). En conditions de stress, une surexpression de l'AOX a été observée chez le colza (Yang *et al.*, 2019).

En effet, les organismes photosynthétiques convertissent l'énergie lumineuse en énergie chimique par la photosynthèse. Or ces organismes sont constamment soumis à des changements des conditions environnementales, entraînant divers stress : changement de températures, de luminosité, de disponibilité en eau ou encore des carences en nutriments. Ces stress environnementaux impactent divers processus métaboliques, dont la photosynthèse. Cela provoque un déséquilibre entre la production et la consommation d'énergie, menant à une surproduction d'électrons au niveau photosynthétique et à un stress oxydatif. Des ROS (*Reactive Oxygen Species*) sont produits et conduisent à des dommages cellulaires.

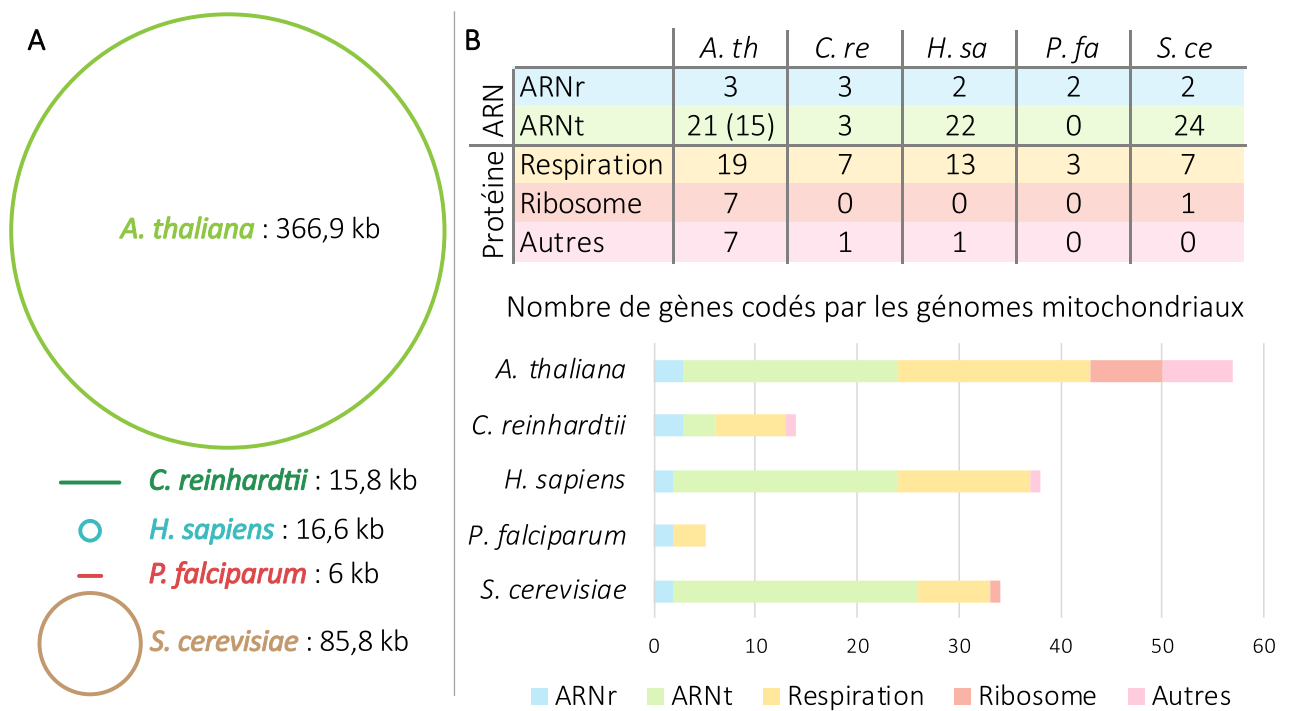


Figure I-8 : Une diversité de génomes mitochondriaux

Comparaison des génomes mitochondriaux de 5 espèces d'Eucaryotes : *Arabidopsis thaliana*, *Chlamydomonas reinhardtii*, *Homo sapiens*, *Plasmodium falciparum* et *Saccharomyces cerevisiae*, en termes de taille (A) et de contenu (B). A : Représentation schématique de la taille et de la structure des génomes mitochondriaux de 4 organismes différents. Les tailles sont indiquées en kilobase (kb) ; B : Tableau et histogramme correspondant au nombre de gènes codés par les 4 génomes mitochondriaux selon si le gène code pour un ARN (ARNr en bleu ou ARNt en vert), pour une protéine des complexes de la chaîne respiratoire en jaune, pour une sous-unité protéique du mitoribosome en rouge ou pour une protéine mitochondriale autre en rose. Le chiffre entre parenthèse pour *Arabidopsis thaliana* est le nombre d'ARNt iso-accepteurs différents codés par le génome.

La voie de l'AOX permet de contrer ces effets négatifs sur les cellules. Un lien fonctionnel entre le chloroplaste et la mitochondrie a été montré (Bennoun, 1982 ; Cardol et al., 2003). Les molécules de NADH produites lors de la photosynthèse peuvent être utilisées pour la respiration mitochondriale, permettant d'éviter la production de ROS et donc des dommages cellulaires. L'AOX court-circuite les complexes III et IV de la chaîne respiratoire et catalyse la réduction de l'oxygène en H₂O grâce à la chaîne de transport d'électrons, sans affecter le gradient de protons. La production d'ATP est donc grandement réduite et l'énergie produite par le transport d'électron est dissipée en chaleur. L'AOX permet donc de réguler la production de ROS et d'améliorer la tolérance des plantes aux stress (Kaye et al., 2019).

VI. Génomes mitochondriaux

Les mitochondries possèdent leur propre génome, vestige de leur ancêtre bactérien pré-endosymbiotique. Il se trouve dans la matrice mitochondriale. Durant l'évolution, la plupart des gènes originels de l' α -protéobactérie ont été transféré au noyau de la cellule hôte ou perdus. Aujourd'hui, les génomes mitochondriaux connus codent pour un nombre très limité de gènes de protéines dont l'expression reste toutefois essentielle. Cependant, ces organites ont suivi des voies évolutivement très divergentes au sein des différentes lignées eucaryotes, résultant en une diversité de génomes en termes de structures, tailles et de composition mais aussi une diversité de mécanismes d'expressions des gènes (Kummer and Ban, 2021).

1. Taille et contenus des génomes mitochondriaux

La taille des génomes mitochondriaux varie énormément selon les organismes, avec 366,9 kb pour la plante à fleurs *A. thaliana* (Marienfeld, Unseld and Brennicke, 1999), 15,8 kb chez l'algue verte *C. reinhardtii* (Salinas-Giegé et al., 2017), 16,6 kb pour l'Homme (Taanman, 1999) ou encore 85 kb pour la levure *S. cerevisiae* (Wolters, Chiu and Fiumera, 2015) (Figure I-8A). Le plus petit génome mitochondrial connu est celui de *Plasmodium falciparum*, avec seulement 6 kb (Feagin et al., 2012). A l'inverse, le plus grand est celui de *Silene conica*, une plante herbacée, avec 11 Mb répartis en plusieurs chromosomes (Sloan et al., 2012). Si la taille des génomes mitochondriaux des mammifères est très conservée, entre 15 et 17 kb, ce n'est pas le cas de toutes les familles d'organismes. En moyenne chez les plantes terrestres, les tailles des génomes mitochondriaux varient entre 300 et 500 kb. Cela s'explique par la présence

d'introns, de longues séquences intergéniques et de grandes séquences répétées, permettant des évènements de recombinaison (Sloan et al., 2012 ; Gualberto and Newton, 2017).

Mais la taille des génomes n'est pas corrélée au nombre de gènes codés (Figure I-8B). Par exemple, le génome mitochondrial de *C. reinhardtii* code pour 14 gènes, avec une taille de 15,8 kb. Chez l'Homme, l'ADNmt de 16,6 kb code pour 38 gènes, le double par rapport à l'algue verte, pour des génomes de tailles équivalentes. Par contre, chez *A. thaliana*, l'ADNmt bien que plus grand, ne contient que 57 gènes, seulement 20 de plus que l'Homme, pour un génome 20 fois plus grand.

Peu importe le nombre de gènes codés par les génomes mitochondriaux, leur expression est essentielle. Il y a 3 types de gènes codés par les génomes mitochondriaux, et ceci est assez conservé parmi les eucaryotes. D'une part, il y a les gènes codant pour des ARNr (ARN ribosomiques) mitochondriaux, qui font partie du peu de gènes retrouvés de manière universelle dans les génomes mitochondriaux (Gray, 2012). Puis il y a les gènes d'ARNt. Le génome mitochondrial peut coder pour un set complet d'ARNt, comme chez l'Homme avec 22 gènes d'ARNt ou chez *S. cerevisiae* avec 24 gènes d'ARNt. D'autres génomes codent pour un set incomplet d'ARNt, comme chez *A. thaliana* qui possède 21 gènes d'ARNt, mais codant pour 15 ARNt iso-accepteurs seulement. Certains organismes ont encore moins de gènes d'ARNt codés dans l'ADNmt, comme *C. reinhardtii* qui n'en possède que 3. Il y a aussi des organismes qui n'ont aucun gènes d'ARNt mitochondriaux, comme *Plasmodium falciparum* ou *Trypanosoma brucei* (Salinas-Giegé, Giegé and Giegé, 2015). Dans ces cas, les ARNt manquants seront importés dans la mitochondrie, afin de pouvoir synthétiser les protéines mitochondriales. Ces dernières appartiennent au troisième type de gènes mitochondriaux : les gènes codant pour des protéines. Il y a les gènes qui codent pour des sous-unités des complexes de la chaîne respiratoire. D'autres codent pour des sous-unités du mitoribosome, et sont retrouvés essentiellement chez les génomes mitochondriaux de plantes et de protistes (Marienfeld, Unseld and Brennicke, 1999 ; Taanman, 1999 ; Wolters, Chiu and Fiumera, 2011 ; Feagin et al., 2012 ; Salinas-Giegé et al., 2017). Enfin, certains génomes mitochondriaux ont des gènes codant pour des protéines avec d'autres fonctions. C'est le cas de *C. reinhardtii* par exemple, dont le génome code pour une protéine *reverse transcriptase-like*, ou d'*A. thaliana* dont le génome code pour des protéines impliquées dans la maturation du cytochrome *c*.

2. Structure et organisation des génomes mitochondriaux

La structure de ces génomes diffère également selon les espèces, comme illustré en [Figure I-8A](#). Classiquement, les génomes mitochondriaux sont représentés sous forme d'une molécule circulaire unique, appelée « *master circle* », comme pour *A. thaliana* ou *H. sapiens* par exemple. *In vivo*, ce n'est pas aussi simple. Chez les plantes, des recombinaisons sont possibles grâce à de grandes séquences répétées, formant des formes subgénomiques circulaires ou linéaires. Dans ces cas-là, la structure de l'ADNmt est hautement dynamique. Il existe aussi des génomes en plusieurs chromosomes circulaires, comme celui du concombre par exemple, qui a 3 chromosomes circulaires (Alverson *et al.*, 2011) ou celui de *Silene conica* qui est composé de 11 petits chromosomes circulaires (Sloan *et al.*, 2012). *Picea sitchensis*, un conifère de la familles des Pinacées, a une structure encore plus complexe, en plusieurs parties assemblées formant différents embranchements (Jackman *et al.*, 2020). Chez les protistes kinétoplastidiés comme *Trypanosoma*, le génome mitochondrial se compose d'une cinquantaine de « *maxicircles* » et d'une centaine de « *minicircles* » (Aphasizhev and Aphasizheva, 2011). Ces *maxi-* et *mini-circles* forment un réseau d'ADN circulaire, appelé kinétoplaste ou ADNk.

Il existe également des génomes linéaires. Cela a été décrit chez des algues comme *Chlamydomonas* ou *Polytomella* (Smith, Hua and Lee, 2010), le protiste *P. falciparum*, certains cnidaires (Lavrov and Pett, 2016 ; Stampar *et al.*, 2019) ou encore un champignon pathogène de la pomme de terre (van de Vossenbergh *et al.*, 2018). Les génomes linéaires sont aussi répandus chez les ciliés (Zhang *et al.*, 2021). Ces génomes contiennent alors des extrémités structurées, similaires à des télomères afin de les protéger de la dégradation (Smith, Hua and Lee, 2010).

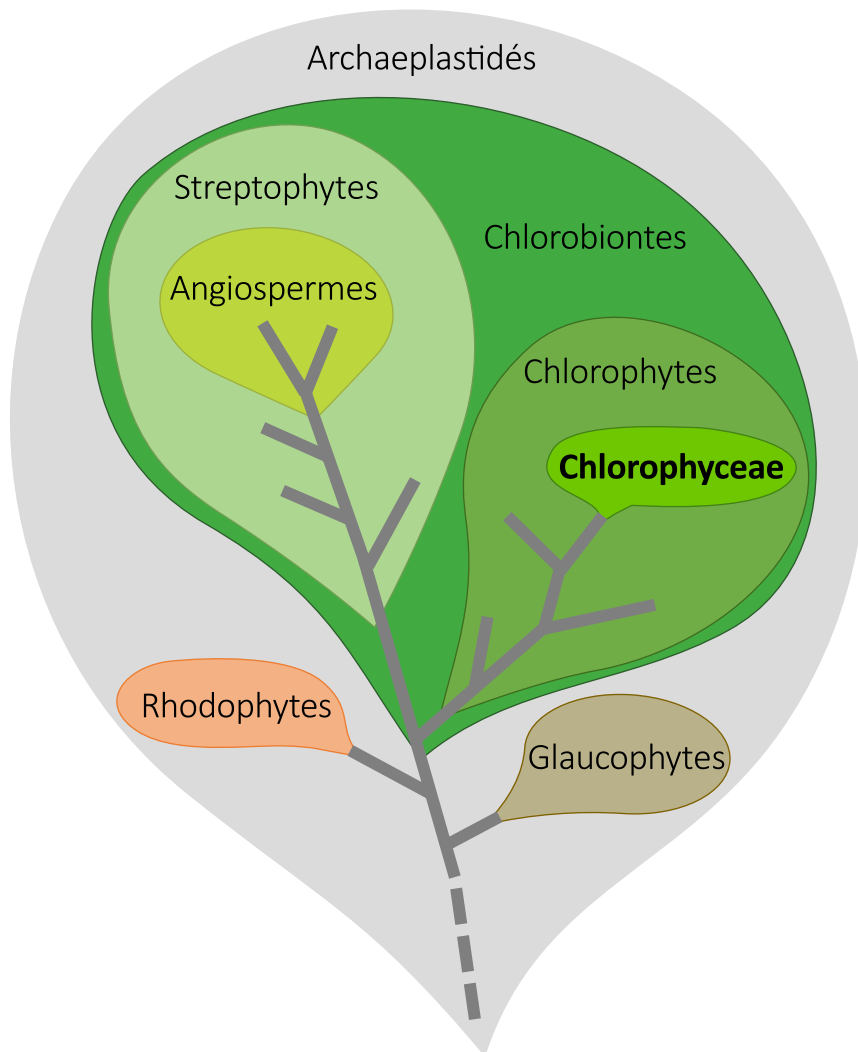


Figure I-9 : Arbre phylogénétique des Archaeplastidés

Schéma de l'arbre phylogénétiques des Archaeplastidés, regroupant les Glaucophytes, les Rhodophytes et des Chlorobiontes. Ces derniers regroupent les Streptophytes, dont font partie les Angiospermes, et les Chlorophytes, dont font partie les Chlorophyceae.

Gras : clade des Chlorophyceae comprenant *C. reinhardtii*.

CHLAMYDOMONAS REINHARDTII

Chlamydomonas reinhardtii est une algue verte unicellulaire. Son appellation provient du grec *chlamys*, le manteau et *monas*, solitaire. Elle a été décrite pour la première fois au 19^{ème} siècle et utilisée pour des études génétiques au 20^{ème} siècle. Son développement en tant qu'organisme modèle date des années 1950, lors de la génération des premiers mutants.

Cette espèce fait partie de la lignée des *Archeplastidia*, la lignées des plantes, et plus précisément du groupe des *Chlorobionta*, de l'embranchement des *Chlorophytes* et de la classe des *Chlorophyceae* (Figure I-9). Dans la nature, elle est présente partout sur le globe, dans les sols humides ainsi que dans les eaux douces.

I. Organisation cellulaire

Mesurant entre 5 et 10 μm , *C. reinhardtii* est une algue unicellulaire photosynthétique. L'organisation cellulaire d'une cellule *C. reinhardtii* est schématisée en Figure I-10. Une paroi cellulaire entoure l'algue, constituée de glycoprotéines, en plusieurs couches. Parmi les composants « classiques » des cellules eucaryotes, sont retrouvés le noyau et les appareils Golgien ainsi que les mitochondries. La membrane nucléaire est en continuité avec le réticulum endoplasmique. Proches du noyau sont retrouvés un à quatre appareils de Golgi. Les mitochondries ont une organisation dynamique en fonction du cycle cellulaire et du rythme circadien, pouvant passer d'une grosse mitochondrie unique proche du noyau à un réseau de mitochondries dispersées dans toute la cellule (Ehara, Osafune and Hase, 1995). Cette dernière conformation est celle qui est classiquement représentée et observée (Figure I-11), puisqu'elle est majoritaire lorsque les cellules sont cultivées en présence de lumière.

De 10 à 12 μm de long, les 2 flagelles antérieurs permettent les déplacements de l'algue. Ils sont aussi essentiels pour sa reproduction sexuée (*mating*). A leurs bases se trouvent les corps basaux, permettant l'assemblage des flagelles. Ce processus est assez dynamique puisque l'algue peut perdre ses flagelles lorsqu'elle est confrontée à divers stress environnementaux (Yueh and Crain, 1993). Ce processus est appelé la déflagellation. La régénération des flagelles se fait en 3h (heures) (Harris, 2001).

Deux vacuoles contractiles, situées proches des corps basaux des flagelles, permettent de contrôler le contenu en eau dans le cytoplasme. Lorsqu'il y a trop d'eau dans le cytoplasme,

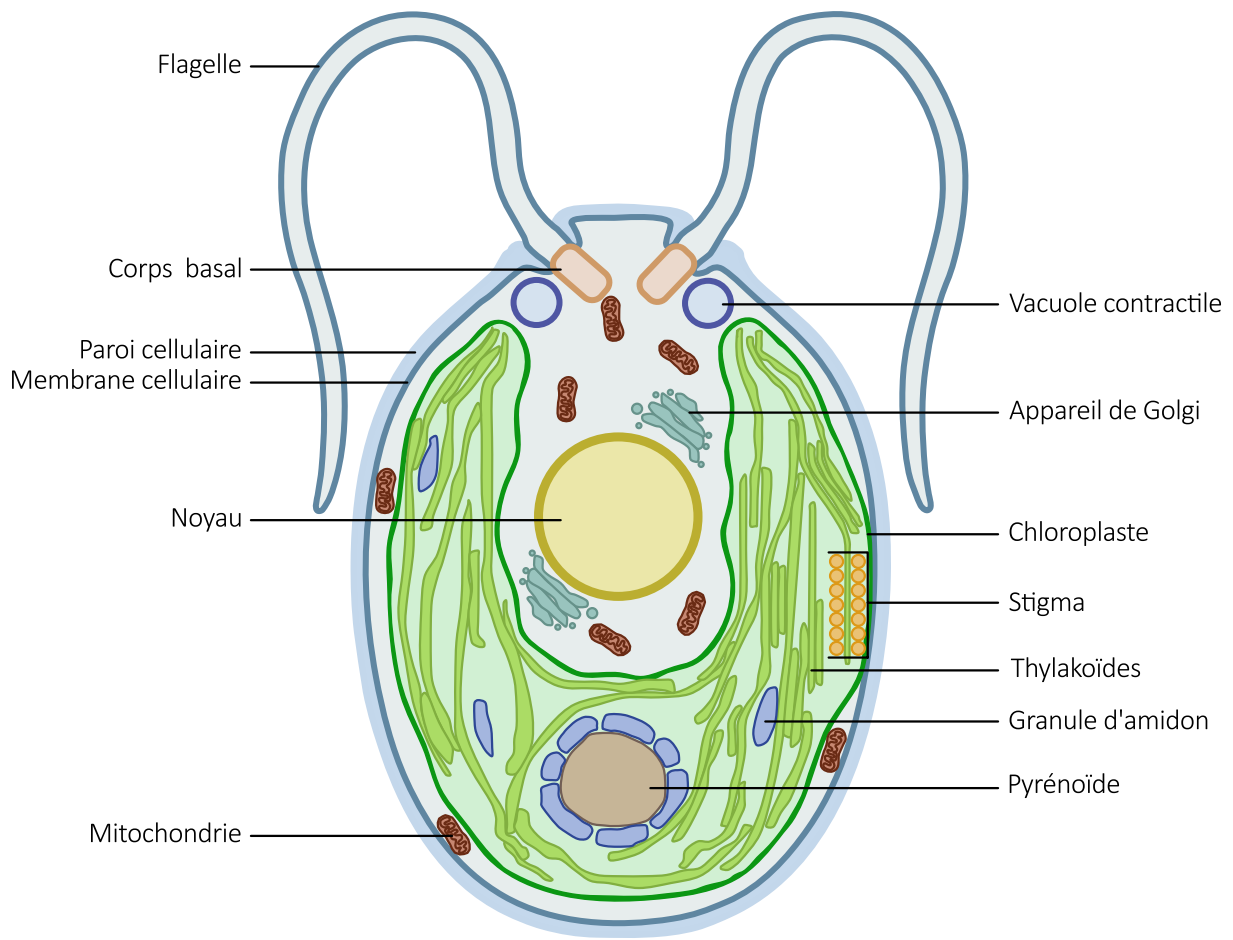


Figure I-10 : Structure d'une cellule végétative de *Chlamydomonas reinhardtii*

D'après (Sasso *et al.*, 2018).

Schéma d'une cellule végétative de *C. reinhardtii*, de sa structure et de ses composants.

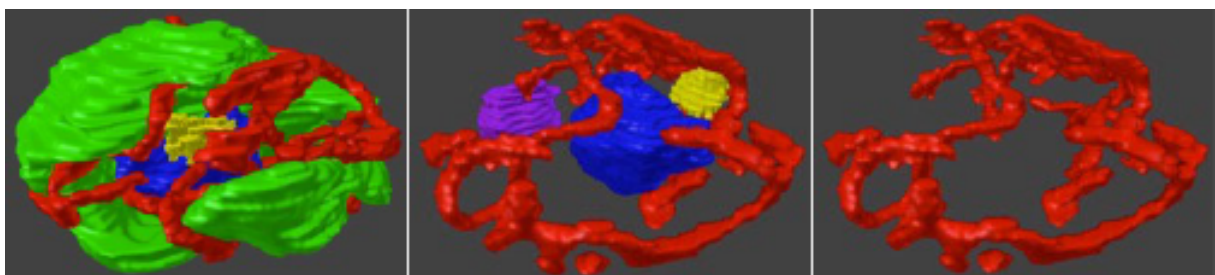


Figure I-11 : Structure en 3D d'une cellule végétative de *C. reinhardtii*.

Modèle de reconstruction en 3D d'une cellule végétative d'une souche sauvage de *C. reinhardtii*, obtenu en microscopie électronique à balayage (appareil : Zeiss SBF-SEM Sigma avec Gatan 3ViewXP2) par Mathieu Erhardt (IBMP (Institut de Biologie Moléculaire des Plantes), plateforme Microscopie et imagerie cellulaire). **Vert** : chloroplaste ; **Rouge** : mitochondries ; **Bleu** : noyau ; **Violet** : pyrénioïde ; **Jaune** : vacuole contractile.

ces vacuoles permettent de pomper l'eau en dehors de la cellule (Komsic-Buchmann, Wösthoff and Becker, 2014).

Une des caractéristiques remarquables de *C. reinhardtii* est la présence d'un chloroplaste unique. Celui-ci, en forme de U, occupe une très grande proportion de la cellule. Le chloroplaste contient la machinerie photosynthétique dans les thylacoïdes, des grains d'amidon mais également un pyrénnoïde et un stigma. Le pyrénnoïde est une structure dans laquelle est concentrée la Rubisco, l'enzyme responsable de la fixation du carbone du CO₂ atmosphérique mais aussi de l'O₂. Les 2 molécules étant en compétition au niveau du site actif de la Rubisco, cela résulte en une perte de fixation du carbone via le processus de photorespiration. Pour contrer cette photorespiration, *C. reinhardtii* a un mécanisme de concentration du carbone (CCM (*Carbon Concentrating Mechanism*), afin d'augmenter le ratio CO₂:O₂ au niveau du site actif de la Rubisco (Mackinder, 2018). Vers l'équateur de la cellule, dans le chloroplaste, se trouve un stigma. C'est un organite photorécepteur permettant à la cellule de s'orienter par rapport au signal lumineux, de s'en approcher ou de s'en éloigner. Ce processus s'appelle la phototaxie. Ce stigma est composé de pigment de caroténoïdes, c'est pourquoi il apparaît orange en microscopie.

II. Cycle cellulaire

Comme pour beaucoup d'organismes unicellulaires, *C. reinhardtii* a 2 moyens de reproduction : asexuée et sexuée (Figure I-12) (Harris, 2001; Goodenough, Lin and Lee, 2007; Sasso *et al.*, 2018). Les cellules ont 2 « sexes » ou *mt* (*mating types*) possibles : *mt*⁺ ou *mt*⁻.

Lors de la reproduction asexuée, les cellules sont haploïdes, à l'état végétatif. Elles se multiplient par mitoses.

La reproduction sexuée commence lorsqu'il y a une carence en azote dans le milieu. Les cellules végétatives des 2 *mating types* se différencient en gamètes. Certaines souches nécessitent, en plus de la carence en azote, la présence de lumière bleue afin de produire des gamètes. La reconnaissance des cellules de *mt* opposés se fait par les protéines des flagelles spécifiques à chaque *mt*. Les flagelles s'associent par paires. L'appariement des flagelles déclenche la dégradation de la paroi cellulaire des 2 gamètes, puis leur fusion. Le zygote diploïde en résultant reste mobile quelques heures, avec 4 flagelles, puis il les perd et devient immobile. Il y a alors la formation d'une zygospore, entourée d'une paroi dure et imperméable. Cela

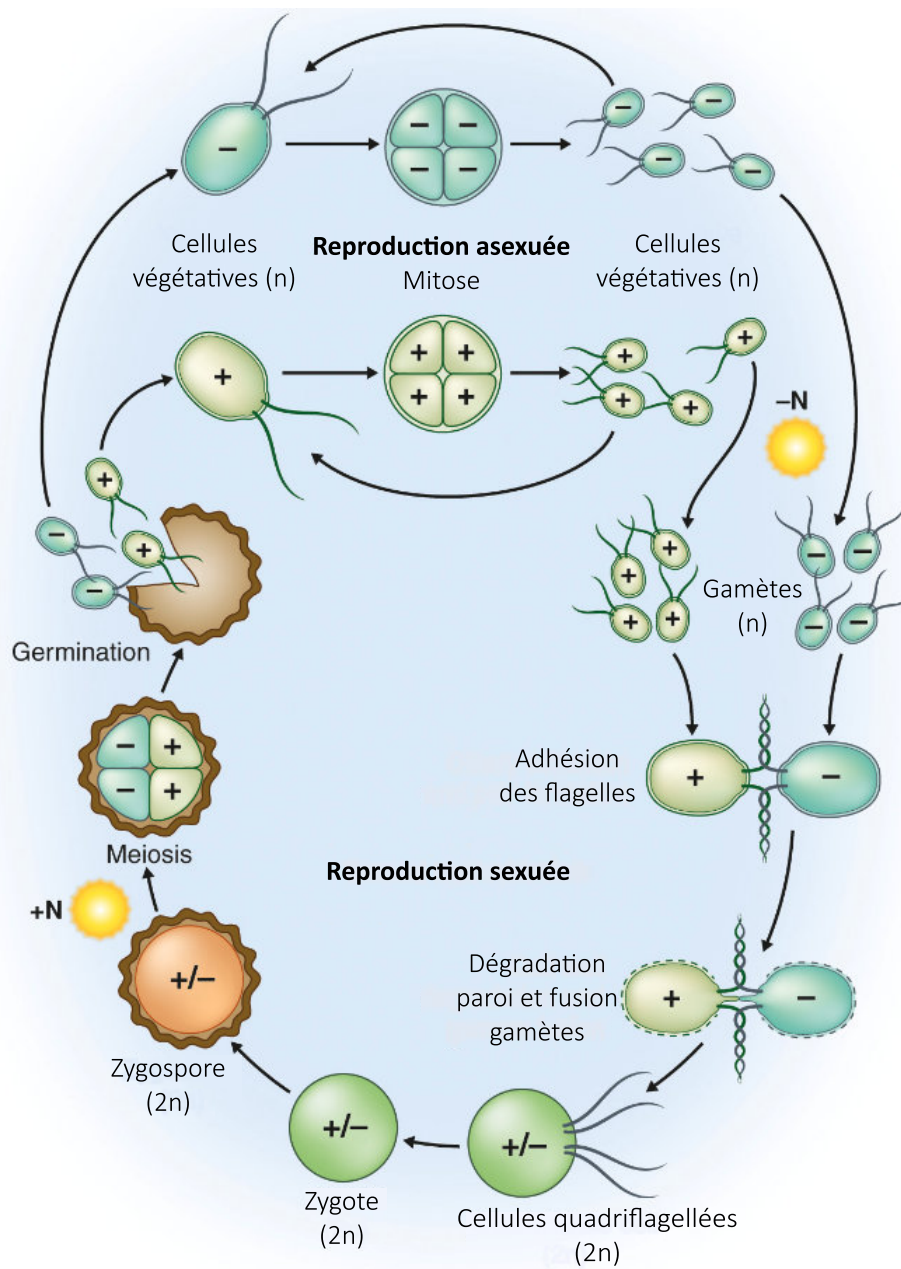


Figure I-12 : Cycle cellulaire de *C. reinhardtii*

Adapté de (Sasso *et al.*, 2018)

Durant la reproduction asexuée, les cellules végétatives haploïdes (n) se multiplient par mitose. Lorsqu'il y a une carence en azote, en présence de lumière, la gamétogenèse est induite : C'est la reproduction sexuée. Deux gamètes opposés (+ (*mt*⁺) et - (*mt*⁻)) fusionnent pour former un zygote diploïde immobile. Celui-ci devient une zygospore résistante et dormante, attendant que les conditions environnementales s'améliorent. Là, elle subira une méiose afin de libérer 4 cellules végétatives haploïdes.

permet une résistance aux conditions environnementale pendant la dormance de la zygospore. Le chloroplaste est dégradé et il y a une accumulation de caroténoïdes, donnant une couleur orangée à la zygospore (Lohr, 2009). La zygospore entre en dormance et peut rester viable pendant quelques années. La dormance est levée lors de l'amélioration des conditions environnementale : la fin de la carence en azote, en présence de lumière. C'est la germination : une méiose génère 4 cellules végétatives haploïdes à partir du zygote $2n$, qui seront libérés de la zygospore.

III. Génomes

Une cellule de *C. reinhardtii* renferme 3 génomes : le nucléaire, le chloroplastique et le mitochondrial. Ceux-ci sont entièrement séquencés et annotés.

Le génome nucléaire a été publié en 2007 (Merchant *et al.*, 2007). Il est constitué de 121 Mb et est divisé en 17 chromosomes, pour 15000 gènes (Figure I-13). Avec plus de 60% de G et de C, on dit qu'il est « G-C riche », ce qui est plus important que pour l'Homme (41% de G-C (Merchant *et al.*, 2007)) et presque le double par rapport à *A. thaliana* (36% de G-C (Mishra *et al.*, 2009)). Les gènes de protéines possèdent un grand nombre d'exons, 8,3 en moyenne par gène, contre 5,2 chez *A. thaliana*. Cette richesse en introns est inhabituelle comparé aux autres eucaryotes unicellulaires ou aux plantes terrestres (Merchant *et al.*, 2007). Cette caractéristique se retrouve aussi chez l'Homme, qui a une moyenne de 8,8 exons par gène. Cela dit, les gènes humains ont une taille moyenne de 27000 pb, contre 4300 pour l'algue ou 2200 pour *A. thaliana*. Cette différence de taille s'explique par la longueur moyenne des introns, qui est bien plus importante pour l'Homme (3300 pb en moyenne) que pour *C. reinhardtii*, avec 373 pb en moyenne, ou *A. thaliana* avec une moyenne de 164 pb.

La première version du génome chloroplastique a été publié en 2002 (Maul *et al.*, 2002). Plusieurs versions ont été publiées depuis, la dernière datant de 2018 (Gallaher *et al.*, 2018). Il est constitué de 205 kb pour un seul chromosome. Structurellement, il y a un mélange de génomes circulaires et linéaires. Il est divisé en 2 régions uniques de 80 kb environ séparés par des séquences répétées inversées de 21 kb. Il code pour 97 gènes : 62 gènes de protéines, 30 d'ARNt et 5 d'ARNr. La majorité de ces gènes ne comporte pas d'introns.

Le génome mitochondrial a été publié en 1993 (Vahrenholz *et al.*, 1993). Il est constitué de 15,8 kb, pour un seul chromosome linéaire. Il ne comporte pas d'introns et les séquences

| | <i>C. reinhardtii</i> | <i>A. thaliana</i> | <i>H. sapiens</i> |
|---------------------------------|-----------------------|--------------------|-------------------|
| Nb chromosomes | 17 | 5 | 23 |
| Taille génome (Mb) | 121 | 140 | 2851 |
| Nb gènes | 15k | 26k | 23k |
| Taille moyenne des gènes (pb) | 4300 | 2200 | 27000 |
| %age GC | 64 | 36 | 41 |
| Nb moyen d'exon/gène | 8,3 | 5,2 | 8,8 |
| Taille moyenne des introns (pb) | 373 | 164 | 3365 |

Figure I-13 : Comparaison des génomes nucléaires de 3 espèces

Tableaux récapitulatifs que certaines caractéristiques des génomes nucléaires de *C. reinhardtii*, *A. thaliana* et *H. sapiens*.

Nb : nombre ; Mb : mégabases, 10^6 bases ; k : 10^3 ; pb : paires de bases ; %age : pourcentage

intergéniques sont courtes. Il code pour huit gènes de protéines, trois d'ARNt, un ARN non codant et trois d'ARNr, fragmentés en 12 modules (Salinas-Giegé *et al.*, 2017; Waltz *et al.*, 2021).

IV. Organisme modèle

C. reinhardtii est aujourd'hui un organisme modèle. Cela est dû à plusieurs caractéristiques. Pour commencer, cette algue est facile à cultiver en laboratoires, même au stade haploïde, ce qui permet l'expression immédiate des phénotypes mutants. Avec un temps de génération entre 8 et 15h (selon les souches et les conditions de cultures), l'expérimentateur peut obtenir rapidement une biomasse suffisante pour son travail. Il est possible de travailler uniquement avec des cellules végétatives haploïdes, puisque la gamétogenèse est induite par une privation en azote dans le milieu (Harris, 2001; Sasso *et al.*, 2018). Les cultures sont donc maintenues avec de l'azote dans le milieu. Cela facilite les études génétiques et l'obtention de mutants, évitant ainsi les étapes de génotypages hétéro/homozygote. De plus, s'il y a de l'acétate dans le milieu, la croissance est possible à la lumière comme à l'obscurité pour des souches sauvages. Enfin, les 3 génomes sont séquencés, annotés et transformables, ce qui permet des études génétiques et l'obtention de mutants.

Puisqu'elle n'a qu'un seul gros chloroplaste, *C. reinhardtii* a été et est toujours très utilisée pour un grand nombre d'études sur le chloroplaste. Le fait qu'une croissance à l'obscurité soit possible a permis l'isolement de mutants de la photosynthèse sensibles à la lumière (Spreitzer and Mets, 1980).

Il existe aussi beaucoup d'étude sur ses flagelles. En effet, ceux-ci partagent des similarités structurales et protéiques avec les cils des mammifères. Du fait de la possible déflagellation, il est aisé de les purifier pour les étudier. *C. reinhardtii* est un organisme de choix pour l'étude des flagelles et de leur mobilité, et cela a permis des avancées remarquables dans la compréhension des ciliopathies (maladie dues aux cils) humaines (Brown and Witman, 2014).

Enfin, cette algue est le seul organisme photosynthétique où l'obtention de mutants mitochondriaux viables est possible (Remacle *et al.*, 2006). Ceux-ci pourront croître en conditions photo-autotrophes, c'est-à-dire à la lumière et en présence d'acétate dans le milieu (Salinas *et al.*, 2013). Cela en fait aussi un modèle de choix pour l'étude des mitochondries, et plus particulièrement celles des organismes photosynthétiques.

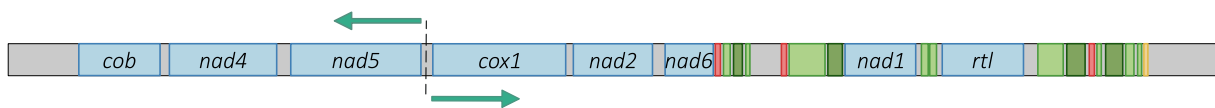


Figure I-14 : Génome mitochondrial de *C. reinhardtii*

Schéma du génome mitochondrial de *C. reinhardtii* avec l'origine de transcription bidirectionnelle représentée par la ligne pointillée et les 2 flèches **turquoise** ; **Gris** : ADN double brin. **Bleu** : gènes de protéines ; **Rouge** : gènes d'ARNt ; **Vert clair** : gènes d'ARNr de la grande sous-unité du mitoribosome ; **Vert foncé** : gènes d'ARNr de la petite sous-unité du mitoribosome ; **Jaune** : gène de l'ARN L2b.

EXPRESSION DU GENOME MITOCHONDRIAL DE *C. REINHARDTII*

I. Génome mitochondrial et réplication

1. Génome mitochondrial

Comme vu précédemment, le génome mitochondrial de *C. reinhardtii*, est une molécule linéaire de 15,8 kb (Figure I-14). Ce génome est dit « compact » : il ne contient pas d'intron et les séquences intergéniques sont courtes. Le génome mitochondrial contient 45% de G-C, ce qui est bien moins que le génome nucléaire de l'algue qui en contient 64% (Gallagher *et al.*, 2018).

Huit gènes codent pour des protéines : 5 pour des sous-unités du complexe I de la chaîne respiratoire (*nad1*, *nad2*, *nad4*, *nad5* et *nad6*), un pour l'apocytochrome *b* complexe III (*cob*), un pour une sous-unité du complexe IV (*cox1*) et le dernier gène code pour la protéine *rtl* (*reverse transcriptase-like protein*), dont la fonction reste inconnue à ce jour. 3 gènes codent les ARNt de la glutamine, du tryptophane et de la méthionine. Enfin, 3 ARNr sont codés en différents modules : le fragment L3a correspond à l'ARNr 5S, 8 modules constituent le 23S et 4 le 16S (Waltz *et al.*, 2021). Nous avons montré que le gène L2b, historiquement annoté comme un ARNr, ne fait pas partie du mitoribosome, bien qu'il soit exprimé (Waltz *et al.*, 2021).

2. Réplication

a) Télomères

Deux télomères forment les extrémités de l'ADNmt de *C. reinhardtii* (Figure I-15A). Pour plus de simplicité, nous nommerons le télomère proche du gène *cob* le « télomère gauche » et celui proche des gènes d'ARNr le « télomère droit ». Ils sont composés de la même séquence de 532 pb, en orientation inversée. Ils ont également une extrémité 3' simple brin libre de 40 nt (nucléotides) (Vahrenholz *et al.*, 1993; Smith and Craig, 2021). De plus, une séquence de 86 nt est répétée dans les 2 télomères et dans la partie codante du génome correspondant à 23 nt en 3' du gène L2b et à la séquence non codante qui le suit. Ces 23 nt sont complémentaires aux 23 nt terminant les extensions 3' simple brin des 2 télomères.

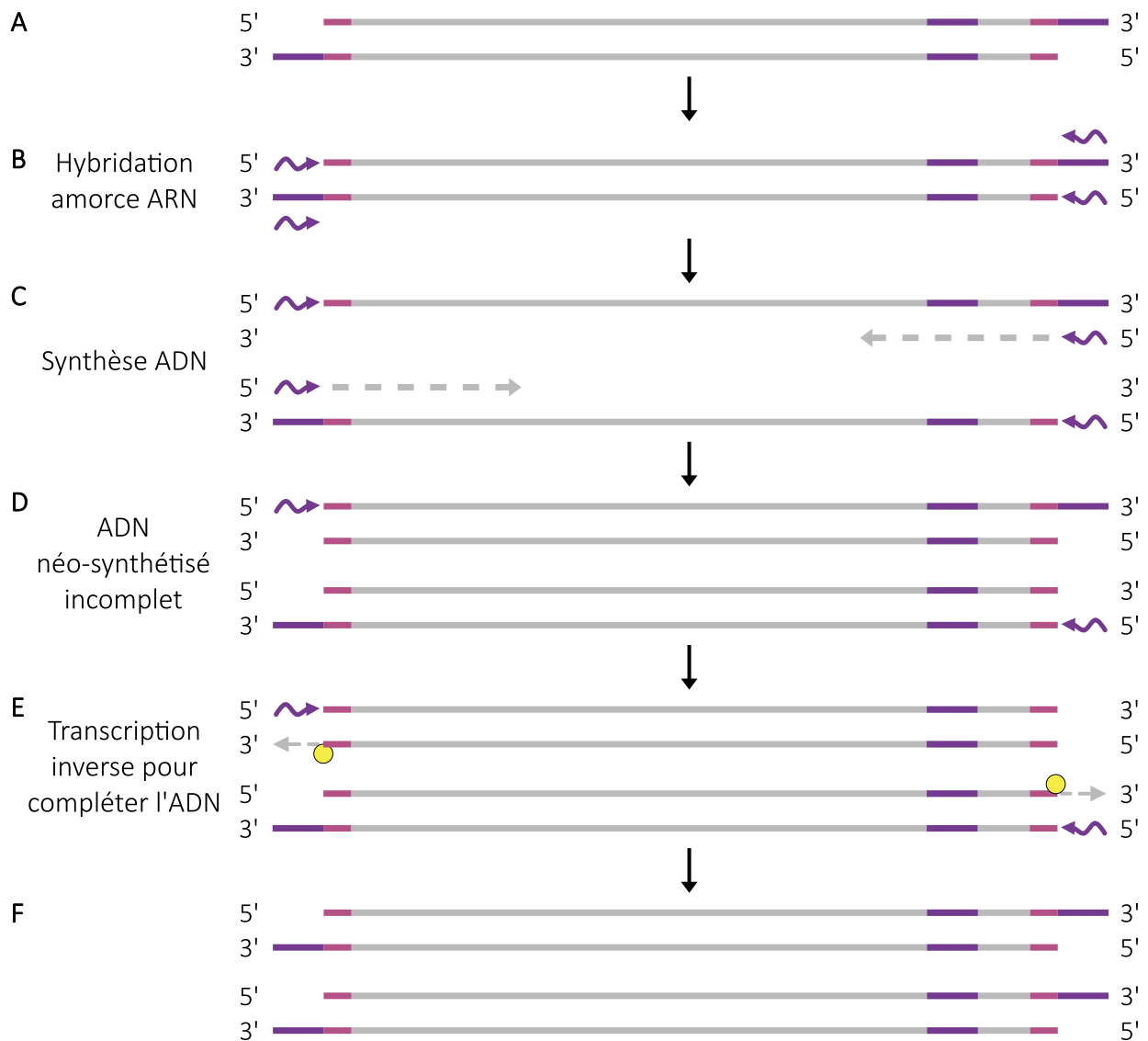


Figure I-15 : Modèle de réplication de l'ADN selon Georg Michaelis

A : Schéma de l'ADN mitochondrial de *C. reinhardtii*. En bouts sont schématisés les télomères, avec en **rose** la séquence commune aux 2 brins d'ADN et en **violet** l'extrémité 3' libre, également répétée dans le génome ; **B à F :** Schéma des différentes étapes de réplication de l'ADN mitochondrial, selon le modèle proposé par (Vahrenholz *et al.*, 1993), jusqu'à l'obtention de 2 copies d'ADN (**F**).

Jaune : protéine Rtl ; **Gris :** ADN ; **Flèche grise en pointillé :** ADN en cours de synthèse ; **flèche violette en zigzag :** amorce ARN issue de la transcription de la séquence répétée interne.

b) Modèle hypothétique

A ce jour, le mécanisme de réplication de l'ADN mitochondrial de *C. reinhardtii* n'a pas encore été élucidé. Toutefois, une hypothèse a été formulée par l'équipe de Georg Michaelis en 1993 (Vahrenholz *et al.*, 1993). Celle-ci se base sur la séquence télomérique répétée dans le génome : l'ARN résultant de la transcription de ces 86 nt permettrait d'initier la réplication, agissant comme une amorce (Figure I-15B). La synthèse d'ADN donnerait une molécule incomplète en 3' (Figure I-15C et D). La réplication de l'ADN mitochondrial se finirait grâce à l'enzyme transcriptase inverse, codée par le gène *rtl*, qui copierait l'amorce ARN en association avec le brin d'ADN (Figure I-15E), afin de compléter le brin néosynthétisé (Figure I-15F).

Si cette hypothèse est plutôt acceptée par la communauté scientifique, il n'y a aucune preuve expérimentale pour la confirmer.

II. Expression du génome mitochondrial

Puisque les mitochondries ont une origine bactérienne et qu'elles ont évolué au sein de la lignée des eucaryotes, l'expression du génome mitochondrial combine des caractéristiques bactériennes héritées de l' α -protéobactérie endocytée avec des caractéristiques eucaryotes, provenant de la cellule hôte (Gray, 2012; Waltz and Giegé, 2020). Cependant, ces organites ont suivi des voies évolutivement très divergentes au sein des différentes lignées eucaryotes, résultant en une hétérogénéité des processus d'expression des ADNmt. Il existe par exemple des différences dans les processus de maturation des ARNm ou dans la structure et de la composition des mitoribosomes et donc aussi des mécanismes de traduction. L'expression du génome des mitochondries est un processus dépendant de facteurs codés par le génome nucléaire, exprimés dans le cytosol et importés dans la mitochondrie.

1. Transcription

L'expression de l'ADNmt commence par la transcription de l'ADN en ARN. Chez *C. reinhardtii*, une origine de transcription bidirectionnelle se trouve entre les gènes *nad5* et *cox1* (Grant and Chiang, 1980) (Figure I-16A). En résulteront 2 co-transcrits primaires polycistroniques (Figure I-16B) qui seront clivés et maturés pour obtenir des ARNm, ARNt et ARNr matures (Figure I-16C) (Gray and Boer, 1988).

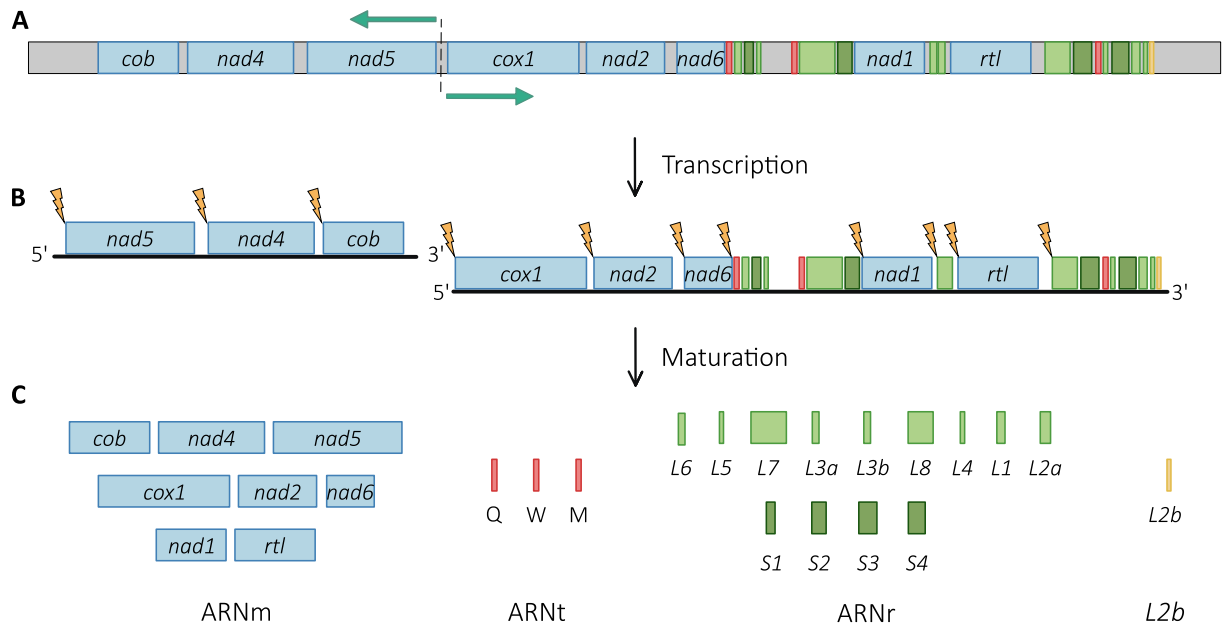


Figure I-16 : Génome mitochondrial de *C. reinhardtii* et clivages des ARNm

A : Schéma du génome mitochondrial de *C. reinhardtii* avec l'origine de transcription bidirectionnelle représentée par la ligne pointillée et les 2 flèches turquoise. **B** : Schéma des 2 transcrits ARN primaires résultants de la transcription du génome ; Éclairs orange : sites de clivages. **C** : Schéma des 4 populations d'ARN issus de la maturation des transcrits ARN du génome mitochondrial de *C. reinhardtii*. Les ARNt sont annotés avec le code à une lettre : Q : glutamine, W : tryptophane et M : méthionine. Bleu : gènes de protéines ; Rouge : gènes d'ARNt ; Vert clair : gènes d'ARNr de la grande sous-unité du mitoribosome ; Vert foncé : gènes d'ARNr de la petite sous-unité du mitoribosome ; Jaune : gène de l'ARN L2b.

Cette initiation de la transcription est bidirectionnelle et génère deux transcrits primaires polycistroniques. Ces transcrits ont été nommés selon leur sens de transcription, vers la gauche ou la droite, par rapport à l'origine de transcription, LTU (*Leftward Transcriptional Unit*) et RTU (*Rightward Transcriptional Unit*) respectivement (Gray and Boer, 1988). La LTU contient les gènes codant pour les ARNm des gènes *cob*, *nad4* et *nad5*, tandis que la RTU comprend les gènes codant pour les 3 ARNt, tous les ARNr fragmentés, l'ARN L2b et les ARNm *cox1*, *nad2*, *nad6*, *nad1* et *rtl* (Figure I-16B). Cette initiation bidirectionnelle de la transcription se fait grâce à 2 séquences promotrices. Pour la LTU, la séquence promotrice, TTATTATTT, se trouve 3 nt en amont du codon initiateur ATG du gène *nad5*. Pour la RTU, la séquence est légèrement différente, TTATTACAT, et se situe 3 nt en amont de l'ATG du gène *cox1* (Duby *et al.*, 2001).

La transcription se fait généralement par une ARN polymérase ADN dépendante. Les ARN polymérases mitochondriales de l'Homme, les plantes et la levure sont toutes codées par le génome nucléaire, en une seule sous-unité et similaires aux bactériophage T3 ou T7 (Richard *et al.*, 1998). Chez l'Homme, l'ARN polymérase reconnaît, avec l'aide de 2 facteurs de la transcription, TFAM et TFB2M, deux séquences promotrices de la transcription (Shutt, Bestwick and Shadel, 2011). Chez la levure, il n'y a besoin que du facteur MTF1 pour l'initiation de la transcription, qui se fait à 11 sites différents (Basu *et al.*, 2020). Puis, avec un facteur d'élongation, se fait la synthèse d'un ARN polycistronique. Chez les plantes terrestres, 3 ARN polymérases ont été identifiées, l'une spécifique aux mitochondries, la deuxième spécifique des chloroplastes et la troisième ayant une action dans les 2 organelles (Hedtke, Börner and Weihe, 2000; Kühn *et al.*, 2009). De plus, des facteurs similaires à ceux utilisés chez l'Homme sont codés par les génomes nucléaires des plantes, mais leurs fonctions n'ont pas été montrées pour le moment (Kühn *et al.*, 2009; Proulex *et al.*, 2021).

Un seul gène d'ARN polymérase a été identifié chez *C. reinhardtii*, qui pourrait coder pour la transcriptase mitochondriale (Maier *et al.*, 2008), mais il n'a pas été étudié expérimentalement.

Le processus de terminaison de la transcription dans les mitochondries n'est pas encore bien compris à ce jour. Chez les animaux et chez les plantes, des protéines de la famille des mTERF (*mitochondrial Transcription Termination Factor*) ont été identifiées (Miranda, Bonekamp and Kühl, 2022). Elles pourraient bloquer physiquement la transcription en se liant avec l'ARN : la machinerie de la transcription ne pourrait plus passer. Chez *C. reinhardtii*, le gène nucléaire *MOC1* a été identifié. Il code pour une protéine de la famille de mTERF qui est localisée dans

les mitochondries. Il a été montré que la protéine MOC1 se lie spécifiquement à une séquence de 8 nt dans le gène de l'ARNr *S3* et a un rôle de terminaison de la transcription (Wobbe and Nixon, 2013).

2. Maturation des transcrits

Les 2 transcrits primaires polycistroniques doivent être maturés afin d'obtenir des transcrits monocistroniques matures. Pour cela, il y a 2 étapes principales : le clivage des ARN polycistroniques en ARN monocistroniques et les modifications post-transcriptionnelles. Celles-ci consistent en l'ajout de nucléotides.

L'ADNmt de *C. reinhardtii* ou de l'Homme ne contiennent pas d'introns. Il n'y a donc pas d'étape d'épissage des transcrits. Cependant, d'autres organismes ont des introns dans leur ADNmt, comme c'est le cas chez les plantes (Zumkeller, Gerke and Knoop, 2020), chez des éponges (Huchon *et al.*, 2015) ou encore chez les levures (Freel, Friedrich and Schacherer, 2015). Là, un processus d'épissage est essentiel.

De plus, chez les plantes terrestres, il existe un processus absent chez *C. reinhardtii*, la modification des ARN (« *RNA editing* »). Des changements de nt ont lieu à des sites spécifiques, après la transcription de l'ADN en ARN (Hammani and Giegé, 2014). La modification la plus fréquente est celle d'un C en un U (Covello and Gray, 1989; Gualberto *et al.*, 1989; Hiesel *et al.*, 1989). Or, à ce jour, chez les algues, aucun évènement d'édition de l'ARN n'a été observé dans les organelles (Cahoon *et al.*, 2017). L'hypothèse actuelle est que ce processus post-transcriptionnel est apparu plus tard dans l'évolution, au cours de la terrestrialisation.

a) Clivages des transcrits primaires

Le génome mitochondrial de *C. reinhardtii* a beaucoup de similitudes avec celui de l'Homme. Ils sont tous les deux petits et compacts, avec de courtes séquences intergéniques. La transcription du génome mitochondrial humain va également générer 2 transcrits primaires polycistroniques. Chez l'Homme, la plupart des gènes de protéines ou d'ARNr sont suivis d'un gène d'ARNt. Ainsi les clivages des transcrits primaires se font via le modèle de la ponctuation des ARNt. Les RNases P et Z clivent les extrémités 5' et 3' (respectivement) chaque ARNt, individualisant les ARNt mais aussi les autres ARN (Ojala, Montoya and Attardi, 1981). Or chez *C. reinhardtii*, seuls 3 ARNt sont codés par le génome mitochondrial, leurs clivages par la PRORP (*protein-only RNase P*) ne permettent donc pas l'individualisation de tous les ARN : seul le

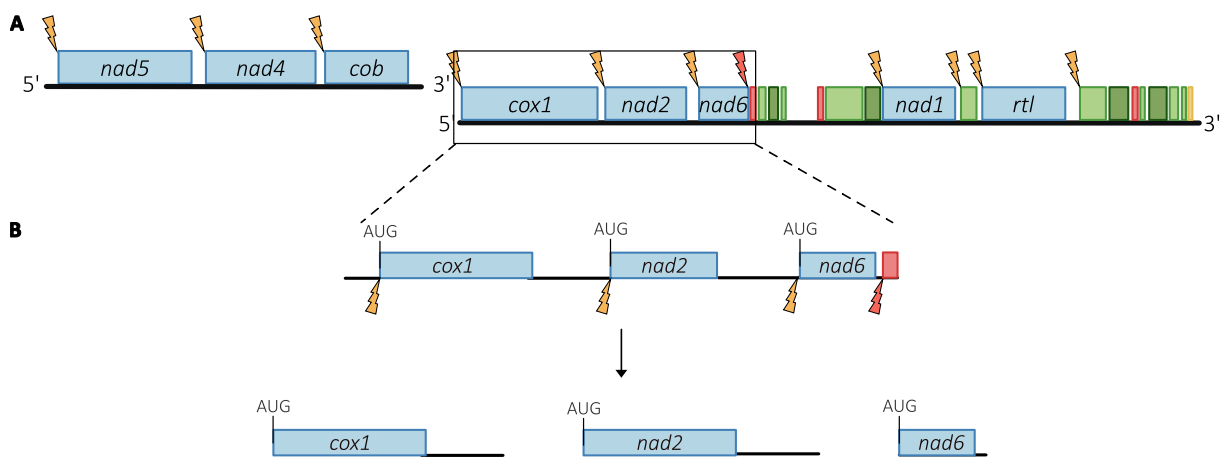


Figure I-17 : Clivages post-transcriptionnels des transcrits de gènes de protéines

A : Schéma des 2 transcrits ARN primaires résultants de la transcription du génome. **B** : Zoom du schéma d'un transcrit primaire avec les clivages des ARNm *cox1*, *nad2* et *nad6*. Éclairs orange : sites de clivages par un ou des enzymes inconnues ; Éclairs rouge : sites de clivages par PRORP ; barre noire : 3'UTR.

transcrit *nad6* est en amont d'un ARNt (Bonnard *et al.*, 2016). Par conséquent, son extrémité 3' sera libérée par la coupure par PRORP (Figure I-17).

Les travaux réalisés dans l'équipe ont montré que les courtes séquences intergéniques, d'une taille allant de 14 et 275 nt, correspondent aux séquences 3'UTR (*Untranslated Transcribed Region*) des ARNm mitochondriaux (Salinas-Giegé *et al.*, 2017). Ainsi un clivage nucléolytique semble générer l'extrémité 3' du gène en amont et l'extrémité 5' du gène en aval (Figure I-17). Ceci implique également que les ARNm mitochondriaux de *C. reinhardtii* n'ont pas de 5'UTR et commencent directement avec le codon initiateur AUG, comme chez l'humain où les transcrits ont soit de très courtes 5'UTR, soit aucune 5'UTR (Brown *et al.*, 2014). On dit qu'ils sont « *leaderless* ».

Les enzymes impliquées dans les clivages des ARNm de *C. reinhardtii* ne sont pas connues. La seule endonucléase mitochondriale connue et étudiée est la RNase P, ou PRORP. Celle-ci agit dans les 3 compartiments, le noyau, le chloroplaste et la mitochondrie (Bonnard *et al.*, 2016). Elle permet le clivage endonucléolytique en 5' des ARNt. D'autres nucléases sont sans doute impliquées, mais leur identité reste à déterminer.

b) Addition post-transcriptionnelle de nt à l'extrémité 3' des transcrits

Dans les bactéries et les archées, la voie de dégradation des ARN se fait grâce à des ajouts de A à l'extrémité 3' des transcrits, par une PAP (*PolyA Polymerase*) ou une PNPase (polynucleotide phosphorylase), en l'absence de PAP (Mohanty and Kushner, 2010). Ces modifications agissent comme des guides pour que les exoribonucléases 3'-5' puissent dégrader l'ARN cible (Levy and Schuster, 2016). De par son origine bactérienne, la polyadénylation des transcrits existe dans la mitochondrie mais son rôle originel a divergé dans certaines espèces (Figure I-18A).

Dans les mitochondries de *S. cerevisiae* et de *Schizosaccharomyces pombe*, l'expression génétique mitochondriale n'implique pas de polyadénylation, tandis que dans les mitochondries de plantes terrestres, le rôle de dégradation des transcrits par la polyadénylation est conservé (Møller, Rasmusson and Van Aken, 2021). En effet, des queues polyA homopolymériques sont ajoutés à l'extrémité 3' des ARNm tronqués, par la PAP AGS1, afin qu'ils soient dégradés (Hirayama *et al.*, 2013). En revanche, chez l'Homme, les ARNm mitochondriaux et les ARNr sont constitutivement adénylés, par la PAPD1 : quelques A dans le cas des ARNr et 40 à 60 adénosines pour les ARNm (Dubin *et al.*, 1982; Pearce *et al.*, 2017).

| A | Nt ajouté | Organisme | Rôle | Enzyme |
|-----|-----------|-----------------------|--|-----------------------|
| A | | <i>A. thaliana</i> | Dégradation | AGS1 |
| | | <i>H. sapiens</i> | Compléter les codons stop Régulation stabilité | PAPD1 |
| | | <i>T. brucei</i> | Stabilisation transcrits en cours d'édition | KPAP1 |
| U | | <i>T. brucei</i> | Uridylation des ARN guides pour l'édition de l'ARN | RET1 |
| A/U | | <i>T. brucei</i> | Permet la traduction des ARNm édités | KPAP1 (A) et RET1 (U) |
| | | <i>C. reinhardtii</i> | Dégradation ? | ?? |
| C | | <i>C. reinhardtii</i> | Stabilisation ? Traduction ? | ?? |

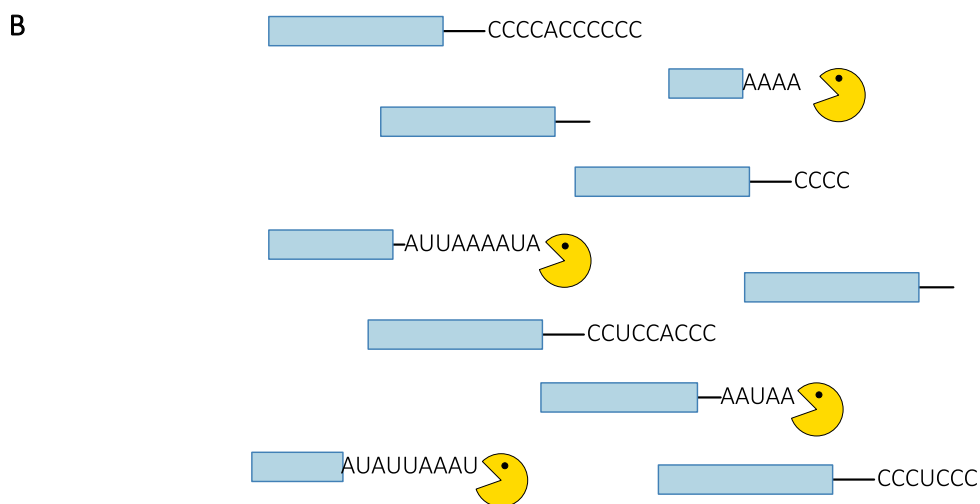


Figure I-18 : Plusieurs types de queues sont retrouvés dans les mitochondries

A : Tableau des différentes queues trouvées dans les mitochondries de 4 organismes, avec leur rôle et le nom de l'enzyme impliquée. Organismes : *A. thaliana* ; *H. sapiens* ; *T. brucei* ; *C. reinhardtii*. ?? : inconnue. **B** : Schéma des différentes extrémités 3' des ARNm mitochondriaux retrouvées dans *C. reinhardtii* : queues polyC, queues polyA/U ou pas de queues. **Bleu** : CDS ; **Noir** : 3'UTR ; **Jaune** : enzyme de dégradation des ARN.

Le clivage de sept ARNm produit des transcrits ayant un codon stop incomplet (Chang and Tong, 2012). La polyadénylation est donc nécessaire pour générer un codon stop fonctionnel. Toutefois, toujours dans les mitochondries humaines, le rôle des queues polyA est plus nuancé et peut, en fonction du transcrit, jouer un rôle dans la stabilité ou dans la dégradation des ARNm (Chang and Tong, 2012; Jedynak-Slyvka, Jabczynska and Szczesny, 2021). La polyadénylation joue un rôle important dans l'expression génétique des mitochondries de *Trypanosoma* (Chang and Tong, 2012). Les transcrits mitochondriaux des trypanosomes sont édités par insertion ou délétion de U et durant l'édition, afin de les stabiliser, ils sont adénylés par une PAP, ajoutant de courtes queues polyA (20-25 nt) (Aphasizhev and Aphasizheva, 2011). Chez *C. reinhardtii*, 3 types d'ajouts post-transcriptionnels de nucléotides ont été identifiés : l'adénylation mais également l'uridylation et la cytidylation (Figure I-18).

L'étude de Zimmer et al. a montré l'existence des polyA homogènes et de queues hétérogènes riches en A à l'extrémité 3' des transcrits mitochondriaux de *C. reinhardtii* (Zimmer et al., 2009). Dans les 2 cas, ces queues semblent être impliquées dans la dégradation des transcrits comme chez les plantes terrestres. Par contre, dans cette même étude, des queues riches en U ont été trouvées. Ces queues polyU-riches sont constituées à 70% de U et 30% de A et ont été principalement observées sur des transcrits tronqués (Zimmer et al., 2009). Ceci suggère que ces queues sont également impliquées dans le processus de dégradation, mais cela reste encore à démontrer. La présence de U à l'extrémité 3' de transcrits avait déjà été référencée chez les mitochondries des trypanosomes (Figure I-18A). En effet, l'achèvement de l'édition des transcrits déclenche l'extension des courtes queues polyA en hétéropolymères riches en A et U, par une TUTase (*terminal uridylate transferases*), permettant la traduction des ARNm (Aphasizheva et al., 2011). Un autre rôle a été attribué à l'uridylation, dans les mitochondries des trypanosomes. Là, l'édition des ARN par ajout ou délétion de U se fait par des ARN guides. Or pour que ces ARN guides soient fonctionnels, il est essentiel qu'ils soient uridylés, à leur extrémité 3' (Aphasizhev, Aphasizheva and Simpson, 2003; Suematsu et al., 2016). Dans ce cas, l'uridylation est essentielle pour l'édition des ARNm.

Enfin, il y a l'ajout de queues riches en C qui ont été observé plus récemment dans les mitochondries de *C. reinhardtii* (Cahoon et al., 2017; Salinas-Giegé et al., 2017; Gallaher et al., 2018). Ce type d'ajout n'avait pas été observé dans l'étude de Zimmer et al. mais cela est probablement dû au fait que pour amplifier les queues polyA complètes, ils ont ligué un

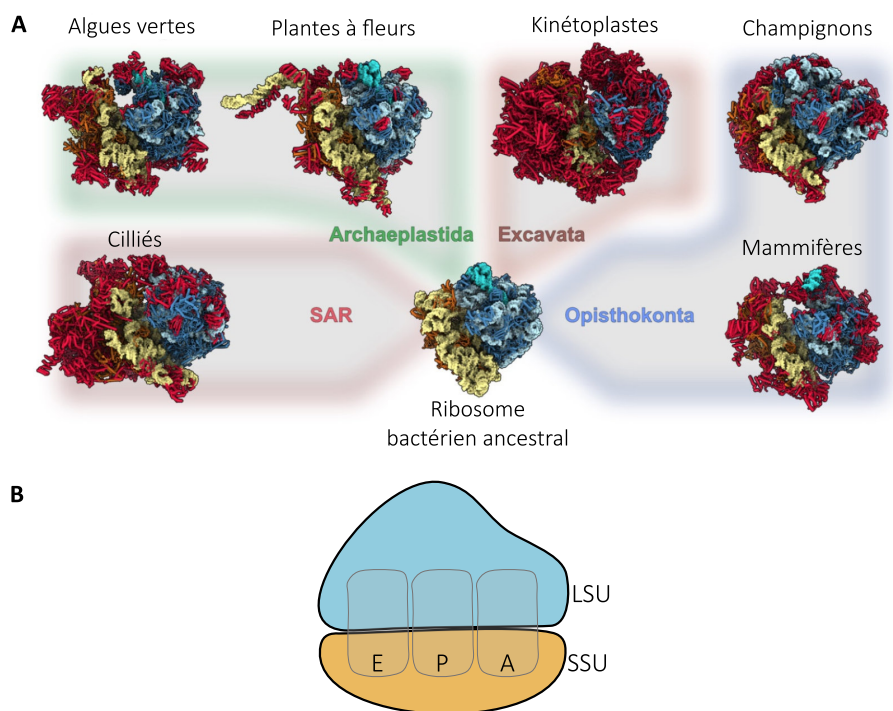
adaptateur polyG en 3' des ARN, ce qui a probablement masqué la présence de ces queues polyC-riches. Ces queues polyC-riches sont également hétérogènes et peuvent contenir des A (20%) ou des U (15%). Leur fonction n'a pas été déterminée pour le moment, mais puisqu'elles sont principalement retrouvées sur les ARNm ayant une séquence UTR complète, elles pourraient par exemple protéger les messagers de la dégradation, améliorant ainsi leur stabilité et générant des ARNm matures et fonctionnels de façon similaire à ce qui est observé pour les queues polyA/U chez les trypanosomes. Il n'y a pas d'autres ajouts naturels de C connus dans la mitochondrie, à ce jour. Seule une étude a montré que l'ajout de C dans les queues d'ARN synthétiques permettaient leur meilleure expression, *in vitro* et *in vivo* chez la souris (Li *et al.*, 2022).

Chez *C. reinhardtii*, les enzymes impliquées dans l'ajout de ces 3 types de queues ne sont pas connues. Il est possible qu'elles fassent partie de la famille des NT (*Nucleotidyl-Transferases*). Ce sont des enzymes qui ajoutent des ribonucléotides, un par un, à l'extrémité 3' des ARN, indépendamment de toute matrice. Parmi elles, on compte notamment les PAP, dont différents membres sont responsables de l'ajout de queues polyA sur les ARNm nucléaires, chloroplastiques ou mitochondriaux (Schuster and Stern, 2009; Hirayama *et al.*, 2013), ainsi que les TUTase ajoutant les queues polyU des ARN nucléaires (Scheer *et al.*, 2016; De Almeida *et al.*, 2018) ou polyA/U des ARNm mitochondriaux des trypanosomes (Aphasizhev, Aphasizheva and Simpson, 2003; Aphasizheva *et al.*, 2011).

Enfin, les PNPases sont également connues pour ajouter des queues hétéropolymériques, notamment dans les chloroplastes (Schuster and Stern, 2009). Si cette dernière paraît être la candidate idéale pour l'ajout des queues polyC-riches hétérogènes, cela ne peut pas être le cas puisque l'unique PNPase de *C. reinhardtii* semble n'agir que dans les chloroplastes (Yehudai-Resheff *et al.*, 2007).

3. Traduction

La traduction est la dernière étape de l'expression du génome mitochondrial, pendant laquelle les ARNm seront lus et traduits en protéines. Elle se fait grâce aux ribosomes, des complexes ribonucléoprotéiques retrouvés chez tous les êtres vivants. Ils lisent les ARN et catalysent l'assemblage des acides aminés afin de former une chaîne peptidique. Cela est possible grâce au code génétique, dans lequel 3 nt forment un codon, et chaque codon correspond à un acide



| C | | <i>S. cerevisiae</i> | <i>H. sapiens</i> | <i>T. brucei</i> | <i>A. thaliana</i> | <i>E. coli</i> |
|----------|-----------|----------------------|---------------------------------|------------------|------------------------------|------------------------------|
| Ribosome | ARN | 2 | 3 | 2 | 3 | 3 |
| | Protéines | 80 | 82 | 127 | 82 | 54 |
| LSU | ARN | 21S (3296 nt) | 16S (1569 nt) ARNt V (75 nt) | 12S (1176 nt) | 26S (3169 nt) 5S (118 nt) | 23S (2904 nt) 5S (120 nt) |
| | Protéines | 46 | 52 | 70 | 42 | 33 |
| SSU | ARN | 15S (1649 nt) | 12S (962 nt) | 9S (620 nt) | 18S (1935 nt) | 16S (1534 nt) |
| | Protéines | 34 | 30 | 57 | 40 | 21 |

Figure I-19 : Diversité des ribosomes mitochondriaux

Adapté de (Scaltsoyiannes *et al.*, 2022a)

A : Comparaison structurale de différents mitoribosomes, avec leur structure en haute résolution et leur composition, selon les lignées eucaryotes. Organismes : ciliés : *T. thermophilus*, algues vertes : *C. reinhardtii*, plantes à fleurs : *A. thaliana*, kinétoplastes : *T. brucei*, champignons : *N. crassa*, mammifères : *H. sapiens* et ribosome bactérien ancestral : *E. coli*. SAR correspond au super-groupe comprenant les sous-groupes des Straménopiles, les Alvéolés et des Rhizaires. LSU : **bleu clair** : ARNr et **bleu foncé** : protéines ribosomales ; SSU : **jaune** : ARNr et **orange** : protéines ribosomales ; **Rouge** : protéines ribosomales spécifiques de chaque organisme. **B** : Schéma d'un ribosome, avec les 2 sous unités (**bleu** : LSU et **orange** : SSU) ainsi que des 3 sites E, P et A en **grisé**. **C** : Tableau de la composition en ARN et en protéines des mitoribosomes de 4 organismes en comparaison du ribosome bactérien de *E. coli*.

aminé. L'association des codons avec leur acides aminés se fait grâce aux ARNt, qui reconnaissent les codons et qui sont chargés spécifiquement avec un acide aminé correspondant à leur séquence anticodon. Les ribosomes aident à la reconnaissance codon-ARNt et ont une activité *peptidyl transferase* permettant de transférer l'acide aminé chargé sur l'ARNt sur la chaîne peptidique en cours de traduction.

a) Mitoribosome

Les mitoribosomes sont les ribosomes mitochondriaux. Ils sont d'origine bactérienne et leur cœur catalytique est composé d'ARNr similaires à ceux des ribosomes bactériens. Cependant, les mitoribosomes ont évolué indépendamment dans les différentes lignées eucaryotes. Il en résulte aujourd'hui des mitoribosomes avec des caractéristiques communes à tous, et d'autres spécifique de chaque lignée ou groupe d'espèces comme leur composition ou leur structure (Figure I-19A) (Waltz and Giegé, 2020).

Ils sont composés de 2 sous-unités : une grande (LSU, *Large SubUnit*) et une petite (SSU, *Small SubUnit*). La petite sous-unité lie l'ARNm et décode l'information génétique, tandis que la grande catalyse la synthèse de la chaîne polypeptidique. Il y a 3 sites de liaison aux ARNt : les sites A (*aminoacyl*), P (*peptidyl*) et E (*exit*) (Figure I-19B). Le site A est le site d'entrée, c'est là que se lient les ARNt chargés après la reconnaissance spécifique du codon correspondant à leur acide aminé. Le site P est d'abord occupé par l'ARNt initiateur, puis à l'ARNt lié à la chaîne peptidique en cours de synthèse. Enfin le site E est le site de sortie des ARNt déchargés de leur acide aminé. Les 2 sous-unités sont des complexes ribonucléoprotéiques, c'est-à-dire qu'elles sont composées d'ARNr et de protéines.

Les ARNr forment le cœur catalytique du mitoribosome. Le contenu en ARNr des mitoribosomes reste similaire aux ribosomes bactériens, bien que leur longueur ou nombre puissent varier. Les gènes d'ARNr sont tous codés par les génomes mitochondriaux (Sloan *et al.*, 2018), et ceux-ci ont subi des insertions ou des délétions. Ces changements de tailles des génomes mitochondriaux se reflètent sur la longueur et le nombre d'ARNr (Figure I-19C). Chez les plantes terrestres où les génomes ont eu des insertions, les ARNr sont plus longs et au nombre de 3 : les ARNr 5S et 26S pour la LSU et l'ARNr 18S pour la SSU (Waltz and Giegé, 2020). A l'inverse, chez les Opisthokontes par exemple, les génomes mitochondriaux ont eu de larges délétions, ils sont donc plus restreints (Amunts *et al.*, 2015; Desai *et al.*, 2017). Les ARNr sont plus courts et au nombre de 2 : 16S (LSU) et 12S (SSU) pour l'Homme et 21S (LSU) et 15S (SSU)

pour la levure. Chez l'Humain, l'ARNr 5S manquant est remplacé par un ARNt valine, tandis que chez la levure, cette perte est compensée par une extension de l'ARNr 21S (Brown *et al.*, 2014; Desai *et al.*, 2017).

De manière générale, les mitoribosomes sont plus larges que les ribosomes bactériens (Figure I-19C). Cela peut être dû à des extensions absentes chez l'ancêtre bactérien, essentiellement au niveau de la SSU, comme chez les trypanosomes ou *A. thaliana*, où elle est finalement plus grande que la LSU (Ramrath *et al.*, 2018; Waltz *et al.*, 2019). De plus, le nombre de protéines associées aux mitoribosomes a grandement augmenté (Waltz and Giegé, 2020). Si 54 sont comptées chez les bactéries *E. coli*, il y en a de 80 chez la levure à 127 chez les trypanosomes. Des pertes de certaines protéines entre les bactéries et les mitoribosomes sont observées, souvent dues à la délétion de l'ARNr avec lequel la protéine était liée. Enfin, des protéines supplémentaires sont trouvées chez les mitoribosomes par rapport aux bactéries. Certaines sont conservées parmi tous les eucaryotes, tandis que d'autres sont spécifiques à des lignées eucaryotiques ou groupe d'espèces (Petrov *et al.*, 2018).

Les mitoribosomes humains et de levures synthétisent essentiellement des protéines membranaires. Celles-ci sont insérées dans la membrane interne au cours de leur traduction, par la protéine translocase OXA1. C'est pourquoi ces mitoribosomes sont attachés à la membrane interne de la mitochondrie. Chez *S. cerevisiae*, cet attachement se fait par une extension d'un ARNr et une protéine de la membrane interne, Mba1 (Jia *et al.*, 2003). Chez l'Homme, l'homologue de Mba1 fait partie de la LSU du mitoribosome (Kummer *et al.*, 2018). Pour *A. thaliana*, cette interaction n'a pas été démontrée, il n'y a pas d'homologue de Mba1, mais le fait que la protéine OXA1 co-purifie avec les mitoribosomes immuno-précipités laisse à penser qu'il y a aussi un attachement à la membrane (Scaltsoyiannes *et al.*, 2022a).

Au début de ma thèse, il y avait peu de connaissances sur le mitoribosome de *C. reinhardtii*. Seuls les ARNr fragmentés étaient connus. En 1988, Boer et Gray montrent que les ARNr mitochondriaux de l'algue sont discontinus, fragmentés en 4 modules pour l'ARNr 16S-like de la SSU et 8 pour l'ARNr 23S-like de la LSU (Boer and Gray, 1988). Ces modules sont dispersés dans le génome mitochondrial, séparés par les gènes d'ARNm et d'ARNt. De plus, contrairement aux gènes des ARNr de la SSU, les gènes des modules de la LSU ne sont pas dans l'ordre dans lequel les domaines correspondant devraient s'assembler. Un modèle de la structure secondaire de ces ARNr assemblés avait été proposé en 1988 (Boer and Gray, 1988).

Dans ce modèle, les fragments d'ARNr sont stabilisés essentiellement par appariement de bases. De plus, il était admis que le mitoribosome de *C. reinhardtii* n'avait pas d'ARNr 5S, puisqu'aucune étude n'avait réussi à identifier un ARNr 5S codé par le génome mitochondrial (Boer and Gray, 1988; Denovan-Wright and Lee, 1995; Nedelcu, 1997; Salinas-Giegé *et al.*, 2017).

b) Code génétique

Le code génétique universel établit la correspondance entre un codon et l'acide aminé pour lequel il code. Seuls 3 codons ne codent pas pour un acide aminé, les codons stop UAA, UGA et UAG permettant aux ribosomes d'arrêter la traduction. Il est dit universel puisqu'il est appliqué dans la très grande majorité des cellules vivantes. Des biais de ce code génétique ont été observés chez les mitochondries des vertébrés et des levures par exemple. Chez les vertébrés et chez les levures, le codon UGA code pour un tryptophane et non un stop, et AUA code pour la méthionine au lieu de l'isoleucine (Barrell, Bankier and Drouin, 1979; Sulo *et al.*, 2017). De plus, chez les levures, 4 codons codent pour la thréonine au lieu de la leucine (Sulo *et al.*, 2017). A l'inverse, chez les mitochondries des plantes, les études montrent que le code génétique universel est appliqué (Jukes and Osawa, 1990; Møller, Rasmusson and Van Aken, 2021).

Le code génétique universel est utilisé dans les mitochondries de *C. reinhardtii*, mais son usage est biaisé (Michaelis, Vahrenholz and Pratje, 1990). En effet, 9 codons ne sont jamais utilisés. De plus, si les 7 gènes codant des protéines de la chaîne respiratoire sont similaires en termes d'usage des codons, le gène *rtl* diffère, avec 4 codons retrouvés uniquement dans ce gène (CTC (Leu), ATA (Ile), ACA (Thr) et GGG (Gly)). Le codon UGA, codant pour un codon stop dans le code génétique universel, est totalement absent. Il reste donc 2 codons stop : UAA et UAG. En revanche, tous les gènes commencent avec le codon d'initiation AUG.

c) ARNt et les aaRS

Le génome mitochondrial de *C. reinhardtii* code pour 3 gènes d'ARNt, correspondant à la glutamine, la méthionine et le tryptophane. De fait, l'importation des ARNt manquants depuis le cytosol est essentiel pour assurer le bon déroulement du processus de traduction. Le génome nucléaire de cette algue code pour 49 ARNt iso-accepteurs, et 34 sont trouvés dans les mitochondries, montrant qu'il y a une sélection des ARNt importés (Vinogradova *et al.*, 2009).

Les taux d'importations sont variables pour chaque iso-accepteur, allant de 0,2% à 95%. Il y a une corrélation entre les populations d'ARNt importés dans la mitochondrie et l'usage de codons de l'organite. Cela serait dû à une coévolution entre l'import des ARNt et l'usage des codons, plutôt qu'à une adaptation dynamique du transport des ARNt cytosoliques dans la mitochondrie (Salinas *et al.*, 2012). Par ailleurs, l'ADNmt de *C. reinhardtii* code pour un ARNt méthionine élongateur et l'ARNt méthionine initiateur codé par le génome nucléaire n'est pas importé (Vinogradova *et al.*, 2009). Un cas similaire a lieu dans les mitochondries de *T. brucei* : certains ARNt méthionine élongateurs sont formylés dans l'organelle afin d'être utilisés en tant qu'ARNt initiateur (Tan *et al.*, 2002). Il est possible que ce processus soit similaire dans les mitochondries de *C. reinhardtii*, puisque l'enzyme ajoutant le groupe formyle à l'ARNt méthionine chargé, une *methionyl-tRNA^{Met} formyltransferase*, a été identifiée dans l'organite (Vinogradova *et al.*, 2009)

Mais les ARNt seuls ne suffisent pas pour la traduction : ils doivent être chargés avec l'acide aminé leur correspondant. Cela est fait par des enzymes spécifiques, les aaRS (*aminoacyl-tRNA synthetase*). Elles sont codées par le génome nucléaire (Sissler *et al.*, 2013). Puisque la synthèse protéique se fait de manière indépendante dans les 2 ou 3 compartiments cellulaires (noyau, mitochondrie et chloroplaste pour les plantes), il faudrait au minimum 60 gènes d'aaRS : 3 gènes pour chaque acide aminé. Or ce n'est pas le cas chez les plantes. Seulement 33 gènes ont été identifiés chez *C. reinhardtii* (Cognat *et al.*, 2013). Pour 13 acides aminés, il y a 2 gènes d'aaRS pour un même acide aminé, tandis que pour les 7 restants, il y a un gène d'aaRS pour un acide aminé. Cela implique un partage des aaRS qui sont doublement ou triplement localisées dans la cellule. Ce modèle, commun aux règne des plantes, contredit ce qui est observé chez les animaux et les champignons, où peu de cas de double localisation ont été observés (Duchêne, Pujol and Maréchal-Drouard, 2009).

d) Déroulement de la traduction

Le déroulement de la traduction peut être décomposé en 3 étapes : l'initiation, l'élongation et la terminaison.

Initiation

Pour l'initiation de la traduction dans les mitochondries, la SSU du ribosome est associée à un ARNt méthionine formylé au niveau de son site P. Le complexe SSU-ARNt se lie à l'extrémité 5'

de l'ARNm et l'ARNt reconnaît le codon AUG initiateur, permettant de positionner correctement la SSU au niveau de ce codon AUG. Puis, la LSU est recrutée et l'élongation peut commencer.

Chez les bactéries, 3 facteurs d'initiations sont nécessaires pour ces étapes, IF1, IF2 et IF3 (*initiation factor*). IF1 est universellement absent dans les mitochondries, alors que des orthologues mitochondriaux de IF2 et IF3 sont universels. Chez les mammifères, une extension dans mIF2 (*mitochondrial IF2*) compense l'absence de IF1. mIF2 a donc l'équivalent fonctionnel des facteurs bactériens IF1 et IF2 (Atkinson *et al.*, 2012). Deux possibles facteurs d'initiation mitochondriaux ont été identifiés chez *C. reinhardtii*, MIF2 (Cre08.g376800) et MIF3 (Cre09.g398215) dans la version 6.1 de l'annotation du génome nucléaire de l'algue, cependant leur fonction n'a pas été étudiée pour le moment.

La fixation de l'ARNm et la reconnaissance du codon AUG sont fait de différentes manières selon les organismes étudiés. Chez les bactéries, des séquences conservées dans les 5'UTR des ARNm, appelées « séquences Shine-Dalgarno », permettent le bon placement de la SSU sur le codon initiateur (Rodnina, 2018). Mais les ARNm mitochondriaux des différents organismes étudiés montrent des extrémités 5' différentes, suggérant des mécanismes différents pour le positionnement de l'ARNm dans la SSU. Chez *S. cerevisiae*, des facteurs activateurs de la traduction reconnaissent les extrémités 5' des ARN et interagissent avec les protéines de la SSU, probablement pour initier la traduction (Herrmann, Woellhaf and Bonnefoy, 2012). Les ARNm mitochondriaux des mammifères n'ont pas de 5'UTR, et leurs 3'UTR sont réduites. Dans ce cas, une protéine PPR (*pentatricopeptide repeat*) permet la liaison entre l'ARNm et la SSU, par son interaction avec une séquence riche en U présente à l'extrémité 5' de la séquence codante (Kummer and Ban, 2021). Chez les plantes terrestres, les ARNm ont de longues séquences 5'UTR ne contenant pas de séquence Shine-Dalgarno (Hammani and Giegé, 2014). Mais il est probable qu'elles contiennent un autre type de séquence afin d'initier la traduction, riches en A (motif AxAAA) qui pourraient être reconnu par une protéine PPR (mS83) de la SSU du mitoribosome (Waltz *et al.*, 2020). Chez *C. reinhardtii*, les ARNm n'ont pas non plus de 5'UTR (Salinas-Giegé *et al.*, 2017). Une étude des années 1990 avait trouvé des RBS (*ribosome binding site*), en amont de l'AUG (Colleaux *et al.*, 1990). Or nous savons maintenant qu'en amont d'un AUG se trouve la 3'UTR du gène précédant avant clivage, et que les ARNm matures n'ont pas de séquence 5'UTR (Salinas-Giegé *et al.*, 2017). Ces RBS seraient donc présentes dans les 3'UTR

des transcrits. De plus, une étude plus récente montre qu'une partie des ARNm associés aux ribosomes sont circularisés : cette circularisation pourrait permettre d'avoir la séquence RBS proche de l'AUG, afin que le ribosome puisse se placer correctement pour l'initiation de la traduction (Cahoon and Qureshi, 2018). Bien qu'intéressante, cette théorie reste à démontrer.

Élongation

Durant la traduction, le ribosome se déplace le long de l'ARNm, qui est traduit codon par codon en la séquence polypeptidique correspondante. L'élongation de la traduction inclus 3 étapes : sélection de l'ARNt chargé avec le bon acide aminé complémentaire à l'ARNm, formation de la liaison peptidique et translocation du module ARNm-ARNt. L'élongation semble être l'étape la plus conservée dans les mitochondries des différents organismes étudié et vis-à-vis des systèmes bactériens ou cytosoliques.

Le facteur mtEFTu (*mitochondrial Elongation Factor Tu*) permet l'acheminement de l'ARNt aminoacylé au site A du mitoribosome, par hydrolyse de GTP. Une fois l'ARNt livré, mtEFTu quitte le mitoribosome pour être recyclé : il sera réassocié avec du GTP grâce à mtEFTs (*mitochondrial Elongation Factor Ts*), et pourra à nouveau se lier à un ARNt chargé afin de l'emmener au mitoribosome (Worjax *et al.*, 1997; Xu, Liu and Song, 2022). Une liaison peptidique se forme entre l'acide aminé nouvellement arrivé dans le site A et la chaîne peptidique naissante présente dans le site P.

Dès que ce lien est formé, le mitoribosome doit se repositionner sur le codon suivant. Il doit bouger de 3 nt précisément, afin qu'il n'y ait pas de décalage du cadre de lecture. C'est la translocation du mitoribosome sur l'ARNm. Cela se fait par un mouvement de rotation de la tête de la SSU qui va permettre le positionnement du dernier acide aminé ajouté à la chaîne peptidique dans le site P, l'évacuation de l'ARNt déchargé par le site E et l'exposition nouveau codon dans le site A, afin qu'il soit décodé. Là, le cycle d'élongation recommence. Ce processus est catalysé par le mtEFG (*mitochondrial Elongation Factor G*) chez les bactéries. Dans les mitochondries humaines et de levures, il y a 2 paralogues de ce facteur, mtEFG1 et mtEFG2 (Hammarsund *et al.*, 2001; Tsuboi *et al.*, 2009). Leurs fonctions sont différentes. mtEFG1 catalyse la translocation du complexe ARNm-ARNt durant l'élongation, tandis que mtEFG2 est impliqué dans le recyclage du mitoribosome, à la fin de la traduction (Tsuboi *et al.*, 2009).

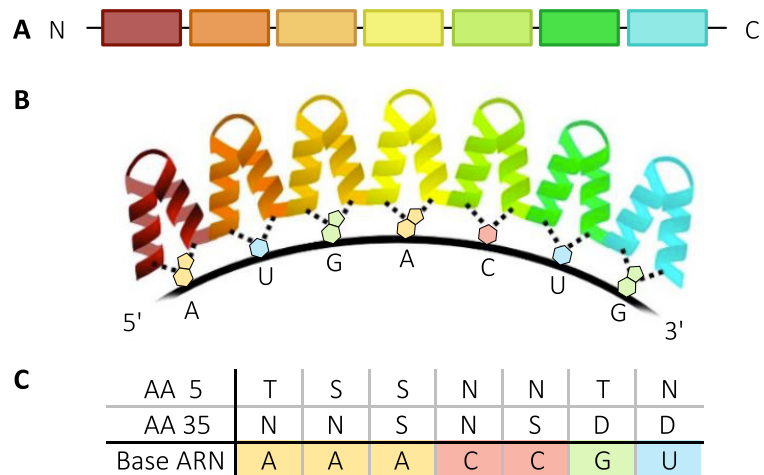


Figure I-20 : Structure des PPR en hélices α et code PPR

Adapté de (Waltz, 2018) et de (Barkan and Small, 2014).

A : Schéma d'une protéine PPR avec un motif répété 7 fois, chaque répétition ayant une couleur différente. **B** : Schéma de l'interaction entre la protéine et son ARN (courbe **noire**) substrat. Chaque motif est replié en 2 hélices α , de couleur différente. Chaque motif reconnaît une base d'ARN spécifique. **C** : Tableau du code PPR avec les combinaisons des 2 aa (en position 5 et 35) pour la reconnaissance des bases ARN (même code couleur que dans **B**). AA : code à une lettre : T : thréonine, S : sérine ; N : asparagine ; D : aspartate.

Des possibles orthologues aux facteurs mtEFTU (Cre06.g259150) et mtEFG (Cre03.g174000) ont été identifiés dans la version 6.1 de l'annotation du génome nucléaire de *C. reinhardtii*, cependant leur fonction n'a pas été étudiée pour le moment.

Terminaison

Lorsque le mitoribosome arrive à la fin d'une ORF (*open reading frame*), la traduction est terminée par la reconnaissance d'un codon stop. Cela se fait par un facteur de relâchement ribosomal (*ribosomal release factor*), qui se lie au codon stop et libère la chaîne peptidique néo-synthétisée du site P. La terminaison de la traduction est l'étape qui a le plus divergé, tant au niveau du code génétique que pour les facteurs impliqués.

Chez les bactéries, il y a 2 facteurs de terminaison, RF1 et RF2 qui reconnaissent les codons stop (RF1 pour UAA et UAG et RF2 pour UAA et UGA), au niveau du site A afin de libérer le polypeptide (Prabhakar *et al.*, 2023). Chez les archées et dans le cytosol eucaryote, il n'y a qu'un seul facteur (aRF1 et eRF1 respectivement) (Chrzanowska-Lightowlers, Pajak and Lightowlers, 2011). Des homologues de RF1 sont identifiés dans les mitochondries. Chez l'Homme, il y en a 2 : mtRF1a qui décode les codons stop UAA et UAG, et mtRF1 qui décode les codons stop non canoniques AGA et AGG, par un décalage du cadre de lecture (Krüger *et al.*, 2023). Chez *S. cerevisiae*, il n'y a que mRF1 qui reconnaît les codons stop (Kutner *et al.*, 2008). Chez les plantes ou les algues, aucune donnée n'est disponible à ce jour.

4. Facteurs *trans* impliqués dans l'expression génétique mitochondriale

a) Les protéines PPR

Les protéines PPR sont des protéines à répétition hélicoïdale, appartenant à la super famille des α solénoïdes. C'est un groupe de protéines codées par le noyau. Elles présentent une architecture modulaire typique, comprenant des répétitions en tandem d'un domaine sous forme de 2 hélices α antiparallèles (Figure I-20). Plusieurs familles de protéines appartenant à ce groupe sont impliquées dans le métabolisme des ARN (Hammani *et al.*, 2013). Une famille est particulièrement impliquée dans le métabolisme des ARN des organelles : les protéines PPR. Grâce aux 2 hélices antiparallèles, les protéines PPR interagissent avec des ARN : chaque motif reconnaît une base d'ARN spécifique (Figure I-20B).

Les protéines PPR contiennent des motifs de 35 acides aminés, dont certains sont conservés. La plupart des protéines PPR ont une activité de liaison à l'ARN simple brin, de manière

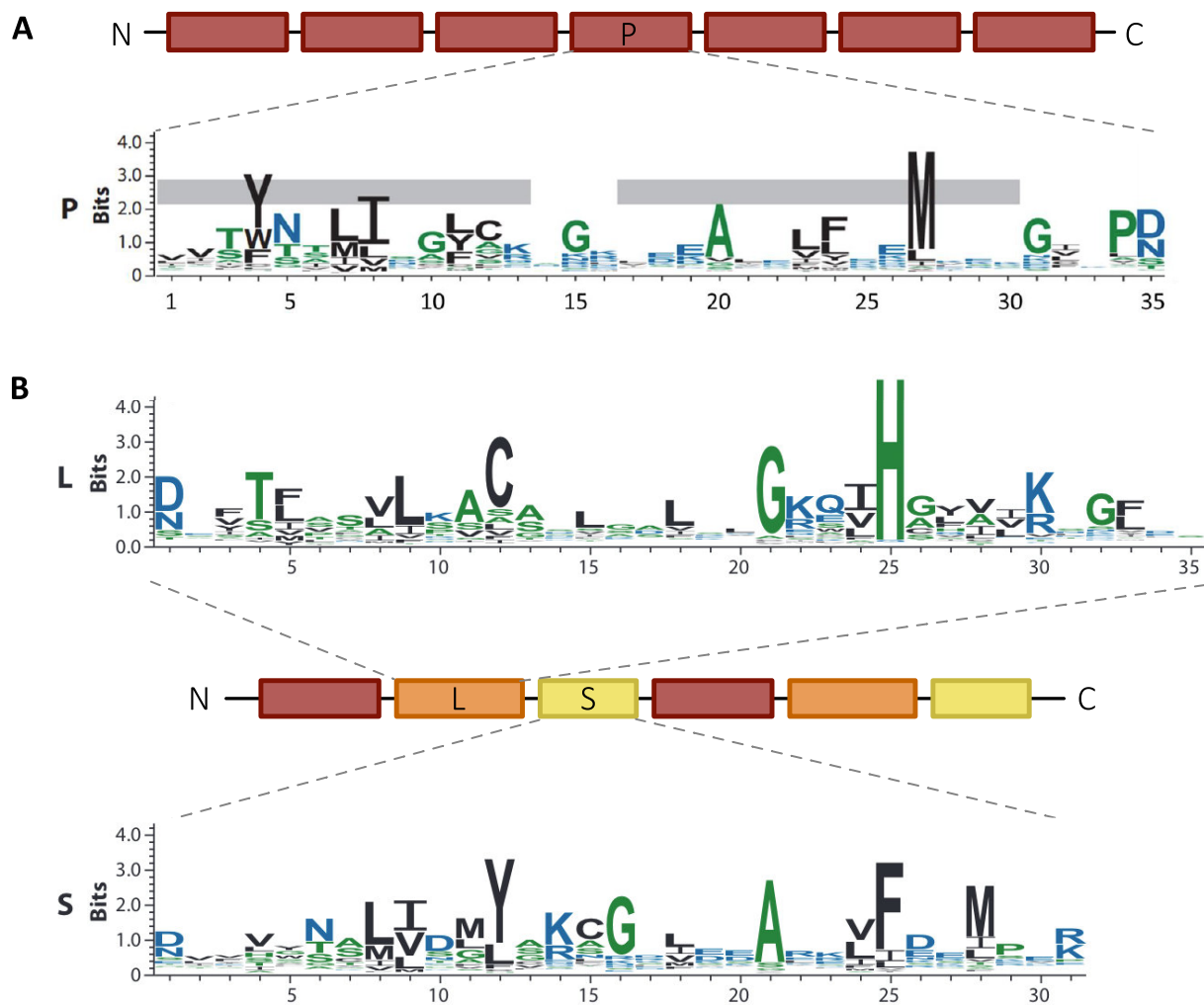


Figure I-21 : Différents types de domaines PPR et leur séquence consensus

Adapté de (Barkan and Small, 2014).

A : Schéma d'une protéine et séquence logo d'un motif PPR de classe P. Les barres grises schématisent les acides aminés impliqués dans les 2 hélices α . **B** : Schéma d'une protéine et séquence logo d'un motif PPR de classe PLS.

Rouge : motif PPR de classe P ; **Orange** : motif PPR de classe PLS, motif L ; **Jaune** : motif PPR de classe PLS, motif S. Plus la lettre du résidu est grande, plus le résidu est abondant.

séquence spécifique. Cela semble corréluer au « code PPR », où chaque répétition peut reconnaître un nucléotide spécifique, grâce à 2 acides aminés en position 5 et 35 (Figure I-20B et C) (Cheng *et al.*, 2016). Par exemple, un résidu thréonine en position 5 combiné à un résidu asparagine en position 35 vont reconnaître un nt adénine, tandis qu'un résidu asparagine en position 5 avec un résidu aspartate en position 35 reconnaîtront une base uracile (Barkan and Small, 2014). Puisqu'il n'y a que 4 bases dans l'ARN, le code PPR est dégénéré : plusieurs combinaisons de résidus en position 5 et 35 peuvent reconnaître la même base.

Le nombre de protéines PPR diffère grandement selon les organismes. Chez *A. thaliana* par exemple, on en compte 496 (Cheng *et al.*, 2016), alors qu'il n'y en a que 7 chez l'Homme (Lightowers and Chrzanowska-Lightowers, 2013), 15 chez les levures (Lipinski *et al.*, 2011) et 14 chez *C. reinhardtii* (Tourasse, Choquet and Vallon, 2013). Leur grand nombre dans les plantes terrestres serait lié à l'expansion de l'édition de C en U des ARNm, processus répandu dans les organelles de ces organismes (Barkan and Small, 2014). L'immense majorité des PPR sont organellaires (Lurin *et al.*, 2004; Colcombet *et al.*, 2013).

Il existe 2 classes de protéines PPR, en fonction de la séquence du motif répété (Cheng *et al.*, 2016). Les types P (*pure PPR*) sont composées strictement de motifs répétés de 35 acides aminés, (Figure I-21A). Les types PLS possèdent des répétitions d'un triplet composés d'un motif P canonique (*pure*, 35 acides aminés) et de 2 variants du motif P, un motif L (*long*) plus grand et un motif S (*short*) plus petit (Figure I-21B) (Cheng *et al.*, 2016). Le motif L possède 35 ou 36 aminés, avec une séquence consensus différente. Le motif S comporte en moyenne 31 acides aminés. Les PPR de type PLS ont souvent des domaines supplémentaires en C-terminal : les domaines E ou DYW. Le domaine DYW est le site catalytique de l'enzyme pour l'édition des ARN des C en U, tandis que la fonction du domaine E n'est pas identifiée (Salone *et al.*, 2007; Takenaka *et al.*, 2021). Alors que les protéines PLS-PPR sont essentiellement impliquées dans l'édition des ARNm, les PPR de type P ont des rôles divers dans l'expression des génomes des organelles (Hammani *et al.*, 2013). De manière non exhaustive, elles peuvent interférer dans la transcription (Lightowers and Chrzanowska-Lightowers, 2013), les clivages (Zhou *et al.*, 2017), l'édition (Okuda and Shikanai, 2012), l'épissage (Wang *et al.*, 2018; Chen *et al.*, 2019), ou la stabilité des ARN (Prikryl *et al.*, 2011; Haili *et al.*, 2013; Zoschke *et al.*, 2016),

la traduction (Prikryl *et al.*, 2011; Herbert, Golik and Bonnefoy, 2013) ou même faire partie des mitoribosomes (Amunts *et al.*, 2015; Waltz *et al.*, 2019).

Chez *C. reinhardtii*, 14 PPR ont été identifiées (Tourasse, Choquet and Vallon, 2013), mais seulement 4 ont été étudiées expérimentalement (Macedo-Osorio, Martínez-Antonio and Badillo-Corona, 2021). Ces 4 PPR sont localisées dans le chloroplaste. MCA1 est impliquée dans la stabilisation de l'ARNm *petA* (Loiselay *et al.*, 2008), MRL1 dans celle de *rbcl* (Johnson *et al.*, 2010). TCB1 est nécessaire pour la traduction de l'ARNm *petB*, en plus de le stabiliser (Cavaiuolo *et al.*, 2017). Enfin, PPR7 participe à la stabilisation de 4 ARNm, ainsi qu'à la maturation des ARN de 3 clusters de gènes (Jalal *et al.*, 2015). Sur les 10 protéines PPR restant à étudier, 3 ont une prédiction de localisation mitochondriale (PPR3, PPR5 et PPR10) (Tourasse, Choquet and Vallon, 2013). Cela dit, TBC1 était prédite mitochondriale alors qu'elle est en fait chloroplastique. Ces prédictions sont donc à confirmer.

b) Les protéines OPR

Comme les protéines PPR, les protéines OPR (*Octatricopeptide Repeat*) sont codées par le génome nucléaire et font aussi partie de la famille des α solénoïdes. Les protéines OPR contiennent des répétitions en tandem d'un domaine de 38 acides aminés sous forme de 2 hélices α antiparallèles (Figure I-22A). Ce sont également des protéines permettant des interactions avec des ARN simple brin (Macedo-Osorio, Martínez-Antonio and Badillo-Corona, 2021). Cette famille de protéines est retrouvée chez les plantes, avec des fonctions similaires aux protéines PPR. Contrairement aux PPR, leur nombre est limité dans les plantes terrestres, avec une seule OPR chez *A. thaliana* par exemple, impliquée dans la maturation de l'ARNr 16S chloroplastique (Kleinknecht *et al.*, 2014). Par contre, les protéines OPR sont majoritairement trouvées dans les Chlorophycées, avec 127 identifiées chez *C. reinhardtii* (Jarrige, 2019). Cela suggère que les protéines OPR et PPR étaient présentes dans l'ancêtre commun aux Chlorophycées et aux plantes terrestres, puis il y a eu une expansion des protéines OPR dans les Chlorophycées et des PPR chez les plantes terrestres (Macedo-Osorio, Martínez-Antonio and Badillo-Corona, 2021). 19 OPR ont été étudiées dans cette algue, elles sont localisées dans le chloroplaste (Macedo-Osorio, Martínez-Antonio and Badillo-Corona, 2021). Leurs fonctions sont similaires aux PPR : la maturation et la stabilisation (Kleinknecht *et al.*, 2014; Wang *et al.*, 2015; Viola *et al.*, 2019), l'épissage (Balczun *et al.*, 2005; Merendino *et al.*, 2006), ou encore

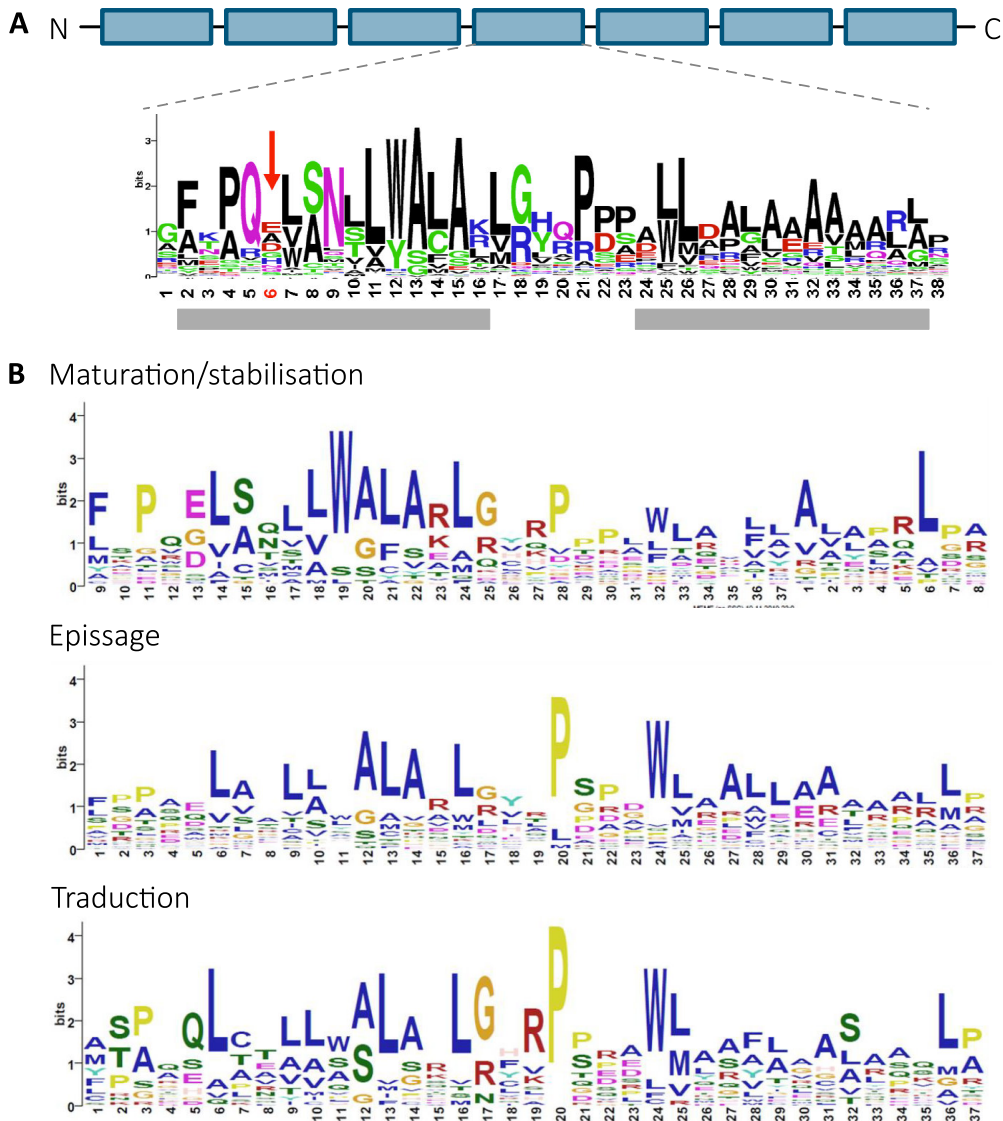


Figure I-22 : Protéines OPR et leurs séquences consensus

Adapté de (Jarrige, 2019).

A : Schéma d'une protéine OPR et séquence logo d'un motif OPR d'organisme photosynthétique. Les barres grises schématisent les acides aminés impliqués dans les 2 hélices α . La flèche rouge indique la position 6, supposée essentielle pour l'interaction avec l'ARN. **B** : Séquences logo des motif des protéines OPR selon leur rôle dans la cellule de *C. reinhardtii*.

Plus la lettre du résidu est grande, plus le résidu est abondant.

l'initiation de la traduction (Auchincloss *et al.*, 2002; Rahire *et al.*, 2012; Lefebvre-Legendre *et al.*, 2015).

A ce jour, un « code OPR » n'a pas été identifié, nous ne connaissons donc pas le fonctionnement des OPR vis-à-vis de la spécificité de leur interaction avec les ARN, mais l'acide aminé en position 6 pourrait être essentiel pour les interactions avec l'ARN cible (Jarrige, 2019). Début des années 2010, une séquence consensus des OPR avait été déterminé comme étant PPPEW (Eberhard *et al.*, 2011; Rahire *et al.*, 2012). Une nouvelle version de cette séquence, obtenue à partir d'un plus grand nombre d'OPR analysées, montre un deuxième motif, encore plus conservé que PPEW : LWALA (Figure I-22A) (Jarrige, 2019). Cette même étude montre des différences entre les séquences consensus des OPR selon leur rôle dans la cellule : facteurs impliqués dans la maturation et la stabilisation des ARNm, dans l'épissage ou dans la traduction (Figure I-22B). Ces différences pourraient venir du fait que les facteurs de la traduction n'interagissent que brièvement avec les ARN, tandis que les facteurs impliqués dans les étapes de maturation ont des interactions plus stables avec les ARN. Enfin, certaines OPR contiennent un domaine supplémentaire, appelé domaine RAP (*RNA binding abundant in Apicomplexa*). Ce dernier a des similarités structurales avec une endoribonucléase bactérienne (Boulouis *et al.*, 2015; Boehm *et al.*, 2017). Les protéines OPR-RAP pourrait être impliquées dans les clivages des ARN (Cartalas, Coudray and Gobert, 2022).

PROJET DE THESE

C'est dans ce contexte de compréhension des mécanismes de l'expression du génome mitochondrial de *C. reinhardtii* que s'inscrit mon sujet de thèse.

Mon projet doctoral a été de chercher à déterminer quelles sont les enzymes impliquées dans l'ajout des queues polyC. Comme indiqué précédemment, l'équipe a montré que les ARNm mitochondriaux matures de cette algue possèdent des queues polyC, ajoutées post-transcriptionnellement à l'extrémité 3'UTR (Salinas-Giegé *et al.*, 2017). Cette particularité a été trouvée uniquement dans les mitochondries des Chlorophycées. Toutefois leur fonction n'a pas encore été déterminée et une première approche était d'identifier la ou les protéines impliquées dans cette modification. Cela nous a amené à nous intéresser à des protéines de la famille des NT. Cette famille de protéines ajoute des ribonucléotides en 3' des ARN, indépendamment de toute matrice. Des études phylogénétiques ont montré qu'il y avait 4 NT spécifiques aux Chlorophycées chez *C. reinhardtii*. Mon objectif était d'étudier ces 4 protéines candidates, afin de déterminer la ou lesquelles sont responsables de l'ajout des queues polyC sur les ARNm mitochondriaux de *C. reinhardtii*.

En parallèle, j'ai participé à un travail collaboratif réalisé entre mon équipe, l'équipe du Docteur Giegé (IBMP, Strasbourg) et l'équipe du Docteur Hashem (Institut Européen de Chimie et Biologie, INSERM, Bordeaux). Au début de ma thèse, pour les plantes, seul le mitoribosome d'*A. thaliana* avait été étudié (Waltz *et al.*, 2019) et aucune donnée n'était disponible pour d'autres organismes photosynthétiques tels que les algues vertes. Or des études de plus en plus nombreuses sur les mitoribosomes de plusieurs espèces ont montré que leur architecture globale, leur composition en ARNr et en protéines divergeaient dans les différentes lignées eucaryotiques (Scaltsoyianes *et al.*, 2022a). *C. reinhardtii* possède des ARNr et des ARNm avec des caractéristiques spécifiques qui rendent l'étude de son mitoribosome attractive. Ce travail nous a permis de reconstituer par des analyses protéomiques et par cryo-EM (*cryo electron microscopy*), les deux sous-unités du mitoribosome et d'en proposer un modèle. Ainsi, nous avons pu déterminer la composition du mitoribosome et cela a mené à l'identification de 11 protéines spécifiques de *C. reinhardtii* dans les LSU et SSU. Mon travail a donc été d'obtenir des mutants pour certaines de ces protéines afin de les caractériser.

MATERIEL & METHODES

MATERIEL

I. Matériel végétal

1. Souches de *Chlamydomonas reinhardtii*

a) CC-4351

cw15 arg7-8 mt+

Cette souche a une haute efficacité de transformation. Ne synthétisant pas l'arginine, cette auxotrophie est utilisée pour sélectionner les transformants. Elle n'a pas de paroi cellulaire, ce qui facilite la plupart des expérimentations.

b) CC-5325

cw15 mt-

Cette souche est la souche d'origine des mutants *dum* (*Dark Uniparental transmission by the Minus parent* (Matagne *et al.*, 1989)) et CLiP (*Chlamydomonas Library Project* - chlamylibrary.org (Li *et al.*, 2019)).

c) Mutants *dum*

Les mutants *dum5* (Cardol, Matagne and Remacle, 2002) et *dum11* (Dorthu *et al.*, 1992) nous ont été donné par la Professeure Claire Remacle. Ce sont des mutants mitochondriaux, avec une mutation dans les gènes *nad5* et *cob*, respectivement.

d) Mutants CLiP

Les mutants ont été commandés sur le site chlamylibrary.org (Li *et al.*, 2016, 2019). Ils ont été obtenus par insertion aléatoire d'une cassette contenant le gène de résistance à la paromomycine dans des souches de *C. reinhardtii* CC-5325. Ils sont répertoriés en [Annexe I](#).

2. *Nicotiana benthamiana*

Des plantes de *Nicotiana benthamiana* âgées de 5 semaines ont été utilisées pour les expériences de transformation transitoire par agro-infiltration.

II. Souches bactériennes

1. Escherichia coli

a) TOP10

ThermoFisher Scientific (#C404010) ; F⁻ mcrA Δ(mrr-hsdRMS-mcrBC) φ80lacZΔM15 ΔlacX74 recA1 araD139 Δ (ara-leu) 7697 galU galK rpsL(Str^R) endA1 nupG

Ces bactéries sont rendues compétentes par un traitement au calcium, effectué par l'ingénieure Mathilde Arrivé (IBMP) et la technicienne Elodie Ubrig (IBMP). Leur efficacité de transformation est de 10⁸-10⁹ cfu/μg plasmide contrôle.

b) DH10-β

NEB (#C3019I) ; F⁻ mcrA Δ(mrr-hsdRMS-mcrBC) φ80lacZΔM15 ΔlacX74 recA1 araD139 Δ (ara-leu) 7697 galU galK λ⁻ rpsL(Str^R) endA1 nupG

Ces bactéries sont rendues compétentes par un traitement au chlorure de rubidium, effectué par le Docteur Patryk Ngondo (IBMC (Institut de Biologie Moléculaire et Cellulaire), anciennement IBMP) et la technicienne Christina Berrissou (IBMP). Leur efficacité de transformation est supérieure à 1x10⁹ cfu/μg plasmide contrôle. Cette souche est adaptée pour les transformations avec de grands plasmides (> 10 kb).

c) Rosetta II (DE3) pLysS

F⁻ ompT hsdSB(RB- mB-) gal dcm λ(DE3 [lacI lacUV5-T7 gene 1 ind1 sam7 nin5]) pLysSRARE (CamR)

Ces bactéries compétentes sont préparées par la plateforme P3P (Purification et Production des Protéines) de l'IBMP. Cette souche a été utilisée pour l'expression et la purification de protéines. Elle dérive de la souche BL21 mais possède un plasmide supplémentaire, le pLysS, codant pour des ARNt reconnaissant des codons fréquents chez les eucaryotes mais rares chez les procaryotes. Ce plasmide code également pour le lysozyme du phage T7, permettant de s'affranchir de l'expression basale hors conditions d'induction. Cette souche est résistante au chloramphénicol.

2. *Agrobacterium tumefaciens*

a) *GV3101 pMP90*

Ces bactéries sont utilisées pour transformer les plantes. Elles sont préparées et rendues compétentes par un traitement au calcium par le Docteur Vianney Poignavent (Serendip innovations, anciennement IBMP) et par la technicienne Christina Berrissou. Elles sont résistantes à la rifampicine et à la gentamycine et contiennent le plasmide Ti (*Tumor induced*) pMP90 désarmé permettant l'expression des gènes de virulence nécessaires au transfert de l'ADN-T dans le génome de la plante hôte.

b) *GV3101 pMP90 P19*

Ces bactéries ont été obtenues par la Docteure Morgane Michaud (LPCV (Laboratoire Physiologie Cellulaire et Végétale), anciennement IBMP) (Michaud *et al.*, 2014). Elles expriment la protéine P19 qui supprime le processus de silencing des ARN *in planta*. Cela permet une meilleure expression des constructions lors des agro-infiltrations de feuilles de *Nicotiana benthamiana*.

c) *GV3101 pMP90 pSU9-RFP*

Ces bactéries ont été obtenues par la Docteure Morgane Michaud (Michaud *et al.*, 2014). Elles expriment la RFP (protéine fluorescente rouge ; *Red Fluorescent Protein*) fusionnée aux 69 premiers acides aminés de la sous-unité 9 de l'ATP (adénosine-triphosphate) synthase. Cela permet d'avoir un contrôle de la localisation mitochondriale en feuille, lors des agro-infiltrations de feuilles de *N. benthamiana*.

III. Vecteurs

1. pChlamiRNA2

Nom abrégé pour cette thèse : pHUN, qui est le nom originel du vecteur avant publication, en référence au prénom de son créateur, Attila.

Ce vecteur a été développé par l'équipe de David Baulcombe (Molnar *et al.*, 2009) pour exprimer un précurseur miARN (micro ARN) ciblant un gène d'intérêt, sous le contrôle du promoteur hybride constitutif *HSP70A-RBCS2* et de la région 3'UTR du gène *RPL12*. Un gène de résistance à l'ampicilline permet de sélectionner les bactéries transformées. Un site unique de

restriction SpeI est situé proche du précurseur miARN pour pouvoir cloner la séquence d'intérêt. Le vecteur contient aussi un site de restriction unique PvuI, pour linéariser le plasmide avant transformation dans *C. reinhardtii*.

2. pGEM[®]-T Easy Vector System I (Promega)

Il est conçu pour cloner des fragments d'ADN amplifiés avec des polymérases ayant une activité terminale transférase, qui ajoute une désoxy-adénosine en 3'. En effet, ce vecteur possède des résidus désoxy-thymidine en 3', permettant une hybridation spécifique A-T. Il contient le gène codant pour la β -lactamase qui confère aux bactéries transformées la résistance à l'ampicilline. La cassette de clonage se trouve à l'intérieur du gène *LacZ*, qui permet la sélection de bactéries transformées selon le principe de l' α -complémentation.

3. pDONR207 (Gateway – Invitrogen)

Il possède les séquences attP permettant le premier évènement de recombinaison avec les sites attB de l'insert nécessaire pour obtenir un clone d'entrée produisant des sites attL. Le gène de résistance à la gentamycine permet de sélectionner les clones d'entrée obtenus. Il contient le gène *ccdB*, inhibiteur de la gyrase bactérienne qui permet la sélection négative des vecteurs non recombinés.

4. pB7FWG2 (Gateway)

Développé par (Karimi, Inzé and Depicker, 2002), il est utilisé pour déterminer la localisation intracellulaire d'une protéine d'intérêt, clonée dans le vecteur. C'est un plasmide binaire qui possède les séquences attR permettant l'évènement de recombinaison nécessaire avec les sites attL du vecteur d'entrée pour l'obtention du clone de destination. Il possède le gène de la EGFP (protéine fluorescente verte améliorée ; *Enhanced Green Fluorescent Protein*) fusionné en 3' de la séquence insérée. L'ensemble est sous le contrôle du promoteur 35S du virus de la mosaïque du chou-fleur. Il contient également le gène de résistance à la spectinomycine.

5. pUPD2 (Golden Braid)

C'est le vecteur utilisé pour le niveau 0 du clonage Golden Braid (Sarrion-Perdigones *et al.*, 2013). Il est produit par la plateforme P3P de l'IBMP. Il comporte un gène de résistance au chloramphénicol. L'intégration des séquences d'intérêt se fait par digestion/ligation avec l'enzyme de restriction Bsmbl.

6. pBACT α 1 (Golden Braid)

C'est le vecteur utilisé pour le niveau 1 du clonage Golden Braid (Sarrion-Perdigones *et al.*, 2013), permettant l'expression de la protéine d'intérêt dans les bactéries. Il est produit par la plateforme P3P de l'IBMP. Il code pour le gène *ccdB*, inhibiteur de la gyrase bactérienne qui permet la sélection négative des vecteurs non recombinés. Il comporte un gène de résistance à l'ampicilline et un promoteur inductible à l'IPTG (isopropyl- β -D-1-thiogalactopyranoside) pour l'expression. L'intégration des séquences d'intérêt se fait par digestion/ligation avec l'enzyme de restriction Bsal.

IV. Oligonucléotides

Les oligos (oligonucléotides) ont été synthétisés par l'entreprise IDT. Ils sont listés selon les expériences pour lesquels ils ont été utilisés en [Annexe II](#).

V. Gène synthétique

La séquence PAP4 pour la production de protéine a été synthétisée par ATG biosynthetics GmbH.

VI. Anticorps

Les différents anticorps utilisés ainsi que leurs conditions d'utilisation sont listés en [Annexe III](#).

1. Goat@Mouse

L'anticorps de chèvre dirigé contre les anticorps de souris provient d'Invitrogen (#G21040). Il est conjugué à la peroxydase du raifort (*horseradish peroxidase*), permettant la révélation du Western blot par émission de lumière (chimiluminescence) lorsqu'elle est en contact avec un agent luminescent.

2. Goat@Rabbit

L'anticorps de chèvre dirigé contre les anticorps de lapins provient d'Invitrogen (#G21234). Il est conjugué à la peroxydase du raifort (*horseradish peroxidase*), permettant la révélation du Western blot par émission de lumière (chimiluminescence) lorsqu'elle est en contact avec un agent luminescent.

3. Mouse@PDH E1

L'anticorps de souris dirigé contre la protéine de la matrice mitochondriale PDH E1 (*pyruvate dehydrogenase E1*) de *Z. mays* a été donné par la Docteur Géraldine Bonnard (IBMP). La taille apparente de la protéine est de 40 kDa.

4. Rabbit@Cytf

L'anticorps de lapin dirigé contre la protéine chloroplastique Cytf (Cytochrome *f*) de *C. reinhardtii* a été acheté à la société Agrisera (#AS06 119). La taille apparente de la protéine est de 31 kDa.

5. Rabbit@NAB1

L'anticorps de lapin dirigé contre la protéine cytoplasmique NAB1 (*Nucleic acid binding protein 1*) de *C. reinhardtii* a été acheté à la société Agrisera (#AS08 333). La taille apparente de la protéine est de 36 kDa.

6. Rabbit@PAP4

L'anticorps dirigé contre la protéine PAP4 (partie centrale, contenant les domaines NTD et PAP) a été produit par la société Covalab. Ce sont des anticorps polyclonaux de lapin immunopurifiés avec la protéine PAP4 utilisée pour immuniser les lapins. La taille apparente est de 110 kDa.

7. Rabbit@PAP6

L'anticorps dirigé contre la protéine PAP6 (partie centrale, contenant les domaines NTD (*nucleotidyl-transferase domain*) et PAPD (*PAP associated domain*)) a été produit par la société Covalab. Ce sont des anticorps polyclonaux de lapin immunopurifiés avec la protéine PAP6 utilisée pour immuniser les lapins. La taille apparente est de 80 kDa.

8. Rabbit@VDAC1

L'anticorps de lapin dirigé contre la protéine des membranes mitochondriales VDAC1 de *C. reinhardtii* a été produit par la Docteure Thalia Salinas-Giegé (Salinas *et al.*, 2012). La taille apparente de la protéine est de 28 kDa.

VII. Outils informatiques

1. MacVector

<https://macvector.com>

Logiciel pour la visualisation des séquences (nucléotidiques et protéiques) et l'analyse des résultats de séquençage. La licence est fournie par l'IBMP.

2. Blast (NCBI)

<https://blast.ncbi.nlm.nih.gov/Blast.cgi>

D'une manière générale, les alignements de séquence ont été réalisés avec les outils BLAST (*Basic Local Alignment Search Tool*) développés et mis à disposition par le NCBI (*National Center for Biotechnology Information*)

- Nucleotide BLAST : alignement de séquences nucléiques
- Protein BLAST : alignement de séquences protéiques
- Smart BLAST : recherche de séquences protéiques similaires ou proche de la séquence d'intérêt, dans tous les organismes modèles.
- Multiple Alignment : alignement de plusieurs séquences simultanément en se basant sur les domaines/séquences en commun.

3. Primer-BLAST

<https://www.ncbi.nlm.nih.gov/tools/primer-blast/>

Pour le design d'amorces de PCR (*Polymerase Chain Reaction*).

4. Universal Probe Library (Roche)

Pour le design des amorces de RT-qPCR (*Reverse Transcription Quantitative PCR*). L'outil n'est à ce jour plus disponible.

5. LightCycler480® Software (Roche)

<https://diagnostics.roche.com/global/en/products/instruments/lightcycler-480-ins-445.html>

Pour le suivi de l'amplification en RT-qPCR et pour le recueil des données : C_t (*threshold cycle*) et T_m (*melting temperature*)

6. LinRegPCR

<https://mybiosoftware.com/linregpcr-analysis-quantitative-pcr-data.html>

Pour le calcul des efficacités des amorces en RT-qPCR, par la méthode de régression linéaire (Ramakers *et al.*, 2003).

7. WMD3

<http://wmd3.weigelworld.org/cgi-bin/webapp.cgi?page=Designer;project=stdwmd>

Pour le choix des régions ciblées par l'amiRNA (*artificial micro RNA*) et pour le design des oligos à commander.

8. Agilent 2100 expert software

<https://www.agilent.com/en/product/automated-electrophoresis/bioanalyzer-systems/bioanalyzer-software/2100-expert-software-228259>

Pour le recueil et l'analyse des données avec le BioAnalyzer (Agilent).

9. Phytozome V13

<https://phytozome-next.jgi.doe.gov>

Base de données des génomes de plus de 300 plantes, dont *C. reinhardtii*, avec les versions d'assemblage du génome *Chlamydomonas reinhardtii* v5.6 et *Chlamydomonas reinhardtii* CC-4532 v6.1

10. PredAlgo

<https://lobosphaera.ibpc.fr/cgi-bin/predalgotdb2.perl?page=main>

Pour les prédictions de localisation des protéines candidates

11. TargetP

<https://services.healthtech.dtu.dk/services/TargetP-2.0/FTAIR>

Pour les prédictions de localisation des protéines candidates

12. ZEN lite (Zeiss)

<https://www.zeiss.fr/microscopie/produits/microscope-software/zen-lite.html>

Pour l'acquisition des données de microscopie, avec le microscope LSM780

13. ImageJ

<https://imagej.nih.gov/ij/index.html>

Le logiciel ImageJ a été utilisé pour l'analyse d'images de microscopie et la quantification de signaux sur gels ou membranes par la méthode de l'aire sous la courbe.

Le plugin FigureJ a permis la réalisation de figures, lisant les images aux formats .LSM et .LIF et gardant les métadonnées associées.

14. Affinity Designer V1

<https://affinity.serif.com/fr/designer/>

Ce logiciel de graphisme a été utilisé pour le dessin de figures.

VIII. Microscopie

Les images de microscopie sont acquises avec un microscope confocal Zeiss LSM780 via le logiciel ZEN Lite (Zeiss). Elles sont traitées avec les logiciels ZEN Lite (Zeiss) et ImageJ.

Le tableau ci-dessous indique les paramètres d'excitation et d'émission utilisés.

| | Chlorophylle A | EGFP | Mitotracker | mRFP |
|-------------------------|----------------|---------|-------------|---------|
| Laser d'excitation (nm) | 633 | 488 | 561 | 561 |
| Spectre d'émission (nm) | 647-721 | 493-558 | 570-650 | 579-640 |

METHODES

Les milieux et solutions en gras dans le texte sont décrits en [Annexe IV](#). Sauf indication contraire, tout est fait à température ambiante. L'eau utilisée pour l'ensemble des expérimentations est de l'eau ultrapure filtrée par des distributeurs Milli-Q® autoclavée.

I. Techniques relatives aux organismes photosynthétiques

1. Cultures de *Chlamydomonas reinhardtii*

Toutes les cultures se font entre 22 et 25°C.

Au début de ma thèse, la lumière continue était émise par des néons, la luminosité variait entre 35 et 60 $\mu\text{mol.m}^{-2}.\text{s}^{-1}$. Puis nous avons changé pour des LED, avec une luminosité de 65 $\mu\text{mol m}^{-2} \text{s}^{-1}$.

Toutes les souches sont repiquées en **milieu TAP** (Tris Acétate Phosphate) **agar** tous les 2 mois, mises en lumière continue (Néons puis LED) pendant 1 semaine puis en lumière très faible (1 $\mu\text{mol.m}^{-2}.\text{s}^{-1}$) jusqu'au prochain repiquage.

Des repiquages frais sont faits lorsque les souches doivent être utilisées : 3 à 4j (jours) avant utilisation, en lumière continue.

Des cultures liquides sont faites en **milieu TAP**. Ces cultures sont alors soit laissées en lumière continue sous agitation, soit mises en alternance lumière 16h (60 $\mu\text{mol.m}^{-2}.\text{s}^{-1}$) et obscurité 8h, en agitation, afin de synchroniser les cellules en une semaine.

2. Quantification des *C. reinhardtii*

Trois méthodes ont été utilisées :

a) Estimation par la chlorophylle

1 mL de culture est centrifugé 1 min (minute) à 16000g. Le culot est repris dans 1 mL d'éthanol 100% puis centrifugé à nouveau. Les absorbances du surnageant sont mesurées à 652 et 664 nm avec un spectrophotomètre UV-1800 (Shimadzu). La concentration en chlorophylle, en $\mu\text{g/mL}$, est déterminée par le calcul suivant :

$$\text{Conc. chlorophylle} = (24,3 \times A_{652\text{nm}}) + (3,26 \times A_{664\text{nm}})$$

Sachant que 25-30 µg de chlorophylle correspondent environ à 10^7 cellules, on en déduit un nombre moyen de cellules/mL.

b) Compteur de cellules

10 µL de culture cellulaire sont déposées sur une lamelle EVE™ cell counting slide (NanoEntek), et la lamelle est insérée dans l'appareil EVE™ Plus (NanoEntek).

Le comptage des cellules se fait avec un focus automatique, avec les paramètres suivants :

- Sensitivity : 1
- Min size : 5
- Max size : 50
- Circularity : 70.

Le protocole du fournisseur recommande de diluer au demi les échantillons avec du bleu trypan, ainsi l'appareil donne le nombre de cellules vivantes, mortes et total. Ce nombre prend en compte la dilution au demi, et donc l'appareil applique un facteur x2. Or, les échantillons de *C. reinhardtii* ne sont pas dilués avec du bleu trypan. Le nombre de cellules/mL total est donc pris en compte et il faut le diviser par 2 afin d'avoir un nombre réel.

c) Estimation par la turbidité

Le nombre de cellule/mL peut aussi être déterminé en fonction de la turbidité de la culture. Pour cela, il faut mesurer les absorbances à 750 nm des cultures. Il faut ensuite appliquer le calcul suivant pour avoir la concentration cellulaire en cellules/mL :

$$\text{Conc. cellulaire} = \frac{DO_{750\text{ nm}} - 0,088}{9 \times 10^{-8}}$$

Cette équation provient du protocole du kit GeneArt® Cryopreservation Kit for Algae (Invitrogen #A24228).

3. Transformation nucléaire de *C. reinhardtii*

Les cellules CC-4351 sont transformées avec l'électroporateur Neon™ transfection system (Invitrogen #MPK5000), avec un Kit de système de transfection Neon™ 100 µL (Invitrogen #MPK10025) et le système GeneArt® MAX Efficiency® Transformation Reagent for Algae (Invitrogen # A24229). Pour se faire, des cultures liquides sont mises en synchronisation pendant une semaine, avec des repiquages tous les 2j en alternance phase lumineuse / phase

d'obscurité. Puis une culture de 200 mL est mise en lumière continue afin d'obtenir 2×10^6 cellules/mL. Les cellules sont culottées par centrifugation 5 min à 1000g, re-suspendues dans 10 mL de GeneArt® MAX Efficiency® Transformation Reagent, puis re-culottées de la même manière. Le culot d'algue est repris dans 1,3 mL de GeneArt® MAX Efficiency® Transformation Reagent (Invitrogen # A24229) afin d'avoir 3×10^8 cellules/mL. Des aliquotes de 120 μ L de cellules sont faits, auxquels est ajouté 1 μ g de vecteur linéarisé (ou un volume similaire de GeneArt® MAX Efficiency® Transformation Reagent pour le contrôle). Le mélange est incubé 5 min à 4°C. L'électroporation se fait comme indiqué dans le protocole fournisseur, sous 2300V, à 13 ms et 3 pulses. Les cellules électroporées sont incubées dans un tube 1,5 mL pendant 15 min, puis transférées dans des erlens de 25 mL avec 4 mL de **milieu TAP-Sucrose** et incubées 24h en lumière continue sous agitation. Après quoi, les cellules sont centrifugées 5 min à 1000g, re-suspendues dans 200 μ L de **milieu TAP-Sucrose**, déposées sur boîtes de **milieu TAP agar** et mises en culture en lumière continue jusqu'à ce que des colonies poussent, après environ 1 semaine et demie.

4. Cryopréservation de *C. reinhardtii*

La cryopréservation a été faite avec le kit GeneArt® Cryopreservation Kit for Algae (Invitrogen #A24228). Pour cela, des cultures des souches à cryogéniser sont faites pour avoir des cultures à 3×10^6 cellules/mL. Les cellules sont repiquées en plaques 6 puits dans 4,5 mL de **milieu TAP** et 100 μ L de Cryopreservation Reagent A (Invitrogen #A24228) pour avoir une absorbance à 750 nm de 0,1. Les plaques sont ainsi incubées pendant 3j sous agitation et en lumière continue. La concentration cellulaire est estimée par la méthode de la turbidité avec la $DO_{750 \text{ nm}}$ ([Méthodes I-2-c](#)). Les cellules sont culottées par centrifugation à 1000g pendant 5 min et re-suspendues dans du Cryopreservation Reagent B (Invitrogen #A24228) pour avoir une concentration finale de $2,5 \times 10^7$ cellules/mL. Le mélange est incubé 45 min puis des échantillons de 240 μ L sont aliquotés dans des tubes CryoPure (Sarstedt #72.379). Ces derniers sont placés dans la boîte Mr. Frosty™ Freezing Container (ThermoScientific #5100-0050), sans l'éponge. Aucun espace vide ne devant rester dans la boîte, il faut remplir avec des tubes vides si besoin. Le tout est placé à -80°C, pendant 4h minimum, sans ouverture du congélateur. Puis les tubes peuvent être rangés dans les boîtes habituelles à -80°C.

5. Tests phénotypiques de *C. reinhardtii*

a) Lumière/obscurité

Les colonies sont repiquées en **milieu TAP agar** pendant 3 à 5j. Puis elles sont re-suspendues dans 2 mL de **milieu TAP liquide** et mises en culture durant 24h à la lumière, sous agitation. La densité des cultures est mesurée par spectrophotomètre à une DO_{750nm} , et les cultures sont diluées pour avoir $DO_{750nm} = 1,5$. Elles sont incubées 1h à 30°C à l'obscurité, en agitation douce. Ces suspensions sont utilisées comme matériel de départ (= 1) pour faire des dilutions en cascade. 10 μ L de chaque dilution sont déposés sur 2 boîtes de **milieu TAP agar**, la première étant placée en culture dans des conditions mixotrophiques (lumière + acétate) durant 5 à 7j et la deuxième durant 10 à 15j dans des conditions hétérotrophiques (obscurité + acétate), afin de tester les capacités de croissance à l'obscurité des transformants.

b) Milieu minimum

A partir des dilutions faites pour les tests lumière/obscurité, la culture de départ non dilué est déposée sur une boîte de **milieu TP agar**. Les boîtes sont placées en lumière continue durant 10 à 15j. Ce milieu ne contient pas de source d'acétate : si la photosynthèse d'une souche est affectée, elle ne pourra pas pousser.

6. Purification de mitochondries de *C. reinhardtii*

Toutes les étapes sont effectuées à 4 °C.

$3 \cdot 10^9$ cellules de *C. reinhardtii* sont centrifugées 10 min à 1000g et re-suspendues dans 10 mL de **solution A**. Dans un erlenmeyer de 50 mL, la suspension cellulaire est incubée dans un bain marie à 50°C de sorte que la température de la suspension atteigne rapidement 30°C. L'erlen est immédiatement transféré dans un bain marie à 30°C pendant 30s puis refroidi 5 min dans de la glace. La suspension cellulaire est alors centrifugée 2 min à 2500g et le culot est lavé dans 30 mL de **solution B**, puis centrifugée 2 min à 1000g. Le culot est repris dans 2 mL de **solution B** et les cellules sont brisées mécaniquement par agitation forte avec un barreau magnétique pendant 45s. 6 mL de **solution D** sont ajoutés et le tout est centrifugé 3 min à 2000g pour culoter les débris cellulaires. Le surnageant est filtré à travers un tissu de filtration en nylon NITEX avec une ouverture de maille de 50 μ m (Dutscher) et centrifugé 10 min à 10000g, en faible décélération. Le culot est repris dans 150 μ L de **solution D**. La suspension est déposée sur un **gradient de percoll (13/21/45%)** et centrifugée 1 h à 40000g, sans frein.

Les mitochondries, entre les couches 21 et 45%, sont prélevées et diluées au dixième dans la **solution MET**. Le tout est centrifugé 10 min, à 11000g. Le culot de mitochondries est lavé deux fois dans 2 mL de **solution MET** et centrifugé 5 min à 9000g. Il est finalement repris dans 400 µL de **solution MET**.

Les mitochondries sont quantifiées en équivalent protéines par la méthode de Bradford ([Méthodes V-1](#)). Les mitochondries sont culotées par centrifugation 5 min à 9000g, et le culot est congelé rapidement dans de l'azote liquide et conservé à -80 °C.

Pour les mutants de la banque CLiP, les mitochondries ont été obtenues sans faire de gradient, ces échantillons sont donc enrichis en mitochondries et non pas purifiés.

7. Cultures de *N. benthamiana*

Les plantes sont cultivées en terre selon un cycle de 8h d'obscurité à 18 °C et 16h de lumière à 22 °C.

8. Agro-infiltration de *N. benthamiana*

Une culture liquide de 10 mL de **milieu LB** contenant les antibiotiques de sélection estensemencée à partir d'une colonie d'agrobactérie et incubée au moins 20h à 28°C sous agitation. La culture bactérienne est centrifugée à 3200g pendant 10 min et le culot bactérien est re-suspendu dans 5 mL de **tampon d'infiltration**. Une dilution au dixième d'une aliquote de la suspension bactérienne permet de mesurer l'absorbance à 600 nm, au spectrophotomètre. Les suspensions bactériennes sont alors diluées pour obtenir une absorbance de 0,5 à 600 nm pour la construction P19, de 0,3 pour la pSU9-RFP et de 1 pour les constructions d'intérêt. Les bactéries sont incubées au moins 1h dans le **tampon d'infiltration**, pour que les gènes de virulence soient exprimés, grâce à l'acétosyringone.

La face abaxiale de 2 à 3 feuilles situées après la première paire apicale de feuilles de *N. benthamiana* âgées de 5 semaines sont agro-infiltrées, avec une seringue de 1 mL, après avoir fait une légère blessure à l'aide d'une aiguille. Les plantes sont ensuite remises en culture pendant 48h avant observation.

9. Observation de feuilles agro-infiltrées de *N. benthamiana*

Un disque de feuille infectée d'environ 8 mm de diamètre est découpé. Il est déposé sur une lame de microscopie en verre, côté adaxial de la feuille contre la lame et recouvert de 10 µL

d'eau. Une lamelle est posée sur la feuille, le tout est scellé par du ruban adhésif. L'espace entre lame et lamelle est rempli d'eau, puis le vide est fait avec une pompe à vide, afin que l'eau pénètre dans le matériel végétal et permette l'observation au microscope.

10. Préparation de protoplastes de *N. benthamiana*

Les protoplastes sont préparés à partir de feuilles agro-infiltrées avec la construction uniquement. Une feuille par construction lacérée finement avec un scalpel est recouverte de **tampon de digestion 1X** (volume variable selon la taille de la feuille) et incubée 4h à 28°C sous agitation. Si nécessaire, les protoplastes sont incubés avec 150 nM de MitoTracker™ Orange CMTMRos (Invitrogen ; abrégé « Mitotracker ») pendant 15 min sous légère agitation à l'obscurité. 15 µL de protoplastes sont montés entre lame et lamelle pour observations.

II. Techniques de cultures et transformation des bactéries

1. Cultures d'*E. coli*

En **milieu LB-agar**, ces bactéries sont cultivées à l'obscurité soit une nuit à 37°C, soit pendant 2 à 3j à température ambiante. Elles peuvent alors être gardées pendant 2 mois à 4°C.

En **milieu LB**, ces bactéries sont cultivées dans 3 à 200 mL selon les besoins, durant une nuit à 37°C, sous agitation et à l'obscurité.

2. Transformation d'*E. coli*

Les trois souches compétentes ([Matériel II-1](#)) sont transformées par choc thermique. 10 µL de plasmide sont ajoutés à 50 µL de bactéries compétentes et laissés 30 min sur glace. Le choc thermique se fait en incubant les bactéries à 42°C pendant 40s puis en les plaçant immédiatement dans de la glace, pendant 2 min. 250 µL de **milieu LB** (*lysogeny broth*) sont ajoutés et le tout est incubé 1h à 37°C, sous agitation. La suspension bactérienne est alors étalée sur des boîtes de Pétri contenant du **milieu LB-agar** avec 100 µg/mL d'ampicilline ou 50 µg/mL de carbénicilline. Les boîtes sont incubées une nuit à 37°C, à l'obscurité.

3. Cultures d'*A. tumefaciens*

En **milieu LB-agar**, ces bactéries sont cultivées 2j à 28°C dans l'obscurité. Elles peuvent alors être gardées pendant 2 mois à 4°C.

En **milieu LB**, ces bactéries sont cultivées dans 3 à 50 mL selon les besoins, durant 24 à 48h à 28°C, en agitation et à l'obscurité.

4. Transformation d'*A. tumefaciens*

800 ng de pB7FWG2-insert purifié sont ajoutés à 100 µL de bactéries compétentes et incubés pendant 10 min sur la glace. Après un choc thermique de 5 min à 37°C, 1 mL de **milieu LB** est ajouté et les bactéries sont mises en culture pendant 2h à 28°C, sous agitation. La suspension bactérienne est ensuite étalée sur du **milieu LB-Agar** contenant 25 µg/mL de rifampicine, 100 µg/mL de spectinomycine et 15 µg/mL de gentamycine. Les boîtes sont incubées pendant 2j à 28°C.

5. Glycérol-stocks de bactéries

Des glycérols-stocks sont faits à partir de cultures liquides de bactéries, *E. coli* ou *A. tumefaciens*. 500 µL de glycérol 50% (v/v) sont ajoutés à 500 µL de culture liquide de bactéries dans un tube de 1,5 mL. Le mélange est refroidi rapidement dans de l'azote liquide puis conservé à -80°C.

III. Techniques liées à l'ADN

1. Extraction d'ADN brut de *C. reinhardtii* pour PCR de génotypage

Un mélange réactionnel de 25 µL est préparé avec 15 µL d'eau, 5 µL de protéinase K (20 mg/mL) et 5 µL de Tampon 5X Colorless GoTaq® Flexi Buffer (Promega # M890A). Sont ajoutés environ 5 mm² de colonie de *C. reinhardtii* fraîches (repiquage entre 1 et 4 jours avant). Le mélange est vortexé 5 secondes à Vmax (vitesse maximale), puis incubé 1h à 58°C et 1h à 96°C. Le mélange est alors centrifugé 1 min à Vmax avec une petite centrifugeuse de paillasse (Heathrow Scientific Sprout) et le surnageant contenant l'ADN est transféré dans un nouveau tube.

2. Amplifications par PCR

Toutes les amorces utilisées pour les différentes PCR sont listées en [Annexe II-1](#).

a) PCR de génotypage sur colonie de *C. reinhardtii*

La PCR est réalisée dans un milieu réactionnel de 25 µL comprenant 5 µL de tampon 5X Green GoTaq® Flexi Buffer (Promega #M891A), 0,5 µL de dNTP 10mM, 1 µL de chaque amorce à

10 μM , 1,25 μL de MgCl_2 25mM (Promega #A351B), 1,25 μL de DMSO, 1 μL de Taq polymérase faite maison. A ceci ont été ajoutés 2,5 μL d'extrait à la protéinase K.

L'amplification est réalisée dans un thermocycleur Biometra® T3000 si l'expérience est faite en barrettes de microtubes 0,2mL ou dans un thermocycleur Bio-Rad T100™ Thermal Cycler si l'expérience est faite en plaque 96 puits.

L'amplification débute par une étape de dénaturation initiale de 5 min à 95°C, suivie de 35 itérations du cycle suivant :

- 30s (secondes) à 95°C, pour dénaturer l'ADN
- 30s à une température = $T_m - 5^\circ\text{C}$, pour hybrider les amorces
- 30s à 2 min (selon la taille du fragment à amplifier) à 72°C.

Une étape d'élongation de 5 min à 72°C termine la réaction.

b) PCR pour les constructions Gateway

Même protocole que pour les PCR de génotypage, mais avec 0,25 μL de GoTaq® G2 Flexi DNA Polymerase (Promega #M780A). La matrice ADN utilisée ici est 1 μL d'ADNc (ADN complémentaire) (voir [Méthodes IV-7](#)) ou 70 ng de plasmide pGEM-insert pour l'ajout des sites attB1 et attB2.

c) PCR sur colonies bactériennes

Même protocole que pour les PCR de génotypage, mais avec 0,25 μL de GoTaq® G2 Flexi DNA Polymerase (Promega #M780A). La colonie de bactéries à analyser est piquée avec un cure-dent en bois, repiquée sur une nouvelle boîte LB agar puis le cure-dent est plongé dans le milieu de PCR pour déposer des bactéries dans le tube. La nouvelle boîte est mise en culture à 37°C sur la nuit.

3. Électrophorèse sur gel d'agarose ADN

Les produits de PCR sont chargés sur un **gel d'agarose ADN** au pourcentage adapté à la taille de fragments à observer. La migration s'effectue dans du tampon **TAE 0,5X**, sous une tension constante de 100V. Les fragments d'ADN sont révélés sous UV grâce au BET (bromure d'éthidium) contenu dans le gel.

4. Purification des produits de PCR

a) Purification sur gel

Après migration sur **gel d'agarose ultra-pure ADN** au pourcentage adapté à la taille des fragments à observer, les fragments d'intérêt sont découpés au scalpel sous UV. L'ADN est ensuite purifié grâce au kit de purification NucleoSpin® Gel and PCR Clean-up (Macherey-Nagel #740609.50) selon les indications du fournisseur.

b) Purification sur colonne

Lorsque l'amplification par PCR ne donne qu'une seule bande d'ADN, il n'est pas nécessaire de purifier sur gel. Le produit de PCR est purifié directement à l'aide du kit de purification NucleoSpin® Gel and PCR Clean-up (Macherey-Nagel #740609.50) selon les indications du fournisseur.

5. Purification de l'ADN au phénol-chloroforme

1v de ROTI®Phénol/Chloroforme/Isopentanol (Carl Roth #A156.1) est ajouté à l'ADN à purifier. Le mélange est vortexé 5 min puis centrifugé à 20000g pendant 5 min. L'ADN présent dans la phase aqueuse est pipeté dans un nouveau tube.

6. Précipitation de l'ADN

L'ADN est précipité en présence de 2,5v d'éthanol 100% et de 0,04v de NaCl 5M, pendant une nuit à 4°C ou 2h à -20°C. Après centrifugation 30 min à 16000g et à 4°C, le culot est lavé 2x avec 1v d'éthanol 75% puis est dissout dans un volume d'eau adéquat.

7. Quantification des acides nucléiques

La quantification de l'ADN se fait avec un spectrophotomètre Nanodrop™ 2000 SpectroPhotometer (ThermoFisher® Scientific), selon le protocole du fournisseur.

8. Clonage avec le système « pGEM T-Easy® vector »

Le système de clonage pGEM T-Easy® vector se fait grâce à une hybridation spécifique A-T : en effet le vecteur contient des T en 3', et les inserts des A, ajoutés par la GoTaq® G2 Flexi DNA Polymerase (Promega #M780A) qui a une activité terminale transférase.

La ligation s'effectue pendant 1h, dans un volume réactionnel de 10 µL comprenant du tampon Rapid Ligation 1X (Promega), 50 ng de pGEM T-Easy vector (Promega), 3 U de T4 DNA ligase (Promega) et 3 µL de produit de PCR purifié.

9. Clonage avec le système Gateway

La technologie Gateway® est une méthode de clonage universelle basée sur la recombinaison site-spécifique du bactériophage lambda. Les amorces utilisées sont listées en [Annexe II-1-b](#).

Tout d'abord la séquence d'ADN d'intérêt est amplifiée par PCR à l'aide d'amorces possédant les sites attB1 et attB2 respectivement en 5' et 3' de la séquence du gène d'intérêt.

Puis une première réaction, la BP, est réalisée dans un volume réactionnel de 5 µL en présence de 100 à 150 ng de produit de PCR purifié, 1 µL de BP Clonase II (ThermoFisher) et 100 ng du pDONR207, pendant 1h. La réaction est stoppée par l'addition de 2 ng de Protéinase K (ThermoFisher) et une incubation de 10 min à 37°C. Cette réaction permet la recombinaison entre les sites *attB* du produit PCR et les sites *attP* du vecteur d'entrée, le pDONR207. Le produit de la réaction est appelé pDONR207-insert, le gène d'intérêt est flanqué de sites *attL*, nécessaires à la deuxième réaction.

La seconde réaction, la LR, est faite dans un volume réactionnel de 10 µL en présence de 100 ng de pDONR207-insert, 2 µL de LR Clonase II (ThermoFisher) et 300 ng de pB7FWG2, pendant 48h. Elle est stoppée de la même manière que la réaction BP. Elle permet la recombinaison entre les sites *attL* du vecteur d'entrée et les sites *attR* d'un vecteur de destination, ici le vecteur pB7FWG2. Le produit de la réaction est appelé B7FWG2-insert.

10. Clonage avec le système GoldenBraid

Cette partie a été réalisée par la Docteure Thalia Salinas-Giegé et l'ingénieur Alexandre Boubegtitène. Les amorces utilisées sont listées en [Annexe II-1-c](#) et les modules en [Annexe V](#).

La technologie GoldenBraid® (Sarrion-Perdigones *et al.*, 2011) est une méthode de clonage basée sur l'utilisation d'enzymes de restriction du type II S qui clivent la séquence d'ADN en dehors de leur site de restriction. Ainsi, quand ces sites sont placés aux extrémités 5' et 3' d'une molécule d'ADN, le clivage élimine le site de restriction et génère des extrémités cohésives sur les séquences adjacentes. Ces séquences sont utilisées pour assembler des « modules » standards dans un ordre précis : chaque module est flanqué de séquences « signatures » qui

génèrent des extrémités cohésives spécifiques lors de la digestion enzymatique, permettant leur assemblage dans l'ordre souhaité lors de la ligation.

D'abord, notre séquence d'intérêt est clonée dans un vecteur d'entrée dépourvu de sites de restriction Bsal, ici le vecteur pUPD2 a été utilisé. Pour ce faire, dans un volume réactionnel de 20 µL sont mélangés 100 ng du vecteur pUPD2, 200 ng de la séquence à insérer (produit de PCR ou séquence de synthèse flanqués des séquences signatures adaptées et du site de restriction de BsmBI), 1 µL d'enzyme BsmBI (ThermoFisher #ER0451), 2 µL de 10X T4 DNA Ligase Buffer et 1 µL de T4 DNA Ligase (5 U/µL) (ThermoFisher #EL0011). Le mélange est ensuite soumis à 26 cycles de digestion-ligation (37°C pendant 3 min pour la digestion, 16°C pendant 4 min pour la ligation) afin de maximiser les rendements, puis 5 min à 50°C suivi de 5 min à 80°C, pour stopper la réaction.

Ensuite, la totalité du milieu réactionnel est utilisée pour transformer des bactéries compétentes. Les colonies obtenues sont mises en culture ([Méthodes II-1](#)) afin de purifier notre construction avec le kit Macherey-Nagel « Nucleospin Plasmid easypure » ([Méthodes III-12](#)). Après vérification par séquençage ([Méthodes III-13](#)), notre construction est utilisée pour la deuxième étape, appelée niveau 1.

Cette étape va permettre l'assemblage dans le vecteur de destination, ici pBactα1, de différents modules avec notre séquence d'intérêt ([Annexe V](#)). Pour cela, dans un volume total de 20 µL, 200 ng des différents modules sont mélangés avec 100 ng de pBactα1, 1 µL d'enzyme Bsal (NEB #R3733), 2 µL de 10X T4 DNA Ligase Buffer et 1 µL de T4 DNA Ligase (5 U/µL) (ThermoFisher #EL0011). La réaction est incubée dans les mêmes conditions que pour l'assemblage de l'insert et du vecteur pUPD2. Des bactéries Rosetta II (DE3) pLysS sont transformées avec le plasmide résultant afin de permettre l'expression des protéines recombinantes ([Méthodes II-2](#)).

11. Digestions enzymatiques

Les digestions enzymatiques sont faites selon les protocoles propre à chaque enzyme. Le produit de digestion est analysé sur gel, et soit purifié sur gel et précipité, soit directement précipité.

12. Purification des plasmides

La purification de plasmides est effectuée avec les kits Macherey-Nagel « Nucleospin Plasmid easypure » et « Nucleobond Xtra Midi », avec 3 mL ou 200 mL de culture bactérienne respectivement, selon le protocole du fournisseur.

13. Séquençage

Au début de ma thèse, les séquences étaient vérifiées par séquençage Sanger par la plateforme AEG (Analyse de l'expression génique) de l'IBMP. Puis, l'IBMP a fait appel à l'entreprise Eurofins Genomics pour le séquençage Sanger.

Les amorces utilisées spécifiquement pour le séquençage sont listées en [Annexe II-2](#).

a) Séquençage à partir de produits de PCR

Pour le séquençage de régions d'ADNg (ADN génomique) ou d'ADNc amplifiées par PCR, 10 pmol d'amorce (même amorce que l'une des deux utilisées pour la réaction de PCR) sont ajoutées à 4 µL de produit de PCR, le tout dans un volume final de 5 µL.

b) Séquençage d'ADN plasmidique

L'ADN plasmidique est séquencé après avoir été exprimé en système bactérien et purifié. Environ 500 à 800 ng d'ADN plasmidique sont ajoutés à 5 pmol d'amorce, le tout dans un volume final de 10 µL.

14. Constructions des micro RNA artificiels pour *C. reinhardtii*

Le protocole d'obtention de mutants par la méthode amiRNA est tiré de (Molnar *et al.*, 2009) ; les amorces utilisées sont listées en [Annexe II-3](#).

a) Choix de la séquence cible

Le choix de la région à cibler par l'amiRNA et des séquences oligonucléotidiques à commander se font à l'aide du site WMD3 (<http://wmd3.weigelworld.org/cgi-bin/webapp.cgi?page=Designer;project=stdwmd>).

Paramètres utilisés : Genome : Creinhardtii_281_v5.6.cds.fa

Min. number of included targets 1

Accepted off-target : 0

Les différentes cibles sont alors testées pour s'assurer qu'il n'y a pas d'autre cible en dehors de celle attendue (*off-targets*).

Paramètres pour "Target search" : Search : target sequence

Genome : Creinhardtii_281_v5.6.cds.fa

Mismatches : 2

Pour avoir un amiRNA fonctionnel dans *C. reinhardtii*, il faut 2 mésappariements entre l'amiRNA et l'ARN cible.

b) Préparation de l'insert

20 µL de **tampon d'hybridation 2X** sont mélangés avec 10 µL de chaque ADNsb (ADN simple brin) à 100 µM (sens et antisens). Ce mélange est chauffé 5 min à 100°C sur une plaque chauffante et est laissé à refroidir doucement sur une nuit. L'ADNdb (ADN double brin) ainsi obtenu sera appelé 90mer.

Le 90mer est purifié à l'aide du kit de purification NucleoSpin® Gel and PCR Clean-up (Macherey-Nagel #740609.50) selon les indications du fournisseur, sauf pour l'élution : 30 µL de tampon NE sont déposés sur la colonne, incubés 1 min et centrifugés 1 min à 11000g. Puis l'éluat est rechargé sur la colonne et l'incubation et la centrifugation sont répétées. L'éluat final est quantifié au NanoDrop ([Méthodes III-7](#)).

La phosphorylation est faite dans un volume final de 10 µL où 700 ng du 90mer sont mélangés avec 1 µL de T4 DNA Ligase 10X Buffer (Promega #C126B) et 1 µL de T4 Polynucleotide Kinase à 10 U/µL (Thermo Fischer #EK0031). Le tout est incubé 30 min à 37°C.

c) Préparation du vecteur pHUN

Le vecteur pHUN est linéarisé par digestion avec l'enzyme BcuI FastDigest (ThermoFischer #FD1253). Dans un volume final de 40 µL, 2 µg sont mélangés à 4 µL de tampon 10X FastDigest Green et 2 µL d'enzyme BcuI FastDigest. Le tout est incubé 1h à 37°C.

Après linéarisation, le vecteur est déphosphorylé. Pour cela, sont ajoutés 2 µL de Phosphatase alcaline thermosensible FastAP à 1 U/µl (Thermo Fischer #EF0651) afin de déphosphoryler le vecteur linéaire. Le mélange est incubé 10 min à 37°C puis sur glace pendant 10 min. Enfin, l'enzyme est inactivée à 65°C pendant 15 min.

Le milieu de déphosphorylation est déposé sur **gel agarose ultra-pure ADN 0,8%**. Après migration, la bande au-dessus de 10 kb (correspondant à la taille du plasmide linéarisé : 12 kb) est découpée et purifiée avec le kit de purification NucleoSpin® Gel and PCR Clean-up (Macherey-Nagel #740609.50) selon les indications du fournisseur. L'élution se fait avec un protocole pour les grands fragments (> 10 kb) : 30 µL de tampon NE à 70°C sont déposés sur la colonne qui est mise à incuber pendant 5 min à 70°C puis à centrifuger d'abord à 50g pendant 1 min et enfin à 11000g pendant 1 min. Cela est répété 3 fois. Les 90 µL récupérés sont précipités ([Méthodes III-6](#)). Après centrifugation 30 min à 16000g et à 4°C, le culot est dissout dans 30 µL d'eau. La solution est quantifiée au NanoDrop ([Méthodes III-7](#)).

d) Clonage du 90mer dans le vecteur pHUN

Dans un volume final de 10 µL, 20 ng de vecteurs sont mélangés avec 1 µL de 90mer dilué soit au 1/10^{ème} soit au 1/100^{ème} dans de l'eau, 5 µL de 2X Rapid Ligation Buffer (Promega #C671A) et 0,5 µL de T4 DNA Ligase (Promega #M180A). La solution de ligation est incubée une heure à température ambiante.

Ensuite, 2,5 µL de réaction de ligation sont mélangés avec 100 µL de bactéries DH10B afin de transformer ces dernières ([Méthodes II-2](#)). Les bactéries sont étalées sur des boîtes LB agar avec 50 µg/mL de carbénicilline et mises à 37°C sur la nuit.

Le lendemain, une dizaine de colonies sont testées par PCR pour savoir si elles ont intégré le plasmide avec le 90mer. Pour cela, un milieu de PCR de 25 µL est préparé selon le protocole décrit précédemment ([Méthodes III-2-c](#)). Si le 90mer est correctement inséré dans le vecteur, une amplification de 188 pb est visible sur le gel. Dans le cas contraire, il n'y aura pas d'amplification. Les colonies ayant la construction correcte sont mises en culture dans 3mL de **milieu LB** ([Méthodes II-1](#)) et les plasmides sont purifiés avec le kit Macherey-Nagel « Nucleospin Plasmid easypure », selon le protocole du fournisseur ([Méthodes III-12](#)). L'insertion est vérifiée par séquençage ([Méthodes III-13](#)).

Après confirmation de la séquence, le vecteur pHUN est linéarisé par digestion avec l'enzyme PvuI FastDigest (Thermo Fischer #FD0624). Dans un volume final de 100 µL, 5 µg de vecteur sont mélangés à 10 µL de tampon 10X FastDigest Green et 2 µL d'enzyme PvuI FastDigest. Le tout est incubé 1h à 37°C. Le vecteur linéarisé est purifié avec un mélange phénol/chloroforme/alcool isoamylique et précipité ([Méthodes III-5 et -6](#)).

IV. Techniques liées à l'ARN

1. Extraction d'ARN total de *C. reinhardtii*

Pour les extractions d'ARN, les colonies de *C. reinhardtii* sont repiquées sur **milieu TAP agar** 3 à 5 jours avant l'extraction.

Entre 1,5 et 2 cm² d'algues sont reprises dans 1,5 mL de TRI-Reagent® en vortexant 10 min, puis centrifugés 10 min à 12000g, à 4°C. Après ajout de 0,2v de chloroforme au mélange, celui-ci est vortexé pendant 15s. Après 5 min, le tout est centrifugé 15 min à 16000g, à 4°C. Le surnageant est alors mélangé à 0,25v de solution **High Salt** et à 0,25v d'isopropanol. L'ensemble est vortexé 5s, incubé 5 min puis centrifugé 10 min, à 16000g, à 4°C. Le culot est lavé avec 500 µL d'éthanol 75% puis centrifugé 5 min, à 16000g, à 4°C. Le culot est repris dans 25 µL d'eau.

La concentration en ARN est estimée au NanoDrop ([Méthodes III-7](#)).

2. Électrophorèse sur gel d'agarose ARN

Un µL d'ARN est mélangé avec 2 µL de **tampon de charge ARN 2X**, dénaturé 5 min à 65°C, mis sur glace durant 5 min puis déposé sur **gel agarose ARN**. La migration est faite dans du **TBE 1X** sous une tension constante de 100 V, pendant 30 min à 1h. Les ARN sont révélés sous UV, grâce au BET contenu dans le tampon de charge.

3. Traitement DNase

Vingt µg d'ARN sont traités dans 100 µL en présence de 10 µL de tampon 10X RQ1 RNase-Free DNase (Promega Cat#M6101), 6U de RQ1 RNase-Free DNase (Promega Cat#M6101), 30 min à 37°C. La réaction est arrêtée par une purification au phénol-chloroforme ([Méthodes IV-4](#)).

4. Purification au phénol-chloroforme

Un volume de ROTI®Aqua-P/C/I (Carl Roth #X985.1) est ajouté aux ARN à purifier. Le mélange est vortexé 5 min puis centrifugé à 20000g pendant 5 min. Les ARN présents dans la phase aqueuse sont transférés dans un nouveau tube.

5. Précipitation

L'ARN est précipité en présence de 2,5v d'éthanol 100% et de 0,1v d'acétate de sodium 3M pH5,3, pendant minimum 1h à -20°C. Après centrifugation 30 min à 16000g et

à 4°C, le culot est lavé avec 1V d'éthanol 75% puis le culot d'acides nucléiques sont dissouts dans 20 µL d'eau.

6. Quantification

Les ARN sont quantifiés au NanoDrop ([Méthodes III-7](#)), comme les ADN.

Pour une quantification plus précise, un fluorimètre Qubit (Thermo Fischer) est utilisé, avec les kits « Qubit RNA HS Assay Kit » (de 4 à 200 ng ; Thermo Fischer #Q32852) ou « Qubit RNA BR Assay Kit » (de 10 à 1200 ng ; Thermo Fischer #Q10210), selon le protocole du fournisseur.

7. Rétrotranscription

La rétrotranscription (RT) des ARN traités DNase en ADNc est faite comme suit : dans un volume de 13 µL, 2,5 µg d'ARN traités à la DNase, 1 µL de Random Hexamers Primers à 500 ng/µL (Promega Cat#C118A), 1 µL d'oligonucléotide dT (voir [Annexe II-4](#)) à 100µM et 1 µL de **dNTP** 10 mM sont incubés 5 min à 65 °C puis refroidis 5 min sur glace. A ce mélange sont rajoutés 1 µL de DTT à 100 mM, 1 µL de RNaseOUT Recombinant Ribonuclease Inhibitor à 40 U/µL (Invitrogen #10777019), 4 µL de tampon SuperScript™ IV Reverse Transcriptase RT 5X (Invitrogen) et 1 µL de SuperScript™ IV Reverse Transcriptase à 200 U/µL (Invitrogen Cat#18090010). Le tout est incubé 10 min à 23°C, 10 min à 50 °C et 10 min à 80 °C.

8. RT-qPCR

Pour les qRT-PCR, 80 µL d'eau sont ajoutés aux 20 µL d'ADNc obtenus par la RT, afin d'avoir un mélange dilué au 5^{ème}, ce qui correspond à 25 ng/µL d'ADNc final.

La RT-qPCR est effectuée à partir d'un mélange réactionnel de 10 µL avec 5 µL de dsDNA-specific dye Takyon™ SYBR® 2X qPCR Mastermix Blue (Eurogentec Cat#UF-FSMT-B0701), 1 µL d'ADNc dilué au 5^{ème} et 2 µL d'un mélange des amorces direct et complémentaire à 25 µM. Les amorces utilisées sont listées en [Annexe II-5](#). Pour chaque plaque, chaque échantillon est déposé 3 fois, afin d'avoir 3 réplicats techniques.

La réaction de PCR ainsi que le suivi de la fluorescence à 530 nm au cours des cycles d'amplification sont réalisés dans le LightCycler 480® (Roche). Après une dénaturation de 7 min à 95°C, les échantillons sont soumis à 45 cycles d'amplification de 15s à 95°C, 15s à 60°C et 15s à 72°C. A la fin de l'amplification, une courbe de dénaturation des produits est réalisée.

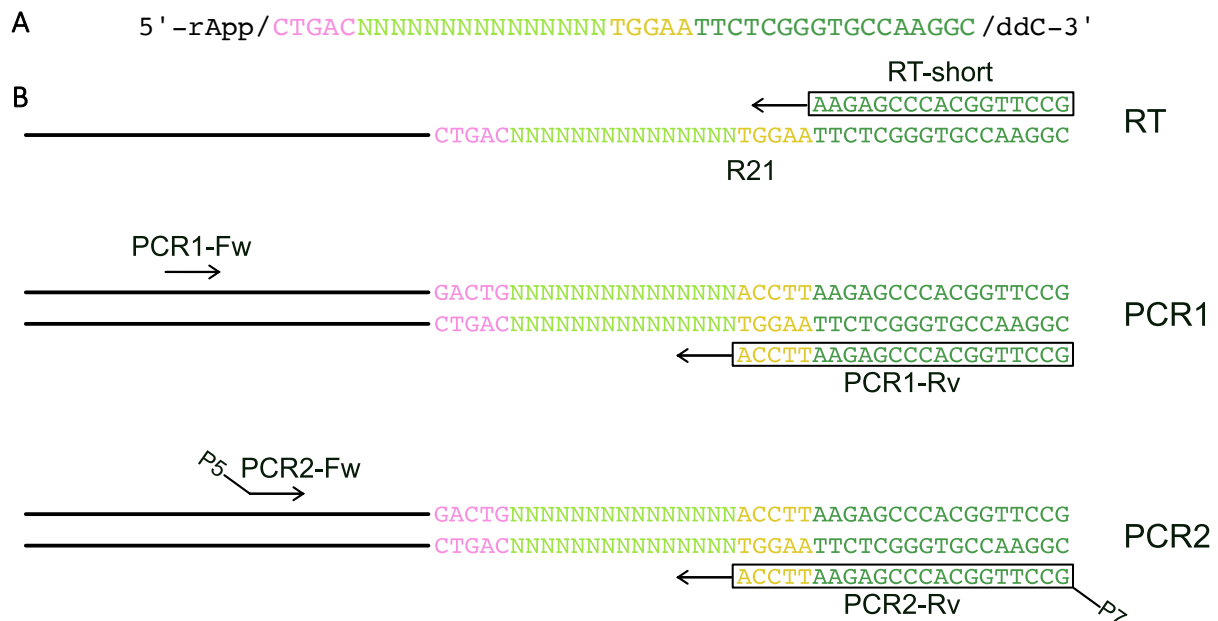


Figure M-1 : Schémas récapitulatifs des différentes amorces utilisés pour le 3'RACE-seq

A : Schéma de l'adaptateur R21. Synthétisé par IDT, il contient une extrémité 5' pré-adénylée (rApp : 5',5'-adenyl-pyrophosphoryl). Cela permettra la liaison avec les groupes 3'OH (3' hydroxyl) à l'extrémité 3' des ARN sans ajout d'ATP dans le mix de ligation. Cela empêche l'adénylation des ARN et donc la ligation de produits indésirables (ARNr, ARNt). Son extrémité 3' avec un nucléotide ddC (3' dideoxy-cytidine) évite une auto-ligation avec le 5' de l'adaptateur. **B** : Schéma des amorces pour les étapes de RT et de PCR. L'oligonucléotide utilisé pour la RT est plus court de 5 nucléotides par rapport à celui utilisé pour la PCR1, afin de favoriser l'amplification PCR d'ADNc provenant d'ARN ligués avec l'adaptateur, et de minimiser les amplifications dues à une liaison non spécifique de l'oligo. De même, l'amorce PCR2 est prise en aval de celle PCR1 afin de favoriser les amplifications spécifiques du gène cible. Les amorces de PCR2 contiennent les séquences P5 et P7 (Illumina), nécessaires pour le séquençage Illumina.

Jaune : 5 nucléotides de la séquence de l'adaptateur manquants dans le RT-short mais présents dans l'oligonucléotide antisens PCR1 ; **Rose** : séquence servant de délimitant lors de l'analyse bio-informatique ; **Vert clair** : séquence aléatoire, permettant d'enlever les duplicats de PCR lors de l'analyse bio-informatique ; **N** = A, G ou T ; **Vert foncé** : séquence du RT-short et du complémentaire au RT-short ; Fw : *forward*, amorce sens ; Rv : *reverse*, amorce antisens.

Le logiciel LightCycler480® Software version 1.5 (Roche) permet ensuite, à partir des courbes d'amplification, de déterminer les C_t pour chaque échantillon. Puis, les efficacités réelles des couples d'amorces sont calculées par le logiciel LinRegPCR. Enfin, la méthode des delta-delta C_t a été utilisée pour calculer l'abondance relative des ARN en normalisant avec la moyenne géométrique de 2 ARN références : MAA7 and CYN19.

9. 3'RACE-seq

Le protocole pour les expériences de 3'RACE-seq (*3' Rapid Amplification of cDNA Ends coupled with sequencing*) est tiré de (Scheer *et al.*, 2020) ; La liste des amorces utilisées est disponible en [Annexe II-6](#).

a) Ligation de de l'adaptateur en 3' des ARN

L'extraction et l'analyse des ARN se fait de la même manière que décrit précédemment ([Méthodes IV-1](#)). Les ARN sont quantifiés au NanoDrop ([Méthodes III-7](#)).

Dans un volume total de 44 μ L, 5 μ g d'ARN sont mélangés avec 1 μ L de l'adaptateur R21 ([Figure M-1A](#)) à 10 μ M. Après 3 min à 65°C et 2 min sur glace, 5 μ L de T4 RNA Ligase Reaction Buffer (NEB #B0216SVIAL) et 1 μ L de T4 RNA Ligase 1 (ssRNA Ligase) (NEB #M0204SVIAL) sont ajoutés au mélange initial. Le tout est incubé à 37°C pendant 1h.

Les ARN ligués sont purifiés à l'aide du kit Nucleospin RNA Clean-up (Macherey-Nagel #740948.50), selon le protocole du fournisseur. Puis ils sont précipités en présence de 0,1v d'acétate de sodium 3M pH5,3, 1 μ L de Glycogène (Thermo Fisher #R0561), et 2,5v d'éthanol 100%, pendant minimum 1h à -20°C. Après centrifugation 30 min à 16000g et à 4°C, le culot est lavé 2x avec 1V d'éthanol 75% puis est re-suspendu dans 11 μ L. Les ARN ligués sont quantifiés au NanoDrop ([Méthodes III-7](#)).

b) RT

L'oligonucléotide utilisé pour la RT a été choisi pour minimiser les produits de PCR non désirés dû à une ligation non spécifique de l'oligonucléotide. Pour cela, l'oligonucléotide pour la RT est plus petit de 5 nucléotides par rapport à l'adaptateur et à l'oligonucléotide qui sera utilisé pour la PCR, cela permet l'amplification uniquement d'ADNc provenant d'ARN ligués avec l'adaptateur. Cette différence est visible en [Figure M-1B](#).

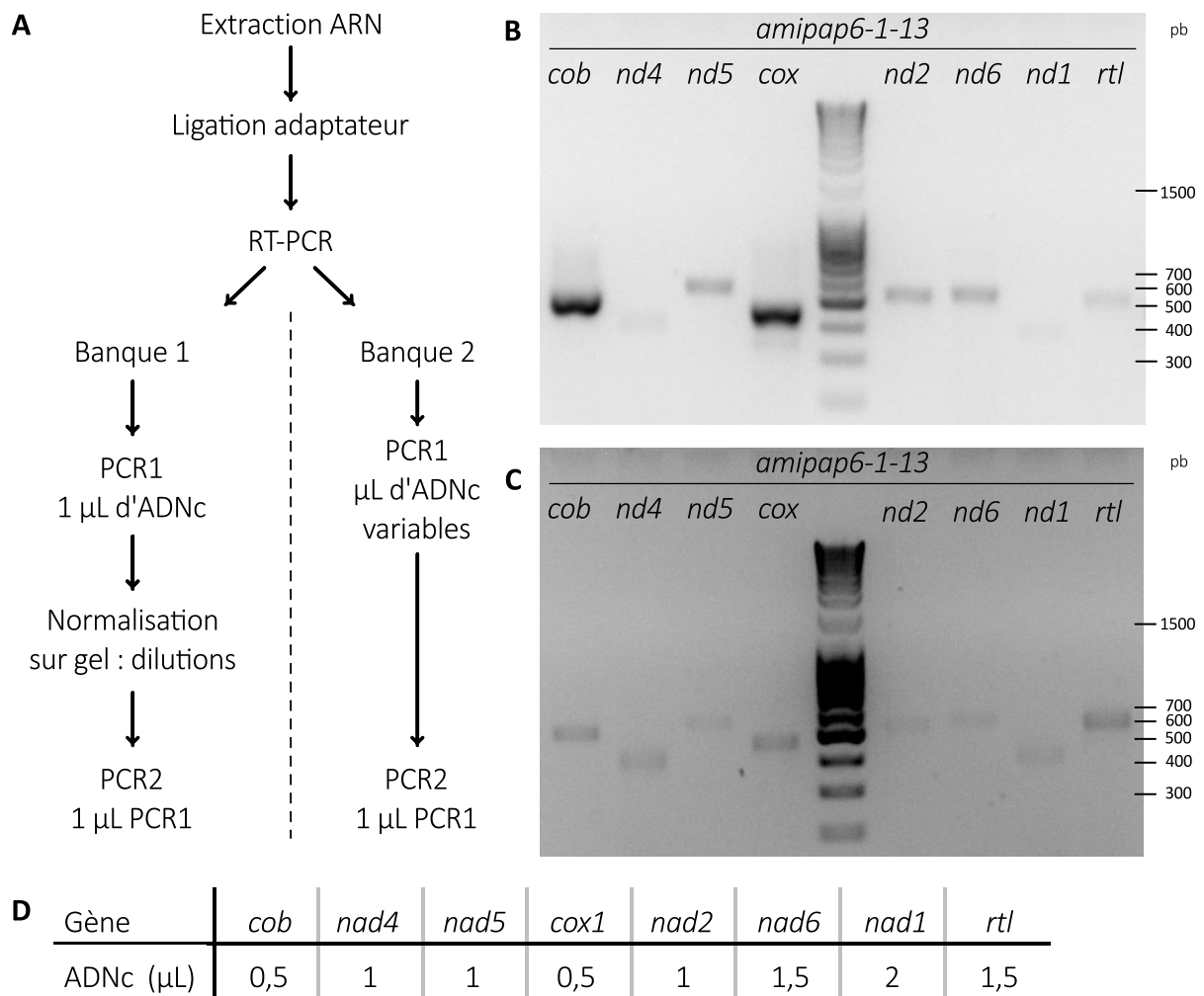


Figure M-2 : Détails du protocole et des résultats de la PCR1

A : Schéma du protocole effectué pour les 2 banques, avec les différences entre les Banques 1 et 2 pour la PCR1; **B** et **C** : Gels d'électrophorèse des produits de la PCR1 pour le mutant *amipap6-1-13* dans la Banque 1, avant (**B**) et après (**C**) la normalisation sur gel des quantités d'amplicons par dilutions ; **D** : Tableau des volumes (en µL) d'ADNc utilisés pour la PCR1 de la Banque 2, en fonction du gène amplifié. Pb : nombre de paires de bases des amplicons PCR.

La RT des ARN ligués est faite comme suit dans un volume de 13 μL : 2 μg d'ARN ligués avec l'adaptateur R21, 0,5 μL de l'oligonucléotide RT-short-primer à 100 μM et 1 μL de **dNTP** 10 mM sont incubés 5 min à 65°C puis refroidis 5 min sur glace. A ce mélange sont rajoutés 1 μL de DTT à 100 mM, 1 μL de RNaseOUT Recombinant Ribonuclease Inhibitor à 40 U/ μL (Invitrogen #10777019), 4 μL de tampon SuperScript™ IV Reverse Transcriptase RT 5X (Invitrogen) et 1 μL de SuperScript™ IV Reverse Transcriptase à 200 U/ μL (Invitrogen Cat#18090010). Le tout est incubé 15 min à 50°C et 15 min à 70°C.

c) Amplification de la région 3' des gènes cibles : PCR1

Les gènes étudiés sont les 8 gènes d'ARNm codés par le génome mitochondrial de *C. reinhardtii* (*cob*, *nad4*, *nad5*, *cox1*, *nad2*, *nad6*, *nad1* et *rtl*). Pour les amplifier, nous réalisons 2 PCR successives, nommées ici PCR1 et PCR2. Le but de la PCR1 est d'amplifier les gènes d'intérêt avec une amorce complémentaire à l'adaptateur et une autre spécifique du gène (Figure M-1B).

Nous avons fait 2 banques d'ADNc séquencées par Illumina, que nous appellerons Banque 1 (non présentée dans cette thèse) et 2 (dont tous les résultats présentés sont issus). Le protocole de la PCR1 diffère entre les 2 banques (Figure M-2A).

Banque 1

La PCR1 est réalisée dans un milieu réactionnel de 20 μL qui comprend 4 μL de tampon 5X Green GoTaq® Flexi Buffer (Promega #M891A), 0,4 μL de **dNTP** 10mM, 1 μL de chaque amorce pour PCR1 à 10 μM , 1,2 μL de MgCl_2 25mM (Promega #A351B), 0,2 μL de GoTaq® G2 Flexi DNA Polymerase (Promega #M780A) et 1 μL d'ADNc. L'amplification est réalisée dans un thermocycleur Biometra® T3000. Elle débute par une étape de dénaturation initiale de 30s à 94°C, suivie de 29 itérations du cycle suivant :

- 20s à 94°C,
- 20s à 56°C,
- 30s à 72°C.

Une étape d'élongation de 30s à 72°C termine la réaction.

Les produits de PCR sont analysés sur gel (Méthodes III-3).

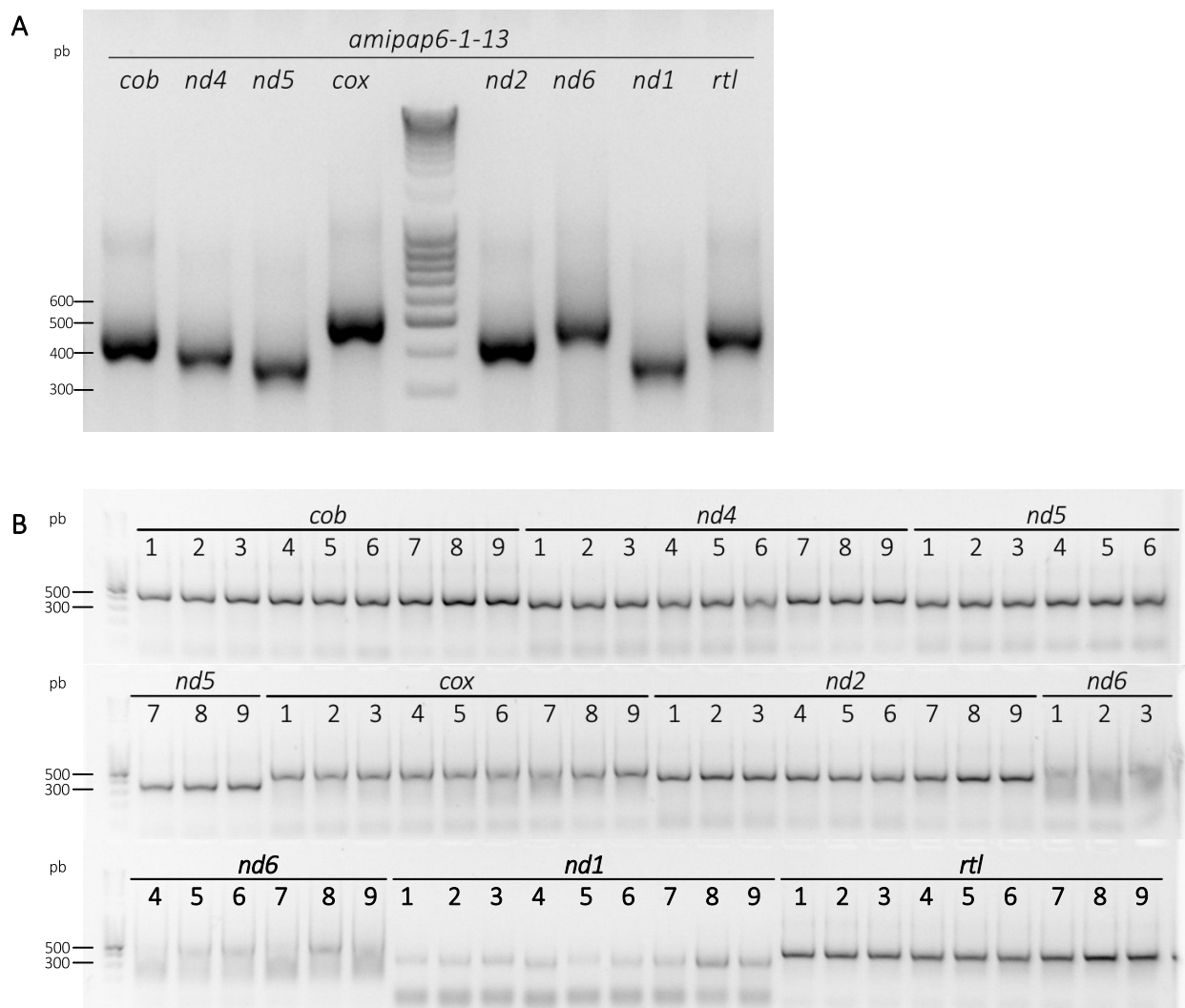


Figure M-3 : Gels d'électrophorèse obtenus après les PCR2

Profils des amplicons après la migration des PCR2 obtenus pour les 8 gènes d'intérêt : **A** : pour la Banque 1, sur la souche *amipap6-1-13* ; **B** : pour la Banque 2, sur 9 souches différentes (de 1 à 9 : *pHUN*, *amipap4-37*, *amipap4-2*, *amipap6-2-54*, *amippap6-2-64*, *amipap6-2-75*, *CC-5325*, *pap4-1::CIB* et *pap4-2::CIB*). Pb : nombre de paires de bases des amplicons PCR.

L'analyse sur gel des produits de la PCR1 montre des intensités variables selon le gène amplifié (Figure M-2B). En se basant sur la bande la plus faible sur gel, une dilution des autres produits de PCR est faite, pour obtenir un profil d'intensités homogènes pour les 8 gènes d'une même souche, comme présenté en Figure M-2C. Cette homogénéité est vérifiée sur **gel d'agarose ADN 2%**, comme expliqué précédemment. Si les intensités des bandes sont correctes, cette dilution est utilisée comme matrice pour la PCR2.

Banque 2 :

La PCR1 est réalisée comme pour la Banque 1, mais les volumes d'ADNc sont différents selon le gène amplifié (Figure M-2D). L'amplification est réalisée en plaques 96 puits, dans un thermocycleur Bio-Rad T100™ Thermal Cycler, selon le même programme que pour la Banque 1. Les produits de la PCR1 ne sont pas analysés sur gel dans ce cas-là.

d) Amplification avec les séquences Illumina : PCR2

Pour la PCR2, les amorces contiennent les séquences P5 (amorces sens) et P7 (amorces complémentaires), nécessaires pour le séquençage Illumina (Figure M-1B). De plus, lors de la PCR2, l'oligonucléotide spécifique au gène d'intérêt est choisi en aval de celui de la PCR1, afin d'éviter les amplifications non spécifiques (Figure M-1B).

La PCR2 est réalisée dans un milieu réactionnel de 20 µL comprenant 4 µL de tampon 5X Green GoTaq® Flexi Buffer (Promega #M891A), 0,4 µL de **dNTP** 10mM, 1 µL de chaque amorce pour PCR2 à 10µM, 1,2 µL de MgCl₂ 25mM (Promega #A351B), 0,2 µL de GoTaq® G2 Flexi DNA Polymerase (Promega #M780A) et 1 µL de PCR1. L'amplification est réalisée en tubes individuels dans un thermocycleur Biometra® T3000 pour la Banque 1 et en plaques 96 puits dans un thermocycleur Bio-Rad T100™ Thermal Cycler pour la Banque 2. Elle débute par une étape de dénaturation initiale de 1 min à 94°C, suivie de 14 itérations du cycle suivant :

- 20s à 94°C,
- 20s à 56°C,
- 30s à 72°C.

Une étape d'élongation de 30s à 72°C termine la réaction.

Les produits de PCR sont analysés sur gel (Méthodes III-3) (Figure M-3).

e) Purification des produits de PCR

Là encore, le protocole diffère selon les banques. Pour la Banque 1, tout a été fait en tubes individuels, alors que pour la banque 2, des plaques 96 puits ont été utilisées.

Banque 1 : en tubes individuels

0,8V de billes magnétiques (produites par la plateforme AEG) sont ajoutées au produit de PCR2. Le mélange est homogénéisé par 10 allers/retours de pipetage puis incubé 5 min. Le tube est centrifugé 10s à 3200g puis placé sur un support magnétique. Après 5 min d'incubation, le surnageant est jeté et 2 lavages successifs d'une minute avec 300 μ L d'éthanol 70% sont réalisés. L'éthanol est enlevé au maximum. 2 éluions successives se font comme suit : 100 μ L d'eau sont ajoutés sur les billes, le mélange est homogénéisé par 10 allers/retours de pipetage, incubé 5 min, puis le tube est placé sur un support magnétique et après 5 minutes d'incubation, l'éluat est prélevé et transféré dans un nouveau tube. Le volume final de 200 μ L d'éluat est précipité en présence de 0,1v d'acétate de sodium 3M pH5,3, 1 μ L de Glycogène (Thermo Fisher #R0561) et 2,5v d'éthanol 100%, pendant une nuit à 4°C. Après centrifugation 30 min à 16000g et à 4°C, le culot est lavé 2x avec 1v d'éthanol 75% puis est dissout dans 11 μ L.

Banque 2 : en plaque 96 puits :

1,25V de billes magnétiques (produites par la plateforme AEG) sont ajoutées au produit de PCR2. Le mélange est homogénéisé par 10 allers/retours de pipetage puis incubé 10 min. La plaque est centrifugée 10s à 3200g puis placée sur un support magnétique. Après 5 min d'incubation, la plaque est centrifugée 10s à 3200g puis remise délicatement sur le support en glissant doucement les puits le long d'un bord de sorte d'avoir un amas de billes d'un côté des puits. Le surnageant est jeté et 2 lavages à l'éthanol 80% sont faits successivement, avec d'abord 120 μ L puis 200 μ L. L'éthanol est enlevé au maximum. L'éluion se fait avec 15 μ L d'eau. Le mélange est homogénéisé par 10 allers/retours de pipetage puis incubé 10 min. La plaque est centrifugée 10s à 3200g puis placée délicatement sur un support magnétique en glissant doucement les puits le long d'un bord de sorte d'avoir un amas de billes d'un côté des puits. 13 μ L sont prélevés et transférés dans une nouvelle plaque, évitant ainsi de prendre des billes.

f) Assemblage des « pool/gènes » et analyses

Les produits de PCR2 purifiés sont quantifiés par un fluorimètre Qubit (Thermo Fischer), avec le kit « Qubit dsDNA HS Assay Kit » (de 0,1 à 120 ng ; Thermo Fischer #Q32851), selon le protocole du fournisseur.

Cette quantification permet de rassembler 10 à 20 ng de chaque PCR2 purifiée d'un même gène pour les différentes souches. Ces mélanges, appelés « pool/gène », sont quantifiés au Qubit avec le kit « Qubit dsDNA HS Assay Kit », comme décrit précédemment.

g) Spike-in

Les spike-in servent de contrôle lors du séquençage MiSeq, afin de vérifier l'efficacité de séquençage lors de séquences avec des répétitions du même résidu. Le spike-in N sert à vérifier l'efficacité du séquençage en général. Quatre types de spike-in ont été utilisés : les spike-in C avec 8C, 16C ou 22C, contenant 8, 16 ou 22 résidus C, respectivement, et le spike-in N, contenant 23 résidus N (= A, T, C ou G), c'est-à-dire une séquence aléatoire de 23 nucléotides ([Annexe II-6](#)).

La PCR Spike-in est réalisée dans un milieu réactionnel de 20 µL qui comprend 4 µL de 5X Green GoTaq® Flexi Buffer (Promega #M891A), 0,4 µL de **dNTP** 10mM, 1 µL de chaque amorce pour PCR1 à 10µM, 1,2 µL de MgCl₂ 25mM (Promega #A351B), 0,2 µL de GoTaq® G2 Flexi DNA Polymerase (Promega #M780A) et 2 µL de spike-in à 100 nM. L'amplification est réalisée dans un thermocycleur Biometra® T3000. Elle débute par une dénaturation initiale de 30s à 94°C, suivie de 10 itérations du cycle suivant :

- 20s à 94°C,
- 20s à 50°C,
- 30s à 72°C.

Une étape d'élongation de 30s à 72°C et une incubation de 5 min à 20°C terminent la réaction.

Les produits de PCR sont déposés sur un **gel acrylamide 6%**. La migration s'effectue dans du **TBE 1X**, à 150V durant 1h15. Puis le gel est incubé dans un bain de BET (10 µL de BET pour 50 mL d'eau) pendant 10 min, lavé une fois à l'eau et révélé sous UV.

Les bandes correspondant aux différents spike-in sont découpées sur gel et mises dans des tubes 0,5 mL préalablement perforés (6 fois) au fond avec une aiguille (0,45 x 12 mm).

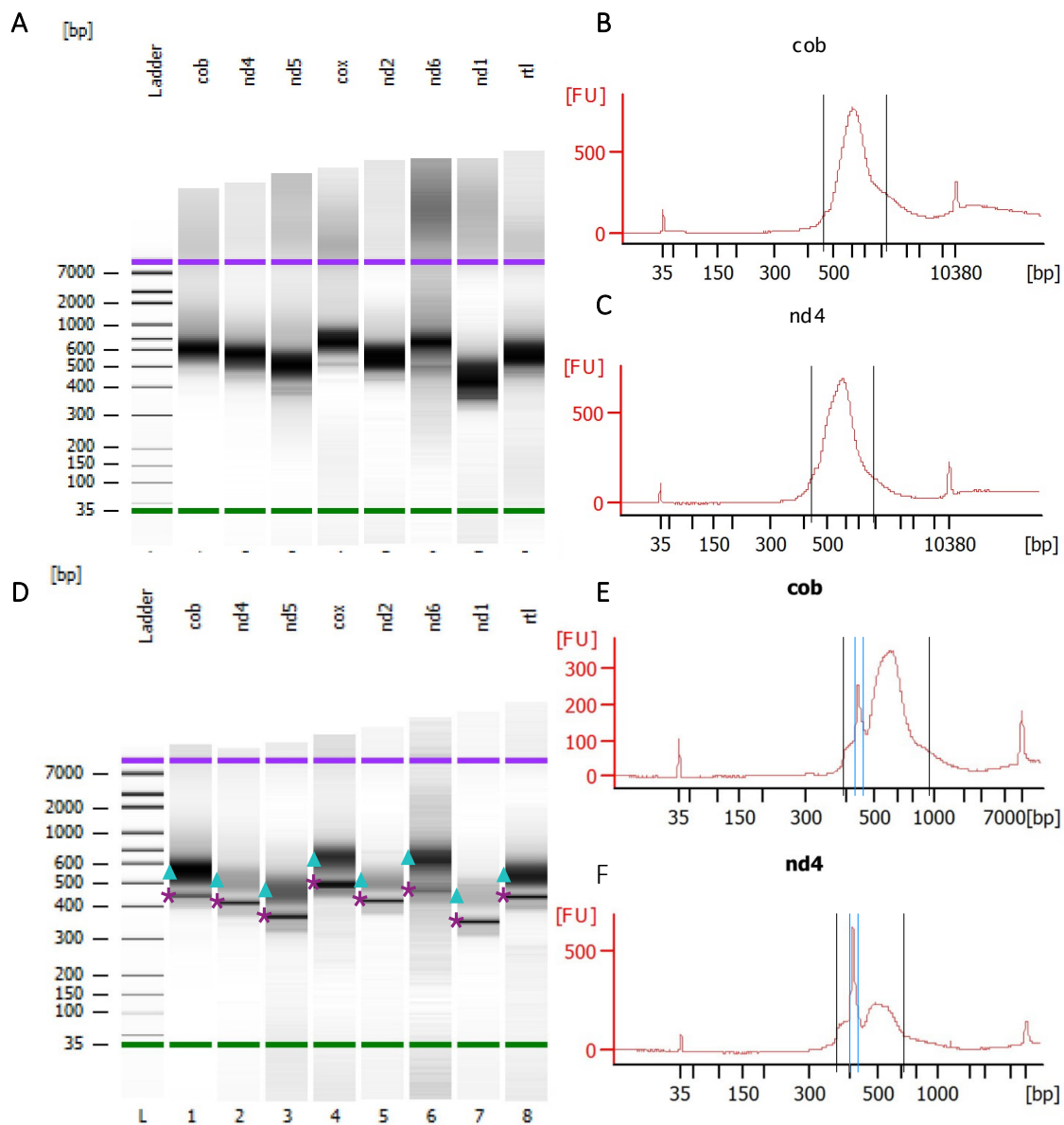


Figure M-4 : Résultats d'analyses des pool/gènes au BioAnalyzer

Exemples de résultats obtenus par le BioAnalyzer pour les Banques 1 (A, B et C) et 2 (D, E et F). A et D : électrophorégrammes obtenus par le BioAnalyzer pour les 8 pool/gènes ; D : les triangles bleus montrent les bandes étalées et les étoiles violettes les bandes fines ; B, C, E et F : exemples de profils obtenus pour *cob* (B et E) et *nd4* (C et F), pour les Banques 1 (B et C) et 2 (E et F) ; les traits noirs encadrent la région prise en compte pour calculer la taille moyenne du profil ; les traits bleus encadrent la région prise en compte pour calculer le pourcentage qu'occupe la bande (= le pic) par rapport à la totalité de l'échantillon. Bp : base pair, nombre de paires de bases ; FU : Fluorescence Units, unités de fluorescence mesurées par le BioAnalyzer.

Ces tubes sont mis dans des tubes 2 mL et l'ensemble est centrifugé 30s à 11000g, afin de fragmenter les morceaux de gel en passant à travers les perforations. 800 µL de NaCl 0,3M sont ajoutés aux débris de gel et le tout est incubé une nuit à 37°C sous agitation (800 rpm avec l'incubateur Thermomixer comfort Eppendorf). Après une centrifugation de 30s à 11000g, le surnageant est réparti dans 2 tubes de 1,5 ml puis précipités par l'ajout de 0,1V d'acétate de sodium 3M pH5,3, de 1 µL de Coprécipitant GlycoBlue™ (15 mg/mL) (Thermo Fisher #AM9515) et de 2,5V d'éthanol 100%, pendant au moins 2h à -20°C. Après centrifugation 30 min à 16000g et à 4°C, le culot est lavé avec 1V d'éthanol 75% puis est dissout dans 11 µL. Enfin, les tubes des mêmes spike-in sont rassemblés en un seul, et reprécipités dans un plus petit volume si besoin. Les Spike-in ainsi purifiés sont quantifiés par un fluorimètre Qubit (Thermo Fischer), avec le kit « Qubit dsDNA HS Assay Kit » (de 0,1 à 120 ng ; Thermo Fischer #Q32851), selon le protocole du fournisseur.

h) Super-pools et envoi au séquençage Illumina

Les pool/gène ainsi que les spike-in sont analysés au BioAnalyzer (Agilent). Pour cela, 5 ng de chaque échantillon sont déposés sur une puce High Sensitivity DNA Analysis (Agilent, #5067-4626) selon le protocole du fournisseur.

Banque 1 :

Chaque pool/gène est dilué pour obtenir une concentration de 4 nM pour 10 µL. Les concentrations molaires sont calculées grâce à la concentration Qubit et avec la taille moyenne du profil observé au BioAnalyzer. Puis, un « Super Pool Gènes » est réalisé par le groupement des 8 pool/gènes en mélangeant 1 µL de chaque dilution à 4 nM. Un exemple est donné en [Figures M-4A, M-4B et M-4C](#).

Il en est de même pour les Spike-in pour obtenir un « Super Pool Spike ».

Banque 2 :

Nous avons constaté une différence des profils des pool-gène par rapport à la Banque 1. Nous avons observé la présence d'une bande fine ([Figure M-4D](#), étoiles violettes) beaucoup plus marquée que dans la Banque 1, en-dessous d'une bande étalée ([Figure M-4D](#), triangles bleus) pour chaque gène. Ces bandes étant à la taille théorique attendue des amplifications PCR pour chaque gène, nous avons voulu normaliser les échantillons en fonction des bandes et non du profil total. Pour cela, nous avons calculé le pourcentage qu'occupe la bande par rapport au

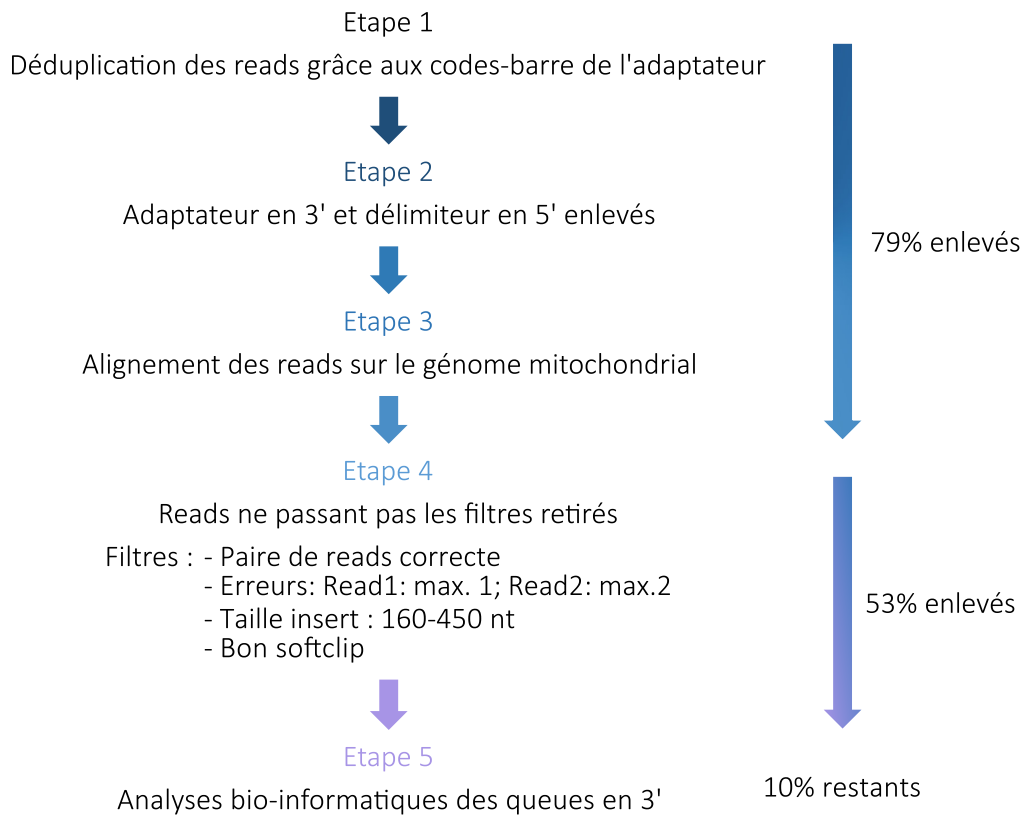


Figure M-5 : Traitement bio-informatique des données de 3'RACE-seq

Schéma du pipeline informatique permettant le traitement des données des 3'RACE-seq, afin de faire les analyses bio-informatiques. Les flèches sur le côté droit montrent le pourcentage du nombre de séquences enlevées suivant les différentes phases du traitement des données.

profil total, avec les données du BioAnalyzer. Nous avons ensuite calculé et appliqué un facteur afin de mettre le même pourcentage de bande dans le « Super Pool Gènes » pour chaque échantillon. Un exemple est donné en [Figures M-4D, M-4E et M-4F](#).

Il en est de même pour les Spike-in et le Super Pool Spike.

L'ensemble des Super-pools sont par la suite confiés à la plateforme Analyse de l'expression génique de l'IBMP, pour dépôt sur une cartouche MiSeq.

i) Séquençage Illumina

La plateforme Analyse de l'expression génique de l'IBMP fait un mélange ADN-Spike avec 96% d'ADN et 4% de Spike. Puis, une cartouche MiSeq reagent kit v3 (150 cycles) (Illumina #MS-102-3001) est remplie avec 30% de PhiX Control v3 (Illumina #FC-110-3001) et 70% du mélange ADN-Spike, et séquencée sur un séquenceur MiSeq (Illumina).

j) Analyses bio-informatiques

Les analyses bio-informatiques ont été faites par l'ingénieur David Pflieger et la Docteure Thalia Salinas-Giegé, comme expliqués en [Figures M-5 et M-6](#). Les résultats des 3 réplicats ont été cumulés pour faire les analyses.

V. Techniques liées aux protéines

1. Quantification au Bradford

Un μL d'échantillon est mélangé à 9 μL d'eau. Les 10 μL sont ensuite additionnés à 1 mL de réactif de Bradford 1X (Bio-Rad Protein Assay Dye Reagent) et, après 5 min d'incubation, l'absorbance est mesurée à 595 nm avec un spectrophotomètre UV-1800 (Shimadzu). Une gamme étalon, préalablement établie avec de la BSA (*bovin serum albumin*), permet de déterminer la concentration en protéines de l'échantillon.

2. Gel de protéines

Les protéines sont séparées par électrophorèse en conditions dénaturantes sur un gel gradient de polyacrylamide 8-16%, comprenant un **gel de concentration** et un **gel de séparation**. La polymérisation du gel s'effectue par addition d'APS (persulfate d'ammonium) 0,1% (p/v) et de TEMED (tétraméthyléthylènediamine) 0,01% (v/v). Pour couler le gel gradient, 2 mélanges pour gels de séparation sont faits, un à 8% et l'autre à 16%. Avec une pipette graduée, 5 mL du 8%

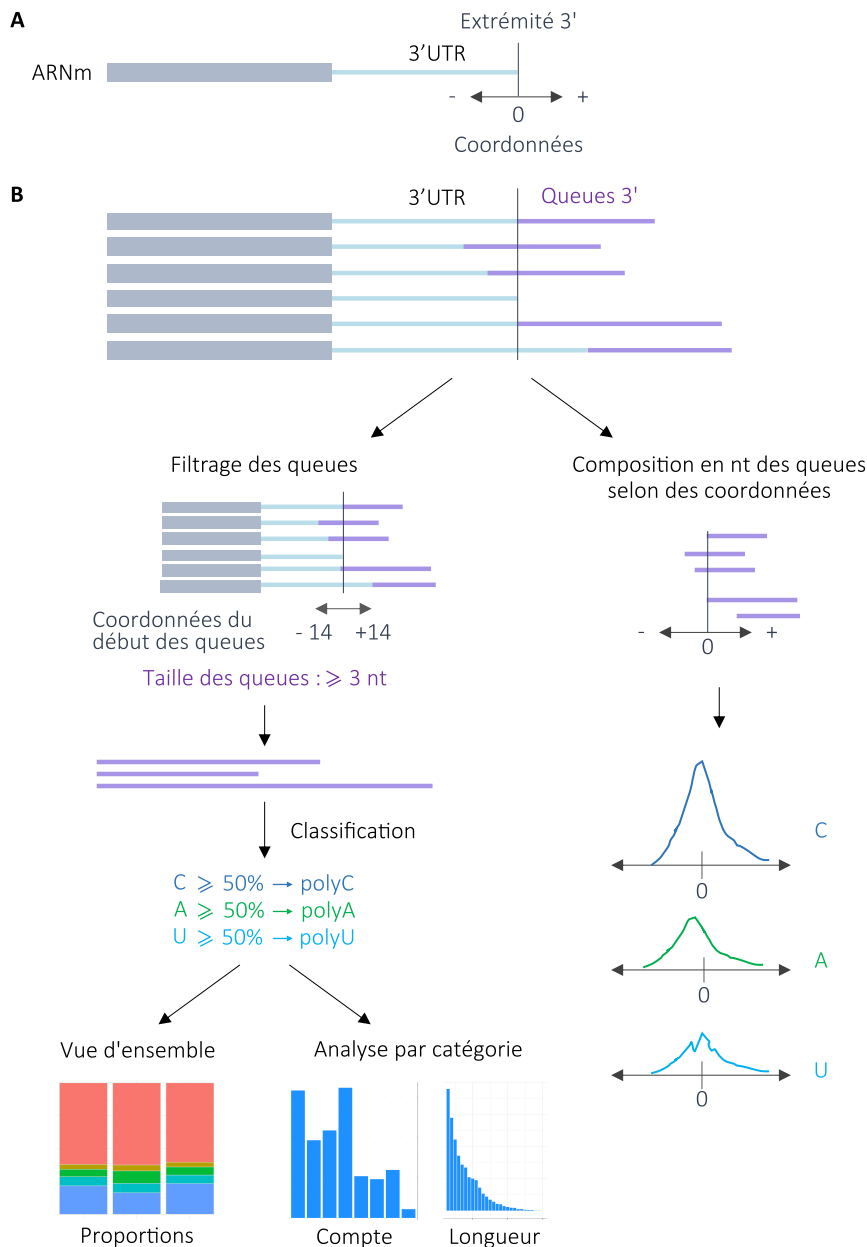


Figure M-6 : Vue d'ensemble des analyses de 3'RACE-seq

Schéma de l'analyse des données de 3'RACE-seq, après le traitement informatique jusqu'à l'obtention des différents graphiques. **A** : Placement de la coordonnée 0, correspondant à l'extrémité 3' de la 3'UTR, selon (Salinas-Giegé *et al.*, 2017). **B** : Illustration des ARNm avec leur 3'UTR et leur queue. La ligne verticale noire représente la coordonnée 0. Deux types d'analyses ont été faites. La première (gauche) se base sur un alignement de tous les ARNm, sans queue ou avec des queues qui commencent entre -14 et +14, en fonction de la coordonnée 0 et ayant une taille supérieure ou égale à 3 nt. Les queues de ces ARNm ont été classées selon leur composition, pour obtenir 3 types de graphiques : les proportions des types de queues, le nombre de chaque type de queue en fonction des gènes et la longueur des queues. La deuxième utilise toutes les queues, sans aucun critère, afin de les aligner. Puis le nombre de chaque nt ajouté à chaque position en amont et en aval de la fin de la 3'UTR (= position 0) est compté, afin d'obtenir des graphiques de la distribution de chaque nt en fonction de la position.

sont prélevés, puis 5 mL du 16%. De l'air est aspiré avec la pipette de sorte qu'une seule bulle remonte dans la pipette, créant ainsi un gradient entre le 8 et le 16%.

Un volume de **Laemmli 5X** est ajouté aux extraits protéiques qui sont ensuite chauffés 5 min à 95°C avant d'être déposé sur gel. La migration se fait à un ampérage constant de 30 mA dans du **tampon TGS**.

3. Western blot

a) Transfert des protéines sur membrane

Les protéines séparées sur gel de polyacrylamide dénaturant sont transférées, par électrophorèse, sur une membrane Immobilon-P 0,45 µm (Millipore) préalablement activée dans de l'éthanol 100%. Le gel est positionné sur la membrane et le tout est entouré de deux couches de papier Whatman et d'une éponge de part et d'autre. Le transfert s'effectue dans du **tampon TGS** durant 1h30 à 110 V.

b) Coloration des protéines sur la membrane

Les protéines transférées sont visualisées par incubation sous agitation de la membrane dans une **solution de coloration WB**. Lorsque la membrane devient bleue, plusieurs bains successifs de **solution de décoloration WB** sont réalisés pour faire apparaître les protéines. Après avoir scanné la membrane, celle-ci est entièrement décolorée dans un bain de méthanol 100%.

c) Immunodétection

Pour saturer les sites de fixation aspécifiques, la membrane est incubée au moins 1h à température ambiante (ou une nuit à 4°C) sous agitation dans du **tampon bloquant WB**. Elle est ensuite incubée 1h à température ambiante ou durant la nuit à 4°C avec l'anticorps primaire spécifique de la protéine à révéler dilué dans du **tampon bloquant WB** (dilutions et conditions données en [Annexe III](#)). Après 3 lavages de 10 min dans du **tampon de lavage WB**, la membrane est incubée 30 min dans du **tampon bloquant WB** avec l'anticorps secondaire dilué au 1/10000^{ème}. La membrane est lavée 3 fois 10 min dans du **tampon de lavage WB**. Les protéines d'intérêt sont finalement révélées à l'aide des kits ECL Clarity ou Clarity Max (Bio-Rad) et visualisées avec l'appareil Fusion FX (Vilber).

4. Production de protéines recombinantes

Cette partie a été réalisée par la Docteure Thalia Salinas-Giegé, l'ingénieur Alexandre Boubegtitene et l'ingénieur Nicolas Baumberger de la plateforme P3P de l'IBMP.

L'ensemble des constructions a été obtenu par clonage Golden Braid, comme décrit précédemment. Elles contiennent toutes une étiquette His-SUMO en N-terminal. Le tag His permet la purification par affinité de la protéine. Le tag SUMO sert à améliorer la solubilité de la protéine à exprimer.

La veille de l'induction, 50 mL de culture sont inoculés avec les bactéries contenant la construction d'intérêt et incubés à 37°C sur la nuit. L'absorbance de la culture est mesurée à une DO_{600nm} et une culture est préparée dans 400 mL de **milieu auto-inductible** pour obtenir une DO_{600nm} finale de 0,1. La culture est incubée pendant 4h à 37°C puis 15h à 25°C, sous agitation. La DO_{600nm} est mesurée et les bactéries sont centrifugées 15 min à 5000g à 4°C et re-suspendues dans du **tampon de lyse bactéries** pour avoir une DO_{600nm} finale de 20. 50 mL de bactéries re-suspendues sont lysées à l'aide du Microfluidizer, deux fois à 1200 bar. La solution lysée est traitée à la Turbonuclease (8 U/mL ; Sigma #T4330), par incubation pendant 20 min. Le tout est ensuite clarifié par centrifugation à 20000g pendant 20 min. Une première étape de purification des protéines par affinité est faite avec le système AKTA pure (Cytiva) et des colonnes HisTrap FF Crude 1ml (Cytiva #29048631) grâce à leur tag His. Quatre tampons sont utilisés : un **tampon de binding PP** (purification protéines) pour lier les protéines à la colonne, un **tampon de lavage PP** et deux **tampons d'élution PP** (1 puis 2) pour détacher les protéines de la colonne avec des concentrations croissantes en imidazole. Grâce au chromatogramme, les fractions présentant un pic à 280 nm sont analysées sur un gel de protéines ([Méthodes V-3](#)). Les fractions contenant la quantité la plus élevée et la plus pure de la protéine d'intérêt sont rassemblées. Le tag SUMO est clivé par une incubation des protéines en présence de SUMO protease (13000 U/mL) (His tagged-Senp2 protéine recombinante humaine, produite par P3P, IBMP) sur la nuit à 4°C. Les échantillons sont concentrés dans 300 μ L par ultrafiltration dans des tubes Amicon Ultra (30 kDa, 15 mL ; Millipore #UFC903008) en centrifugeant à 7500g, durant un temps variable jusqu'à l'obtention du volume désiré. Ensuite les échantillons sont centrifugés à 18000g pendant 20 min. Enfin, les protéines sont séparées par gel filtration avec le système AKTA Pure (Cytiva), sur une colonne HiLoad 16/600 Superdex 200 pg (Cytiva #28989335), avec un **tampon gel filtration PP**. Grâce au chromatogramme, les fractions

présentant un pic à 280 nm sont analysées sur un gel de protéines (donné le nom du paragraphe). Les fractions contenant la protéine d'intérêt sont rassemblées, quantifiées au Nanodrop™ 2000 SpectroPhotometer (ThermoFisher® Scientific) et concentrés dans 150 µL, par ultrafiltration dans des tubes Amicon Ultra (30 kDa, 0,5 mL ; Millipore #C82301) en centrifugeant à 7500g, durant un temps variable jusqu'à l'obtention du volume désiré.

5. Test d'activités

Les tests d'activité ont été effectués par la Docteure Thalia Salinas-Giegé. Les ARN utilisés sont listés en [Annexe VI](#).

Les tests d'activité sont réalisés dans un volume final de 10 µL, dans le **tampon d'activité** avec 1 µM d'ARN et 2 µM de protéine purifiée. Le milieu réactionnel est ensuite incubé à 37°C pendant 30 min. La réaction est arrêtée par l'addition de 10 µL de **solution STOP** et déposée dans un **gel acrylamide 20% dénaturant**. La visualisation sous UV des acides nucléiques est faite par coloration du gel au SYBR-Gold (Invitrogen™ S11494).

RESULTATS

| C-rich tails | A/U-rich tails |
|----------------------|-----------------------|
| UUUUACCCCCCCCCCCCCC | UUUUAAAAAAAAAAAAAAAAA |
| CCCCCCCCCCCCCCCCCCCC | UUCUAAUAAACCC |
| UUUUAAUUCCCCCCCCCCC | AUUUUAAAAA |
| CAUUUAAACCCCCCCCCCA | UUAAAAACCA |
| CCCCCCCUUCCCCCCC | AAUAAAAA |
| UUUCCCCCCCCCCCCC | CCAUA AAAA |
| CCCCCCCCCCCCC | AAACCCAAA |
| CACCCCCCCCCC | UAUACC |
| CCCCCCCCCCCCA | AAAAAAA |
| CCCCCCCCCU | AUUUU |
| CCCCCCCCC | AUAAA |
| CCCCCCCCC | AAAAA |
| CCCCCCCCCU | UUAAA |
| AACCCAC | AUAAA |
| AACCCAC | CCAAA |
| AUCCCCG | CCAAA |
| ACCCCC | AAAAA |
| CCCCC | AGAC |
| CCCCC | AGAC |
| CUCC | AAUU |

Figure R-1 : Exemples de queues obtenues par des expériences de 3'RACE

De (Salinas-Giegé *et al.*, 2017).

Exemples aléatoires de queues polyC (gauche) et polyA/U (droite) observées aux extrémités 3' des ARNm mitochondriaux de *C. reinhardtii*

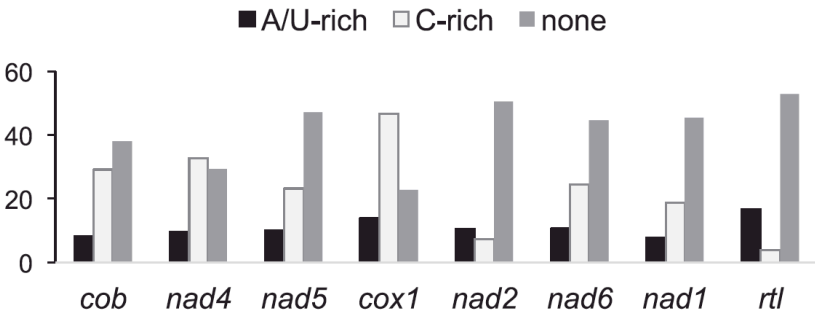


Figure R-2 : Proportions des différentes queues ajoutées pour chaque transcrit

De (Salinas-Giegé *et al.*, 2017).

Histogramme récapitulatif des résultats de 3'RACE, avec les pourcentages de transcrits avec des queues polyA/U (noir), polyC (blanc) ou sans queue (gris), pour chaque gène mitochondrial.

CHAPITRE I : ÉTUDE DE DEUX NUCLEOTIDYL-TRANSFERASES :

PAP4 ET PAP6 : ACTRICES DE LA POLYCYTIDYLATION ?

I. Mise en contexte

1. Analyses des extrémités 3' des ARNm mitochondriaux

Plusieurs études ont montré des ajouts de queues riches en cytidines dans les mitochondries de *C. reinhardtii* (Salinas-Giegé *et al.*, 2017; Cahoon and Qureshi, 2018; Gallaher *et al.*, 2018). Celles-ci sont ajoutées de manière post-transcriptionnelle, aux extrémités 3' UTR des transcrits. Cet ajout est principalement trouvé pour les ARNm des mitochondries de la classe des Chlorophycées. Ces queues sont appelées queues polyC-riches, ou plus simplement queues polyC, car elles contiennent au moins 63,7% de C, mais également des A et des U (entre 15 et 20% selon les études (Salinas-Giegé *et al.*, 2017; Gallaher *et al.*, 2018)) (Figure R-1). Des queues uniquement composées de C ont été observées également, mais elles ne sont pas majoritaires. Très peu de résidus G ont été trouvés. En plus de ces queues polyC, des queues A/U riches ont été observées (Figure R-1). Selon les transcrits, le taux d'ajout d'un type de queues ou de l'autre varie. Par exemple, il y a la même proportion de queues polyC que de A/U pour les transcrits du gène *nad2*, tandis que pour l'ARNm du gène *rtl*, les queues A/U sont majoritaires. Par contre, pour les transcrits des 6 autres gènes (*cob*, *cox1*, *nad1*, *nad4*, *nad5* et *nad6*), les queues polyC sont plus fréquentes (Figure R-2).

Ces queues polyC sont présentes sur 35% des ARN mitochondriaux étudiés en 3'RACE (Salinas-Giegé *et al.*, 2017). Les plus polycytidylés sont les ARNm matures, pour lesquels les ratios de polycytidylation varient selon les gènes de 3,2% à 46,8%, avec en moyenne 21,7% des ARNm matures possédant des queues polyC-riches (Figure R-2). Les transcrits des gènes *nad4* et *cox1* présentent le plus de queues polyC (32,8 et 46,8%, respectivement). La taille des queues polyC-riches, différente selon les transcrits, varie entre 10 et 29 nt, avec une moyenne de 14 nt (Salinas-Giegé *et al.*, 2017; Cahoon and Qureshi, 2018). Il a été observé que les ARNm tronqués au niveau de la CDS (*coding DNA sequence*) n'ont pas d'ajout de séquence riche en C. En revanche, si la séquence 3'UTR est tronquée, il y a aussi des ajouts de queues polyC-riches, mais en moindre proportion et celles-ci sont plus courtes par rapport à celles des ARNm non tronqués.

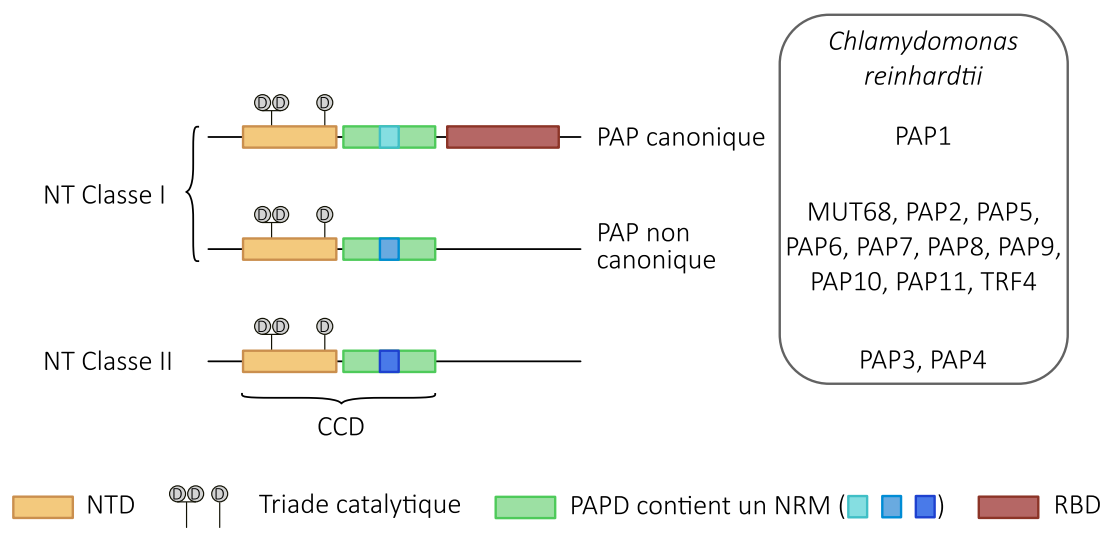


Figure R-3 : Classification des NT

Schéma de la classification des NT. Dans le rectangle à droite sont listées les 13 NT identifiées chez *C. reinhardtii*, selon leur classification.

L'ajout de queues riches en C n'avait jamais été observé auparavant. Si des ajouts de queues riches en U ont déjà été vu chez les trypanosomes par exemple, ils restent rares et peu connus (Aphasizhev *et al.*, 2016). Afin de comprendre par quelle enzyme et comment sont ajoutées ces queues, nous nous sommes intéressées à une famille de protéines, les NT. L'une d'elles pourrait être impliquée dans ces ajouts post-transcriptionnels mitochondriaux.

2. Les Nucleotidyl-transférases : à l'origine de la polycytidylation ?

a) Présentation et classification des NT

Les NT font partie de la super famille des *DNA pol β-like NT* (polymérase β). Ce sont des enzymes qui ajoutent des ribonucléotides, un par un, à l'extrémité 3' des ARN, indépendamment de toute matrice. Cet ajout se fait uniquement sur une extrémité 3'OH libre. Certaines NT n'ajoutent qu'un seul nt, tandis que d'autres en ajoutent une centaine voire plus (Martin, Doublé and Keller, 2008). Elles sont distinguées en 2 classes, I et II (Figure R-3). En effet, le site actif des NT de classe I ou II est différent : chez les NT de classe II, il y a 4 acides aminés supplémentaires, permettant ainsi la distinction entre les enzymes des 2 classes (Yue, Maizels and Weiner, 1996). Dans la classe I, les plus connues sont les PAP canoniques, qui sont responsables de la polyadénylation des ARNm cytosoliques. Il y a aussi des PAP non canoniques qui peuvent ajouter d'autres nucléotides, comme les TUTases ajoutant des U (Scheer *et al.*, 2016). Les PAP non canoniques de la classe I comprennent aussi les CCAses (*CCA-adding enzymes*) des archées, qui ajoutent le CCA à l'extrémité 3' des ARNt. La classe II regroupe les PAP bactériennes et les CCAses bactériennes et eucaryotes (Martin and Keller, 2007; Martin, Doublé and Keller, 2008).

Au niveau structurel, elles contiennent toutes un domaine pol β NTD conservé, avec 3 résidus aspartate (D selon le code à une lettre) qui sont indispensables pour l'activité de l'enzyme. Ces 3 résidus sont aussi appelés la « triade catalytique » (Figure R-3), et leur substitution conduit à l'inactivation de l'enzyme (Scheer *et al.*, 2016). Elles ont aussi un domaine PAPD. Ce dernier possède une région qui varie selon la catégorie de l'enzyme (classe I/II, canonique ou non) : le NRM (*nucleotide recognition motif*). Le NRM des PAP canoniques, est presque invariant dans toutes les lignées eucaryotes (Martin, Doublé and Keller, 2008). Ensemble, les domaines NTD et PAPD forment le CCD (*core catalytic domain*) (De Almeida *et al.*, 2018). Les PAP canoniques de classe I ont en plus un domaine RBD (*RNA binding domain*), permettant une liaison à l'ARN.

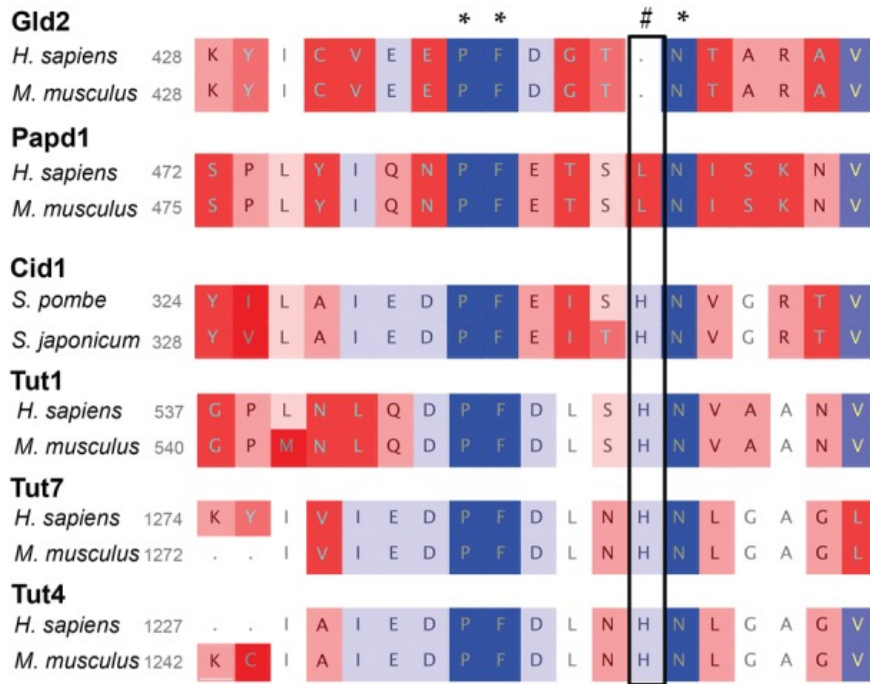


Figure R-4 : La préférence des NT pour leur substrat dépend d'un acide aminé

De (Chung, Hyung Suk Jo and Heinemann, 2016)

Alignement des séquences protéiques du motif NRM des enzymes GLD2 et Papd1 (polyadénylation) avec 3 TUTases (Tut1, Tut7 et Tut4 – polyuridylation) de *H. sapiens* et *M. musculus* ainsi que de Cid1 (polyuridylation) de 2 levures. * : 3 positions strictement conservées ; # : position d'intérêt, conférant la spécificité pour un A ou un U.

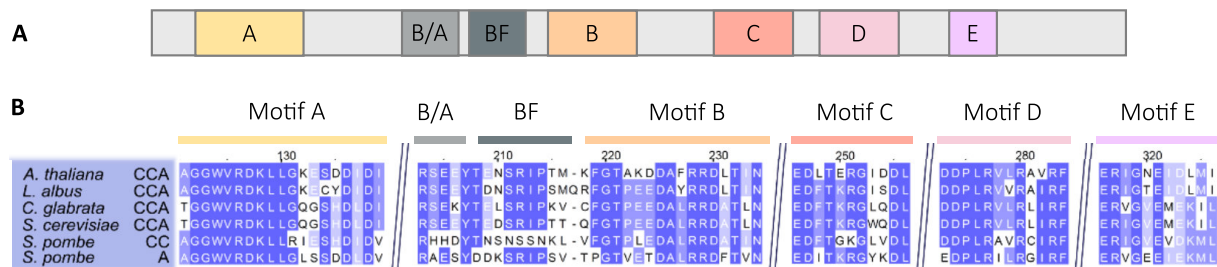


Figure R-5 : Structure et spécificité des CCAses

Adapté de (Erber *et al.*, 2020).

A : Schéma des 7 motifs caractéristiques des CCAses ; B/A : basique et acide ; BF : boucle flexible. **B** : Alignement des séquences des motifs de CCAses (CCA) de 4 organismes (2 plantes : *Arabidopsis thaliana*, *Lupinus albus*, et 2 levures : *Candida glabrata*, *Saccharomyces cerevisiae*) avec les enzymes ajoutant les C (CC) ou le A (A) de la levure *Schizosaccharomyces pombe* ; du bleu foncé au blanc : résidus plus ou moins conservés, respectivement.

b) Spécificité NT-substrat

Le substrat de ces enzymes consiste en générale en un seul type de résidus, comme pour les queues polyA ou polyU. La spécificité d'une enzyme pour un nt substrat dépend de la séquence de la région NRM (Figure R-4). Par exemple, un résidu histidine, situé proche du site actif, a été montré indispensable pour la sélectivité des U par la TUTase Cid1 de *S. pombe* (Yates *et al.*, 2015). Même si les autres résidus sont présents en excès par rapport aux U, Cid1 choisira toujours le U, grâce au résidu histidine, capable de le discriminer parmi toutes les bases disponibles. Par contre, si cette histidine est mutée en une alanine ou une asparagine, l'activité TUTase est convertie en une activité PAP (Lunde, Magler and Meinhart, 2012; Yates *et al.*, 2012). Ce résidu histidine est conservé chez d'autres TUTases (Chung, Hyung Suk Jo and Heinemann, 2016). A l'inverse, l'enzyme humaine GLD2, bien qu'ayant une forte similarité de séquence avec les TUTases, et longtemps considérée comme telle (Chung, Hyung Suk Jo and Heinemann, 2016) possède une activité PAP. Bien qu'elle puisse ajouter tous les nt, elle a une préférence pour les A. Au niveau moléculaire, GLD2 n'a pas l'histidine conservé des TUTases mais l'insertion d'une histidine dans la séquence de GLD2 lui confère une spécificité pour les U. Enfin, l'adényltransférase Papd1 de l'Homme et de la souris ont une leucine à la place de l'histidine, leur conférant leur activité d'adénylation (Chung, Hyung Suk Jo and Heinemann, 2016). Ces données montrent que le changement d'un acide aminé joue sur la spécificité de l'enzyme pour son substrat : une histidine conférant la spécificité aux U, et une leucine ou aucun acide aminé pour la spécificité aux A.

Néanmoins, certaines NT comme les CCAsés peuvent ajouter plusieurs types de nt. Chez la plupart des organismes, une seule enzyme, la CCAsé, ajoute le triplet CCA à l'ARNt. Or chez certaines bactéries et certains eucaryotes, l'ajout du CCA est fait par 2 enzymes : une qui ajoute les 2 C à l'extrémité 3' de l'ARNt (*CC-adding enzyme*) puis l'autre qui ajoute la A (*A-adding enzyme*), en une parfaite collaboration (Tomita and Weiner, 2002; Erber, Betat and Mörl, 2020). Seuls les ARNt possédant les 2 C à leur extrémités 3' peuvent entrer dans le cœur catalytique de l'enzyme ajoutant le A (Tretbar *et al.*, 2011). Si les CCAsés comportent les domaines des NT de Classe II décrits précédemment, elles contiennent en général 7 motifs caractérisés par des acides aminés conservés à certaines positions (Figure R-5) (Neuenfeldt *et al.*, 2008; Erber *et al.*, 2020). Les CCAsés et les enzymes n'ajoutant qu'un A ont des séquences similaires (Figure R-5B) (Tretbar *et al.*, 2011). La différence entre les CCAsés et les enzymes

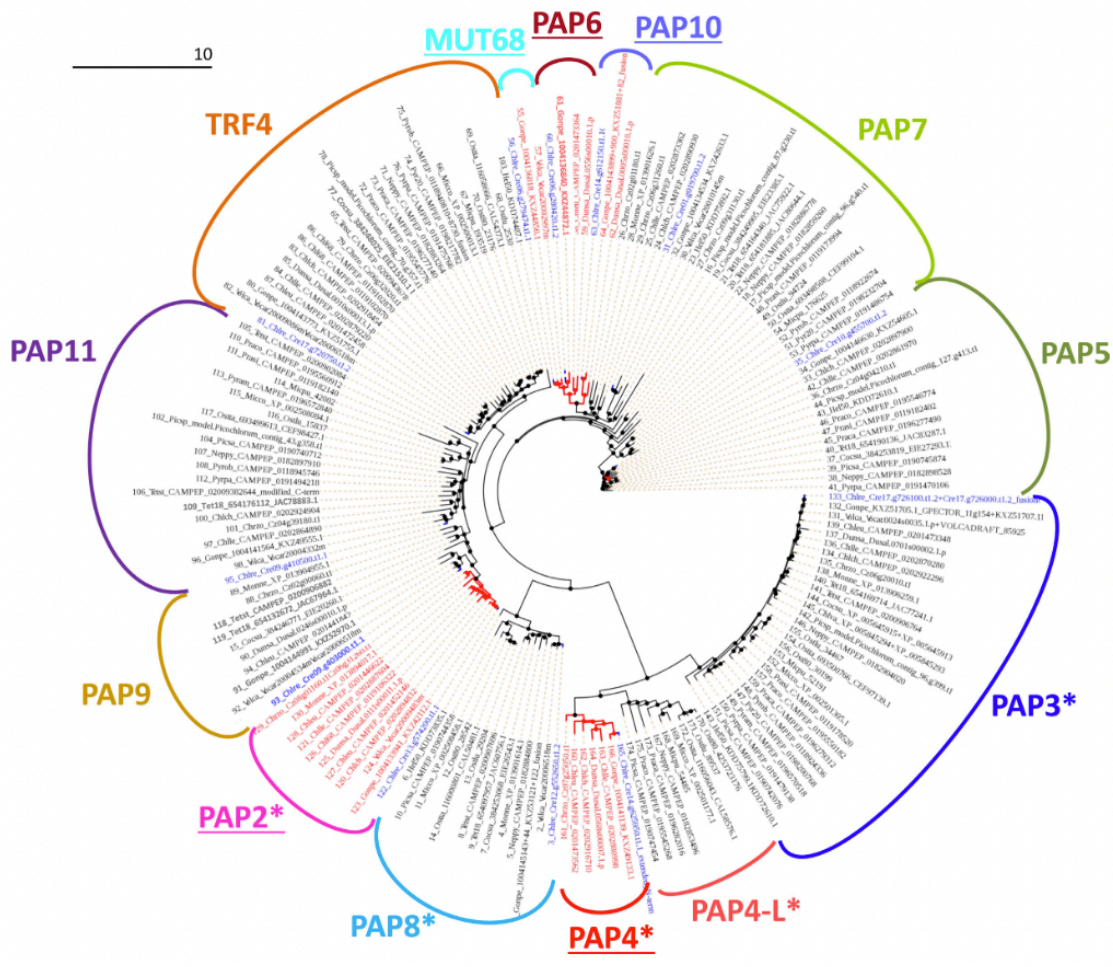


Figure R-6 : Arbre phylogénique des NT des Chlorophytes

De (Salinas-Giegé *et al.*, 2017).

Arbre phylogénique des NT des Chlorophytes non canoniques de classe I et de classe II. En **rouge** sont représentés les espèces appartenant au groupe de Chlorophyceae, en **noir** celles n'en faisant pas parti. Les séquences de *C. reinhardtii* sont en **bleu**. Le nom de chaque groupe se base sur l'orthologue de *C. reinhardtii* sur le site Phytozome (Chlamydomonas reinhardtii v5.5, <http://phytozome.jgi.doe.gov/>) et ceux ne comportant que des membres Chlorophyceae sont soulignés. Les astérisques représentent les groupes pour lesquels le plus de séquences ont une prédiction d'adressage aux organites, par TargetP (Armenteros *et al.*, 2019) ou Predalgo (Tardif *et al.*, 2012).

n'ajoutant que 2 C réside en 2 régions, présente dans les premières mais absentes dans les dernières. Ces régions, appelées motif B/A (composé de résidus basiques et acides) et motif BF (*flexible loop*), sont essentielles pour l'ajout de A, en plus des C. En effet, le motif B/A permet le positionnement correct des 2 C en 3' de l'ARNt au niveau du cœur catalytique de l'enzyme, afin que le A puisse être ajouté. La BF permet quant à elle de moduler la spécificité pour les C et les A, chez les CCAases, afin d'ajouter d'abord des C puis un A. Une étude montre que si ces régions sont délétées dans une CCAase, celle-ci ne pourra plus ajouter de A. En revanche, si elles sont insérées dans une enzyme n'ajoutant que des C, un gain de fonction de l'enzyme est observé, avec un ajout de A possible (Neuenfeldt *et al.*, 2008). Enfin, une étude a montré qu'un résidu glutamate (E avec le code à une lettre) dans le motif E est spécifique des CCAases, permettant de les distinguer des PAP (Martin and Keller, 2004). Cet acide aminé pourrait également avoir un rôle dans le mécanisme de sélection C/A.

c) Quelles candidates à l'ajout des polyC ?

Chez *C. reinhardtii*, 13 NT ont été identifiées (Figure R-3). Parmi elles, la PAP canonique de classe I, PAP1, est responsable de la polyadénylation des ARNm nucléaires. Les NT non canoniques de classe I ou celles de classe II sont donc de bonnes candidates pour une activité d'ajout des queues polyC.

L'enzyme que nous cherchons doit être adressée aux mitochondries et est possiblement spécifique à la classe des Chlorophycées, comme l'ajout de queues polyC. C'est pour cela que des analyses phylogéniques et de prédictions d'adressage ont été faites, sur les 12 NT candidates restantes de *C. reinhardtii* (Figure R-6) (Salinas-Giegé *et al.*, 2017). Parmi les 12 autres candidates possibles, TRF4 et MUT68, des PAP non canoniques de classe I ont déjà été décrites. TRF4 adényle les précurseurs ARNr au niveau nucléaire (Zimmer, Fei and Stern, 2008; Sikorski *et al.*, 2015), tandis que MUT68 permet l'uridylation des petits ARN non codants impliqués dans l'ARN interférence au niveau cytosolique (Ibrahim *et al.*, 2010). Ces deux enzymes ne semblent donc pas posséder d'activité de polycytidylation. Bien que prédites mitochondriales, PAP3 et PAP8 n'ont pas été retenues comme candidate pour notre étude. En effet, PAP3 serait la CCAase triplement adressée (noyau, chloroplaste et mitochondries) (Zimmer *et al.*, 2009), tandis que PAP8 n'est pas spécifique des Chlorophycées.

En plus de MUT68 déjà caractérisée, 4 enzymes sont spécifiques à la classe des Chlorophycées : PAP2, PAP4, PAP6 et PAP10. Seules PAP2 et PAP4 sont prédites

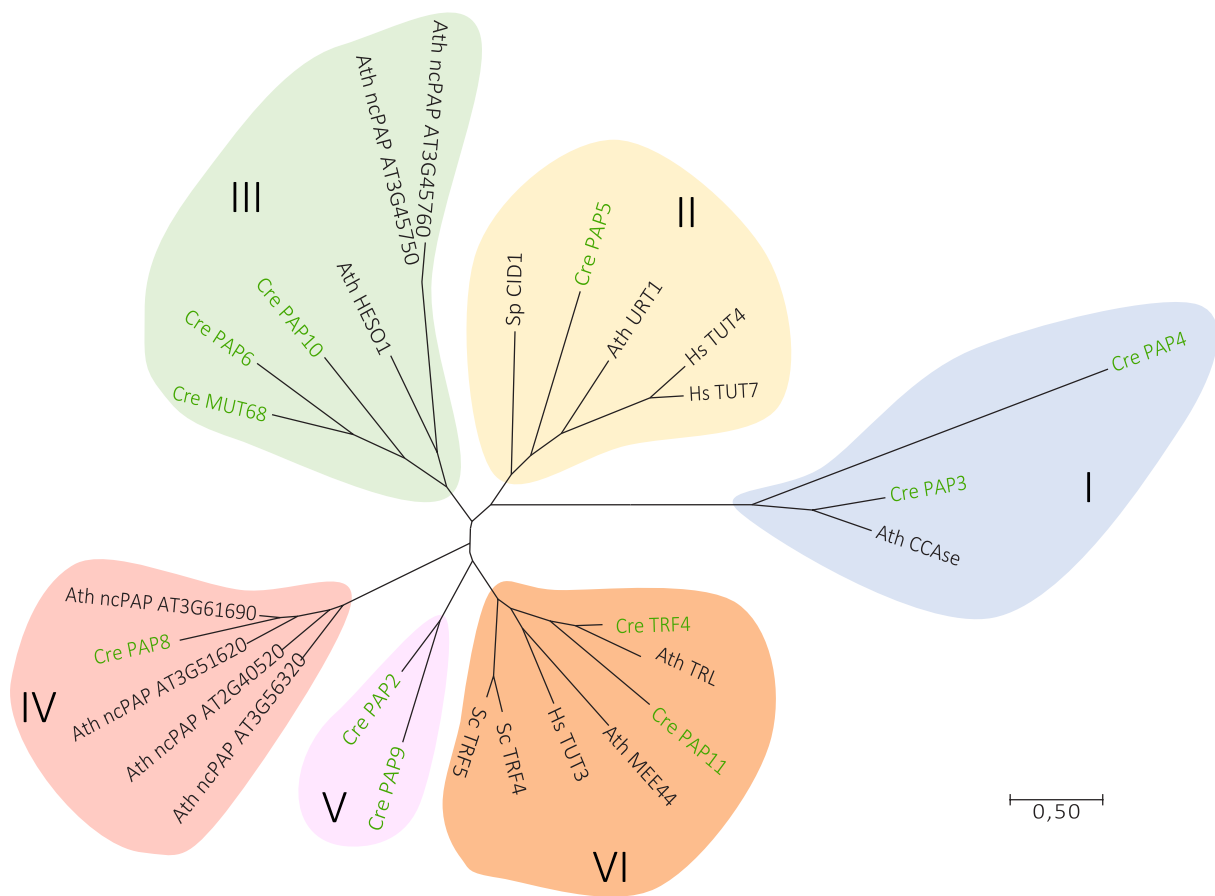


Figure R-7 : Analyse phylogénique des NT de *C. reinhardtii*

Arbre phylogénique obtenu à partir des séquences des domaine NTD et PAPD de 28 protéines, qui ont été alignées avec le programme Muscle, puis l'arbre a été réalisé par MEGA7, par la méthode de la vraisemblance maximale (Kumar, Stecher and Tamura, 2016). Les protéines alignées sont de 5 organismes différents : Ath : *Arabidopsis thaliana*, Cre ; *Chlamydomonas reinhardtii* (en vert), Hs : *Homo sapiens*, Sc : *Saccharomyces cerevisiae* et Sp : *Schizosaccharomyces pombe*. Les différents groupes trouvés sont différenciés par un fond de couleur différent et un chiffre romain pour les numéroter leur a été attribué.

mitochondriales. De plus, l'étude de Zimmer et *al.* (Zimmer *et al.*, 2009) a montré que la séquence N-terminale de PAP4 permet d'adresser l'EGFP à la mitochondrie, en protoplastes de tomate. Ils ont également observé que, *in vitro*, PAP4 peut ajouter des nucléotides en 3' d'un substrat ARN composé de 20 A, et que, *in vivo*, elle complémente une souche bactérienne délétée de la PAP endogène (*E. coli* Δ pcnB). Toutefois, son rôle dans la mitochondrie de *C. reinhardtii* n'a pas été déterminé.

Une autre analyse phylogénique a été réalisée avant mon arrivée par la Docteure Thalia Salina-Giegé, avec les 12 NT de classes I et II, afin de les classer entre elles selon leurs identités (Figure R-7). Cette analyse préliminaire est à but informatif, elle pourrait nous donner des indices sur les fonctions des différentes PAP de l'algue. Pour cela, des séquences de PAP d'autres organismes ont également été analysées avec celles de *C. reinhardtii*. L'analyse a permis de distinguer 6 groupes, dans lesquels sont répartis les différentes PAP. Nous pouvons voir que PAP4 est dans le même groupe que PAP3 et la CCAsé d'*A. thaliana*. Cependant, PAP4 semble assez éloignée des 2 autres, indiquant peut-être une origine commune mais une divergence par la suite. Cela indique un possible ajout de C par PAP4. PAP6 et PAP10 sont quant à elles regroupées avec MUT68 et HES01 d'*A. thaliana*, indiquant peut-être une activité de polyuridylation. Cela dit, elles semblent plus éloignées du groupe II dans lequel sont retrouvées plusieurs TUTases. Cela pourrait être dû à une divergence évolutive, ou à une divergence de fonction.

II. PAP4 et PAP6 sont adressées aux mitochondries

Afin de valider les prédictions, nous avons voulu déterminer la localisation subcellulaire des 4 protéines candidates : PAP2, PAP4, PAP6 et PAP10. Pour étudier la localisation de ces protéines, deux approches complémentaires ont été mises en œuvre : l'expression de la partie N-terminale des protéines en fusion avec une protéine fluorescente en système hétérologue et des analyses protéomiques en spectrométrie de masse sur des fractions de mitochondries purifiées.

1. Colocalisation de PAP4 et PAP6 avec les mitochondries

Pour déterminer la localisation intracellulaire des protéines candidates, nous avons d'abord choisi d'exprimer transitoirement nos constructions en système hétérologue, dans des feuilles de *Nicotiana benthamiana*. Ce choix a été fait en fonction de plusieurs paramètres. Avant mon

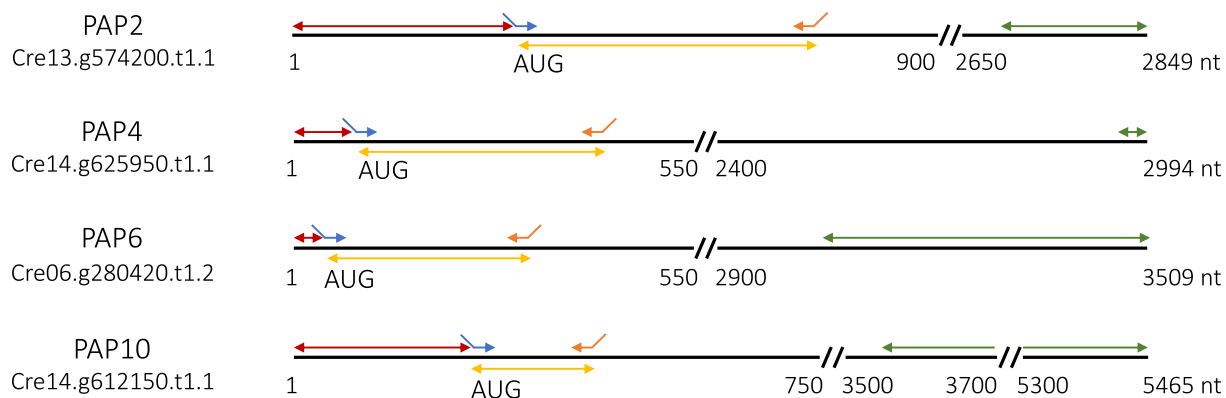


Figure R-8 : Séquences des protéines candidates et régions clonées

Schéma des séquences transcrites correspondant aux 4 protéines candidates et des régions amplifiées clonées dans le vecteur pB7FWG2, avec l'emplacement des amorces, afin d'observer la localisation des protéines de fusion dans *N. benthamiana*. Le nom des protéines est indiqué avec le numéro d'accèsion du gène dans Phytozome. La taille des séquences est indiquée en nt.

PAP2 : séquence N-terminale clonée de 420 pb ; PAP4 : séquence N-terminale clonée de 348 pb ; PAP6 : séquence N-terminale clonée de 332 pb ; PAP10 : séquence N-terminale clonée de 424 pb.

Rouge : 5' UTR ; **Jaune** : MTS putative, appelée séquence Nter ; **Vert** : 3' UTR ; **Bleu** : amorce sens avec le site attB1 ; **Orange** : amorce complémentaire avec le site attB2.

arrivée, l'équipe avait tenté d'exprimer des constructions en transformant *C. reinhardtii*, selon le protocole de l'équipe du Dr. Mayfield de l'Université de Californie (Rasala *et al.*, 2014). Ces expériences n'ont pas abouti, le transgène ne s'exprimant pas dans les cellules. Nous nous sommes donc tournés vers le système hétérologue, *N. benthamiana*. Celui-ci a l'avantage d'être déjà maîtrisé, utilisé et fonctionnel dans le laboratoire. De plus, la localisation mitochondriale de la protéine PAP4 de *C. reinhardtii* en système hétérologue, en protoplastes de tomate (*Solanum lycopersicum*), a déjà été publiée (Zimmer *et al.*, 2009). Cela permet d'utiliser PAP4 pour valider la technique.

Pour les 4 candidates, les prédictions d'adressage ont d'abord été vérifiées par le logiciel PredAlgo (Tardif *et al.*, 2012). La protéine PAP4 est prédite mitochondriale, alors que PAP2 est prédite chloroplastique. PAP6 a une prédiction d'adressage aux voies de sécrétion et PAP10 un adressage autre, pouvant inclure le cytosol, le flagelle, le peroxysome, etc. Lorsqu'un adressage est obtenu pour une protéine, le logiciel PredAlgo donne la séquence d'adressage putative. Les constructions réalisées correspondent aux séquences MTS putatives fusionnées avec le gène codant pour l'EGFP. Pour cela, l'extrémité 5' des ADNc correspondant aux phases ouvertes de lecture, entre 332 et 424 nucléotides (nt) pour chaque protéine candidate, a été clonée dans le plasmide pB7FWG2, contenant le gène codant pour l'EGFP (Figure R-8). Généralement, les MTS de plantes ont une longueur qui varie entre 20 et 60 aa (acides aminés), donc entre 60 et 180 nt (Pujol, Maréchal-Drouard and Duchêne, 2007). Nous avons cloné des séquences un peu plus longues que cela, les protéines chimériques Nter-EGFP en résultant devraient comprendre la MTS, s'il y en a une.

Pour voir la localisation des protéines d'intérêt, des *A. tumefaciens* ont été transformées avec les constructions Nter-EGFP. Des feuilles de *N. benthamiana* ont été agro-infiltrées afin d'exprimer les constructions. Deux contrôles mitochondriaux ont été utilisés : le vecteur pSU9-RFP, qui a la sous-unité 9 de l'ATP synthase mitochondriale en fusion avec la RFP, utilisé pour les observations en feuilles (Michaud *et al.*, 2014) et du Mitotracker, utilisé pour les observations en protoplastes.

Sur la Figure R-9 sont présentées des images de microscopie confocale observées pour les 4 candidates PAP2, PAP4, PAP6 et PAP10. Ainsi, nous avons observé la co-localisation de PAP4-EGFP et du signal du Mitotracker confirmant la localisation de PAP4 dans la mitochondrie (Figure R-9A), comme préalablement publié (Zimmer *et al.*, 2009). Pour PAP6, le signal de EGFP

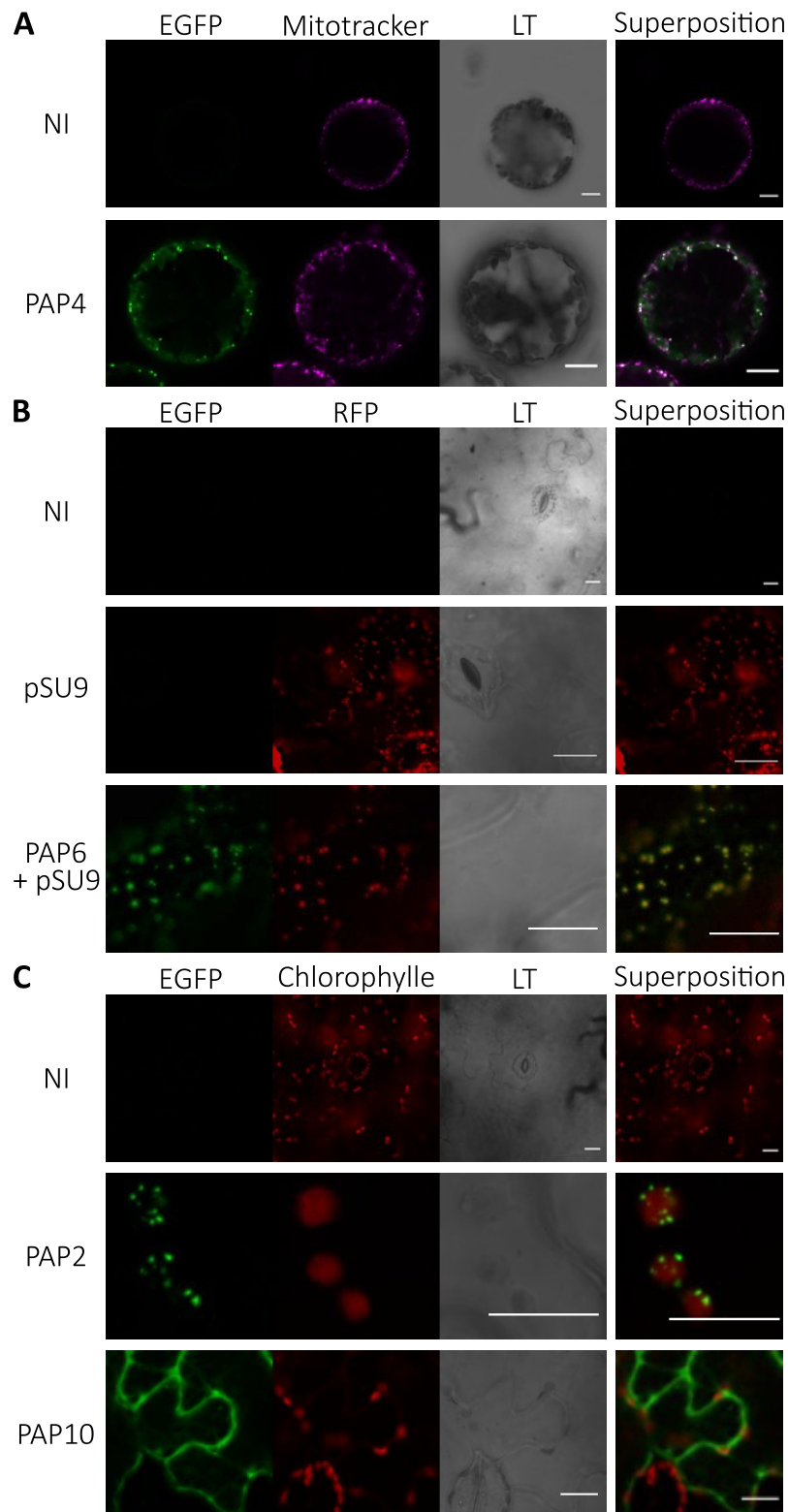


Figure R-9 : Localisation subcellulaire de PAP2, PAP4, PAP6 et PAP10

Localisation dans les protoplastes (**A**) ou cellules (**B** et **C**) de feuilles de *N. benthamiana* des protéines d'intérêt, en microscopie confocale. Barre d'échelle : 15 μ m.

NI : non infiltrée ; LT : lumière transmise ; Superposition : canaux EGFP et Mitotracker (**A**) / RFP (**B**) / Chlorophylle (**C**) superposés.

co-localise avec celui de pSU9-RFP dans les feuilles de *N. benthamiana*, indiquant que cette protéine est mitochondriale (Figure R-9B). Enfin, nous avons noté la formation de petits granules verts à proximité des chloroplastes pour la construction PAP2 alors que pour PAP10, une localisation cytoplasmique a été observée (Figure R-9C).

Ainsi, nos résultats montrent que seules les protéines PAP4 et PAP6 sont mitochondriales.

2. Essais d'expression *in alga*

Nous avons souhaité exprimer les protéines candidates en fusion avec une protéine fluorescente, afin d'observer leur localisation *in vivo* dans les algues et de confirmer les résultats obtenus en système hétérologue avec PAP4 et PAP6.

Pour cela, il a fallu mettre au point un protocole de transformation reproductible et donnant un grand nombre de transformants. J'ai voulu transformer des cellules de *C. reinhardtii* avec des plasmides construits spécialement dans ce but par l'équipe du Dr. Pringle (Onishi and Pringle, 2016). Ceux-ci contiennent un gène codant pour une protéine fluorescente, la NeonGreen, en fusion avec un Flag et le gène *aphVIII*, conférant la résistance à la paromomycine aux cellules transformées. Quatre protocoles d'électroporation différents ont été testés, avec deux souches sauvages différentes de *C. reinhardtii*. La concentration optimale d'antibiotique nécessaire pour la sélection des colonies transformées a été déterminée. Enfin, il a fallu tester plusieurs protocoles d'extraction d'ADN des algues, pour choisir le plus rapide avec un bon rendement en ADN, afin de faire un criblage des colonies par PCR pour déterminer quels sont les transformants ayant intégré le plasmide. Ces tests de transformation ont été effectués avec les vecteurs contrôles du laboratoire du Dr. Pringle ayant le gène de résistance à la paromomycine mais pas la protéine fluorescente. J'ai obtenu des transformants avec les vecteurs contrôles, mais j'ai rencontré des difficultés pour les transformations avec les vecteurs exprimant la protéine fluorescente, puisque qu'aucune fluorescence n'était observable en microscopie confocale. Par Western blot avec un anticorps dirigé contre l'étiquette Flag, aucune expression n'a été observée dans les transformants. Le séquençage des vecteurs n'a décelé aucune mutation pouvant altérer l'expression de la protéine fluorescente.

Une hypothèse était que les vecteurs expriment bien la fluorescence, mais que nous n'arrivons pas à la détecter. Pour vérifier cela, j'ai contacté le Docteur Oliver Caspari (IBPC (Institut de Biologie Physico-Chimique) - Paris), qui nous a procuré 2 souches de *C. reinhardtii* exprimant la

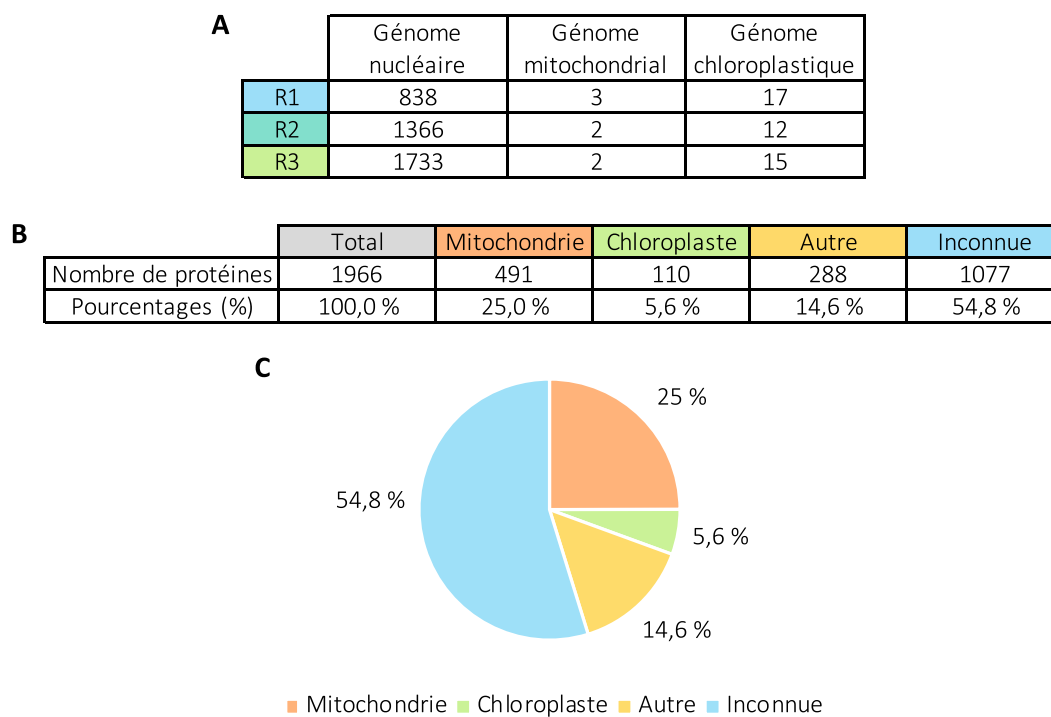


Figure R-10 : Résultats des analyses de spectrométrie de masse

A : Tableau dénombrant les protéines identifiées qui sont codées par le génome nucléaire, mitochondrial et chloroplastique pour chaque réplicat. **Bleu** : R1, réplicat 1 ; **Turquoise** : R2, réplicat 2 ; **Vert** : R3, réplicat 3. **B** et **C** : Tableau (**B**) et diagramme (**C**) dénombrant et montrant la répartition des protéines en fonction de leur localisation, dans les 3 réplicats. Autre : cytoplasme, peroxysome ou flagelle ; Inconnue : localisation non connue.

protéine Venus soit dans le cytosol, soit dans les mitochondries (Caspari, 2022). Plusieurs tentatives ont été faites pour visualiser en microscopie confocale la fluorescence de ces souches, mais cela n'a pas été possible dans nos mains. Nous avons donc des soucis pour visualiser la fluorescence en microscopie confocale.

Par manque de temps durant ma thèse, nous avons laissé cet aspect du projet de côté.

3. Détection de PAP4 et PAP6 dans des mitochondries purifiées

En parallèle de ces analyses, nous avons également réalisé des analyses protéomiques sur des mitochondries de *C. reinhardtii* purifiées, afin d'essayer de déterminer la localisation mitochondriale ou non des 4 protéines candidates d'une autre manière. Une partie des données brutes est présentée en [Annexe VII](#).

Trois extraits mitochondriaux (R1-R3) ont été réalisés de manière indépendante ([Figure R-10](#)). Les protéines qui nous intéressent sont *à priori* solubles et souvent faiblement exprimées. Or, les protéines mitochondriales les plus abondantes sont celles des complexes respiratoires au niveau des membranes mitochondriales. Pour l'analyse protéomique, il fallait donc enrichir la fraction soluble matricielle. C'est ainsi que chaque échantillon a été fractionné en 2 extraits. Pour ce faire, les mitochondries ont été cassées par 3 congélations/décongelations successives afin de libérer les protéines contenues dans la matrice. Une étape de centrifugation a permis de culoter les membranes mitochondriales et les protéines associées à ces membranes. Le surnageant contenant les protéines solubles a été prélevé formant l'extrait « matrice ». Le culot membranaire a ensuite été traité avec 1% Triton afin de solubiliser les protéines. Après centrifugation, le surnageant a été prélevé formant l'extrait « membrane ». Les fractions matricielles et membranaires ont été analysées à la plateforme de protéomique de l'IBMC, par nanoLC-MS/MS (*nano liquid chromatography coupled to tandem mass spectrometry*), avec la version 5.5 du génome de *C. reinhardtii*. Les résultats sont transmis sous forme de tableaux avec le nombre de spectres identifiés pour chaque protéine.

Au total, le nombre de protéines identifiées varie de 838 (R1) à 1766 (R3) ([Figure R-10A](#)). Trois protéines seulement sur les 8 codées par le génome mitochondrial ont été identifiées (*nad4*, *nad5* et *cox1*). Cela n'est pas surprenant car ces protéines sont fortement hydrophobes et ne sont souvent pas détectées en MS (*mass spectrometry*) (Peters, Belt and Braun, 2013). En ce qui concerne les protéines codées par le génome chloroplastique, 17 sur les 64 ont été

A

| | Matrice | Membrane | Ma et Mb |
|----|---------|----------|----------|
| R1 | 163 | 305 | 370 |
| R2 | 395 | 472 | 499 |
| R3 | 257 | 845 | 631 |

B

| R1 | | | |
|------------------------------|-------------|---|-------------|
| Matrice | | Membrane | |
| Protéine | Nb spectres | Protéine | Nb spectres |
| Aconitate hydratase | 274 | ATP synthase F1F0 beta chain | 195 |
| 2-oxoglutarate dehydrogenase | 131 | Alpha subunit of the mitochondrial ATP synthase | 187 |
| Glycine dehydrogenase | 114 | Mitochondrial ATP synthase associated protein 1 | 144 |
| HSP70C | 104 | Mitochondrial ATP synthase associated protein 2 | 135 |
| Pyruvate-formate lyase | 98 | Adenine Nucleotide Translocator 1 | 132 |

| R2 | | | |
|------------------------------|-------------|---|-------------|
| Matrice | | Membrane | |
| Protéine | Nb spectres | Protéine | Nb spectres |
| Aconitate hydratase | 343 | Alpha subunit of the mitochondrial ATP synthase | 146 |
| HSP60 | 171 | ATP synthase associated protein 1 | 94 |
| HSP70C | 161 | ATP synthase F1F0 beta chain | 88 |
| 2-oxoglutarate dehydrogenase | 150 | Aconitate hydratase | 78 |
| Carbonic anhydrase | 139 | Core II subunit of mitochondrial Complex III | 76 |

| R3 | | | |
|---------------------------------------|-------------|---|-------------|
| Matrice | | Membrane | |
| Protéine | Nb spectres | Protéine | Nb spectres |
| Aconitate hydratase | 358 | ATP synthase F1F0 beta chain | 213 |
| HSP70C | 204 | Alpha subunit of the mitochondrial ATP synthase | 203 |
| HSP60 | 183 | Aconitate hydratase | 184 |
| 2-oxoglutarate dehydrogenase | 141 | Mitochondrial ATP synthase associated protein 2 | 170 |
| Cre06.g263250.t1.1 (non caractérisée) | 131 | Mitochondrial ATP synthase associated protein 1 | 145 |

Figure R-11 : Résultats des analyses de protéomique des fractions matricielles et membranaires

A : Tableau correspondant au nombre de protéines trouvées uniquement dans les fractions matricielles (**Jaune**, Matrice), uniquement dans les fractions membranaires (**Rouge**, Membrane) ou dans les deux types de fractions (**Orange**, Ma et Mb), pour les 3 réplicats. **B** : Tableaux recensant les 5 protéines majoritaires identifiées dans chaque fraction (**Jaune**, Matrice ou **Rouge**, Membrane) pour les 3 réplicats, avec le nombre de spectres obtenus. Elles sont toutes codées par le génome nucléaire.

retrouvées dans nos analyses (Figure R-10A). Cela montre une contamination de nos échantillons en chloroplastes.

Pour notre analyse, nous nous sommes basés sur des données protéomiques mitochondriales (Atteia *et al.*, 2009) et chloroplastiques (Terashima, Specht and Hippler, 2011) ainsi que sur les prédictions d'adressage subcellulaire des protéines données par le logiciel PredAlgo (Tardif *et al.*, 2012). Au total, 1966 protéines non redondantes ont été détectées pour les 3 réplicats (Figure R-10B). Parmi elles, 25 % des protéines ont été identifiées comme étant mitochondriales, 5,6 % comme chloroplastiques et 14,6 % auraient une autre localisation (cytoplasmique, flagellaire ou peroxysomale). Les autres protéines, qui représentent 54,8 %, n'ont pas de localisation connue (Figure R-10C). Ainsi un quart des protéines auraient une localisation mitochondriale mais ce chiffre est très probablement sous-estimé car les protéines peuvent être adressées à 2, voire 3, compartiments différents et près de 55 % des protéines identifiées n'ont pas de localisation caractérisée. Cela montre une contamination chloroplastique et cytoplasmique de nos échantillons.

Pour chaque réplicat, les protéines identifiées ont été classées en 3 catégories (Figure R-11A) : les protéines uniquement trouvées dans la fraction matricielle, les protéines uniquement trouvées dans la fraction membranaire et celles qui ont été identifiées dans les 2 fractions. En moyenne, par réplicat, ont été identifiées 272 protéines dans les fractions matricielles, 541 dans les membranaires et 500 présentes dans les deux. Les 3 réplicats sont plutôt homogènes, avec des valeurs proches pour R1 et R2 et plus variables pour R3. Néanmoins, les expériences sont reproductibles et comparables. Les protéines les plus abondantes dans les extraits matriciels sont des protéines mitochondriales matricielles (Figure R-11B), comme par exemple l'aconitate hydratase et la malate déshydrogénase impliquées dans le cycle de Krebs (Gardner, Nguyen and White, 1994; Lee *et al.*, 2001), les protéines chaperonnes mtHSP70C et 60 impliquées dans le maintien de la bonne conformation des protéines (Böttinger *et al.*, 2015), notamment en cas de stress thermique, ou encore la glycine déshydrogénase mitochondriale intervenant dans le système de clivage de la glycine (Douce *et al.*, 2001). De même, dans les extraits membranaires, sont retrouvées principalement des protéines mitochondriales membranaires, comme par exemple des sous-unités ou des protéines associées à l'ATP synthase mitochondriale. Mais il y a aussi des protéines mitochondriales matricielles, comme l'aconitate hydratase ou la malate déshydrogénase. Ainsi, nos échantillons sont bien enrichis

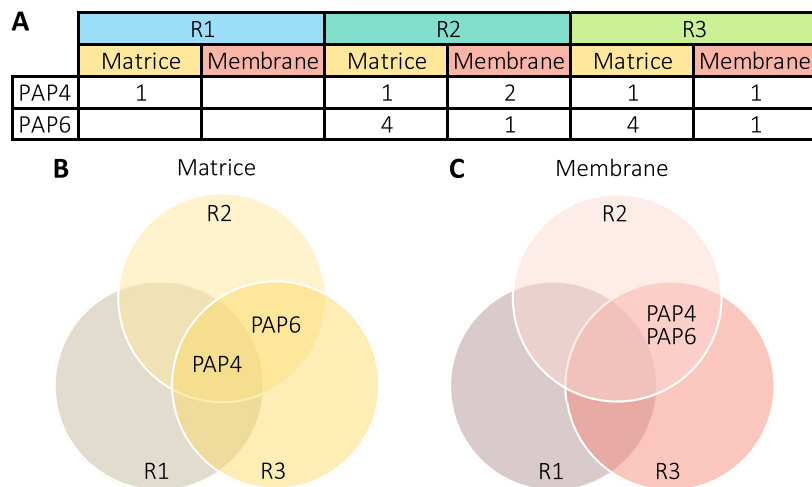


Figure R-12 : PAP4 et PAP6 identifiées dans les extraits mitochondriaux

A : Tableau des protéines d'intérêt trouvées dans les fractions matricielles et/ou membranaires avec le nombre de spectres obtenus pour chacune. **B** et **C** : Diagrammes de Venn avec les protéines matricielles (**B**) ou membranaires (**C**) des 3 échantillons. Si une section est vide, il n'y a pas de protéine commune.

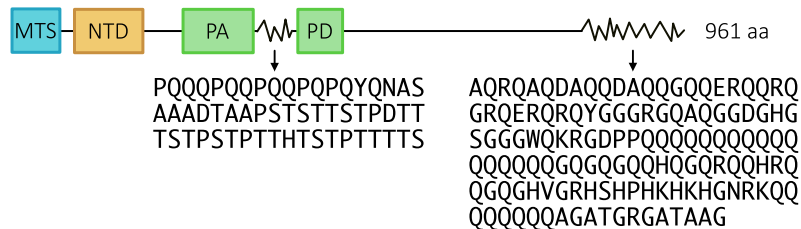


Figure R-13 : Structure de la protéine PAP4

Schéma de la protéine PAP4. Elle possède les 2 domaines caractéristiques des NT, les domaines NTD et PAPD. En N-terminal, elle possède une MTS, dont la séquence exacte n'est pas déterminée. Les régions en zig-zag sont des régions riches en Q (glutamine) et T (thréonine) pour la première, Q pour la deuxième. La séquence de ces régions est indiquée en dessous.

en protéines mitochondriales, mais le fractionnement matrice/membrane n'est pas stricte, il y a des contaminations de l'un dans l'autre et inversement.

Nous avons recherché nos protéines d'intérêt dans nos données protéomiques et seule PAP4 a été identifiée dans les 3 échantillons, alors que PAP6 n'a été identifiée que dans les échantillons R2 et R3 (Figure R-12). L'identification avec un aussi petit nombre de spectre montre que ces protéines sont très peu exprimées. PAP2, PAP10 ou aucune autre NT n'ont été identifiées. Cela est en cohérence avec les données de microscopie, indiquant PAP4 et PAP6 comme mitochondriales, tandis que PAP2 et PAP10 ne le sont pas. Au niveau des extraits membranaires, aucune protéine n'est commune aux 3 échantillons, car ni PAP4 ni PAP6 n'a été identifiée dans la fraction membranaire de R1 (Figure R-12C). Néanmoins, PAP4 et PAP6 sont identifiées dans les échantillons R2 et R3. Cela indique une plus grande hétérogénéité de ces extraits, qui peut être liée à la qualité de purification des mitochondries et/ou la solubilisation des protéines membranaires. Ces résultats ne nous permettent pas de savoir si PAP4 et PAP6 sont matricielles ou membranaires, puisque le nombre de spectres dans les 2 fractions sont similaires.

4. Conclusion

L'ensemble de ces analyses semblent indiquer que, parmi les 4 candidates, seules PAP4 et PAP6 sont mitochondriales. Pour la suite de mes travaux je me suis donné la tâche de déterminer le rôle de ces 2 protéines. Bien que les études de Zimmer et ses collaborateurs semblent indiquer que PAP4 ajoute des A, dans ces mêmes études, les queues polyC n'ont pas été identifiées et le rôle de PAP4 dans la mitochondrie n'avait pas été clairement déterminé (Zimmer *et al.*, 2009). Nous avons donc fait le choix de poursuivre nos analyses sur cette protéine.

III. PAP4, impliquée dans la polycytidylation ?

1. Architecture et classification

La protéine PAP4 est composée de 961 aa (Figure R-13). Nos analyses ainsi que celles publiées par Zimmer *et al.* indiquent que la protéine PAP4 fait partie des NT de classe II, regroupant les PAP bactériennes et les CCAses bactériennes et eucaryotiques (Zimmer *et al.*, 2009). Elle possède 4 des domaines caractéristiques des CCAses (domaines A, C, D et E - Figure R-14) (Neuenfeldt *et al.*, 2008; Erber *et al.*, 2020), identifiés en alignant la séquence protéique de

PAP4 avec celles d'autres CCAsés. Comme cela avait déjà été vu (Zimmer *et al.*, 2009), le résidu E du motif E qui est caractéristique des CCAsés est remplacé par un résidu S (sérine) chez PAP4, indiquant que ce n'est pas une CCAsé canonique (Figure R-14B). De plus, les domaines B/A, BF et B n'ont pas été clairement identifiés.

Cela suggère que cette protéine pourrait être impliquées dans l'ajout de C.

2. Localisation uniquement mitochondriale de PAP4

L'ensemble de nos données pointe vers une localisation mitochondriale de PAP4. Nous avons voulu confirmer ceci par des expériences de western blot. Ainsi, des anticorps dirigés contre la protéine PAP4 ont été produits. Pour cela, une partie de la protéine PAP4 a été produite par la plateforme P3P de l'IBMP. La région C-terminale de la protéine est riche en Q et désordonnée, ce qui a rendu le clonage impossible dans nos mains. C'est pourquoi la protéine purifiée utilisée pour la production d'anticorps n'est pas entière. Elle comporte le domaine NTD et le domaine PAPD, pour un total de 375 aa (Figure R-15). La production des anticorps a été faite par la société Covalab, dans des lapins et les anticorps ont été testés par la Docteure Thalia Salinas-Giegé.

Des mitochondries et des chloroplastes purifiés ainsi que des extraits de protéines totales de *C. reinhardtii* ont été analysés en western blot avec les anticorps @PAP4, par la technicienne Christina Berrissou (Figure R-16). Les anticorps @VDAC1, @Cytf et @NAB1 servent de contrôles, afin de vérifier que nos échantillons d'organelles purifiés sont bien enrichis. Ce qui est en effet le cas pour l'échantillon de mitochondries purifiées, pour lequel nous pouvons observer un signal plus fort pour l'échantillon Mito avec l'anticorps @VDAC1 que pour les échantillons Chloro et Total. Il n'y a aucun signal pour l'échantillon Mito avec l'anticorps @Cytf et un très faible avec l'anticorps @NAB1. Cela montre que nos échantillons Mito sont bien purifiés.

Avec l'anticorps @PAP4, nous pouvons voir qu'un signal n'est visible que pour les échantillons Mito, indiquant une localisation de PAP4 dans nos échantillons de mitochondries purifiées. Cela est en cohérence avec les précédents résultats : PAP4 est bien mitochondriale. Puisque nous ne voyons aucun signal dans les échantillons Chloro, nous pouvons en conclure qu'elle n'est pas également adressée au chloroplaste. Enfin, l'absence de signal dans les échantillons Total peut s'expliquer par une faible expression de la protéine. Cela concorde également avec

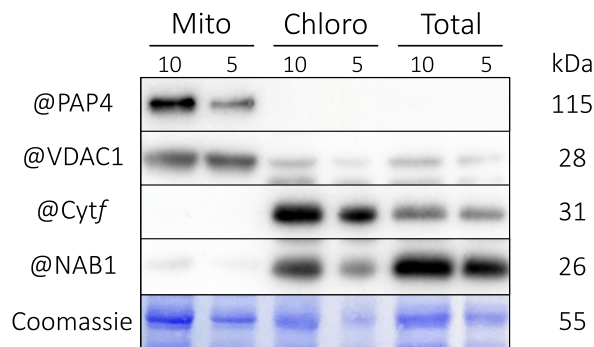


Figure R-16 : PAP4 localisée dans les mitochondries en Western Blot

Échantillons de mitochondries (mito) ou de chloroplastes (chloro) purifiés, ainsi que des protéines totales (total) déposés sur gel SDS-PAGE 8-16% (gradient), en 2 quantités différentes (5 et 10 µg). Après transfert sur membrane, la membrane a été colorée au bleu de Coomassie (Coomassie) et incubée avec différents anticorps : @PAP4, @VDAC1 (protéine mitochondriale) @Cytf (protéine chloroplastique) et @NAB1 (protéine cytosolique).

kDa : taille des protéines

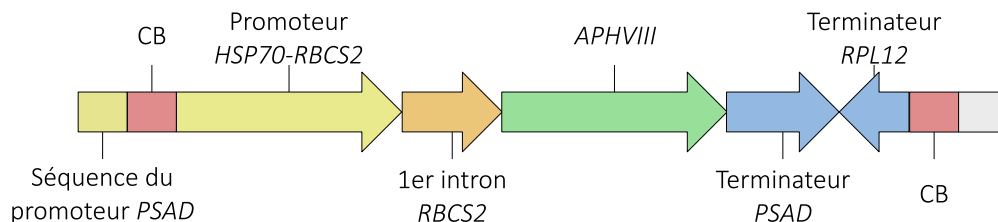


Figure R-17 : Cassette insérée dans la banque de mutants CLiP

D'après <https://www.chlamylibrary.org/showCassette?cassette=CIB1>.

Schéma de la construction de la cassette CIB1, de 2223 pb. Les séquences CB (code barre) sont des séquences aléatoires, chaque insert et donc chaque mutant a des séquences CB différentes, cela permet de les identifier et de les recenser. Les séquences du promoteur *PSAD* et celle en grisée (non codante) servent à délimiter les séquences CB.

Jaune : séquences de promoteur ; **Rouge** : CB ; **Orange** : séquence d'intron, non codante ; **Vert** : séquence codante ; **Bleu** : séquence de terminateurs ; **Gris** : séquence non codante.

le fait que peu de spectres ont été détectés pour PAP4 lors des analyses de spectrométrie de masse.

Ces données nous permettent de confirmer la localisation mitochondriale de PAP4. De plus, ces résultats indiquent que l'anticorps @PAP4 n'est pas utilisable sur des extraits de protéines totales et une purification des mitochondries est nécessaire.

3. Obtention de lignées mutantes

Afin d'étudier le rôle de PAP4 dans la cellule, nous avons initié des analyses de génétique inverse et pour ce faire, nous avons étudié deux types de mutants de PAP4 : des mutants d'insertion et des mutants amiRNA.

a) Mutants d'insertion PAP4

Une banque de mutants de *C. reinhardtii* existe : la banque CLiP (Li *et al.*, 2019). Les mutants ont été obtenus par électroporation de cellules haploïdes avec une cassette qui est insérée de manière aléatoire dans le génome. Cette cassette contient un gène de résistance à la paromomycine (*APHVIII*) précédé d'un promoteur, 2 terminateurs et des séquences aléatoires, servant de code barre pour le mutant afin de les identifier et recenser (Figure R-17). Un des terminateurs est dans le sens du gène *APHVIII*, l'autre en antisens, afin que la transcription soit arrêtée par la cassette peu importe dans quel sens celle-ci s'insère.

Pour la protéine PAP4, au début de ma thèse, il y avait 2 mutants disponibles dans la banque. Il y en a maintenant 5. Tous les mutants de la banque ont une prédiction d'insertion dans un intron. Nous avons obtenu les 2 mutants disponibles à l'époque, les souches LMJ.RY0402.079134 et LMJ.RY0402.125943, que nous renommerons *pap4-1::CIB* et *pap4-2::CIB*, respectivement. La banque CLiP prédit une insertion de la cassette dans l'intron 2 pour *pap4-1::CIB* et dans l'intron 3 pour *pap4-2::CIB* (Figure R-18).

Afin de vérifier la bonne insertion de la cassette, à l'endroit prédit par la banque, des PCR sur l'ADN génomique des souches ont été réalisées, à 2 loci différents. Le « Locus PAP4 » désigne la zone théorique d'insertion renseignée par la banque CLiP, la PCR est faite avec des amorces de part et d'autre du site d'insertion, dans le gène *PAP4* (Figure R-18). Le « Locus contrôle » désigne une région entre les gènes *HRP4* et *OAT1*, cela constitue un contrôle de l'intégrité de l'ADN génomique utilisé comme matrice pour les PCR. Pour chaque mutant, la souche reçue de la banque a été striée fin d'obtenir des colonies isolées. Trois colonies ont été testées pour

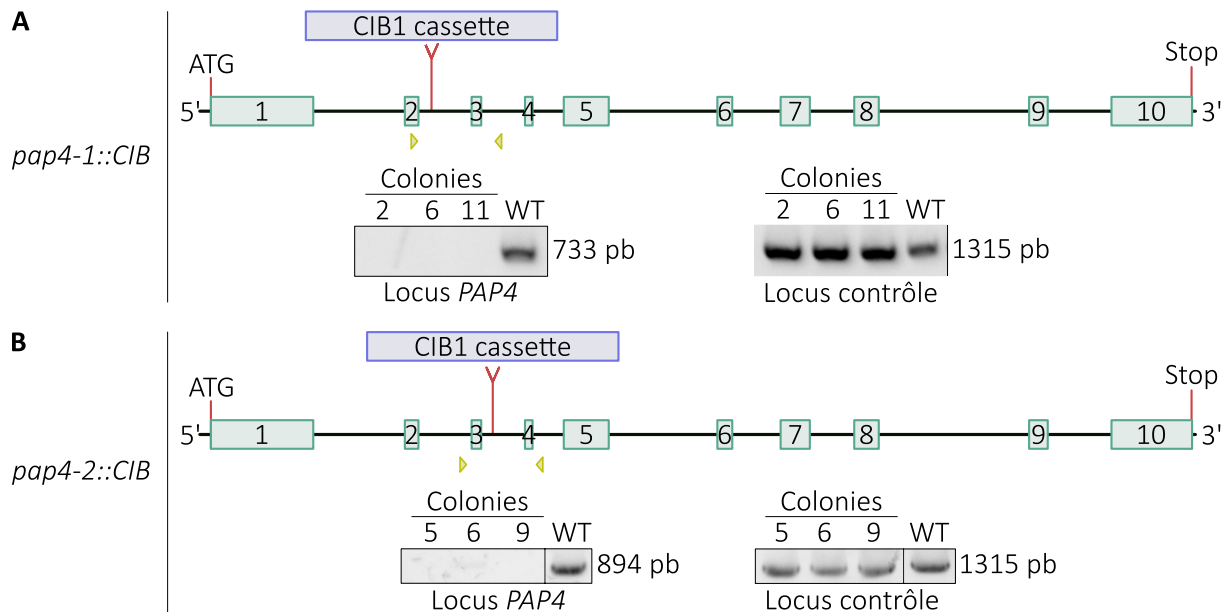


Figure R-18 : Vérification de l'insertion de la cassette dans les mutants CLiP

Schéma du gène *PAP4* et de l'insertion théorique de la cassette CIB, pour *pap4-1::CIB* (A) et *pap4-2::CIB* (B) avec les gels d'électrophorèse montrant les résultats des PCR, pour le locus *PAP4* et le locus contrôle. **Turquoise** : exons numérotés ; **Jaune** : amorces pour la PCR ; pb : taille.

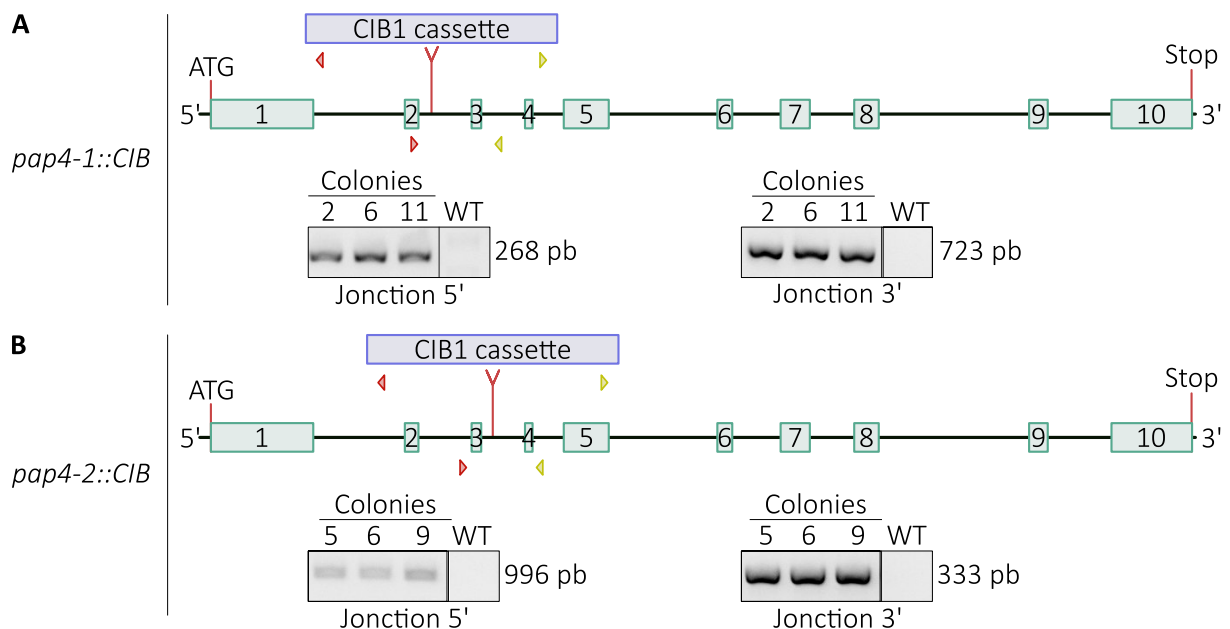


Figure R-19 : Amplification des jonctions gène - cassette dans les mutants CLiP

Schéma du gène *PAP4* et de l'insertion théorique de la cassette CIB, pour *pap4-1::CIB* (A) et *pap4-2::CIB* (B) avec les gels d'électrophorèse montrant les résultats des PCR des jonctions 5' et 3'. **Turquoise** : exons numérotés ; **Rouge** : amorces couvrant la jonction 5' gène - cassette ; **Jaune** : amorces couvrant la jonction 3' gène - cassette ; pb : taille.

chaque mutant, en comparaison avec la souche de référence CC-5325 fournie par la banque, nommée WT (*wild type*).

La [Figure R-18A](#) présente les résultats obtenus pour *pap4-1::CIB*. Des amplifications dans un locus contrôle (entre les gènes *HRP4* et *OAT1*) vérifient que la réaction de PCR a bien fonctionné, et que l'ADN matrice est intègre. Pour les 3 colonies testées, aucune amplification PCR n'est observable à la taille attendue pour le locus *PAP4*. Cela indique que le gène est interrompu au niveau de l'intron 2. Des amplifications sont observables pour les 3 colonies pour le locus contrôle. Il en est de même pour *pap4-2::CIB* ([Figure R-18B](#)). Il n'y a pas d'amplification du locus *PAP4*, indiquant que le gène est interrompu.

Puisque le gène *PAP4* est interrompu dans les 2 mutants, j'ai vérifié que cela est dû à l'insertion de la cassette CIB1. Pour cela, des PCR ont été faites avec des amorces dans le gène *PAP4* et dans la cassette, pour l'ensemble des colonies mutantes ainsi que pour le WT ([Figure R-19](#)). Ces PCR sont appelées « Jonction 5' » lorsque l'amorce est en amont du site d'insertion dans le gène *PAP4* et « Jonction 3' » lorsqu'elle est en aval. Pour les 3 colonies *pap4-1::CIB* et *pap4-2::CIB* et pour les 2 types de PCR, des amplifications sont observables pour les colonies testées alors qu'il n'y en a pas pour le WT.

Ensemble, ces résultats confirment que l'insertion est présente dans les souches mutantes, dans l'intron 2 pour les colonies *pap4-1::CIB* et dans l'intron 3 pour les *pap4-2::CIB*.

Ces dernières amplifications ont été séquencées, afin d'obtenir une cartographie précise des mutants, avec l'endroit d'insertion exact d'insertion dans le gène *PAP4*. Dans les 2 cas, la cassette est insérée en antisens. Pour les colonies *pap4-1::CIB*, il manque 12 nt par rapport à la séquence entière de la cassette CIB1. Pour les colonies *pap4-2::CIB*, il manque 5 nt à l'extrémité 5' et 4 nt à l'extrémité 3' de la cassette CIB1. Or cela n'affecte pas l'expression de la cassette, puisque les mutants ont été sélectionnés par la banque sur leur résistance à la paromomycine, conférée par la cassette.

Les 3 colonies testées par souche mutantes étant similaires, seule l'une d'entre elle a été utilisée pour la suite des expériences : la colonie 2 pour *pap4-1::CIB*, et la colonie 5 pour *pap4-2::CIB*.

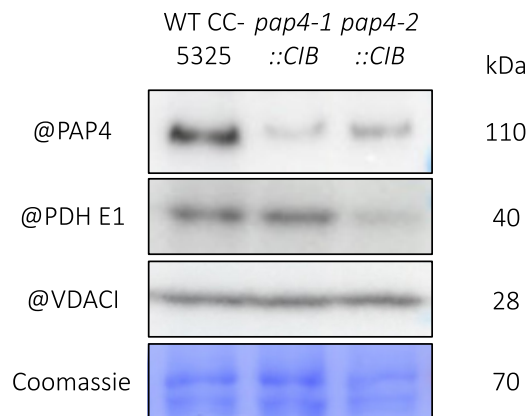


Figure R-20 : Les mutants d'insertion PAP4 sont des mutants *knock-down*

Échantillons de mitochondries enrichies déposés sur gel SDS-PAGE 8-16% (gradient), pour les mutants d'insertion PAP4 et la souche WT correspondante. Après transfert sur membrane, la membrane a été colorée au bleu de Coomassie (Coomassie) et incubée avec différents anticorps : @PAP4, @PDH E1 et @VDAC1. kDa : taille des protéines

b) L'épissage est altéré dans les mutants d'insertion *pap4*

Les insertions étant dans des introns, il faut vérifier que cela impacte l'expression de la protéine. En effet, les introns chez *C. reinhardtii* ont une taille moyenne de 373 pb (Merchant *et al.*, 2007). Or la cassette CIB de plus de 2 kb, est beaucoup plus grande. Le processus d'épissage est sans doute perturbé par une si grande insertion. Mais il est également possible que l'épissage se fasse correctement et que l'expression de la protéine se fasse tout à fait normalement par rapport à une souche WT, l'insertion n'ayant alors aucun effet.

Nous avons donc testé la présence de la protéine PAP4 dans les mutants par western blot. Ces analyses ont été faites sur des échantillons de mitochondries enrichies. En effet, les souches de la banque CLiP ont une paroi cellulaire, rendant la purification des mitochondries beaucoup plus compliquée. C'est pour cette raison que nous avons choisi de travailler avec des mitochondries enrichies.

J'ai donc réalisé des Western blot avec les anticorps @PAP4 (Figure R-20). Les anticorps dirigés contre la protéine PDH E1 matricielle et la protéine VDACL de la membrane externe ont été utilisés comme contrôles afin de vérifier respectivement l'intégrité et l'enrichissement en mitochondries. *pap4-2::CIB* montre un signal moins fort avec l'anticorps @PDH E1 que les échantillons WT ou *pap4-1::CIB*. Cette souche mutante semble donc contenir moins de protéines matricielles que les 2 autres échantillons, mais autant de protéines membranaires. Les analyses de Western blot montrent une diminution, mais pas une absence totale, de la protéine PAP4 dans les mutants. Cette diminution est plus forte pour *pap4-1::CIB* que pour *pap4-2::CIB*. Nous avons donc des mutants *KD* (*knock-down*). Au vu de la différence entre les contrôles matriciels et membranaires, il est possible qu'il y ait plus de protéine PAP4 exprimée dans le mutant que ce que les analyses de western blot montrent avec l'anticorps @PAP4. Ces analyses sont à refaire, avec des quantités de protéines matricielles et membranaires équivalentes pour le mutant *pap4-2::CIB*, afin de vérifier la réelle expression de la protéine PAP4 dans cette souche.

Afin de mieux comprendre ce résultat, j'ai analysé les transcrits du gène *PAP4* par RT-qPCR. Six couples d'amorces ont été choisis à différentes positions du transcrit, avant, autour et après l'insertion des cassettes (Figure R-21). Avant l'insertion, l'expression relative du gène *PAP4* dans le mutant *pap4-1::CIB* ne semble pas altérée, contrairement à celle de *pap4-2::CIB*, qui

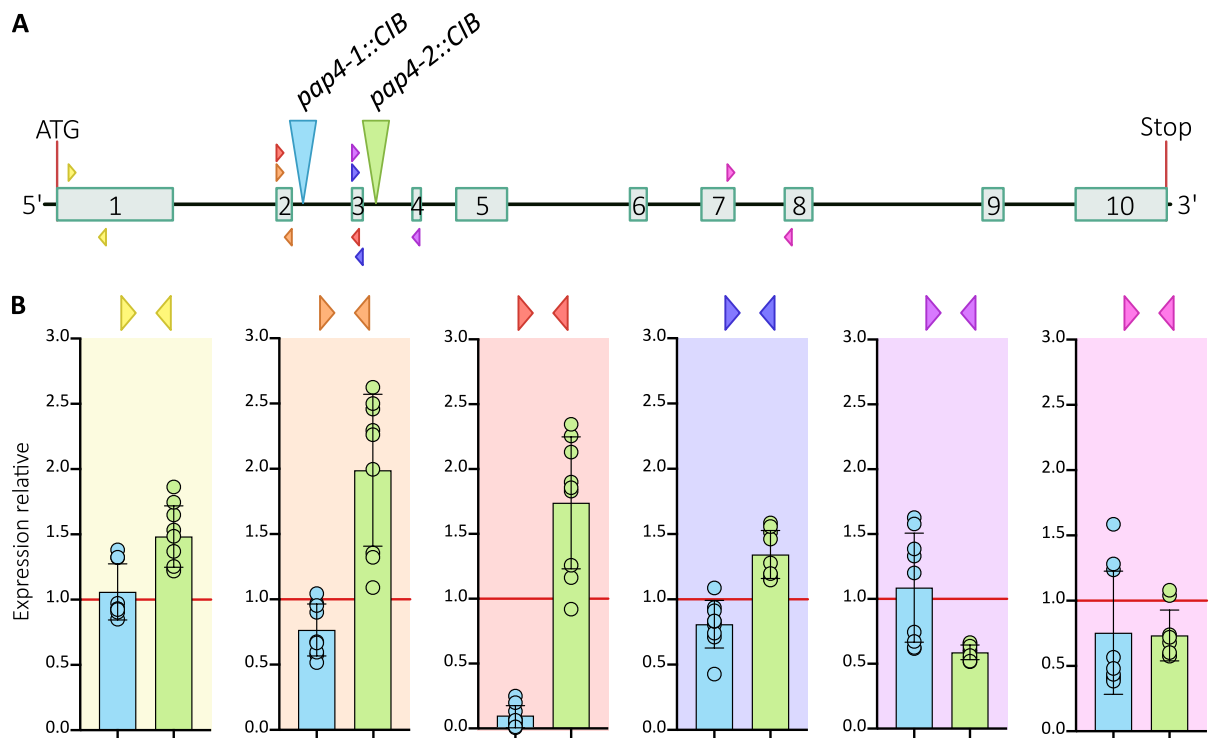


Figure R-21 : Vérification de l'expression du transcrite du gène *PAP4* dans les mutants CLiP

A : Schéma du gène *PAP4* et de l'insertion de la cassette CIB, pour *pap4-1::CIB* (bleu) et *pap4-2::CIB* (vert). Les couples d'amorces utilisés en RT-qPCR sont schématisés par des triangles de différentes couleurs. **Turquoise** : exons numérotés. **B** : Histogrammes de l'expression relative de l'ARNm de *PAP4* dans les 2 mutants CLiP, mesurée par RT-qPCR. Chaque couple d'amorce est associé à une couleur dans (**A**) qui est reportée en tant que couleur de fond. Le trait rouge horizontal indique le niveau d'expression relative du WT normalisé à 1. Les histogrammes présentent la moyenne de 3 réplicats biologiques ($n = 3$, barre verticale). Pour chaque réplicat biologique, 3 réplicats techniques ont été faits. Toutes les valeurs obtenues sont représentées par les ronds sur les histogrammes.

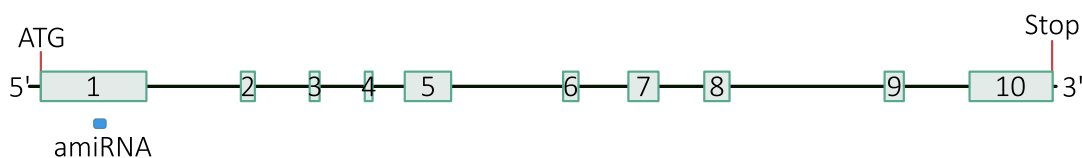


Figure R-22 : Génération de mutants amiRNA de *PAP4*

Schéma du gène *PAP4* avec les exons numérotés (turquoise) et la région ciblée par l'amiRNA (bleu).

est augmentée, jusqu'à 200%. Une diminution de l'expression du transcrite *PAP4* est observée avec des amorces choisies de part et d'autre de l'insertion, indiquant que peu de transcrits sont correctement épissés.

Ces résultats révèlent un défaut d'épissage des introns dans lesquels sont insérées les cassettes CIB. Pour le mutant *pap4-1::CIB*, seulement 9% des transcrits ont un épissage correct de l'intron 2. Pour le mutant *pap4-2::CIB*, il y a un défaut d'épissage de l'intron 3, avec seulement 58% des transcrits correctement épissés. Pour les 2 souches mutantes, il y a donc encore des transcrits qui sont correctement maturés, c'est pourquoi il n'y a pas d'absence totale de la protéine en western blot. Ces résultats expliquent l'obtention de mutants KD, puisqu'un défaut d'épissage des introns avec l'insertion de la cassette a été observé mais il n'est pas total.

En conclusion, il y a encore des transcrits *PAP4* qui sont correctement épissés dans les lignées mutantes *pap4-1::CIB* et *pap4-2::CIB*, permettant ainsi la synthèse d'une protéine mature et fonctionnelle mais en moindres quantités.

c) Génération de mutants amiRNA

En parallèle du génotypage des mutants CLiP, nous avons également initié la production de d'autres souches mutantes pour PAP4 par la méthode des amiRNA. Cette méthode était déjà maîtrisée dans le laboratoire par la Docteure Thalia Salinas-Giegé, et elle était mieux établie chez *C. reinhardtii* que la technologie CRISPR-Cas au début de ma thèse. La méthode amiRNA s'appuie sur le processus d'ARN *interférence*, naturellement présent dans les cellules, en introduisant des miARN artificiels choisis pour cibler un gène d'intérêt. Cela résulte, en général, en une diminution de l'expression du gène et non une extinction complète (Molnar *et al.*, 2009).

J'ai donc généré des mutants amiRNA, ciblant une séquence du premier exon du gène (Figure R-22).

Afin d'avoir une souche contrôle pour les analyses, j'ai également transformé des algues avec le vecteur pHUN vide, le vecteur utilisé pour les constructions amiRNA. Les transformants ont subi différents tests afin de ne garder que ceux qui se comportent comme une souche sauvage : se développent correctement à l'obscurité, croissent en milieu minimum, possèdent des flagelles. Sur les 20 colonies testées, les 14 colonies validant ces critères ont été mélangées.

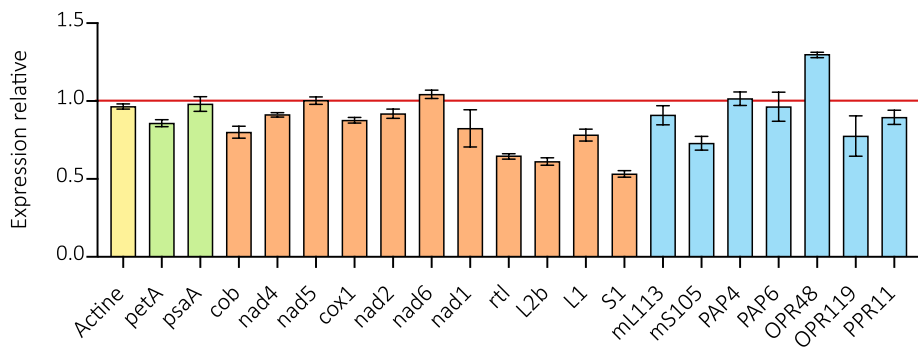


Figure R-23 : Analyses des transcrits de la souche contrôle pHUN

Histogrammes des résultats d'analyses de RT-qPCR sur la souche pHUN (résultant du mélange de 14 colonies transformées avec le vecteur pHUN vide), avec l'expression relative pour chaque transcrit. Le trait **rouge** horizontal indique le niveau d'expression relative du WT normalisée à 1. n=1. **Jaune** : gène nucléaire ; **Vert** : gènes chloroplastiques ; **Orange** : gènes mitochondriaux ; **Bleu** : gènes nucléaires dont les protéines sont localisées dans les mitochondries.

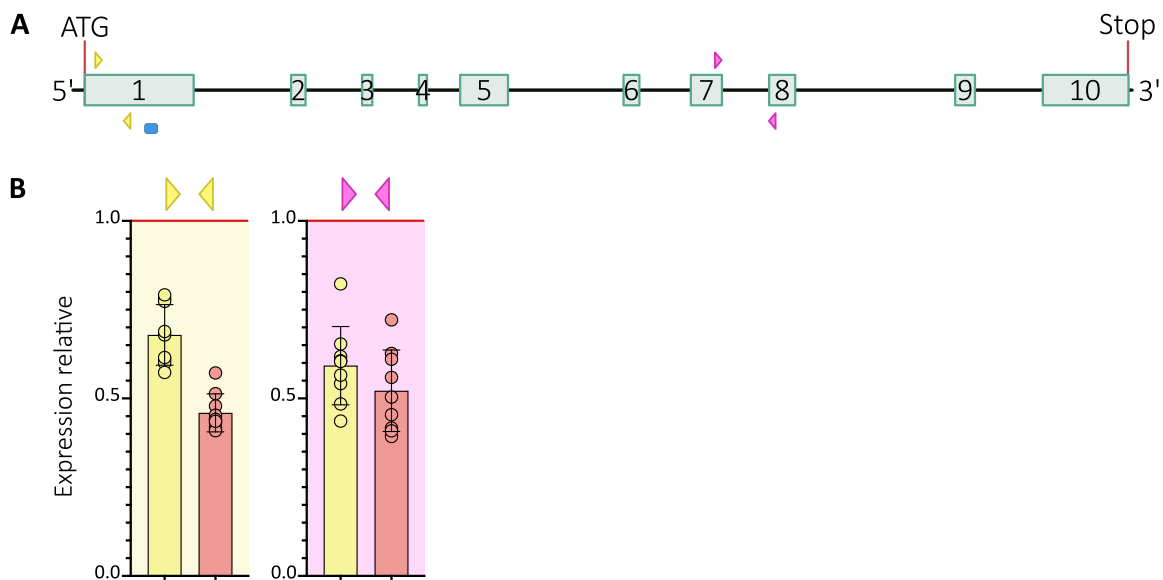


Figure R-24 : Vérification de l'expression du transcrit du gène *PAP4* dans les mutants amiRNA

A : Schéma du gène *PAP4* et de la région ciblée par l'amiRNA (**bleu**). Les différents couples d'amorces utilisés en RT-qPCR sont schématisés par des triangles de différentes couleurs. **Turquoise** : exons numérotés. **B** : Histogrammes de l'expression relative de l'ARNm de *PAP4* dans les mutants *amiRNA amipap4-2* (**jaune**) et *amipap4-37* (**rouge**), mesurée par RT-qPCR. Chaque couple d'amorce est associé à une couleur dans (**A**) qui est reportée en tant que couleur de fond. Le trait **rouge** horizontal indique le niveau d'expression relative du contrôle normalisé à 1. Les histogrammes présentent la moyenne de 3 réplicats biologiques (n = 3), avec l'écart-type (+/-, barre verticale). Pour chaque réplicat biologique, 3 réplicats techniques ont été faits. Toutes les valeurs obtenues sont représentées par les ronds sur les histogrammes.

Cela permet d'atténuer les effets liés à l'insertion aléatoire du vecteur dans le génome nucléaire. Ce mélange de colonies sera appelé plus simplement « pHUN » ou contrôle. Des analyses par RT-qPCR ont permis de vérifier les taux de différents transcrits, pour valider son utilisation en tant que contrôle (Figure R-23). J'ai choisi d'analyser l'expression de gènes nucléaires et chloroplastiques n'agissant pas dans la mitochondrie, ainsi que de gènes nucléaires et mitochondriaux ayant un rôle dans cette organelle. Ces derniers comprennent les gènes des 8 protéines, de *L2b* et deux gènes d'ARNr (*L1* et *S1*) codés par le génome mitochondrial, ainsi que les gènes nucléaires de deux protéines du mitoribosome (*mL113* et *mS105*), de nos protéines candidates PAP4 et PAP6, ainsi que deux OPR et une PPR identifiées dans les extraits mitochondriaux lors des analyses de protéomique (données non montrées). Ces analyses ont été faites en utilisant la souche sauvage CC-4351, transformée pour l'obtention du contrôle pHUN et des mutants amiRNA, en tant que référence, en un seul réplicat, avec 21 transcrits de gènes différents ciblés. Globalement, l'expression relative des gènes testés est similaire à la souche contrôle, sauf pour les transcrits des gènes *rtl*, *L2b* et *S1* dont le niveau du transcrit est diminué entre 35 et 45%. Ces variations entre la souche pHUN et la souche WT CC-4351 peuvent être dues à l'insertion aléatoire du vecteur dans le génome nucléaire, et pourrait se retrouver dans tous les mutants amiRNA. Cela montre l'importance d'avoir une souche contrôle pour l'analyse des mutants amiRNA. Cette souche pHUN sera donc utilisée comme contrôle pour les analyses des mutants amiRNA.

En tout, 83 transformants ont été obtenus sur 2 transformations indépendantes. Des analyses de RT-qPCR sur 20 transformants ont permis d'isoler 2 mutants intéressants : *amipap4-2* et *amipap4-37* (Figure R-24). Pour le mutant *amipap4-2*, le niveau du transcrit est globalement diminué de 36,5%, en moyenne avec les 2 couples d'amorces. Pour le mutant *amipap4-37*, le taux de transcrits est diminuée de 51%. Les différences d'expression relative entre 2 couples d'amorces peuvent s'expliquer par le fait que les différents couples d'amorces n'ont pas la même efficacité d'hybridation. De plus, la répression de l'expression du gène ciblé par l'amiRNA peut se faire par clivage et dégradation du transcrit. Il est donc possible que 2 régions du transcrit ne soient pas affectées de la même manière, l'une étant plus rapidement dégradée que l'autre par exemple. Néanmoins, les résultats pour chaque mutant sont du même ordre de grandeur pour les 2 couples d'amorces utilisées.

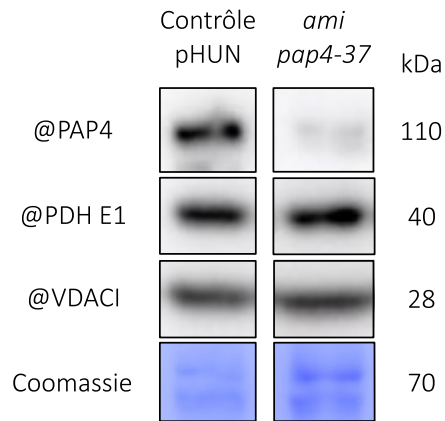


Figure R-25 : Le mutant *amipap4-37* est un mutant *knock-down*

Échantillons de mitochondries purifiées (*mito pures*) déposés sur gel SDS-PAGE 8-16% (gradient), pour le mutant amiRNA de PAP4 et la souche WT correspondante. Après transfert sur membrane, la membrane a été colorée au bleu de Coomassie (Coomassie) et incubée avec différents anticorps : @PAP4, @PDH E1 et @VDAC1. kDa : taille des protéines

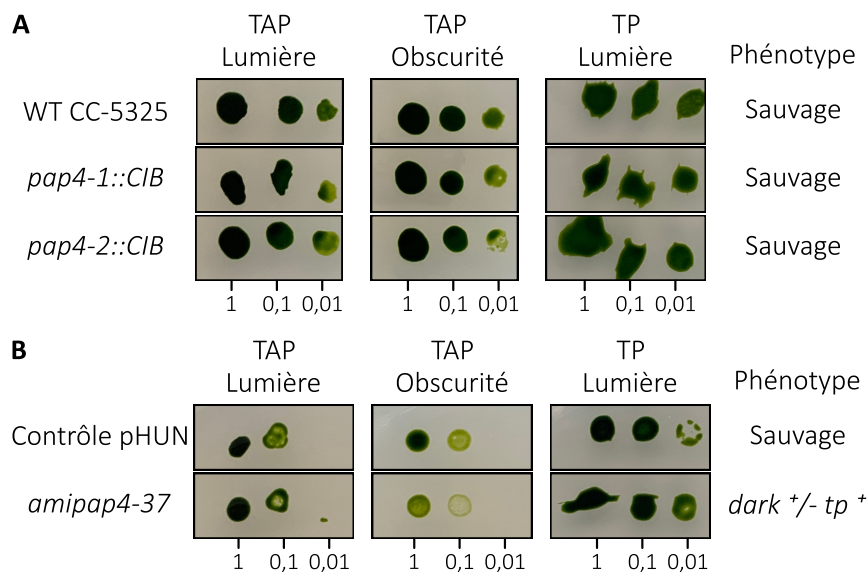


Figure R-26 : Tests phénotypiques des mutants de PAP4

Photos des phénotypes de croissance selon la luminosité (Lumière/Obscurité) ou le milieu (TAP/TP), pour les mutants CLiP (A) et amiRNA (B). La concentration par rapport à la culture de départ (1 = 6×10^6 cellules/mL) sont indiquées en dessous des photos. Les phénotypes observés sont notés à droite des photos.

Nous pouvons donc en conclure que le niveau de l'expression relative des transcrits du gène *PAP4* est affecté dans les transformants *amipap4-2* et *amipap4-37*. Il y a 63,5% d'expression du transcrit de *PAP4* pour *amipap4-2* et 49% pour *amipap4-37* par rapport à la souche contrôle pHUN. Puisque le mutant *amipap4-2* est moins affecté que *amipap4-37*, nous n'avons gardé que ce dernier pour les analyses qui suivront.

Comme pour les mutants CLiP, le niveau de la présence de la protéine a été vérifié par western blot avec les anticorps @PAP4, @PDH E1 et @VDAC1 (Figure R-25). Les analyses de western blot montrent une forte diminution de la protéine PAP4 pour le mutant *amipap4-37*, qui est donc un mutant KD, comme les mutants CLiP.

4. État des mitochondries dans les mutants de PAP4

a) Phénotypes des mutants

Afin de déterminer si les mitochondries des 3 lignées mutantes (*pap4-1::CIB*, *pap4-2::CIB* et *amipap4-37*) sont affectées, des tests phénotypiques ont été menés, en regardant la croissance des colonies transformées en conditions hétérotrophes (absence de lumière et présence d'acétate comme source de carbone). Dans ces conditions, les mutants mitochondriaux ont soit un phénotype *dark*⁻ s'ils ne poussent pas à l'obscurité, soit un phénotype *dark*^{+/-} s'ils ont un retard de croissance (Salinas *et al.*, 2013). Seul le mutant *amipap4-37* présente un retard de croissance à l'obscurité, il a donc un phénotype *dark*^{+/-} (Figure R-26). Les 2 mutants CLiP ne présentent pas de phénotype particulier à l'obscurité.

Leur croissance en absence de source de carbone (TP, milieu minimum sans acétate) et à la lumière a également été étudiée, afin de vérifier l'état des chloroplastes. En effet, normalement, la photosynthèse suffit à fournir le carbone nécessaire aux cellules. Or si le chloroplaste ou la photosynthèse sont impactés, cet apport n'est plus assuré : ce phénotype s'appellera *tp*⁻. Des travaux sur *C. reinhardtii* ayant montré l'impact de la respiration mitochondriale sur le transport photosynthétique des électrons (Cardol *et al.*, 2003), ce test est complémentaire à celui à l'obscurité. Aucun des 3 mutants ne présente de phénotype en milieu minimum, leur croissance se fait correctement (Figure R-26).

Pour conclure, les mutants *pap4-1::CIB* et *pap4-2::CIB* présentent un phénotype similaire aux souches sauvages. Seul le mutant *amipap4-37* présente un phénotype *dark*^{+/-}. Cela pose la question de savoir si le phénotype est lié à l'insertion aléatoire du vecteur amiRNA dans le

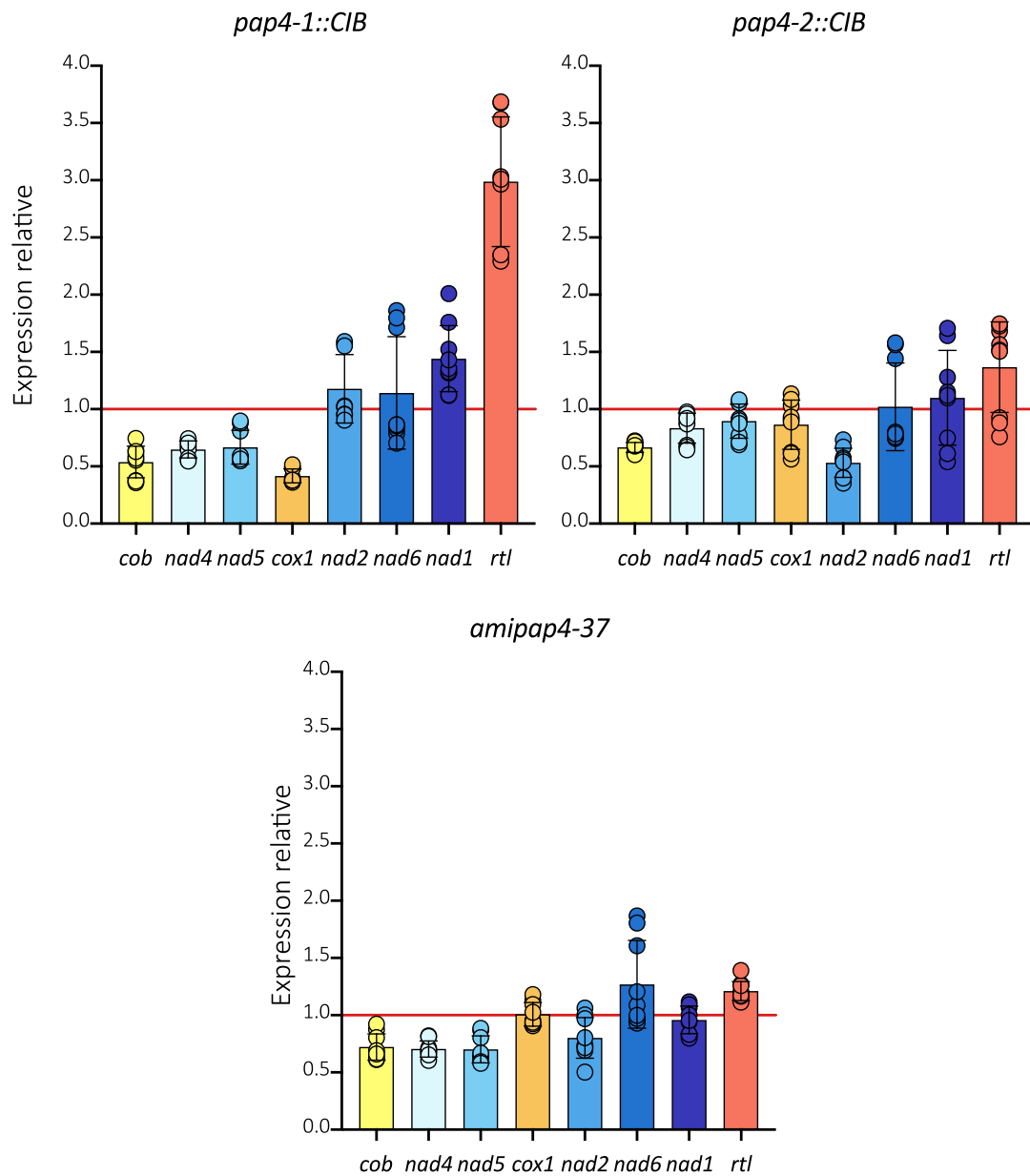


Figure R-27 : Analyses de l'expression des transcrits mitochondriaux des mutants de PAP4

Histogrammes de l'expression relative des 8 ARNm mitochondriaux, dans l'ordre de succession des gènes dans le génome, dans les mutants PAP4 ClIP et amiRNA, mesurée par RT-qPCR. Le trait **rouge** horizontal indique le niveau d'expression relative de la souche contrôle pHUN normalisé à 1. Les histogrammes présentent la moyenne de 3 réplicats biologiques (n = 3), avec l'écart-type (+/-, barre verticale). Pour chaque réplicat biologique, 3 réplicats techniques ont été faits. Toutes les valeurs obtenues sont représentées par les ronds sur les histogrammes.

génomique nucléaire ou s'il est dépendant de la souche, puisque la souche de référence n'est pas la même pour les 2 types de mutants (CC-5325 pour les CLiP, CC-4351 pour l'amiRNA).

b) Le niveau des transcrits mitochondriaux est affecté dans les mutants

Au niveau moléculaire, le taux des transcrits mitochondriaux a été analysé dans les trois lignées mutantes PAP4, par RT-qPCR, afin de voir si l'expression des gènes mitochondriaux est affectée dans ces mutants (Figure R-27).

Pour le mutant *pap4-1::CIB*, une diminution de l'expression des gènes *cob*, *nad4*, *nad5* et *cox1* a été observée, de 47, 35, 34 et 59% respectivement. Les transcrits des gènes *nad2* et *nad6* ne sont pas affectés, tandis que ceux des gènes *nad1* et *rtl* sont augmentés, avec respectivement une expression relative de 144 et 298% en moyenne.

Pour le mutant *pap4-2::CIB*, une diminution de l'expression des gènes *cob*, *nad4* et *nad2* a été observée, de 34, 17 et 47% respectivement. Les transcrits des gènes *nad5*, *cox1*, *nad6*, *nad1* et *rtl* ne sont statistiquement pas affectés, bien que les résultats divergent entre les 3 réplicats, notamment pour les gènes *nad6*, *nad1* et *rtl*.

Enfin, pour le mutant *amipap4-37*, une diminution de l'expression des gènes *cob*, *nad4*, *nad5* et *nad2* a été observée, de 28, 30, 30 et 20% respectivement. Les transcrits des gènes *cox1* et *nad6* ne sont pas affectés, bien que les résultats divergent entre les 3 réplicats pour le gène *nad6*. Enfin, Le taux à l'équilibre du gène *rtl* est augmenté, avec une expression relative moyenne de 121%.

Mises ensembles, ces analyses nous montrent que les ARNm n'ont pas les mêmes comportements selon les souches mutantes. Dans les mutants *pap4-1::CIB*, *pap4-2::CIB* et *amipap4-37*, le taux des ARNm correspondant à *nad6* est similaire à la souche sauvage, alors que ceux correspondant à *cob* et *nad4* sont diminués de 37% en moyenne. Les trois autres ARNm correspondant à *cox1*, *nad5* et *nad2* sont différemment affectés selon le mutant étudié. Par exemple, dans les lignées *pap4-1::CIB* et *amipap4-37*, les taux des ARNm correspondant à *nad5* sont diminués de 30% et ceux de *rtl* sont augmentés de 298% et 121% respectivement, alors que dans la souche *pap4-2::CIB*, ils sont similaires à la souche sauvage.

Les différences entre les différents mutants sont difficilement explicables. Cela pourrait être dû au fond génétique différent entre les mutants CLiP (CC-5325) et les mutants amiRNA (CC-4351). Nous ne pouvons pas exclure un possible effet d'insertion du vecteur dans le mutant amiRNA,

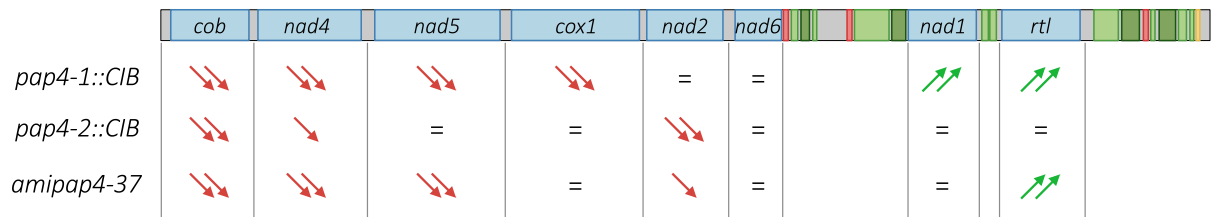


Figure R-28 : Effets de la mutation sur les ARNm mitochondriaux selon leur position

Schéma du génome mitochondrial de *C. reinhardtii* et résultats des analyses de RT-qPCR des transcrits du génome mitochondrial selon les mutants. Flèches : **rouges** : diminution par rapport au contrôle ; **vertes** : augmentation par rapport au contrôle ; = : similaire au contrôle, moins de 10% de différences avec le contrôle ; 1 flèche : entre 10% et 25% de différence avec le contrôle ; 2 flèches : au-delà de 25% de différence avec le contrôle. **Bleu** : gènes de protéines ; **Rouge** : gènes d'ARNt ; **Vert clair** : gènes d'ARNr de la grande sous-unité du mitoribosome ; **Vert foncé** : gènes d'ARNr de la petite sous-unité du mitoribosome ; **Jaune** : gène de l'ARN L2b.

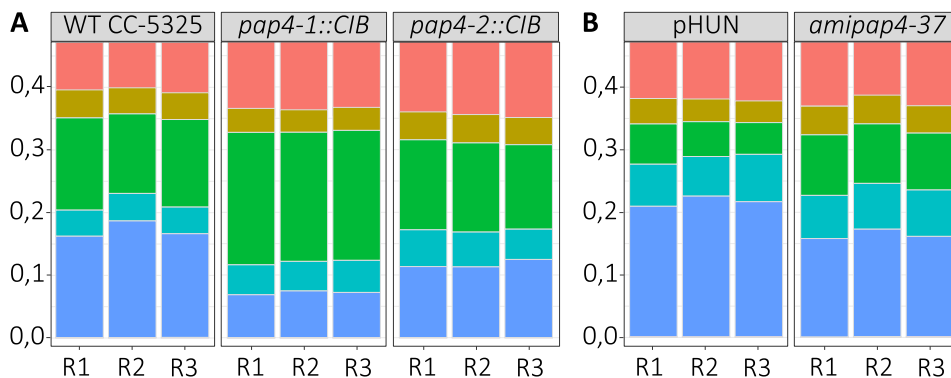


Figure R-29 : Les résultats des triplicats sont similaires pour les mutants de PAP4

Analyses des expériences de 3'RACE-seq des mutants CLiP (A) et amiRNA (B) de PAP4 ainsi que de leur souche de référence associée, en fonction de la catégorie des queues, pour les 3 réplicats biologiques. En abscisse sont indiquées les proportions des différentes catégories de queues, selon les souches et selon les 3 réplicats (R1, R2 et R3). Zoom des graphiques présentés en Annexe VIII-2.

Corail : pas de queues ; **Jaune** : queues riches en A, U et C ; **Vert** : queues riches en A ; **Bleu clair** : queues riches en U ; **Bleu foncé** : queues riches en C ; **Fuchsia** : autres.

puisque le site d'insertion dans le génome nucléaire n'est pas connu. Cependant, il y a aussi un point commun majeur à ces 3 souches : une diminution globale des taux des ARNm qui ne sont pas suivis par des ARNt ou des ARNr est observée (Figure R-28). Le niveau du transcrite de *nad6*, suivi d'un ARNt, n'est pas affecté, alors que l'expression des ARNm *nad1* et *rtl*, qui sont suivis d'un ARNr, est soit augmentée, soit non affectée. Nous savons que les clivages des ARNt se fait par la RNase P (Bonnard *et al.*, 2016), mais les enzymes de clivage ne sont pas connues pour les ARNm et les ARNr. Il se peut alors que l'activité de PAP4 soit influencée par le type de clivage et donc par l'environnement génique du gène ciblé.

5. Il y a moins de queues polyC dans les mutants de PAP4

Pour poursuivre la caractérisation des mutants et les investigations quant au rôle de PAP4 dans les mitochondries, des analyses de 3'RACE-seq ont été menées. Cela consiste en l'amplification des extrémités 3' des ADNc des 8 transcrits mitochondriaux puis en leur séquençage par la technologie Illumina, afin de voir l'état de la polycytidylation dans les mutants de PAP4.

Les expériences de 3'RACE-seq à partir d'ARN totaux ont été faites sur des triplicats biologiques pour chaque souche étudiée. Les analyses bio-informatiques ont été faites par l'ingénieur David Pflieger de la plateforme de bio-informatique de l'IBMP et la Docteure Thalia Salinas-Giegé (Méthodes - Figures M-5 et M-6). Puisque les analyses montrent des résultats similaires pour les 3 réplicats biologiques pour chaque souche (Figure R-29), les données des 3 réplicats ont été cumulées pour faire les analyses. Les détails des résultats obtenus par gène pour chaque réplicat de chaque souche est disponible en Annexe VIII-1. Cela nous a permis d'avoir bien plus de *reads* pour faire les analyses et donc d'avoir plus de profondeur dans nos résultats.

a) Analyses des mutants CLiP

Les Figures R-30, R-31 et R-32 présentent les résultats obtenus pour les 2 mutants CLiP, *pap4-1::CIB* et *pap4-2::CIB*.

Concernant la souche sauvage CC-5325, environ 60% des transcrits analysés n'ont pas de queue ajoutée à leur extrémité 3'. Cette souche présente des proportions similaires d'ajouts de queues polyA ou polyC-riches (entre 13,5 et 17%), mais très peu de polyuridylation, avec seulement 4,5% des transcrits (Figure R-30A). L'analyse individuelle des catégories de queues ajoutées pour chaque transcrite mitochondrial montre des différences entre les ARNm. Les transcrits des gènes *cob* et *cox1* sont les plus cytidylés (27% et 26% respectivement), tandis que

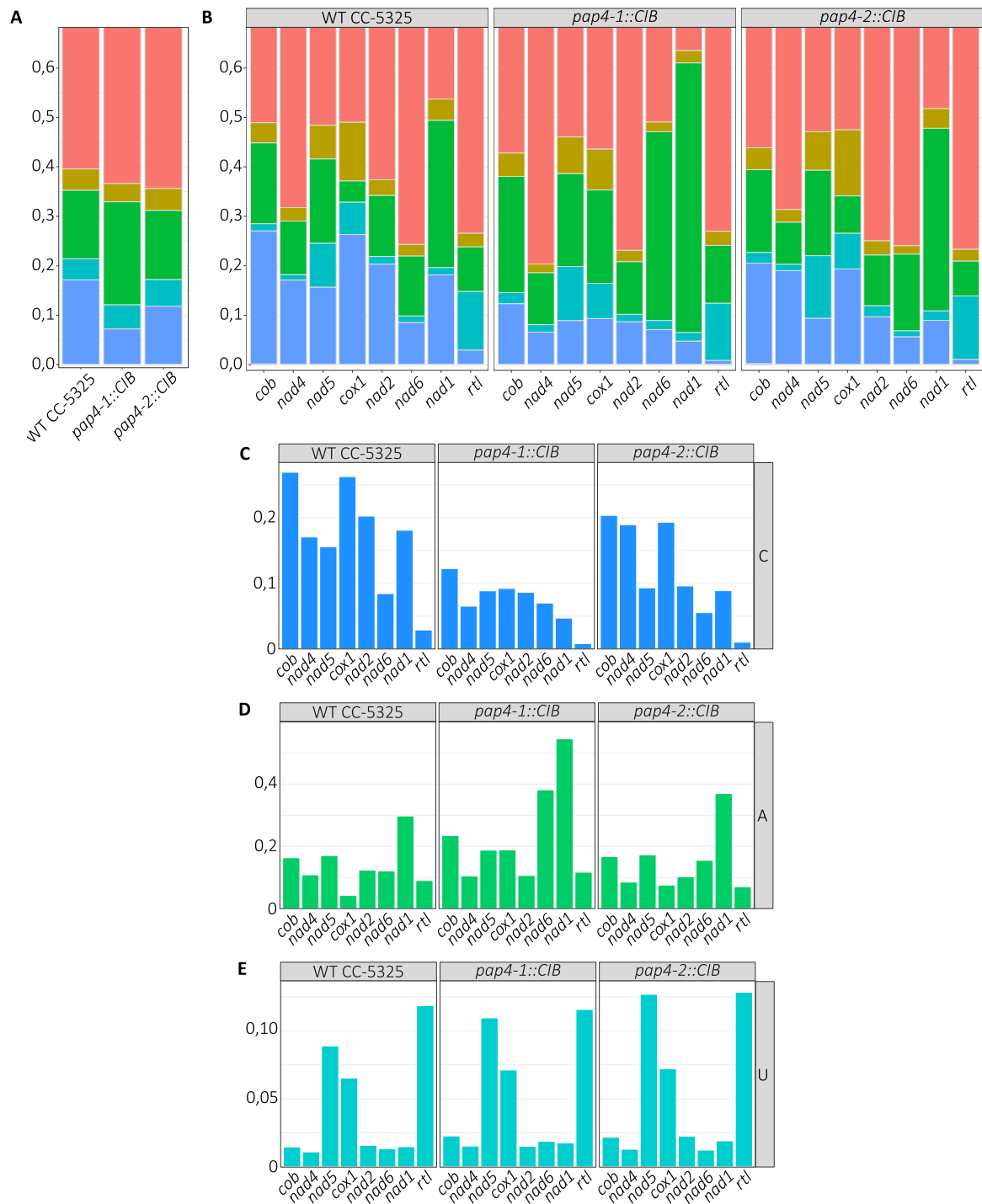


Figure R-30 : Comparaison des 3 catégories de queues ajoutées entre les mutants CLiP de PAP4 et la souche WT CC-5325

Analyses des expériences de 3'RACE-seq des mutants CLiP de PAP4 en fonction de la catégorie des queues. Résultats cumulés des 3 réplicats biologiques. **A** et **B** : En abscisse sont indiquées les proportions des différentes catégories de queues, selon les souches (**A**) et selon les gènes pour chaque souche (**B**). Zoom des graphiques présentés en [Annexe VIII-2](#). **C**, **D** et **E** : En abscisse sont indiquées les proportions selon les gènes, pour les queues polyC-riches (**C**), polyA-riches (**D**) ou polyU-riches (**E**). **Corail** : pas de queues ; **Jaune** : queues riches en A, U et C ; **Vert** : queues riches en A ; **Bleu clair** : queues riches en U ; **Bleu foncé** : queues riches en C.

nad6 et *rtl* ne le sont que très peu (8,5% et 3% respectivement) (Figure R-30B et C). Le processus d'adénylation est assez similaire pour les différents transcrits, avec entre 15 et 20%, à part pour *cox1* qui a le moins de queues polyA (4%) et pour *nad1* qui en a le plus, avec 30% (Figure R-30B et D). De manière générale, il y a beaucoup moins de queues polyU-riches (Figure R-30B et E). Il n'y a que 3 transcrits qui sont fortement uridylés, *nad5*, *cox1* et *rtl* (entre 7 et 12%), tandis que les 5 autres ont un taux d'ajout de queues polyU similaire, entre 1 et 2%. Ces résultats diffèrent de ce qui avait été trouvé dans l'étude de 2017 (Salinas-Giegé *et al.*, 2017) (Annexe IX). En 2017, la distinction n'avait été faite qu'entre les queues polyA/U-riches et les queues polyC-riches. Dans cette nouvelle étude, nous avons scindé les queues polyA/U-riches en 2 catégories : les queues polyA-riches et les polyU-riches. Cela dit, en additionnant nos résultats pour les A et les U, nous trouvons des profils assez similaires, bien qu'avec des proportions plus importantes, pour la plupart des transcrits. Certains transcrits ont plus de différences entre les 2 études. Par exemple, dans l'étude de 2017, les transcrits du gène *nad1* avaient environ 10% de queues polyA/U-riches, alors que nous en trouvons presque 30%. Pour les queues polyC-riches, l'étude de 2017 montrait une polycytidylation équivalente pour les transcrits *cob*, *nad4*, *nad5*, *nad6* et *nad1*, avec une moyenne de 25%. L'ARNm de *cox1* était le plus polycytidylé, avec 45%, et ceux de *nad2* et *rtl* les moins polycytidylés, avec moins de 10%. Or notre étude montre des profils différents, avec des valeurs plus basses en moyenne. Ces différences peuvent être dues aux souches utilisées mais également à la méthode employée (CC-5325 et 3'RACE-seq pour nous, CC-4351 et 3'RACE puis clonage et séquençage Sanger pour l'étude de 2017), puisque nos analyses se basent sur un nombre bien plus important de séquences analysées (millions de *reads* pour nous, centaines de séquences pour l'étude de 2017).

Les différences entre les mutants et la souche de référence sont calculées en pourcentages de différences entre les 2 souches comparées. Ces valeurs sont disponibles en Annexe X. L'analyse des résultats globaux montre une augmentation de 56% des ajouts de queues polyA dans le mutant *pap4-1::CIB* par rapport à la souche WT CC-5325, tandis que ces proportions sont similaires à la souche de référence pour le mutant *pap4-2::CIB* (Figure R-30A). Dans les 2 mutants, les queues polyU-riches sont augmentées de 11% par rapport à la souche sauvage, tandis qu'une diminution de 59 et 29% est constatée concernant l'addition des queues polyC-riches entre les mutants *pap4-1::CIB* et *pap4-2::CIB*, respectivement, et la souche WT.

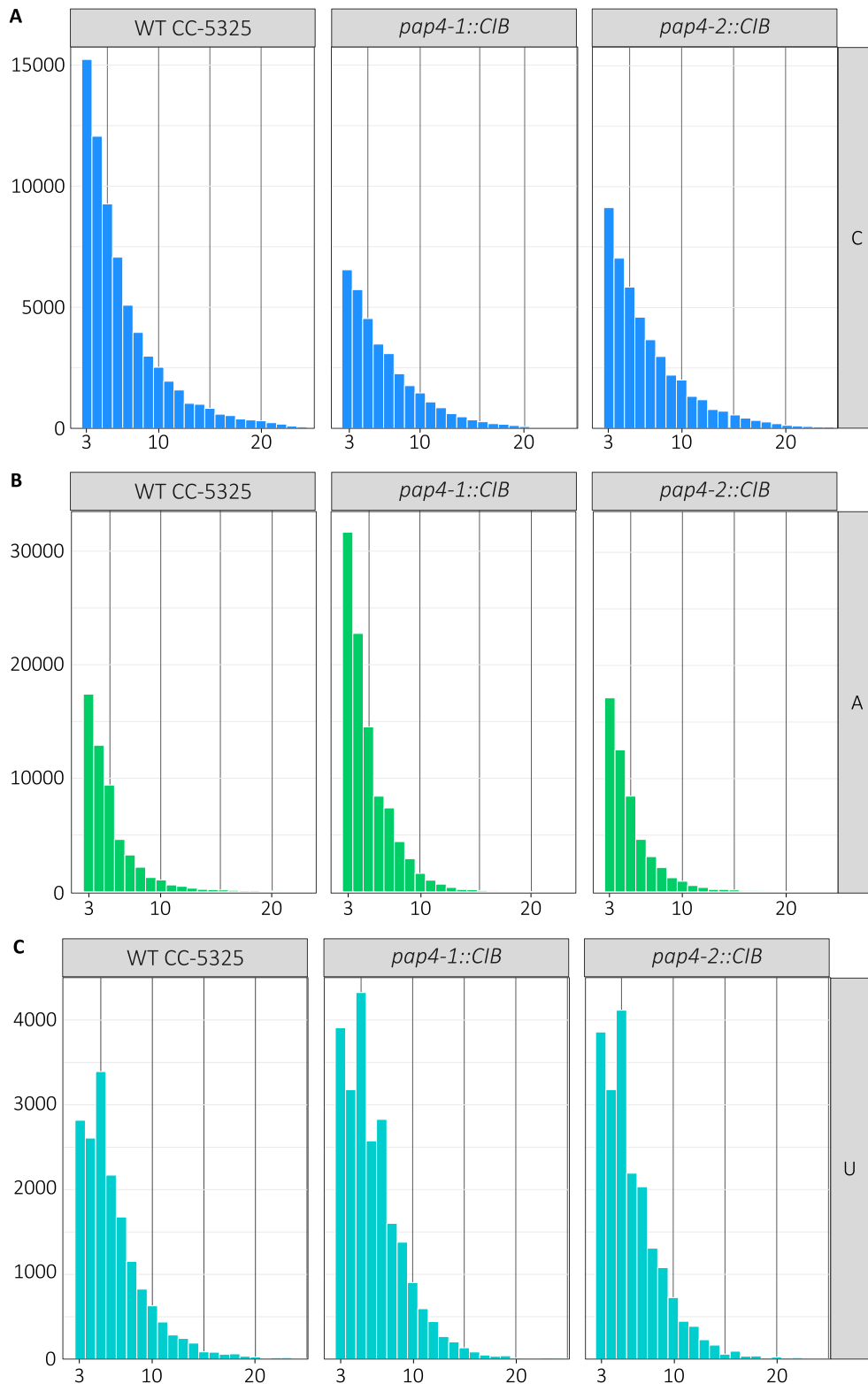


Figure R-31 : Le nombre de queues ajoutées selon leur longueur

Analyses des expériences de 3'RACE-seq des mutants CLiP de PAP4 en fonction de la catégorie des queues. Résultats cumulés des 3 réplicats biologiques. **A**, **B** et **C** : histogrammes du nombre de queues polyC-riches (**A**), polyA-riches (**B**) ou polyU-riches (**C**) ajoutées selon leur longueur, pour chaque souche. **Bleu foncé** : queues riches en C ; **Vert** : queues riches en A ; **Bleu clair** : queues riches en U.

L'analyse individuelle des catégories de queues ajoutées pour chaque transcrite mitochondrial montre une diminution de la polycytidylation entre 18 et 75%, pour tous les transcrits des 2 mutants par rapport au WT CC-5325, à l'exception de l'ARNm de *nad4* dans la lignée *pap4-2::CIB* pour lequel une augmentation de 12% est constatée (Figure R-30B et C). Le processus d'adénylation diffère également pour certains transcrits dans les mutants par rapport à la souche CC-5325 (Figure R-30B et D). Dans la souche *pap4-1::CIB*, l'adénylation des transcrits *nad4*, *nad5* et *nad2* reste similaire à la souche WT. Il en est de même pour les transcrits *cob* et *nad5* du mutant *pap4-2::CIB*. Par contre, les transcrits *cob*, *cox1*, *nad6*, *nad1* et *rtl* de la souche *pap4-1::CIB* voient leurs queues polyA augmentées, entre 28 et 363% de plus. Idem pour les transcrits *cox1*, *nad6* et *nad1* de la souche *pap4-2::CIB*, bien que l'augmentation par rapport aux données de la souche WT CC-5325 soit plus modérée, variant entre 23 et 75%. Enfin, les ARNm *nad4*, *nad2* et *rtl* du mutant *pap4-2::CIB* ont une diminution des proportions de queues polyA-riche ajoutées par rapport à la souche de référence, avec respectivement 18, 19 et 22% de moins. Enfin, pour certains transcrits, il n'y a pas de changement des taux d'uridylation, comme pour les ARNm *cox1* et *rtl* des 2 mutants, *nad6* de la souche *pap4-1::CIB* ainsi que *nad2* et *nad1* du mutant *pap4-2::CIB* (Figure R-30B et E). Cependant, des augmentations de la polyuridylation sont constatées pour les transcrits *cob*, *nad4*, *nad5* dans les 2 souches mutantes, ainsi que pour les ARNm *nad1* du mutant *pap4-1::CIB*, variant entre 22 et 67% de plus que pour la souche de référence CC-5325. Pour finir, les transcrits *nad2* et *nad1*, des souches *pap4-1::CIB* et *pap4-2::CIB* respectivement, sont diminués de 50 et 25%.

La Figure R-31 présente le nombre de queues en fonction de leur longueur. Pour les queues polyC-riche, malgré la diminution globale visible dans les 2 mutants, aucune variation n'est observée dans les profils du nombre de queues ajoutées en fonction de leur longueur (Figure R-31A). Concernant les queues polyA-riche, les données pour le mutant *pap4-2::CIB* est très semblables à celles de la souche WT CC-5325 (Figure R-31B). Cependant, bien qu'une augmentation globale soit observée pour le mutant *pap4-1::CIB*, il n'y a pas de variation dans les profils du nombre de queues ajoutées en fonction de leur longueur. Enfin, concernant la polyuridylation, les profils du nombre de queues ajoutés en fonction de leur longueur sont similaires entre les 2 mutants et la souche de référence, avec une légère augmentation globale pour les 2 mutants (Figure R-31B).

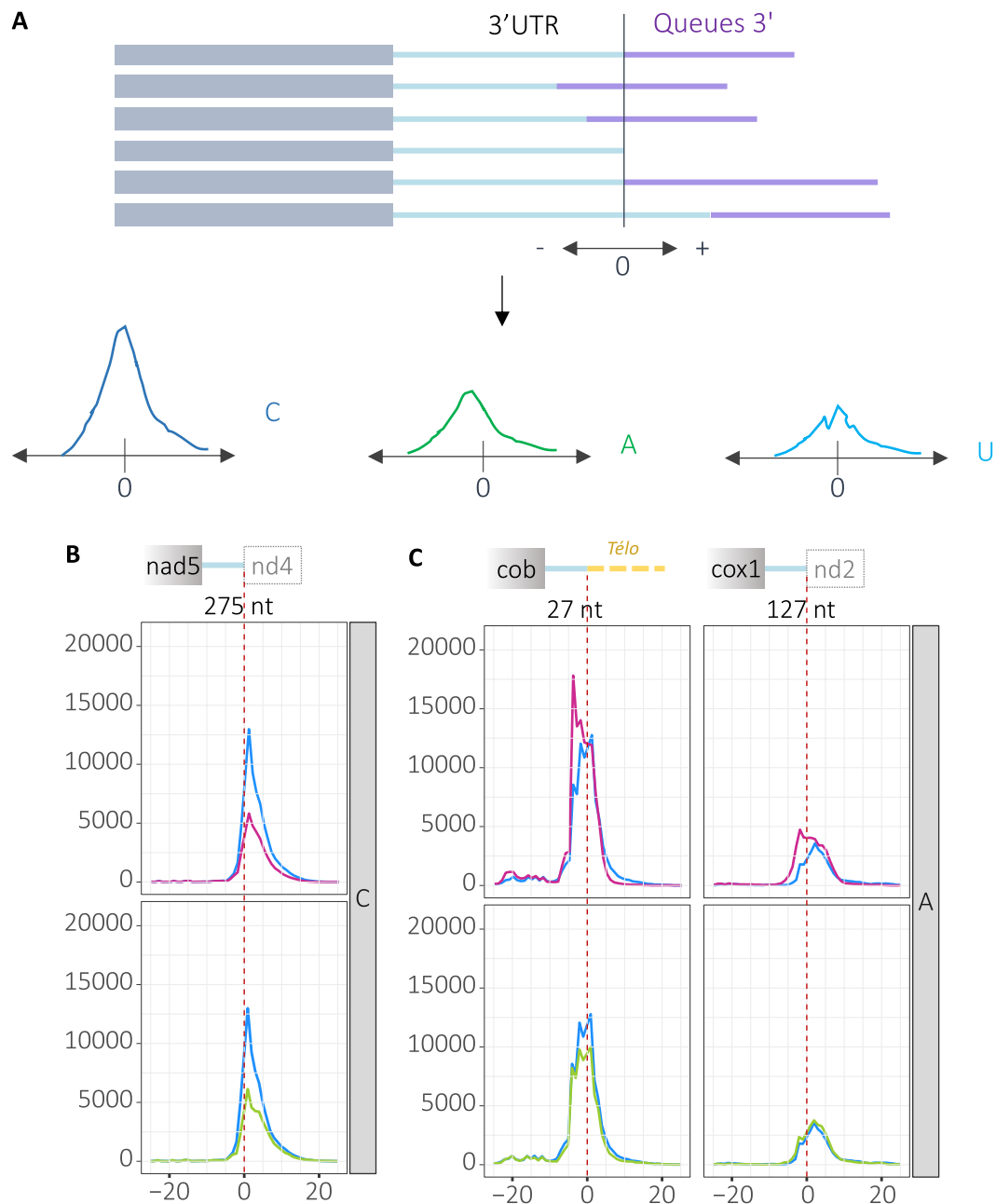


Figure R-32 : Distribution des nt différents dans les mutants, pour certains transcrits

Analyses des expériences de 3'RACE-seq des mutants amiRNA de PAP6. Résultats cumulés des 3 réplicats biologiques. **A** : Schéma de l'obtention des graphiques présentés en **B** et **C**. Tous les *reads*, sans aucun critère de longueur ou de composition, sont alignés. Puis le nombre de chaque nt ajouté à chaque position en amont et en aval de la fin de la 3'UTR (= position 0) est compté, afin d'obtenir des graphiques de la distribution de chaque nt en fonction de la position. **B** et **C** : graphiques du nombre de nt C (**B**) ou A (**C**) trouvés dans les queues, en fonction de leur position vis-à-vis de la fin de la 3'UTR (position 0, trait rouge pointillé) pour certains transcrits mitochondriaux. Au-dessus des graphiques sont schématisés la séquence codante du transcrit (gris), suivie de sa séquence 3'UTR (bleu, avec la taille indiquée en dessous en nt) et l'environnement de ce transcrit avec le transcrit suivant (gris : ARNm et jaune : télomère). Bleu : WT CC-5325 ; Rose : *pap4-1::CIB* ; Vert : *pap4-2::CIB*.

Il y a peut-être un biais de la méthode employée. En effet, nous voyons un grand nombre de queues entre 3 et 7 nt, pour les 3 catégories de queues observées. Or dans l'article de 2017, une moyenne de 14 nt avait été montrée pour les queues polyC-riches (Salinas-Giegé *et al.*, 2017). Lors de nos analyses de 3'RACE-seq, nous avons également regardé des spike-in, c'est-à-dire des séquences d'une longueur fixe en C : 8C, 16C et 22C. Or, avec les 16C et 22C, nous avons obtenu moins de *reads* exploitables que pour le 8C. Il est donc possible que le séquenceur ne puisse lire correctement les longues séquences homopolymériques. Nous n'avons pas testé avec des séquences homopolymériques en A ou U, mais la taille des queues étant similaire à celle des C, il y a peut-être le même problème. De fait, nous ne pouvons pas exclure que les queues ajoutées post-transcriptionnellement aux ARNm mitochondriaux soient plus longues que ce que nous avons observé. Néanmoins, même s'il y a un biais de séquençage, des différences entre notre souche de référence et nos mutants sont observables.

Enfin, nous avons regardé la distribution des 3 types de nt ajoutés, en fonction de la position, par rapport à la fin de la séquence 3'UTR de chaque transcrit (Figure R-32A). Pour cela, nous avons analysé toutes les séquences obtenues, sans appliquer de filtres de taille ou de composition. La position 0 est la fin de la 3'UTR de chaque transcrit. Là, chaque nucléotide ajouté a été compté, afin d'obtenir le nombre de chaque nt ajouté à chaque position, après la 3'UTR, mais également avant, afin de prendre en compte les ARN tronqués. Les résultats pour chaque nt et chaque transcrit sont disponibles en Annexe XI, ne seront présentés ici que les résultats présentant une différence entre le WT et la souche mutante.

Concernant l'ajout de C, nous constatons des différences de profils entre les 2 mutants et la souche de référence pHUN pour le transcrits *nad5* (Figure R-32B). En effet, un changement de la position correspondant à l'ajout majoritaire de C est observé. Pour la souche de référence, il y a une majorité d'ajouts de C à la position 1 pour *nad5*, tandis que la fréquence d'ajout de C est majoritaire entre les positions 1 et 5 dans les 2 mutants. Pour l'ajout de A, il y a des différences entre le mutant *pap4-1::CIB* et la souche WT CC-5325, pour les transcrits *cob* et *cox1* (Figure R-32C). Dans la souche de référence, il y a une majorité d'ajout de A entre les positions -3 et 2, alors que pour le mutant, cela se fait entre les positions -5 et -2, plus précocement. De même pour l'ARNm *cob*, dans la souche sauvage, le maximum d'ajout de A est observé aux positions 1 à 3, tandis que pour le mutant *pap4-1::CIB*, ce maximum s'étend des positions -3 à +3. Concernant les mutant *pap4-2::CIB*, aucune différence de profil n'est

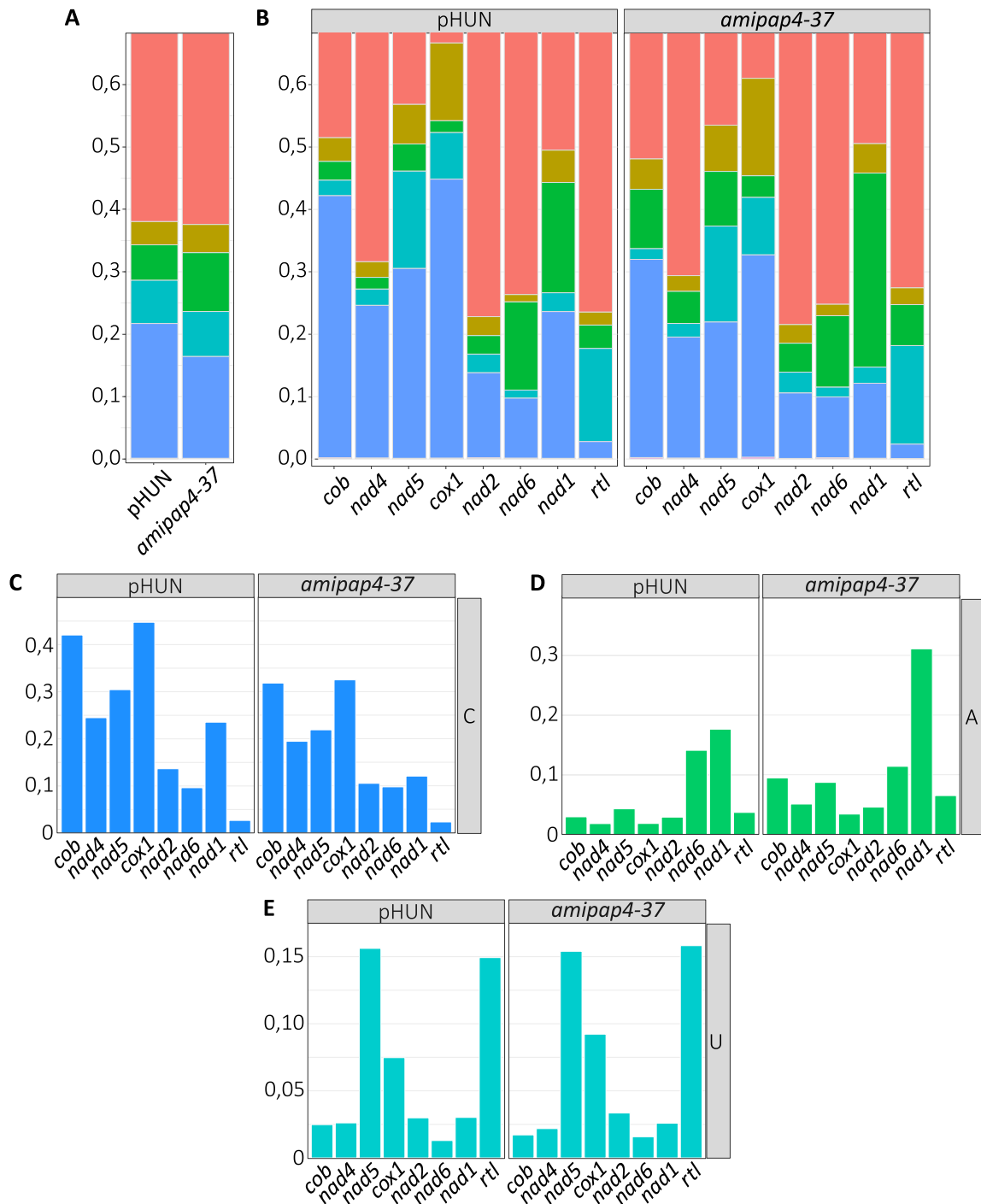


Figure R-33 : Comparaison des 3 catégories de queues ajoutées entre le mutant amiRNA de PAP4 et la souche contrôle pHUN

Analyses des expériences de 3'RACE-seq des mutants amiRNA pour PAP4 en fonction de la catégorie des queues. Résultats cumulés des 3 réplicats biologiques. **A** et **B** : En abscisse sont indiquées les proportions des différentes catégories de queues, selon les souches (**A**) et selon les gènes pour chaque souche (**B**). Zoom des graphiques présentés en [Annexe VIII-2](#). **C**, **D** et **E** : En abscisse sont indiquées les proportions selon les gènes, pour les queues polyC-riches (**C**), polyA-riches (**D**) ou polyU-riches (**E**). **Corail** : pas de queues ; **Jaune** : queues riches en A, U et C ; **Vert** : queues riches en A ; **Bleu clair** : queues riches en U ; **Bleu foncé** : queues riches en C.

observée pour ces 2 transcrits. Enfin, les profils d'ajouts de U en fonction de la position sont strictement similaires dans la souche sauvage CC-5325 et dans les 2 souches mutantes CLiP. Tous ensemble, ces résultats montrent une diminution globale de la polycytidylation dans les mutants de PAP4, ainsi qu'une augmentation de la polyadénylation et de la polyuridylation qui diffère selon les transcrits. Ces effets sont moindres dans la souche *pap4-2::CIB* par rapport à la souche *pap4-1::CIB*. Ils varient également selon les gènes, certains sont plus affectés par la mutation de PAP4 que d'autres.

b) Analyses du mutant *amiRNA*

Les résultats obtenus pour le mutant *amiRNA amipap4-37* sont présentés en [Figures R-33, R-34](#) et [R-35](#).

Concernant la souche de référence pHUN, environ 60% des transcrits analysés n'ont pas de queue ajoutée à leur extrémité 3'. Nous constatons que le processus de polycytidylation est majoritaire, avec 22% des transcrits et qu'il y a des proportions plus faibles mais similaires d'ajouts de queues polyA et polyU-riches (entre 6,5 et 7%) ([Figure R-33A](#)).

L'analyse individuelle des catégories de queues ajoutées pour chaque transcrit mitochondrial montre des différences entre les ARNm. Les transcrits des gènes *cob* et *cox1* sont les plus cytidylés (42% et 45% respectivement), tandis que *rtl* ne l'est que très peu (3%) ([Figure R-33B](#) et [C](#)). Pour les 5 autres transcrits, les proportions d'ajout de séquences riches en C varient entre 10 et 30%. Le processus d'adénylation est assez similaire pour les différents transcrits, avec entre 2 et 4,5% de transcrits adénylés, à l'exception de *nad6* et *nad1* qui ont des taux plus élevés (14 et 17% respectivement) ([Figure R-33B](#) et [D](#)). Il n'y a que 3 transcrits qui sont fortement uridylés, *nad5*, *cox1* et *rtl* (entre 7 et 16%), tandis que les 5 autres ont un taux d'ajout de queues polyU similaire, entre 1 et 3% ([Figure R-33B](#) et [E](#)). Ces résultats sont proches de ceux qui avait été trouvés dans l'étude de 2017 (Salinas-Giegé *et al.*, 2017) ([Annexe IX](#)). Pour les queues polyC-riches, les 2 études montrent l'ARNm de *cox1* comme le plus polycytidylé, avec plus de 40% et celui de *rtl* comme le moins polycytidylé, avec moins de 5%. Pour les autres transcrits, il y a des différences de profil, mais les valeurs restent du même ordre de grandeur. En additionnant nos résultats pour les A et les U, nous trouvons des profils assez similaires, bien qu'avec des proportions plus importantes, pour la plupart des transcrits. Puisque les souches analysées dans les 2 études ont le même fond génétique (CC-4153), il est normal d'avoir des résultats plus cohérents. Il se peut donc que les différences vues précédemment entre les

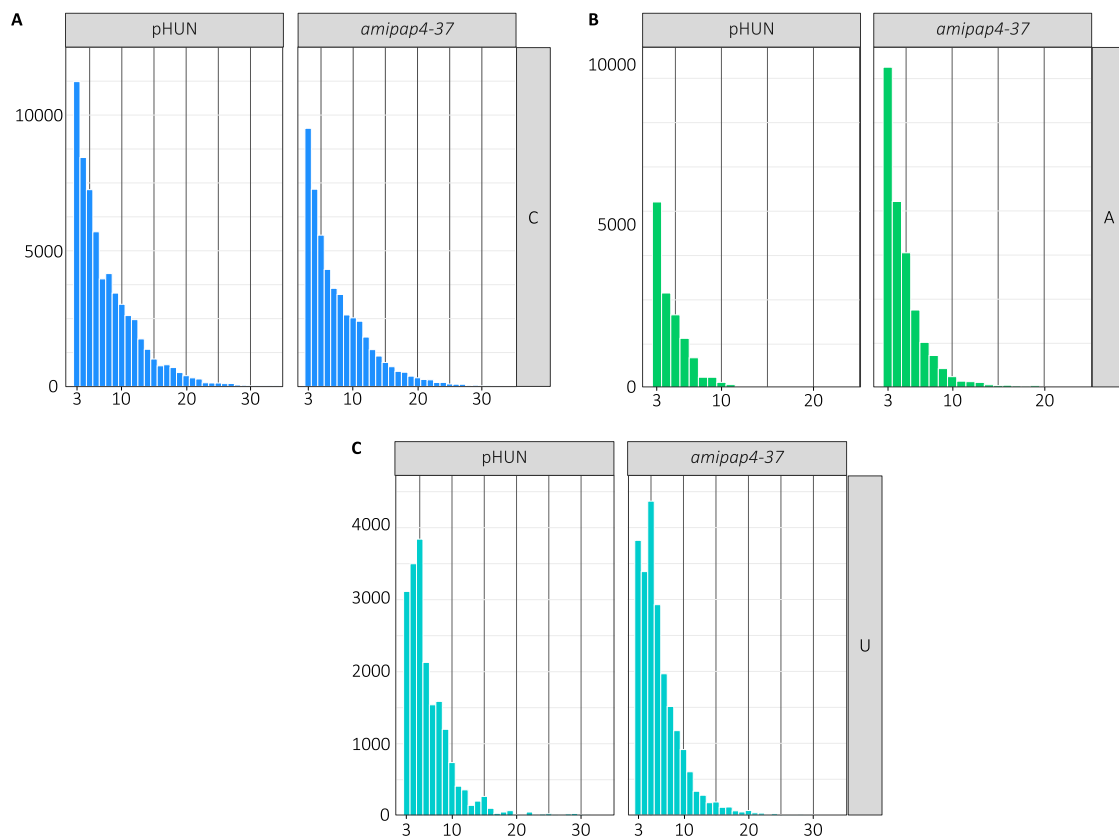


Figure R-34 : Le nombre de queues ajoutées selon leur longueur

Analyses des expériences de 3'RACE-seq du mutant amiRNA de PAP4 en fonction de la catégorie des queues. Résultats cumulés des 3 réplicats biologiques. **A**, **B** et **C** : histogrammes du nombre de queues polyC-riches (**A**), polyA-riches (**B**) ou polyU-riches (**C**) ajoutées selon leur longueur, pour chaque souche. **Bleu foncé** : queues riches en C ; **Vert** : queues riches en A ; **Bleu clair** : queues riches en U.

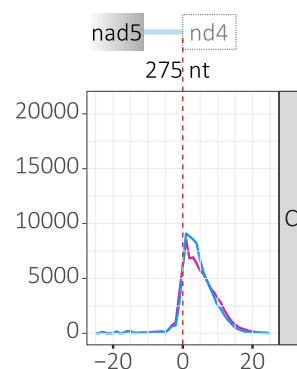


Figure R-35 : Distribution différente des nt dans le mutant amiRNA de PAP4, pour l'ARNm *nad5*

Analyses des expériences de 3'RACE-seq des mutants amiRNA de PAP6. Résultats cumulés des 3 réplicats biologiques. Graphiques du nombre de nt C trouvés dans les queues du transcript *nad5*, en fonction de leur position vis-à-vis de la fin de la 3'UTR (position 0, trait **rouge** pointillé) Au-dessus du graphique est schématisée la séquence codante du transcript (gris), suivie de sa séquence 3'UTR (bleu, avec la taille indiquée en dessous en nt) et l'environnement de ce transcript avec le transcript suivant (gris : ARNm) **Bleu** : contrôle pHUN ; **Rose** : *amipap4-37*.

souches CC-5325 et CC-4153 montrent une réalité biologique, avec des ajouts de queues variables selon les souches. Les différences entre les mutants et la souche de référence sont calculées en pourcentages de différence entre les 2 souches comparées. Ces valeurs sont disponibles en [Annexe X](#).

L'analyse des résultats globaux ([Figure R-33A](#)) ne montre aucune variation des queues polyU-riches par rapport à la souche pHUN. Par contre, nous pouvons observer une forte diminution de 25% des queues polyC-riches dans le mutant *amipapa4-37* par rapport à la souche de référence pHUN. A l'opposé, les queues polyA-riches subissent une augmentation de 38% dans le mutant.

L'analyse de l'ajout de queues polyC-riches pour chaque transcrite mitochondrial montre une réduction globale entre la lignée mutante et la souche de référence, variant entre 20 et 50% pour tous les transcrits, sauf *nad6* qui reste similaire à la souche contrôle pHUN ([Figures R-33B et C](#)). A l'inverse, les profils de polyadénylation du mutant sont augmentés entre 63 et 217% dans le mutant, en comparaison de la souche contrôle pHUN, à l'exception encore de *nad6*, qui a une diminution des ajouts de polyA de 21% ([Figures R-33B et D](#)). Enfin, l'ajout des queues polyU-riches n'est pas affecté pour les transcrits *nad5*, *nad2* et *rtl* de la lignée mutante *amipap4-37*, par rapport à la souche contrôle pHUN ([Figures R-33B et E](#)). En revanche, les proportions d'ajouts de queues polyU diminuent pour les transcrits *cob*, *nad4* et *nad1* dans le mutant, entre 20 et 33% de réduction. Enfin, les taux de polyuridylation sont augmentés de 29 et 100% pour les transcrits *cox1* et *nad6* respectivement dans la souche *amipap4-37*, en comparaison avec la souche pHUN.

La [Figure R-34](#) présente le nombre de queues en fonction de leur longueur. Pour les queues polyC-riches, une diminution du nombre de petites queues est constatée dans le mutant *amipap4-37*, avec moins de queues de 3 à 6 nt, après quoi, le profil est similaire à celui de la souche contrôle pHUN ([Figure R-34A](#)). Pour les queues polyA et polyU-riches, des profils similaires sont observés entre le mutant et la souche de référence, bien qu'il y ait une augmentation globale du nombre de queues ajoutées ([Figure R-34B et C](#)).

Enfin, nous avons regardé la distribution des 3 types de nt ajoutés, en fonction de la position, par rapport à la fin de la séquence 3'UTR de chaque transcrite ([Figure R-35](#)). Pour ces analyses, toutes les queues identifiées ont été prises en compte, sans appliquer de filtre de taille ou de composition. Comme précédemment, les résultats complets sont disponibles en [Annexe XI](#),

| | | <i>cob</i> | <i>nad4</i> | <i>nad5</i> | <i>cox1</i> | <i>nad2</i> | <i>nad6</i> | | <i>nad1</i> | <i>rtl</i> | |
|--------------------|---------------------|------------|-------------|-------------|-------------|-------------|-------------|--|-------------|------------|--|
| <i>pap4-1::CIB</i> | Expression relative | ↘↘ | ↘↘ | ↘↘ | ↘↘ | = | = | | ↗↗ | ↗↗ | |
| | Queues polyC-riches | ↘↘ | ↘↘ | ↘↘ | ↘↘ | ↘↘ | ↘ | | ↘↘ | ↘↘ | |
| | Queues polyA-riches | ↗↗ | = | = | ↗↗ | = | ↗↗ | | ↗↗ | ↗↗ | |
| | Queues polyU-riches | ↗↗ | ↗↗ | ↗↗ | = | ↘↘ | = | | ↗↗ | = | |
| | | | | | | | | | | | |
| <i>pap4-2::CIB</i> | Expression relative | ↘↘ | ↘ | = | = | ↘↘ | = | | = | = | |
| | Queues polyC-riches | ↘↘ | ↗ | ↘↘ | ↘↘ | ↘↘ | ↘↘ | | ↘↘ | = | |
| | Queues polyA-riches | = | ↘ | = | ↗↗ | ↘ | ↗↗ | | ↗↗ | ↘↘ | |
| | Queues polyU-riches | ↗↗ | ↗↗ | ↗↗ | = | = | ↘↘ | | = | = | |
| | | | | | | | | | | | |
| <i>amipap4-37</i> | Expression relative | ↘↘ | ↘↘ | ↘ | = | ↘ | = | | = | ↗ | |
| | Queues polyC-riches | ↘↘ | ↘ | ↘↘ | ↘↘ | ↘↘ | = | | ↘↘ | ↘↘ | |
| | Queues polyA-riches | ↗↗ | ↗↗ | ↗↗ | ↗↗ | ↗↗ | ↘ | | ↗↗ | ↗↗ | |
| | Queues polyU-riches | ↘↘ | ↘ | = | ↗↗ | = | ↗↗ | | ↘↘ | = | |
| | | | | | | | | | | | |

Figure R-36 : Effets de la mutation de PAP4 sur les ARNm mitochondriaux

Schéma du génome mitochondrial de *C. reinhardtii* et résultats des analyses de RT-qPCR (« Expression relative ») et de 3'RACE-Seq (« Queues polyA, U ou C-riches ») sur les transcrits du génome mitochondrial selon les mutants, en comparaison avec les contrôles.

Flèches : **rouges** : diminution par rapport au contrôle ; **vertes** : augmentation par rapport au contrôle ; = : similaire au contrôle ; 2 flèches : à partir de 20% de différence avec le contrôle pour l'expression relative et à partir de 5% de différence pour les queues ; 1 flèche : en dessous de 20% de différence significative selon les tests statistiques (Figure R-27) avec le contrôle pour l'expression relative et entre 2 et 5% de différence avec le contrôle pour les queues.

Bleu : gènes de protéines ; **Rouge** : gènes d'ARNt ; **Vert clair** : gènes d'ARNr de la grande sous-unité du mitoribosome ; **Vert foncé** : gènes d'ARNr de la petite sous-unité du mitoribosome ; **Jaune** : gène de l'ARN L2b.

seuls les profils différents entre la souche de référence et le mutant sont montrés, par souci de lisibilité. Seul le transcrite *nad5* présente des différences entre le mutant *amipap4-37* et la souche de référence pHUN. En effet, dans la souche contrôle, le maximum d'ajout de C se fait entre les positions 1 et 4, tandis que cela est majoritaire à la position 1 uniquement, avec un 2^{ème} pic plus petit à la position 3. Cela révèle un petit décalage de la fréquence d'ajout des C pour ce transcrite. Pour les autres transcrits, les profils sont strictement similaires entre la souche mutée et la souche de référence.

Ensemble, ces résultats montrent, comme pour les mutants CLiP, une diminution globale de la polycytidylation dans le mutant amiRNA de PAP4, sauf pour *nad6*. Il y a également une augmentation de la polyadénylation et des effets plus variables quant à la polyuridylation.

c) Conclusions des expériences de 3'RACE-Seq

Lorsque les 2 types de mutants sont comparés, nous voyons que la souche *pap4-1::CIB* est la plus affectée, en termes d'expression relative des transcrits mitochondriaux comme pour les niveaux de polynucléotides ajoutés en 3', suivie de *amipap4-37* puis *pap4-2::CIB*, avec des conséquences plus légères de la mutation de la protéine PAP4 (Figure R-36). Cela est en cohérence avec les résultats des RT-qPCR et des Western blot, qui montraient que le mutant *pap4-1::CIB* a l'expression de PAP4 la plus diminuée, tandis que le mutant *pap4-2::CIB* semble le moins affecté. La souche *amipap4-37* présente une expression intermédiaire, entre les 2 souches CLiP. Néanmoins, dans tous les cas, une diminution globale de la polycytidylation est observée dans les mutants de PAP4, ainsi qu'une augmentation de la polyadénylation. Les changements de la polyuridylation sont plus variables dans les mutants, selon les transcrits analysés.

Les diverses analyses montrent que le transcrite *nad6* ne semble moins affecté par la mutation de PAP4, contrairement aux 7 autres, tant au niveau de son expression relative que dans l'analyse des queues ajoutées après sa 3'UTR. Cela pourrait être lié à l'environnement du gène, puisque c'est le seul à être suivi par un ARNt.

Les gènes *nad1* et *rtl* sont eux suivis d'une courte séquence intergénique puis d'un gène d'ARNr. Pour ces 2 transcrits, il y a une augmentation de la proportion des queues polyA-riches ajoutées et les ajouts de queues polyU-riches sont inchangées pour *rtl* par rapport aux souches contrôle. Concernant *nad1*, les résultats sont variables, avec une augmentation dans le mutant

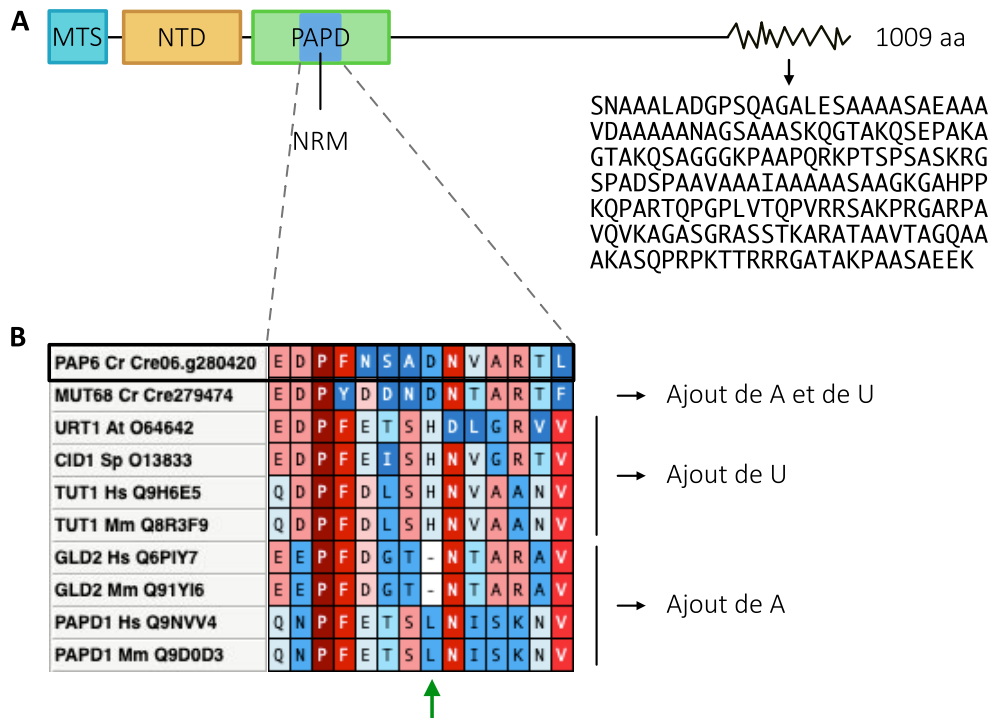


Figure R-37 : Domaines de la protéine PAP6

A : Représentation schématique des domaines de la protéine. La protéine PAP6 possède les 2 domaines caractéristiques des NT, les domaines NTD (jaune) et PAPD (vert). La séquence d'adressage mitochondriale putative (MTS – bleu clair) ainsi que le motif de reconnaissance des nucléotides (NRM – bleu foncé) dans le domaine PAPD sont schématisés. La région en zig-zag correspond à la région riche en A (alanine), dont la séquence est indiquée en dessous.

B : Alignement (Muscle) de la séquence protéique des motifs NRM de PAP6 (encadrée en noir) avec 9 NT de classe I de différents organismes indiqués à gauche de l'alignement de la façon suivante : nom de la protéine, nom de l'espèce et numéro du gène ou d'accèsion UniProt ; Cr : *Chlamydomonas reinhardtii*, At : *Arabidopsis thaliana*, Hs : *Homo sapiens*, Mm : *Mus musculus*, et Sp : *Schizosaccharomyces pombe*. A droite, par une flèche noire, sont indiquées les sélectivités des protéines : MUT68 : ajoute des U et des A *in vitro* (Ibrahim *et al.*, 2010) ; URT1, CID1 et TUT1 : TUTases, ajoutent des U ; GLD2 et PAPD1 : PAP non canoniques, ajoutent des A. Flèche verte : position du résidu de la sélectivité des nucléotides : résidu histidine responsable de la sélectivité pour les uridines ; résidu leucine ou absence de résidu responsable de la sélectivité pour les adénines ; alignement : du rouge au bleu : de la plus forte à la plus faible identité entre les résidus des différentes séquences.

pap4-1::CIB, une diminution dans le mutant *amipap4-37* et aucun effet pour le mutant *pap4-2::CIB*. Pour l'ajout des queues polyC-riches, seul *nad1* est affecté dans les 3 mutants, avec une diminution. Pour *rtl*, une diminution des queues polyC-riches est observée pour les mutants *pap4-1::CIB* et *amipap4-37*, tandis qu'il n'y a pas de changement pour *pap4-2::CIB* par rapport à la souche WT. Ce gène a également une expression relative augmentée dans les 3 mutants.

Enfin, les 5 autres gènes, *cob*, *nad4*, *nad5*, *cox1* et *nad2*, sont eux suivis de courtes séquences intergéniques puis d'un gène de protéine. Dans les 3 mutants, l'expression de *cob* est diminuée, de même que la proportion de queues riches en C ajoutées. Il y a une augmentation de l'ajout de queues riches en A dans les souches *pap4-1::CIB* et *amipap4-37*, mais pas dans le mutant *pap4-2::CIB*. Pour les autres transcrits, les résultats divergent plus selon les souches mutantes, il est donc plus difficile de conclure. L'ARNm de *nad4* est moins exprimé dans les 3 mutants, mais la polycytidylation n'est diminuée que dans les mutants *pap4-1::CIB* et *amipap4-37*, pas dans le mutant *pap4-2::CIB*. Dans les mutants CLiP, le processus d'adénylation ne semble pas affecté ou peu, tandis qu'une augmentation est observée dans le mutant amiRNA. Pour le transcrit *nad5*, une expression diminuée est observée dans les mutants *pap4-1::CIB* et *amipap4-37*, pas dans le mutant *pap4-2::CIB*. Or dans ce dernier, il y a une diminution de l'ajout des queues riches en C, comme pour les 2 autres souches. C'est uniquement dans le mutant *amipap4-37* qu'il y a une légère augmentation des queues polyA pour ce transcrit. Concernant l'ARNm *cox1*, son expression n'est diminuée que dans le mutant *pap4-1::CIB*. Cela dit, une augmentation du taux de polyadénylation est remarquée dans les 3 mutants ainsi qu'une diminution de l'ajout des queues polyC. Enfin, le transcrit *nad2* est moins exprimé dans les mutants *amipap4-37* et *pap4-2::CIB*. Cependant, la polycytidylation est diminuée dans les 3 mutants, tandis que les résultats pour la polyadénylation et la polyuridylation varient selon les souches mutantes, ne nous permettant pas de conclure.

Globalement, il n'y a pas de corrélation entre la diminution de l'expression relative du transcrit et la diminution de la proportion de queues riches en C qui lui sont ajoutées. Aucune corrélation entre la diminution de l'ajout de C et l'augmentation de l'ajout de A ne peut être affirmée comme une règle générale à tous les transcrits. Par exemple, dans le mutant *pap4-1::CIB*, le transcrit *nad6* n'a presque aucun changement au niveau de l'ajout des queues polyC-riches dans les mutants, alors que la proportion d'ajout de queues polyA-riches est bien plus

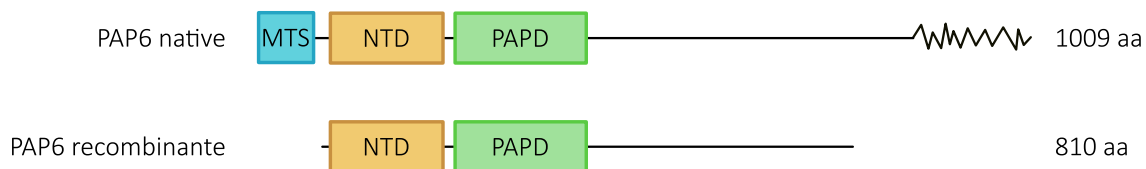


Figure R-38 : Protéine PAP6 recombinante pour la production des anticorps polyclonaux

La partie supérieure correspond à la représentation schématique de la protéine PAP6, la partie inférieure à celle correspondant à la protéine PAP6 recombinante utilisée pour la production des anticorps. Zig-zag : région désordonnée ; à droite : tailles en aa.

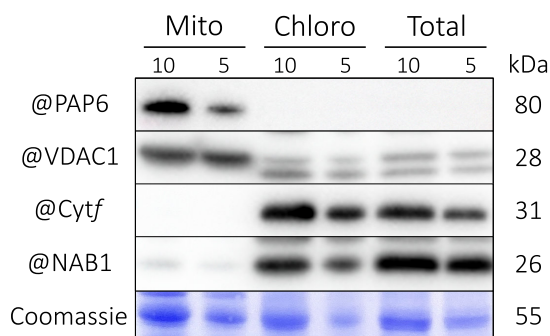


Figure R-39 : La protéine PAP6 est localisée uniquement dans les mitochondries

Expériences de western blot sur des extraits mitochondriaux (mito) ou chloroplastiques (chloro) purifiés, et des extraits de protéines totales (total). Ils ont été déposés sur gel SDS-PAGE 8-16%, en 2 quantités différentes (5 et 10 µg). Après transfert sur membrane, celle-ci a été colorée au bleu de Coomassie (Coomassie) puis incubée avec différents anticorps : @PAP6, @VDAC1, @Cyt f et @NAB1. A droite sont indiquées les tailles en kDa, auxquelles les signaux sont observés.

importante. En revanche, une importante diminution de la polycytidylation est observée pour le transcrit *cob*, mais il a une légère augmentation de l'ajout de queues polyA-riches. Tout cela laisse à penser que PAP4 est impliqué dans la polycytidylation, puisque ce processus est altéré lors de la mutation de PAP4. La fonction des queues polyC, polyA et polyU restent du moins à déterminer, ainsi que le lien entre la mutation de PAP4 et l'augmentation des queues riches en A et en U.

IV. PAP6, impliquée dans l'uridylation des ARNm mitochondriaux

1. Architecture de PAP6

Des analyses phylogénétiques réalisées dans le laboratoire montrent que la protéine PAP6 est proche des PAP non canoniques et possède les domaines caractéristiques des NT de classe I (Figure R-37A). L'alignement de la séquence protéique du motif NRM, qui est responsable de la sélectivité de la protéine pour les nucléotides, de PAP6 avec celles des TUTases et des PAP non canoniques montrent que le motif NRM de la protéine PAP6 est plutôt proche de celui de MUT68, URT1 et CID1, qui sont des enzymes ajoutant des U (Figure R-37B). Lors de l'analyse phylogénétique présentée en Figure R-7, nous avons déjà vu que PAP6 semble proche de TUTases, comme MUT68 ou HESO1. Chez les TUTases, au milieu du domaine NRM se trouve le résidu H responsable de leur sélectivité pour les U. Il a été également montré que si ce résidu était remplacé en un résidu L (leucine) ou absent, l'enzyme avait plutôt une sélectivité pour les A (Lunde, Magler and Meinhart, 2012; Yates *et al.*, 2012, 2015; Chung, Hyung Suk Jo and Heinemann, 2016). Chez *C. reinhardtii*, la protéine MUT68 possède un résidu D (acide aspartique) à cette position, or des études ont montré qu'elle est responsable de l'uridylation des ARN d'interférence matures (Ibrahim *et al.*, 2010). De plus, des tests d'activité *in vitro* ont également montré qu'en plus des U, elle pouvait ajouter des A. La protéine PAP6 contenant également un résidu D à cette position, il est donc possible d'émettre l'hypothèse qu'elle soit aussi capable d'ajouter des U ou des A, comme le fait MUT68.

2. Localisation uniquement mitochondriale de PAP6

Les expériences de localisation *in vivo* dans *N. benthamiana* montrent une localisation mitochondriale de la protéine PAP6 (Figure R-9). Nous avons toutefois souhaité confirmer cette localisation mitochondriale chez *C. reinhardtii* en réalisant des expériences de western blot sur des fractions cellulaires.

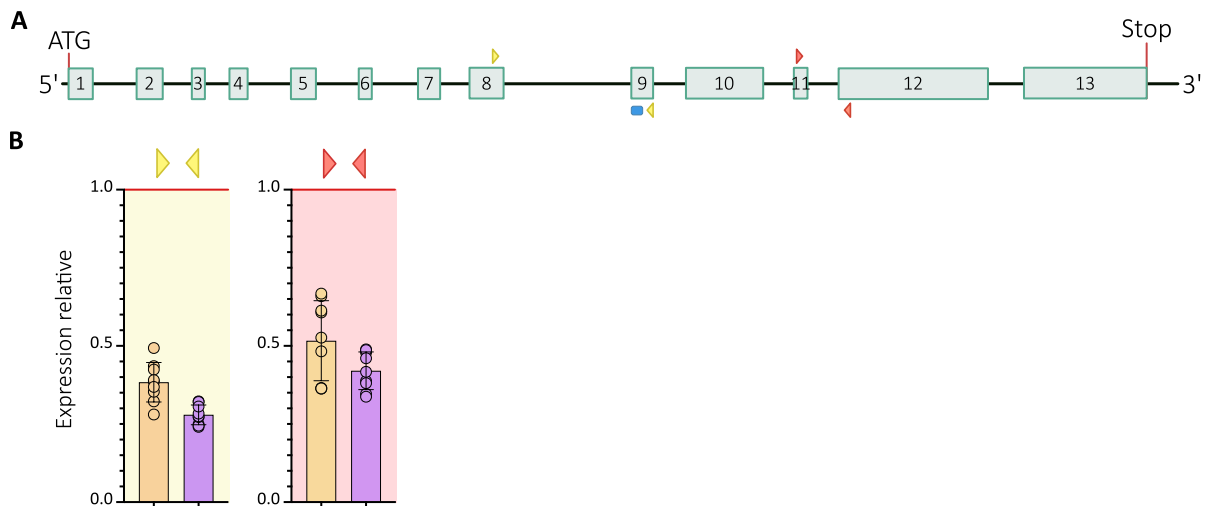


Figure R-40 : Expression relative du transcrit *PAP6* dans les mutants amiRNA

A : Schéma du transcrit de *PAP6* avec ses 13 exons numérotés (turquoise). La région ciblée par la construction amiRNA est indiquée (bleu). Les différents couples d'amorces utilisés en RT-qPCR sont schématisés par des triangles de différentes couleurs. **B** : Histogrammes de l'expression relative de l'ARNm de *PAP6* dans les mutants amiRNA *amipap6-64* (orange) et *amipap6-75* (violet), mesurée par RT-qPCR. Chaque couple d'amorce est associé à une couleur dans (A) qui est reportée en tant que couleur de fond. Le trait rouge horizontal indique le niveau d'expression relative du contrôle normalisé à 1. Les histogrammes présentent la moyenne de 3 réplicats biologiques (n = 3), avec l'écart-type (+/-, barre verticale). Pour chaque réplicat biologique, 3 réplicats techniques ont été faits. Toutes les valeurs obtenues sont représentées par les ronds sur les histogrammes.

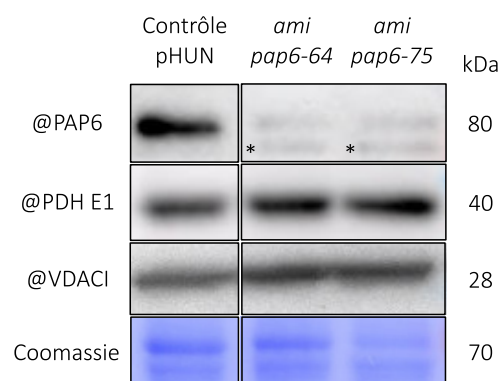


Figure R-41 : Les mutants amiRNA *PAP6* sont des mutants *knock-down*

Expériences de western blot sur des extraits de mitochondries purifiées. Les extraits mitochondriaux ont été déposés sur gel SDS-PAGE 8-16%, pour les 2 mutants *PAP6* et la souche contrôle. Après transfert, la membrane a été colorée au bleu de Coomassie (Coomassie), découpée et incubée avec différents anticorps : @*PAP6*, @*PDH E1* et @*VDAC1*. A droite sont indiquées les tailles en kDa, auxquelles les signaux sont observés. * : signaux aspécifiques.

Une première étape a été de produire des anticorps dirigés contre la protéine PAP6, comme cela a été fait pour PAP4. Une protéine recombinante de 810 aa a été produite, elle comporte le domaine NTD et le domaine PAPD (Figure R-38). La production des anticorps polyclonaux de lapins a été faite par la société Covalab. Les anticorps ont été testés par la Docteure Thalia Salinas-Giegé.

Les expériences de western Blot avec les anticorps PAP6 ont été faites comme pour PAP4 (Figure R-39). Avec les anticorps dirigés contre PAP6, le signal est uniquement visible pour les extraits correspondants aux mitochondries purifiées. Cela est cohérent avec une localisation mitochondriale de PAP6. PAP6 ayant une taille théorique d'environ 100 kDa, il est surprenant d'observer un signal autour de 80 kDa. Plusieurs tests et réplicats de western blot semblent indiquer que le signal observé à cette taille est celui qui correspond à PAP6. Aucun signal n'est observé sur les extraits correspondants aux chloroplastes purifiés, excluant ainsi le double adressage mitochondrial et chloroplastique de la protéine PAP6.

Bien que PredAlgo prédise une localisation de PAP6 aux voies de sécrétion, les localisations en *N. benthamiana*, les analyses de protéomiques ainsi que les western blot sur les fractions cellulaires indiquent que la protéine PAP6 est mitochondriale.

3. Obtention de lignées mutantes

Aucun mutant de PAP6 n'était disponible dans la banque CLiP au début de ma thèse, c'est pourquoi des mutants ont été générés par la technique amiRNA, ciblant une région au début de l'exon 9 (Figure R-40A).

Nous avons obtenu 108 transformants et des analyses de RT-qPCR ont été réalisées sur 20 transformants à l'aide de 2 couples d'amorces, l'un situé entre les exons 8 et 9, comprenant la région ciblée par l'amiRNA, et l'autre entre les exons 11 et 12, en aval de la séquence ciblée par l'amiRNA. Comme pour les mutants amiRNA de PAP4, la souche pHUN a été utilisée comme contrôle. Ceci a permis d'isoler 2 mutants intéressants : *amipap6-64* et *amipap6-75* pour lesquels une diminution de l'expression relative du transcrit *PAP6* de 55,5% pour le mutant *amipap6-64* et de 65% pour le mutant *amipap6-75* (en moyenne pour les 2 couples d'amorces) a été observée (Figure R-40B).

Le niveau de protéine PAP6 présent dans ces mutants a été ensuite analysé par Western blot sur des extraits de mitochondries purifiées de ces mutants (Figure R-41). Les anticorps dirigés

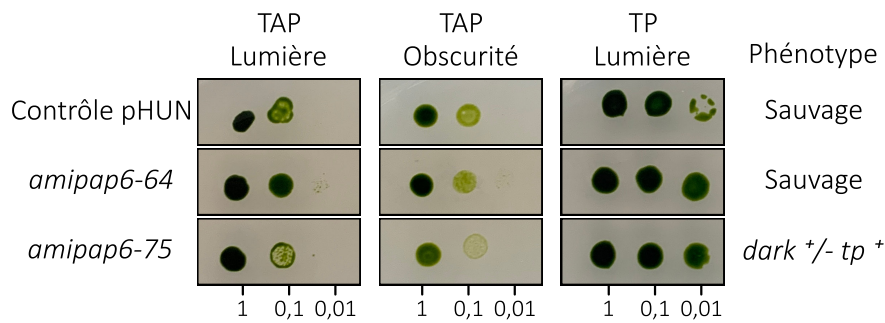


Figure R-42 : Tests phénotypiques des mutants de PAP6

Phénotypes de croissance des mutants amiRNA à la lumière ou à l'obscurité dans un milieu en présence (TAP) ou en absence (TP) de source de carbone. La concentration par rapport à la culture de départ (1 = 6×10^6 cellules/mL) sont indiquées en dessous des photos. Les phénotypes observés sont notés à droite des photos.

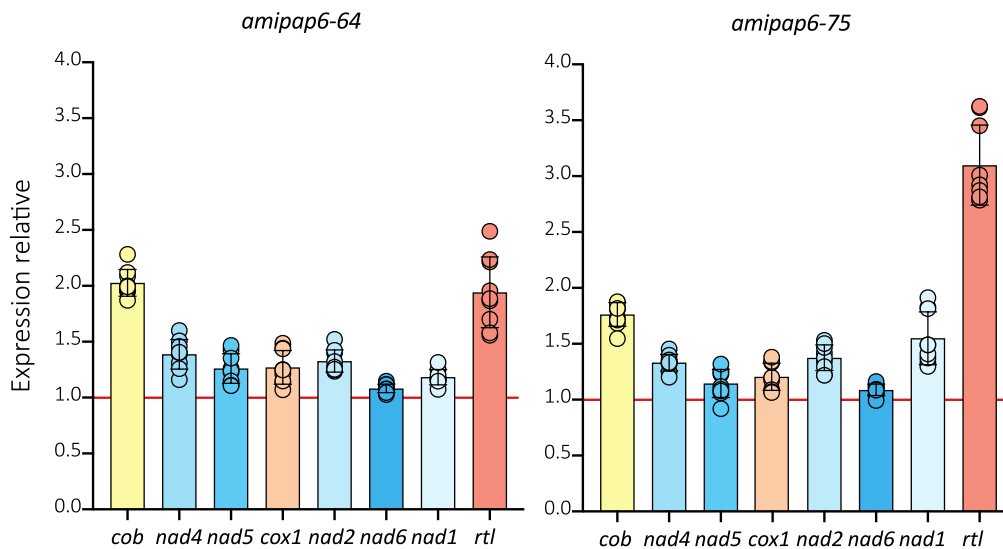


Figure R-43 : Analyses de l'expression relative des transcrits mitochondriaux des mutants PAP6

Histogrammes de l'expression relative des 8 ARNm mitochondriaux, dans l'ordre de succession des gènes dans le génome, dans les mutants *amipap6-64* et *amipap6-75*, mesurée par RT-qPCR. Le trait rouge horizontal indique le niveau d'expression relative de la souche contrôle pHUN normalisé à 1. Les histogrammes présentent la moyenne de 3 réplicats biologiques (n = 3), avec l'écart-type (+/-, barre verticale). Pour chaque réplicat biologique, 3 réplicats techniques ont été faits. Toutes les valeurs obtenues sont représentées par les ronds sur les histogrammes.

contre la protéine PDH E1 matricielle et la protéine VDACL de la membrane externe ont été utilisés comme contrôles afin de vérifier respectivement l'intégrité et l'enrichissement en mitochondries. Le signal avec les anticorps PAP6 montre une très forte diminution de la protéine PAP6 pour les 2 souches mutantes *amipap6-64* et *amipap6-75*. Ce sont donc des mutants KD, fortement affectés.

4. Caractérisation des mutants de PAP6

a) Phénotypes des mutants

Pour déterminer si les mitochondries sont affectées dans ces 2 lignées mutantes, leur croissance à l'obscurité a été testée. Les résultats montrent que seul le mutant *amipap6-75* présente un retard de croissance à l'obscurité, il a donc un phénotype *dark*^{+/-} (Figure R-42). Leur croissance a été également testée en milieu minimum à la lumière mais aucun des mutants amiRNA de PAP6 n'a présenté de phénotype (Figure R-42).

Ainsi, seul le mutant *amipap6-75* semble être légèrement affecté au niveau du métabolisme mitochondrial.

Comme pour les mutants de PAP4, l'expression relative des transcrits mitochondriaux a été analysée par RT-qPCR, afin de voir si la diminution de la présence de PAP6 dans les mutants affecte leur expression (Figure R-43). Les données montrent une augmentation globale de l'expression relative des transcrits mitochondriaux pour les 2 mutants à l'exception du transcrit *nad6* qui semble être similaire à la souche contrôle. L'augmentation est plus marquée pour les transcrits *cob* et *rtl* qui se trouvent aux 2 extrémités du génome mitochondrial (Figure R-44). L'augmentation du transcrit *rtl* est d'ailleurs beaucoup plus importante dans le mutant *amipap6-75* que dans le mutant *amipap6-64* avec une augmentation de 309% et 194% respectivement. Pour les autres transcrits cette augmentation varie entre 117% et 142%.

Nous avons ainsi obtenu deux mutants pour PAP6 qui présentent une forte diminution de la protéine. Bien que cette diminution de la protéine semble similaire dans les 2 mutants, les tests de croissance à l'obscurité montrent que le mutant *amipap6-75* est plus affecté que le mutant *amipap6-64*. Dans les mutants de PAP4, le niveau du transcrit *rtl* est également augmenté mais par contre les transcrits *cob*, *nad4*, *nad5*, *cox1* et *nad2*, présentent une diminution (Figure R-44). Le comportement des mutants PAP4 versus ceux de PAP6 diffèrent en ce point

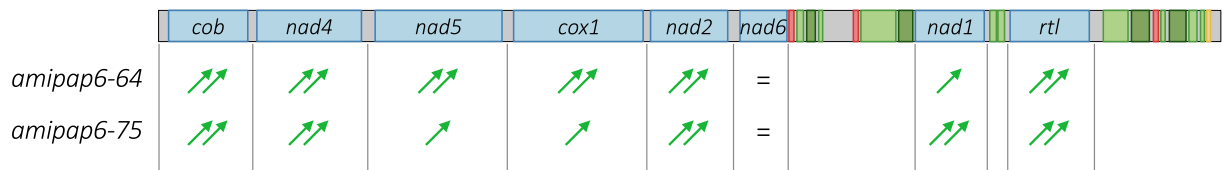


Figure R-44 : Effets de la mutation de PAP6 sur les ARNm mitochondriaux selon leur position

Schéma du génome mitochondrial de *C. reinhardtii* et résultats des analyses de RT-qPCR des transcrits du génome mitochondrial selon les mutants amiRNA PAP6.

Flèches **vertes** : augmentation par rapport au contrôle ; = : similaire au contrôle, moins de 10% de différences avec le contrôle ; 1 flèche : entre 10% et 25% de différence avec le contrôle ; 2 flèches : au-delà de 25% de différence avec le contrôle.

Bleu : gènes de protéines ; **Rouge** : gènes d'ARNt ; **Vert clair** : gènes d'ARNr de la grande sous-unité du mitoribosome ; **Vert foncé** : gènes d'ARNr de la petite sous-unité du mitoribosome ; **Jaune** : gène de l'ARN L2b.

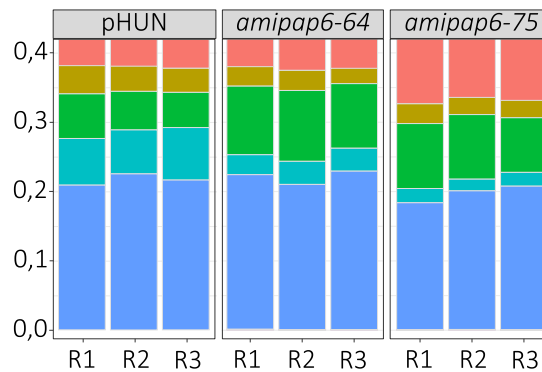


Figure R-45 : Les résultats des triplicats sont similaires pour les mutants amiRNA de PAP6

Analyses des expériences de 3'RACE-seq des mutants amiRNA de PAP6 et de la souche de référence pHUN, en fonction de la catégorie des queues, pour les 3 réplicats biologiques. En abscisse sont indiquées les proportions des différentes catégories de queues, selon les souches et selon les 3 réplicats (R1, R2 et R3). Zoom des graphiques présentés en [Annexe VIII-2](#).

Corail : pas de queues ; **Jaune** : queues riches en A, U et C ; **Vert** : queues riches en A ; **Bleu clair** : queues riches en U ; **Bleu foncé** : queues riches en C.

et cela semble indiquer une fonction différente de ces 2 protéines dans la mitochondrie. La question se pose alors de savoir quels types de queues sont affectés dans les mutants PAP6.

5. Queues polyU-riches affectées dans les mutants de PAP6

Comme pour les mutants PAP4, des analyses de 3'RACE-seq ont été menées sur les 2 mutants amiRNA de PAP6, *amipap6-64* et *amipap6-75*. De même, le mélange de souches pHUN a été utilisé comme référence pour nos analyses. Comme pour PAP4, les résultats des 3 réplicats étant similaires, nous avons cumulé les données obtenues dans les réplicats afin d'en faire une analyse globale (Figure R-45). Les Figures R-46, R-47 et R-48 présentent les résultats obtenus.

Les différences entre les mutants et la souche de référence sont calculées en pourcentages de différence entre les 2 souches comparées. Ces valeurs sont disponibles en Annexe X. L'analyse des résultats globaux (Figure R-46A) montre une faible variation des queues polyC-riches avec une augmentation de 2% pour le mutant *amipap6-64* et une diminution de 11% pour le mutant *amipap6-75*, par rapport à la souche pHUN. Par contre, nous pouvons observer une forte diminution des queues polyU-riches dans les 2 mutants, qui est beaucoup plus marquée pour le mutant *amipap6-75* avec une réduction de 72% par rapport à la souche de référence pHUN. A l'opposé, les queues polyA-riches subissent une augmentation de 81% et 63% pour *amipap6-64* et *amipap6-75*, respectivement. L'analyse de l'ajout de queues polyC-riches pour chaque transcrite mitochondrial montre des profils globalement similaires entre les 2 mutants et la souche de référence (Figures R-46B et C). Néanmoins, des augmentations de 50% et 36% sont constatées pour le transcrite *nad2* dans les mutants *amipap6-64* et *amipap6-75*, respectivement. Des diminutions sont observées pour le mutant *amipap6-75*, concernant les transcrits *nad4*, *nad5*, *cox1* et *nad6*, d'environ 22%. De même, les profils de polyadénylation des 2 mutants restent similaires à ceux de la souche contrôle pHUN, avec une augmentation globale de la polyadénylation de tous les transcrits, variant entre 25% et 100% selon les ARNm (Figures R-46B et D). Enfin, l'ajout des queues polyU-riches est moindre pour les mutants que pour la souche contrôle pHUN, avec des profils similaires (Figures R-46B et E). En effet, les proportions d'ajouts de queues polyU diminuent pour tous les transcrits des 2 mutants, entre 25 et 85% de réduction, à part pour *nad6* dans le mutant *amipap6-64*, qui a une uridylation équivalente à la souche contrôle pHUN.

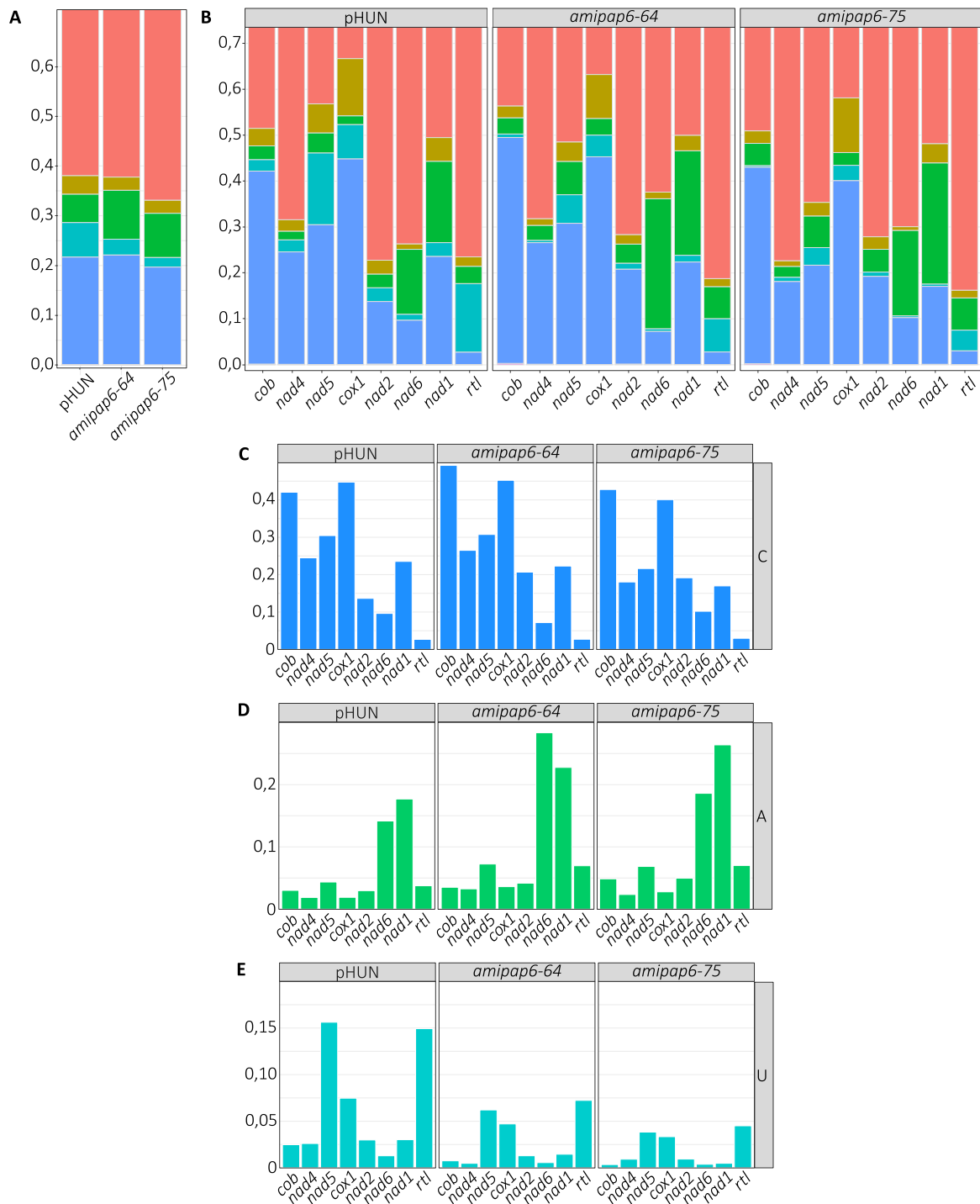


Figure R-46 : Comparaison des 3 catégories de queues ajoutées entre les mutants amiRNA de PAP6 et la souche contrôle pHUN

Analyses des expériences de 3'RACE-seq des mutants amiRNA pour PAP6 en fonction de la catégorie des queues. Résultats cumulés des 3 réplicats biologiques. **A** et **B** : En abscisse sont indiquées les proportions des différentes catégories de queues, selon les souches (**A**) et selon les gènes pour chaque souche (**B**). Zoom des graphiques présentés en [Annexe VIII-2](#). **C**, **D** et **E** : En abscisse sont indiquées les proportions selon les gènes, pour les queues polyC-riches (**C**), polyA-riches (**D**) ou polyU-riches (**E**). **Corail** : pas de queues ; **Jaune** : queues riches en A, U et C ; **Vert** : queues riches en A ; **Bleu clair** : queues riches en U ; **Bleu foncé** : queues riches en C.

La longueur des queues ajoutées a également été analysée (Figure R-47). Pour les queues polyC et polyA-riches, malgré l'augmentation globale visible dans les 2 mutants pour l'ajout des queues polyA-riches, aucune variation n'est observée dans les profils du nombre de queues ajoutées en fonction de leur longueur (Figure R-47A et B). Concernant les queues polyU-riches, le mutant *amipap6-75* a un profil différent de la souche de référence, avec une forte diminution globale mais également une forte diminution du nombre de petites queues ajoutées (entre 3 et 5 nt) (Figure R-47C). En revanche, le mutant *amipap6-64* a un profil similaire à la souche pHUN, avec une diminution globale du nombre de queues ajoutées.

Enfin, nous avons regardé la distribution des 3 types de nt ajoutés, en fonction de la position, par rapport à la fin de la séquence 3'UTR de chaque transcrite (Figure R-48). Pour ces analyses, toutes les queues identifiées ont été prises en compte, sans appliquer de filtre de taille ou de composition. Comme pour PAP4, les résultats complets sont disponibles en Annexe XI, seuls les profils différents entre la souche de référence et les mutants sont montrés, par soucis de lisibilité. Concernant l'ajout de C, nous constatons des différences de profils entre les 2 mutants et la souche de référence pHUN pour les transcrits *nad5* et *cox1* (Figure R-48A). En effet, dans les 2 cas, un changement de la position correspondant à l'ajout majoritaire de C est observé. Pour la souche contrôle pHUN, il y a une majorité d'ajouts de C entre les positions 1 et 5 pour *nad5*, tandis que la fréquence d'ajout de C est majoritaire à la position 1 uniquement dans les 2 mutants. De même, pour le transcrite *cox1*, il y a le plus d'ajouts de C entre les positions 4 et 8 dans la souche pHUN, tandis que cela se fait majoritairement entre les positions 4 et 6 dans les 2 mutants. Enfin, une différence est également observée pour le transcrite *nad1*, mais entre les souches pHUN et *amipap6-75* uniquement. Là, la fréquence majoritaire d'ajouts de C se fait entre les positions 0 et 5 dans le mutant et entre les positions 4 et 5 uniquement dans la souche contrôle. Pour l'ajout de A, les résultats sont plus homogènes entre les 2 mutants, mais les profils sont différents de la souche contrôle pour les transcrits *nad5*, *nad1* et *rtl* (Figure R-48B). Dans les 3 cas, l'ajout majoritaire de A se fait à des positions plus en amont que pour la souche de référence pHUN. Enfin pour l'ajout de U, il n'y a qu'au niveau du transcrite *rtl* que des différences de profils sont observées (Figure R-48C), avec un décalage de la position de 1 en amont pour l'ajout majoritaire des U. Même si certains gènes, notamment *nad5*, *cox1*, *nd1* et *rtl*, ont des profils différents, avec un ajout majoritaire des C, A ou U à des positions plus en amont vis-à-vis de la souche contrôle, globalement, ces analyses montrent que les profils de

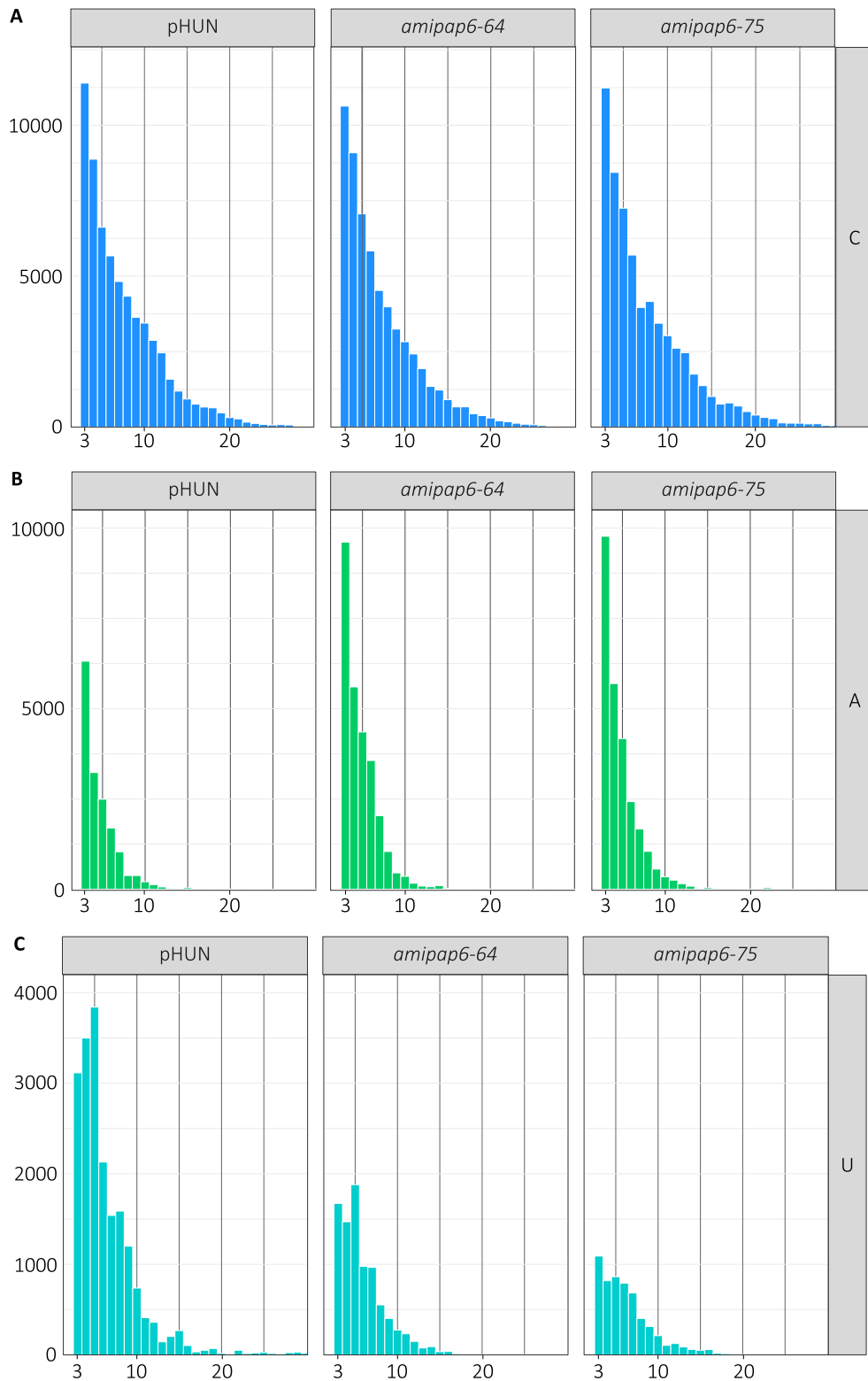


Figure R-47 : La longueur des queues ajoutées varie dans les mutants

Analyses des expériences de 3'RACE-seq des mutants amiRNA pour PAP6 en fonction de la catégorie des queues. Résultats cumulés des 3 réplicats biologiques. **A**, **B** et **C** : histogrammes du nombre de queues polyC-riches (**A**), polyA-riches (**B**) ou polyU-riches (**C**) ajoutées selon leur longueur, pour chaque souche.

Bleu foncé : queues riches en C ; **Vert** : queues riches en A ; **Bleu clair** : queues riches en U.

distribution d'ajouts de nt en fonction de la position sont plutôt similaires pour l'ensemble des gènes entre les mutants et la souche contrôle.

En résumé, dans les deux mutants *amipap6-64* et *amipap6-75*, malgré quelques faibles variations, la très forte réduction de PAP6 ne semble pas affecter la polycytidylation. Par contre, les données montrent d'une part une diminution de la polyuridylation dans les mutants de PAP6, et d'autre part une augmentation de la polyadénylation (Figure R-49). Hormis *nad6* pour lequel la stabilité de l'ARNm reste inchangée, tous les autres transcrits mitochondriaux voient leur expression relative inchangée.

Ces résultats suggèrent que PAP6 puisse être impliquée dans l'addition des queues polyU-riches, puisque ce processus est altéré lors de la mutation de PAP6. De plus, cela entraîne une augmentation des queues polyA-riches. Une telle augmentation, bien que plus modérée, a aussi été observée pour les mutants de PAP4. Néanmoins, ne connaissant la fonction d'aucune de ces 3 additions dans les mitochondries de *C. reinhardtii*, il est difficile de faire un lien entre PAP4, PAP6 et l'augmentation des queues polyA-riches.

6. Activité *in vitro* de PAP6

Enfin, en parallèle de l'analyse des mutants, nous avons souhaité tester l'activité d'ajout de nt par la protéine PAP6. Pour cela, la Docteure Thalia Salinas-Giegé a fait des tests d'activité *in vitro* avec une protéine PAP6 recombinante, afin de tester son activité NT. La protéine PAP6 recombinante, nommée « PAP6 », correspond à celle utilisée pour la production des anticorps (Figure R-38 et R-50). Pour ces tests d'activité, une deuxième protéine recombinante mutée, dont le site catalytique a été produite. Dans cette version mutée deux résidus aspartate conservés dans le site catalytique des Nucléotidyl-transférases (D96 et D98 pour PAP6) ont été mutés en alanines.

Différents tests d'activité *in vitro* ont été faits par la Docteure Thalia Salinas-Giegé et l'ingénieur Alexandre Boubegtitene (Figure R-51).

Dans un premier temps, les activités de la protéine PAP6 et de sa version mutante ont été testées sur un oligoribonucléotide de 19 nt correspondant à la séquence 3'UTR de *nad6*, en présence d'un mélange des 4 rNTP (ribonucléotides) (Figure R-51A). Les résultats sur gel montrent une incorporation de rNTP à l'extrémité 3' de l'ARN en présence de PAP6, confirmant l'activité NT de la protéine PAP6. De plus, en présence de PAP6 mutée, il n'y a pas de

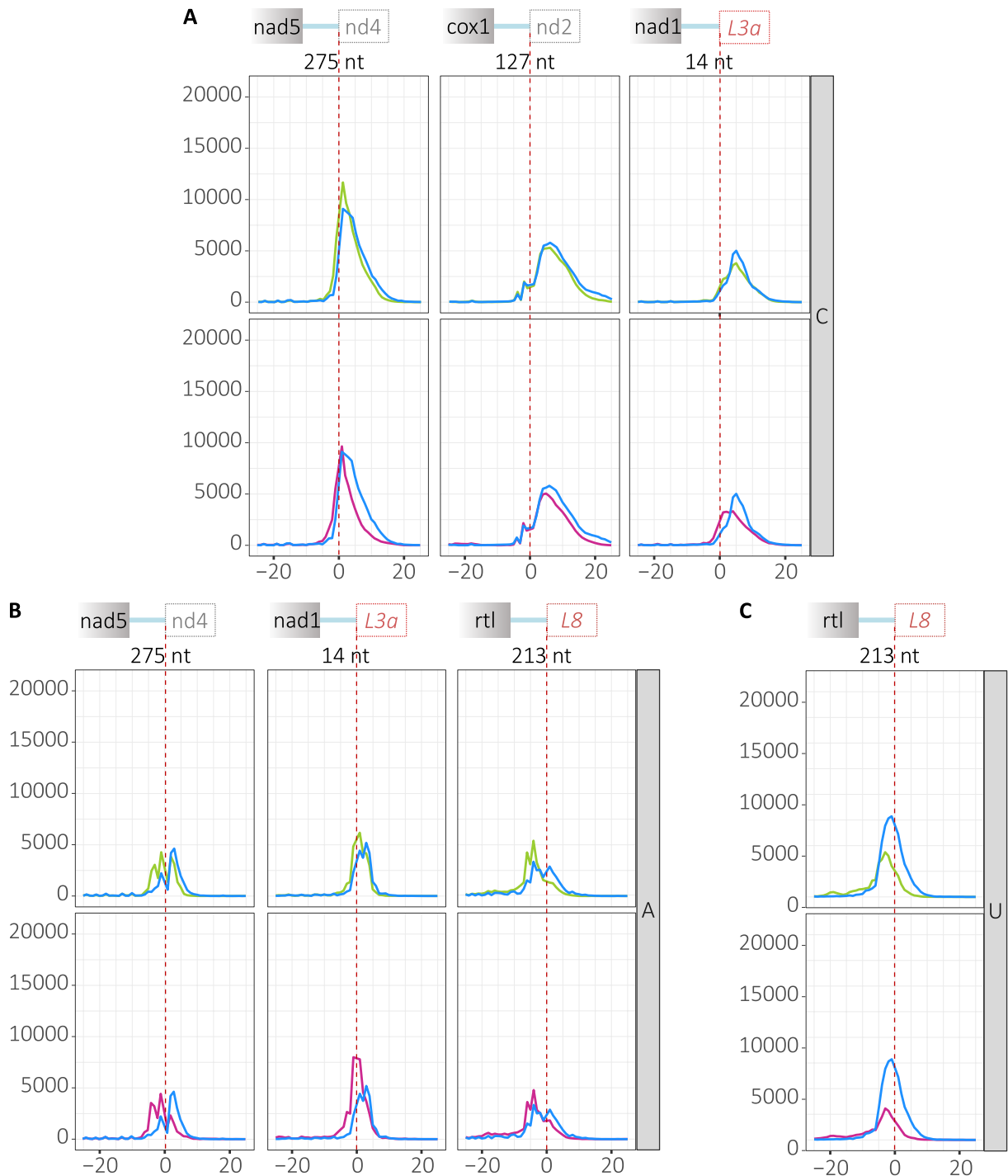


Figure R-48 : Distribution des nt différents dans les mutants, pour certains transcrits

Analyses des expériences de 3'RACE-seq des mutants amiRNA de PAP6. Résultats cumulés des 3 réplicats biologiques. **A**, **B** et **C** : graphiques du nombre de nt C (**A**), A (**B**) ou U (**C**) trouvés dans les queues, en fonction de leur position vis-à-vis de la fin de la 3'UTR (position 0, trait rouge pointillé) pour certains transcrits mitochondriaux. Au-dessus des graphiques sont schématisés la séquence codante du transcrit (gris), suivie de sa séquence 3'UTR (bleu, avec la taille indiquée en dessous en nt) et l'environnement de ce transcrit avec le transcrit suivant (gris : ARNm ; rouge : ARNr et vert : ARNt). Bleu : souche contrôle pHUN ; Vert : *amipap6-64* ; Rose : *amipap6-75*.

changement de taille. Ceci montre qu'en absence des 2 résidus D de la triade catalytique, il n'y a plus d'activité NT de PAP6. Ensuite, l'activité de PAP6 a été testée sur 2 autres oligoribonucléotides de 21 nt, correspondant à la séquence 3'UTR de *nad4* et à la séquence CDS de *nad6*, respectivement (Figure R-51B). PAP6 est capable d'ajouter des ribonucléotides sur les 2 ARN montrant ainsi que l'ajout des rNTP par PAP6 n'est pas séquence-spécifique. Par la suite, des tests ont été faits avec le mélange des 4 rNTP ou en présence d'un seul rNTP (rATP, rCTP, rGTP et rUTP) (Figure R-51C). On observe un changement de taille de l'ARN soit en présence du mélange de rNTP (10mM de chaque), soit en présence de rUTP seul (10 mM). Ceci montre que la protéine PAP6 a vraisemblablement une sélectivité pour les rUTP. Une différence de taille est observée entre l'activité en présence du mélange de rNTP (environ 33 pb en moyenne) et de rUTP uniquement (environ 24 pb en moyenne), avec des bandes plus hautes en présence du mélange de rNTP. Cela semble indiquer qu'en présence de rUTP, d'autres nucléotides peuvent également être ajoutés. Pour confirmer cela, des tests ont été faits en présence d'un mélange constitué de rUTP et d'un deuxième rNTP (rATP, rCTP ou rGTP) (Figure R-51D). Les résultats sur gel montrent qu'il y a plus de ribonucléotides ajoutés avec les mélanges de 2 rNTP qu'avec le rUTP seuls. Il semble alors que PAP6 peut ajouter d'autres rNTP, à condition d'être en présence de rUTP.

Ces tests d'activité indiquent que la protéine PAP6 ajoute préférentiellement des U. Cela confirme les résultats de 3'RACE-seq, qui montraient une diminution de l'uridylation dans les mutants de PAP6. Est-ce que PAP6 doit d'abord ajouter une série de rUTP avant d'ajouter d'autres rNTP ou bien est-ce qu'elle fait une alternance entre les rUTP et les autres rNTP ? Il est également possible que, *in vivo*, PAP6 n'ajoute que des U. En effet, il a déjà été montré que des expériences *in vitro* peuvent ne pas refléter la spécificité qu'a une enzyme envers son substrat *in vivo* (Yehudai-Resheff and Schuster, 2000). C'est une question à laquelle nous devons répondre à l'avenir.

| | | <i>cob</i> | <i>nad4</i> | <i>nad5</i> | <i>cox1</i> | <i>nad2</i> | <i>nad6</i> | | <i>nad1</i> | <i>rtl</i> | |
|-------------------|---------------------|------------|-------------|-------------|-------------|-------------|-------------|--|-------------|------------|--|
| <i>amipap6-64</i> | Expression relative | ↗↗ | ↗↗ | ↗↗ | ↗↗ | ↗↗ | = | | ↗ | ↗↗ | |
| | Queues polyC-riches | ↗ | = | = | = | ↗↗ | ↘↘ | | = | = | |
| | Queues polyA-riches | ↗↗ | ↗↗ | ↗↗ | ↗↗ | ↗↗ | ↗↗ | | ↗↗ | ↗↗ | |
| | Queues polyU-riches | ↘↘ | ↘↘ | ↘↘ | ↘↘ | ↘↘ | = | | ↘↘ | ↘↘ | |
| | | | | | | | | | | | |
| <i>amipap6-75</i> | Expression relative | ↗↗ | ↗↗ | ↗ | ↗ | ↗↗ | = | | ↗↗ | ↗↗ | |
| | Queues polyC-riches | = | ↘ | ↘↘ | ↘ | ↗↗ | = | | ↘↘ | = | |
| | Queues polyA-riches | ↗↗ | ↗↗ | ↗↗ | ↗↗ | ↗↗ | ↗↗ | | ↗↗ | ↗↗ | |
| | Queues polyU-riches | ↘↘ | ↘↘ | ↘↘ | ↘↘ | ↘↘ | ↘↘ | | ↘↘ | ↘↘ | |
| | | | | | | | | | | | |

Figure R-49 : Effets de la mutation de PAP6 sur les ARNm mitochondriaux

Schéma du génome mitochondrial de *C. reinhardtii* et résultats des analyses de RT-qPCR (« Expression relative ») et de 3'RACE-Seq (« Queues polyC, A ou U-riches ») sur les transcrits du génome mitochondrial selon les mutants, en comparaison avec les contrôles.

Flèches : **rouges** : diminution par rapport au contrôle ; **vertes** : augmentation par rapport au contrôle ; = : similaire au contrôle, moins de 10% de différences avec le contrôle ; 1 flèche : entre 10% et 25% de différence avec le contrôle ; 2 flèches : au-delà de 25% de différence avec le contrôle.

Bleu : gènes de protéines ; **Rouge** : gènes d'ARNt ; **Vert clair** : gènes d'ARNr de la grande sous-unité du mitoribosome ; **Vert foncé** : gènes d'ARNr de la petite sous-unité du mitoribosome ; **Jaune** : gène de l'ARN L2b.

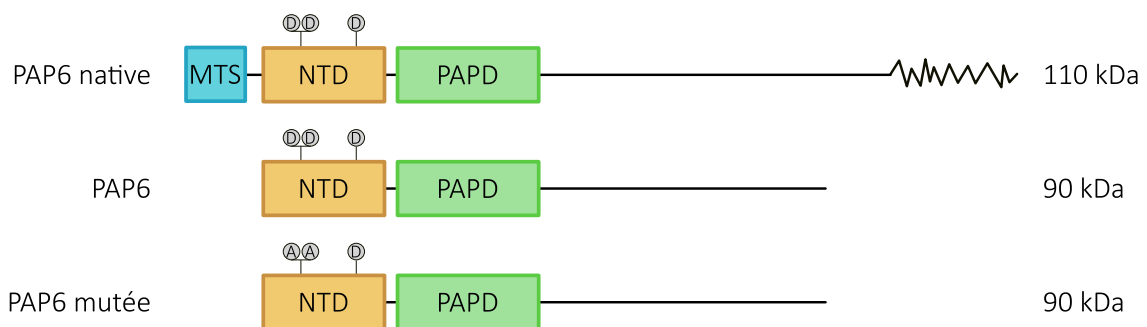


Figure R-50 : Production des protéines recombinantes PAP6 et PAP6 mutée

En haut, la protéine PAP6 native. Au milieu, la portion de PAP6 qui a été clonée et produite pour obtenir la protéine recombinante « PAP6 ». En bas, la portion de PAP6 qui a été clonée, mutée (2 premiers D de la triade catalytique remplacés par des A) et produite pour obtenir la protéine recombinante « PAP6 mutée ». A droite sont indiquées les tailles en kDa. La région en zig-zag correspond à la région C-terminale riche en A (alanine).

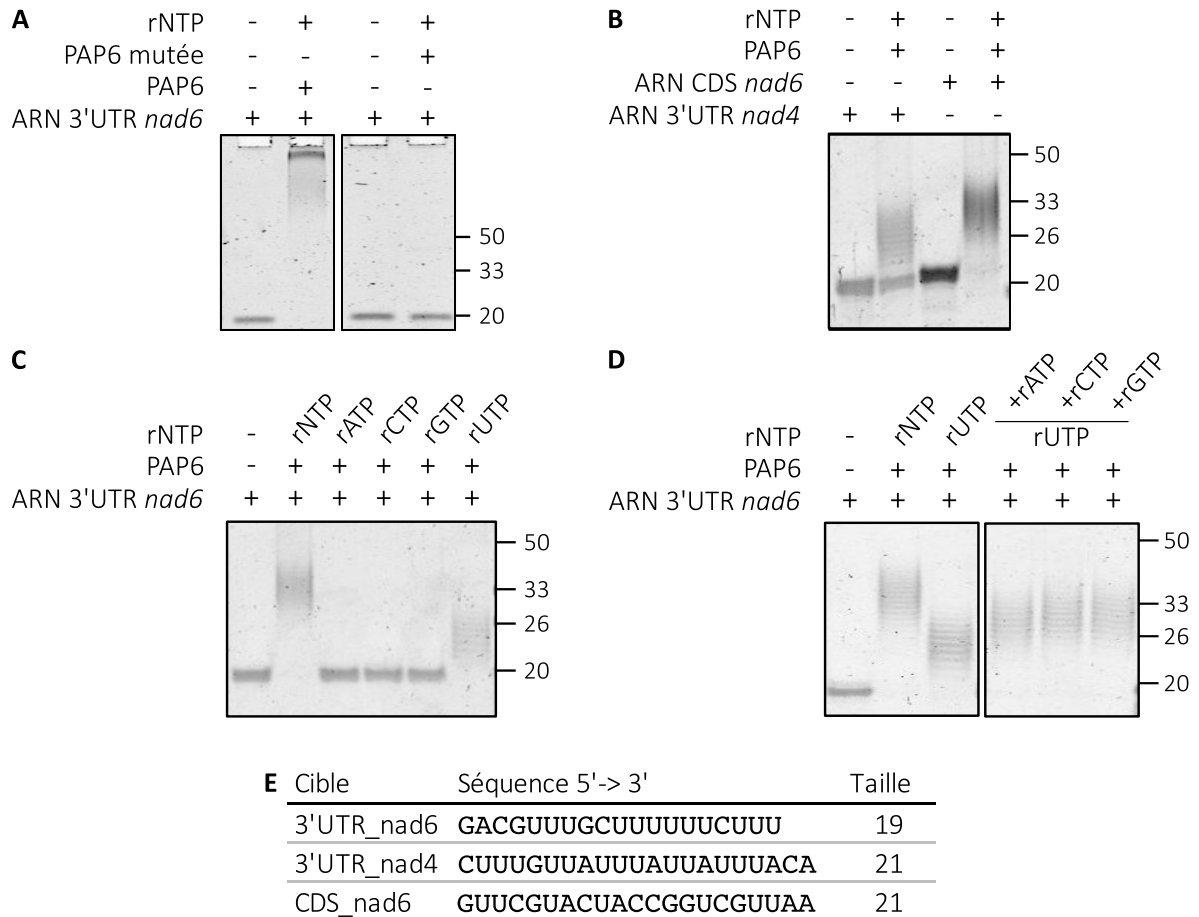


Figure R-51 : Test d'activité *in vitro* de PAP6

Résultats des tests d'activité *in vitro* avec la protéine recombinante PAP6 (A, B, C et D) et la protéine recombinante PAP6 mutée (A), en présence d'un mélange de rNTP (mélange équimolaire des 4 rNTP) ou de rNTP seuls. Ces tests ont été fait avec la séquence ARN de la 3'UTR du gène mitochondrial *nad6* (A, C et D) ou les séquences ARN de la 3'UTR du gène mitochondrial *nad4* ou de la séquence codante de *nad6* (B). Les séquences et tailles de ces 3 ARN différents sont indiqués en E.

+ : présence ; - : absence. Les tailles sont indiquées en nt à droite des gels.

CHAPITRE II : CARACTERISATION DU MITORIBOSOME DE *C. REINHARDTII*

I. Mise en contexte et résultats complémentaires à l'article

Le deuxième axe de travail de ma thèse s'articule autour de la caractérisation du mitoribosome de *C. reinhardtii*. Ces travaux ont été publiés en décembre 2021, dans *Nature Communications* (Waltz *et al.*, 2021). C'est pourquoi après une brève introduction et des précisions sur mon implication dans ce projet, les résultats seront présentés dans l'article joint ci-après.

De nombreuses études ont été publiées récemment décrivant les mitoribosomes dans différents organismes, détaillées dans la revue de (Scaltsoyiannes *et al.*, 2022b). Ces investigations ont montré que l'architecture globale du mitoribosome, leur composition en ARNr et en protéines ribosomales divergeaient dans les différentes lignées eucaryotiques. Au début de ma thèse, pour les organismes photosynthétiques, seul le mitoribosome d'*Arabidopsis thaliana* avait été étudié (Waltz *et al.*, 2019) et aucune donnée n'était disponible pour d'autres organismes photosynthétiques tels que les algues vertes.

C. reinhardtii possède des ARNr et des ARNm caractéristiques qui rendent l'étude de son mitoribosome très attractive. En effet, les ARNr sont fragmentés en 9 modules pour la LSU et 4 pour la SSU (Boer and Gray, 1988). De plus, l'ARNr 5S n'était pas identifié. D'autre part, les ARNm mitochondriaux n'ont pas de 5'UTR, et cela pose la question de savoir comment se fait l'initiation de la traduction (voir paragraphe [Initiation](#) dans [Introduction - Expression du génome mitochondrial de *C. reinhardtii* – II-3-d](#)). Enfin, puisque les queues polyC-riches sont majoritairement trouvées sur les ARNm matures, il est possible qu'elles soient nécessaires pour la traduction des ARNm. Peut-être ont-elles un rôle dans l'initiation de la traduction ?

Connaitre la structure et la composition du mitoribosome avait comme objectif d'aider à répondre à ces questions. C'est pourquoi un travail collaboratif entre notre équipe, celle du Docteur Philippe Giégé (IBMP), du Docteur Yaser Hashem (IECB – Bordeaux), et celle du Docteur Ben Engel (Université de Bâle – Bâle, Suisse) a été mis en place pour obtenir la structure du mitoribosome de cette algue. Des analyses par Cryo-EM et par spectrométrie de masse de mitoribosomes purifiés ont permis de reconstituer la composition et de proposer un modèle de la structure du mitoribosome de *C. reinhardtii*. Cette reconstitution a révélé des caractéristiques structurales uniques avec des domaines protéiques complètement nouveaux

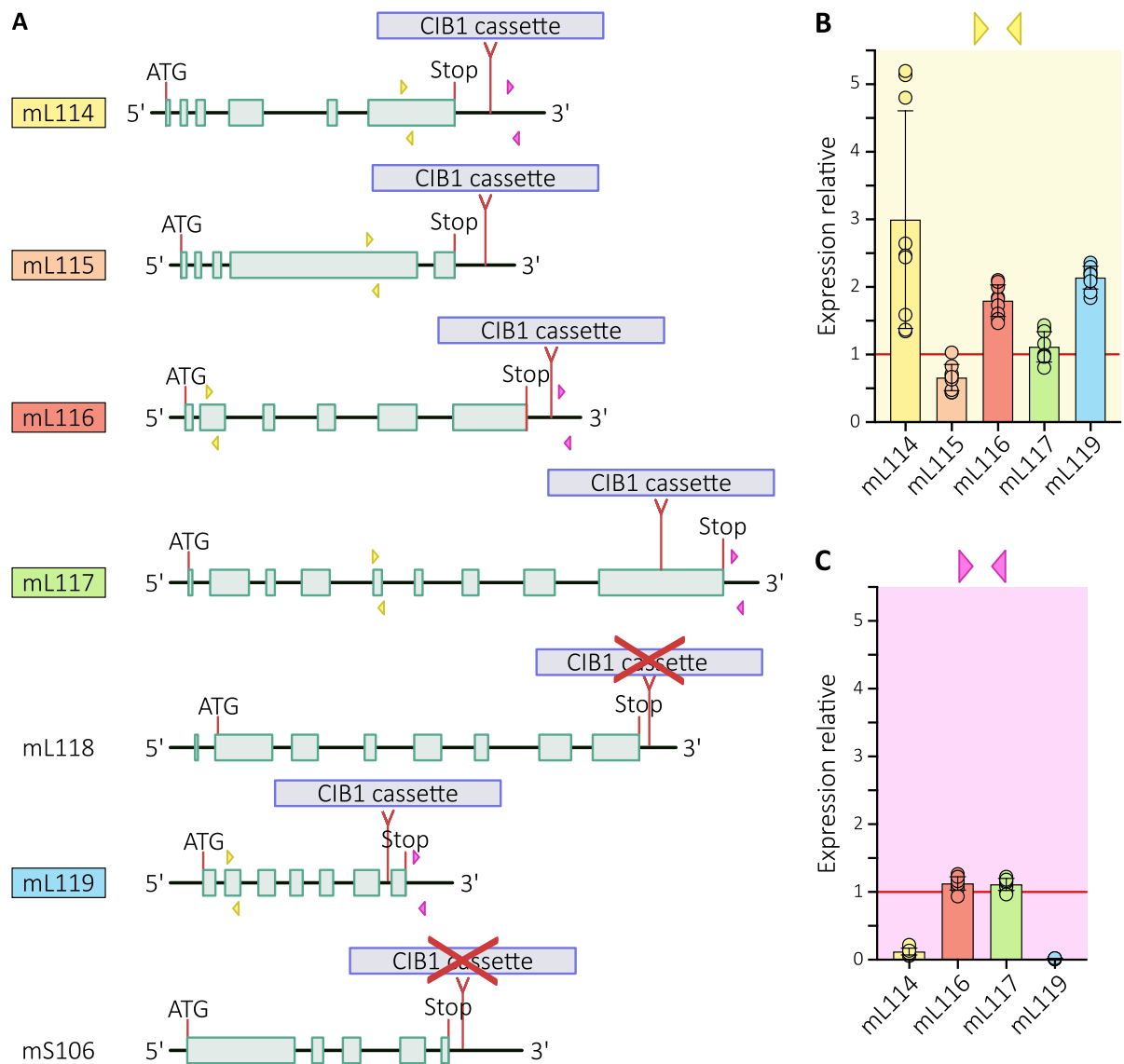


Figure R-52 : Mutants d'insertion des protéines mitoribosomales

A : Schéma des gènes *mL114* à *mL119*, ainsi que de *mS106* et de l'insertion de leur cassette CIB, aux endroits prédits par la banque CLiP pour les différentes lignées. Aucune insertion n'a été trouvée à l'endroit indiqué pour *mL118::CIB* et *mS106::CIB* (croix rouges sur les cassettes). Les couples d'amorces utilisés en RT-qPCR sont schématisés par des triangles de différentes couleurs. **Turquoise** : exons. **B** : Histogrammes de l'expression relative du gène ciblé par l'insertion dans les mutants correspondants, mesurée par RT-qPCR. Chaque couple d'amorce est associé à une couleur dans (A) qui est reportée en tant que couleur de fond, **jaune** pour les amorces avant l'insertion, **rose** pour celles après. Le trait **rouge** horizontal indique le niveau d'expression relative du WT CC-5325 normalisé à 1. Les histogrammes présentent la moyenne de 3 colonies indépendantes, en 3 réplicats techniques. Toutes les valeurs obtenues sont représentées par les ronds sur les histogrammes. N=1.

et une composition protéique qui diffère considérablement de celle de tous les autres mitoribosomes connus, y compris les mitoribosomes des plantes terrestres. Dans ce mitoribosome, 47 protéines ont été identifiées dans la LSU et 36 dans la SSU. Parmi elles, 11 sont spécifiques à *C. reinhardtii* (8 pour la LSU (mL113 à mL119 et une protéine inconnue qui semble être une PPR) et 3 pour la SSU (mS105 à mS107)). A l'exception de mL119 et mS105, ces protéines sont toutes de la famille des α -solénoïdes, avec notamment une protéine mTERF, plusieurs protéines OPR et deux PPR. Des données de cryo-tomographie semblent indiquer que les protéines mL113 et mL119 ancrent le mitoribosome à la membrane mitochondriale interne, tandis que les protéines mL115 à mL118 interagissent avec des fragments d'ARNr.

En ce qui concerne les ARNr, les annotations du génome mitochondrial prédisaient 14 gènes d'ARNr, 4 pour la SSU et 10 pour la LSU (Boer and Gray, 1988; Gray and Boer, 1988). Pour la SSU, les 4 ARNr prédits ont été identifiés. Néanmoins, sur les 10 prédits comme ARNr de la LSU, seuls 9 ont été trouvés. Nous avons pu déterminer que le fragment d'ARNr L2b, initialement annoté comme étant un module ARNr de la LSU n'en faisait pas parti et sa fonction reste inconnue. Le fragment d'ARNr L3a a été identifié comme étant l'ARNr 5S, jusqu'alors considéré comme absent, bien qu'il soit hautement divergent par rapport au 5S bactérien.

Mon travail a été tout d'abord de caractériser les protéines spécifiques du mitoribosome de *C. reinhardtii*, à savoir, les protéines mL114, mL115, mL116, mL117, mL118, mL119 et mS106. Pour se faire, des mutants CLiP disponibles pour ces protéines ont été commandés. Par des PCR ciblant le locus des insertions théoriques pour chaque mutant, j'ai vérifié la présence des cassettes CIB dans les loci prédits par la banque CLiP (Figure R-52A). Les souches CLiP pour les protéines mL118 et mS106 n'ont pas d'insertion, elles ne sont donc pas mutantes. Les souches *ml114::CIB*, *ml115::CIB* et *ml116::CIB* possèdent une insertion, dans la séquence 3'UTR. *ml117::CIB* et *ml119::CIB* ont quant à elles, une insertion dans le dernier exon et dans le 6^{ème} intron, respectivement. En RT-qPCR, j'ai pu voir que *ml115::CIB* présente une diminution de l'expression du transcrit, avec 65% d'expression, lorsque cela est mesuré avec des amorces avant l'insertion (Figure R-51B). Pour les souches *ml114::CIB*, *ml116::CIB* et *ml119::CIB*, l'expression mesurée avant l'insertion semble augmentée, variant entre 179 et 300%. Enfin, seules *ml114::CIB* et *ml119::CIB* ont une expression altérée du gène cible dans les mutants, après l'insertion de la cassette, avec 12 et 2% d'expression (Figure R-52C). Deux couples d'amorces ont été testés en aval de l'insertion pour *ml115::CIB*, mais cela n'a pas

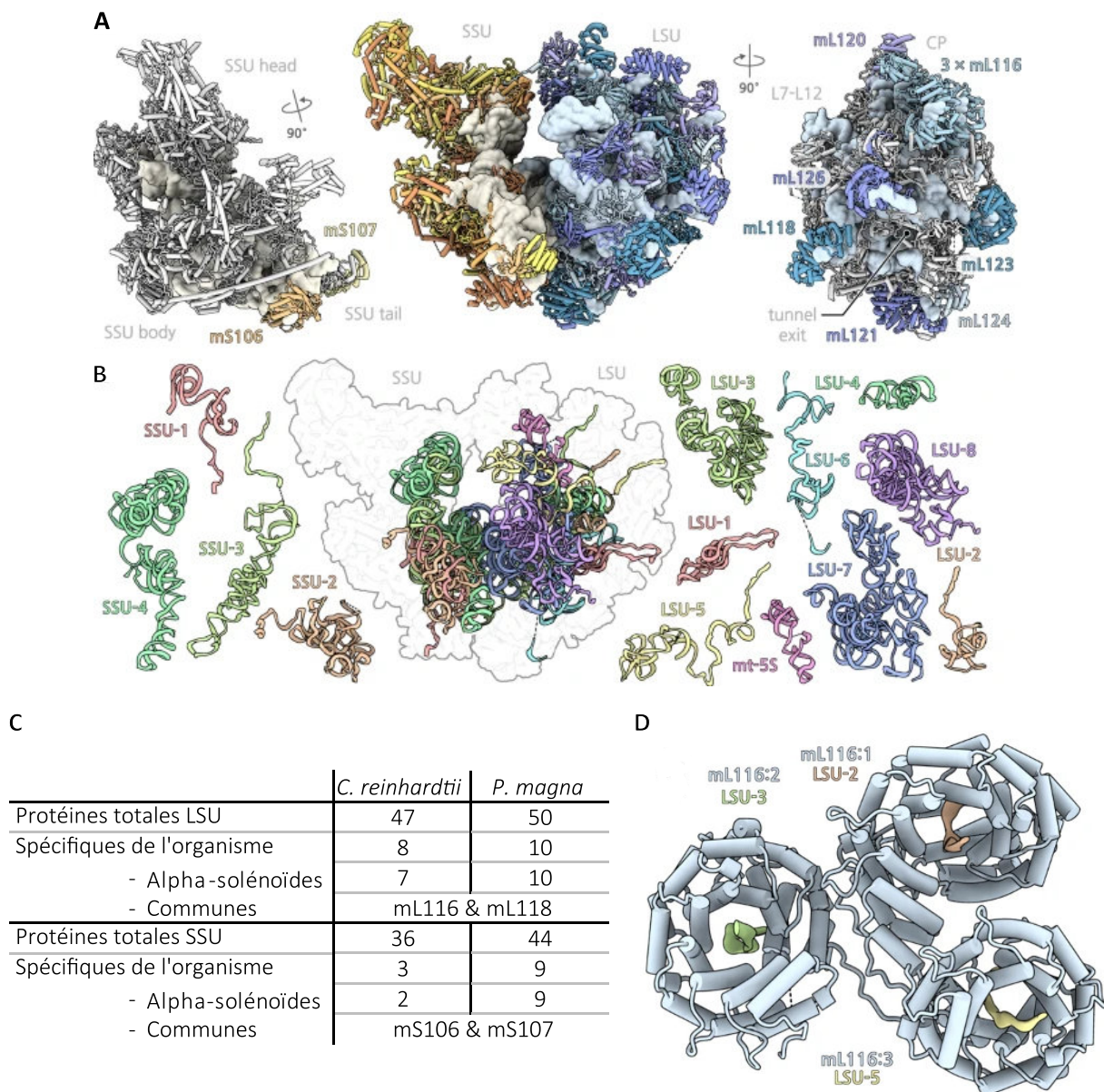


Figure R-53 : Mitoribosome de *Polytomella magna*

Adapté de (Tobiasson, Berzina and Amunts, 2022).

A : Vue d'ensemble du mitoribosome de *P. magna*, avec au centre un monosome avec les protéines de la SSU (jaune – orange) et de la LSU (bleu – violet) et sur les côtés les SSU (gauche) et LSU (droite) avec les protéines HEAT colorées. **B** : Arrangement des fragments d'ARNr assemblés (au centre) et individuels sur les côtés (SSU à gauche et LSU à droite). **C** : Tableau comparatif du nombre de protéines totales et spécifiques trouvées dans les LSU et SSU des mitoribosomes des algues *C. reinhardtii* et *P. magna*. Parmi les protéines spécifiques, sont dénombrées celles qui font partie de la famille des α -solénoïdes (OPR, PPR ou mTERF), ainsi que celles qui sont communes aux 2 mitoribosomes. **D** : Zoom sur le trimère de la protéine mL116, donc chaque copie lie un fragment d'ARNr (LSU-2/3/5).

donné de résultats exploitables en RT-qPCR. Or, sans anticorps dirigés contre ces protéines, il nous était impossible de savoir si l'insertion de la cassette en 3'UTR (*ml114::CIB*, *ml115::CIB* et *ml116::CIB*) affecte la synthèse des protéines correspondantes. C'est pour cela que ces mutants n'ont pas été d'avantage exploités et pour poursuivre l'étude j'ai généré puis caractérisé des lignées mutantes avec la technologie amiRNA, mutants présentés dans l'article.

Les mutants amiRNA ont été obtenus pour les protéines mL113, mL116, mL117, mL118 et mS105. Les analyses de RT-qPCR montrent une réduction du transcrite ciblé par l'amiRNA pour les lignées *amiml113*, *amiml116*, *amiml117*, *amiml118* et *amims105*, de 78%, 48%, 84%, 51% et 77% respectivement. Tous les mutants présentent un phénotype *dark*^{+/−}, indiquant que les mitochondries sont affectées dans les différentes lignées mutantes.

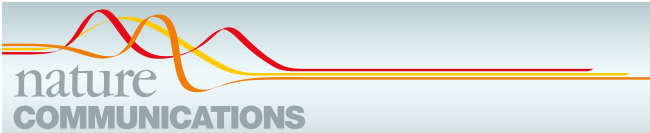
J'ai ensuite analysé l'expression des fragments d'ARNr mitochondriaux par RT-qPCR dans les différentes lignées mutantes. Ces analyses montrent que les niveaux à l'état d'équilibre des ARNr des LSU et SSU sont peu affectés dans les mutants de mL116 et de mS105. A contrario, les analyses indiquent une diminution de l'accumulation des ARNr de la LSU de 13%, 39% et 14% en moyenne pour les mutants des protéines mL113, mL117, mL118 respectivement, suggérant que ces protéines spécifiques de *C. reinhardtii* sont nécessaires à la stabilité de ces ARNr. La structure montre des interactions entre les protéines spécifiques et les ARNr et cela suggère un rôle de ces protéines dans l'assemblage des différents fragments d'ARNr.

Depuis la publication de l'article, la structure du mitoribosome d'une autre algue, *Polytomella magna*, a été résolue et publiée (Figure R-53) (Tobiasson, Berzina and Amunts, 2022). Elle fait partie des algues vertes, bien qu'elle ait perdu son chloroplaste, ce qui lui vaut l'appellation d'« algue blanche ». Comme *C. reinhardtii*, cette algue a un génome mitochondrial linéaire, qui code pour des ARNr fragmentés en 12 modules, 8 pour la LSU et 4 pour la SSU (Figure R-53B). De plus, un ARNr 5S a été identifié dans la structure du mitoribosome, bien qu'aucun gène correspondant n'ait été trouvé pour le moment. Dans ce mitoribosome, 94 protéines ont été identifiées, 50 protéines dans la LSU et 44 dans la SSU (Figure R-53C). Parmi elles, 19 sont spécifiques à *P. magna* (10 pour la LSU (mL120 à mL129) et 9 pour la SSU (mS108 à mS116)). Ces protéines sont toutes des protéines à répétitions hélicoïdales, appelées HEAT dans cette publication. Les protéines HEAT peuvent former des alpha solénoïdes, comme les OPR ou les PPR par exemple. En plus de ces 19 protéines, il y en a 4 autres de la même famille. Celles-ci font parties des 11 protéines spécifiques identifiées dans le mitoribosome de *C. reinhardtii* : mL116,

mL118, mS106 et mS107 (Figure R-53A). A l'exception de mS107 qui est une PPR, les 3 autres sont des OPR. Il est intéressant de noter que mL116 est présente en 3 exemplaires chez *P. magna*, s'associant en trimère pour former des structures rondes, en forme de donut comme mS105 chez *C. reinhardtii*, chaque copie liant un ARNr différent (L2, L3, et L5) (Figure R-53D). En tout, il y a donc 13 protéines *HEAT* dans la LSU, et 11 dans la SSU. Enfin, le mitoribosome de *P. magna* semble également ancré à la membrane, par la protéine mL105.

Pour conclure, l'étude de la structure et de la composition du mitoribosome de *C. reinhardtii* a permis l'identification de nouvelles protéines mitoribosomales, spécifiques à cette algue. Nous avons pu répondre à la question de l'assemblage et de la stabilisation des ARNr fragmenté dans le génome, ainsi que celui de l'ARNr 5s qui n'avait pas été identifié jusqu'alors. De plus, grâce à la structure du mitoribosome de *P. magna* publiée peu après, force est de constater qu'il y a des points communs entre ces 2 structures, qui sont peut-être spécifiques des algues vertes, avec par exemple un grand nombre de protéines α -solénoïdes identifiées. Cependant, des questions restent en suspens, notamment celle de l'initiation de la traduction qui n'a pas pu être résolue pour le moment.

II. Article : How to build a ribosome from RNA fragments in *Chlamydomonas* mitochondria



ARTICLE


<https://doi.org/10.1038/s41467-021-27200-z>

OPEN

How to build a ribosome from RNA fragments in *Chlamydomonas* mitochondria

Florent Waltz ^{1,2,3,8}, Thalia Salinas-Giegé ^{2,8}, Robert Englmeier⁴, Herrade Meichel², Heddy Soufari¹, Lauriane Kuhn ⁵, Stefan Pfeffer ⁶, Friedrich Förster ⁴, Benjamin D. Engel ^{3,7}, Philippe Giegé ^{2✉}, Laurence Drouard ^{2✉} & Yaser Hashem ^{1✉}

Mitochondria are the powerhouse of eukaryotic cells. They possess their own gene expression machineries where highly divergent and specialized ribosomes, named hereafter mitoribosomes, translate the few essential messenger RNAs still encoded by mitochondrial genomes. Here, we present a biochemical and structural characterization of the mitoribosome in the model green alga *Chlamydomonas reinhardtii*, as well as a functional study of some of its specific components. Single particle cryo-electron microscopy resolves how the *Chlamydomonas* mitoribosome is assembled from 13 rRNA fragments encoded by separate non-contiguous gene pieces. Additional proteins, mainly OPR, PPR and mTERF helical repeat proteins, are found in *Chlamydomonas* mitoribosome, revealing the structure of an OPR protein in complex with its RNA binding partner. Targeted amiRNA silencing indicates that these ribosomal proteins are required for mitoribosome integrity. Finally, we use cryo-electron tomography to show that *Chlamydomonas* mitoribosomes are attached to the inner mitochondrial membrane via two contact points mediated by *Chlamydomonas*-specific proteins. Our study expands our understanding of mitoribosome diversity and the various strategies these specialized molecular machines adopt for membrane tethering.

¹Institut Européen de Chimie et Biologie, U1212 Inserm, Université de Bordeaux, 2 rue R. Escarpit, 33600 Pessac, France. ²Institut de biologie moléculaire des plantes, CNRS, Université de Strasbourg, 12 rue du général Zimmer, 67084 Strasbourg, France. ³Helmholtz Pioneer Campus, Helmholtz Zentrum München, Ingolstädter Landstraße 1, 85764 Neuherberg, Germany. ⁴Structural Biochemistry, Bijvoet Centre for Biomolecular Research, Utrecht University, Universiteitsweg 99, 3584 CG Utrecht, The Netherlands. ⁵Plateforme protéomique Strasbourg Esplanade FRC1589 du CNRS, Université de Strasbourg, 67084 Strasbourg, France. ⁶Zentrum für Molekulare Biologie der Universität Heidelberg, DKFZ-ZMBH Alliance, Im Neuenheimer Feld 282, 69120 Heidelberg, Germany. ⁷Department of Chemistry, Technical University of Munich, Lichtenbergstraße 4, 85748 Garching, Germany. ⁸These authors contributed equally: Florent Waltz, Thalia Salinas-Giegé. ✉email: philippe.giege@ibmp-cnrs.unistra.fr; laurence.drouard@ibmp-cnrs.unistra.fr; yaser.hashem@inserm.fr

Mitochondria are essential organelles of eukaryotic cells that act as metabolic hubs and powerhouses, producing energy through aerobic respiration. They still possess their own genome and gene expression machineries, vestiges of their once free-living bacterium ancestor^{1,2}. Due to the evolutionary drift of eukaryotes, mitochondrial complexes involved in metabolism and gene expression combine features from their bacterial ancestor with traits that evolved in eukaryotes^{3,4}. The final step of gene expression, translation, is carried out by specialized mitochondrial ribosomes (mitoribosomes). They synthesize the few proteins still encoded by the mitochondrial genome, most of which are hydrophobic components of the respiratory chain. Despite their shared prokaryotic origin⁵, mitoribosome structure and composition were shown to be highly divergent across eukaryotes. They systematically acquired numerous additional ribosomal proteins (r-proteins), and their ribosomal RNAs (rRNAs) were either greatly reduced, like in animals and kinetoplastids^{6–9}, or expanded, like in plants and fungi^{10–13}.

Among the most prominent unresolved questions in the mitochondria biology field is the long-standing debate over the peculiar organization of the mitochondrial genome in the unicellular green alga *Chlamydomonas reinhardtii* and the biogenesis of its mitoribosome. This organism is widely used to study photosynthesis and cilia¹⁴, but it is also an excellent model to investigate mitochondrial biology. It is one of the few organisms where mitochondrial transformation is possible¹⁵, and mitochondrial mutants are viable in photoautotrophic conditions¹⁶. In contrast to vascular plants, or Viridiplantae in general, which are characterized by gene-rich and largely expanded mitochondrial (mt)-genomes, *C. reinhardtii* possess a small linear mt-genome of 16 kb. It only encodes eight proteins (all membrane-embedded components of the respiratory chain), three transfer RNAs (tRNAs) and, most intriguingly, non-contiguous pieces of the large subunit (LSU) and small subunit (SSU) ribosomal RNAs (rRNAs), scrambled across the genome^{17–19}. The mt-genome is transcribed as two polycistronic primary transcripts synthesized from opposite strands^{17,19}. Individual transcripts are then generated from the primary transcripts to produce mature, functional RNAs. When initially characterized 30 years ago¹⁷, *Chlamydomonas* mitoribosome rRNA fragmentation represented the earliest example that an rRNA does not need to be continuous in order to be functional²⁰. Although it was predicted that the rRNA fragments would somehow be integrated into a functional ribosome¹⁸, it is enigmatic how these fragments are recruited, interact with each other, and are stabilized to form the 3D mitoribosome structure.

Here, we combine cryo-electron microscopy (cryo-EM) with in situ cryo-electron tomography (cryo-ET) to resolve the structure of a green algal mitoribosome, stunningly different from both its prokaryotic ancestor, as well as from the flowering plant mitoribosome¹¹, but also from all other characterized mitoribosomes across diverse species³. Our structure reveals how the reduced and fragmented rRNAs are organized and stabilized in the mitoribosome via numerous *Chlamydomonas*-specific r-proteins. Cryo-ET resolves the native structure and organization of *Chlamydomonas* mitoribosomes inside mitochondria, revealing that these mitoribosomes are exclusively bound to the inner mitochondrial membrane. Our study provides an example of a mitoribosome composed of numerous rRNA fragments, revealing a strikingly divergent blueprint for building this conserved molecular machine.

Results

Isolation, mass spectrometry, and cryo-EM of mitoribosomes.

To analyze the *C. reinhardtii* mitoribosome, mitochondria were purified and used for mitoribosome isolation following a procedure based on sucrose density gradient separation (see Methods)

(Supplementary Fig. 1). Collected fractions were systematically analyzed by nano-LC MS/MS (Supplementary Table 1) and screened by cryo-EM to determine their composition. This approach allowed us to identify fractions containing the two mitoribosome subunits, which were subsequently used for data collection (Supplementary Fig. 1). Proteomic analysis identified putative *Chlamydomonas*-specific r-proteins that were then confirmed by the corresponding cryo-EM reconstructions. Following image processing and extensive particle sorting, reconstructions of both dissociated subunits were obtained. The large subunit (LSU) was resolved to 2.9 Å, while the small subunit (SSU) was reconstructed at 5.49 Å and further refined to 4.19 Å for the body and 4.47 Å for the head using a focused refinement approach (Supplementary Fig. 2). Fully assembled mitoribosomes were identified by nano-LC MS/MS in the cytoribosome fraction (Supplementary Fig. 1), but cryo-EM investigation revealed aggregates in this fraction, most likely corresponding to mitoribosomes. Nevertheless, the individual subunit reconstructions were docked into the map of the entire *C. reinhardtii* mitoribosome obtained from subtomogram averaging of the in situ cryo-ET data (see below), allowing accurate positioning of both subunits relative to each other in the context of a fully assembled native mitoribosome. The isolated subunit reconstructions were similar to the in situ subtomogram average, demonstrating that they represent the mature LSU and SSU and not assembly intermediates. Notably, all densities corresponding to *Chlamydomonas*-specific r-proteins were present in both single particle and subtomogram average reconstructions.

Overall structure of the *Chlamydomonas* mitoribosome. Our cryo-EM reconstructions, along with our extensive MS/MS analyses, allowed us to build atomic models of both *C. reinhardtii* mitoribosome subunits (see “Methods” section) (Figs. 1 and 2). The overall architecture of this mitoribosome (Fig. 1) is clearly distinct from both its bacterial ancestor and the flowering plant mitoribosome¹¹. *Chlamydomonas*-specific proteins and domains largely reshape both subunits. Similar to all previously described mitoribosomes, the *Chlamydomonas* mitoribosome has more r-proteins compared to its bacterial counterpart^{3,4}. These proteins include ancestral r-proteins conserved with bacteria, mitoribosome-specific r-proteins shared with other mitoribosomes, and *Chlamydomonas*-specific r-proteins. In total, the *Chlamydomonas* mitoribosome contains 47 r-proteins in the LSU and 36 in the SSU. These include 11 new r-proteins (not accounting for unknown densities), 8 in the LSU and 3 in the SSU. The total of 83 r-proteins greatly exceeds the 54 r-proteins in bacterial ribosomes (Fig. 2 and Supplementary Table 1). As a result, very few rRNAs are exposed to the solvent, with proteins coating the entire mitoribosome and stabilizing the fragmented rRNAs (Supplementary Movie 1). Proteins follow the classical r-protein nomenclature²¹, and newly identified proteins are numbered according to the last inventory of mitoribosomal r-proteins²².

Reconstruction of the LSU (Fig. 1d) revealed eight additional r-proteins, named mL113 to mL119, plus PPR*, a putative Pentatricopeptide Repeat (PPR) protein (Figs. 1, 2 and Supplementary Fig. 4, Supplementary Table 1). They are distributed across the whole LSU, where they extend into the solvent and are anchored to the ribosome by interacting with both conserved r-proteins and rRNA fragments. With the exception of mL119 at the exit of the peptide channel, all these proteins are relatively large RNA binders composed of repeated alpha-helical folds, including a mitochondrial TERmination Factor (mTERF) protein, several Octotricopeptide Repeat (OPR) proteins, and PPR*.

The small subunit (Fig. 1c) reconstruction highlights several distinctive features. Most strikingly, the SSU is shaped by two

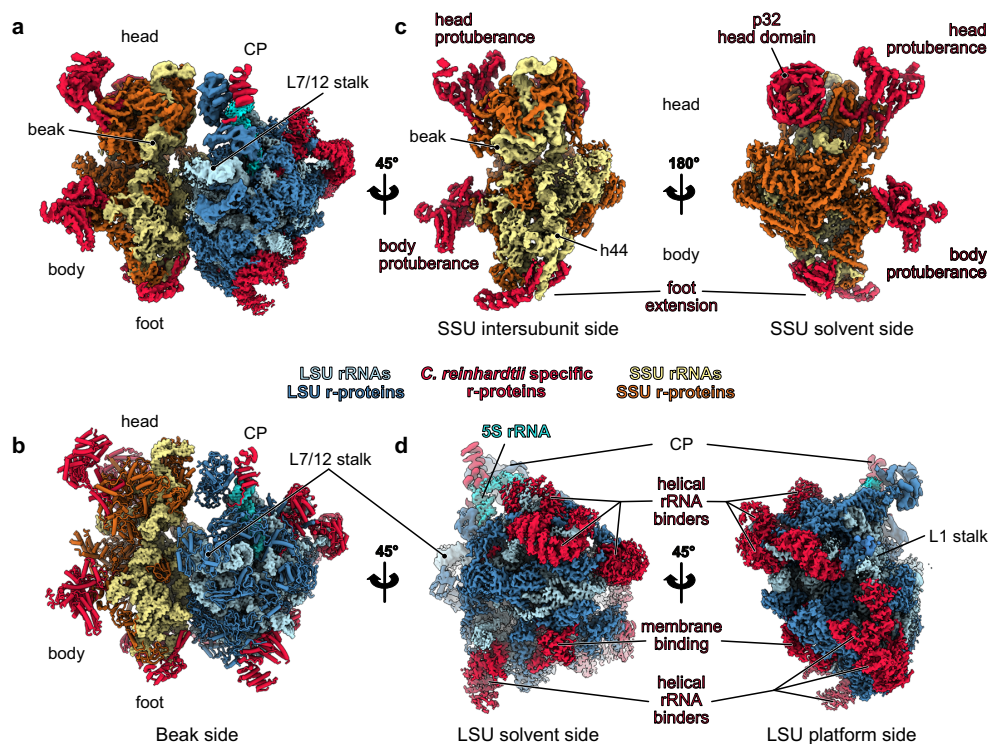


Fig. 1 Overall structure of the *Chlamydomonas* mitochondrial ribosome. **a** Composite cryo-EM map of the *Chlamydomonas reinhardtii* mitochondrial ribosome and **b** the resulting atomic model. The large subunit (LSU) components are depicted in blue shades, the small subunit (SSU) components are shown in yellow shades, and the specific r-proteins and domains are displayed in red. **c**, **d** Different views of the cryo-EM reconstructions of the SSU (**c**) and the LSU (**d**).

large protuberances positioned on its beak side, one on the head and one on the body, both formed by helical-rich proteins. The body protuberance, located close to the mRNA entrance, is mainly formed by a specific extension of more than 400 aa in the mitochondria-specific r-protein mS45 (Supplementary Fig. 5), a highly variable r-protein²². The head protuberance density could not be assigned due to the low resolution of this area. However, several conserved proteins of the SSU head located nearby, notably uS3m, uS10m, and mS35, present large extensions. Therefore, it is likely that these extensions could come together and form the head protuberance, as no other apparent candidates could be identified by MS/MS analyses. On the solvent side of the head, a large torus-shaped domain protrudes in the solvent. This additional domain is a homotrimeric complex formed by three copies of mS105, also called p32 (Supplementary Fig. 5a). This MAM33-family protein is seemingly conserved in all eukaryotes and has been described to have several functions in mitochondria, some related to mitoribosome assembly^{23–26}. However, mS105/p32 was never before reported as a core component of a ribosome. Additionally, the head of the SSU is characterized by its missing beak, which is typically formed by helix 33 at the junction site of rRNA fragments S3 and S4 (Fig. 3c and Supplementary Fig. 7). The foot of the SSU is reshaped by *Chlamydomonas*-specific r-proteins. The extension is formed by two super-helical proteins, one PPR (mS106) and one OPR (mS107) identified by MS/MS and confirmed by AlphaFold²⁷. The mS106 protein occupies a position similar to mS27 in humans^{6,7} and fungi¹², but it does not appear to interact with RNA, nor does it share any sequence identity with mS27 (Supplementary Fig. 5d). On the

other hand, the OPR mS107 directly interacts with rRNA fragment S2, where it encapsulates the tip of helix 11 (Supplementary Fig. 5c).

Fragmented ribosomal RNAs are assembled to reconstitute the core of *Chlamydomonas* mitoribosome. In contrast to flowering plants, where rRNAs are largely expanded, the *C. reinhardtii* mitoribosome is characterized by its reduced and fragmented rRNAs (Fig. 3). These rRNAs are scrambled in the mitochondrial genome (Fig. 3a), where they are expressed as a single polycistron that is then further processed into matured transcripts by currently unknown endonucleases¹⁷. The “23S” and “16S” rRNAs are respectively split into eight fragments totaling 2035 nt, and four fragments totaling 1200 nt (Fig. 3c and Supplementary Figs. 6, 7, Supplementary Supplementary Movie 1). This corresponds to 30% and 22% reductions compared to bacteria (Fig. 3f). Among all the rRNA pieces predicted to be integrated into the mature mitoribosome, all but one (L2b) could be identified in our cryo-EM reconstructions. To confirm the absence of the L2b fragment, we performed comparative RNAseq analyses of mitochondrial and mitoribosomal fractions (Fig. 3b). Consistently, all rRNA fragments could be identified in the mitoribosome fraction except L2b. However, this analysis confirmed that the fragment is indeed expressed and accumulates in the purified mitochondria fraction (Fig. 3b), which is in line with previous transcriptomic analyses^{19,28,29}. These results were further confirmed by RNA blots hybridized against L2b and an L2a control found in the ribosome and stabilized by the r-protein mL116 (Fig. 3d). Therefore, the L2b RNA is not associated with the mitoribosome,

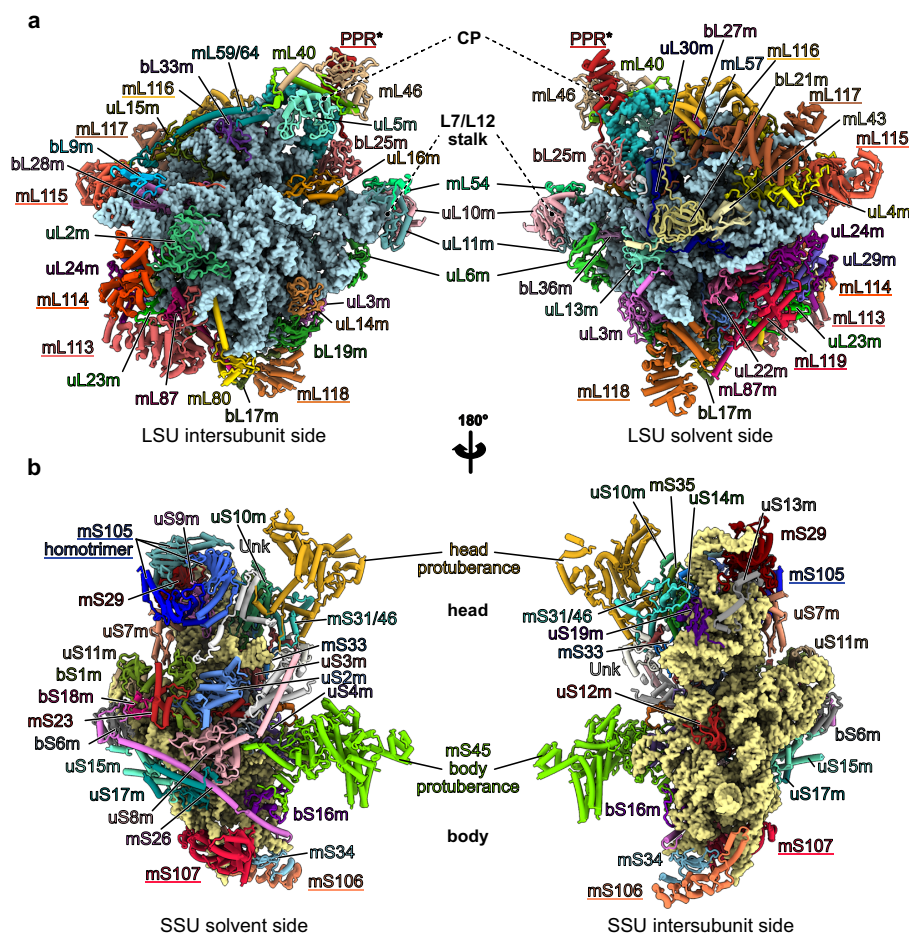


Fig. 2 Ribosomal proteins of the *Chlamydomonas* mitoribosome. **a, b** The atomic models of the *Chlamydomonas* mitoribosome LSU (**a**) and SSU (**b**), with mitoribosomal proteins shown in cartoon representation and individually colored and annotated. rRNAs are shown in surface representation and colored in light blue (LSU) and beige (SSU). Newly identified r-proteins are highlighted with underlined names.

suggesting that this small RNA has an independent function that remains to be elucidated.

In the SSU, the fragmented rRNAs form only a few interactions with the additional r-proteins and are mainly stabilized by base-pairing with each other (with the exception the S2 fragment's h11), which is encapsulated by mS107 (Supplementary Figs. 5 and 7). Fragments S1, S2, and a small portion of S3 form the 5' domain, with the rest of S3 making up domain C. The 3' end of S3 and the entirety of S4 constitute domains 3'M and 3'm, with S4 largely contributing to linking the head and body of the SSU. The region most conserved with bacteria is the decoding center, made by h1-2 and h27-28 (Fig. 3 and Supplementary Fig. 7).

In contrast to the SSU, the nine rRNA fragments in the LSU are all stabilized by the newly identified *Chlamydomonas*-specific r-proteins. These fragments reconstitute the different domains of the large subunit. L1 forms the highly reduced domain I of the LSU. L2a, L3b, L4, L5, and part of L6 together form domain II. Portions of L6 and L7 form the highly reduced—almost deleted—domain III. Fragments L7 and L8, the largest of all, make up domains IV, V, and VI, which form the catalytic core of the ribosome. These three domains are the least altered, with only a few helices missing and two expansion segments ES-66 and ES-94 (Fig. 4 and Supplementary

Fig. 6). The peptidyl transferase center (PTC) formed by H89 to H93 is particularly conserved with bacteria. The conservation of these domains is most likely due to the high selective pressure to conserve the catalytic region of the ribosome³⁰. These rRNA fragments are held together by base-pairing with each other, and their extremities are stabilized by base-pairing with other fragments, e.g., L2a, L3b, and L4 (Fig. 3e) or with themselves. These results corroborate the initial predictions made 30 years ago by Boer and Gray¹⁷. However, several single-stranded rRNA extremities are also stabilized by the *Chlamydomonas*-specific r-proteins (see below). Surprisingly, while it was anticipated that 5S rRNA should be absent from *Chlamydomonas* mitoribosome, we identified an RNA density at the typical position of the 5S rRNA in the central protuberance (CP) (Fig. 1 and Supplementary Fig. 8). This rRNA density could be attributed to the L3a rRNA fragment (Supplementary Fig. 8). We further confirmed its association with the mitoribosome by comparative RNAseq analysis of mitochondrial and mitoribosomal fractions (Fig. 3b). Previous studies of the *C. reinhardtii* mitochondrial genome and rRNAs always failed to identify this rRNA as a putative 5S^{17-19,31}. While L3a likely derives from an ancestral bacterial 5S rRNA, it has highly diverged; very little of the primary sequence is conserved with other 5S rRNAs (Supplementary Fig. 8d),

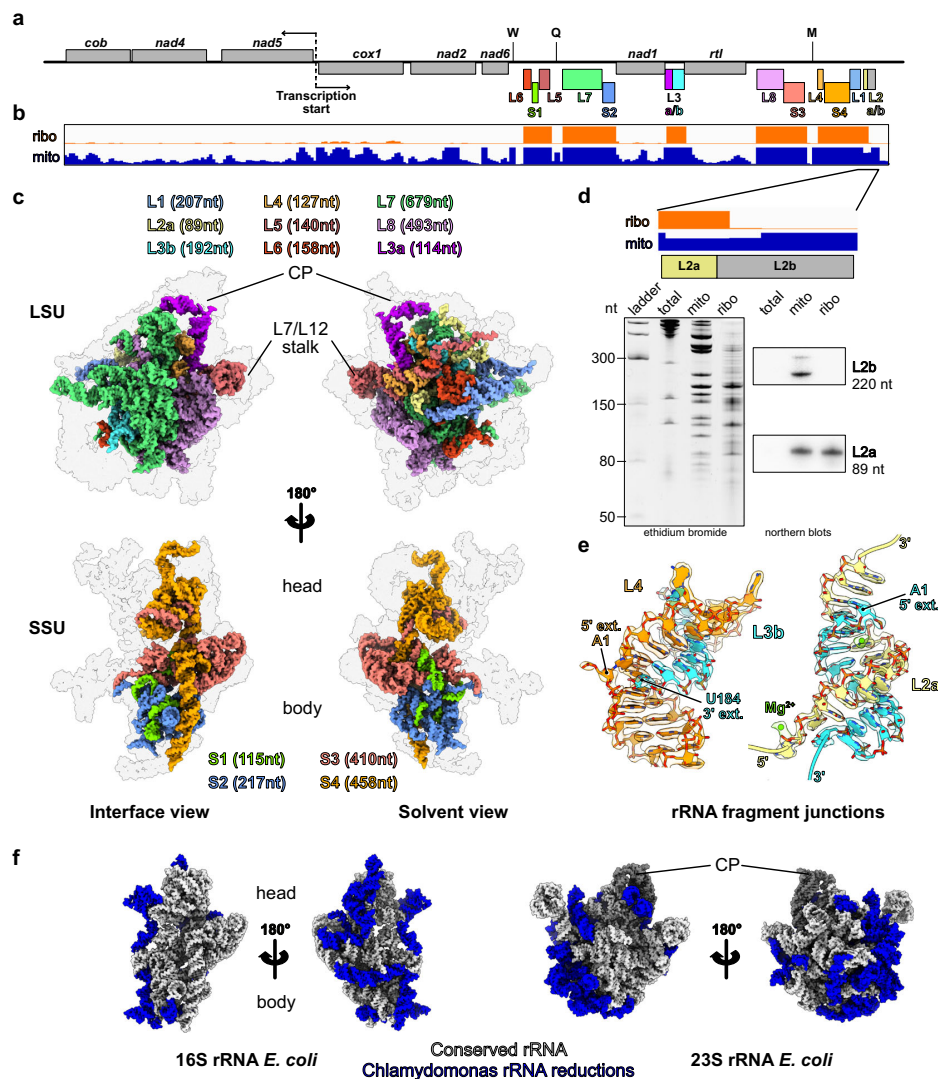


Fig. 3 The ribosomal RNAs of the *Chlamydomonas* mitoribosome are fragmented. **a** Schematic representation of the entire *C. reinhardtii* mitochondrial genome. The protein-coding genes are displayed in gray, and the rRNA fragments incorporated in the mitoribosome are individually colored. The tRNA genes are indicated by letters. **b** Browser view of the RNAseq data of libraries built from purified mitochondria (blue) and purified mitoribosome (orange) fractions. Data range is [0-700], mapped to the *Chlamydomonas* mitochondrial genome in **a**. **c** rRNA fragment 3D organization in the LSU and the SSU. The r-proteins are shown as a gray silhouette. The fragment colors match the color code used in **a**. **d** Zoomed view of the L2a/b coverage (top). The L2b fragment is absent in the ribosomal fraction in the RNAseq analysis, which was confirmed by northern blots (bottom) and the cryo-EM reconstruction. Northern blot was performed for the L2a and L2b fragments on total cell extracts (total), mitochondrial (mito), and mitoribosomal fraction (ribo) on one biological sample. **e** Detailed view of the rRNA extremities of the L3b fragment and its pairing with the L2a and L4 fragments. The atomic models are shown mapped into the cryo-EM densities. **f** Comparison of the *Chlamydomonas* mitoribosome rRNAs with the *E. coli* ribosome. rRNA reductions are shown in blue.

but a consensus of 6 consecutive nucleotides confirmed its origin. Its overall structure is also weakly conserved, with only domain γ retaining its characteristic structure to interact with H38. Domain β is angled differently relative to the domain γ stem, which allows the interaction of the terminal loop of domain β with H87, in contrast to other known ribosome structures. Additionally, domain α could not be fully resolved, but most likely interacts with the putative PPR protein (labeled “PPR*”), possibly stabilizing its 3' and 5' termini. In conclusion, when compared with the flowering plant

mitoribosome¹¹, the overall structure of the CP appears similar. However, in terms of composition, the *Chlamydomonas* mitoribosome CP includes a divergent 5S, lacks the uL18m protein, and contains an additional protein, PPR*.

Specific r-proteins stabilize the rRNAs by highly intertwined protein-RNA interactions. In the LSU, all the *Chlamydomonas*-specific r-proteins except mL119 are predicted to be nucleic-acid

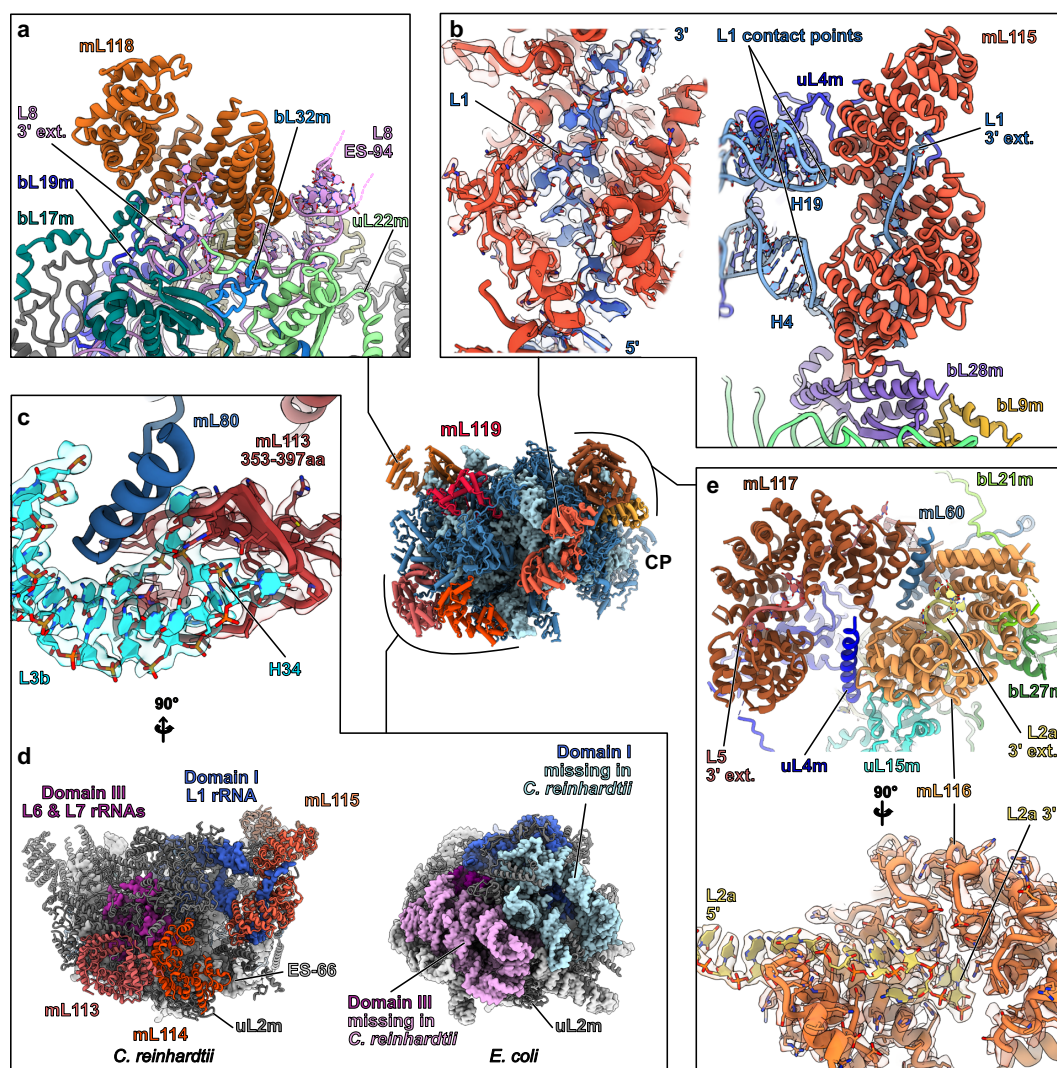


Fig. 4 *Chlamydomonas*-specific proteins stabilize the fragmented rRNAs via highly intertwined interactions. Magnified views of the *Chlamydomonas*-specific r-proteins involved in rRNA stabilization. **a** mL118 stabilizes the 3' end of the L8 fragment and contacts the expansion segment 94 (L8 ES-94). **b** mL115 stabilizes the highly reduced domain I, which is formed by the L1 fragment. mL115 binds the 3' end of the L1 fragment and also contacts the L1 fragment at two specific points corresponding to H4 and H19. The single-stranded portion of L1 interacting with mL115 is shown in its density. **c** Detailed view of the mL113 contact with the L3b fragment. An inter-repeat domain of mL113 (red) formed by amino acids 353–397 clamps the tip of H34 (cyan). The models are shown in their densities. Contrary to the rest of the r-proteins, mL113 does not enlase single-stranded rRNA. **d** Structural compensation for the loss of large portions of domain I and III. The missing rRNA regions are depicted on the *E. coli* model in pink and light blue, and the compensating proteins mL113, mL114, and mL115 (red shades) are shown on the *Chlamydomonas* model. mL113 and mL114 compensate for domain III reduction, and mL115 stabilizes and compensates for the reduced domain I. **e** mL116 and mL117 interact with each other and with several surrounding proteins. mL117 is involved in stabilizing the 3' end of the L5 rRNA fragment, and mL116 stabilizes the L2a 3' extremity, but also makes additional contacts with rRNAs (Supplementary Fig. 4). The single-stranded portion of L2a in mL116 is shown in its density. With the exception of the mTERF protein mL114, the rest of the proteins belong to the same class of ASA2-like/OPR proteins. Further detailed views are shown in Supplementary Fig. 4.

binders. mL114 is an mTERF protein, but mL113, mL115, mL116, mL117, and mL118 all appear to belong to the same protein family, as they all have similar tertiary structures and resemble ASA2/OPR proteins. These OPR proteins (Octatricopeptide Repeat), are predicted to fold into repeated pairs of α -helices, forming a super-helical solenoid, similar to PPR (Pentatricopeptide Repeat) and TPR (Tetratricopeptide Repeat)

proteins^{32,33}. Both PPR and TPR are widespread in eukaryotes and have previously been found in mitribosomes, notably in the flowering plant mitribosome, which includes 8 PPR proteins (rPPR proteins) that stabilize the numerous rRNA expansions¹¹. In *Chlamydomonas*, OPR proteins were previously described to be involved in gene expression regulation, notably in the chloroplast^{32,34–37}. In the mitribosome, these proteins stabilize

the many rRNA fragments by different modes of RNA interaction. This is, to our knowledge, the sole structural description of this kind of protein in interaction with RNA. Proteins mL115, mL116 and mL117 stabilize the 3' extremities of L1, L2a, and L5, respectively, by binding the single-stranded rRNA fragments in their inner groove (Fig. 4). The stabilization is primarily mediated by positive/negative charge interactions, where the inner grooves of the proteins are largely positively charged, filled with lysine and arginine that interact with the negatively charged phosphate backbone of the RNA (Supplementary Fig. 4). Unlike the rest of the RNA binders, mL113 does not interact with RNA in its inner groove. An inter-repeat domain formed by amino acids 353–397 clamps the tip of H34 from the L3b fragment (Fig. 4c and Supplementary Fig. 4a). Here, the interaction does not stabilize the rRNA itself, but rather constitutes an anchor point between mL113 and the ribosome. mL118 acts similarly to the SSU's OPR mS107, as it binds the 3' extremity of the L8 fragment, which forms a loop inside the inner groove of the protein (Fig. 4a). Additionally, these proteins also interact with RNA via the convex side of their super-helical fold. This is the case for mL115, which interacts with H4 and H19 of the L1 fragment. mL116 interacts with H28 of the L2a fragment and H38 formed by the L3b/L4 duplex, while mL118 interacts with ES-94 of the L8 fragment (Supplementary Fig. 4). Moreover, the mTERF protein mL114 stabilizes the ES-66 via its C-terminal region, not its inner groove (Supplementary Fig. 4a). All these proteins largely interact with conserved proteins as well as with each other, e.g., mL113 with mL114 and mL116 with mL117 (Fig. 4). Furthermore, mL113, mL114, and mL115 structurally compensate for missing rRNA on the back side of the LSU (Fig. 4d). mL113 and mL114 compensate for the almost wholly deleted domain III, while mL115 both stabilizes and compensates for the missing parts of domain I. Interestingly, mL113 and mL114 are similarly positioned compared to mL101 and mL104 of flowering plants, which stabilize the remodeled domain III¹¹.

Knockdown of *Chlamydomonas*-specific r-proteins affects cell fitness and rRNA stability. Next, we used targeted gene silencing to investigate the importance of the *Chlamydomonas*-specific r-proteins for ribosome integrity. We explored the *Chlamydomonas* CLiP mutant library³⁸, but no mutant strains for the genes of interest could be confirmed. Hence, we generated artificial miRNA (amiRNA) strains for these factors. This method reduces targeted protein expression at the transcript level³⁹. Strains were generated for mL113, mL116, mL117, mL118, and mS105 (p32). The physiological phenotype of each amiRNA strain was analyzed, particularly the capacity to grow under heterotrophic conditions (dark + acetate), which is typically defective in *Chlamydomonas* mutants impaired in mitochondrial respiration¹⁶. Some transformants revealed growth retardation when cultivated in heterotrophic conditions, and the ones presenting the most severe phenotypes were selected (Fig. 5a). The expression of targeted mRNAs was monitored by quantitative RT-PCR, showing reductions of 80%, 49%, 82%, 51%, and 83% on average for mL113, mL116, mL117, mL118, and mS105, respectively (Fig. 5b).

The levels of rRNA fragments in these downregulated amiRNA strains were then monitored by quantitative RT-PCR to determine the effect on mitochondrial rRNAs stability. This analysis showed that the overall levels of LSU rRNAs decreased by 13%, 40%, and 14% in the mL113, mL117, mL118 knockdown strains, respectively, while the level of L2b RNA, which is not present in mitoribosomes, followed a different behavior (Fig. 5c). In contrast, relative rRNA levels were not significantly affected in the mL116 and mS105 knockdown

strains, with the exception of L2b, which was reduced to about 67% in mS105. Altogether, it appears that *Chlamydomonas*-specific r-proteins, in particular mL113, mL117, and mL118, are required for the proper stability of the LSU rRNAs. In addition, the accumulation of mitochondria-encoded proteins was investigated by protein immunoblots. Two mitochondrial-encoded components of respiratory complex I, Nad4 and Nad6, were analyzed alongside two controls, the nuclear-encoded subunit NUO7 of complex I and the mitochondrial porin VDAC (Fig. 5d). Analysis from 3 to 4 technical replicates from 2 biological replicates showed that Nad4 and Nad6 levels were decreased in the mL113, mL117, and mS105-mutant strains compared to wild-type; in contrast, NUO7 and VDAC had unchanged levels (Supplementary Fig. 9 and Source Data file). Finally, the accumulation of assembled respiratory complexes, which contain mitochondria-encoded proteins, was investigated by blue native PAGE (BN-PAGE) coupled to in-gel activity assays (Fig. 5e). These tests revealed that the mL113 strain is impaired in complex I and IV activity, whereas the mL117 strain also appears to be affected but to a lesser extent, which correlates with the immunoblot assays. Collectively, these analyses show different impacts on the knockdown strains, suggesting non-redundant functions for these r-proteins.

The *Chlamydomonas* mitoribosome is tethered to the inner mitochondrial membrane via two protein contact sites.

The LSU reconstruction revealed the presence of a specific r-protein, mL119, precisely located at the exit of the peptide channel (Fig. 6). There, this protein forms several contacts with r-proteins uL22m, uL24m, uL29m, and bL32m, and with nucleotides 114–124 of the L6 fragment via its C-terminal part, which anchors the protein to the ribosome (Fig. 6e). The N-terminal part of mL119, which constitutes most of the protein's mass, is exposed to the solvent. This protein has no apparent homolog and appears to be restricted to the Chlorophyceae (green alga) lineage. Similar to humans, the *Chlamydomonas* mitochondrial genome only codes for membrane components of the respiratory chain, except for the *rtl* gene, whose expression and function remain uncertain. In humans and yeast, it was previously shown that mitoribosomes contact the membrane protein insertase Oxal1 via r-protein mL45 in human and the linker protein Mba1 in yeast^{40–44}. These two homologous proteins are positioned at the exit of the peptide channel, where they link the ribosome to the membrane by binding Oxal1, allowing direct insertion of nascent proteins into the membrane. Given the position of mL119, one would expect this protein to fulfill a function similar to mL45 and Mba1. To assess the role of mL119 in membrane binding, *Chlamydomonas* mitoribosomes were directly visualized inside cells using in situ cryo-ET (Fig. 6 and Supplementary Movie 2).

Whole *Chlamydomonas reinhardtii* cells were vitrified, thinned by cryo-focused ion beam (FIB) milling, and imaged by cryo-ET. A representative tomogram depicting a section of a native mitochondrion within a *Chlamydomonas* cell is shown in Fig. 6a–d. ATP synthase dimers, cytosolic ribosomes, and mitoribosomes were automatically localized by template matching, structurally resolved by subtomogram averaging, and then mapped back into the native cellular environment (Fig. 6b). Contrary to cytosolic ribosomes, which crowd the cytoplasm, mitoribosomes have a very low abundance and are localized to the inner mitochondrial membrane. Their low copy number and membrane association highlight the difficulty of purifying these complexes compared to cytosolic ribosomes. Alignment of subtomograms containing mitoribosomes yielded a structure of the native membrane-bound mitoribosome at 31.5 Å resolution (Supplementary Fig. 3). With the exception of dynamic flexible

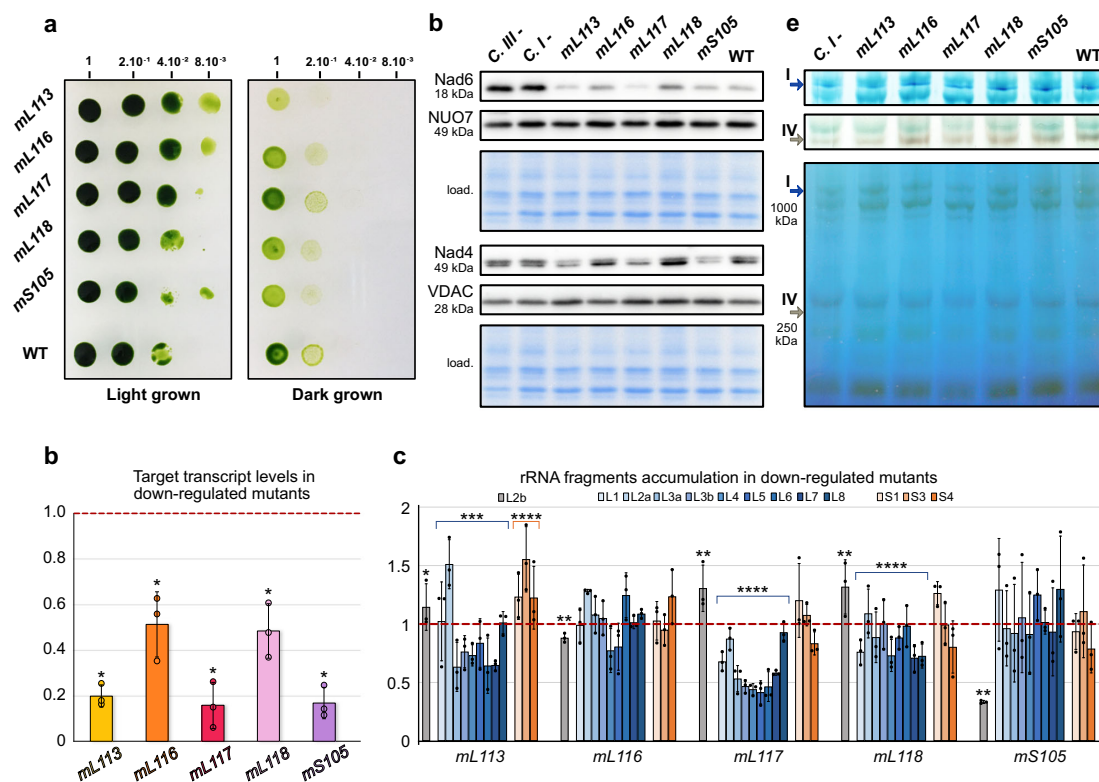


Fig. 5 Downregulation of *Chlamydomonas*-specific r-proteins affects fitness, rRNA accumulation, and mitochondrial proteins synthesis. **a** Growth phenotype of amiRNA knockdown strains. After a first-round selection, transformants were obtained for *mL113*, *mL116*, *mL117*, *mL118*, and *mS105*. Growth phenotypes in the dark were investigated by 5-fold dilution series. Dilutions were spotted on two identical TAP plates, one placed in light and the other in the darkness. **b** Relative levels of amiRNA-targeted mRNAs investigated by RT-qPCR. All strains show clear downregulation compared to WT. Data are presented as mean values \pm SD from $n = 3$ biological independent experiments, each analyzed in three technical replicates. The statistical difference between WT and each amiRNA strain for each signal protein was calculated with the two-tailed Mann-Whitney test ($*p < 0.05$). Statistics and the exact p -values are detailed in the Source Data file. **c** Relative steady-state levels of rRNA fragments in the different *Chlamydomonas*-specific r-protein amiRNA strains. rRNA fragments of the LSU are shown in blue shades, rRNA fragments of the SSU are shown in orange shades, and L2b (not present in the ribosome) is gray. Data are presented as mean values \pm SD from $n = 3$ biological independent experiments, each analyzed in three technical replicates. For statistical analysis, the rRNAs were divided into three groups: the LSU rRNA fragments, the SSU rRNA fragments, and the L2b rRNA fragment. The statistical differences between WT and amiRNA strain for each group was calculated with the two-tailed Mann-Whitney test ($****p < 0.0005$; $***p < 0.005$; $**p < 0.05$; $*p < 0.5$). Statistics and the exact p -values are detailed in the Source Data file. For **b** and **c**, the WT levels are normalized to 1 on the y axis. **d** Steady-state levels of mitochondria-encoded Nad6 and Nad4, as well as nuclear-encoded NUO7 and the mitochondrial porin VDAC, are shown by protein immunoblots corresponding to an example of a typical result from $n = 2$ biological replicates performed in 3–4 technical replicates (Supplementary Fig. 9 and Source Data file). **e** Blue native PAGE and in-gel activity assays for complexes I and IV. Arrows indicate the resulting bands. The activity assay for complex I corresponds to a typical result from two independent experiments. The activity assay for complex IV corresponds to the result of one biological experiment. For **d** and **e**, Wild-type (WT), complex I-mutant *dum5* (C I-), and complex III-mutant *dum11* (C III-) were used as controls.

regions (e.g., L7/L12 and L1 stalks), the in situ subtomogram average is highly similar to our single-particle reconstructions, as revealed by molecular fitting. This structural agreement confirms that the single-particle reconstructions very likely correspond to mature forms of the mitoribosome subunits (Fig. 6f). The in situ subtomogram average had one additional density located at the mitoribosome's mRNA exit channel. Although we do not know the identity of this density, we speculate that it may correspond to exiting mRNAs or the recruitment of additional factors during active translation (Fig. 6g).

The in situ subtomogram average reveals how the mitoribosome is tethered to the inner mitochondrial membrane. Membrane-bound mitoribosomes were previously described by cryo-ET of mitochondria isolated from yeast⁴⁵ and humans⁴⁶.

Similar to yeast, but not humans, the *Chlamydomonas* mitoribosome makes two distinct contacts with the membrane. Superposition with the atomic model reveals that one contact is located at the precise position of mL119, supporting the hypothesis that this protein could directly interact with the ribosome binding domain of Oxal1 in vivo. Therefore, it appears that mL119 constitutes a functional analog of mL45 and Mba1. However, mL119 and mL45/Mba1 are not evolutionary related, but rather appear to have convergently evolved to fulfill the same function. The mitoribosome's second membrane contact is mediated via the C-terminal part of mL113 (Fig. 6f). This region of mL113 had poorly resolved density in our cryo-EM map and was thus not modeled, but we could still observe its position at low resolution (Supplementary Fig. 4a). The mL113 contact mimics the rRNA

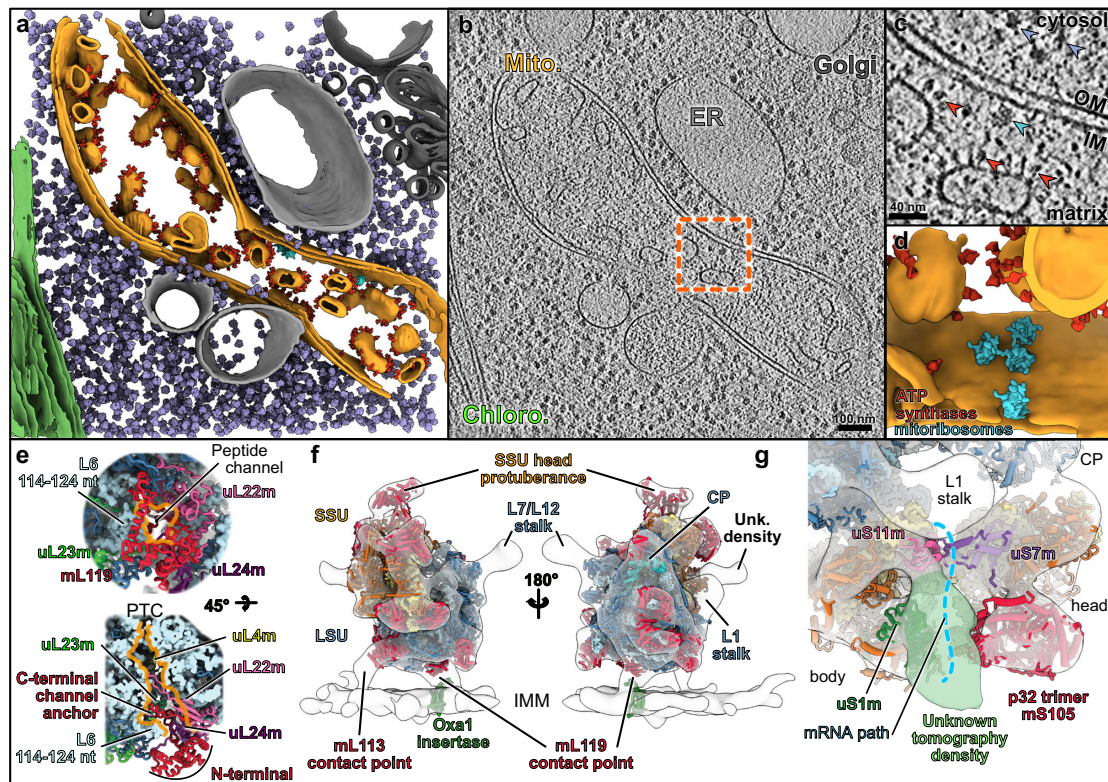


Fig. 6 The mitoribosome is attached to the mitochondrial inner membrane via specific proteins. **a** Segmentation of an in situ cryo-tomogram, depicting a mitochondrion within a native *C. reinhardtii* cell. The mitochondrion is shown in orange, chloroplast in green, ER and peroxisome in light gray, and Golgi in dark gray. Subtomogram averages of ATP synthase dimers (red), cytosolic ribosomes (purple) and mitoribosomes (cyan) are mapped into the volume. **b** A slice through the corresponding raw tomogram (one representative tomogram from $n = 47$ in total). **c** Close-up view of a membrane-bound mitoribosome at the inner membrane (IM), boxed in **b**. Arrowheads point to ATP synthase, cytosolic ribosomes, and a mitoribosome. **d** Close-up view of the segmented area presented in **c**, highlighting a cluster of three mitoribosomes that possibly form a polysome. Scale bars are indicated on **b** and **c**. **e** The peptide channel of the *Chlamydomonas* LSU from a solvent view (top) and as well as a cut view (bottom). The mL119 protein is located at the exit of the peptide channel. The C-terminal part of mL119 anchors the protein to the ribosome, while the N-terminal part is exposed to the solvent where it could interact with Oxa1. Orange lines delimit the peptide channel. Ribosomal proteins are depicted as cartoons, and rRNAs are depicted as surface representations in light blue. **f** The atomic model of the mitoribosome is fitted into the subtomogram average density of the native membrane-bound mitoribosome. The mitoribosome contacts the inner mitochondrial membrane (IMM) at two specific points. These contacts are mediated by two *Chlamydomonas*-specific proteins, mL113 and mL119, the latter of which is located at the exit of the peptide channel, adjacent to the Oxa1 insertase. The model of the bacterial homolog of Oxa1 is shown for illustration purposes. **g** Close-up view of the mRNA exit channel. The molecular model is fitted into the subtomogram average, highlighting the presence of an additional unknown density (green) located close to uS1m.

expansion segment ES-96 that forms the second contact site with the membrane in yeast⁴⁵.

Discussion

Our study describes the structure and composition of the *Chlamydomonas* mitochondrial ribosome. The cryo-EM reconstructions show that this green alga mitoribosome differs significantly from prokaryotic ribosomes as well as its flowering plant counterpart^{47,48}. In both the small and the large subunits, the mitoribosome has acquired several additional r-proteins that significantly reshape its overall architecture. These specific r-proteins combined with the fragmented rRNAs (four pieces in the SSU and nine pieces in the LSU) constitute an extreme case of ribosome divergence, even among the exceptionally diverse mitoribosomes.

One striking feature of the SSU is the presence of two large protuberances on the head and the body. The body protuberance

could be assigned to a large insertion in the mitoribosome-specific protein mS45 (Supplementary Fig. 5b), which shows high structural variability between mitoribosomes in different species despite the conservation of its core domain²². The head protuberance was poorly resolved in our density map but is most likely composed of long extensions of the head's conserved r-proteins (uS3m, uS10m, mS35). In vascular plant mitoribosomes, the uS3m r-protein was shown to form a similar large protuberance on the SSU head, suggesting a common origin of these protrusions¹¹. The roles of these two protuberances are unknown. The position of the body protuberance, close to the mRNA entry channel, might suggest a species-specific mechanism of mRNA recruitment, similar to that mediated by mS39 in humans^{4,41,46,49}. This protuberance, in conjunction with the additional density observed in the subtomogram average next to bS1m at the mRNA exit channel (Fig. 6g), might point to the existence of specific translation processes in *Chlamydomonas*.

Translation initiation in *Chlamydomonas* mitochondria shares some features with that of human mitochondria, as the mRNAs lack 5' untranslated regions in both organisms. However, *Chlamydomonas* most probably has a specific mechanism for translation initiation, as its mitochondrial mRNAs do not have the U-rich motif downstream of the AUG that was proposed to interact with mS39 in humans. Furthermore, mature *Chlamydomonas* mRNAs have poly-C rich 3' tails that might be required for translation initiation¹⁹.

Another key feature of the SSU is the homotrimeric mS105 (p32) protein forming a torus-shaped protuberance on the back of the SSU head, reminiscent of RACK1 on the cytosolic ribosome⁵⁰. This protein belongs to the MAM33 family, which appears to be eukaryote-specific and is characterized by its quaternary structure: a doughnut-shaped trimer that is highly negatively charged^{26,51}. The p32 protein has been the subject of many studies, as its mutations result in severe diseases in humans^{25,52–54}. However, despite decades of research, its precise functions remain elusive. Recent studies suggest that MAM33-family proteins might be involved in mitoribosome biogenesis. Indeed, they are linked to LSU biogenesis in yeast²⁴, the recent structure of the *Trypanosoma* SSU “assembosome” includes a heterotrimeric p22 (homolog of p32), directly highlighting its role in mitoribosome biogenesis^{55,56}, and in humans, the YBEY protein forms a complex with p32 and is involved in SSU biogenesis²³. We observed that p32 is an integral component of the *Chlamydomonas* mitoribosome. However, it does not appear to have an obvious function related to the translation process, as it does not bind rRNA and only makes a few contacts with the adjacent r-proteins. Likewise, its downregulation does not impair the accumulation of mitochondrial rRNAs. Taking these results together with the above-mentioned studies, it seems that p32 could act in mitoribosome maturation. Given its overall negative charge, it might serve as a binding platform that scaffolds other factors during ribosome biogenesis. It is unclear why p32 would be kept as a constitutive ribosomal protein in *Chlamydomonas* and not in other eukaryotes, but this may point to species-specific functions.

One of the most prominent features of the *Chlamydomonas* mitoribosome is its extensively fragmented rRNAs, with four pieces in the SSU and nine pieces in the LSU. It is interesting to note that Kinetoplastida and Euglenozoa also have fragmented rRNA in their cytosolic ribosomes^{57–62}. However, only the LSU rRNA is fragmented, and the fragments are continuous in the genome. Among mitoribosomes, the ciliate mitoribosome also contains LSU and SSU rRNAs that are each split into two pieces, contrasting with the extensive fragmentation observed here with *Chlamydomonas*^{63–65}.

It has been proposed that the fragmentation and scrambling of the *Chlamydomonas* rRNA genes are the result of several mitochondrial genome recombination events between short repeated sequences³¹. We assigned all the previously identified rRNA fragments in the mitoribosome except one, L2b. This fragment is not incorporated into the mature mitoribosome and is thus not an rRNA. Nevertheless, its transcript has reproducibly been found in the total mitochondrial fraction^{19,28}. Its function is unclear, but it might be involved in mitochondrial genome maintenance, which involves telomere-like structures in *Chlamydomonas*, as the L2b sequence is highly similar to both ends of the linear *Chlamydomonas* mitochondrial genome^{66,67}. Alternatively, the observation that the L2b RNA level is decreased in the mS105-mutant strain may hint at a function related to mS105, and thus, possibly to mitoribosome or SSU biogenesis. We cannot exclude the possibility that the mS105 protein plays a role outside the mitoribosome and is involved in mitochondrial genome maintenance via an interaction with L2b.

Importantly, we reveal that the L3a rRNA fragment is a 5S rRNA, which previously escaped identification because of its highly divergent primary sequence. Even compared to closely related Chlorophytes and Chlorophyceae species, *Chlamydomonas* L3a is particularly different at the sequence level (Supplementary Fig. 8d). L3a occupies the same position as a classical 5S, but its overall structure, notably the domain α and β structures, are quite different from other known 5S structures, rendering it one of the most divergent 5S rRNA described to date. Putative highly divergent mitochondrial 5S rRNAs have also been previously identified in various amoebozoan species^{68,69}. Given the high divergence of the 5S rRNA in *Chlamydomonas*, together with the aforementioned studies, a wider phylogenetic distribution of mitochondrial 5S rRNA might be suggested.

The rest of the rRNA fragments form the core of the mitoribosome and are globally conserved, yet reduced. In the ribosome core, these fragments are stabilized by base-pairing with each other. In contrast, on the outer shell of the ribosome, especially in the LSU, the rRNA fragments are stabilized by the *Chlamydomonas*-specific r-proteins. These proteins are all alpha-helical repeats belonging to nucleic-acid binder families PPR, OPR, and mTERF. In our structures, they form highly intertwined interfaces with single- and double-stranded RNA, all involving positive/negative charge interactions. These proteins stabilize the 3' end of rRNA fragments L1, L2a, L5, and L8 by enlacing their single-stranded extremities and also contact and stabilize additional rRNA helices via their convex surfaces.

The function of these proteins was investigated by analyzing downregulation mutants (Fig. 5). All mutants for LSU r-proteins, except the *mL116* mutant, showed reduced levels of LSU rRNA fragments, indicating that these r-proteins are important for the stability of the rRNAs, and thus integrity of the mitoribosome. They seem to play a chaperone-like role, stabilizing and perhaps contributing to the recognition and recruitment of the different rRNA pieces during assembly. In *mS105*, where the protein does not directly interact with rRNAs, the rRNA levels are almost unaffected. Additionally, in downregulation mutants *mL113* and *mL117*, steady-state levels of mitochondria-encoded proteins and active respiratory complexes are decreased, highlighting their importance in translation and impact on mitochondrial metabolism.

Our cryo-ET analysis reveals that *Chlamydomonas* mitoribosomes are bound to the inner mitochondrial membrane. In animals and most eukaryotes, the mitochondrial genome encodes almost exclusively components of the respiratory chain, which are all membrane-embedded proteins. These proteins are co-translationally inserted into the inner mitochondrial membrane to reduce the probability of protein aggregation during transport⁷⁰. To facilitate this process, mitoribosomes are consistently found attached to the inner membrane^{45,46}. In mammals, the mitoribosome attachment is mediated by a specific r-protein, mL45, located at the exit of the peptide channel, which links the ribosome to the main insertase of the inner membrane, Oxa1^{42,43}. In the yeast *S. cerevisiae*, where one of the mitochondria-encoded proteins is soluble⁷¹, the association is mediated by an mL45 homolog, Mba1, which is not an integral constituent of the mitoribosome, and an expansion segment of H96 directly contacting the membrane^{13,40,44}. This is most likely also the case in the yeast *N. crassa*¹². In *Chlamydomonas*, similarly to mammals, all proteins encoded in the mitochondrial genome are components of the respiratory chain. Therefore, it is not surprising that the *Chlamydomonas* mitoribosome would have acquired a specific r-protein to tether translation to Oxa1. Interestingly, the membrane interaction in *Chlamydomonas* is mediated via two contact points. mL119 forms one contact, and mL113 create a second contact point directly with the membrane,

similar to ES-H96 in yeast. Notably, despite the similar location of the mL113 and ES-H96 contact sites, they are of different molecular nature (protein vs. rRNA) and have been acquired via different evolutionary mechanisms: an expansion of the nuclear genome in case of mL113, compared to the expansion of the mitochondrial gene coding for 23S rRNA in yeast. In light of recent literature suggesting an early expansion of mitoribosomal proteins in eukaryotes⁷², this raises the question of whether the second contact site is an isolated case of convergent evolution between green algae and yeast, or whether it is a more universal feature of mitoribosomes that was either replaced (yeast) or lost (mammals) throughout evolution. The fact that membrane association is mediated by different proteins in each organism, yet the Oxa1 contact is conserved, indicates that this interface is particularly critical⁴⁰. In flowering plant mitochondria, which still encode a large number of soluble proteins, accessory factors might recruit mitoribosomes to the membrane, similar to Mba1 in yeast. However, in *Tetrahymena*, the mitochondrial genome encodes numerous soluble proteins, but the mitoribosome has still acquired a probable permanent anchor to Oxa1, the r-protein mL105⁶⁵. Finally, contrary to mammalian mL45 blocking the peptide channel until mitoribosome association with the membrane^{41,43}, or kinetoplastid assembly factor mL71 blocking the peptide channel during ribosome maturation⁹, the *Chlamydomonas* peptide channel is not blocked by mL119.

In conclusion, our structural and functional characterization of *Chlamydomonas* mitoribosome provides a new perspective on mitoribosome evolution and membrane binding. It delivers essential information to broader questions on the function and evolution of rRNA and the ribosome. This work paves the way for future investigations of mitoribosomes in other species, in particular Apicomplexa such as *Plasmodium* and *Toxoplasma*, where fragmented mitochondrial rRNAs also occur^{3,73}. Indeed, the structure of the *Chlamydomonas* mitoribosome demonstrates that despite the extreme fragmentation of rRNAs, the functionally important regions are well preserved. Moreover, it indicates that rRNAs do not have to be covalently continuous if the 3D ribosome structure can be recreated via RNA-RNA and RNA-protein interactions. Interestingly, Gray et al. proposed that long, covalently continuous conventional rRNAs might have evolved from short, non-covalently interacting ancestors²⁰. The *Chlamydomonas* mitoribosome might thus represent a relic from an ancestral form of rRNA organization.

Overall, the structure reported here provides further insights into the evolution of mitoribosomes and the elaboration of independent strategies to accomplish and regulate translation. Together with recent work on mitoribosomes from other species, our study demonstrates how the mitoribosome is truly one of nature's most eclectic playgrounds for evolving diverse strategies to regulate a fundamental cellular process.

Methods

Chlamydomonas reinhardtii mitochondria and mitoribosome purification.

Chlamydomonas reinhardtii cell wall-less strain CC-4351 (*cw15-325 arg7-8 mt+*) was used for mitochondria purification and transformation. The mitochondrial mutants *dums5*⁷⁴ and *dum11*⁷⁵ were used as a control for the phenotypic growth analysis in the dark, kindly provided by Dr. Remacle (University of Liège) and, respectively, annotated on figures as CI- and CIII-. The strains were grown on Tris-Acetate Phosphate (TAP) solid or liquid medium¹⁴, supplemented with 100 µg/ml of arginine when necessary, under continuous white light (50 µE/m²/s), or in the dark. Mitochondria were isolated from liquid cell cultures grown up to the exponential phase⁷⁶. Cells were harvested by centrifugation 10 min 1000 × g, and resuspended in 10 ml ice-cold 25 mM phosphate buffer pH 6.5 containing 6% PEG 6000, 0.4% (w/v) bovine serum albumin (BSA), and 0.016% (w/v) digitonin to a final concentration of 3 × 10⁸ cells/ml. The suspension was warmed rapidly to 30 °C, kept at this temperature for 30 seconds, and cooled to 4 °C. Then the broken cells were pelleted at 2500 × g and washed with 40 ml of ice-cold 20 mM Hepes-KOH pH 7.2 containing 0.15 M mannitol, 2 mM EDTA, 0.1% (w/v) BSA, and 1 mM MgCl₂. After a 2 min 1000 × g centrifugation, the pellet was resuspended in

2 ml of the same solution, stirred vigorously for 45 s, and then 6 ml of 20 mM Hepes-KOH buffer pH 7.2 containing 0.15 M mannitol, 0.8 mM EDTA, and 4 mM MgCl₂ were added. Mitochondria were collected at 12,000 × g for 10 min, resuspended in the same last buffer, and then loaded on a discontinuous Percoll gradient (13%/21%/45%) in MET buffer (280 mM Mannitol, 10 mM Tris-HCl pH 6.8, 0.5 mM EDTA, and 0.1% BSA) and centrifuged for 60 min at 40,000 × g. Purified mitochondria were recovered at the 45/21 interface and washed two times in MET buffer by centrifugation at 12,000 × g for 10 min and stored at -80 °C.

Mitoribosome purification was conducted as previously^{10,11}. In brief, purified mitochondria were resuspended in Lysis buffer (20 mM HEPES-KOH, pH 7.6, 100 mM KCl, 30 mM MgCl₂, 1 mM DTT, 1.6% Triton X-100, 0.5% *n*-DDM, supplemented with proteases inhibitors (CComplete EDTA-free)) to a 1 mg/ml concentration and incubated for 15 min in 4 °C. Lysate was clarified by centrifugation at 25,000 × g, 20 min at 4 °C. The supernatant was loaded on a 40% sucrose cushion in Monosome buffer (Lysis buffer without Triton X-100 and 0.1% *n*-DDM) and centrifuged at 235,000 × g, 3 h, 4 °C. The crude ribosomes pellet was resuspended in Monosome buffer and loaded on a 10–30% sucrose gradient in the same buffer and run for 16 h at 65,000 × g. Fractions corresponding to mitoribosomes were collected, pelleted, and resuspended in Monosome buffer and analyzed by nanoLC-ESI-MS/MS and cryo-EM (Supplementary Fig. 1).

Grid preparation. For the single-particle analyses, 4 µl of the samples at a protein concentration of 1.5 µg/µl was applied onto Quantifoil R2/2 300-mesh holey carbon grid, coated with thin home-made continuous carbon film and glow-discharged (2.5 mA for 20 s). The sample was incubated on the grid for 30 s and then blotted with filter paper for 2 s in a temperature and humidity-controlled Vitrobot Mark IV (T = 4 °C, humidity 100%, blot force 5) followed by vitrification in liquid ethane.

Cryo-electron microscopy data collection. The single-particle data collection was performed on a Talos Arctica instrument (ThermoFisher Company) at 200 kV using the SerialEM software for automated data acquisition. Data were collected at a nominal underfocus of -0.5 to -2.5 µm, at magnifications of 36,000× with a pixel size of 1.13 Å for the SSU, and 45,000× with a pixel size of 0.9 Å for the LSU. Micrographs were recorded as movie stacks on a K2 direct electron detector (GATAN Company); each movie stack was fractionated into 65 frames, for a total exposure of 6.5 s corresponding to an electron dose of 45 e⁻/Å².

Electron microscopy image processing. Drift and gain correction and dose weighting were performed using MotionCorr²⁷. A dose-weighted average image of the whole stack was used to determine the contrast transfer function with the software Gctf⁷⁸. The following workflow was processed using RELION 3.0⁷⁹. Initial analyses were performed in CryoSPARC⁸⁰ to assess sample composition and to generate ab-initio cryo-EM map. After reference-free 2D classification, for the LSU 346,994 particles were extracted and used for 3D classification into 6 classes (Supplementary Fig. 2). Ab-initio cryo-EM reconstruction generated in CryoSPARC was low-pass filtered to 30 Å, and used as an initial reference for 3D classification. Two subclass depicting high-resolution features was selected for refinement with 101,291 particles. After Bayesian polishing, the LSU reconstruction reached 3.00 Å resolution. For the SSU reconstruction, a similar workflow was applied. After 2D classification, 445,469 particles were extracted and used for 3D classification into 6 classes. A single subclass depicting high-resolution features was selected for refinement with 40,131 particles. After focus refinement using masks for the head and body of the small subunit, the SSU reconstruction reached a resolution of 4.19 Å for the body and 4.47 Å for the head. Determination of the local resolution of the final density map was performed using ResMap⁸¹.

Structure building and model refinement. The atomic model of the *C. reinhardtii* LSU was built into the high-resolution maps using Coot, Phenix, and Chimera. Atomic models from *E. coli* (PDB: 5KCR) and the *A. thaliana* mitoribosome (PDB: 6XYW) were used as starting points for protein identification and modelization, as well as rRNA modelization. The online SWISS-MODEL⁸², as well as AlphaFold²⁷ through the ColabFold service⁸³, were used to generate initial models for bacterial and mitochondria conserved r-proteins. Models were then rigid body fitted to the density in Chimera⁸⁴ and all subsequent modeling was done in Coot⁸⁵. Extensions were built as polyalanine and mutated to the adequate sequences. *Chlamydomonas*-specific proteins for which no model could be generated were first built entirely as polyalanine, then the sequence-from-map Phenix tool⁸⁶ was used to identify each of the proteins, and the correct sequences were placed in the densities. For the SSU, due to the lower resolution in comparison to the LSU, all extensions of the homology models were built as polyalanine and unknown densities were built as Ala residues. For refinement, a combination of regularization and real-space refine was performed in Coot for each protein. The global atomic model was then subjected to real-space refinement cycles using *phenix.real_space_refine* Phenix⁸⁶ function, during which protein secondary structures, Ramachandran, and side-chain rotamer restraints were applied. Several rounds of refinement (manual in Coot and automated using the *phenix.real_space_refine*) were performed to obtain the final models, which were validated using the built-in validation tool of

Phenix, based on MolProbity. Refinement and validation statistics are summarized in Supplementary Table 2.

Cell vitrification and Cryo-FIB milling. We used *Chlamydomonas reinhardtii* mat3-4 cells (strain CC-3994)⁸⁷, which exhibit superior vitrification due to their small size. The strain was acquired from the Chlamydomonas Resource Center, University of Minnesota, St. Paul. Cells were grown until the mid-log phase in Tris-acetate-phosphate (TAP) medium under constant light exposure and bubbling with a normal atmosphere. Vitrification and FIB milling were performed as previously described^{88,89}. Using a Vitrobot Mark 4 (FEI), cells in suspension (4 μ l of \sim 1000 cells per μ l) were blotted onto R2/1 carbon-coated 200-mesh copper grids (Quantifoil Micro Tools) and plunge frozen in a liquid ethane/propane mixture. Grids were then mounted into Autogrid supports (FEI) and transferred into either a FEI Scios or FEI Quanta dual-beam FIB/SEM instrument. The grids were coated with an organometallic platinum layer by the gas injection system (FEI), and cells were thinned from both sides with a gallium ion beam to a final thickness of \sim 100–200 nm.

Cryo-electron tomography data acquisition. Cellular tomograms were acquired on a 300 kV Titan Krios microscope (FEI), equipped with a post-column energy filter (Quantum, Gatan) and a direct detector camera (K2 summit, Gatan). Tilt series were recorded using SerialEM software⁹⁰ with 2° tilt increments from -60° to $+60^\circ$ (in two halves separated at either 0° or -20°), an object pixel size of 3.42 Å, 12 frames per second, a defocus of -4 to -5.5 μ m, and a total accumulated dose of \sim 100 e⁻/Å.

Cryo-electron tomography data processing. Frames from the K2 detector were motion corrected with MotionCor2⁷⁷. Using IMOD software, tilt series were aligned with patch-tracking, and tomograms were reconstructed with weighted back projection. Out of \sim 130 tomograms, 47 tomograms containing mitochondria were selected. The following workflow is described in Supplementary Fig. 3. An initial structure of the *C. reinhardtii* mitoribosome was obtained by manually picking 103 mitoribosomes from 27 tomograms following template-free alignment by spherical harmonics⁹¹. The initial map was then used as a template for automated template matching on 47 tomograms with a voxel size of 2.1 nm using PyTOM⁹². To reduce false positives, the highest correlation peaks of the resulting 6-D cross-correlation function localized in the mitochondrial matrix were manually inspected in UCSF Chimera⁸⁴, and a set of 222 subvolumes from 27 tomograms was obtained. The subvolumes were reconstructed at a voxel size of 6.84 Å, aligned using PyTOM's real-space refinement, and subjected to one round of classification with a mask encompassing the membrane region. This yielded a class of 73 mitoribosomes with a clear membrane density that was subjected to one more round of real-space refinement in PyTOM. For the resulting average, a resolution of 31.6 Å (large ribosomal subunit) and 31.5 Å (small ribosomal subunit) was determined by fourier-shell cross-resolution of the two maps against the maps obtained by single-particle analysis (FSC = 0.33). For the localization of ATP synthases, an initial structure of the *C. reinhardtii* ATP synthase was obtained by manually picking 417 subvolumes from four tomograms and aligning them using spherical harmonics. The obtained map was used as a template for automated template matching on tomograms with a voxel size of 2.1 nm using PyTOM, and the highest correlation peaks were then manually inspected in Chimera to remove false positives. Cytosolic 80S ribosomes were localized in one tomogram by filtering EMDB-1780⁹³ to 40 Å and using it as a template for template matching on the deconvolved tomogram (tom_deconv; https://github.com/dtegunov/tom_deconv) using PyTOM. Subvolumes were extracted for the 2100 highest correlation peaks, of which 1700 subvolumes were classified as ribosomes by unsupervised, auto-focused 3D classification⁹⁴.

Proteomic analyses of *C. reinhardtii* mitoribosome composition. Mass spectrometry analyses of the total, mitochondrial and ribosomal fractions of *C. reinhardtii* were done at the Strasbourg-Esplanade proteomic platform and performed as previously¹⁰. In brief, proteins were trypsin digested, mass spectrometry analyses and quantitative proteomics were carried out by nanoLC-ESI-MS/MS analysis on a QExactive+(T hermo) mass spectrometer. Data were searched against the UniProtKB (Swissprot + trEMBL) database restricted to the *C. reinhardtii* taxonomy with a target-decoy strategy (UniProtKB release 2020_03, taxon 3055, 31246 forward protein sequences). Proteins were validated respecting FDR < 1% (false discovery rate) and quantitative label-free analysis was performed through in-house bioinformatics pipelines.

Artificial miRNA *C. reinhardtii* strain generation and analyses. Artificial microRNAs constructs were created according to Molnar et al.³⁹ as follows: the oligonucleotides were designed using the WMD3 Web MicroRNA Designer software v3.2 [<http://wmd3.weigelworld.org/cgi-bin/webapp.cgi>] and genome release *Chlamydomonas* CDS *reinhardtii* 281 v5.6.cds (Phytozome) (Supplementary Table 3). The oligonucleotides were annealed, phosphorylated, and ligated into a *Spe*I-digested pChlamiRNA2 containing the ARG7 gene as a selection marker. The resulting plasmids were linearized and transformed into *Chlamydomonas* CC-4351 strain by the Neon® Transfection System (LifeTechnologies) according to the

GeneArt® MAX Efficiency® Transformation protocol for Algae (LifeTechnologies Cat#A24229). Cells with integrated plasmid were selected on TAP plates without arginine. Colonies (16–48 depending on the transformation) were picked to grow to logarithmic phase on liquid TAP medium. They were then spotted on two identical TAP plates to test their capacity to grow in the dark. One plate was placed in a mixotrophic condition (light + acetate) for 5–7 days, and the other one in a heterotrophic condition (dark + acetate), for 10–15 days. For the dilution series, cells were grown for 3–4 days on TAP plates and were resuspended in 2 ml of liquid TAP medium. The cell density was measured spectrophotometrically at OD750 and diluted to an OD750 = 1.5. This normalized suspension was used as the starting material (set to 1) for making three serial 5-fold dilutions (2.10^{-1} , 4.10^{-2} , and 8.10^{-3}). A volume of 10 μ l for each dilution was then spotted on two identical TAP plates.

rRNA analysis by RNA sequencing. The RNAs were prepared from cells using TRI Reagent® (Molecular Research Center) according to the manufacturer's instructions.

For northern blots, 1 μ g of total, mitochondrial and mitoribosome fraction RNA, were separated on 7 M Urea - 8% polyacrylamide gel, transferred onto Amersham Hybond™-N + membrane (GE Healthcare Cat#RPN203B), and hybridized to radiolabelled oligonucleotide probes (Supplementary Table 3) in 6 \times SSC, 0.5% SDS at 45 °C. Washing conditions were: 2 times 10 min in 2 \times SSC and 1 time 30 min in 2 \times SSC, 0.1% SDS at the hybridization temperature. For each specific probe, the signal was detected with the Amersham Typhoon laser scanner (Amersham).

For the quantitative real-time RT-PCR analyses, RNAs were treated with RQ1 RNase-Free DNase (Promega Cat#M6101) according to Promega's protocol, using 0.2 U/ μ g of RNA. To obtain cDNA, reverse transcription assays were performed according to the manufacturer's instructions with 2.5 μ g of total RNA in the presence of 5 μ M of oligo(dT) primer (Supplementary Table 3) and 25 ng/ μ l of Random Primers (Promega Cat#C118A) using the SuperScript™ IV Reverse Transcriptase (Invitrogen Cat#18090010). The RT-qPCR amplification was carried out with the dsDNA-specific dye Takyon™ SYBR® 2X qPCR Mastermix Blue (Eurogentec Cat#UF-FSMT-B0701) and monitored in real-time with a LightCycler 480 instrument (Roche). The primers used are listed in (Supplementary Table 3). The delta-delta Ct method was used to calculate the relative RNA abundance with respect to the geometric mean of two RNA references MAA7 and CYN19-3⁹⁵.

For the RNA sequencing, the p204 library was built with total mitochondrial RNA. The RNA was first chemically fragmented (4 min) and then enzymatically treated with Antarctic Phosphatase (NEB#M0289S) and T4 Polynucleotide Kinase (NEB#M0201S). Library preparation was done according to the TruSeq Small RNA Sample Preparation Guide #15004197 Rev. F February 2014. The library was sequenced on the Illumina MiSeq sequencer in a paired-end mode of 2 \times 75 nt reads. The NGS192-small library was built with the mitoribosome fraction. The RNA was also enzymatically treated with Antarctic Phosphatase and T4 Polynucleotide Kinase. The library was then constructed with the NEBNext multiplex small RNA Library set for Illumina reference E7580 following the manufacturer's instructions. Following PCR amplification, a size selection was performed on a 6% TBE gel to recover the 160–350 bp PCR fragments for sequencing. The NGS192-total library was prepared according to the TruSeq Stranded Total RNA with Ribozero Plant kit, starting from the first-strand cDNA synthesis step and omitting the two first depletion and fragmentation steps. The library was sequenced on the Illumina MiSeq sequencer in a paired-end mode of 2 \times 110 nt reads. Both libraries were sequenced at the IBMP platform. The reads were mapped to the *Chlamydomonas* mitochondrial genome (EU306622) using Bowtie2 version 2.4.1 with the following options -end-to-end -very-sensitive -N 0 -L 22. Alignments were displayed with the Integrative Genomics Viewer (IGV) with the bigWig format.

Protein analyses. The *Chlamydomonas* crude total membrane fractions were obtained according to Remacle et al.⁹⁶ as follows: *Chlamydomonas* cells TAP liquid cultures were collected and resuspended to $2\text{--}1.6 \times 10^8$ cells/ml in MET buffer (280 mM mannitol, 0.5 mM EDTA, 10 mM Tris-HCl pH 7) with 1 \times cOmplete™ Protease Inhibitor Cocktail and then disrupted by sonication (four times 30 s of sonication and 30 s of pause; Bioruptor® Pico, Diagenode). The suspension was centrifugated (10 min at 500 \times g, followed by 4 min at 3000 \times g) and the protein content of the supernatant was determined by the Bradford method. Equal amounts of protein were separated using 15% SDS-polyacrylamide gel electrophoresis (PAGE), and transferred to a 0.45 μ m PVDF membrane (Immobilon®-P Transfer Membrane; Merck Millipore Cat.#IPVH00010). Specific antibodies were used in immunoblotting and were detected using chemiluminescence (Clarity Western ECL Substrate, Bio-Rad). We used rabbit sera obtained against *Chlamydomonas reinhardtii* mitochondrial-encoded subunits complex I, Nad4 (1:1000) and Nad6 (1:100), nuclear-encoded subunit complex I, NUO7 (1:2000), and the nuclear-encoded mitochondrial protein VDAC1 (1:25000). The expected / apparent molecular weight are as follows 49 kDa/50–55 kDa for Nad4, 18 kDa/18 kDa for Nad6, 49 kDa/38 kDa for NUO7, and 28 kDa/28 kDa for VDAC1. Blue native polyacrylamide gel electrophoresis (BN-PAGE) analyses were conducted according to Schägger et al.⁹⁷ as follows: the crude total membrane fractions were prepared as above with an additional centrifugation at high speed (27,000 \times g for 15 min) and

the final pellet was suspended in ACA buffer (375 mM 6-aminohexanoic acid, 25 mM Bis-Tris, pH 7, and, 250 mM EDTA). 0.5 mg of the crude total membrane were first solubilized in the presence of 1.5% (w/v) *n*-dodecyl- β -D-maltoside and then centrifuged for 40 min 14,200 \times g at 4 °C to remove insoluble matters. 0.65% (w/v) of coomassie serva blue G was then added to the supernatant prior to separation by electrophoresis on a 5% to 12% polyacrylamide gradient BN gel. In-gel detection of Complex I (NADH dehydrogenase) activity was performed using a 100 mM Tris-HCl pH 7.4 buffer containing 200 μ M NADH and 0.2% nitro blue tetrazolium (NBT). In-gel detection of Complex IV (cytochrome *c* oxydase) activity was performed using a 10 mM MOPS-KOH pH 7.4 buffer containing 7.5% saccharose, 19 U/ml catalase from bovine liver, 0.1% cytochrome *c*, and 0.01% 3,3'-diaminobenzidine (DAB).

Figure preparation and data visualization. Three-dimensional segmentation of ER, mitochondrial, and chloroplast membranes in the cryo-tomogram was performed using EMAN's convolutional neural network for automated annotation⁹⁸. Using the TOM toolbox in matlab⁹⁹, the averages of the cytoribosomes, mitoribosomes, and ATP synthases were pasted into the tomogram at the refined coordinates and angles determined by subtomogram analysis. Figures featuring cryo-EM densities as well as atomic models were visualized with UCSF ChimeraX¹⁰⁰ and Chimera⁸⁴.

Statistical information. Data are presented as mean values \pm SD (standard deviation), calculated using Microsoft Excel version 16.43 and GraphPad Prism 8 version 8.4. The *p*-value < 0.05 was considered the threshold for statistical significance. The *p*-value significance intervals (*) are provided within each figure legend, together with the statistical test performed for each experiment: the two-tailed Mann-Whitney. For Fig. 5b, c, derived statistics correspond to the analysis of mean values of *n* = 3 biological replicates. Statistics detailed data (means, standard deviation, *n* values, exact *p*-values) are provided in the Source Data file.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The data that support this study are available from the corresponding authors upon reasonable request. The cryo-EM maps of *C. reinhardtii* mitoribosome have been deposited at the Electron Microscopy Data Bank (EMDB): EMD-13480 for the LSU, EMD-13481 for the head of the SSU, EMD-13477 for the body of the SSU, and EMD-13578 for the subtomogram averaging of the whole ribosome. The corresponding atomic models have been deposited in the Protein Data Bank (PDB) under the accession 7PKT for the LSU and 7PKQ for the SSU. Mass spectrometric data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD024708. RNAseq data were deposited in the NCBI Gene Expression Omnibus under accession number GSE171125. Source data are provided with this paper.

Received: 8 July 2021; Accepted: 8 November 2021;
Published online: 09 December 2021

References

- Eme, L., Spang, A., Lombard, J., Stairs, C. W. & Ettema, T. J. G. Archaea and the origin of eukaryotes. *Nat. Rev. Microbiol.* **15**, 711–723 (2017).
- Gray, M. W. Mitochondrial evolution. *Cold Spring Harb. Perspect. Biol.* **4**, a011403 (2012).
- Waltz, F. & Giegé, P. Striking diversity of mitochondria-specific translation processes across eukaryotes. *Trends Biochem. Sci.* **45**, 149–162 (2020).
- Kummer, E. & Ban, N. Mechanisms and regulation of protein synthesis in mitochondria. *Nat. Rev. Mol. Cell Biol.* **1–19**, <https://doi.org/10.1038/s41580-021-00332-2> (2021).
- Martijn, J., Vosseberg, J., Guy, L., Offre, P. & Ettema, T. J. G. Deep mitochondrial origin outside the sampled alphaproteobacteria. *Nature* **557**, 101–105 (2018).
- Amunts, A., Brown, A., Toots, J., Scheres, S. H. W. & Ramakrishnan, V. The structure of the human mitochondrial ribosome. *Science* **348**, 95–98 (2015).
- Greber, B. J. et al. The complete structure of the 55S mammalian mitochondrial ribosome. *Science* **348**, 303–308 (2015).
- Ramrath, D. J. F. et al. Evolutionary shift toward protein-based architecture in trypanosomal mitochondrial ribosomes. *Science* **362**, eaau7735 (2018).
- Soufari, H. et al. Structure of the mature kinetoplasts mitoribosome and insights into its large subunit biogenesis. *Proc. Natl Acad. Sci. USA* **117**, 29851–29861 (2020).
- Waltz, F. et al. Small is big in *Arabidopsis* mitochondrial ribosome. *Nat. Plants* **5**, 106–117 (2019).

- Waltz, F., Soufari, H., Boehler, A., Giegé, P. & Hashem, Y. Cryo-EM structure of the RNA-rich plant mitochondrial ribosome. *Nat. Plants* **6**, 377–383 (2020).
- Itoh, Y., Naschberger, A., Mortezaei, N., Herrmann, J. M. & Amunts, A. Analysis of translating mitoribosome reveals functional characteristics of translation in mitochondria of fungi. *Nat. Commun.* **11**, 5187 (2020).
- Desai, N., Brown, A., Amunts, A. & Ramakrishnan, V. The structure of the yeast mitochondrial ribosome. *Science* **355**, 528–531 (2017).
- Harris, E. H., Stern, D. B. & Witman, G. B. in *The Chlamydomonas Sourcebook* Vols 1–3, iii (Elsevier, 2009).
- Remacle, C., Cardol, P., Coosemans, N., Gaisne, M. & Bonnefoy, N. High-efficiency biolistic transformation of *Chlamydomonas* mitochondria can be used to insert mutations in complex I genes. *Proc. Natl Acad. Sci. USA* **103**, 4771–4776 (2006).
- Salinas, T., Larosa, V., Cardol, P., Maréchal-Drouard, L. & Remacle, C. Respiratory-deficient mutants of the unicellular green alga *Chlamydomonas*: a review. *Biochimie* **100**, 207–218 (2014).
- Boer, P. H. & Gray, M. W. Scrambled ribosomal RNA gene pieces in *Chlamydomonas reinhardtii* mitochondrial DNA. *Cell* **55**, 399–411 (1988).
- Denovan-Wright, E. M. & Lee, R. W. Evidence that the fragmented ribosomal RNAs of *Chlamydomonas* mitochondria are associated with ribosomes. *FEBS Lett.* **370**, 222–226 (1995).
- Salinas-Giegé, T. et al. Polycytidylation of mitochondrial mRNAs in *Chlamydomonas reinhardtii*. *Nucleic Acids Res.* **45**, 12963–12973 (2017).
- Gray, M. W. & Gopalan, V. Piece by piece: building a ribozyme. *J. Biol. Chem.* **295**, 2313–2323 (2020).
- Ban, N. et al. A new system for naming ribosomal proteins. *Curr. Opin. Struct. Biol.* <https://doi.org/10.1016/j.sbi.2014.01.002> (2014).
- Valach, M. et al. An unexpectedly complex mitoribosome in *Andalucia godoyi*, a protist with the most bacteria-like mitochondrial genome. *Mol. Biol. Evol.* **38**, 788–804 (2021).
- Summer, S. et al. YBEY is an essential biogenesis factor for mitochondrial ribosomes. *Nucleic Acids Res.* **48**, 9762–9786 (2020).
- Hillman, G. A. & Henry, M. F. The yeast protein Mam33 functions in the assembly of the mitochondrial ribosome. *J. Biol. Chem.* **294**, 9813–9829 (2019).
- Yagi, M. et al. p32/gC1qR is indispensable for fetal development and mitochondrial translation: importance of its RNA-binding ability. *Nucleic Acids Res.* **40**, 9717–9737 (2012).
- Jiang, J., Zhang, Y., Krainer, A. R. & Xu, R.-M. Crystal structure of human p32, a doughnut-shaped acidic mitochondrial matrix protein. *Proc. Natl Acad. Sci. USA* **96**, 3572–3577 (1999).
- Jumper, J. et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
- Wobbe, L. & Nixon, P. J. The mTERF protein MOC1 terminates mitochondrial DNA transcription in the unicellular green alga *Chlamydomonas reinhardtii*. *Nucleic Acids Res.* **41**, 6553–6567 (2013).
- Gallagher, S. D. et al. High-throughput sequencing of the chloroplast and mitochondrion of *Chlamydomonas reinhardtii* to generate improved de novo assemblies, analyze expression patterns and transcript speciation, and evaluate diversity among laboratory strains and wild isolates. *Plant J.* **93**, 545–565 (2018).
- Petrov, A. S. et al. Structural patching fosters divergence of mitochondrial ribosomes. *Mol. Biol. Evol.* **36**, 207–219 (2019).
- Nedelcu, A. M. Fragmented and scrambled mitochondrial ribosomal RNA coding regions among green algae: a model for their origin and evolution. *Mol. Biol. Evol.* **14**, 506–517 (1997).
- Hammani, K. et al. Helical repeats modular proteins are major players for organelle gene expression. *Biochimie* **100**, 141–150 (2014).
- Barkan, A. & Small, I. Pentatricopeptide repeat proteins in plants. *Annu. Rev. Plant Biol.* **65**, 415–442 (2014).
- Rahire, M., Laroche, F., Cerutti, L. & Rochoix, J.-D. Identification of an OPR protein involved in the translation initiation of the Psab subunit of photosystem I. *Plant J.* **72**, 652–661 (2012).
- Eberhard, S. et al. Dual functions of the nucleus-encoded factor TDA1 in trapping and translation activation of atpA transcripts in *Chlamydomonas reinhardtii* chloroplasts. *Plant J.* **67**, 1055–1066 (2011).
- Wang, F. et al. Two *Chlamydomonas* OPR proteins stabilize chloroplast mRNAs encoding small subunits of photosystem II and cytochrome b 6 f. *Plant J.* **82**, 861–873 (2015).
- Boulouis, A. et al. Spontaneous dominant mutations in *Chlamydomonas* highlight ongoing evolution by gene diversification. *Plant Cell* **27**, 984–1001 (2015).
- Li, X. et al. A genome-wide algal mutant library and functional screen identifies genes required for eukaryotic photosynthesis. *Nat. Genet.* **51**, 627–635 (2019).
- Molnar, A. et al. Highly specific gene silencing by artificial microRNAs in the unicellular alga *Chlamydomonas reinhardtii*. *Plant J.* **58**, 165–174 (2009).
- Ott, M. & Herrmann, J. M. Co-translational membrane insertion of mitochondrially encoded proteins. *Biochim. Biophys. Acta* **1803**, 767–775 (2010).

41. Kummer, E. et al. Unique features of mammalian mitochondrial translation initiation revealed by cryo-EM. *Nature* **560**, 263–267 (2018).
42. Desai, N. et al. Elongational stalling activates mitoribosome-associated quality control. *Science* **370**, 1105–1110 (2020).
43. Itoh, Y. et al. Mechanism of membrane-tethered mitochondrial protein synthesis. *Science* **371**, 846–849 (2021).
44. Ott, M. et al. Mba1, a membrane-associated ribosome receptor in mitochondria. *EMBO J.* **25**, 1603–1610 (2006).
45. Pfeffer, S., Woellhaf, M. W., Herrmann, J. M. & Förster, F. Organization of the mitochondrial translation machinery studied in situ by cryoelectron tomography. *Nat. Commun.* **6**, 6019 (2015).
46. Englmeier, R., Pfeffer, S. & Förster, F. Structure of the human mitochondrial ribosome studied in situ by cryoelectron tomography. *Structure* **25**, 1574–1581.e2 (2017).
47. Waltz, F., Corre, N., Hashem, Y. & Giegé, P. Specificities of the plant mitochondrial translation apparatus. *Mitochondrion* **53**, 30–37 (2020).
48. Tomal, A., Kwasiak-Owczarek, M. & Janska, H. An update on mitochondrial ribosome biology: the plant mitoribosome in the spotlight. *Cells* **8**, 1562 (2019).
49. Aibara, S., Singh, V., Modelska, A. & Amunts, A. Structural basis of mitochondrial translation. *Elife* **9**, 1–17 (2020).
50. Johnson, A. G. et al. RACK1 on and off the ribosome. *RNA* **25**, 881–895 (2019).
51. Seytter, T., Lottspeich, F., Neupert, W. & Schwarz, E. Mam33p, an oligomeric, acidic protein in the mitochondrial matrix of *Saccharomyces cerevisiae* is related to the human complement receptor gC1q-R. *Yeast* **14**, 303–310 (1998).
52. Saha, P. & Datta, K. Multi-functional, multicompartmental hyaluronan-binding protein 1 (HABP1/p32/gC1qR): implication in cancer progression and metastasis. *Oncotarget* **9**, 10784–10807 (2018).
53. Hu, M. et al. p32 protein levels are integral to mitochondrial and endoplasmic reticulum morphology, cell metabolism and survival. *Biochem. J.* **453**, 381–391 (2013).
54. Fogal, V. et al. Mitochondrial p32 protein is a critical regulator of tumor metabolism via maintenance of oxidative phosphorylation. *Mol. Cell. Biol.* **30**, 1303–1318 (2010).
55. Saurer, M. et al. Mitoribosomal small subunit biogenesis in trypanosomes involves an extensive assembly machinery. *Science* **365**, 1144–1149 (2019).
56. Sprehe, M., Fisk, J. C., McEvoy, S. M., Read, L. K. & Schumacher, M. A. Structure of the *Trypanosoma brucei* p22 protein, a cytochrome oxidase subunit II-specific RNA-editing accessory factor*. *J. Biol. Chem.* **285**, 18899–18908 (2010).
57. Matzov, D. et al. Cryo-EM structure of the highly atypical cytoplasmic ribosome of *Euglena gracilis*. *Nucleic Acids Res.* **48**, 11750–11761 (2020).
58. Hashem, Y. et al. High-resolution cryo-electron microscopy structure of the *Trypanosoma brucei* ribosome. *Nature* **494**, 385–389 (2013).
59. Greenwood, S. J. & Gray, M. W. Processing of precursor rRNA in *Euglena gracilis*: identification of intermediates in the pathway to a highly fragmented large subunit rRNA. *Biochim. Biophys. Acta* **1443**, 128–138 (1998).
60. Schnare, M. N. & Gray, M. W. Sixteen discrete RNA components in the cytoplasmic ribosome of *Euglena gracilis*. *J. Mol. Biol.* **215**, 73–83 (1990).
61. Gray, M. W. Unusual pattern of ribonucleic acid components in the ribosome of *Crithidia fasciculata*, a trypanosomatid protozoan. *Mol. Cell. Biol.* **1**, 347–357 (1981).
62. Spencer, D. F., Collings, J. C., Schnare, M. N. & Gray, M. W. Multiple spacer sequences in the nuclear large subunit ribosomal RNA gene of *Crithidia fasciculata*. *EMBO J.* **6**, 1063–1071 (1987).
63. Schnare, M. N., Heinonen, T. Y. K., Young, P. G. & Gray, M. W. A discontinuous small subunit ribosomal RNA in *Tetrahymena pyriformis* mitochondria. *J. Biol. Chem.* **261**, 5187–5193 (1986).
64. Heinonen, T. Y., Schnare, M. N., Young, P. G. & Gray, M. W. Rearranged coding segments, separated by a transfer RNA gene, specify the two parts of a discontinuous large subunit ribosomal RNA in *Tetrahymena pyriformis* mitochondria. *J. Biol. Chem.* **262**, 2879–2887 (1987).
65. Tobiasson, V. & Amunts, A. Ciliate mitoribosome illuminates evolutionary steps of mitochondrial translation. *Elife* **9**, e59264 (2020).
66. Smith, D. R. & Craig, R. J. Does mitochondrial DNA replication in *Chlamydomonas* require a reverse transcriptase? *N. Phytol.* **229**, 1192–1195 (2021).
67. Smith, D. R., Hua, J. & Lee, R. W. Evolution of linear mitochondrial DNA in three known lineages of *Polytomella*. *Curr. Genet.* **56**, 427–438 (2010).
68. Bullerwell, C. E., Schnare, M. N. & Gray, M. W. Discovery and characterization of *Acanthamoeba castellanii* mitochondrial 5S rRNA. *RNA* **9**, 287–292 (2003).
69. Bullerwell, C. E. et al. Abundant 5S rRNA-like transcripts encoded by the mitochondrial genome in amoebozoans. *Eukaryot. Cell* **9**, 762–773 (2010).
70. Ott, M., Amunts, A. & Brown, A. Organization and Regulation of mitochondrial protein synthesis. *Annu. Rev. Biochem.* **85**, 77–101 (2016).
71. Freil, K. C., Friedrich, A. & Schacherer, J. Mitochondrial genome evolution in yeasts: an all-encompassing view. *FEMS Yeast Res.* **15**, 23 (2015).
72. Gray, M. W. et al. The draft nuclear genome sequence and predicted mitochondrial proteome of *Andalucia godoyi*, a protist with the most gene-rich and bacteria-like mitochondrial genome. *BMC Biol.* **18**, 1–35 (2020).
73. Feagin, J. E. et al. The fragmented mitochondrial ribosomal RNAs of *Plasmodium falciparum*. *PLoS ONE* **7**, e38320 (2012).
74. Cardol, P. et al. Photosynthesis and state transitions in mitochondrial mutants of *Chlamydomonas reinhardtii* affected in respiration. *Plant Physiol.* **133**, 2010–2020 (2003).
75. Dorthu, M.-P. et al. Biochemical, genetic and molecular characterization of new respiratory-deficient mutants in *Chlamydomonas reinhardtii*. *Plant Mol. Biol.* **18**, 759–772 (1992).
76. Salinas, T. et al. Co-evolution of mitochondrial tRNA import and codon usage determines translational efficiency in the green alga *Chlamydomonas*. *PLoS Genet.* **8**, e1002946 (2012).
77. Zheng, S. Q. et al. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331–332 (2017).
78. Zhang, K. Gctf: Real-time CTF determination and correction. *J. Struct. Biol.* **193**, 1–12 (2016).
79. Zivanov, J. et al. New tools for automated high-resolution cryo-EM structure determination in RELION-3. *Elife* **7**, e42166 (2018).
80. Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, 290–296 (2017).
81. Kucukelbir, A., Sigworth, F. J. & Tagare, H. D. Quantifying the local resolution of cryo-EM density maps. *Nat. Methods* **11**, 63–65 (2014).
82. Waterhouse, A. et al. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res.* **46**, W296–W303 (2018).
83. Mirdita, M., Ovchinnikov, S. & Steinegger, M. ColabFold - making protein folding accessible to all. Preprint at *bioRxiv* <https://doi.org/10.1101/2021.08.15.456425> (2021).
84. Pettersen, E. F. et al. UCSF Chimera-a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
85. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66**, 486–501 (2010).
86. Liebschner, D. et al. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr. Sect. D Struct. Biol.* **75**, 861–877 (2019).
87. Umen, J. G. & Goodenough, U. W. Control of cell division by a retinoblastoma protein homolog in *Chlamydomonas*. *Genes Dev.* **15**, 1652–1661 (2001).
88. Schaffer, M. et al. Optimized cryo-focused ion beam sample preparation aimed at in situ structural studies of membrane proteins. *J. Struct. Biol.* **197**, 73–82 (2017).
89. Schaffer, M. et al. Cryo-focused ion beam sample preparation for imaging vitreous cells by cryo-electron tomography. *Bio-protocol* **5**, e1575 (2015).
90. Mastronarde, D. N. Automated electron microscope tomography using robust prediction of specimen movements. *J. Struct. Biol.* **152**, 36–51 (2005).
91. Chen, Y., Pfeffer, S., Hrabe, T., Schuller, J. M. & Förster, F. Fast and accurate reference-free alignment of subtomograms. *J. Struct. Biol.* **182**, 235–245 (2013).
92. Hrabe, T. et al. PyTom: a python-based toolbox for localization of macromolecules in cryo-electron tomograms and subtomogram analysis. *J. Struct. Biol.* **178**, 177–188 (2012).
93. Armache, J.-P. et al. Localization of eukaryote-specific ribosomal proteins in a 5.5-A cryo-EM map of the 80S eukaryotic ribosome. *Proc. Natl Acad. Sci. USA* **107**, 19754–19759 (2010).
94. Chen, Y., Pfeffer, S., Fernández, J. J., Sorzano, C. O. S. & Förster, F. Autofocused 3D classification of cryoelectron subtomograms. *Structure* **22**, 1528–1537 (2014).
95. Livak, K. J. & Schmittgen, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the 2^{-ΔΔCT} method. *Methods* **25**, 402–408 (2001).
96. Remacle, C., Duby, F., Cardol, P. & Matagne, R. F. Mutations inactivating mitochondrial genes in *Chlamydomonas reinhardtii*. *Biochem. Soc. Trans.* **29**, 442–446 (2001).
97. Schägger, H. & von Jagow, G. Blue native electrophoresis for isolation of membrane protein complexes in enzymatically active form. *Anal. Biochem.* **199**, 223–231 (1991).
98. Chen, M. et al. Convolutional neural networks for automated annotation of cellular cryo-electron tomograms. *Nat. Methods* **14**, 983–985 (2017).
99. Nickell, S. et al. TOM software toolbox: acquisition and analysis for electron tomography. *J. Struct. Biol.* **149**, 227–234 (2005).
100. Goddard, T. D. et al. UCSF ChimeraX: meeting modern challenges in visualization and analysis. *Protein Sci.* **27**, 14–25 (2018).

Acknowledgements

We are grateful to Pr. Claire Remacle (University of Liège) for her kind gifts of strains and antibodies. We also thank, J. Chicher and P. Hamman of the Strasbourg Esplanade proteomic platform for the proteomic analysis, and S. Graindorge and D. Pflieger of the IBMP bioinformatics core facility for the bioinformatic analysis. We thank Miroslava Schaffer and Wojciech Wietrzynski for help with FIB milling and cryo-ET acquisition. We thank Jürgen Plitzko and Wolfgang Baumeister for access to cryo-EM instrumentation and support. This work has benefitted from the facilities and expertise of the Biophysical and Structural Chemistry platform (BPCS) at IECB, CNRS UMS3033, Inserm US001, University of Bordeaux. The mass spectrometry instrumentation was funded by the University of Strasbourg, IdEx "Equipement mi-lourd" 2015. This work was supported by a European Research Council Starting Grant (TransTryp ID:759120) to Y.H., by the LabEx consortium 'MitoCross' (ANR-11-LABX-0057_MITOCROSS), by the ITI 2021-2028 program of the University of Strasbourg, CNRS, and Inserm supported by IdEx Unistra (ANR-10-IDEX-0002), and EUR IMCBio (ANR-17-EURE-0023) under the framework of the French Investments for the Future Program to P.G. and L.D., and by the Agence Nationale de la Recherche (ANR) grants [MITRA, ANR-16-CE11-0024-02], [DAMIA, ANR-20-CE11-0021], and [ARAMIS, ANR-21-CE12] to Y.H., P.G., and L.D. Additional funding was provided by the Max Planck Society and the Helmholtz Zentrum München to F.W. and B.E., and the Nederlandse Organisatie voor Wetenschappelijke Onderzoek (Vici 724.016.001) to F.F., as well as by an Alexander von Humboldt Postdoctoral Fellowship to F.W.

Author contributions

F.W., L.D., P.G., and Y.H. designed and coordinated the experiments. H.M. and T.S.-G. generated the amiRNA strains and analyzed them. T.S.-G. purified mitochondria. F.W. purified the mitochondrial ribosomes. H.S. and F.W. acquired the cryo-EM data and performed the single-particle analysis. R.E., S.P., F.F., and B.E. collected and analyzed the cryo-ET data, including subtomogram averaging. F.W. built the atomic models and interpreted the structures. L.K. performed the mass spectrometry experiments. F.W. prepared and assembled the figures. F.W. wrote the manuscript, which was edited by T.S.-G., R.E., H.S., H.M., L.K., S.P., F.F., B.E., P.G., L.D., and Y.H.

Competing interests

The authors declare no competing interests.

Additional information


Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-021-27200-z>.

Correspondence and requests for materials should be addressed to Philippe Giegé, Laurence Drouard or Yaser Hashem.

Peer review information *Nature Communications* thanks the anonymous reviewers for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021

DISCUSSION & PERSPECTIVES

VERS LA CARACTERISATION DES AJOUTS POST-TRANSCRIPTIONNELS DES TRANSCRITS MITOCHONDRIAUX DE *C. REINHARDTII*

Des études antérieures à ma thèse ont montré qu'il existe trois types d'ajouts de nt en 3' des ARNm mitochondriaux chez l'algue verte *C. reinhardtii* (Zimmer *et al.*, 2009; Salinas-Giegé *et al.*, 2017; Cahoon and Qureshi, 2018; Gallaher *et al.*, 2018). Ces ajouts se font après la transcription, indépendamment de toute matrice. Les premiers observés sont les queues polyA- et polyU-riches, dont les sites d'addition peuvent être dans la séquence codante ou dans la 3'UTR, donc plutôt sur les ARNm tronqués (Zimmer *et al.*, 2009). Plus récemment, des queues polyC-riches ont été identifiées, majoritairement en 3' des ARNm matures (Salinas-Giegé *et al.*, 2017; Gallaher *et al.*, 2018). Ces trois types de queues ont en commun d'être homo- ou hétéro-polymériques : d'autres nt y sont ajoutés, en moindre fréquence. C'est pour cela qu'elles sont appelées polyX-riches.

Si des queues polyA ou polyU avaient déjà été observées dans les mitochondries d'autres organismes, ce n'était pas le cas des queues polyC-riches, spécifiques des mitochondries des algues de la classe des Chlorophycées (Salinas-Giegé *et al.*, 2017). Cependant, les enzymes responsables de l'ajout de ces trois types de queues ne sont pas connues, pas plus que les rôles biologiques de ces queues. L'objectif principal de mon travail de thèse a été d'identifier la ou les enzymes responsables de la polycytidylation des ARNm mitochondriaux de *C. reinhardtii* et potentiellement d'identifier les enzymes responsables de leur polyadénylation et de leur polyuridylation.

Préalablement à mon arrivée, l'équipe a mené des analyses phylogéniques sur les enzymes de *C. reinhardtii* de la famille des NT (Salinas-Giegé *et al.*, 2017). Ces analyses ont permis d'établir une liste de quatre NT candidates (PAP2, PAP4, PAP6 et PAP10), spécifiques aux Chlorophycées, potentiellement capables d'ajouter des queues riches en C dans les mitochondries. Des études protéomiques et biochimiques ont permis de démontrer que PAP4 et PAP6 sont exclusivement mitochondriales. Des mutants dans lesquels l'expression des gènes *PAP4* et *PAP6* est diminuée ont été générés. Ceux-ci nous ont permis d'une part d'analyser la stabilité des ARNm mitochondriaux par des expériences de RT-qPCR et d'autre part de quantifier leur niveau de

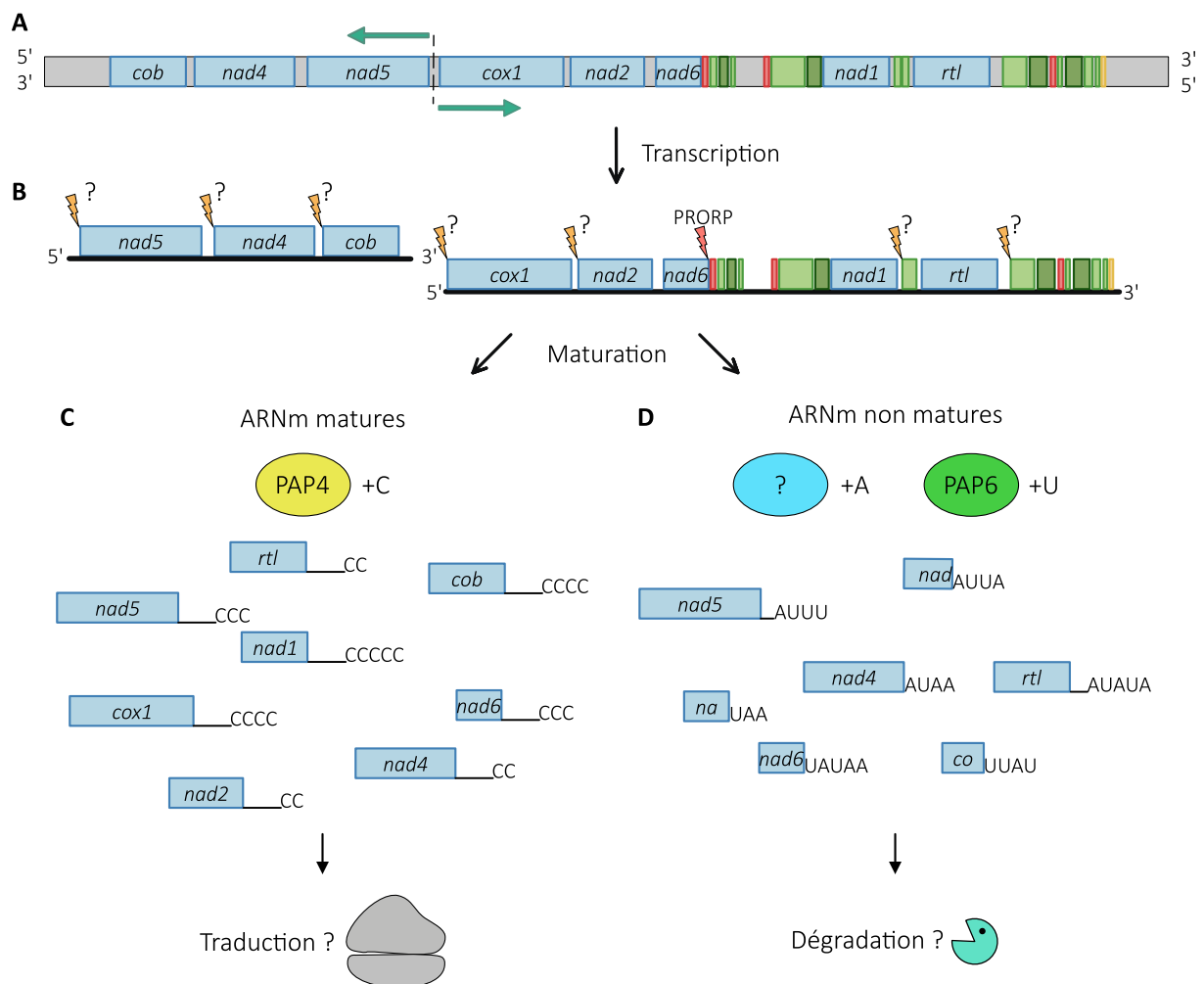


Figure D-1 : Modèle des ajouts de queues sur les ARNm dans les mitochondries de *C. reinhardtii*
Schéma du modèle d'addition des différentes queues proposé dans cette thèse.

A : Génome mitochondrial de *C. reinhardtii* avec l'origine de transcription bidirectionnelle représentée par la ligne pointillée et les 2 flèches **turquoise**. **B** : Deux transcrits d'ARN primaires résultants de la transcription du génome, qui seront clivés en 5' des ARNm et ARNr par une ou plusieurs enzymes inconnues (éclair **orange**) et en 5' de l'ARNt Q par la PRORP (éclair **rouge**). **C** : les ARNm matures semblent polycytidylés par la protéine PAP4 (**jaune anisé**), puis seront ensuite traduits grâce aux mitoribosomes (**gris**). **D** : les ARNm non matures auront des queues polyA (enzyme inconnue – **bleu**) et/ou polyU-riches (par PAP6 – **vert**), servant de signal de dégradation (par un ou des enzymes inconnues - **turquoise**).

Bleu : gènes et ARNm de protéines ; **Rouge** : gènes d'ARNt ; **Vert clair** : gènes d'ARNr de la grande sous-unité du mitoribosome ; **Vert foncé** : gènes d'ARNr de la petite sous-unité du mitoribosome ; **Jaune** : gène de l'ARN L2b.

polynucléotidylation par des expériences de 3'RACE-seq, dans ces souches mutantes. Nos résultats suggèrent un rôle différent des protéines PAP4 et PAP6 dans les mitochondries.

I. PAP4 et la polycytidylation pour la stabilisation des transcrits ?

L'analyse des trois mutants de PAP4 a montré trois types d'effets sur les différents transcrits mitochondriaux. Une diminution globale de l'expression relative des transcrits *cob*, *nad4*, *nad5*, *cox1* et *nad2* est observée. Le transcrit *nad6* n'est pas affecté, avec une expression similaire à la souche de référence. En revanche, pour les deux derniers transcrits, *nad1* et *rtl*, les résultats sont plus divergents entre les trois mutants, avec soit aucun effet soit une augmentation de l'expression du transcrit. Il est intéressant de noter que ces trois « catégorisations » des résultats corrélient avec trois environnements géniques différents (Figure D-1B). Dans le transcrit primaire polycistronique, les transcrits dont l'expression baisse sont suivis d'un autre ARNm en aval (ou la séquence télomérique pour *cob*). Le transcrit *nad6*, qui n'est pas affecté, est en amont d'un ARNt. Enfin, les ARNm des gènes *nad1* et *rtl*, dont l'expression est soit non affectée soit augmentée, sont en amont d'ARNr. Ces premiers résultats suggèrent une implication de PAP4 dans la stabilisation des transcrits suivis d'un ARNm.

Puisque ces trois groupes ne sont pas affectés pareillement dans les mutants de PAP4, il se pourrait qu'il y ait un lien entre le clivage du transcrit et l'activité de PAP4. En effet, lors de la maturation des deux transcrits primaires polycistroniques, ceux-ci vont être clivés afin d'obtenir des ARN monocistroniques. La seule enzyme de clivage connue aujourd'hui dans les mitochondries de *C. reinhardtii* est la PRORP qui permet la coupure de l'extrémité 5' des ARNt (Bonnard *et al.*, 2016). Puisque le transcrit *nad6* se situe en amont d'un ARNt, cela implique que la coupure par PRORP de l'extrémité 5' de l'ARNt permet aussi la libération de l'extrémité 3' de l'ARNm *nad6*. La ou les enzymes impliquées dans les coupures des ARNm et des ARNr ne sont en revanche pas encore identifiées. Concernant la maturation des ARNm, il a été montré que la séquence 3'UTR de chaque transcrit correspond à la séquence intergénique le séparant du transcrit en aval. Cela implique que la coupure au niveau de l'AUG d'un transcrit libère l'extrémité 3' du transcrit précédent (Figure D-1B). Au vu des différences de taille et de structure entre les ARNm et les ARNr, il est possible que deux enzymes différentes soient responsables de leurs clivages respectifs. Il y aurait donc trois enzymes permettant le clivage de l'extrémité 3' des trois « groupes » d'ARNm.

Pour confirmer cette hypothèse, il est donc nécessaire d'identifier et d'étudier la ou les enzymes impliquées dans les clivages des ARNm et ARNr. Une première approche pourrait être de réaliser des expériences de co-immunoprécipitations avec la protéine PAP4 comme appât, puis d'étudier par des approches biochimiques et par génétique inverse les interactants protéiques pouvant avoir une fonction ribonucléase. Ces analyses permettraient d'améliorer notre compréhension des processus de maturation des ARN et son déroulement.

Lors de l'analyse des ajouts post-transcriptionnels des queues des transcrits dans les mutants de PAP4, une diminution globale des queues polyC-riche a été observée. Il y a des divergences entre les différents transcrits des trois mutants concernant les ajouts de A et de U. Nous constatons toutefois une augmentation générale des queues polyA-riche. Cependant les proportions d'ajouts de U sont trop différentes entre les transcrits des trois souches mutantes pour en déduire un motif de polyuridylation dans les mutants de PAP4. Contrairement aux résultats de l'expression des transcrits, nous ne voyons pas d'effet spécifique d'un transcrit ou d'un groupe de transcrits. Néanmoins, la diminution des proportions de polyC pour tous les messagers dans les mutants PAP4 indique que cette protéine semble responsable de l'ajout de ces séquences riches en C.

Selon une étude, PAP4 a une activité de polyadénylation *in vivo* dans des bactéries délétées de leur PAP endogène (Zimmer *et al.*, 2009). Les auteurs avaient observé qu'un quart des queues ajoutées étaient composées uniquement de A, les autres étant hétérogènes, riches en A/U avec quelques C et G. De plus, ils ont montré que PAP4 peut ajouter *in vitro* des A en 3' d'une séquence ARN, bien que l'ajout des trois autres nt n'ait pas été testé. Or, d'après nos analyses de 3'RACE-seq, PAP4 n'est pas impliquée dans l'addition de A, puisqu'en son absence les proportions de A ajoutés augmentent. De plus, la séquence de PAP4 a des similarités avec les NT de classe II, et plus particulièrement les *CC-adding enzyme* par l'absence du motif BF, et non avec les PAP, canoniques ou non, de classe I (Neuenfeldt *et al.*, 2008). De plus, il a déjà été montré que des expériences *in vitro* peuvent ne pas refléter la spécificité d'une enzyme pour son substrat *in vivo* (Yehudai-Resheff and Schuster, 2000). Des tests d'activité *in vitro* sont en cours dans le laboratoire, afin de déterminer la potentielle activité de polycytidylation de PAP4. L'addition de queues polyC-riche n'a été observée que dans les mitochondries des Chlorophycées. Aucun ajout naturel de ce type n'a été vu ailleurs dans le règne du vivant. Dès leur découverte, il a été proposé que ces queues soient impliquées dans la stabilisation

et/ou la traduction des ARNm (Salinas-Giegé *et al.*, 2017; Cahoon and Qureshi, 2018). En effet, ces queues se trouvent majoritairement sur les ARNm matures non tronqués et sont ajoutées essentiellement juste après la 3'UTR, au site de clivage. Ceci implique que la polycytidylation se fasse rapidement après le clivage des transcrits polycistroniques, afin de stabiliser le transcrit et/ou éviter sa dégradation par des ribonucléases.

Nous avons vu que, dans une souche sauvage, tous les transcrits ont des queues polyC en proportions différentes. Par exemple, le transcrit *rtl* est le moins polycytidylé (moins de 5%), tandis que ceux de *cob* et *cox1* sont ceux avec le plus d'ajouts de C (entre 30 et 40%, selon le fond génétique de la souche étudiée). Il est donc possible que tous les transcrits ne nécessitent pas les mêmes proportions de polyC pour un rôle équivalent. Dans les mutants de PAP4, une diminution de l'expression relative de certains ARNm mitochondriaux a été observée, bien que les proportions de queues polyC ajoutées soient moindres pour tous les transcrits. Sachant que les trois souches mutées étudiées ici sont des mutants KD, la proportion de protéines PAP4 résiduelles dans les mitochondries est peut-être suffisante pour ajouter assez de queues polyC à certains transcrits. Ainsi, leur expression relative ne sera pas affectée contrairement à d'autres, dont l'expression relative sera alors plus basse. Pour les mutants CLiP uniquement, il est possible de compléter les lignées afin de confirmer que la mutation de PAP4 est bien à l'origine des résultats décrits. De plus, l'utilisation d'un outil tel CRISPR-Cas permettrait l'obtention de mutants KO (*knock-out*) où les effets seraient donc plus marqués que dans les mutants KD. Grâce à cela, nous pourrions savoir si l'absence totale d'activité de PAP4 provoque une diminution de l'expression relative observée uniquement pour certains transcrits mitochondriaux comme dans les mutants KD, ou pour tous les transcrits mitochondriaux. Cela nous permettrait donc de conclure quant au fait que les queues polyC soient un facteur général de la stabilité des transcrits mitochondriaux ou spécifique à certains ARNm chez *C. reinhardtii*.

Ces résultats sont cohérents avec un rôle de stabilisation ou de protection des transcrits par les queues polyC-riche (Figure D-1C) : la diminution de la proportion de queue polyC-riche rendrait les transcrits moins stables, peut-être plus vulnérables aux ribonucléases. Enfin, les queues polyA-riche sont supposées être impliquées dans la dégradation des transcrits, comme chez les plantes terrestres (Zimmer *et al.*, 2009; Levy and Schuster, 2016), et leur augmentation dans les mutants de PAP4 pourrait signifier que si les transcrits ne sont pas polycytidylés,

ils sont dégradés. Lors de l'étude sur la circularisation des ARNm mitochondriaux, les auteurs montrent que les ARNm mitochondriaux *C. reinhardtii* sont présents sous forme linéaire et circulaire (Cahoon and Qureshi, 2018). D'après leurs expériences, les ARNm circulaires peuvent être traduits, puisque certains transcrits sous forme circularisée ont été trouvés en association avec des mitoribosomes. Les auteurs ont identifié des ARNm circularisés avec des ajouts de C et de U, ils suggèrent donc que ces queues permettent la protection du transcrit contre la dégradation, entre leur clivage et leur circularisation. Ces queues seraient ensuite clivées, puisque la majorité des ARNm circularisés en association avec les ribosomes n'ont pas d'addition de nt. Les auteurs suggèrent que les ARNm circularisés possédant encore des ajouts de C voient leur traduction inhibée. Toutefois, ils n'ont pas testé l'association des ARNm mature linéaires avec les mitoribosomes. Bien que ces résultats soient intéressants, leurs expériences montrent une très faible quantité d'ARNm circularisés, avec une nécessité de faire deux PCR nichées successives, à partir d'ARN extraits d'échantillons enrichis en mitochondries. En comparaison, pour nos expériences de 3'RACE-seq, les ARNm mitochondriaux sont détectés avec une PCR classique à partir d'ARN total. Or, si seuls les ARNm circularisés peuvent être traduits, leur quantité est-elle suffisante pour synthétiser efficacement les protéines nécessaires aux complexes respiratoires ? Il serait intéressant de regarder si les ARNm matures linéaires sont également associés aux mitoribosomes, et si oui, s'ils contiennent des séquences polyC-riches. En effet, il pourrait y avoir une coexistence de la traduction des ARNm circulaires et linéaires. Cela permettrait également de déterminer si les queues polyC servent uniquement à stabiliser les transcrits ou si elles sont aussi impliquées dans l'initiation de la traduction.

Par ailleurs, lors de l'étude sur la circularisation des ARNm mitochondriaux, des ajouts de U ont été trouvés parmi les ARNm circulaires, mais majoritairement en tandem avec des queues polyC (Cahoon and Qureshi, 2018). Dans l'étude de l'équipe sur les queues polyC-riches, il a aussi été noté que celles-ci sont souvent précédées par des A/U (Salinas-Giegé *et al.*, 2017). Dans nos analyses, ce type de queues aura été considéré comme polyC-riches, puisque contenant une majorité de C. Il est donc nécessaire d'approfondir nos analyses afin de déterminer le nombre de queues polyC-riches commençant par des U, afin de mieux caractériser cet ajout.

II. PAP6 et polyuridylation pour la dégradation des transcrits ?

Concernant les mutants de PAP6, l'analyse de l'expression des transcrits mitochondriaux des deux mutants montre une augmentation de l'expression de tous les ARNm mitochondriaux, à l'exception de *nad6* qui n'est pas affecté. À nouveau, *nad6* a de différent son environnement génique, étant le seul transcrit suivi d'un ARNt. Là aussi, nous pouvons imaginer un lien entre l'activité de PAP6 et l'environnement du transcrit, mais cela ne pourra être démontré que par la connaissance et l'étude de la ou des enzymes impliquées dans les clivages des ARNm et ARNr.

Dans les mutants de PAP6, les taux de polycytidylation varient selon les transcrits et les souches étudiées : certains sont diminués, d'autres augmentés et pour d'autres encore, il n'y a pas d'effets. De telles variations peuvent indiquer qu'il n'y a pas de lien entre les queues polyC et l'activité de PAP6. En revanche, dans les deux mutants, l'adénylation est augmentée pour tous les transcrits, tandis que la polyuridylation diminue également pour tous les transcrits, à l'exception de *nad6* dans la lignée *amipap6-64* dont les taux de polyuridylation sont similaires à ceux de la souche de référence. Or le transcrit *nad6* est le moins polyuridylé, avec seulement 1% chez la souche de référence. Des variations d'une si petite proportion seront donc plus difficilement mesurables. Ces résultats montrent une implication de PAP6 dans l'ajout de queues polyU-riches. Ceci est confirmé par nos tests d'expression *in vitro*, qui montrent que PAP6 peut ajouter tous les nt, à condition qu'elle ait d'abord ajouté un U. Elle possède donc une activité de polyuridylation. Ceci est également cohérent avec les analyses phylogéniques montrant que PAP6 est proche des TUTases. Toutefois, comme nous l'avons vu précédemment, les tests d'activité *in vitro* ne reflètent pas nécessairement la spécificité d'une enzyme pour son substrat *in vivo*. Néanmoins, cela est aussi cohérent avec l'hétérogénéité de la composition des queues trouvées en 3' des transcrits mitochondriaux.

Contrairement à l'addition de C qui n'avait jamais été décrite, quelques cas d'addition de U ont été trouvés dans les mitochondries d'autres organismes, mais avec des rôles qui semblent spécifiques à chaque organisme étudié. Chez les trypanosomes, les ARN guides permettant l'édition des transcrits doivent être uridylés afin d'être fonctionnels (Aphasizhev, Aphasizheva and Simpson, 2003; Suematsu *et al.*, 2016). Chez les myxomycètes, des queues polyU ont été identifiées sur les ARNm édités (Horton and Landweber, 2000). Chez l'Homme, des queues polyU ont été trouvées dans des mutants de la PNPase uniquement, suggérant un rôle dans la dégradation (Slomovic and Schuster, 2008; Jedynak-Slyvka, Jabczynska and Szczesny, 2021).

Il semble donc qu'au cours de l'évolution, l'uridylation soit apparue dans les mitochondries plusieurs fois avec des rôles distincts selon les organismes. La PAP6, spécifique de la classe des Chlorophycées, pourrait donc présenter sa propre spécificité quant à l'ajout des queues polyU-riques. Quel serait alors le rôle des queues polyU-riques dans les mitochondries de *C. reinhardtii* ?

Lors de la première observation de queues polyU-riques dans les mitochondries de l'algue, il a été proposé que leur addition soit un processus similaire aux queues polyA, agissant comme signal ou intermédiaire de dégradation (Zimmer *et al.*, 2009). En effet, les additions de U ont été trouvées essentiellement en 3' d'ARNm tronqués, et leur détection avait nécessité un protocole de PCR en plusieurs étapes. Ces difficultés sont corroborées par le peu de séquences polyU-riques que nous avons obtenus lors de nos expériences de 3'RACE-seq. Cela pourrait indiquer que les queues polyU-riques sont des intermédiaires dans le processus de dégradation, et sont donc peu détectées à l'équilibre. Les résultats présentés dans cette thèse soutiennent ces hypothèses (Figure D-1D). En effet, dans les mutants de PAP6, nos analyses de l'expression relative des transcrits mitochondriaux montrent une augmentation de l'accumulation des différents transcrits, à l'exception de *nad6*. Ces résultats sont similaires à ceux obtenus lors de l'étude du système AGS1-AHG2 responsable de la régulation de la polyadénylation des transcrits mitochondriaux chez *A. thaliana* (Hirayama *et al.*, 2013). Dans ce système, AGS1, une PAP de type bactérien, polyadényle les transcrits mitochondriaux et AHG2, une PARN (*poly(A) specific ribonuclease*), dégrade les ARN ciblés. Dans cette étude, des mutants de AHG2 montrent une augmentation de l'accumulation relative des transcrits en absence de l'enzyme de dégradation. Aucune étude d'un mutant de AGS1 n'a été faite à ce jour, il serait alors intéressant de faire des expériences similaires dans le mutant de AGS1, afin de voir si les résultats obtenus sont identiques. En effet, si en l'absence de l'enzyme de dégradation il y a une accumulation des transcrits, nous pouvons supposer qu'en l'absence de AGS1, ces transcrits non polyadénylés ne seront pas non plus dégradés, entraînant une augmentation de leur expression relative qui sera mesurable par RT-qPCR. Cette logique peut être transposée à notre cas, aux queues polyU-riques ajoutées par PAP6 qui serviraient donc de signal de dégradation dans les mitochondries de *C. reinhardtii*. Cette hypothèse doit être confirmée par des expériences de co-immunoprécipitations qui nous permettront d'étudier l'interactome de PAP6 et donc d'identifier une ou plusieurs enzymes candidates pour la dégradation des ARN

mitochondriaux polyadénylés dans cette espèce. Il serait également intéressant d'étudier la polyuridylation des ARNr ou des ARNt par PAP6 *in vivo* ou *in vitro* puisque cela n'a jusqu'alors pas été réalisé.

Les études précédentes évoquent des queues A/U-riches (Zimmer *et al.*, 2009; Salinas-Giegé *et al.*, 2017). Chez les trypanosomes, les transcrits mitochondriaux matures ont de longues queues A/U-riches (entre 200 et 300 nt) permettant leur reconnaissance par les mitoribosomes et activant ainsi leur traduction (Aphasizhev *et al.*, 2016). Au vu de l'hétérogénéité des queues des transcrits mitochondriaux de *C. reinhardtii*, il est difficile de les catégoriser strictement. Néanmoins, nos résultats suggèrent que la protéine PAP6 est responsable de l'ajout de U, mais pas de A, puisque dans le mutant il y a une diminution des proportions de queues polyU-riches mais une augmentation des queues polyA-riches. Mais alors, quel est le rôle de l'addition des queues polyA ? Des analyses approfondies de nos résultats de 3'RACE-seq sont nécessaires afin de mieux étudier les queues polyA et polyU-riches. Cela nous permettra de conclure à une coexistence des queues polyA et polyU-riches dans les mitochondries de *C. reinhardtii*, ou à une hétérogénéité des séquences polyA/U. Aussi, l'identification et l'étude de l'enzyme ajoutant les A aidera à déterminer la fonction réelle des queues polyA, et donc permettra de nous éclairer quant à l'existence ou non des queues polyA/U.

III. Qu'en est-il de la polyadénylation des transcrits ?

La conservation du rôle des queues polyA comme signal de dégradation dans les mitochondries de *C. reinhardtii* est discutée par plusieurs études (Zimmer *et al.*, 2009; Cahoon and Qureshi, 2018). Les queues polyA servent de signal de dégradation dans les bactéries et les archées. De par leur origine bactérienne, les mitochondries des plantes terrestres ont conservé ce rôle. Il est donc probable que ce rôle ait été conservé aussi dans les mitochondries des algues. Cela est renforcé par l'observation des ajouts de A essentiellement sur des ARNm tronqués, à de multiples sites dans la séquence 3'UTR (Zimmer *et al.*, 2009; Cahoon and Qureshi, 2018). Lors de l'analyse des ARNm circularisés, aucun ajout de A n'a été identifié (Cahoon and Qureshi, 2018). Pour les auteurs, c'est un argument de plus en faveur du rôle des queues polyA dans la dégradation, puisque les ARNm polyadénylés seraient dégradés avant la circularisation.

Nos analyses montrent une augmentation de la proportion de queues polyA dans les mutants

de PAP4 et de PAP6. Pour les mutants de PAP4, cette augmentation peut être corrélée avec la diminution de l'expression relative de certains transcrits et avec la diminution des proportions d'ajouts de C. Il est possible qu'il y ait une plus forte addition de queues polyA dans le mutant de PAP4 afin que les transcrits non polycytidylés soient dégradés. Dans les mutants PAP6, la diminution des queues polyU semble être compensée par l'augmentation des queues polyA, mais une augmentation de l'expression relative des ARNm mitochondriaux est observée en RT-qPCR. Il est alors possible qu'il y ait deux voies de dégradation, l'une activée par les queues polyA, l'autre par les queues polyU, n'ayant pas la même efficacité. Cela expliquerait l'augmentation de l'expression relative des transcrits dans les mutants de PAP6. Il se peut que la voie de dégradation par les A soit secondaire, expliquant son efficacité moindre. Une différence de dégradation par les A que par les U expliquerait également le fait qu'il y ait plus de queues polyA-riches détectées dans nos analyses que de polyU. En effet, comme nous l'avons vu précédemment, les intermédiaires de dégradation ne sont généralement presque pas détectables à l'équilibre, car le *turnover* des ARN est un processus rapide. Enfin, l'apparition de la polyuridylation comme deuxième signal de dégradation a peut-être permis à ces organismes de renforcer le processus de dégradation dans les mitochondries. Cependant, nous ne pouvons pas exclure que leurs fonctions soient associées, et que les ajouts des queues polyA/U-riches se fassent en tandem par deux enzymes, bien qu'ayant la même fonction. L'identification et l'étude des enzymes responsables de la polyadénylation et de la dégradation pourraient aider à la compréhension de ces processus.

IV. Les limites du 3'RACE-seq

Lors de cette thèse, nous avons opté pour la technique du 3'RACE-seq afin d'analyser les extrémités 3' des gènes d'intérêt. Brièvement, cette technique débute par la ligation d'un adaptateur en 3' des ARN qui sont ensuite rétro-transcrits en ADNc. Ceux-ci sont utilisés comme matrices afin d'amplifier par PCR les extrémités 3' des gènes d'intérêt. Pour ajouter les séquences nécessaires au séquençage Illumina, nous devons passer par 2 PCR successives. Enfin, les amplicons de PCR sont séquencés par la technologie Illumina. Cependant, cette technique comporte de nombreux biais. Premièrement, il n'est pas possible de vérifier l'efficacité de la ligation. Or lors de cette réaction, l'ARN doit posséder une extrémité 3'OH accessible pour l'enzyme, ce qui implique que les bases modifiées ou les structures secondaires

des ARN pourraient inhiber la ligation et donc fausser les résultats obtenus. De plus, la ligation de l'adaptateur ne peut sans doute pas se faire pour les ARNm circularisés. Bien qu'à priori restreinte, cette population d'ARN pourrait donc ne pas être incluse dans nos analyses. Outre les biais associés à cette réaction de ligation, la technique de 3'RACE-seq repose sur de multiples PCR. Les deux PCR successives évoquées précédemment sont réalisées avec la GoTaq DNA polymérase, moins fidèle que certaines polymérases à haute-fidélité. Nous avons choisi cette enzyme car les amplifications obtenues sont de petites tailles, variant entre 300 et 500 pb, ce qui limite le risque d'apparition d'une mutation. Or, au vu de la composition des queues trouvées (riches en A, U ou C), il est probable que ces motifs répétés de nucléotides favorisent les erreurs dans la réaction de polymérisation en provoquant le « glissement » de l'enzyme polymérase sur la séquence. Enfin, le séquençage Illumina a également des limites, notamment dans la lecture de longues séquences homopolymériques, puisque la technique comprend également des étapes de polymérisation qui peuvent introduire des erreurs. Nous avons testé cela avec des séquences Spike-in, ayant une longueur de 8C, 16C ou 22C, et le séquençage est bien moins efficace pour les séquences 16C et 22C. Des soucis similaires avec les longues queues polyA ont déjà été reportés (Chang *et al.*, 2014). Tout cela nous indique que les tailles que nous observons pour les trois types de queues ne sont peut-être pas les tailles réelles *in vivo*. De plus, il est possible que la composition exacte des queues soit biaisée, par les étapes de PCR et de séquençage notamment.

Pour contourner ces biais, le séquençage Nano3P-seq (*Nanopore 3'end-capture sequencing*) offre une alternative enthousiasmante (Oxford Nanopore Technologie) (Begik *et al.*, 2023). Cette méthode permet de coupler le séquençage des extrémités 3' *via* la technologie Nanopore à la quantification de l'abondance des ARN, et ne nécessite aucune étape de ligation ou d'amplification. De plus, elle semble efficace sur des ARN polyadénylés ou non, peu importe la composition en nt des queues en 3', contrairement au dRNA-seq (*direct RNA sequencing*), méthode basée sur la présence de queues polyA d'au moins 10nt (Workman *et al.*, 2019; Tudek *et al.*, 2021). Des tests menés dans l'étude du Nano3P-seq ont montré que cette méthode permet d'identifier correctement des changements de nt dans des queues hétéropolymériques (Begik *et al.*, 2023). Une hétérogénéité des queues polyA a pu être observée, avec la présence de G et de U. Puisque cela n'a pas été fait, il serait intéressant d'essayer cette technique avec des queues riches en C comme celles que nous étudions. Afin d'obtenir une meilleure

caractérisation des queues ajoutées en 3' des ARNm mitochondriaux, de leur composition comme de leur longueur, nous pourrions donc utiliser cette technique. De plus, parvenir à quantifier l'abondance des ARN serait essentiel afin de vérifier si la balance entre les ARN nucléaires, chloroplastiques et mitochondriaux est affectée dans nos mutants.

Pour conclure, nos analyses nous ont permis d'identifier PAP4 comme étant l'enzyme de polycytidylation et PAP6 celle de polyuridylation des ARNm mitochondriaux. De plus, nous démontrons que ces deux processus sont spécifiques des mitochondries des algues de la classe des Chlorophycées. Néanmoins, de nombreuses questions restent en suspens concernant le mécanisme de maturation des ARNm mitochondriaux de *C. reinhardtii*. Par quels mécanismes moléculaires les additions de C et de U se réalisent-elles ? PAP4 et PAP6 agissent-elles seules ? Comment sont recrutées ces enzymes au niveau de leur ARN cible ? Quel lien y-a-t-il entre la polyuridylation et la polyadénylation, et quelle est l'enzyme qui réalise cette dernière réaction ? Comment se fait-il que ces queues ajoutées soient si hétérogènes ? Ces réactions se produisent-elles avant ou à la suite des événements de clivages de l'ARN, et quelles sont les enzymes associées à ce dernier processus ? Répondre à ces questions permettra une meilleure compréhension de l'expression des gènes mitochondriaux chez *C. reinhardtii*.

CARACTERISATION DU MITORIBOSOME DE *C. REINHARDTII*

Jusque récemment, les connaissances sur les mitoribosomes des plantes se limitaient au mitoribosome d'*A. thaliana*, publié en 2019 (Waltz *et al.*, 2019). Peu de données étaient disponibles concernant ceux des algues vertes. Chez *C. reinhardtii*, il avait été montré que les ARNr sont fragmentés en différents modules, dispersés sur le génome (Boer and Gray, 1988). Un modèle avait été proposé, dans lequel les fragments d'ARNr s'assembleraient par des appariements de bases, mais cela n'avait jamais été prouvé. De plus, le gène de l'ARNr 5S n'avait pas été identifié dans le génome mitochondrial, il pourrait être importé depuis le noyau, ou remplacé comme c'est le cas chez l'Homme où l'ARNr 5S est remplacé par un ARNt (Brown *et al.*, 2014; Greber *et al.*, 2014, 2015). D'autre part, les ARNm mitochondriaux n'ont pas de 5'UTR, et cela pose la question de savoir comment se fait l'initiation de la traduction. Par ailleurs, puisque les queues polyC-riches sont majoritairement trouvées sur les ARNm matures, il est possible qu'elles soient nécessaires pour la traduction des ARNm, et plus précisément pour l'initiation de la traduction. Enfin, les diverses structures des mitoribosomes qui ont été obtenues au fil des années montrent des différences entre les mitoribosomes et le ribosome bactérien, mais également entre les mitoribosomes des organismes étudiés (Scaltsoyiannes *et al.*, 2022a). Il y a une diversité de mitoribosomes différents, avec des spécificités propres à chacun. Connaître celui de *C. reinhardtii* apporterait non seulement des connaissances structurales et une compréhension supplémentaire concernant le processus de traduction dans les mitochondries de cette algue, mais cela permettrait aussi de compléter l'Histoire évolutive des mitoribosomes, très divergents entre les lignées eucaryotes.

Par des analyses de protéomique, de cryo-EM et de cryo-tomographie, nous avons pu décrire la structure et la composition du mitoribosome de *C. reinhardtii* (Waltz *et al.*, 2021). Celui-ci diffère du ribosome bactérien ancestral mais également du mitoribosome d'*A. thaliana* (Waltz *et al.*, 2019), avec des domaines et une composition protéiques nouveaux. Ont été identifiées 47 protéines pour la LSU et 36 pour la SSU. Cela donne un total de 83 protéines, ce qui est bien plus que les 54 protéines des ribosomes bactériens (Waltz and Giegé, 2020; Kummer and Ban, 2021). Les 54 protéines bactériennes ancestrales sont retrouvées dans le mitoribosome de *C. reinhardtii*, ainsi que 18 protéines en commun avec d'autres mitoribosomes et 11 qui sont spécifiques à l'algue (8 pour la LSU (mL113 à mL119 et une protéine inconnue qui semble être

une PPR) et trois pour la SSU (mS105 à mS107)). Il est important de noter que la plupart des protéines spécifiques de l'algue sont de la famille des α -solénoïdes, avec 6 OPR, 1 mTERF et 2 PPR. Cette famille de protéines est connue pour son implication dans l'expression des gènes des organelles. Cette caractéristique semble être partagée par *P. magna*, qui possède 19 protéines α -solénoïdes spécifiques à son mitoribosome, ainsi que 4 protéines identifiées parmi les protéines spécifiques de *C. reinhardtii*. Mis ensemble, cela reflète l'origine bactérienne des mitoribosomes, leur évolution, leur histoire eucaryotique commune mais également leurs divergences au sein des différentes lignées eucaryotes.

D'après la structure tomographique, les protéines mL113 et mL119 semblent ancrer le mitoribosome à la membrane interne mitochondriale, tandis que les protéines mL115, mL116, mL117 et mL118 interagissent avec des fragments d'ARNr. Nous avons étudié des mutants de certaines protéines spécifiques au mitoribosome de *C. reinhardtii*, à savoir de mL113, mL116, mL117, mL118 et mS105. Dans les mutants des protéines mL113, mL117 et mL118, une diminution de l'accumulation des ARNr de la LSU a été observée. Cela laisse à penser que ces protéines sont essentielles pour l'assemblage ou la stabilisation de la LSU, ce qui est cohérent avec les données structurales. Afin de mieux comprendre leur rôle dans le mitoribosome, il serait intéressant d'obtenir des mutants pour les autres protéines, mL114, mS106 et mS107, puis de purifier et d'analyser les mitoribosomes de tous ces mutants par cryo-EM.

Un autre point commun intéressant entre les mitoribosomes de ces deux algues est que les gènes d'ARNr sont fragmentés en plusieurs modules non contigus sur le génome mitochondrial. Bien que particulière, cette fragmentation des ARNr n'est pas spécifique de ces algues, puisqu'elle est également retrouvée chez *P. falciparum*, qui possède 39 fragments d'ARNr (Feagin *et al.*, 2012; Dass *et al.*, 2022). La question s'est longtemps posée de savoir comment ces fragments sont assemblés en un ARNr fonctionnel. La résolution des structures a permis de voir que les ARNr du centre catalytique (*core*) du mitoribosome sont stabilisés par des appariements de bases, conformément au modèle proposé en 1988 (Boer and Gray, 1988). Néanmoins, les fragments plus en périphérie des sous-unités sont stabilisés par les protéines spécifiques à l'algue. Cela permet de faire un lien entre la réduction et la fragmentation des ARNr et le gain en protéines de ces mitoribosomes. Aussi, pour les mitoribosomes de *C. reinhardtii* et *P. magna*, les gènes de l'ARNr 5S n'avaient jamais été identifiés, ils étaient donc considérés comme absents. Or, pour les deux mitoribosomes, les structures ont montré la

présence de l'ARNr 5S. Le gène a été identifié comme celui du fragment d'ARNr L3a pour *C. reinhardtii*, tandis que celui de *P. parva* n'a pas été trouvé pour le moment.

Enfin, la question de l'initiation de la traduction chez *C. reinhardtii* reste ouverte. En effet, les ARNm mitochondriaux de cette algue n'ont pas de séquence 5'UTR, généralement essentielle pour le recrutement du mitoribosome et l'initiation de la traduction. Les ARNm mitochondriaux humains n'ont pas non plus de 5'UTR, mais ils ont en aval de l'AUG des séquences riches en U, qui sont supposées interagir avec la protéine mS39 de la SSU, afin d'initier la traduction. Ces séquences ne sont pas retrouvées chez *C. reinhardtii*. Comme nous l'avons vu, il est possible que les queues polyC-riches soient impliquées d'une manière ou d'une autre dans l'initiation de la traduction. Il est possible d'imaginer un mécanisme similaire aux ARNm nucléaires, où le facteur d'initiation eIF4G recrute une protéine se liant aux queues polyA permettant la circularisation de l'ARNm et la stabilisation du complexe d'initiation. Cependant, nous ne connaissons pas les facteurs mitochondriaux impliqués dans l'initiation de la traduction mitochondriale. Une autre possibilité est que la circularisation des ARNm permette de « recréer » des séquences 5'UTR afin d'avoir une interaction avec la SSU (Cahoon and Qureshi, 2018). Deux possibles facteurs d'initiation mitochondriaux ont été identifiés chez *C. reinhardtii*, MIF2 (Cre08.g376800) et MIF3 (Cre09.g398215) dans la version 6.1 de l'annotation du génome nucléaire de l'algue. L'étude de ces deux protéines pourrait aider à identifier tous les facteurs impliqués et le déroulement de l'initiation de la traduction mitochondriale dans l'algue. Par exemple, un tag fusionné à la protéine d'intérêt pourra être exprimé dans des algues, afin de co-immunopurifier les partenaires de ces facteurs, qui seront vraisemblablement les différents composants du complexe d'initiation.

Pour conclure, l'obtention de la structure du mitoribosome de *C. reinhardtii* a permis d'obtenir des réponses quant à l'assemblage et la stabilisation des ARNr. Néanmoins, il reste des questions sans réponses, notamment en ce qui concerne l'initiation de la traduction. Une longue histoire reste encore à compléter afin d'élucider entièrement le processus d'expression des gènes mitochondriaux de l'algue verte *C. reinhardtii*.

En conclusion, mon travail de thèse a permis d'identifier les enzymes responsables de la polytydylation et de la polyuridylation des ARNm mitochondriaux de l'algue verte *C. reinhardtii*, PAP4 et PAP6, respectivement. Ces deux processus semblent être spécifiques des mitochondries des algues de la classe des Chlorophycées. De plus, la caractérisation de la composition et de la structure du mitoribosome de cette algue a permis la compréhension de l'assemblage des fragments d'ARNr dans les sous-unités du ribosome, stabilisés notamment par des protéines spécifiques de cet organisme. Une étude plus approfondie de ces protéines pourrait permettre de mieux comprendre l'initiation de la traduction par exemple, et peut-être de faire le lien entre les ajouts de queues polyC-riche et la traduction des transcrits.

De par ces deux aspects, nous avons plus d'indications sur les mécanismes d'expression des gènes mitochondriaux et leur spécificité dans *C. reinhardtii* par rapport aux autres organismes déjà étudiés. De nombreuses questions persistent cependant, auxquelles il faudra répondre afin d'avoir une compréhension complète des différentes étapes de l'expression du génome mitochondrial de *C. reinhardtii*.

BIBLIOGRAPHIE

- Akil, C. and Robinson, R.C. (2018) 'Genomes of Asgard archaea encode profilins that regulate actin', *Nature*, 562, pp. 439–443.
- De Almeida, C. *et al.* (2018) 'RNA uridylation: a key posttranscriptional modification shaping the coding and noncoding transcriptome', *Wiley Interdisciplinary Reviews: RNA*, 9(1).
- Alverson, A.J. *et al.* (2011) 'Origins and Recombination of the Bacterial-Sized Multichromosomal Mitochondrial Genome of Cucumber', *The Plant Cell*, 23, pp. 2499–2513.
- Amunts, A. *et al.* (2015) 'The structure of the human mitochondrial ribosome', *Science*, 348(6230), pp. 95–98.
- Aphasizhev, R. *et al.* (2016) 'Constructive edge of uridylation-induced RNA degradation', *RNA Biology*, 13(11), pp. 1078–1083.
- Aphasizhev, R. and Aphasizheva, I. (2011) 'Mitochondrial RNA processing in trypanosomes', *Research in Microbiology*, 162(7), pp. 655–663.
- Aphasizhev, R., Aphasizheva, I. and Simpson, L. (2003) 'A tale of two TUTases', *Proceedings of the National Academy of Sciences of the United States of America*, 100(19), pp. 10617–10622.
- Aphasizheva, I. *et al.* (2011) 'Pentatricopeptide Repeat Proteins Stimulate mRNA Adenylation/Uridylation to Activate Mitochondrial Translation in Trypanosomes', *Molecular Cell*, 42(1), pp. 106–117.
- Del Arco, A. and Satrústegui, J. (2005) 'New mitochondrial carriers: an overview', *Cellular and Molecular Life Sciences*, 62, pp. 2204–2227.
- Armenteros, J.J.A. *et al.* (2019) 'Detecting sequence signals in targeting peptides using deep learning', *Life Science Alliance*, 2(5).
- Atkinson, G.C. *et al.* (2012) 'Evolutionary and genetic analyses of mitochondrial translation initiation factors identify the missing mitochondrial IF3 in *S. cerevisiae*', *Nucleic Acids Research*, 40(13), pp. 6122–6134.
- Atteia, A. *et al.* (2009) 'A Proteomic Survey of *Chlamydomonas reinhardtii* Mitochondria Sheds New Light on the Metabolic Plasticity of the Organelle and on the Nature of the α -Proteobacterial Mitochondrial Ancestor', *Molecular Biology and Evolution*, 26(7), pp. 1533–1548.

- Auchincloss, A.H. *et al.* (2002) 'Characterization of Tbc2, a nucleus-encoded factor specifically required for translation of the chloroplast psbC mRNA in *Chlamydomonas reinhardtii*', *The Journal of Cell Biology*, 157(6), pp. 953–962.
- Avendaño-Monsalve, M.C., Ponce-Rojas, J.C. and Funes, S. (2020) 'From cytosol to mitochondria: the beginning of a protein journey', *Biological Chemistry*, 401(6–7), pp. 645–661.
- Backes, S. and Herrmann, J.M. (2017) 'Protein Translocation into the Intermembrane Space and Matrix of Mitochondria: Mechanisms and Driving Forces', *Frontiers in Molecular Biosciences*, 4(83).
- Balczun, C. *et al.* (2005) 'Two adjacent nuclear genes are required for functional complementation of a chloroplast trans-splicing mutant from *Chlamydomonas reinhardtii*', *The Plant Journal*, 43(5), pp. 636–648.
- Barkan, A. and Small, I. (2014) 'Pentatricopeptide Repeat Proteins in Plants', *Annual Review of Plant Biology*, 65, pp. 415–442.
- Barrell, B.G., Bankier, A.T. and Drouin, J. (1979) 'A different genetic code in human mitochondria', *Nature*, 282, pp. 189–194.
- Barros, M.H. and McStay, G.P. (2020) 'Modular biogenesis of mitochondrial respiratory complexes', *Mitochondrion*, 50, pp. 94–114.
- Basu, U. *et al.* (2020) 'Structure, mechanism, and regulation of mitochondrial DNA transcription initiation', *Journal of Biological Chemistry*, 295(52), pp. 18406–18425.
- Begik, O. *et al.* (2023) 'Nano3P-seq: transcriptome-wide analysis of gene expression and tail dynamics using end-capture nanopore cDNA sequencing', *Nature Methods*, 20, pp. 75–85.
- Bennoun, P. (1982) 'Evidence for a respiratory chain in the chloroplast', *Proceedings of the National Academy of Sciences of the United States of America*, 79, pp. 4352–4356.
- Boehm, E. *et al.* (2017) 'FASTKD1 and FASTKD4 have opposite effects on expression of specific mitochondrial RNAs, depending upon their endonuclease-like RAP domain', *Nucleic Acids Research*, 45(10), pp. 6135–6146.
- Boer, P.H. and Gray, M.W. (1988) 'Scrambled Ribosomal RNA Gene Pieces in *Chlamydomonas reinhardtii* Mitochondrial DNA', *Cell*, 55, pp. 399–411.

- Bonnard, G. *et al.* (2016) 'Transfer RNA maturation in *Chlamydomonas* mitochondria, chloroplast and the nucleus by a single RNase P protein', *The Plant Journal*, 87(3), pp. 270–280.
- Böttinger, L. *et al.* (2015) 'Mitochondrial Heat Shock Protein (Hsp) 70 and Hsp10 Cooperate in the Formation of Hsp60 Complexes', *The Journal of Biological Chemistry*, 290(18), pp. 11611–11622.
- Boulouis, A. *et al.* (2015) 'Spontaneous Dominant Mutations in *Chlamydomonas* Highlight Ongoing Evolution by Gene Diversification', *The Plant Cell*, 27, pp. 984–1001.
- Brown, A. *et al.* (2014) 'Structure of the large ribosomal subunit from human mitochondria', *Science*, 346(6210), pp. 718–722.
- Brown, J.M. and Witman, G.B. (2014) 'Cilia and Diseases', *BioScience*, 64(12), pp. 1126–1137.
- Burns, J.A., Pittis, A.A. and Kim, E. (2018) 'Gene-based predictive models of trophic modes suggest Asgard archaea are not phagocytotic', *Nature Ecology and Evolution*, 2, pp. 697–704.
- Cahoon, A.B. *et al.* (2017) 'Deep Transcriptome Sequencing of Two Green Algae, *Chara vulgaris* and *Chlamydomonas reinhardtii*, Provides No Evidence of Organellar RNA Editing', *Genes*, 8(80).
- Cahoon, A.B. and Qureshi, A.A. (2018) 'Leaderless mRNAs are circularized in *Chlamydomonas reinhardtii* mitochondria', *Current Genetics*, 64(6), pp. 1321–1333.
- Cardol, P. *et al.* (2003) 'Photosynthesis and State Transitions in Mitochondrial Mutants of *Chlamydomonas reinhardtii* Affected in Respiration', *Plant Physiology*, 133(4), pp. 2010–2020.
- Cardol, P., Matagne, R.F. and Remacle, C. (2002) 'Impact of Mutations Affecting ND Mitochondria-encoded Subunits on the Activity and Assembly of Complex I in *Chlamydomonas*. Implication for the Structural Organization of the Enzyme', *Journal of Molecular Biology*, 319(5), pp. 1211–1221.
- Carrie, C., Murcha, M.W. and Whelan, J. (2010) 'An in silico analysis of the mitochondrial protein import apparatus of plants', *BMC Plant Biology*, 10(249).
- Cartalas, J., Coudray, L. and Gobert, A. (2022) 'How RNases Shape Mitochondrial Transcriptomes', *International Journal of Molecular Sciences*, 23(6141).

- Caspari, O.D. (2022) 'Transit Peptides Often Require Downstream Unstructured Sequence for Efficient Chloroplast Import in *Chlamydomonas reinhardtii*', *Frontiers in Plant Science*, 13.
- Cavaiuolo, M. *et al.* (2017) 'Small RNA profiling in *Chlamydomonas*: insights into chloroplast RNA metabolism', *Nucleic Acids Research*, 45(18), pp. 10783–10799.
- Cavalier-Smith, T. (2002) 'The phagotrophic origin of eukaryotes and phylogenetic classification of Protozoa', *International Journal of Systematic and Evolutionary Microbiology*, 52, pp. 297–354.
- Chang, H. *et al.* (2014) 'TAIL-seq: Genome-wide Determination of Poly(A) Tail Length and 3' End Modifications', *Molecular Cell*, 53(6), pp. 1044–1052.
- Chang, J.H. and Tong, L. (2012) 'Mitochondrial poly(A) polymerase and polyadenylation', *Biochimica et Biophysica Acta*, 1819, pp. 992–997.
- Chaudhuri, M. *et al.* (2020) 'Tim17 Updates: A Comprehensive Review of an Ancient Mitochondrial Protein Translocator', *Biomolecules*, 10(1643).
- Chen, Z. *et al.* (2019) 'PPR-SMR1 is required for the splicing of multiple mitochondrial introns, interacts with Zm-mCSF1, and is essential for seed development in maize', *Journal of Experimental Botany*, 70(19), pp. 5245–5258.
- Cheng, S. *et al.* (2016) 'Redefining the structural motifs that determine RNA binding and RNA editing by pentatricopeptide repeat proteins in land plants', *The Plant Journal*, 85, pp. 532–547.
- Chrzanowska-Lightowlers, Z.M.A., Pajak, A. and Lightowlers, R.N. (2011) 'Termination of Protein Synthesis in Mammalian Mitochondria', *Journal of Biological Chemistry*, 286(40), pp. 34479–34485.
- Chung, C.Z., Hyung Suk Jo, D. and Heinemann, I.U. (2016) 'Nucleotide specificity of the human terminal nucleotidyltransferase Gld2 (TUT2)', *RNA*, 22(8), pp. 1239–1249.
- Cognat, V. *et al.* (2013) 'PlantRNA, a database for tRNAs of photosynthetic eukaryotes', *Nucleic Acids Research*, 41(D273-9).
- Colcombet, J. *et al.* (2013) 'Systematic study of subcellular localization of Arabidopsis PPR proteins confirms a massive targeting to organelles', *RNA Biology*, 10(9), pp. 1557–1575.

- Cole, L.W. (2016) 'The Evolution of Per-cell Organelle Number', *Frontiers in Cell and Developmental Biology*, 4(85).
- Colleaux, L. *et al.* (1990) 'The apocytochrome b gene of *Chlamydomonas smithii* contains a mobile intron related to both *Saccharomyces* and *Neurospora* introns', *Molecular & General Genetics*, 223, pp. 288–296.
- Colombini, M. (2016) 'The VDAC channel: Molecular basis for selectivity', *Biochimica et Biophysica Acta - Molecular Cell Research*, 1863(10), pp. 2498–2502.
- Covello, P.S. and Gray, M.W. (1989) 'RNA editing in plant mitochondria', *Nature*, 341, pp. 662–666.
- Dass, S. *et al.* (2022) 'Transcriptional changes in *Plasmodium falciparum* upon conditional knock down of mitochondrial ribosomal proteins RSM22 and L23', *PLoS ONE*, 17(10).
- Denovan-Wright, E.M. and Lee, R.W. (1995) 'Evidence that the fragmented ribosomal RNAs of *Chlamydomonas* mitochondria are associated with ribosomes', *FEBS Letters*, 370, pp. 222–226.
- D'Erchia, A.M. *et al.* (2015) 'Tissue-specific mtDNA abundance from exome data and its correlation with mitochondrial transcription, mass and respiratory activity', *Mitochondrion*, 20, pp. 13–21.
- Desai, N. *et al.* (2017) 'The structure of the yeast mitochondrial ribosome', *Science*, 355, pp. 528–531.
- Dorrell, R.G. and Howe, C.J. (2012) 'What makes a chloroplast? Reconstructing the establishment of photosynthetic symbioses', *Journal of Cell Science*, 125(8), pp. 1865–1875.
- Dorthu, M.-P. *et al.* (1992) 'Biochemical, genetic and molecular characterization of new respiratory-deficient mutants in *Chlamydomonas reinhardtii*', *Plant Molecular Biology*, 18(4), pp. 759–772.
- Douce, R. *et al.* (2001) 'The glycine decarboxylase system : a fascinating complex', *Trends in Plant Science*, 6(4), pp. 167–176.
- Dubin, D.T. *et al.* (1982) 'Sequence Analysis and Precise Mapping of the 3' Ends of HeLa Cell Mitochondrial Ribosomal RNAs', *Journal of Molecular Biology*, 157, pp. 1–19.

- Duby, F. *et al.* (2001) 'Structure of the telomeric ends of mt DNA, transcriptional analysis and complex I assembly in the dum24 mitochondrial mutant of *Chlamydomonas reinhardtii*', *Molecular Genetics and Genomics*, 266, pp. 109–114.
- Duchêne, A.-M., Pujol, C. and Maréchal-Drouard, L. (2009) 'Import of tRNAs and aminoacyl-tRNA synthetases into mitochondria', *Current Genetics*, 55, pp. 1–18.
- Eberhard, S. *et al.* (2011) 'Dual functions of the nucleus-encoded factor TDA1 in trapping and translation activation of atpA transcripts in *Chlamydomonas reinhardtii* chloroplasts', *The Plant Journal*, 67(6), pp. 1055–1066.
- Edwards, R., Eaglesfield, R. and Tokatlidis, K. (2021) 'The mitochondrial intermembrane space: the most constricted mitochondrial sub-compartment with the largest variety of protein import pathways', *Open Biology*, 11.
- Ehara, T., Osafune, T. and Hase, E. (1995) 'Behavior of mitochondria in synchronized cells of *Chlamydomonas reinhardtii* (Chlorophyta)', *Journal of Cell Science*, 108, pp. 499–507.
- Erber, L. *et al.* (2020) 'Divergent Evolution of Eukaryotic CC-and A-Adding Enzymes', *International Journal of Molecular Sciences*, 21(462).
- Erber, L., Betat, H. and Mörl, M. (2020) 'CCA-Addition Gone Wild: Unusual Occurrence and Phylogeny of Four Different tRNA Nucleotidyltransferases in *Acanthamoeba castellanii*', *Molecular Biology and Evolution*, 38(3), pp. 1006–1017.
- Feagin, J.E. *et al.* (2012) 'The Fragmented Mitochondrial Ribosomal RNAs of *Plasmodium falciparum*', *PLoS ONE*, 7(6).
- Freel, K.C., Friedrich, A. and Schacherer, J. (2015) 'Mitochondrial genome evolution in yeasts: an all-encompassing view', *FEMS Yeast Research*, 15.
- Fuchs, P. *et al.* (2020) 'Single organelle function and organization as estimated from *Arabidopsis* mitochondrial proteomics', *The Plant Journal*, 101, pp. 420–441.
- Gabaldón, T. (2018) 'Relative Timing of Mitochondrial Endosymbiosis and the "Pre-Mitochondrial Symbioses" Hypothesis', *IUBMB Life*, 70(12), pp. 1188–1196.
- Gallaher, S.D. *et al.* (2018) 'High-throughput sequencing of the chloroplast and mitochondrion of *Chlamydomonas reinhardtii* to generate improved de novo assemblies, analyze expression

patterns and transcript speciation, and evaluate diversity among laboratory strains and wild isolates', *The Plant Journal*, 93(3), pp. 545–565.

Gardner, P.R., Nguyen, D.-D.H. and White, C.W. (1994) *Aconitase is a sensitive and critical target of oxygen poisoning in cultured mammalian cells and in rat lungs*, *Proc. Natl. Acad. Sci. USA*, pp. 12248–12252.

Goodenough, U., Lin, H. and Lee, J.-H. (2007) 'Sex determination in *Chlamydomonas*', *Seminars in Cell & Developmental Biology*, 18, pp. 350–361.

Grant, D. and Chiang, K.S. (1980) 'Physical mapping and characterization of *Chlamydomonas* mitochondrial DNA molecules: Their unique ends, sequence homogeneity, and conservation', *PLASMID*, 4(1), pp. 82–96.

Gray, M.W. (2012) 'Mitochondrial Evolution', *Cold Spring Harbor Perspectives in Biology*, 4.

Gray, M.W. (2015) 'Mosaic nature of the mitochondrial proteome: Implications for the origin and evolution of mitochondria', *Proceedings of the National Academy of Sciences of the United States of America*, 112(33), pp. 10133–10138.

Gray, M.W. and Boer, P.H. (1988) 'Organization and expression of algal (*Chlamydomonas reinhardtii*) mitochondrial DNA', in *Philosophical Transactions of the Royal Society of London*, pp. 135–147.

Greber, B.J. *et al.* (2014) 'The complete structure of the large subunit of the mammalian mitochondrial ribosome', *Nature*, 515, pp. 283–286.

Greber, B.J. *et al.* (2015) 'The complete structure of the 55S mammalian mitochondrial ribosome', *Science*, 348(6232), pp. 303–308.

Griparic, L. and Van Der Blik, A.M. (2001) 'The Many Shapes of Mitochondrial Membranes', *Traffic*, 2, pp. 235–244.

Gualberto, J.M. *et al.* (1989) 'RNA editing in wheat mitochondria results in the conservation of protein sequences', *Nature*, 341, pp. 660–662.

Gualberto, J.M. and Newton, K.J. (2017) 'Plant Mitochondrial Genomes: Dynamics and Mechanisms of Mutation', *Annual Review of Plant Biology*, 68, pp. 225–252.

- Haïli, N. *et al.* (2013) 'The pentatricopeptide repeat MTSF1 protein stabilizes the nad4 mRNA in Arabidopsis mitochondria', *Nucleic Acids Research*, 41(13), pp. 6650–6663.
- Hammani, K. *et al.* (2013) 'Helical repeats modular proteins are major players for organelle gene expression', *Biochimie*, 100, pp. 141–150.
- Hammani, K. and Giegé, P. (2014) 'RNA metabolism in plant mitochondria', *Trends in Plant Science*, 19(6), pp. 380–389.
- Hammarsund, M. *et al.* (2001) 'Identification and characterization of two novel human mitochondrial elongation factor genes, hEFG2 and hEFG1, phylogenetically conserved through evolution', *Human Genetics*, 109, pp. 542–550.
- Harris, E.H. (2001) 'Chlamydomonas as a model organism', *Annual Review of Plant Physiology and Molecular Biology*, 52, pp. 363–406.
- Hedtke, B., Börner, T. and Weihe, A. (2000) 'One RNA polymerase serving two genomes', *EMBO Reports*, 1(5), pp. 435–440.
- Herbert, C.J., Golik, P. and Bonnefoy, N. (2013) 'Yeast PPR proteins, watchdogs of mitochondrial gene expression', *RNA Biology*, 10(9), pp. 1477–1494.
- Herrmann, J.M., Woellhaf, M.W. and Bonnefoy, N. (2012) 'Control of protein synthesis in yeast mitochondria: The concept of translational activators', *Biochimica et Biophysica Acta*, 1833, pp. 286–294.
- Hiesel, R. *et al.* (1989) 'RNA editing in plant mitochondria', *Science*, 246(4937), pp. 1632–1634.
- Hirayama, T. *et al.* (2013) 'A poly(A)-specific ribonuclease directly regulates the poly(A) status of mitochondrial mRNA in Arabidopsis', *Nature Communications*, 4, pp. 1–9.
- Hoch, F.L. (1992) 'Cardiolipins and biomembrane function', *Biochimica et Biophysica Acta*, 1113(1), pp. 71–133.
- Hodge, T. and Colombini, M. (1997) 'Regulation of Metabolite Flux through Voltage-Gating of VDAC Channels', *Journal of Membrane Biology*, 157, pp. 271–279.
- Horton, T.L. and Landweber, L.F. (2000) 'Mitochondrial RNAs of myxomycetes terminate with non-encoded 3' poly(U) tails.', *Nucleic Acids Research*, 28(23), pp. 4750–4754.

- Huchon, D. *et al.* (2015) 'Mitochondrial group I and group II introns in the sponge orders Agelasida and Axinellida', *BMC Evolutionary Biology*, 15(278).
- Ibrahim, F. *et al.* (2010) 'Uridylation of mature miRNAs and siRNAs by the MUT68 nucleotidyltransferase promotes their degradation in *Chlamydomonas*', *Proceedings of the National Academy of Sciences of the United States of America*, 107(8), pp. 3906–3911.
- Imachi, H. *et al.* (2020) 'Isolation of an archaeon at the prokaryote-eukaryote interface', *Nature*, 577, pp. 519–525.
- Jackman, S.D. *et al.* (2020) 'Complete Mitochondrial Genome of a Gymnosperm, Sitka Spruce (*Picea sitchensis*), Indicates a Complex Physical Structure', *Genome Biology and Evolution*, 12(7), pp. 1174–1179.
- Jalal, A. *et al.* (2015) 'A Small Multifunctional Pentatricopeptide Repeat Protein in the Chloroplast of *Chlamydomonas reinhardtii*', *Molecular Plant*, 8(3), pp. 412–426.
- Jarrige, D. (2019) *Deciphering the 'OPR code' to further assess the physiological role of OPR proteins.*
- Jedynak-Slyvka, M., Jabczynska, A. and Szczesny, R.J. (2021) 'Human Mitochondrial RNA Processing and Modifications: Overview', *International Journal of Molecular Science*, 22(7999).
- Jia, L. *et al.* (2003) 'Yeast Oxa1 interacts with mitochondrial ribosomes: the importance of the C-terminal region of Oxa1', *EMBO Journal*, 22(24), pp. 6438–6447.
- Johnson, X. *et al.* (2010) 'MRL1, a Conserved Pentatricopeptide Repeat Protein, Is Required for Stabilization of *rbcl* mRNA in *Chlamydomonas* and *Arabidopsis*', *The Plant Cell*, 22(1), pp. 234–248.
- Jukes, T.H. and Osawa, S. (1990) 'The genetic code in mitochondria and chloroplasts', *Experientia*, 46(11–12), pp. 1117–1126.
- Karimi, M., Inzé, D. and Depicker, A. (2002) 'GATEWAY™ vectors for *Agrobacterium*-mediated plant transformation', *Trends in Plant Science*, 7(5), pp. 193–195.
- Karnkowska, A. *et al.* (2016) 'A Eukaryote without a Mitochondrial Organelle', *Current Biology*, 26(10), pp. 1274–1284.

- Kaye, Y. *et al.* (2019) 'The mitochondrial alternative oxidase from *Chlamydomonas reinhardtii* enables survival in high light', *Journal of Biological Chemistry*, 294(4), pp. 1380–1395.
- Kleinknecht, L. *et al.* (2014) 'RAP, the Sole Octotricopeptide Repeat Protein in Arabidopsis, Is Required for Chloroplast 16S rRNA Maturation', *The Plant Cell*, 26, pp. 777–787.
- Klingenberg, M. (1989) 'Molecular Aspects of the Adenine Nucleotide Carrier from Mitochondria', *Archives of Biochemistry and Biophysics*, 270(1), pp. 1–14.
- Komsic-Buchmann, K., Wöstehoff, L. and Becker, B. (2014) 'The Contractile Vacuole as a Key Regulator of Cellular Water Flow in *Chlamydomonas reinhardtii*', *Eukaryotic Cell*, 13(11), pp. 1421–1430.
- Krüger, A. *et al.* (2023) 'Human mitochondria require mtRF1 for translation termination at non-canonical stop codons', *Nature Communications*, 14(30).
- Kühn, K. *et al.* (2009) 'Phage-Type RNA Polymerase RPOTmp Performs Gene-Specific Transcription in Mitochondria of *Arabidopsis thaliana*', *The Plant Cell*, 21, pp. 2762–2779.
- Kumar, S., Stecher, G. and Tamura, K. (2016) 'MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets', *Molecular Biology and Evolution*, 33(7), pp. 1870–1874.
- Kummer, E. *et al.* (2018) 'Unique features of mammalian mitochondrial translation initiation revealed by cryo-EM', *Nature*, 560, pp. 263–267.
- Kummer, E. and Ban, N. (2021) 'Mechanisms and regulation of protein synthesis in mitochondria', *Nature Reviews Molecular Cell Biology*, 22, pp. 307–325.
- Kutner, J. *et al.* (2008) 'Mitochondrial release factor in yeast: interplay of functional domains', *Current Genetics*, 53, pp. 185–192. Available at: <https://doi.org/10.1007/s00294-008-0177-y>.
- Lang, B.F., Gray, M.W. and Burger, G. (1999) 'Mitochondrial Genome Evolution and the Origin of Eukaryotes', *Annual Review of Genetics*, 33, pp. 351–97.
- Lavrov, D. V. and Pett, W. (2016) 'Animal Mitochondrial DNA as We Do Not Know It: mt-Genome Organization and Evolution in Nonbilaterian Lineages', *Genome Biology and Evolution*, 8(9), pp. 2896–2913.

- Lee, B. II *et al.* (2001) 'Crystal Structure of the MJ0490 Gene Product of the Hyperthermophilic Archaeobacterium *Methanococcus jannaschii*, a Novel Member of the Lactate/Malate Family of Dehydrogenases', *Journal of Molecular Biology*, 307, pp. 1351–1362.
- Lefebvre-Legendre, L. *et al.* (2015) 'A nucleus-encoded Chloroplast Protein Regulated by Iron Availability Governs Expression of the Photosystem I Subunit PsaA in *Chlamydomonas reinhardtii*', *Plant Physiology*, 167(4), pp. 1527–1540.
- Levy, S. and Schuster, G. (2016) 'Polyadenylation and degradation of RNA in the mitochondria', *Biochemical Society Transactions*, 44, pp. 1475–1482.
- Li, C.Y. *et al.* (2022) 'Cytidine-containing tails robustly enhance and prolong protein production of synthetic mRNA in cell and in vivo', *Molecular Therapy - Nucleic Acids*, 30, pp. 300–310.
- Li, X. *et al.* (2016) 'An Indexed, Mapped Mutant Library Enables Reverse Genetics Studies of Biological Processes in *Chlamydomonas reinhardtii*', *Plant Cell*, 28(2), pp. 367–387.
- Li, X. *et al.* (2019) 'A genome-wide algal mutant library and functional screen identifies genes required for eukaryotic photosynthesis', *Nature Genetics*, 51, pp. 627–635.
- Lightowers, R.N. and Chrzanowska-Lightowers, Z.M.A. (2013) 'RNA Biology Human pentatricopeptide proteins: Only a few and what do they do?', *RNA Biology*, 10(9), pp. 1433–1438.
- Lipinski, K.A. *et al.* (2011) 'Revisiting the Yeast PPR Proteins-Application of an Iterative Hidden Markov Model Algorithm Reveals New Members of the Rapidly Evolving Family', *Molecular Biology and Evolution*, 28(10), pp. 2935–2948.
- Lister, R. *et al.* (2007) 'Functional Definition of Outer Membrane Proteins Involved in Preprotein Import into Mitochondria', *Plant Cell*, 19, pp. 3739–3759.
- Logan, D.C. (2006) 'The mitochondrial compartment', *Journal of Experimental Botany*, 57(6), pp. 1225–1243.
- Lohr, M. (2009) 'Carotenoids', in E. Harris, D. Stern, and G. Witman (eds) *The Chlamydomonas Sourcebook*. Amsterdam : Academic Press, pp. 799–817.
- Loiselay, C. *et al.* (2008) 'Molecular Identification and Function of cis- and trans-Acting Determinants for petA Transcript Stability in *Chlamydomonas reinhardtii* Chloroplasts', *Molecular and Cellular Biology*, 28(17), pp. 5529–5542.

- Lunde, B.M., Magler, I. and Meinhart, A. (2012) 'Crystal structures of the Cid1 poly (U) polymerase reveal the mechanism for UTP selectivity', *Nucleic Acids Research*, 40(19), pp. 9815–9824.
- Lurin, C. *et al.* (2004) 'Genome-Wide Analysis of Arabidopsis Pentatricopeptide Repeat Proteins Reveals Their Essential Role in Organelle Biogenesis', *The Plant cell*, 16, pp. 2089–2103.
- Lyons, T.W., Reinhard, C.T. and Planavsky, N.J. (2014) 'The rise of oxygen in Earth's early ocean and atmosphere', *Nature*, 506, pp. 307–315.
- Macedo-Osorio, K.S., Martínez-Antonio, A. and Badillo-Corona, J.A. (2021) 'Pas de Trois: An Overview of Penta-, Tetra-, and Octo-Tricopeptide Repeat Proteins From *Chlamydomonas reinhardtii* and Their Role in Chloroplast Gene Expression', *Frontiers in Plant Science*, 12(775366).
- Mackinder, L.C.M. (2018) 'The *Chlamydomonas* CO₂-concentrating mechanism and its potential for engineering photosynthesis in plants', *New Phytologist*, 217, pp. 54–61.
- Maier, U.G. *et al.* (2008) 'Complex chloroplast RNA metabolism: Just debugging the genetic programme?', *BMC Biology*, 6(36), pp. 1–9.
- Marienfeld, J., Unseld, M. and Brennicke, A. (1999) 'The mitochondrial genome of *Arabidopsis* is composed of both native and immigrant information', *Trends in Plant Science*, 4(12), pp. 495–502.
- Martin, G., Doublé, S. and Keller, W. (2008) 'Determinants of substrate specificity in RNA-dependent nucleotidyl transferases', *Biochimica et Biophysica Acta*, 1779(4), pp. 206–216.
- Martin, G. and Keller, W. (2004) 'Sequence motifs that distinguish ATP(CTP):tRNA nucleotidyl transferases from eubacterial poly(A) polymerases', *RNA*, 10, pp. 899–906.
- Martin, G. and Keller, W. (2007) 'RNA-specific ribonucleotidyl transferases', *RNA*, 13, pp. 1834–1839.
- Matagne, R.F. *et al.* (1989) 'Induction and Characterization of Mitochondrial DNA Mutants in *Chlamydomonas reinhardtii*', *The Journal of Cell Biology*, 108, pp. 1221–1226.
- Maul, J.E. *et al.* (2002) 'The *Chlamydomonas reinhardtii* Plastid Chromosome: Islands of Genes in a Sea of Repeats', *The Plant Cell*, 14, pp. 2659–2679.

- McBride, H.M., Neuspiel, M. and Wasiak, S. (2006) 'Mitochondria: More Than Just a Powerhouse', *Current Biology*, 16(14), pp. R551–R560.
- Merchant, S.S. *et al.* (2007) 'The Chlamydomonas genome Reveals the Evolution of Key Animal and Plant Functions', *Science*, 318(5848), pp. 245–250.
- Merendino, L. *et al.* (2006) 'A novel multifunctional factor involved in trans-splicing of chloroplast introns in Chlamydomonas', *Nucleic Acids Research*, 34(1), pp. 262–274.
- Michaelis, G., Vahrenholz, C. and Pratje, E. (1990) 'Mitochondrial DNA of Chlamydomonas reinhardtii: The gene for apocytochrome b and the complete functional map of the 15.8 kb DNA', *Molecular & General Genetics*, 223, pp. 211–216.
- Michaud, M. *et al.* (2014) 'Differential targeting of VDAC3 mRNA isoforms influences mitochondria morphology', *Proceedings of the National Academy of Sciences of the United States of America*, 111(24), pp. 8991–8996.
- Miranda, M., Bonekamp, N.A. and Kühl, I. (2022) 'Starting the engine of the powerhouse: Mitochondrial transcription and beyond', *Biological Chemistry*, 403(8–9), pp. 779–805.
- Miranda-Astudillo, H. *et al.* (2022) 'Beyond being an energy supplier, ATP synthase is a sculptor of mitochondrial cristae', *Biochimica et Biophysica Acta - Bioenergetics*, 1863(148569).
- Mishra, A.K. *et al.* (2009) 'High GC content: Critical parameter for predicting stress regulated miRNAs in Arabidopsis thaliana', *Bioinformatics*, 4(4), pp. 151–154.
- Mohanty, B.K. and Kushner, S.R. (2010) 'Bacterial/archaeal/organelle polyadenylation', *Wiley interdisciplinary reviews. RNA*, 2(2), pp. 256–276.
- Møller, I.M., Rasmusson, A.G. and Van Aken, O. (2021) 'Plant mitochondria – past, present and future', *The Plant Journal*, 108, pp. 912–959. Available at: <https://doi.org/10.1111/TPJ.15495>.
- Molnar, A. *et al.* (2009) 'Highly specific gene silencing by artificial microRNAs in the unicellular alga Chlamydomonas reinhardtii', *The Plant Journal*, 58(1), pp. 165–174.
- Nedelcu, A.M. (1997) 'Fragmented and Scrambled Mitochondrial Ribosomal RNA Coding Regions Among Green Algae: A Model for Their Origin and Evolution', *Molecular Biology and Evolution*, 14(5), pp. 506–517.

- Neuenfeldt, A. *et al.* (2008) 'Evolution of tRNA nucleotidyltransferases: A small deletion generated CC-adding enzymes', *Proceedings of the National Academy of Sciences of the United States of America*, 105(23), pp. 7953–7958.
- Ojala, D., Montoya, J. and Attardi, G. (1981) 'tRNA punctuation model of RNA processing in human mitochondria', *Nature*, 290, pp. 470–474.
- Okuda, K. and Shikanai, T. (2012) 'A pentatricopeptide repeat protein acts as a site-specificity factor at multiple RNA editing sites with unrelated cis-acting elements in plastids', *Nucleic Acids Research*, 40(11), pp. 5052–5064.
- Onishi, M. and Pringle, J.R. (2016) 'Robust Transgene Expression from Bicistronic mRNA in the Green Alga *Chlamydomonas reinhardtii*', *Genes/Genomes/Genetics*, 6, pp. 4115–4125.
- O'Rourke, B. (2007) 'Mitochondrial Ion Channels', *Annual Review of Physiology*, 69, pp. 19–49.
- Osafune, T. *et al.* (1976) 'Electron Microscope Studies of the Vegetative Cellular Life Cycle of *Chlamydomonas reinhardtii* Dangeard in Synchronous Culture IV. Mitochondria in Dividing Cells', *Journal of Electron Microscopy*, 25(4), pp. 261–269.
- Pearce, S.F. *et al.* (2017) 'Maturation of selected human mitochondrial tRNAs requires deadenylation', *eLife*, 6(e27596).
- Peters, K., Belt, K. and Braun, H.-P. (2013) '3D Gel Map of Arabidopsis Complex I', *Frontiers in Plant Science*, 4, pp. 1–9.
- Petrov, A.S. *et al.* (2018) 'Structural Patching Fosters Divergence of Mitochondrial Ribosomes', *Molecular Biology and Evolution*, 36(2), pp. 207–219.
- Philippe, H. and Forterre, P. (1999) 'The Rooting of the Universal Tree of Life Is Not Reliable', *Journal of Molecular Evolution*, 49, pp. 509–523.
- Picault, N. *et al.* (2004) 'The growing family of mitochondrial carriers in Arabidopsis', *Trends in Plant Science*, 9(3), pp. 138–146.
- Poole, A.M. and Gribaldo, S. (2014) 'Eukaryotic Origins: How and When Was the Mitochondrion Acquired?', *Cold Spring Harbor Perspectives in Biology*, 6.
- Prabhakar, A. *et al.* (2023) 'Dynamics of release factor recycling during translation termination in bacteria', *Nucleic Acids Research*, 51(1).

- Prikryl, J. *et al.* (2011) 'Mechanism of RNA stabilization and translational activation by a pentatricopeptide repeat protein', *Proceedings of the National Academy of Sciences of the United States of America*, 108(1), pp. 415–420.
- Proulex, G.C.R. *et al.* (2021) 'Mitochondrial mRNA Processing in the Chlorophyte Alga *Pediastrum duplex* and Streptophyte Alga *Chara vulgaris* Reveals an Evolutionary Branch in Mitochondrial mRNA Processing', *Plants*, 10(576).
- Pujol, C., Maréchal-Drouard, L. and Duchêne, A.M. (2007) 'How Can Organellar Protein N-terminal Sequences Be Dual Targeting Signals? In silico Analysis and Mutagenesis Approach', *Journal of Molecular Biology*, 369, pp. 356–367.
- Rahire, M. *et al.* (2012) 'Identification of an OPR protein involved in the translation initiation of the PsaB subunit of photosystem I', *The Plant Journal*, 72, pp. 652–661.
- Ramakers, C. *et al.* (2003) 'Assumption-free analysis of quantitative real-time polymerase chain reaction (PCR) data', *Neuroscience Letters*, 339, pp. 62–66.
- Ramrath, D.J.F. *et al.* (2018) 'Evolutionary shift toward protein-based architecture in trypanosomal mitochondrial ribosomes', *Science*, 362(6413).
- Rasala, B.A. *et al.* (2014) 'Enhanced Genetic Tools for Engineering Multigene Traits into Green Algae', *PLoS ONE*, 9(4), pp. 1–8. Available at: <https://doi.org/10.1371/journal.pone.0094028>.
- Remacle, C. *et al.* (2006) 'High-efficiency biolistic transformation of *Chlamydomonas* mitochondria can be used to insert mutations in complex I genes', *Proceedings of the National Academy of Sciences of the United States of America*, 103(12), pp. 4771–4776.
- Richard, O. *et al.* (1998) 'Transcription Initiation and RNA Processing in the Mitochondria of the Red Alga *Chondrus crispus*: Convergence in the Evolution of Transcription Mechanisms in Mitochondria', *Journal of Molecular Biology*, 283, pp. 549–557.
- Rodnina, M. V. (2018) 'Translation in Prokaryotes', *Cold Spring Harbor Perspectives in Biology*, 10(a032664).
- Salinas, T. *et al.* (2006) 'The voltage-dependent anion channel, a major component of the tRNA import machinery in plant mitochondria', *Proceedings of the National Academy of Sciences of the United States of America*, 103(48), pp. 18362–18367.

- Salinas, T. *et al.* (2012) 'Co-Evolution of Mitochondrial tRNA Import and Codon Usage Determines Translational Efficiency in the Green Alga *Chlamydomonas*', *PLoS Genetics*, 8(9), pp. 1–11.
- Salinas, T. *et al.* (2013) 'Respiratory-deficient mutants of the unicellular green alga *Chlamydomonas*: A review', *Biochimie*, 100, pp. 207–218.
- Salinas-Giegé, T. *et al.* (2017) 'Polycytidylation of mitochondrial mRNAs in *Chlamydomonas reinhardtii*', *Nucleic Acids Research*, 45(22), pp. 12 963-12973.
- Salinas-Giegé, T., Giegé, R. and Giegé, P. (2015) 'tRNA Biology in Mitochondria', *International Journal of Molecular Sciences*, 16, pp. 4518–4559.
- Salone, V. *et al.* (2007) 'A hypothesis on the identification of the editing enzyme in plant organelles', *FEBS Letters*, 581(22), pp. 4132–4138.
- Sarrion-Perdigones, A. *et al.* (2011) 'GoldenBraid: An Iterative Cloning System for Standardized Assembly of Reusable Genetic Modules', *PLoS ONE*, 6(7).
- Sarrion-Perdigones, A. *et al.* (2013) 'GoldenBraid 2.0: A Comprehensive DNA Assembly Framework for Plant Synthetic Biology', *Plant Physiology*, 162, pp. 1618–1631.
- Sassera, D. *et al.* (2011) 'Phylogenomic Evidence for the Presence of a Flagellum and cbb 3 Oxidase in the Free-Living Mitochondrial Ancestor', *Molecular Biology and Evolution*, 28(12), pp. 3285–3296.
- Sasso, S. *et al.* (2018) 'From molecular manipulation of domesticated *Chlamydomonas reinhardtii* to survival in nature', *eLife*, 7.
- Scaltsoyiannes, V. *et al.* (2022a) 'Types and Functions of Mitoribosome-Specific Ribosomal Proteins across Eukaryotes', *International Journal of Molecular Sciences*, 23(3474).
- Scaltsoyiannes, V. *et al.* (2022b) 'Types and Functions of Mitoribosome-Specific Ribosomal Proteins across Eukaryotes', *International Journal of Molecular Sciences*. MDPI.
- Schatz-Daas, D. (2021) *OEX1, une exonucléase/flap-endonucléase nécessaire à la stabilité du génome mitochondrial d'Arabidopsis thaliana*. Université de Strasbourg.
- Scheer, H. *et al.* (2016) 'Uridylation Earmarks mRNAs for Degradation... and More', *Trends in Genetics*, 32(10), pp. 607–619.

- Scheer, H. *et al.* (2020) 'High-Resolution Mapping of 3' Extremities of RNA Exosome Substrates by 3' RACE-Seq', *Methods in Molecular Biology*, 2062, pp. 147–167.
- Schmidt, O., Pfanner, N. and Meisinger, C. (2010) 'Mitochondrial protein import: from proteomics to functional mechanisms', *Nature Reviews Molecular Cell Biology*, 11, pp. 655–667.
- Schuster, G. and Stern, D. (2009) 'RNA Polyadenylation and Decay in Mitochondria and Chloroplasts', *Progress in Molecular Biology and Translational Science*, 85, pp. 393–422.
- Scott, I. and Logan, D.C. (2008) 'Mitochondria and cell death pathways in plants: Actions speak louder than words', *Plant Signaling & Behavior*, 3(7), pp. 475–477.
- Selinski, J. *et al.* (2018) 'Alternative Oxidase Is Positive for Plant Performance', *Trends in Plant Science*, 23(7), pp. 588–597.
- Shutt, T.E., Bestwick, M. and Shadel, G.S. (2011) 'The core human mitochondrial transcription initiation complex: It only takes two to tango', *Transcription*, 2(2), pp. 55–59.
- Sikorski, P.J. *et al.* (2015) 'Distinct 18S rRNA precursors are targets of the exosome complex, the exoribonuclease RRP6L2 and the terminal nucleotidyltransferase TRL in *Arabidopsis thaliana*', *The Plant Journal*, 83, pp. 991–1004.
- Sissler, M. *et al.* (2013) 'Mitochondrial Aminoacyl-tRNA Synthetases', in *Madame Curie Bioscience Database*. Landes Bioscience, pp. 2000–2013.
- Sloan, D.B. *et al.* (2012) 'Rapid Evolution of Enormous, Multichromosomal Genomes in Flowering Plant Mitochondria with Exceptionally High Mutation Rates', *PLoS Biology*, 10(1).
- Sloan, D.B. *et al.* (2018) 'Cytonuclear integration and co-evolution', *Nature Reviews Genetics*, 19, pp. 635–648.
- Slomovic, S. and Schuster, G. (2008) 'Stable PNPase RNAi silencing: Its effect on the processing and adenylation of human mitochondrial RNA', *RNA*, 14, pp. 310–323.
- Smith, D.R. and Craig, R.J. (2021) 'Does mitochondrial DNA replication in *Chlamydomonas* require a reverse transcriptase?', *New Phytologist*, 229, pp. 1192–1195.
- Smith, D.R., Hua, J. and Lee, R.W. (2010) 'Evolution of linear mitochondrial DNA in three known lineages of *Polytomella*', *Current Genetics*, 56, pp. 427–438.

- Spinelli, J.B. and Haigis, M.C. (2018) 'The multifaceted contributions of mitochondria to cellular metabolism', *Nature Cell Biology*, 20, pp. 745–754. Available at: <https://doi.org/10.1038/s41556-018-0124-1>.
- Spreitzer, R.J. and Mets, L.J. (1980) 'Non-mendelian mutation affecting ribulose-1, 5-bisphosphate carboxylase structure and activity', *Nature*, 285, pp. 114–115.
- Stampar, S.N. *et al.* (2019) 'Linear Mitochondrial Genome in Anthozoa (Cnidaria): A Case study in Ceriantharia', *Nature Scientific Reports*, 9(6094).
- Suematsu, T. *et al.* (2016) 'Antisense Transcripts Delimit Exonucleolytic Activity of the Mitochondrial 3' Processome to Generate Guide RNAs', *Molecular Cell*, 61(3), pp. 364–378.
- Sulo, P. *et al.* (2017) 'The evolutionary history of *Saccharomyces* species inferred from completed mitochondrial genomes and revision in the "yeast mitochondrial genetic code"', *DNA Research*, 24(6), pp. 571–583.
- Taanman, J.-W. (1999) 'The mitochondrial genome: structure, transcription, translation and replication', *Biochimica et Biophysica Acta*, 1410, pp. 103–123.
- Takenaka, M. *et al.* (2021) 'DYW domain structures imply an unusual regulation principle in plant organellar RNA editing catalysis', *Nature Catalysis*, 4, pp. 510–522.
- Tan, T.H.P. *et al.* (2002) 'Eukaryotic-type elongator tRNAMet of *Trypanosoma brucei* becomes formylated after import into mitochondria', *Proceedings of the National Academy of Sciences of the United States of America*, 99(3), pp. 1152–1157.
- Tardif, M. *et al.* (2012) 'Predalgo: A New Subcellular Localization Prediction Tool Dedicated to Green Algae', *Molecular Biology and Evolution*, 29(12), pp. 3625–3639.
- Terashima, M., Specht, M. and Hippler, M. (2011) 'The chloroplast proteome: a survey from the *Chlamydomonas reinhardtii* perspective with a focus on distinctive features', *Current Genetics*, 57, pp. 151–168.
- Tobiasson, V., Berzina, I. and Amunts, A. (2022) 'Structure of a mitochondrial ribosome with fragmented rRNA in complex with membrane-targeting elements', *Nature Communications*, 13(6132).

- Tomita, K. and Weiner, A.M. (2002) 'Closely Related CC-and A-adding Enzymes Collaborate to Construct and Repair the 3'-Terminal CCA of tRNA in *Synechocystis* sp. and *Deinococcus radiodurans*', *The Journal of Biological Chemistry*, 227(50), pp. 48192–48198.
- Tourasse, N.J., Choquet, Y. and Vallon, O. (2013) 'PPR proteins of green algae', *RNA Biology*, 10(9), pp. 1526–1542.
- Tretbar, S. *et al.* (2011) 'An inhibitory C-terminal region dictates the specificity of A-adding enzymes', *Proceedings of the National Academy of Sciences of the United States of America*, 108(52), pp. 21040–21045.
- Tsuboi, M. *et al.* (2009) 'EF-G2mt Is an Exclusive Recycling Factor in Mammalian Mitochondrial Protein Synthesis', *Molecular cell*, 35, pp. 502–510.
- Tudek, A. *et al.* (2021) 'Global view on the metabolism of RNA poly(A) tails in yeast *Saccharomyces cerevisiae*', *Nature Communications*, 12(1), pp. 1–14.
- Vahrenholz, C. *et al.* (1993) 'Mitochondrial DNA of *Chlamydomonas reinhardtii*: the structure of the ends of the linear 15.8-kb genome suggests mechanisms for DNA replication', *Current Genetics*, 24, pp. 241–247.
- Verner, K. (1993) 'Co-translational protein import into mitochondria: an alternative view', *Trends in Biochemical Sciences*, 18(10), pp. 366–371.
- Vincent, T. *et al.* (2017) 'A genome-scale analysis of mRNAs targeting to plant mitochondria: upstream AUGs in 5' untranslated regions reduce mitochondrial association', *The Plant Journal*, 92, pp. 1132–1142.
- Vinogradova, E. *et al.* (2009) 'Steady-state levels of imported tRNAs in *Chlamydomonas* mitochondria are correlated with both cytosolic and mitochondrial codon usages', *Nucleic Acids Research*, 37(5), pp. 1521–1528.
- Viola, S. *et al.* (2019) 'MDA1, a nucleus-encoded factor involved in the stabilization and processing of the *atpA* transcript in the chloroplast of *Chlamydomonas*', *The Plant Journal*, 98(6), pp. 1033–1047.
- van de Vossenbergh, B.T.L.H. *et al.* (2018) 'The linear mitochondrial genome of the quarantine chytrid *Synchytrium endobioticum*; insights into the evolution and recent history of an obligate biotrophic plant pathogen', *BMC Evolutionary Biology*, 18(136).

- Walther, D.M. and Rapaport, D. (2009) 'Biogenesis of mitochondrial outer membrane proteins', *Biochimica et Biophysica Acta - Molecular Cell Research*, 1793(1), pp. 42–51.
- Waltz, F. (2018) *Characterization of the mitochondrial translation apparatus of Arabidopsis thaliana*. Université de Strasbourg.
- Waltz, F. *et al.* (2019) 'Small is big in Arabidopsis mitochondrial ribosome', *Nature Plants*, 5, pp. 106–117.
- Waltz, F. *et al.* (2020) 'Specificities of the plant mitochondrial translation apparatus', *Mitochondrion*, 53, pp. 30–37.
- Waltz, F. *et al.* (2021) 'How to build a ribosome from RNA fragments in Chlamydomonas mitochondria', *Nature Communications*, 12(1).
- Waltz, F. and Giegé, P. (2020) 'Striking Diversity of Mitochondria-Specific Translation Processes across Eukaryotes', *Trends in Biochemical Sciences*, 45(2), pp. 149–162.
- Wang, C. *et al.* (2018) 'Three new pentatricopeptide repeat proteins facilitate the splicing of mitochondrial transcripts and complex I biogenesis in Arabidopsis', *Journal of Experimental Botany*, 69(21), pp. 5131–5140.
- Wang, F. *et al.* (2015) 'Two Chlamydomonas OPR proteins stabilize chloroplast mRNAs encoding small subunits of photosystem II and cytochrome b6f', *The Plant Journal*, 82, pp. 861–873.
- Wang, Z. and Wu, M. (2015) 'An integrated phylogenomic approach toward pinpointing the origin of mitochondria', *Scientific Reports*, 5.
- Williams, C.C., Jan, C.H. and Weissman, J.S. (2014) 'Targeting and plasticity of mitochondrial proteins revealed by proximity-specific ribosome profiling', *Science*, 346(6210), pp. 748–751.
- Wobbe, L. and Nixon, P.J. (2013) 'The mTERF protein MOC1 terminates mitochondrial DNA transcription in the unicellular green alga Chlamydomonas reinhardtii', *Nucleic Acids Research*, 41(13), pp. 6553–6567.
- Wolters, J.F., Chiu, K. and Fiumera, H.L. (2015) 'Population structure of mitochondrial genomes in Saccharomyces cerevisiae', *BMC Genomics*, 16(451).
- Worriax, V.L. *et al.* (1997) 'Mechanistic studies of the translational elongation cycle in mammalian mitochondria', *Biochimica et Biophysica Acta*, 1352, pp. 91–101.

- Workman, R.E. *et al.* (2019) 'Nanopore native RNA sequencing of a human poly(A) transcriptome', *Nature Methods*, 16(12), pp. 1297–1305.
- Xu, B., Liu, L. and Song, G. (2022) 'Functions and Regulation of Translation Elongation Factors', *Frontiers in Molecular Biosciences*, 8.
- Yang, H. *et al.* (2019) 'Overexpression of BnaAOX1b Confers Tolerance to Osmotic and Salt Stress in Rapeseed', *Genes/Genomes/Genetics*, 9, pp. 3501–3511.
- Yates, L.A. *et al.* (2012) 'Structural basis for the activity of a cytoplasmic RNA terminal uridylyl transferase', *Nature Structural & Molecular Biology*, 19(8), pp. 782–787.
- Yates, L.A. *et al.* (2015) 'Structural plasticity of Cid1 provides a basis for its distributive RNA terminal uridylyl transferase activity', *Nucleic Acids Research*, 43(5), pp. 2968–2979.
- Yehudai-Resheff, S. *et al.* (2007) 'Integration of Chloroplast Nucleic Acid Metabolism into the Phosphate Deprivation Response in *Chlamydomonas reinhardtii*', *The Plant Cell*, 19, pp. 1023–1038.
- Yehudai-Resheff, S. and Schuster, G. (2000) 'Characterization of the E.coli poly(A) polymerase: nucleotide specificity, RNA-binding affinities and RNA structure dependence', *Nucleic Acids Research*, 28(5), pp. 1139–1144.
- Yue, D., Maizels, N. and Weiner, A.M. (1996) 'CCA-adding enzymes and poly(A) polymerases are all members of the same nucleotidyltransferase superfamily: Characterization of the CCA-adding enzyme from the archaeal hyperthermophiles *Sulfolobus shibatae*', *RNA*, 2(9), pp. 895–908.
- Yueh, Y.G. and Crain, R.C. (1993) 'Deflagellation of *Chlamydomonas reinhardtii* Follows a Rapid Transitory Accumulation of Inositol 1,4,5,-trisphosphate and Requires Ca²⁺ Entry', *The Journal of Cell Biology*, 123(4), pp. 869–875.
- Zaremba-Niedzwiedzka, K. *et al.* (2017) 'Asgard archaea illuminate the origin of eukaryotic cellular complexity', *Nature*, 541, pp. 353–358.
- Zhang, T. *et al.* (2021) 'Characterization and Comparative Analyses of Mitochondrial Genomes in Single-Celled Eukaryotes to Shed Light on the Diversity and Evolution of Linear Molecular Architecture', *International Journal of Molecular Sciences*, 22(5).

- Zhao, F. and Zou, M.-H. (2021) 'Role of the Mitochondrial Protein Import Machinery and Protein Processing in Heart Disease', *Frontiers in Cardiovascular Medicine*, 8(749756).
- Zhou, W. *et al.* (2017) 'PPR-SMR protein SOT1 has RNA endonuclease activity', *Proceedings of the National Academy of Sciences of the United States of America*, 114(8), pp. E1554–E1563.
- Zimmer, S.L. *et al.* (2009) 'Polyadenylation in Arabidopsis and Chlamydomonas organelles: The input of nucleotidyltransferases, poly(A) polymerases and polynucleotide phosphorylase', *The Plant Journal*, 59(1), pp. 88–99.
- Zimmer, S.L., Fei, Z. and Stern, D.B. (2008) 'Genome-Based Analysis of Chlamydomonas reinhardtii Exoribonucleases and Poly(A) Polymerases Predicts Unexpected Organellar and Exosomal Features', *Genetics*, 179, pp. 125–136.
- Zizi, M. *et al.* (1995) 'Oriented channel insertion reveals the motion of a transmembrane beta strand during voltage gating of VDAC', *The Journal of Membrane Biology*, 144(2), pp. 121–129.
- Zoschke, R. *et al.* (2016) 'The PPR-SMR protein PPR53 enhances the stability and translation of specific chloroplast RNAs in maize', *The Plant Journal*, 85, pp. 594–606.
- Zumkeller, S., Gerke, P. and Knoop, V. (2020) 'A functional twintron, "zombie" twintrons and a hypermobile group II intron invading itself in plant mitochondria', *Nucleic Acids Research*, 48(5), pp. 2661–2675.

ANNEXES

I. Mutants CLiP

Tableau des mutants de la banque CLiP, achetés sur le site de la banque CLiP ou donnés par le Docteur Olivier Vallon. Prédiction des sites d'insertion dans : **jaune** : exon ; **orange** : intron ; **bleu** : 3'UTR.

| Gène | Gene ID | Référence achat | Prédiction d'insertion |
|---------|--------------------|-------------------|------------------------|
| mL114 | Cre06.g303850.t1.1 | LMJ.RY0402.068576 | 3'UTR |
| mL115 | Cre09.g392579.t1.1 | LMJ.RY0402.139756 | 3'UTR |
| mL116 | Cre03.g179901.t1.1 | LMJ.RY0402.119390 | 3'UTR |
| mL117 | Cre14.g615450.t1.2 | LMJ.RY0402.123594 | Exon 9 |
| mL118 | Cre04.g217904.t1.1 | LMJ.RY0402.114494 | 3'UTR |
| mL119 | Cre12.g521050.t1.2 | LMJ.RY0402.143618 | Intron 6 |
| mS106 | Cre10.g447100.t1.2 | LMJ.RY0402.123403 | 3'UTR |
| OPR119 | Cre06.g278138.t1.1 | LMJ.RY0402.134295 | Intron 4 |
| OPR14 | Cre02.g146900.t1.1 | LMJ.RY0402.197733 | Intron 13 |
| OPR44 | Cre10.g426200.t1.1 | LMJ.RY0402.064348 | 3'UTR |
| OPR48-1 | Cre10.g441600.t1.1 | LMJ.RY0402.098352 | Exon 3 |
| OPR48-2 | Cre10.g441600.t1.1 | LMJ.RY0402.194194 | Exon 1 |
| PAP4-1 | Cre06.g280420.t1.2 | LMJ.RY0402.079134 | Intron 2 |
| PAP4-2 | Cre06.g280420.t1.2 | LMJ.RY0402.125943 | Intron 3 |
| PPR11 | Cre09.g412150.t1.1 | LMJ.RY0402.129802 | Intron 15 |

II. Oligonucléotides

1. PCR

a) Mutants CLiP

Tableaux des amorces utilisées pour le génotypage des mutants CLiP. Elles se situent de part et d'autre du site théorique d'insertion fourni par la banque. Le couple « Contrôle » sert de contrôle positif pour les PCR, amplifiant la région entre les gènes HRP4 et OAT1.

Locus d'insertion de la cassette

| Gène | N° | Sens | Séquence 5' -> 3' | Gène | N° | Sens | Séquence 5' -> 3' |
|----------|-----|------|-----------------------|---------|------|------|----------------------|
| Contrôle | 653 | F | ATGCTTCTCTGCATCCGTCT | OPR119 | 758 | F | GGGTTGTGGATGCAGAAGAT |
| | 654 | R | ATGTTTTACGTCCAGTCCGC | | 759 | R | GGCCATTCAATACCTGACCA |
| mL114 | 841 | F | AGGAGGAGGACGAAGAGGAG | OPR14 | 799 | F | GGACGGCGACGGGCTGGC |
| | 842 | R | AAGGATTTCATCGTGTGGGAG | | 798 | R | CACACGCCGACGACCTC |
| mL115 | 755 | F | CTGCTGCAGACCCCTCTTAG | OPR44 | 938 | F | AGGAGGATGAGGAGGAGAGC |
| | 754 | R | AGGGGTATGCTGTATGGCTG | | 939 | R | CGAAGTTCTCTTGGGCTGAC |
| mL116 | 843 | F | GTCACAATGGACCTGACGG | OPR48-1 | 651 | F | GAGCTGTTTGACGCAGTGAC |
| | 844 | R | TGCCGTACATAAGTGGGTGA | | 652 | R | ACTGCGTGTCTTGGCAAC |
| mL117 | 845 | F | CCGGGTGTTGTCTATTTCAC | OPR48-2 | 752 | F | AACACGACAGGGAAGGTGAC |
| | 846 | R | CACACAACCTCCATACGGTGC | | 753 | R | AGTAGGCCACAGGATGTTG |
| mL118 | 888 | F | GCTCGCTGGCTGCACGTC | PAP4-1 | 751 | F | CACCAACGCCTCCATCACC |
| | 889 | R | GCCGTCTTACCCATCGTCCG | | 750 | R | GCGTTTGGGTTTTGTGTTGG |
| mL119 | 849 | F | CTGAATGATGCATGTCCCTG | PAP4-2 | 649 | F | AGCATTCTGATACTGCGGCT |
| | 850 | R | GATTTTGGCCTCATTCTCCA | | 650 | R | CACCTCACTCCAAAGGTGT |
| mS106 | 871 | F | TTGGTTTGGTTTAGCTTGGG | PPR11 | 940 | F | TAGTCCCAACCAGACCAAGG |
| | 872 | R | CACGTAGACGGCCAGTATT | | 1017 | R | CCTGTGGGAGAATGAGGAGG |

Tableau des amorces utilisées pour la cartographie des mutants CLiP. Une amorce se situe avant l'insertion de la cassette, dans le gène, tandis que la deuxième est dans la cassette, en 3' ou en 5'.

Jonction 3' de la cassette avec le locus d'insertion

| Gène | N° | Sens | Séquence 5' -> 3' |
|---------|------|------|-----------------------|
| OPR119 | 656 | F | GACGTTACAGCACACCCCTTG |
| | 759 | R | GGCCATTCAATACCTGACCA |
| OPR44 | 938 | F | AGGAGGATGAGGAGGAGAGC |
| | 656 | R | GACGTTACAGCACACCCCTTG |
| OPR48-1 | 656 | F | GACGTTACAGCACACCCCTTG |
| | 652 | R | ACTGCGTGTCTTGGCAAC |
| OPR48-2 | 656 | F | GACGTTACAGCACACCCCTTG |
| | 753 | R | AGTAGGCCACAGGATGTTG |
| PAP4-1 | 751 | F | CACCAACGCCTCCATCACC |
| | 656 | R | GACGTTACAGCACACCCCTTG |
| PAP4-2 | 656 | F | GACGTTACAGCACACCCCTTG |
| | 650 | R | CACCTCACTCCAAAGGTGT |
| PPR11 | 656 | F | GACGTTACAGCACACCCCTTG |
| | 1017 | R | CCTGTGGGAGAATGAGGAGG |

Jonction 5' de la cassette avec le locus d'insertion

| Gène | N° | Sens | Séquence 5' -> 3' |
|---------|-----|------|--------------------------|
| OPR119 | 758 | F | GGGTTGTGGATGCAGAAGAT |
| | 655 | R | GCACCAATCATGTCAAGCCT |
| OPR44 | 655 | F | GCACCAATCATGTCAAGCCT |
| | 939 | R | CGAAGTTCTCTTGGGCTGAC |
| OPR48-1 | 651 | F | GAGCTGTTTGACGCAGTGAC |
| | 709 | R | TAGTCGCGTATGTATAAGTGCTCG |
| OPR48-2 | 752 | F | AACACGACAGGGAAGGTGAC |
| | 655 | R | GCACCAATCATGTCAAGCCT |
| PAP4-1 | 655 | F | GCACCAATCATGTCAAGCCT |
| | 750 | R | GCGTTTGGGTTTTGTGTTGG |
| PAP4-2 | 649 | F | AGCATTCTGATACTGCGGCT |
| | 709 | R | TAGTCGCGTATGTATAAGTGCTCG |
| PPR11 | 940 | F | TAGTCCCAACCAGACCAAGG |
| | 655 | R | GCACCAATCATGTCAAGCCT |

b) Clonage Gateway

Tableau des amorces utilisées pour les clonages avec la technologie Gateway, pour la construction des vecteurs exprimant la partie 5' des gènes d'intérêt en fusion avec la EGFP.

Bleu : site attB1 ; **Orange** : site attB2.

| Gène | N° | Sens | Séquence 5' -> 3' | Description |
|--------|-----|------|--|---------------------------------|
| OPR10 | 521 | F | ATGTCAATCGCGTTGCTCTGC | Amplification de l'extrémité 5' |
| | 578 | R | CAGGTGCTCCAGCCGCTTTAG | |
| | 602 | F | GGGGACAAGTTTGTACAAAAAAGCAGGCTAAAAATGTCAATCGCGTTGCTCTGC | Ajout des sites attB1 et attB2 |
| OPR15 | 603 | R | GGGGACCACTTTGTACAAGAAAGCTGGGTGGTGTCTCCAGCCGCTTTAG | Amplification de l'extrémité 5' |
| | 579 | F | ATGACGTTGGGACGCAGCCGAG | |
| | 581 | R | GCAGCAGCATCTCATCGAACAGC | Ajout des sites attB1 et attB2 |
| OPR48 | 604 | F | GGGGACAAGTTTGTACAAAAAAGCAGGCTAAAAATGACGTTGGGACGCAGCCGAG | Amplification de l'extrémité 5' |
| | 605 | R | GGGGACCACTTTGTACAAGAAAGCTGGGTCTTGTGTCTCCTCCCTCGC | |
| | 583 | F | ATGGCTGCGTGCCTCCGCCAATTC | Ajout des sites attB1 et attB2 |
| OPR107 | 585 | R | CCTGCCGCTAGCTTCTGCCGACC | Amplification de l'extrémité 5' |
| | 606 | F | GGGGACAAGTTTGTACAAAAAAGCAGGCTAAAAATGGCTGCGTGCCTCCGCC | |
| | 607 | R | GGGGACCACTTTGTACAAGAAAGCTGGGTGTGTGCACGGTTCTCGCGC | Ajout des sites attB1 et attB2 |
| OPR109 | 587 | F | ATGGCCTCTAAACAACCTGCTGAGGCG | Amplification de l'extrémité 5' |
| | 589 | R | GGTGGTGCCGAGCGGGATC | |
| | 608 | F | GGGGACAAGTTTGTACAAAAAAGCAGGCTAAAAATGGCCTCTAAACAACCTGCTG | Ajout des sites attB1 et attB2 |
| OPR109 | 609 | R | GGGGACCACTTTGTACAAGAAAGCTGGGTGGTGGCCGAGCGGGATCGC | Amplification de l'extrémité 5' |
| | 591 | F | ATGAGTACATCGGCATTCTGCTTACGCC | |
| | 593 | R | TGGCGGACTCTGGATGGCG | Ajout des sites attB1 et attB2 |
| OPR119 | 610 | F | GGGGACAAGTTTGTACAAAAAAGCAGGCTAAAAATGAGTACATCGGCATTCTGCTG | Amplification de l'extrémité 5' |
| | 611 | R | GGGGACCACTTTGTACAAGAAAGCTGGGTGCAGCTTGGGCTGCAGCGG | |
| | 595 | F | ATGTCTGTGCCAAAGCGCTTGG | Ajout des sites attB1 et attB2 |
| PAP2 | 597 | R | TCCACCCCGGCCCTCTC | Amplification de l'extrémité 5' |
| | 612 | F | GGGGACAAGTTTGTACAAAAAAGCAGGCTAAAAATGTCTGTGCCAAAGCGCTTGG | |
| | 613 | R | GGGGACCACTTTGTACAAGAAAGCTGGGTGGCGGCCACACGTTTCTG | Ajout des sites attB1 et attB2 |
| PAP4 | 394 | F | ATGCTGCCAGCTGCAGCAGCCTTG | Amplification de l'extrémité 5' |
| | 396 | R | CGGCAGGCGCACAGAAGC | |
| | 483 | F | GGGGACAAGTTTGTACAAAAAAGCAGGCTAAAAATGCTGCCAGCTGCAGCAGC | Ajout des sites attB1 et attB2 |
| PAP6 | 484 | R | GGGGACCACTTTGTACAAGAAAGCTGGGTGCAGGCGCACAGAAGCTGTC | Amplification de l'extrémité 5' |
| | 402 | F | ATGCCAGCTTCGGCCTC | |
| | 404 | R | TAGCGCATGTGGTTGAAAC | Ajout des sites attB1 et attB2 |
| PAP10 | 487 | F | GGGGACAAGTTTGTACAAAAAAGCAGGCTAAAAATGCCAGCTTCGGCCTC | Amplification de l'extrémité 5' |
| | 488 | R | GGGGACCACTTTGTACAAGAAAGCTGGGTGCGCATGTGGGTGAAACTTG | |
| | 557 | F | ATGCTGCCCCGCGCTGC | Ajout des sites attB1 et attB2 |
| PAP6 | 362 | R | TTCTACGGGCAGAGGGCAC | Amplification de l'extrémité 5' |
| | 489 | F | GGGGACAAGTTTGTACAAAAAAGCAGGCTAAAAATGCTGCCCGCGCTGCTAGG | |
| | 490 | R | GGGGACCACTTTGTACAAGAAAGCTGGGTGGTCAGACTTGGCGTCATAG | Ajout des sites attB1 et attB2 |
| PAP10 | 378 | F | ATGGCGACGCCCCGCTC | Amplification de l'extrémité 5' |
| | 380 | R | GGACGACTCCACGAACCTTG | |
| | 491 | F | GGGGACAAGTTTGTACAAAAAAGCAGGCTAAAAATGGCGACGCCCCGCTCGG | Ajout des sites attB1 et attB2 |
| | 492 | R | GGGGACCACTTTGTACAAGAAAGCTGGGTGCTGGCTGTAGCAGGAGCTC | |

c) Clonage Golden Braid

Tableau des amorces utilisées lors des PCR pour les clonages par la technologie Golden Braid, pour les constructions permettant l'expression des protéines en vue de les purifier. **Bleu** : séquences pour l'assemblage des modules Golden Braid ; **Jaune** : site de restriction Bsmbl ; **Rouge** : nucléotides mutés dans le vecteur pUPD2-HSUM-PAP6, afin d'avoir la construction pUPD2-HSUM-PAP6MUT ; **Turquoise** : partie commune entre les amorces F et R ; **Vert** : nouveau site créé pour l'insertion de l'extension de PAP4 dans le vecteur pUPD2-HSUM-PAP4, afin d'avoir la construction pUPD2-HSUM-EXTPAP4.

| Cible | N° | Sens | Séquence 5' -> 3' | Description |
|-----------------|---------|------|---|---|
| pUPD2-HSUM | 1102 | F | ATGCC CGTCTCG CTCG AGGT GCCACCTATACCACACGCA | Oligos utilisés pour cloner la séquence PAP6 dans le vecteur pUPD2-HSUM |
| | 1103 | R | GCAT CGTCTCG CTCA CGAA GCCTCGGCGGATGCGGC | |
| pUPD2-HSUM-PAP6 | 1107 | F | CTGCCCTAGCCCTGGT CTATCGG | Oligos utilisés pour muter la séquence de PAP6 |
| | 1108 | R | CCAGGGCTAGGGCAGACT TGGCG | |
| pUPD2-HSUM-PAP4 | 1117 | F | GCCGTCTCGAGGT ACTGCCGTGACAGCTCAC | Oligos de part et d'autre de l'extension de PAP4 à ajouter entre |
| | 1116 | R | GCCGTCTCGCTCA ACCT GATAAAATGCCGCGGTTGTG | |
| | P3P-389 | F | | |
| | 1118 | R | GCAT CGTCTCG ACCT CCC GTCTGCTGCTG | Oligos de part et d'autre du tag HSUMO pour intégrer l'extension PAP4 |

2. Séquençage

Tableau des amorces spécifiques utilisés pour le séquençage. Les amorces décrites pour les PCR ont également pu être utilisées pour du séquençage. F : *forward*, amorce sens ; R : *reverse*, amorce antisens.

| Cible | N° | Sens | Séquence 5' -> 3' | Description |
|------------------|-----------|------|----------------------------|---|
| pGEM | LD738 | F | CAGGAAACAGCTATGAC | Séquençage des pGEM-insert ; Amorces de part et d'autre de l'insert |
| | LD739 | R | TAAAACGACGGCCAG | |
| pDONR207 | Heike1052 | F | TCGCGTTAACGCTAGCATGGA | Séquençage des pDONR207-MTS ; Amorces de part et d'autre du site d'insertion de |
| | Heike1053 | R | GTAACATCAGAGATTTTGAGA | |
| pB7FWG2 | AM595 | F | CCTTCCTCTATATAAGGAAGTTCA | Séquençage des pB7FWG2-MTS ; Amorces de part et d'autre du site d'insertion de |
| | 514 | R | ACTTGTGGCCGTTTACGTCG | |
| | Heike129 | F | ACAAGTTTGTACAAAAAAGCAGGCT | Séquençage des pB7FWG2-MTS ; Amorces à la jonction vecteur/sites att de l'insert |
| | Heike130 | R | ACCACTTTGTACAAAGAAAGCTGGGT | |
| pUPD2 | P3P-352 | F | GCTTTCGCTAAGGATGATTTCTGG | Séquençage des pUPD2-insert ; Amorces de part et d'autre du site de restriction Bsmbl |
| | P3P-353 | R | CAGGGTGGTGACACCTTGCC | |
| pBACT α 1 | P3P-348 | F | CCTTTGAGTGAGCTGATACC | Séquençage des pBACT α 1-insert ; Amorces de part et d'autre du site de restriction Bsa1 |
| | P3P-349 | R | AGGCGTATCACGAGGCAG | |

3. Constructions pour les mutants *amiRNA*

Tableau des séquences et oligos utilisés pour l'obtention des mutants *amiRNA*. **Bleu** : mismatch entre la séquence ciblée et la séquence de l'*amiRNA* ; minuscules : séquence commune à tous les *amiRNA*, permet la structure en tête d'épingle.

| Gène | Séquence ciblée 5' -> 3' | Où ? | N° | Sens | Séquence 5' -> 3' |
|--------|--------------------------|-----------------|------|------|--|
| mLXX1 | GCGCATCAGTATAACTTGATT | Début Exon 3 | 892 | F | ctagtTCGCATCAGTATAACTAGATctctcgtgatcggcaccatgggggtggtggtgacagcgcctaTATCAAGTTTATACTGATCGGAg |
| mLXX2 | CCGGCTCACGCATTGATTAC | Exon 3 | 893 | R | ctagcTCGCATCAGTATAACTTTGATAtagcgtgatcaccaccaccaccatggtgcgacagcagaTATCTAGTTTATACTGATCGGA |
| mLXX3 | TTGTCTCCATATGGGCGTTTGG | Début Exon 5 | 896 | F | ctagtACGGCTCACGCATTGTTTAAtctcgtgatcggcaccatgggggtggtggtgacagcgcctaTAAACAATCGGTGAGCGGTTg |
| mLXX4 | ATGATCAACACCTCTGAAAGC | Exon 2 | 897 | R | ctagcACGGCTCACGCATTGTTTAAtagcgtgatcaccaccaccaccatggtgcgacagcagaTTAAACAATCGGTGAGCGGTTa |
| mLXX5 | CAGGCTGCACAAACTCAAAACA | Début Exon 5 | 890 | F | ctagtTTGCTCCATATGGCGATTGATctcgtgatcggcaccatgggggtggtggtgacagcgcctaTCAAACGCCCATATGGACCAAG |
| mLXX6 | CTGAAAGGAGGCTGTGTCATTA | Exon 5 | 891 | R | ctagcTTGCTCCATATGGCGTTTGAtagcgtgatcaccaccaccaccatggtgcgacagcagaTCAAATCGCCATATGGACCAAA |
| mLXX7 | CGGAAATGCGAAAGGTTCAATG | Exon 5 | 898 | F | ctagtAAGATCAACACCTCTGTAAGAtctcgtgatcggcaccatgggggtggtggtgacagcgcctaTCTTTACAGAGTGTGATCTTg |
| mSXX1 | GCGATCGCGTATGCTGTTTG | Début Exon 2 | 899 | R | ctagcAAGATCAACACCTCTGAAAGAtagcgtgatcaccaccaccaccatggtgcgacagcagaTCTTTACAGAGTGTGATCTTt |
| mSXX2 | GCGGTTATTAGCACGTTTCT | Autour Intron 5 | 904 | F | ctagtCAGGCTGACAAACTTGAACAtctcgtgatcggcaccatgggggtggtggtgacagcgcctaTGTTCAGAGTGTGATCTTt |
| mSXX3 | CTGGAGGCCCTACACGAGGAAC | Début Exon 1 | 905 | R | ctagcCAGGCTGACAAACTTGAACAtagcgtgatcaccaccaccaccatggtgcgacagcagaTGTTCAGAGTGTGATCTTg |
| PAP6.1 | TCGGCACACGCGTCTTAAGTTG | Début Exon 13 | 906 | F | ctagtCCGAAAGGAGGCTGTGATtctcgtgatcggcaccatgggggtggtggtgacagcgcctaTAAACAGCAATACCGCATCAc |
| PAP6.2 | TTGGCATCTCTCCCTTTTAT | Début Exon 9 | 907 | R | ctagcCCGAAAGGAGGCTGTGATttagcgtgatcaccaccaccaccatggtgcgacagcagaTAAACAGCAATACCGCATCAc |
| PAP4 | CCGGCATTTTATCCAGTTG | Début Exon 1 | 909 | F | ctagtCGGCAATGCGGATTCATtctcgtgatcggcaccatgggggtggtggtgacagcgcctaTAAAGAGCAATACCGCATCAc |
| | | | 910 | F | ctagtCGGCTATTAGCAGCTTCAAtctcgtgatcggcaccatgggggtggtggtgacagcgcctaTGAACAGCTGTAAATAGCGCGc |
| | | | 911 | R | ctagcCGGCTATTAGCAGGTTCAAtagcgtgatcaccaccaccaccatggtgcgacagcagaTGAAGACGTTGTAATACGGCGCa |
| | | | 912 | F | ctagtCTGGAAGTCTACAGGAAAtctcgtgatcggcaccatgggggtggtggtgacagcgcctaTAAACAGCAATACCGCATCAc |
| | | | 913 | R | ctagcCTGGAAGTCTACAGGAAAttagcgtgatcaccaccaccaccatggtgcgacagcagaTAAACAGCAATACCGCATCAc |
| | | | 714 | F | ctagtTCGCCACAGCTTCTGGTTAAtctcgtgatcggcaccatgggggtggtggtgacagcgcctaTAACTAGAACGCTGTGTGGCGGAg |
| | | | 715 | R | ctagcTCGCCACAGCTTCTGGTTAAtagcgtgatcaccaccaccaccatggtgcgacagcagaTAACTAGAACGCTGTGTGGCGGAA |
| | | | 1088 | F | ctagtTTGCGCATCTCTCCCTTTTATtctcgtgatcggcaccatgggggtggtggtgacagcgcctaTAAAGAGGAGAGATGGCGCAAg |
| | | | 1089 | R | ctagcTTGCGCATCTCTCCCTTTTATttagcgtgatcaccaccaccaccatggtgcgacagcagaTAAATAGGGAGAGATGGCGCAAA |
| | | | 1086 | F | ctagtCCGGGGCATTATCTGGTTAAtctcgtgatcggcaccatgggggtggtggtgacagcgcctaTAACTGGATAAATATGCCCCGGGg |
| | | | 1087 | R | ctagcCCGGGGCATTATCTCCAGTTAAtagcgtgatcaccaccaccaccatggtgcgacagcagaTAACTGGATAAATATGCCCCGGGg |

Tableau des oligos utilisés pour les PCR sur colonies transformées avec le pHUN-90mer.

| Cible | N° | Sens | Séquence 5' -> 3' | Description |
|------------|----|------|----------------------|--|
| pHUN-90mer | 17 | F | GCTAGAGGTTGGGTCGGTG | Cible la fin du promoteur et le début de la séquence du précurseur miARN |
| | 18 | R | TAGGCTGATCACCAACCACC | Cible le 90mer |

4. RT

Tableau de l'amorce utilisée pour les RT d'ARN en ADNc.

| RT | | |
|----------|--------------------------------|-----|
| Nom | Séquence 5' -> 3' | N° |
| Oligo dT | TTTTTTTTTTTTTTTTTTTTTTTTTTTTTT | 221 |

5. RT-qPCR

Tableaux des oligos utilisés en qPCR, selon la catégorie de leur cible. **Orange** : gènes de référence ; **Jaune** : gènes nucléaires codant pour une protéine du mitoribosome ; **Turquoise** : gènes mitochondriaux.

| Gène | N° | Génome | Séquence 5' -> 3' | Description |
|-------|------|-----------|-------------------------|-----------------------------|
| Actin | 1035 | Nucléaire | F CTGACTCTGCGCTACCCCATT | Gène nucléaire de référence |
| | 1036 | | R CCTCAGTCAGCAGCACGGG | |
| CYN19 | 800 | Nucléaire | F CAAGACTGCGGAGAACTTCC | Gène nucléaire de référence |
| | 801 | | R CACATGAAGGAGGGGATGAC | |
| MAA7 | 96 | Nucléaire | F GTTCGGCGGCAAGTACGT | Gène nucléaire de référence |
| | 97 | | R CGCAATCGCCTCGTTGTAGT | |

| Gène | N° | Génome | Séquence 5' -> 3' | Description |
|-------|------|-----------|-------------------------------|---|
| mL113 | 953 | Nucléaire | F GACAGCTATCTCTGTCTTTGTTAGCTC | Gène nucléaire codant pour une protéine de la grande sous unité du mitoribosome |
| | 954 | | R ACTCCTCTCCCAGACAGCAC | |
| mL116 | 1001 | Nucléaire | F CTCAGACCGGCAACATC | Gène nucléaire codant pour une protéine de la grande sous unité du mitoribosome |
| | 1002 | | R TGAATCTCAGCGCGTAGG | |
| mL117 | 1003 | Nucléaire | F CAGCTGCAGCAACACCAG | Gène nucléaire codant pour une protéine de la grande sous unité du mitoribosome |
| | 1004 | | R CGACACGTTGGACTGGAAC | |
| mL118 | 1005 | Nucléaire | F ATCCCAGCGTCCAATCT | Gène nucléaire codant pour une protéine de la grande sous unité du mitoribosome |
| | 1006 | | R TTCGGAGCTGTTGGTGGT | |
| mS105 | 959 | Nucléaire | F TGATGAGCTGGATGACAACCT | Gène nucléaire codant pour une protéine de la petite sous unité du mitoribosome |
| | 960 | | R GAGGAAGCGCAGGTACTCC | |

| Gène | N° | Génome | Séquence 5' -> 3' | Description |
|------|-----|---------------|------------------------------|--|
| nd1 | 180 | Mitochondrial | F ACGCGACCCATTGATCT | Gène mitochondrial codant pour la sous unité 1 du complexe I de la chaîne respiratoire |
| | 181 | | R GGGCAAACCCAAACTACTG | |
| nd2 | 926 | Mitochondrial | F GCTTGCCACCATTTGCAGGTTTC | Gène mitochondrial codant pour la sous unité 2 du complexe I de la chaîne respiratoire |
| | 927 | | R CGCTCATAGCATGCCAGAAGATCC | |
| nd4 | 102 | Mitochondrial | F TGATGCCTGTACACTTGTGGTT | Gène mitochondrial codant pour la sous unité 4 du complexe I de la chaîne respiratoire |
| | 103 | | R CACTACCAGCAGTTGGAGCA | |
| nd5 | 172 | Mitochondrial | F CGCAGTCCAGTAAACCCAAAT | Gène mitochondrial codant pour la sous unité 5 du complexe I de la chaîne respiratoire |
| | 173 | | R TTAGCAAACCGCGTCCAC | |
| nd6 | 104 | Mitochondrial | F GTACTTTGCTGACTTGTGATTTTG | Gène mitochondrial codant pour la sous unité 6 du complexe I de la chaîne respiratoire |
| | 105 | | R CGAAACGACCAAACAGAGCTA | |
| cob | 333 | Mitochondrial | F CAAAGCAGCCTAATTTCTTTCG | Gène mitochondrial codant pour l'apocytochrome <i>b</i> du complexe IV de la chaîne respiratoire |
| | 334 | | R GAAGAAGACTAGAATGCTGAACACAA | |
| cox1 | 174 | Mitochondrial | F ACGAGGGTCGTGCTATGGTA | Gène mitochondrial codant pour la sous unité 1 du complexe III de la chaîne respiratoire |
| | 175 | | R CAGCCAAACCCAAGAAGT | |
| rtl | 928 | Mitochondrial | F TACGCCGACAGTACATCCAG | Gène mitochondrial codant pour une protéine reverse transcriptase-like |
| | 929 | | R TACCAAACCCAGGACGGAAG | |

Tableaux des oligos utilisés en qPCR, selon la catégorie de leur cible. **Bleu** : gènes des ARN ribosomiques du mitoribosome ; **Vert** : gènes nucléaires codant pour les PAP.

| Gène | N° | Génome | Séquence 5' -> 3' | Description |
|------|------------------------|-----------|---------------------------------------|---|
| PAP6 | 1015 | Nucléaire | F AGGTCATCAACGACGACAAGT | Gène nucléaire codant pour la protéine PAP6 |
| | 1016 | | R ACAGTCCTTGGACTGCCAAC | Couple d'amorces encadrant l'intron 11 |
| | 1011 | Nucléaire | F CCTTCACAACAAGAAGCCACT | Gène nucléaire codant pour la protéine PAP6 |
| | 1012 | | R CGAATGTCGCCATAAAAGG | Couple d'amorces encadrant l'intron 8 |
| PAP4 | 770 | Nucléaire | F GATTCCTTTGGGGCGACT | Gène nucléaire codant pour la protéine PAP4 |
| | 771 | | R ATGAAGCCGAGGATAGAACG | Couple d'amorces au début de l'exon 1 |
| | 826 | Nucléaire | F TGGAGCTGTGCCTGGAGT | Gène nucléaire codant pour la protéine PAP4 |
| | 827 | | R CATCACGTCCCTGCGACAC | Couple d'amorces encadrant l'intron 7 |
| | 1150 | Nucléaire | F CGGTCTCACGGCTTTACGG | Gène nucléaire codant pour la protéine PAP4 |
| | 1151 | | R GGTGATGGAGGCGTTGGTG | Couple d'amorces dans l'exon 2 |
| | 1150 | Nucléaire | F CGGTCTCACGGCTTTACGG | Gène nucléaire codant pour la protéine PAP4 |
| 625 | R CGAAACACCTCATACACCTC | | Couple d'amorces encadrant l'intron 2 | |

| Gène | N° | Génome | Séquence 5' -> 3' | Description |
|------|------|---------------|----------------------------|---|
| L1 | 993 | Mitochondrial | F AGGGATTGCTGACTGATAGTGC | Gène mitochondrial codant pour un ARN ribosomique de la grande sous unité du mitoribosome |
| | 994 | | R ATGCGCGAATGCTCACGTATTG | |
| L2a | 204 | Mitochondrial | F GCTTTCTGCTTAACTCACG | Gène mitochondrial codant pour un ARN ribosomique de la grande sous unité du mitoribosome |
| | 194 | | R GCGTACCTTTTGCATTATGG | |
| L2b | 995 | Mitochondrial | F GCGTGCAACAACACACTAAAGGG | Gène mitochondrial codant pour un ARN ribosomique de la grande sous unité du mitoribosome |
| | 996 | | R TGTGAGAGTGTGGGTTGACCAG | |
| L3a | 981 | Mitochondrial | F ACTCAATTTCCGTGCTCCGTGA | Gène mitochondrial codant pour un ARN ribosomique de la grande sous unité du mitoribosome |
| | 982 | | R GGGTCCCAGTACTAAGCCAGTAA | |
| L3b | 983 | Mitochondrial | F AACCGATCGATCTAGCGACGTG | Gène mitochondrial codant pour un ARN ribosomique de la grande sous unité du mitoribosome |
| | 984 | | R TCTTGCAACAGACACGGGTTCCG | |
| L4 | 989 | Mitochondrial | F ATACCGGTAGCGCAAGCGAATG | Gène mitochondrial codant pour un ARN ribosomique de la grande sous unité du mitoribosome |
| | 990 | | R TGACCTTACCGCCGATTGTCTG | |
| L5 | 975 | Mitochondrial | F AAGGTAGAGTCTGCGTTGAACACC | Gène mitochondrial codant pour un ARN ribosomique de la grande sous unité du mitoribosome |
| | 976 | | R TTCCAAGCACAGTGGCTGCTTC | |
| L6 | 973 | Mitochondrial | F ACGTTCTGGAACAATTACGCTGA | Gène mitochondrial codant pour un ARN ribosomique de la grande sous unité du mitoribosome |
| | 974 | | R GGGCTTACGGCCATGTTATCCA | |
| L7 | 977 | Mitochondrial | F ACGTTTAATTGGCGTCCCTTG | Gène mitochondrial codant pour un ARN ribosomique de la grande sous unité du mitoribosome |
| | 978 | | R AGCTGCATAGGGTCTCTTCG | |
| L8 | 1075 | Mitochondrial | F CGCCTATGCACAGTTAGCC | Gène mitochondrial codant pour un ARN ribosomique de la grande sous unité du mitoribosome |
| | 1076 | | R TGACAACAAGTGTACCAGAGGTTT | |
| S1 | 1037 | Mitochondrial | F GTTTGGTGCCTGGCTCAGCTTTC | Gène mitochondrial codant pour un ARN ribosomique de la petite sous unité du mitoribosome |
| | 1038 | | R ATTACGCACTCGTGGCTTAATC | |
| S2 | 979 | Mitochondrial | F CCCATGGCTTATCCTTAGCCTGTC | Gène mitochondrial codant pour un ARN ribosomique de la petite sous unité du mitoribosome |
| | 980 | | R ATCCGCCGTTCAAGCGATTGTC | |
| S3 | 1071 | Mitochondrial | F TTCAAATCTATAACGCCTGGTG | Gène mitochondrial codant pour un ARN ribosomique de la petite sous unité du mitoribosome |
| | 1072 | | R GCTGCCAATTGCTTTGATTT | |
| S4 | 991 | Mitochondrial | F ACGGTCGAGCAAGTCATCATGG | Gène mitochondrial codant pour un ARN ribosomique de la petite sous unité du mitoribosome |
| | 992 | | R TGGTATCCAACCTGTGGGTGTC | |

6. 3'RACE-seq

Tableaux des oligos utilisés pour le 3'RACE-seq, en fonction de leur utilité. **Bleu clair** : séquence Illumina ajoutée aux oligos forward ; **Jaune** : 5 nucléotides de la séquence de l'adaptateur manquant dans le RT-short mais présents dans l'oligonucléotide reverse PCR1 ; **Rose** : séquence servant de délimitant lors de l'analyse bio-informatique ; **Rouge** : séquence spécifique du gène à amplifier ; **Vert clair** : séquence aléatoire, permettant d'enlever les duplicats de PCR lors de l'analyse bio-informatique ; N = A, G, C ou T ; **Vert foncé** : séquence du RT-short et du complémentaire au RT-short.

Ligation

| Nom | Séquence 5' -> 3' |
|-----|--|
| R21 | /5App/ CTGACNNNNNNNNNNNNNNNNNTGGAAT TCTCGGGTGCCAAGGC /3ddC/ |

RT

| Nom | Séquence 5' -> 3' | N° |
|----------|--------------------------|------|
| RT-short | GCCTTGGCACCCGAGAA | 1099 |

Oligos reverse pour la PCR1

| Nom | Séquence 5' -> 3' | N° |
|---------|------------------------------|------|
| PCR1-Rv | CCTTGGCACCCGAGAATTCCA | 1100 |

Oligos forward pour la PCR1

| Nom | Séquence 5' -> 3' | N° | Taille |
|------|--------------------------------|-----|--------|
| cob | ACCCAGACTTGTGGGTCAC | 108 | 477 |
| nd4 | CTATACTTCAGCGGTCTAGGTG | 282 | 388 |
| nd5 | CCTTTGGTTTGATGGTTGC | 122 | 524 |
| cox1 | CACAACACTTCTTGGGTTTGGC | 285 | 431 |
| nd2 | CATTTGGGTATCTACATTATTAC | 287 | 496 |
| nd6 | GCATTGCCGTAGGATACACC | 289 | 507 |
| nd1 | ACGCGACCCATTTGATCT | 180 | 363 |
| rtl | CTCTGCTACAAAAAGCG | 127 | 492 |

Oligos forward avec les séquences Illumina (oligos spécifiques des gènes) pour PCR2

| Nom | Séquence 5' -> 3' | N° | Taille |
|------|---|-----|--------|
| cob | AATGATACGGCGACCACCGAGATCTACACGTTTCAGAGTTCTACAGTCCGACGATCGGGC GTATTGGCTATTGGTC | 281 | 455 |
| nd4 | AATGATACGGCGACCACCGAGATCTACACGTTTCAGAGTTCTACAGTCCGACGATCCCA CTATCCCTAACTTC | 283 | 416 |
| nd5 | AATGATACGGCGACCACCGAGATCTACACGTTTCAGAGTTCTACAGTCCGACGATCTGC GTGCTTGTTTCAATGTTG | 284 | 371 |
| cox1 | AATGATACGGCGACCACCGAGATCTACACGTTTCAGAGTTCTACAGTCCGACGATCGAT TATGCTGACTGCTTTGC | 286 | 496 |
| nd2 | AATGATACGGCGACCACCGAGATCTACACGTTTCAGAGTTCTACAGTCCGACGATGGCT GCTGCTTGTCTGC | 288 | 437 |
| nd6 | AATGATACGGCGACCACCGAGATCTACACGTTTCAGAGTTCTACAGTCCGACGATGGTG CATTGGCTGTACTATTCCTATTCG | 350 | 481 |
| nd1 | AATGATACGGCGACCACCGAGATCTACACGTTTCAGAGTTCTACAGTCCGACGATGGCT GTAATGAGTGCTATTGC | 292 | 366 |
| rtl | AATGATACGGCGACCACCGAGATCTACACGTTTCAGAGTTCTACAGTCCGACGATCCAA CACTCCCACAAGTATG | 293 | 446 |

Tableaux des amorces Illumina pour les PCR Spike et pour l'ajout des RPI (*Read Primer Index*), utilisés pour le 3'RACE-seq. **Bleu clair** : séquence Illumina ajoutée aux oligos forward ; **Bleu foncé** : séquence Illumina ajoutée aux oligos reverse ; **Orange** : code-barre spécifique à chaque RPI ; **Violet** : région de liaison à l'oligonucléotide du séquençage Illumina.

Oligo forward pour les PCR Spike-in

| Nom | Séquence 5' -> 3' |
|-----|---|
| RP1 | AATGATACGGCGACCACCGAGATCTACACGTTCTACAGTCCGA |

Oligos reverse avec le code-barre (oligos TruSeq) pour la PCR2 et les PCR Spike-in

| Nom | Séquence 5' -> 3' | Code-barre |
|------|--|------------|
| RP1 | CAAGCAGAAGACGGCATAACGAGATCGTGATGTGACTGGAGTTCCTTGGCACCCGAGAATTCCA | ATCACG |
| RP12 | CAAGCAGAAGACGGCATAACGAGATACATCGGTGACTGGAGTTCCTTGGCACCCGAGAATTCCA | CGATGT |
| RP13 | CAAGCAGAAGACGGCATAACGAGATGCCAAGTGACTGGAGTTCCTTGGCACCCGAGAATTCCA | TTAGGC |
| RP14 | CAAGCAGAAGACGGCATAACGAGATGGTTCAGTGACTGGAGTTCCTTGGCACCCGAGAATTCCA | TGACCA |
| RP15 | CAAGCAGAAGACGGCATAACGAGATCACTGTGTGACTGGAGTTCCTTGGCACCCGAGAATTCCA | ACAGTG |
| RP16 | CAAGCAGAAGACGGCATAACGAGATATTGGCGTGACTGGAGTTCCTTGGCACCCGAGAATTCCA | GCCAAAT |
| RP17 | CAAGCAGAAGACGGCATAACGAGATGATCTGGTGACTGGAGTTCCTTGGCACCCGAGAATTCCA | CAGATC |
| RP18 | CAAGCAGAAGACGGCATAACGAGATTCAAGTGTGACTGGAGTTCCTTGGCACCCGAGAATTCCA | ACTTGA |
| RP19 | CAAGCAGAAGACGGCATAACGAGATCTGATCGTGACTGGAGTTCCTTGGCACCCGAGAATTCCA | GATCAG |
| RP10 | CAAGCAGAAGACGGCATAACGAGATAAGCTAGTGACTGGAGTTCCTTGGCACCCGAGAATTCCA | TAGCTT |
| RP11 | CAAGCAGAAGACGGCATAACGAGATGTAGCCGTGACTGGAGTTCCTTGGCACCCGAGAATTCCA | GGCTAC |
| RP12 | CAAGCAGAAGACGGCATAACGAGATTACAAGTGACTGGAGTTCCTTGGCACCCGAGAATTCCA | CTTGTA |
| RP13 | CAAGCAGAAGACGGCATAACGAGATTTGACTGTGACTGGAGTTCCTTGGCACCCGAGAATTCCA | AGTCAA |
| RP14 | CAAGCAGAAGACGGCATAACGAGATGGAACGTGACTGGAGTTCCTTGGCACCCGAGAATTCCA | AGTTCC |
| RP15 | CAAGCAGAAGACGGCATAACGAGATTGACATGTGACTGGAGTTCCTTGGCACCCGAGAATTCCA | ATGTCA |
| RP16 | CAAGCAGAAGACGGCATAACGAGATGGACGGGTGACTGGAGTTCCTTGGCACCCGAGAATTCCA | CCGTCC |
| RP17 | CAAGCAGAAGACGGCATAACGAGATCTCTACGTGACTGGAGTTCCTTGGCACCCGAGAATTCCA | GTAGAG |
| RP18 | CAAGCAGAAGACGGCATAACGAGATGCGGACGTGACTGGAGTTCCTTGGCACCCGAGAATTCCA | GTCCGC |
| RP19 | CAAGCAGAAGACGGCATAACGAGATTTTCACGTGACTGGAGTTCCTTGGCACCCGAGAATTCCA | GTGAAA |
| RP20 | CAAGCAGAAGACGGCATAACGAGATGGCCACGTGACTGGAGTTCCTTGGCACCCGAGAATTCCA | GTGGCC |
| RP21 | CAAGCAGAAGACGGCATAACGAGATCGAAACGTGACTGGAGTTCCTTGGCACCCGAGAATTCCA | GTTTCG |
| RP22 | CAAGCAGAAGACGGCATAACGAGATCGTACGGTGACTGGAGTTCCTTGGCACCCGAGAATTCCA | CGTACG |
| RP23 | CAAGCAGAAGACGGCATAACGAGATCCACTCGTGACTGGAGTTCCTTGGCACCCGAGAATTCCA | GAGTGG |
| RP24 | CAAGCAGAAGACGGCATAACGAGATGCTACCCTGACTGGAGTTCCTTGGCACCCGAGAATTCCA | GGTAGC |
| RP25 | CAAGCAGAAGACGGCATAACGAGATATCAGTGTGACTGGAGTTCCTTGGCACCCGAGAATTCCA | ACTGAT |
| RP26 | CAAGCAGAAGACGGCATAACGAGATGCTCATGTGACTGGAGTTCCTTGGCACCCGAGAATTCCA | ATGAGC |
| RP27 | CAAGCAGAAGACGGCATAACGAGATAGGAATGTGACTGGAGTTCCTTGGCACCCGAGAATTCCA | ATTCTT |
| RP28 | CAAGCAGAAGACGGCATAACGAGATCTTTTGGTGACTGGAGTTCCTTGGCACCCGAGAATTCCA | CAAAAG |
| RP29 | CAAGCAGAAGACGGCATAACGAGATTAGTTGGTGACTGGAGTTCCTTGGCACCCGAGAATTCCA | CAACTA |
| RP30 | CAAGCAGAAGACGGCATAACGAGATCCGGTGGTGACTGGAGTTCCTTGGCACCCGAGAATTCCA | CACCGG |
| RP31 | CAAGCAGAAGACGGCATAACGAGATATCGTGGTGACTGGAGTTCCTTGGCACCCGAGAATTCCA | CACGAT |

www.chlamycollection.org) ; pH 7,2 avec de l'acide acétique glacial

Milieu TAP agar : agar 1,25% (p/v) ; milieu TAP

Milieu TAP-Sucrose : milieu TAP ; sucrose 40 mM

Milieu TP : Tris 293 M ; NH₄Cl 7 mM ; MgSO₄.7H₂O 406 μM ; CaCl₂.2H₂O 340 μM ; K₂HPO₄ 537 μM ; KH₂PO₄ 460 μM ; Hutner's trace elements 0,1% (v/v) (Chlamydomonas Resource Center www.chlamycollection.org) ; pH 7,2 avec de l'acide chlorhydrique

Milieu TP agar : agar 1,25% (p/v) ; milieu TP

Tampon d'infiltration : MgCl₂ 1 mM ; acétosyringone 50 μM

Tampon de digestion 1X : MSO222 (430 mg/100mL ; Duchefa) ; MES 2,4 mM ; sucrose 58,5 mM ; mannitol 440 mM ; macérozyme (40 mg/100 mL ; Yakult Pharmaceutical Industry Co., Ltd.) ; cellulase R10 (200 mg/100 mL ; Yakult Pharmaceutical Industry Co., Ltd.) ; driselase (100 mg/100 mL ; Sigma)

2. Milieux et solutions relatifs aux ADN

dNTP : dATP 10mM ; dCTP 10 mM ; dGTP 10 mM ; dTTP 10 mM

Gel agarose ADN : agarose (p/v, selon pourcentage souhaité) ; TAE 0,5X ; BET (0,7 mg/mL pour 100 mL de gel)

Gel agarose ultra-pure ADN : agarose ultra-pure (Invitrogen) (p/v, selon pourcentage souhaité) ; TAE 0,5X ; BET (0,7 mg/mL pour 100 mL de gel)

TAE 0,5X : Tris-acétate 20 mM pH 8 ; EDTA 1 mM

Tampon d'hybridation 2X : Tris 20 mM pH 8.0 ; EDTA 2 mM ; NaCl 100 mM

3. Milieux et solutions relatifs aux ARN

Gel agarose ARN : agarose 1,2% (p/v) ; TBE 1X (v/v)

Gel acrylamide 6% : Acryl 19/1 6% ; TBE 1X (v/v) ; APS 0,1% (p/v) ; TEMED 0,01% (v/v)

High salt : Citrate de Sodium 0,8 M ; NaCl 1,2 M

Tampon de charge ARN : formamide 54% (v/v) ; formaldéhyde 21,5% (v/v) ; TBE 1X ; EDTA 1 mM pH 8 ; BET 13,4 ng/μL ; bleu de bromophénol 0,035% (p/v) ; xylène cyanol 0,035% (p/v)

TBE 1X : Tris 90 mM ; EDTA 2 mM pH 8,0 ; acide borique 90 mM

4. Milieux et solutions relatifs aux protéines

Gel acrylamide 20% dénaturant : acrylamide 20% / N,N' méthylène bisacrylamide (19/1) ; urée 7 M ; TBE 1X

Gel de concentration : Acrylamide/bisacrylamide 37,5/1 5% (v/v) ; Tris-HCl 100 mM pH6,8 ; SDS 0,1% (v/v) ; APS 0,1% (p/v) ; TEMED 0,01% (v/v)

Gel de séparation : Acrylamide/bisacrylamide 37,5/1 (v/v, selon pourcentage souhaité) ; Tris-HCl 380 mM pH8,8 ; SDS 0,1% (v/v)

Laemmli 5X : Tris-HCl 250 mM, pH 6,8 ; Glycérol 37,5% (v/v) ; SDS 10% (p/v) ; β -mercaptoéthanol 10% (v/v) ; Bleu de Bromophénol 0,125% (p/v)

Milieu auto-inductible : NZ amines AS 0,9% ; extraits de levures 0,45% ; glycérol 54,3 mM ; glucose 2,8 mM ; α -lactose 5,8 mM ; Na_2HPO_4 25 mM ; KH_2PO_4 25 mM ; NH_4Cl 50,1 mM ; Na_2SO_4 774,4 μM ; FeCl_3 50 μM ; HCl 50 μM ; CaCl_2 40 μM ; MnCl_2 40 μM ; ZnSO_4 20 μM ; CoCl_2 4 μM ; CuCl_2 4 μM ; NiCl_2 4 μM ; Na_2MoO_4 4 μM ; Na_2SeO_3 4 μM ; H_3BO_3 4 μM

rNTP : rATP 10mM ; rCTP 10 mM ; rGTP 10 mM ; rUTP 10 mM

Solution A : KPO_4 25 mM ; PEG 6000 6% ; BSA 0,4% ; digitonine 0,016% ; pH 6,5

Solution B : HEPES 20 mM ; Mannitol 150 mM ; EDTA 2 mM ; MgCl_2 1 mM ; BSA 0,1% ; pH 7,2

Solution D : HEPES 20 mM ; Mannitol 150 mM ; EDTA 0,8 mM ; MgCl_2 4 mM ; pH 7,2

Solution de coloration WB : Bleu de Coomassie R250 0,1% (p/v) ; éthanol 50% (v/v) ; acide acétique 7% (v/v)

Solution de décoloration WB : éthanol 50% (v/v) ; acide acétique 7% (v/v)

Solution MET : Mannitol 280 mM ; EDTA 0,5 mM ; Tris-HCl 10 mM ; pH 7

Solution percoll 13% : Percoll 13% ; solution MET+BSA 87% ; Mannitol 250 mM

Solution percoll 21% : Percoll 21% ; solution MET+BSA 79% ; Mannitol 250 mM

Solution percoll 45% : Percoll 45% ; solution MET+BSA 55% ; Mannitol 250 mM

Solution STOP : formamide désionisée 95% (v/v) ; EDTA 20 mM ; xylène cyanol 0,05% (p/v) ; bleu de bromophénol 0,05% (p/v)

Tampon bloquant WB : Tris-HCl 25 mM pH 7,6 ; NaCl 150 mM ; Tween-20 0,2% (v/v) ; lait écrémé 5% (p/v)

Tampon d'activité : HEPES-KOH 20 mM pH 7.5 ; MgCl_2 5 mM ; DTT 5 mM ; Glycérol 10% (v/v) ; BSA 0,1 $\mu\text{g}/\mu\text{l}$; mix rNTP 1 mM

Tampon d'élution 1 PP : Tris 50 mM pH8 ; NaCl 300 mM ; Glycérol 5% ; Imidazole 250 mM

Tampon d'élution 2 PP : Tris 50 mM pH8 ; NaCl 300 mM ; Glycérol 5% ; Imidazole 500 mM

Tampon d'élution IP : Tris HCl 50 mM (pH 6.8) ; DTT 50 mM ; SDS 1% ; bleu de bromophénol 0,005% ; Glycérol 10%

Tampon de binding PP : Tris 50mM pH8 ; NaCl 300mM ; Glycérol 5% ; Imidazole 25 mM

Tampon de lavage IP : HEPES-KOH 20 mM pH 7,6 ; KCl 100 mM ; MgCl_2 30 mM ; DTT 1 mM ; Dodécyl- β -D-maltoside 0,01% ; chloramphenicol 100 $\mu\text{g}/\text{mL}$; Triton X-100 0,1% ; cOmplete™ EDTA-free Protease Inhibitor Cocktail 1X (Roche #04693132001)

Tampon de lavage PP : Tris 50mM pH8 ; NaCl 300mM ; Glycérol 5% ; Imidazole 50 mM

Tampon de lavage WB : Tris-HCl 25 mM, pH 7,6 ; NaCl 150 mM ; Tween-20 0,2% (v/v)

Tampon de lyse bactéries : Tris 50 mM pH8 ; NaCl 300 mM ; Glycérol 5% ; CHAPS 0.5% ; DTT 5 mM

Tampon de lyse IP : HEPES-KOH 20 mM pH 7,6 ; KCl 100 mM ; MgCl₂ 30 mM ; DTT 1 mM ; Dodécyl-β-D-maltoside 0,3% ; chloramphénicol 100 µg/mL ; Triton X-100 1% ; cOmplete™ EDTA-free Protease Inhibitor Cocktail 1X (Roche #04693132001)

Tampon gel filtration PP : Tris 50mM pH8 ; NaCl 300mM ; Glycérol 5%

Tampon TGS : Tris 25 mM ; glycine 250 mM ; SDS 0,1% (p/v)

V. Modules Golden Braid

Tableau des modules utilisés lors du clonage Golden Braid. Pour le tag 6xHis-SUMO, c'est la séquence SUMO humaine qui est utilisée.

| Nom | N° | Description |
|-----------|---------|---|
| pT7 A1B1 | P3P-394 | Promoteur T7 ; Séquence pour assemblage : 5' : A1 et 3' : B1 |
| HSUM B2B3 | P3P-396 | Tag 6xHis-SUMO ; Séquences pour assemblage : 5' : B2 et 3' : B3 |
| tT7 B5C1 | P3P-388 | Terminateur T7 ; Séquence pour assemblage : 5' : B5 et 3' : C1 |

VI. Activités *in vitro*

Tableau des séquences ARN utilisées pour les tests d'activités *in vitro*. Tailles en pb.

| Cible | Séquence 5' -> 3' | Taille |
|------------|-----------------------|--------|
| 3'UTR_nad6 | GACGUUUUGCUUUUUUCUUU | 19 |
| 3'UTR_nad4 | CUUUGUUUUUUUUUUUUACA | 21 |
| 3'CDS_nad6 | GUUCGUACUACCGGUCGUUAA | 21 |

VII. Données brutes des analyses de spectrométrie de masse

Tableau 1 : récapitulatif des données obtenues dans les 3 réplicats des 2 fractions

| Identification par Mascot et validation à FDR<1% par Proline (ProFi consortium) | Matrix | | | Membrane | | |
|--|--------|------|------|----------|------|-------|
| | #1 | #2 | #3 | #1 | #2 | #3 |
| Nombre total de protéines identifiées | 894 | 505 | 888 | 969 | 673 | 1474 |
| Nombre total de spectres identifiés | 9898 | 5367 | 9416 | 7010 | 7471 | 12634 |

Légende des tableaux 2 et 3 suivants :

Bleu : localisation mitochondriale des protéines
Vert : localisation chloroplastique des protéines
Noir : localisation non organellaire des protéines

| |
|----------------|
| 1 spectre |
| 2-5 spectres |
| 6-10 spectres |
| 11-30 spectres |
| >= 31 spectres |

Présentation des 3 premières pages de données (202 protéines) pour les 2 tableaux, en ordre décroissant, selon le nombre de spectres identifiés dans le réplicat avec le plus de protéines identifiées (Tableau 1) : réplicat 1 de la fraction matricielle (Tableau 2, gauche) et réplicat 3 de la fraction membranaire (Tableau 3, droite)

Tableau 2 : classement décroissant des protéines selon le nombre de spectres identifiés dans le réplicat 1 de l'échantillon matrice

| Accession | Matrice | | | Membrane | | |
|--------------------|---------|-----|-----|----------|-----|-----|
| | #1 | #2 | #3 | #1 | #2 | #3 |
| Cre01.g042750.t1.2 | 343 | 274 | 358 | 78 | 129 | 184 |
| Cre06.g309100.t1.2 | 171 | 92 | 183 | 74 | 10 | 78 |
| Cre09.g393200.t1.2 | 161 | 104 | 204 | 62 | 3 | 93 |
| Cre12.g537200.t1.2 | 150 | 131 | 141 | 68 | 34 | 89 |
| Cre05.g248400.t1.2 | 139 | 71 | 111 | 56 | 68 | 85 |
| Cre01.g044800.t1.2 | 124 | 98 | 105 | 24 | 53 | 63 |
| Cre10.g451950.t1.2 | 124 | 81 | 89 | 88 | 195 | 213 |
| Cre09.g387726.t1.1 | 117 | 87 | 120 | 26 | 44 | 46 |
| Cre07.g353450.t1.2 | 115 | 85 | 104 | 30 | 36 | 54 |
| Cre12.g534800.t1.1 | 115 | 114 | 95 | 94 | 144 | 145 |
| Cre16.g675650.t1.2 | 107 | 57 | 107 | 66 | 31 | 93 |
| Cre16.g664550.t1.2 | 92 | 60 | 78 | 52 | 88 | 125 |
| Cre12.g483950.t1.2 | 89 | 71 | 102 | 11 | 16 | 25 |
| Cre12.g500150.t1.1 | 86 | 54 | 78 | 60 | 63 | 83 |
| Cre18.g749847.t1.1 | 85 | 65 | 71 | 45 | 92 | 92 |
| Cre03.g193850.t1.2 | 84 | 38 | 56 | 74 | 116 | 127 |
| Cre06.g263250.t1.1 | 78 | 2 | 131 | 58 | 135 | 170 |
| Cre06.g294650.t1.2 | 77 | 60 | 65 | 19 | 35 | 48 |
| Cre17.g743897.t1.1 | 73 | 34 | 68 | 146 | 187 | 203 |
| Cre07.g343700.t1.2 | 72 | 59 | 58 | 24 | 44 | 67 |
| Cre04.g214500.t1.1 | 72 | 80 | 53 | 76 | 108 | 113 |
| Cre11.g467350.t1.2 | 70 | 43 | 50 | 47 | 52 | 69 |
| Cre03.g193750.t1.1 | 69 | 43 | 68 | 33 | 55 | 72 |
| Cre07.g321400.t1.1 | 69 | 16 | 61 | 34 | 4 | 62 |
| Cre02.g143250.t1.2 | 68 | 47 | 59 | 69 | 132 | 110 |
| Cre06.g254400.t1.1 | 65 | 42 | 54 | 48 | 20 | 62 |
| Cre09.g392867.t1.1 | 61 | 28 | 69 | 8 | 23 | 30 |
| Cre03.g149100.t1.2 | 61 | 45 | 52 | 35 | 36 | 42 |
| Cre01.g005150.t1.1 | 60 | 46 | 57 | 12 | 26 | 32 |
| Cre06.g295450.t1.2 | 60 | 35 | 47 | 22 | 23 | 36 |
| Cre17.g703700.t1.1 | 58 | 42 | 52 | 17 | 7 | 38 |
| Cre16.g695050.t1.2 | 57 | 23 | 45 | 23 | 42 | 56 |
| Cre06.g259150.t1.2 | 57 | 51 | 43 | 2 | | 7 |
| Cre06.g253350.t1.2 | 55 | 29 | 77 | 1 | 4 | 12 |
| Cre17.g699000.t1.2 | 53 | 22 | 38 | 30 | 37 | 39 |
| Cre11.g477950.t1.1 | 52 | 1 | 69 | 53 | 69 | 66 |
| Cre12.g495850.t1.2 | 52 | 37 | 52 | 38 | 16 | 54 |
| Cre06.g251400.t1.2 | 52 | 38 | 46 | 1 | 2 | 9 |
| Cre10.g447100.t1.2 | 51 | 34 | 69 | 56 | 62 | 68 |
| Cre08.g370450.t1.2 | 50 | 18 | 80 | 36 | 9 | 20 |
| Cre03.g199000.t1.2 | 49 | 3 | 18 | 32 | 90 | 74 |
| Cre02.g080650.t1.2 | 48 | 4 | 18 | 15 | | 31 |
| Cre02.g080700.t1.2 | 47 | 13 | 49 | 19 | 25 | 32 |
| Cre16.g677500.t1.2 | 47 | 29 | 39 | 30 | 36 | 49 |
| Cre03.g178450.t1.2 | 46 | 23 | 78 | 23 | 30 | 40 |
| Cre06.g268750.t1.2 | 46 | 34 | 35 | 15 | 16 | 25 |
| Cre05.g232150.t1.2 | 46 | 31 | 29 | 1 | 6 | 9 |
| Cre02.g116750.t1.1 | 46 | 77 | 27 | 36 | 71 | 63 |
| Cre17.g727300.t1.2 | 45 | 35 | 39 | 19 | 24 | 34 |
| Cre02.g080600.t1.2 | 45 | 11 | 36 | 33 | 12 | 53 |
| Cre17.g723650.t1.2 | 44 | 19 | 47 | 8 | 4 | 24 |
| Cre13.g605150.t1.2 | 44 | 28 | 34 | 45 | 19 | 47 |
| Cre07.g337650.t1.2 | 44 | 26 | 33 | 7 | 11 | 15 |
| Cre09.g386735.t1.1 | 43 | 20 | 41 | 7 | 31 | 60 |
| Cre06.g252550.t1.1 | 43 | 21 | 32 | 4 | 15 | 19 |
| Cre02.g107450.t1.2 | 42 | 27 | 33 | 20 | 4 | 36 |
| Cre17.g728800.t1.2 | 41 | 36 | 41 | 32 | 38 | 48 |
| Cre12.g514750.t1.2 | 41 | 23 | 34 | 28 | 40 | 44 |
| Cre11.g474800.t1.2 | 41 | 22 | 24 | 27 | 5 | 40 |
| Cre06.g288700.t1.2 | 41 | 13 | 14 | 41 | 68 | 86 |
| Cre02.g088200.t1.2 | 40 | 8 | 40 | 1 | | 3 |
| Cre03.g148300.t1.2 | 39 | 13 | 35 | 29 | 39 | 52 |
| Cre06.g277200.t1.2 | 39 | 12 | 16 | 40 | 47 | 61 |
| Cre14.g630859.t1.1 | 38 | 30 | 44 | 7 | 3 | 22 |

Tableau 3 : classement décroissant des protéines selon le nombre de spectres identifiés dans le réplicat 3 de l'échantillon membrane

| Accession | Matrice | | | Membrane | | |
|--------------------|---------|-----|-----|----------|-----|-----|
| | #1 | #2 | #3 | #1 | #2 | #3 |
| Cre17.g698000.t1.2 | 27 | 51 | 22 | 88 | 195 | 213 |
| Cre02.g116750.t1.1 | 46 | 77 | 27 | 146 | 187 | 203 |
| Cre01.g042750.t1.2 | 343 | 274 | 358 | 78 | 129 | 184 |
| Cre09.g415550.t1.2 | 23 | 20 | 16 | 58 | 135 | 170 |
| Cre07.g340350.t1.1 | 8 | 29 | 8 | 94 | 144 | 145 |
| Cre13.g581600.t1.2 | 21 | 23 | 17 | 74 | 116 | 127 |
| Cre07.g353450.t1.2 | 115 | 85 | 104 | 52 | 88 | 125 |
| Cre12.g509750.t1.2 | 21 | 25 | 12 | 76 | 108 | 113 |
| Cre09.g386650.t1.2 | 14 | 23 | 21 | 69 | 132 | 110 |
| Cre07.g321400.t1.1 | 69 | 16 | 61 | 66 | 31 | 93 |
| Cre11.g477950.t1.1 | 52 | 1 | 69 | 62 | 3 | 93 |
| Cre12.g535950.t1.2 | 29 | 22 | 12 | 45 | 92 | 92 |
| Cre09.g392867.t1.1 | 61 | 28 | 69 | 68 | 34 | 89 |
| Cre12.g523850.t1.2 | 23 | 17 | 20 | 41 | 68 | 86 |
| Cre09.g393200.t1.2 | 161 | 104 | 204 | 56 | 68 | 85 |
| Cre06.g262700.t1.1 | 3 | 2 | 3 | 48 | 63 | 84 |
| Cre06.g288700.t1.2 | 41 | 13 | 14 | 60 | 63 | 83 |
| Cre06.g263250.t1.1 | 78 | 2 | 131 | 74 | 10 | 78 |
| Cre03.g144807.t1.1 | 35 | 17 | 29 | 32 | 90 | 74 |
| Cre16.g664550.t1.2 | 92 | 60 | 78 | 33 | 55 | 72 |
| Cre16.g672650.t1.2 | 9 | 16 | 8 | 64 | 62 | 72 |
| Cre11.g467707.t1.1 | 7 | 5 | 8 | 27 | 30 | 71 |
| Cre12.g483950.t1.2 | 89 | 71 | 102 | 47 | 52 | 69 |
| Cre16.g690431.t1.1 | 16 | 11 | 15 | 56 | 62 | 68 |
| Cre12.g534800.t1.1 | 115 | 114 | 95 | 24 | 44 | 67 |
| Cre05.g241950.t1.2 | 16 | 15 | 22 | 53 | 69 | 66 |
| Cre15.g635850.t1.2 | 10 | 11 | 7 | 38 | 51 | 65 |
| Cre06.g309100.t1.2 | 171 | 92 | 183 | 24 | 53 | 63 |
| Cre01.g013700.t1.2 | 18 | 11 | 15 | 36 | 71 | 63 |
| Cre02.g080700.t1.2 | 47 | 13 | 49 | 48 | 20 | 62 |
| Cre01.g005050.t1.1 | 7 | | 12 | 34 | 4 | 62 |
| Cre12.g500500.t1.1 | 13 | 7 | 13 | 40 | 47 | 61 |
| Cre06.g263450.t1.2 | 31 | 28 | 31 | 7 | 31 | 60 |
| Cre03.g149100.t1.2 | 61 | 45 | 52 | 24 | 40 | 58 |
| Cre15.g639150.t1.1 | 7 | 1 | 5 | 23 | 42 | 56 |
| Cre07.g338050.t1.2 | 7 | 12 | 5 | 39 | 45 | 55 |
| Cre12.g537200.t1.2 | 150 | 131 | 141 | 30 | 36 | 54 |
| Cre02.g080600.t1.2 | 45 | 11 | 36 | 38 | 16 | 54 |
| Cre06.g293850.t1.2 | 2 | 3 | 1 | 36 | 55 | 54 |
| Cre02.g088200.t1.2 | 40 | 8 | 40 | 33 | 12 | 53 |
| Cre14.g619133.t1.1 | 35 | 20 | 32 | 29 | 39 | 52 |
| Cre05.g248400.t1.2 | 139 | 71 | 111 | 32 | 26 | 51 |
| Cre09.g415850.t1.2 | | 4 | | 42 | 55 | 51 |
| Cre09.g393150.t1.2 | 8 | 5 | 11 | 18 | 15 | 50 |
| Cre03.g172300.t1.2 | 1 | 5 | 3 | 37 | 49 | 50 |
| Cre09.g387726.t1.1 | 117 | 87 | 120 | 30 | 36 | 49 |
| Cre16.g695050.t1.2 | 57 | 23 | 45 | 19 | 35 | 48 |
| Cre16.g680000.t1.2 | 9 | 11 | 7 | 30 | 45 | 48 |
| Cre03.g144967.t1.1 | 2 | 2 | 2 | 32 | 38 | 48 |
| Cre02.g143250.t1.2 | 68 | 47 | 59 | 17 | 26 | 47 |
| Cre12.g507400.t1.2 | 22 | 12 | 23 | 22 | 29 | 47 |
| Cre02.g081050.t1.2 | 15 | 2 | 17 | 45 | 19 | 47 |
| Cre10.g434450.t1.2 | 5 | 5 | 3 | 20 | 38 | 47 |
| Cre01.g044800.t1.2 | 124 | 98 | 105 | 26 | 44 | 46 |
| Cre13.g567600.t1.2 | 7 | 10 | 15 | 53 | 34 | 45 |
| Cre16.g691552.t1.1 | 2 | 2 | 2 | 16 | 34 | 45 |
| Cre17.g723650.t1.2 | 44 | 19 | 47 | 28 | 40 | 44 |
| Cre02.g119550.t1.2 | 31 | 9 | 43 | 23 | 26 | 44 |
| Cre09.g416150.t1.2 | 5 | 8 | 5 | 25 | 43 | 43 |
| Cre18.g749847.t1.1 | 85 | 65 | 71 | 35 | 36 | 42 |
| Cre12.g522600.t1.2 | 28 | 10 | 33 | 34 | 18 | 42 |
| Cre16.g675650.t1.2 | 107 | 57 | 107 | 23 | 30 | 40 |
| Cre12.g512000.t1.1 | 6 | | 12 | 27 | 5 | 40 |
| Cre03.g193750.t1.1 | 69 | 43 | 68 | 30 | 37 | 39 |

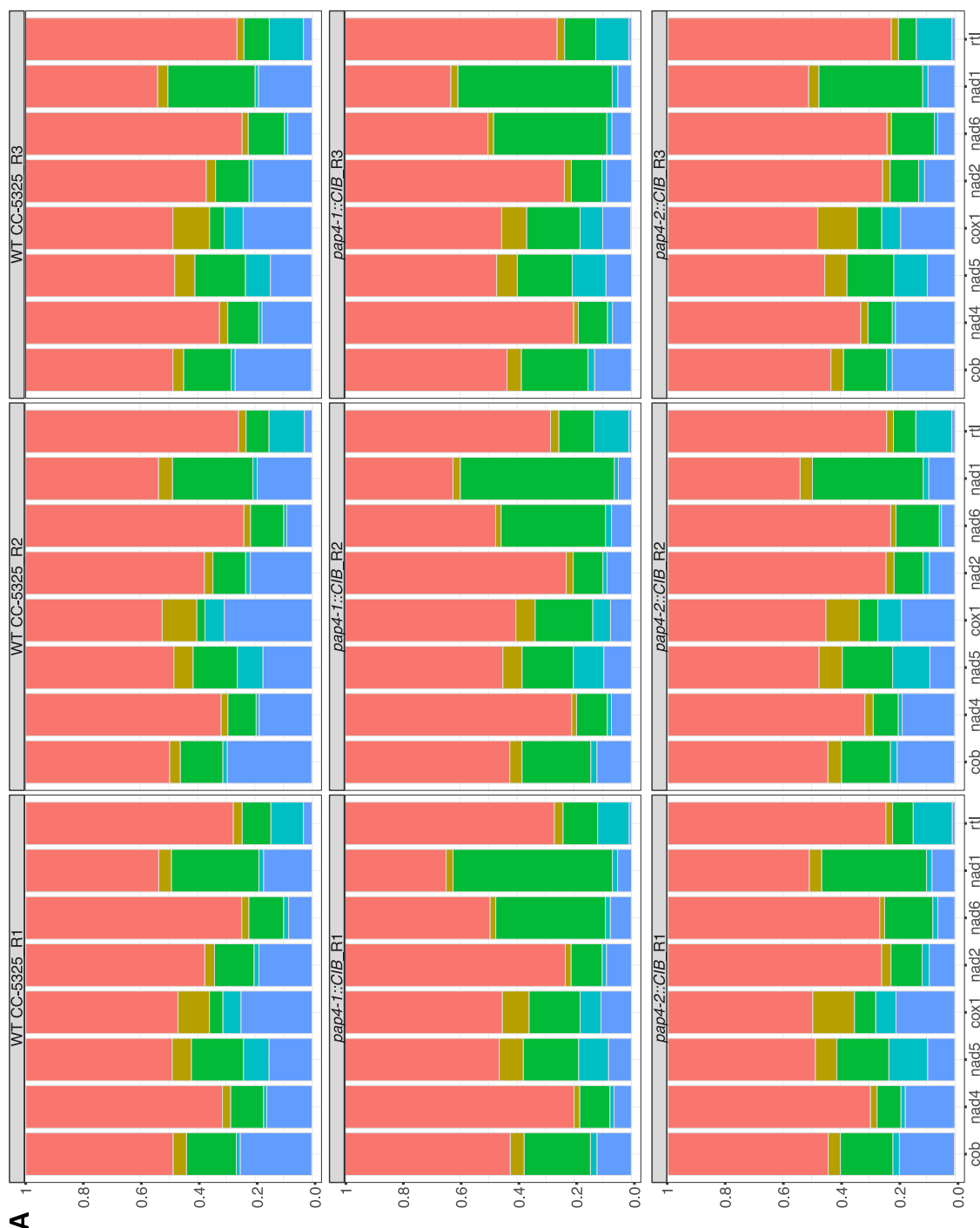
| | | | | | | | | | | | | | |
|--------------------|----|----|----|----|----|----|--------------------|-----|----|----|----|----|----|
| Cre04.g222800.t1.2 | 38 | 4 | 14 | 3 | 10 | 12 | Cre12.g537450.t1.2 | | | | 20 | 21 | 39 |
| Cre06.g281350.t1.2 | 36 | 25 | 27 | 17 | 19 | 23 | Cre02.g080650.t1.2 | 48 | 4 | 18 | 17 | 7 | 38 |
| Cre03.g145867.t1.1 | 36 | 24 | 27 | 17 | 26 | 47 | Cre12.g508500.t1.2 | 7 | | 3 | 19 | 6 | 37 |
| Cre08.g381702.t1.1 | 35 | 22 | 34 | 18 | 26 | 36 | Cre04.g214500.t1.1 | 72 | 80 | 53 | 22 | 23 | 36 |
| Cre14.g619133.t1.1 | 35 | 20 | 32 | 5 | 6 | 11 | Cre11.g467350.t1.2 | 70 | 43 | 50 | 18 | 26 | 36 |
| Cre09.g417150.t1.2 | 35 | 18 | 30 | 24 | 40 | 58 | Cre09.g417150.t1.2 | 35 | 18 | 30 | 25 | 32 | 36 |
| Cre03.g144807.t1.1 | 35 | 17 | 29 | 18 | 15 | 50 | Cre10.g461050.t1.2 | 34 | 3 | 21 | 20 | 4 | 36 |
| Cre02.g080900.t1.2 | 35 | 15 | 22 | 2 | 5 | 13 | Cre06.g294650.t1.2 | 77 | 60 | 65 | 19 | 24 | 34 |
| Cre10.g461050.t1.2 | 34 | 3 | 21 | 6 | 12 | 24 | Cre09.g394436.t1.1 | 13 | 9 | 8 | 26 | 23 | 34 |
| Cre01.g020918.t1.1 | 34 | 20 | 20 | 24 | 35 | 34 | Cre02.g145800.t1.2 | 7 | 7 | 10 | 24 | 35 | 34 |
| Cre01.g038400.t1.2 | 32 | 10 | 40 | 21 | | 19 | Cre15.g638500.t1.2 | 2 | 6 | 3 | 20 | 39 | 34 |
| Cre07.g322000.t1.2 | 32 | 20 | 25 | | 2 | 7 | Cre02.g088000.t1.2 | 1 | | | 25 | 34 | 33 |
| Cre02.g119550.t1.2 | 31 | 9 | 43 | 5 | 18 | 18 | Cre10.g451950.t1.2 | 124 | 81 | 89 | 12 | 26 | 32 |
| Cre12.g514850.t1.2 | 31 | 19 | 33 | 39 | 45 | 55 | Cre17.g743897.t1.1 | 73 | 34 | 68 | 19 | 25 | 32 |
| Cre06.g263450.t1.2 | 31 | 28 | 31 | 20 | 22 | 29 | Cre10.g447100.t1.2 | 51 | 34 | 69 | 16 | 17 | 32 |
| Cre14.g615450.t1.2 | 31 | 18 | 20 | 15 | 8 | 30 | Cre06.g250902.t1.1 | 12 | | 6 | 13 | 3 | 32 |
| Cre09.g388800.t1.2 | 30 | 17 | 17 | 64 | 62 | 72 | Cre13.g566650.t2.1 | 12 | | 5 | 21 | 12 | 32 |
| Cre16.g687350.t1.2 | 29 | 17 | 21 | 5 | 11 | 23 | Cre10.g422600.t1.1 | 8 | 8 | 3 | 23 | 35 | 32 |
| Cre12.g535950.t1.2 | 29 | 22 | 12 | 32 | 26 | 51 | Cre07.g340450.t1.1 | 10 | 1 | 1 | 15 | | 31 |
| Cre12.g522600.t1.2 | 28 | 10 | 33 | 1 | 1 | 4 | Cre12.g519350.t1.2 | 2 | 4 | 1 | 22 | 38 | 31 |
| Cre13.g562550.t1.1 | 28 | 15 | 28 | 26 | | 20 | Cre06.g254400.t1.1 | 65 | 42 | 54 | 8 | 23 | 30 |
| Cre10.g423250.t1.2 | 28 | 19 | 27 | 12 | 19 | 26 | Cre09.g386735.t1.1 | 43 | 20 | 41 | 14 | 17 | 30 |
| Cre06.g254600.t1.1 | 28 | 11 | 16 | 16 | 34 | 45 | Cre12.g514750.t1.2 | 41 | 23 | 34 | 13 | 20 | 30 |
| Cre09.g407700.t1.2 | 27 | 17 | 33 | 14 | 17 | 30 | Cre02.g076350.t1.2 | 27 | 4 | 20 | 15 | 8 | 30 |
| Cre02.g146250.t1.2 | 27 | 22 | 29 | 9 | 4 | 29 | Cre09.g401886.t1.1 | 6 | 3 | 9 | 22 | 29 | 30 |
| Cre16.g677026.t1.1 | 27 | 18 | 24 | 20 | 4 | 26 | Cre03.g146247.t1.1 | | 3 | 1 | 27 | 23 | 30 |
| Cre17.g698000.t1.2 | 27 | 51 | 22 | 13 | 20 | 23 | Cre01.g038400.t1.2 | 32 | 10 | 40 | 33 | 11 | 29 |
| Cre16.g650700.t1.1 | 27 | 8 | 21 | 22 | 29 | 47 | Cre10.g439900.t1.2 | 16 | 1 | 17 | 9 | 4 | 29 |
| Cre02.g076350.t1.2 | 27 | 4 | 20 | 20 | 22 | 29 | Cre10.g459200.t1.2 | 5 | | 15 | 22 | 5 | 29 |
| Cre16.g687950.t1.2 | 26 | 8 | 30 | 33 | 11 | 29 | Cre06.g278750.t1.2 | 5 | 2 | 2 | 23 | 36 | 29 |
| Cre03.g199150.t1.2 | 26 | 13 | 24 | 21 | 12 | 32 | Cre08.g375200.t1.2 | | | | 20 | 22 | 29 |
| Cre12.g485550.t1.2 | 25 | 20 | 25 | 2 | 4 | 12 | Cre10.g444800.t1.2 | | | | 20 | 22 | 29 |
| Cre10.g424100.t1.2 | 25 | 11 | 24 | | 2 | 9 | Cre07.g324200.t1.2 | | | | 15 | 1 | 29 |
| Cre03.g179901.t1.1 | 25 | 12 | 23 | 42 | 55 | 51 | Cre09.g407700.t1.2 | 27 | 17 | 33 | 15 | 11 | 28 |
| Cre03.g213313.t1.1 | 25 | 15 | 20 | 1 | 6 | 11 | Cre10.g450400.t1.2 | 17 | 8 | 8 | 27 | 30 | 28 |
| Cre06.g291150.t1.1 | 25 | 5 | 5 | 9 | 17 | 21 | Cre09.g405850.t1.1 | 8 | 6 | 2 | 14 | 30 | 27 |
| Cre10.g428550.t1.2 | 24 | 16 | 26 | 16 | 16 | 25 | Cre16.g696000.t1.2 | 1 | | | 5 | 1 | 27 |
| Cre06.g267350.t1.2 | 24 | 10 | 21 | 5 | 1 | 3 | Cre08.g358950.t1.1 | | | | 17 | 26 | 27 |
| Cre17.g709850.t1.2 | 23 | 12 | 27 | 22 | 5 | 29 | Cre12.g558600.t1.2 | | | | 5 | | 27 |
| Cre06.g257950.t1.2 | 23 | 12 | 22 | 20 | | 19 | Cre07.g343700.t1.2 | 72 | 59 | 58 | 12 | 19 | 26 |
| Cre12.g523850.t1.2 | 23 | 17 | 20 | 22 | 29 | 30 | Cre12.g495850.t1.2 | 52 | 37 | 52 | 17 | 20 | 26 |
| Cre16.g660800.t1.2 | 23 | 10 | 20 | 7 | 18 | 18 | Cre03.g194350.t1.2 | 21 | | 32 | 20 | 4 | 26 |
| Cre16.g676600.t1.1 | 23 | 9 | 17 | 1 | | 1 | Cre03.g190500.t1.2 | 14 | 2 | 6 | 5 | | 26 |
| Cre09.g415550.t1.2 | 23 | 20 | 16 | 5 | 9 | 10 | Cre10.g424450.t1.2 | 3 | 2 | 4 | 17 | 29 | 26 |
| Cre18.g750047.t1.1 | 23 | 5 | 12 | 6 | 14 | 15 | Cre08.g378050.t1.2 | | 2 | 2 | 15 | 24 | 26 |
| Cre12.g507400.t1.2 | 22 | 12 | 23 | | 1 | 4 | Cre12.g555150.t1.2 | | | | 21 | 25 | 26 |
| Cre02.g147850.t1.2 | 22 | 16 | 21 | 30 | 45 | 48 | Cre12.g500150.t1.1 | 86 | 54 | 78 | 11 | 16 | 25 |
| Cre13.g605650.t1.2 | 22 | 18 | 20 | 6 | 9 | 24 | Cre17.g703700.t1.1 | 58 | 42 | 52 | 16 | 16 | 25 |
| Cre17.g746597.t1.1 | 22 | 2 | 13 | 15 | 7 | 22 | Cre17.g699000.t1.2 | 53 | 22 | 38 | 15 | 16 | 25 |
| Cre03.g182800.t1.1 | 22 | 15 | 9 | 11 | 13 | 24 | Cre16.g677026.t1.1 | 27 | 18 | 24 | 12 | 19 | 25 |
| Cre12.g557600.t1.1 | 21 | 6 | 36 | 24 | 5 | 18 | Cre03.g199150.t1.2 | 26 | 13 | 24 | 21 | 21 | 25 |
| Cre03.g194350.t1.2 | 21 | | 32 | 3 | 3 | 3 | Cre04.g217929.t1.1 | 5 | 7 | 4 | 8 | 19 | 25 |
| Cre01.g044600.t1.2 | 21 | | 23 | 1 | | 5 | Cre15.g635067.t1.1 | 3 | | 7 | 16 | 5 | 25 |
| Cre13.g562501.t1.1 | 21 | 10 | 22 | 14 | 12 | 21 | Cre01.g019850.t1.1 | | | | 9 | 21 | 25 |
| Cre09.g413750.t1.1 | 21 | 10 | 18 | 31 | 6 | 6 | Cre06.g295450.t1.2 | 60 | 35 | 47 | 8 | 4 | 24 |
| Cre13.g581600.t1.2 | 21 | 23 | 17 | | 6 | 5 | Cre06.g252550.t1.1 | 43 | 21 | 32 | 6 | 9 | 24 |
| Cre03.g184250.t1.2 | 21 | 5 | 16 | 7 | 14 | 21 | Cre14.g630859.t1.1 | 38 | 30 | 44 | 6 | 12 | 24 |
| Cre12.g520350.t1.2 | 21 | 12 | 14 | | | | Cre01.g007051.t1.2 | 5 | | 3 | 7 | 8 | 24 |
| Cre12.g510900.t1.2 | 21 | 10 | 13 | 16 | 24 | 24 | Cre14.g621751.t1.1 | 4 | 1 | 7 | 11 | 13 | 24 |
| Cre12.g509750.t1.2 | 21 | 25 | 12 | 9 | 21 | 25 | Cre10.g452450.t1.2 | 3 | | 3 | 13 | | 24 |
| Cre07.g352800.t1.2 | 21 | 3 | 11 | 19 | 14 | 23 | Cre01.g054500.t1.2 | 1 | 2 | 2 | 16 | 24 | 24 |
| Cre11.g467682.t1.1 | 20 | 9 | 35 | 13 | | 24 | Cre05.g234661.t1.1 | 1 | | 1 | 11 | | 24 |
| Cre06.g278227.t1.1 | 20 | 7 | 26 | 3 | 7 | 14 | Cre09.g388750.t1.2 | | | | 16 | 2 | 24 |
| Cre09.g412150.t1.1 | 20 | 14 | 23 | 36 | 55 | 54 | Cre10.g434200.t1.1 | | | | 22 | 23 | 24 |
| Cre17.g743597.t1.1 | 20 | 6 | 16 | 38 | 51 | 65 | Cre01.g005150.t1.1 | 60 | 46 | 57 | 17 | 19 | 23 |
| Cre09.g388550.t1.2 | 20 | 11 | 11 | 37 | 49 | 50 | Cre13.g562550.t1.1 | 28 | 15 | 28 | 5 | 11 | 23 |
| Cre03.g211409.t1.1 | 20 | 6 | 8 | 1 | 2 | 8 | Cre12.g557600.t1.1 | 21 | 6 | 36 | 19 | 14 | 23 |
| Cre11.g476950.t1.1 | 19 | 5 | 28 | 24 | | 12 | Cre03.g199423.t1.1 | 1 | | | 14 | 17 | 23 |
| Cre17.g725350.t1.1 | 19 | 2 | 22 | 25 | 32 | 36 | Cre15.g637761.t1.2 | | | 2 | 13 | 19 | 23 |

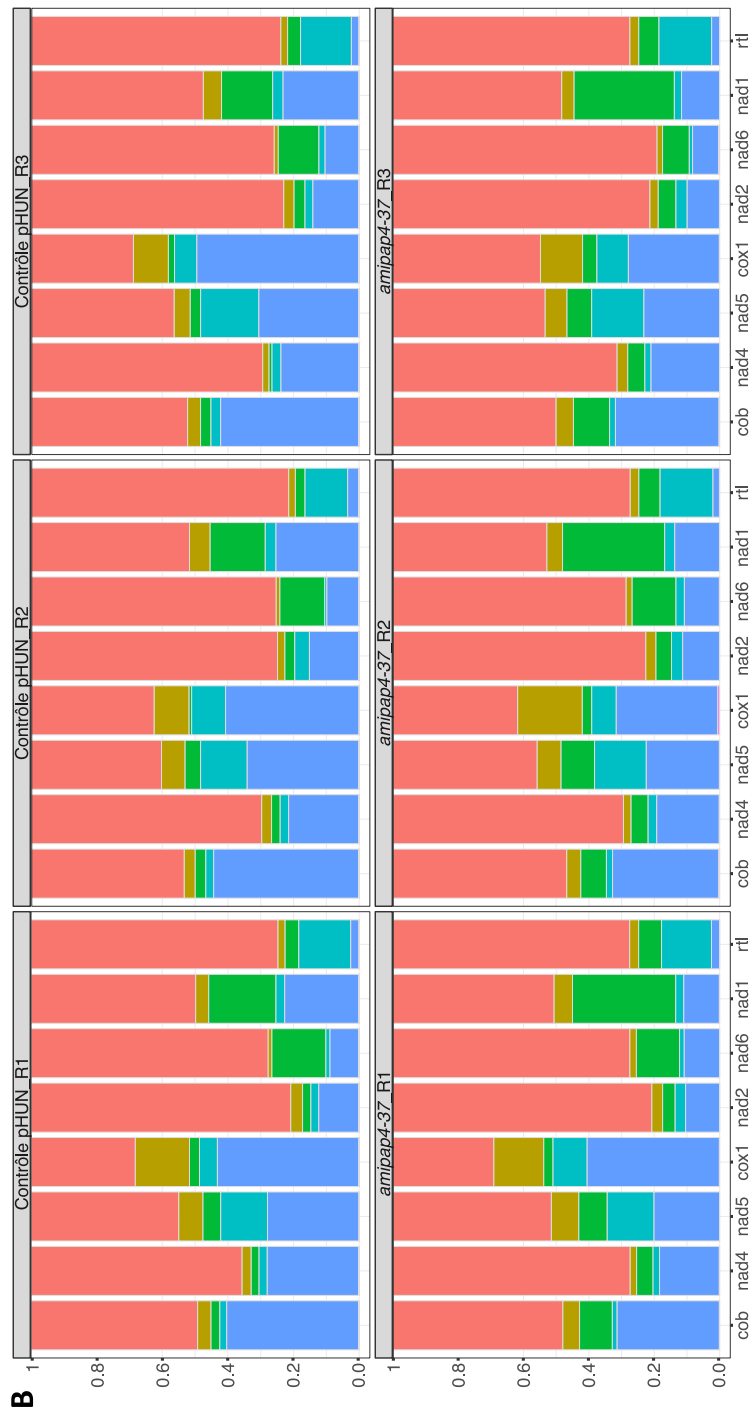
| | | | | | | | | | | | | | | | | | |
|--------------------|----|----|----|----|----|----|--|--|--|--------------------|----|----|----|----|----|----|----|
| Cre07.g313122.t1.1 | 19 | 9 | 19 | | | 1 | | | | Cre06.g267700.t1.2 | | 2 | | | 13 | 20 | 23 |
| Cre05.g241750.t1.2 | 19 | 12 | 18 | 1 | 2 | 5 | | | | Cre03.g199000.t1.2 | 49 | 3 | 18 | 7 | 3 | | 22 |
| Cre06.g296400.t1.2 | 19 | 11 | 17 | 25 | 34 | 33 | | | | Cre06.g264200.t1.2 | 17 | 8 | 12 | 22 | 20 | | 22 |
| Cre03.g191850.t1.1 | 19 | 7 | 16 | 23 | 36 | 29 | | | | Cre17.g707900.t1.2 | 6 | | 5 | 11 | 4 | | 22 |
| Cre02.g099800.t1.1 | 18 | 13 | 26 | 6 | 8 | 12 | | | | Cre06.g296500.t1.2 | 4 | | 9 | 15 | 7 | | 22 |
| Cre12.g549450.t1.2 | 18 | 8 | 23 | 53 | 34 | 45 | | | | Cre17.g721300.t1.2 | | 4 | 2 | 20 | 20 | | 22 |
| Cre07.g355600.t1.1 | 18 | 13 | 18 | 12 | 19 | 25 | | | | Cre02.g094250.t1.1 | | | | 18 | 10 | | 22 |
| Cre16.g689050.t1.1 | 18 | 15 | 16 | | | 1 | | | | Cre16.g677500.t1.2 | 47 | 29 | 39 | 14 | 12 | | 21 |
| Cre01.g013700.t1.2 | 18 | 11 | 15 | 4 | 5 | 8 | | | | Cre16.g687350.t1.2 | 29 | 17 | 21 | 12 | 14 | | 21 |
| Cre01.g010900.t1.2 | 18 | 10 | 14 | 17 | 20 | 26 | | | | Cre10.g423250.t1.2 | 28 | 19 | 27 | 9 | 17 | | 21 |
| Cre01.g009350.t1.2 | 17 | 15 | 25 | 5 | 1 | 27 | | | | Cre16.g676600.t1.1 | 23 | 9 | 17 | 7 | 14 | | 21 |
| Cre02.g096400.t1.2 | 17 | 7 | 24 | 15 | 11 | 28 | | | | Cre11.g481126.t1.1 | 15 | 14 | 12 | 14 | 16 | | 21 |
| Cre06.g278215.t1.1 | 17 | 5 | 17 | 15 | | 13 | | | | Cre01.g010400.t1.2 | 7 | 1 | 15 | 14 | 6 | | 21 |
| Cre06.g264200.t1.2 | 17 | 8 | 12 | 3 | 47 | 14 | | | | Cre06.g284900.t1.2 | 7 | 2 | 9 | 6 | 2 | | 21 |
| Cre10.g450400.t1.2 | 17 | 8 | 8 | 8 | 19 | 25 | | | | Cre07.g327400.t1.1 | 5 | 5 | 5 | 18 | 26 | | 21 |
| Cre09.g403034.t1.1 | 16 | 5 | 23 | 16 | 18 | 19 | | | | Cre11.g468950.t1.2 | 3 | 3 | 1 | 12 | 21 | | 21 |
| Cre05.g241950.t1.2 | 16 | 15 | 22 | 13 | 20 | 30 | | | | Cre03.g196900.t1.2 | 2 | | 2 | 10 | 1 | | 21 |
| Cre03.g180750.t1.2 | 16 | 4 | 18 | 13 | 3 | 32 | | | | Cre11.g467575.t1.1 | | | 2 | 7 | | | 21 |
| Cre03.g145647.t1.1 | 16 | | 18 | 2 | 3 | 5 | | | | Cre16.g691850.t1.2 | | 1 | | 21 | 24 | | 21 |
| Cre10.g439900.t1.2 | 16 | 1 | 17 | 9 | 9 | 13 | | | | Cre11.g467668.t1.1 | | | | 13 | 27 | | 21 |
| Cre12.g508550.t1.1 | 16 | 7 | 16 | 13 | 19 | 23 | | | | Cre18.g750047.t1.1 | 23 | 5 | 12 | 36 | 9 | | 20 |
| Cre16.g690431.t1.1 | 16 | 11 | 15 | 13 | 5 | 4 | | | | Cre01.g044600.t1.2 | 21 | | 23 | 26 | | | 20 |
| Cre07.g350500.t1.2 | 16 | 9 | 14 | 8 | 9 | 10 | | | | Cre01.g051900.t1.2 | 4 | 9 | 5 | 15 | 21 | | 20 |
| Cre02.g081450.t1.2 | 16 | 5 | 14 | | | 1 | | | | Cre09.g402775.t1.1 | 2 | | | 8 | 5 | | 20 |
| Cre02.g106750.t1.2 | 16 | 3 | 14 | | 3 | 5 | | | | Cre12.g557250.t1.2 | | | | 4 | | | 20 |
| Cre05.g239100.t1.1 | 16 | 6 | 9 | 1 | 1 | 5 | | | | Cre17.g727300.t1.2 | 45 | 35 | 39 | 4 | 15 | | 19 |
| Cre03.g211745.t1.1 | 16 | 4 | 9 | 16 | 5 | 25 | | | | Cre17.g697650.t1.1 | 9 | | 3 | 3 | | | 19 |
| Cre16.g695100.t1.1 | 16 | 6 | 8 | 15 | 1 | 29 | | | | Cre12.g525150.t1.2 | 8 | | 16 | 21 | | | 19 |
| Cre02.g081050.t1.2 | 15 | 2 | 17 | | 1 | 2 | | | | Cre17.g745697.t1.1 | 4 | | 8 | 20 | | | 19 |
| Cre14.g612633.t1.1 | 15 | 1 | 17 | 11 | | 24 | | | | Cre07.g319100.t1.1 | 2 | 3 | 3 | 5 | 14 | | 19 |
| Cre01.g028600.t1.2 | 15 | 5 | 16 | | 6 | 10 | | | | Cre03.g164600.t1.2 | 1 | | 4 | 13 | | | 19 |
| Cre11.g481126.t1.1 | 15 | 14 | 12 | | 2 | 2 | | | | Cre19.g750847.t1.1 | 1 | 2 | 1 | 16 | 18 | | 19 |
| Cre05.g236050.t1.1 | 15 | | 12 | 12 | 14 | 21 | | | | Cre11.g467767.t1.1 | 1 | 5 | | 12 | 21 | | 19 |
| Cre03.g171300.t1.2 | 15 | 9 | 11 | 7 | 9 | 17 | | | | Cre01.g049500.t1.2 | 1 | 2 | | 21 | 22 | | 19 |
| Cre09.g386650.t1.2 | 14 | 23 | 21 | 15 | | 11 | | | | Cre06.g295400.t1.2 | 1 | 1 | | 17 | 10 | | 19 |
| Cre01.g002050.t1.1 | 14 | 13 | 20 | 9 | 11 | 10 | | | | Cre05.g242350.t1.2 | | | | 17 | 3 | | 19 |
| Cre11.g467553.t1.1 | 14 | 7 | 13 | | | 6 | | | | Cre13.g568400.t1.2 | | | | 7 | 16 | | 19 |
| Cre06.g278098.t1.1 | 14 | 2 | 10 | | | 15 | | | | Cre09.g390604.t1.1 | | | | 6 | 7 | | 19 |
| Cre05.g247400.t1.2 | 14 | 4 | 8 | 5 | 8 | 16 | | | | Cre06.g251400.t1.2 | 52 | 38 | 46 | 5 | 18 | | 18 |
| Cre03.g190500.t1.2 | 14 | 2 | 6 | 9 | 11 | 15 | | | | Cre07.g337650.t1.2 | 44 | 26 | 33 | 7 | 18 | | 18 |
| Cre11.g467702.t1.1 | 13 | 3 | 19 | 19 | 6 | 37 | | | | Cre17.g725350.t1.1 | 19 | 2 | 22 | 24 | 5 | | 18 |
| Cre01.g035350.t1.2 | 13 | 14 | 17 | 16 | 17 | 32 | | | | Cre03.g180750.t1.2 | 16 | 4 | 18 | 5 | 2 | | 18 |
| Cre06.g250100.t1.2 | 13 | 11 | 17 | 10 | 10 | 14 | | | | Cre12.g496750.t1.2 | 2 | | 2 | 11 | 15 | | 18 |
| Cre12.g545177.t1.1 | 13 | 3 | 16 | 26 | 23 | 34 | | | | Cre10.g439850.t1.2 | 1 | | | 12 | 3 | | 18 |
| Cre09.g388282.t1.1 | 13 | 1 | 16 | 1 | 1 | 4 | | | | Cre15.g641200.t1.2 | | 2 | | 12 | 18 | | 18 |
| Cre08.g385900.t1.2 | 13 | 1 | 15 | | 1 | 9 | | | | Cre08.g370450.t1.2 | 50 | 18 | 80 | 11 | 5 | | 17 |
| Cre09.g393543.t1.1 | 13 | 13 | 14 | 11 | 5 | 17 | | | | Cre10.g456050.t1.2 | 12 | | 9 | 13 | 2 | | 17 |
| Cre12.g500500.t1.1 | 13 | 7 | 13 | 16 | 2 | 24 | | | | Cre16.g675850.t1.2 | 11 | 4 | 9 | 7 | 9 | | 17 |
| Cre02.g090850.t1.1 | 13 | 5 | 13 | | | | | | | Cre01.g036850.t1.2 | 10 | 3 | 5 | 6 | 9 | | 17 |
| Cre07.g348850.t1.2 | 13 | 8 | 12 | 1 | | 4 | | | | Cre17.g734500.t1.2 | 6 | | 8 | 11 | | | 17 |
| Cre12.g530500.t1.2 | 13 | 4 | 12 | 20 | 39 | 34 | | | | Cre05.g230850.t1.1 | 6 | | 6 | 8 | | | 17 |
| Cre05.g248500.t1.2 | 13 | 4 | 12 | 17 | 26 | 27 | | | | Cre06.g304350.t1.2 | 4 | 5 | 5 | 10 | 19 | | 17 |
| Cre16.g677950.t1.1 | 13 | 6 | 11 | 8 | | 17 | | | | Cre09.g390763.t1.1 | 2 | | | 11 | 4 | | 17 |
| Cre12.g534350.t1.2 | 13 | 3 | 11 | 20 | 38 | 47 | | | | Cre02.g094300.t1.2 | | | | 8 | 12 | | 17 |
| Cre02.g114250.t1.2 | 13 | 3 | 11 | | 3 | 8 | | | | Cre09.g402300.t1.2 | | | | 12 | 9 | | 17 |
| Cre09.g394436.t1.1 | 13 | 9 | 8 | 12 | 5 | 15 | | | | Cre01.g020350.t1.2 | | | | 13 | 17 | | 17 |
| Cre05.g232002.t1.1 | 13 | 10 | 6 | 10 | | 2 | | | | Cre16.g677950.t1.1 | 13 | 6 | 11 | 5 | 8 | | 16 |
| Cre01.g071662.t1.1 | 13 | 9 | 6 | 17 | 3 | 19 | | | | Cre16.g679500.t1.2 | 8 | | 4 | 9 | 17 | | 16 |
| Cre06.g261150.t1.2 | 13 | 3 | 6 | 3 | 3 | 5 | | | | Cre13.g568800.t1.2 | 5 | 3 | 9 | 10 | 17 | | 16 |
| Cre09.g392500.t1.2 | 13 | | 6 | 22 | 38 | 31 | | | | Cre01.g002500.t1.2 | 2 | | | 18 | | | 16 |
| Cre12.g511700.t1.2 | 13 | 2 | 5 | | 32 | 2 | | | | Cre16.g664600.t1.2 | | 1 | 1 | 3 | 21 | | 16 |
| Cre07.g332150.t1.2 | 12 | 4 | 20 | 1 | 2 | 7 | | | | Cre06.g259150.t1.2 | 57 | 51 | 43 | 7 | 11 | | 15 |
| Cre09.g398215.t1.1 | 12 | | 19 | 5 | | 26 | | | | Cre16.g660800.t1.2 | 23 | 10 | 20 | 9 | 11 | | 15 |
| Cre01.g015750.t1.2 | 12 | 11 | 15 | 17 | 29 | 26 | | | | Cre01.g002050.t1.1 | 14 | 13 | 20 | 6 | 14 | | 15 |
| Cre03.g154550.t1.1 | 12 | 11 | 13 | | | 1 | | | | Cre03.g194850.t1.2 | 7 | 1 | 6 | 5 | 4 | | 15 |
| Cre17.g702800.t1.2 | 12 | 7 | 10 | 3 | 2 | 3 | | | | Cre03.g182050.t1.2 | 6 | | 2 | 12 | 5 | | 15 |
| Cre10.g456050.t1.2 | 12 | | 9 | | 1 | | | | | Cre10.g439600.t1.1 | 1 | | | 2 | 1 | | 15 |
| Cre04.g217904.t1.1 | 12 | 2 | 8 | 2 | 1 | 15 | | | | Cre12.g560450.t1.1 | | | | | | | 15 |
| Cre06.g250902.t1.1 | 12 | | 6 | 48 | 63 | 84 | | | | Cre02.g106400.t1.1 | | | | 7 | 1 | | 15 |

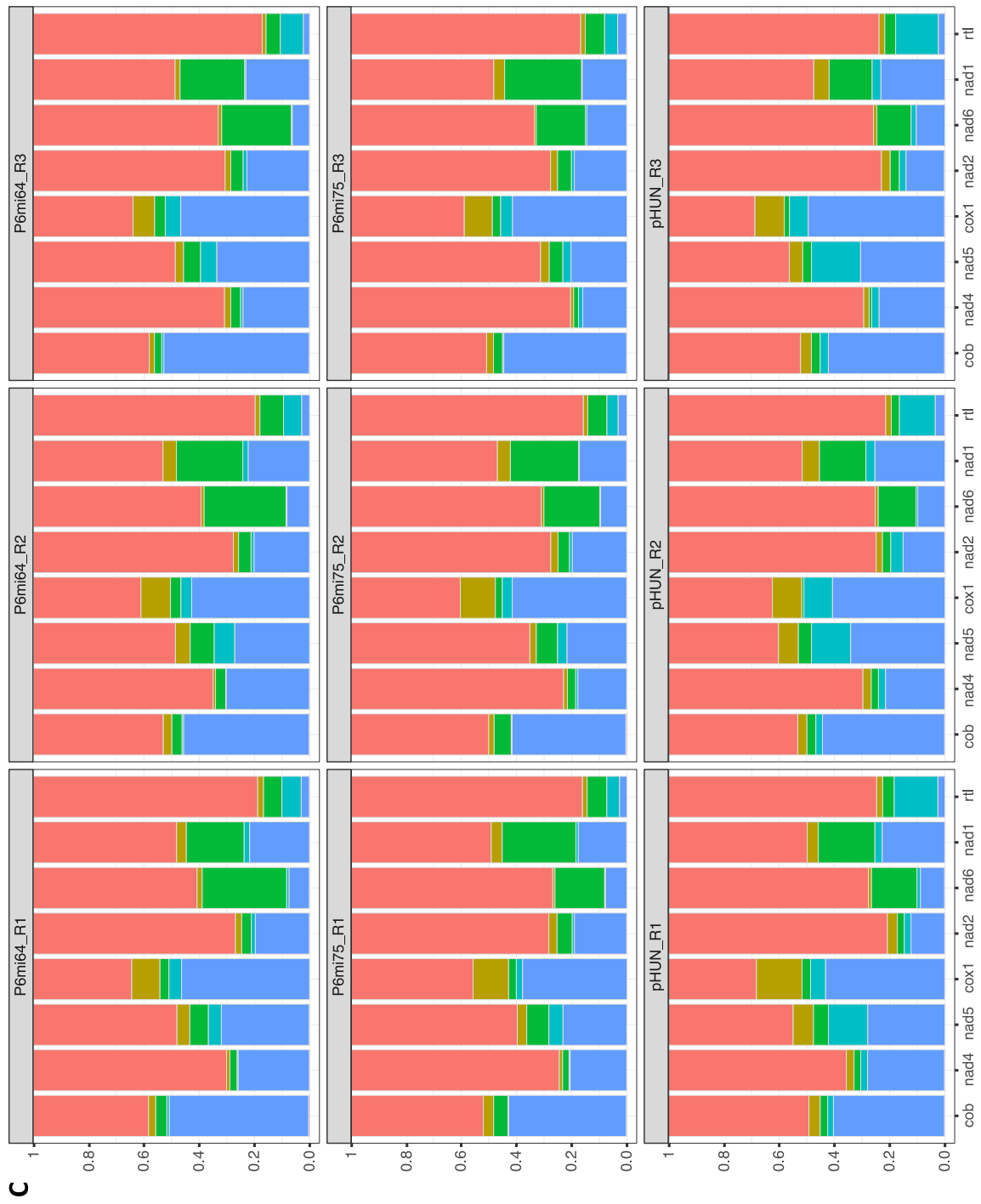
VIII. Données de 3'RACE-seq

1. Données des triplicats biologiques

Analyses des expériences de 3'RACE-seq des triplicats biologiques pour chaque souche en fonction de la catégorie des queues. En abscisse sont indiquées les proportions des différentes catégories de queues selon les gènes pour chaque souche. **A** : mutants de PAP4 CLiP ; **B** : mutant amiRNA de PAP4 ; **C** : mutants amiRNA de PAP6. **Corail** : pas de queues ; **Jaune** : queues riches en A, U et C ; **Vert** : queues riches en A ; **Bleu clair** : queues riches en U ; **Bleu foncé** : queues riches en C ; **Fuchsia** : autres.



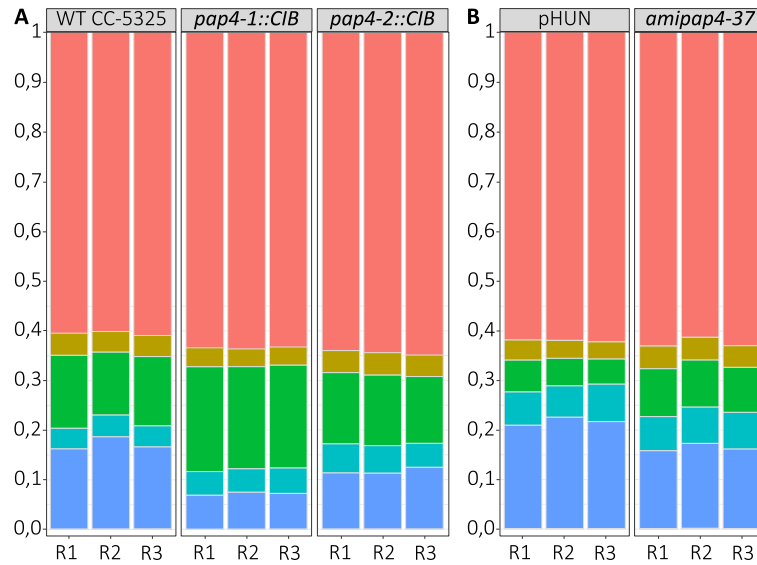




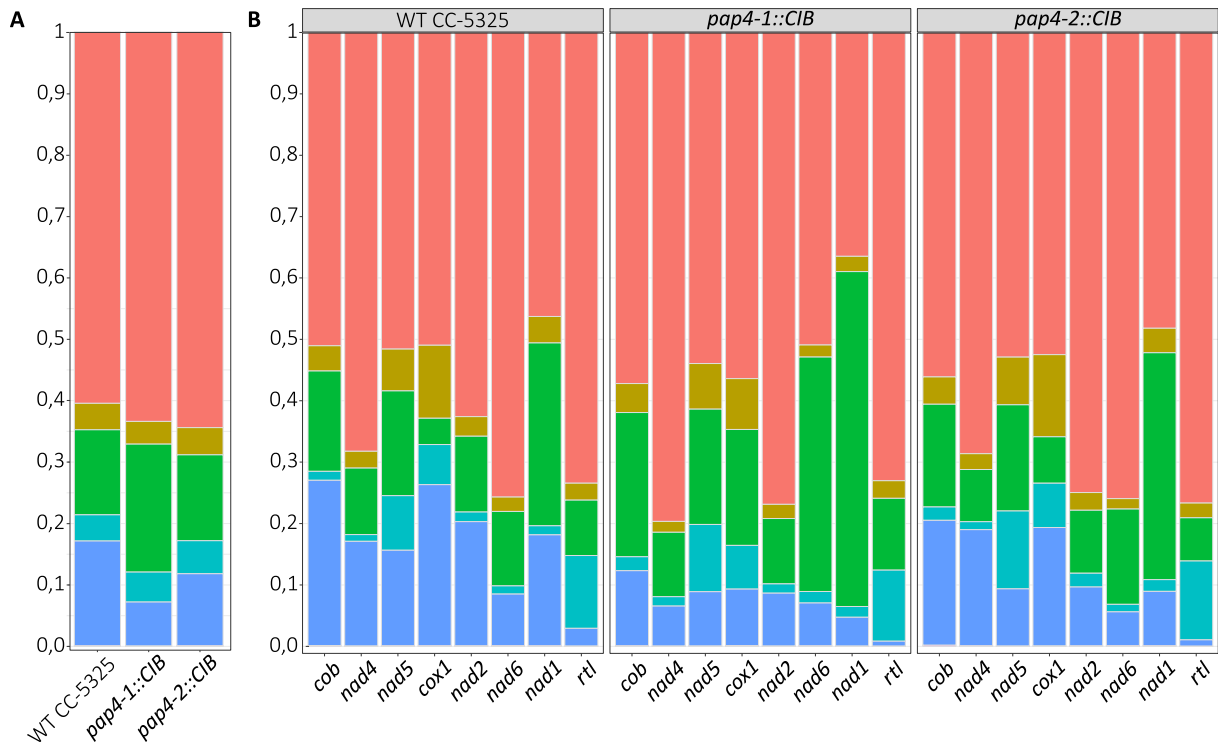
2. Dé-zoom des graphiques présentés dans les Résultats

Analyses des expériences de 3'RACE-seq des mutants amiRNA pour PAP6 en fonction de la catégorie des queues. Résultats cumulés des 3 réplicats biologiques. En abscisse sont indiquées les proportions des différentes catégories de queues, selon les souches (A et D) et selon les gènes pour chaque souche (B, C et E). **Corail** : pas de queues ; **Jaune** : queues riches en A, U et C ; **Vert** : queues riches en A ; **Bleu clair** : queues riches en U ; **Bleu foncé** : queues riches en C.

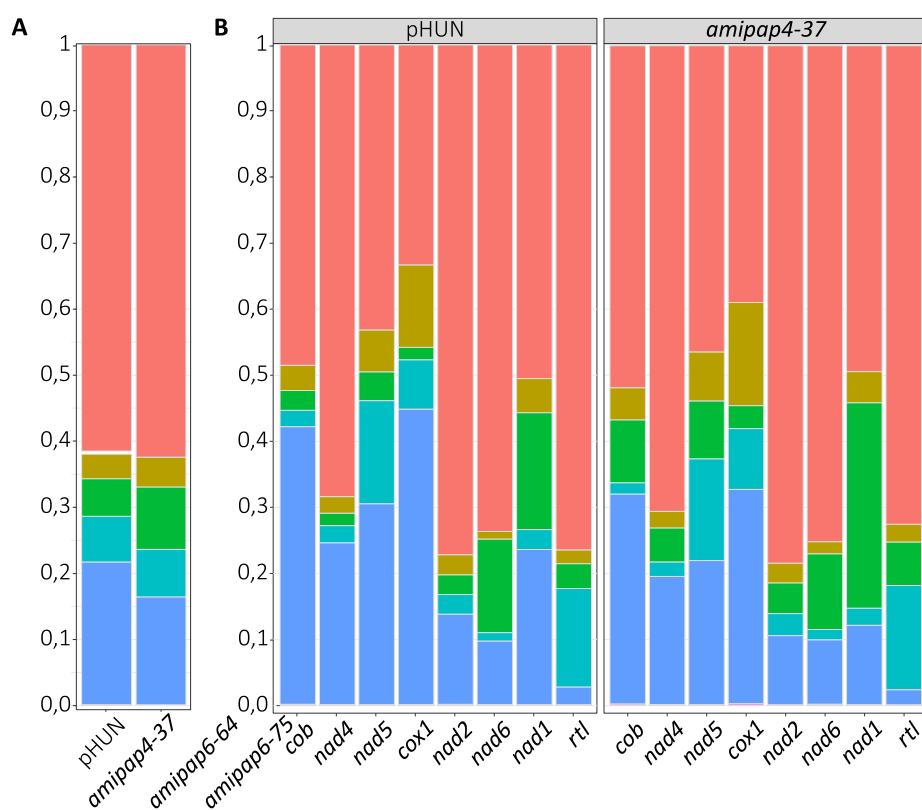
A : Figure R-29 :



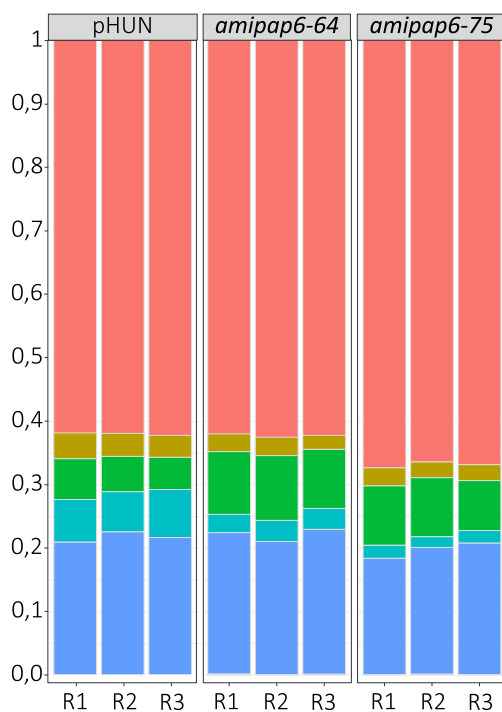
B : Figure R-30 :



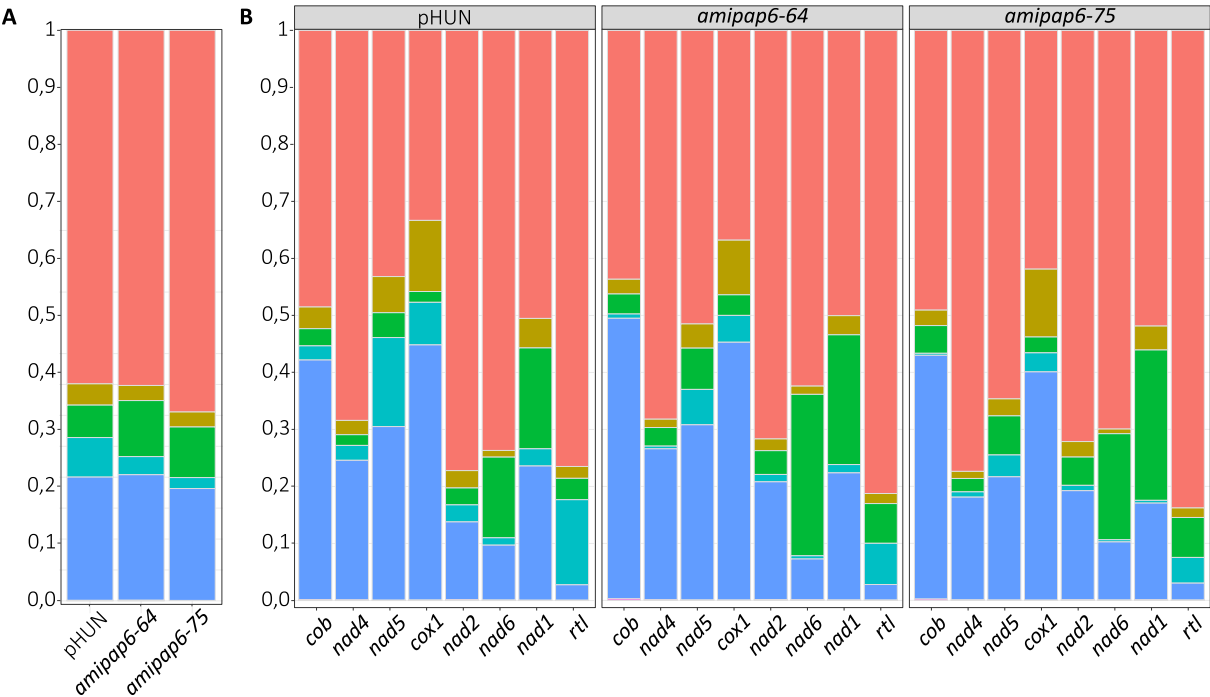
C : Figure R-33 :



D : Figure R-45 :



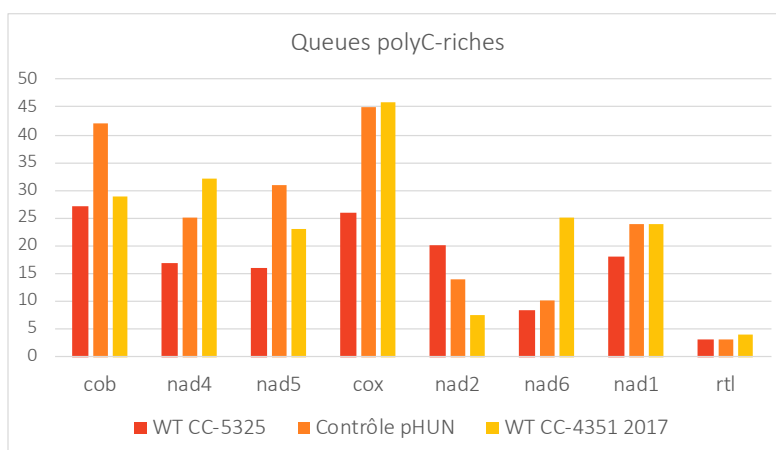
E : Figure R-46 :



IX. Comparaison des données de 3'RACE et 3'RACE-seq

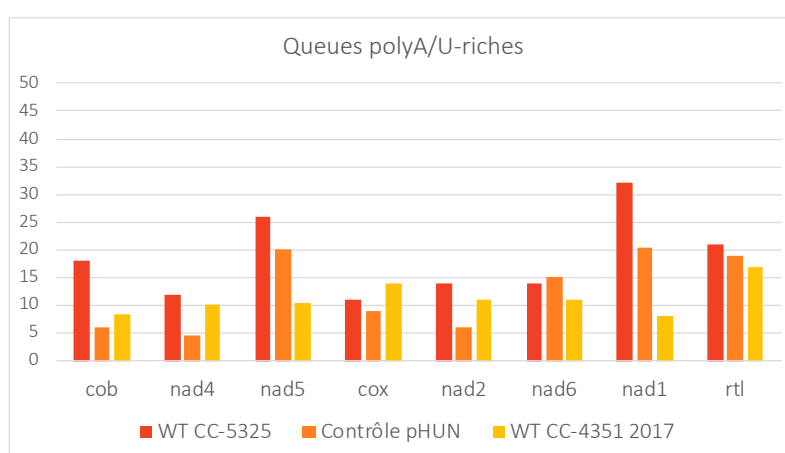
Graphique et tableau des proportions de queues polyC-riche et polyA/U-riche ajoutées à chaque transcrite mitochondrial dans 3 souches : WT CC-5325 (résultats présentés dans cette thèse), Contrôle pHUN (résultats présentés dans cette thèse) et WT CC-4351 (résultats publiés en 2017 (Salinas-Giegé *et al.*, 2017)). Pour les souches présentées dans cette thèse, les proportions de queues polyA-riche et polyU-riche ont été additionnées.

Queues polyC-riche



| | <i>cob</i> | <i>nad4</i> | <i>nad5</i> | <i>cox</i> | <i>nad2</i> | <i>nad6</i> | <i>nad1</i> | <i>rtl</i> |
|-----------------|------------|-------------|-------------|------------|-------------|-------------|-------------|------------|
| WT CC-5325 | 27 | 17 | 16 | 26 | 20 | 8,5 | 18 | 3 |
| Contrôle pHUN | 42 | 25 | 31 | 45 | 14 | 10 | 24 | 3 |
| WT CC-4351 2017 | 29 | 32 | 23 | 46 | 7,5 | 25 | 24 | 4 |

Queues polyA/U-riche



| | <i>cob</i> | <i>nad4</i> | <i>nad5</i> | <i>cox</i> | <i>nad2</i> | <i>nad6</i> | <i>nad1</i> | <i>rtl</i> |
|-----------------|------------|-------------|-------------|------------|-------------|-------------|-------------|------------|
| WT CC-5325 | 18 | 12 | 26 | 11 | 14 | 14 | 32 | 21 |
| Contrôle pHUN | 6 | 4,5 | 20 | 9 | 6 | 15 | 20,5 | 19 |
| WT CC-4351 2017 | 8,5 | 10 | 10,5 | 14 | 11 | 11 | 8 | 17 |

X. Pourcentages de différences – 3'RACE-seq

Tableaux de pourcentages de différences entre la chaque souche mutante et la souche de référence qui lui est associée. Colonne Total : données présentées en Figures R30A, R33A et R46A. *cob* à *rtl* : données présentées en Figures R30B à E, R33 B à E et R46 B à E.

A : mutants de PAP4 CLiP ; **B** : mutant amiRNA de PAP4 ; **C** : mutants amiRNA de PAP6.

Blanc : jusqu'à 10% d'augmentation ou de diminution dans le mutant vis-à-vis de la souche de référence ; Vert pâle : entre 10 et 20% d'augmentation ; Vert : plus de 20% d'augmentation ; Rouge pâle : entre 10 et 20% de diminution ; Rouge : plus de 20% de diminution.

| A | | Total | <i>cob</i> | <i>nad4</i> | <i>nad5</i> | <i>cox</i> | <i>nad2</i> | <i>nad6</i> | <i>nad1</i> | <i>rtl</i> | % |
|------------------|--------------------|-------|------------|-------------|-------------|------------|-------------|-------------|-------------|------------|---|
| C | WT CC-5325 | 17 | 27 | 17 | 16 | 26 | 20 | 8,5 | 18 | 3 | % |
| | <i>pap4-1::CIB</i> | 7 | 12 | 6,5 | 9 | 9,5 | 8,5 | 7 | 4,5 | 1 | % |
| | Différence mt/WT | 41% | 44% | 38% | 56% | 37% | 43% | 82% | 25% | 33% | |
| | <i>pap4-2::CIB</i> | 12 | 20,5 | 19 | 9,5 | 19,5 | 9,5 | 5,5 | 9 | 1 | % |
| Différence mt/WT | 71% | 76% | 112% | 59% | 75% | 48% | 65% | 50% | 33% | | |
| A | WT CC-5325 | 13,5 | 16,5 | 11 | 17 | 4 | 12 | 12 | 30 | 9 | % |
| | <i>pap4-1::CIB</i> | 21 | 23 | 10,5 | 18,5 | 18,5 | 11 | 38 | 54,5 | 11,5 | % |
| | Différence mt/WT | 156% | 139% | 95% | 109% | 463% | 92% | 317% | 182% | 128% | |
| | <i>pap4-2::CIB</i> | 14 | 16,5 | 9 | 17 | 7 | 10 | 15 | 37 | 7 | % |
| Différence mt/WT | 104% | 100% | 82% | 100% | 175% | 83% | 125% | 123% | 78% | | |
| U | WT CC-5325 | 4,5 | 1,5 | 1 | 9 | 7 | 2 | 2 | 2 | 12 | % |
| | <i>pap4-1::CIB</i> | 5 | 2,5 | 1,5 | 11 | 7 | 1 | 2 | 2,5 | 11,5 | % |
| | Différence mt/WT | 111% | 167% | 150% | 122% | 100% | 50% | 100% | 125% | 96% | |
| | <i>pap4-2::CIB</i> | 5 | 2 | 1,5 | 12,5 | 7 | 2 | 1,5 | 2 | 13 | % |
| Différence mt/WT | 111% | 133% | 150% | 139% | 100% | 100% | 75% | 100% | 108% | | |

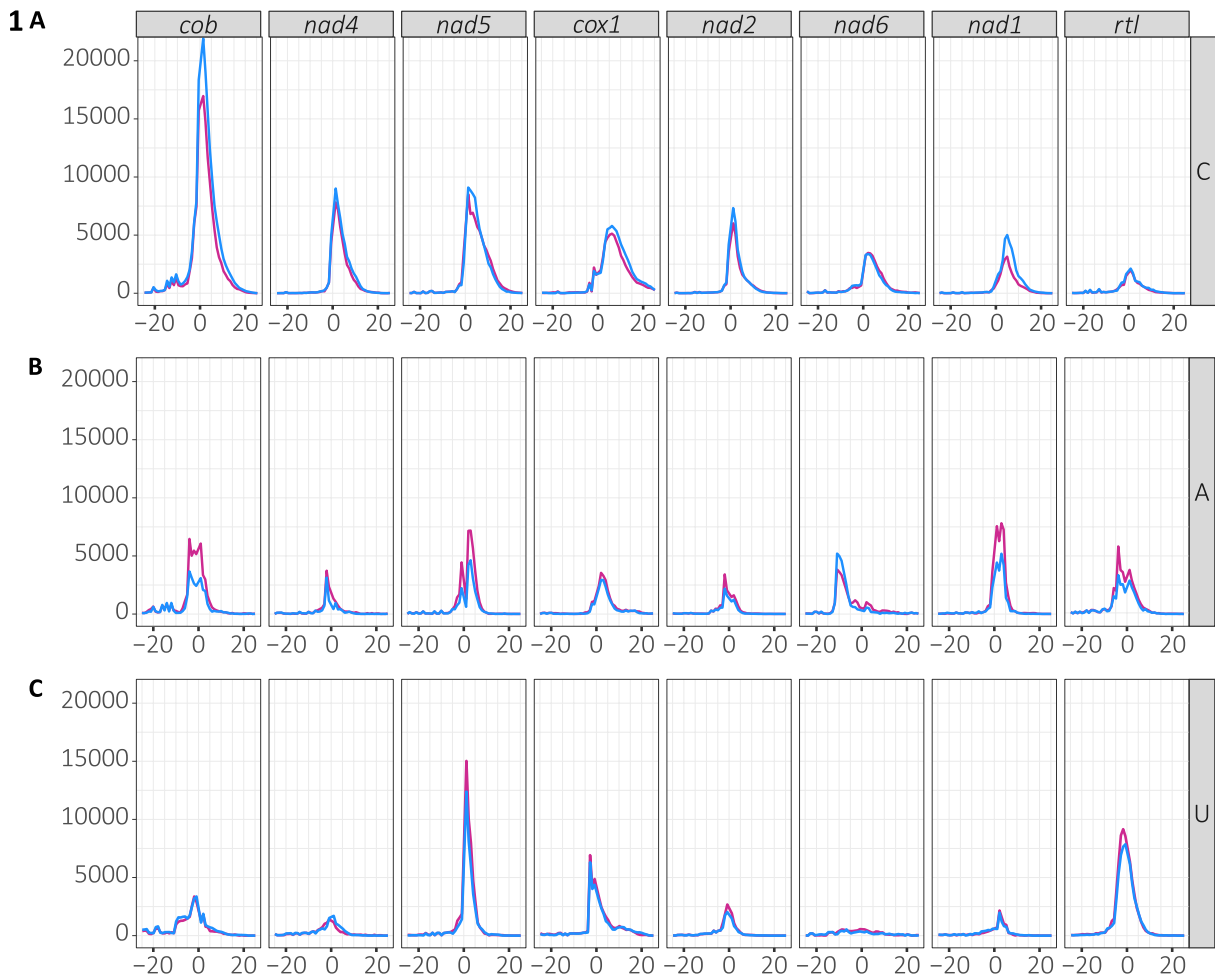
| B | | Total | <i>cob</i> | <i>nad4</i> | <i>nad5</i> | <i>cox</i> | <i>nad2</i> | <i>nad6</i> | <i>nad1</i> | <i>rtl</i> | % |
|---|--------------------|-------|------------|-------------|-------------|------------|-------------|-------------|-------------|------------|---|
| C | Contrôle pHUN | 22 | 42 | 25 | 30 | 45 | 14 | 10 | 23,5 | 3 | % |
| | <i>amipap4-37</i> | 16,5 | 32 | 20 | 22 | 33 | 11 | 10 | 12 | 2 | % |
| | Différence mt/ctrl | 75% | 76% | 80% | 73% | 73% | 79% | 100% | 51% | 67% | |
| A | Contrôle pHUN | 6,5 | 3 | 2 | 4,5 | 2 | 3 | 14 | 17 | 3,5 | % |
| | <i>amipap4-37</i> | 9 | 9,5 | 5 | 9 | 4 | 5 | 11 | 31 | 6,5 | % |
| | Différence mt/ctrl | 138% | 317% | 250% | 200% | 200% | 167% | 79% | 182% | 186% | |
| U | Contrôle pHUN | 7 | 3 | 2 | 16 | 7 | 3 | 1 | 3 | 15 | % |
| | <i>amipap4-37</i> | 7,5 | 2 | 2 | 15 | 9 | 3 | 2 | 2 | 16 | % |
| | Différence mt/ctrl | 107% | 67% | 100% | 94% | 129% | 100% | 200% | 67% | 107% | |

| C | | Total | <i>cob</i> | <i>nad4</i> | <i>nad5</i> | <i>cox</i> | <i>nad2</i> | <i>nad6</i> | <i>nad1</i> | <i>rtl</i> | % |
|------------------|-------------------|-------|------------|-------------|-------------|------------|-------------|-------------|-------------|------------|---|
| C | Contrôle pHUN | 22 | 42 | 25 | 30 | 45 | 14 | 10 | 23,5 | 3 | % |
| | <i>amipap6-64</i> | 22 | 49,5 | 26,5 | 31,5 | 45 | 21 | 7 | 22 | 3 | % |
| | Différence mt/WT | 100% | 118% | 106% | 105% | 100% | 150% | 70% | 94% | 100% | |
| | <i>amipap6-75</i> | 20 | 43 | 19 | 22 | 40 | 19 | 10 | 17 | 3 | % |
| Différence mt/WT | 91% | 102% | 76% | 73% | 89% | 136% | 100% | 72% | 100% | | |
| A | Contrôle pHUN | 6,5 | 3 | 2 | 4,5 | 2 | 3 | 14 | 17 | 3,5 | % |
| | <i>amipap6-64</i> | 10 | 4 | 3 | 7 | 3,5 | 4 | 18 | 22,5 | 7 | % |
| | Différence mt/WT | 154% | 133% | 150% | 156% | 175% | 133% | 129% | 132% | 200% | |
| | <i>amipap6-75</i> | 9 | 5 | 2,5 | 7 | 2,5 | 5 | 18 | 26 | 7 | % |
| Différence mt/WT | 138% | 167% | 125% | 156% | 125% | 167% | 129% | 153% | 200% | | |
| U | Contrôle pHUN | 7 | 3 | 2 | 16 | 7 | 3 | 1 | 3 | 15 | % |
| | <i>amipap6-64</i> | 3 | 1 | 0,5 | 6 | 5 | 1 | 1 | 2 | 7 | % |
| | Différence mt/WT | 43% | 33% | 25% | 38% | 71% | 33% | 100% | 67% | 47% | |
| | <i>amipap6-75</i> | 2 | 0,5 | 1 | 3,5 | 3,5 | 1 | 0,5 | 0,5 | 4,5 | % |
| Différence mt/WT | 29% | 17% | 50% | 22% | 50% | 33% | 50% | 17% | 30% | | |

XI. Distribution des nt selon la position – 3'RACE-seq

Analyses des expériences de 3'RACE-seq des mutants amiRNA de PAP4 (mutant amiRNA (1) puis mutants CLiP (2)) et de PAP6 (3). Résultats cumulés des 3 réplicats biologiques. Graphiques du nombre de nt C (A), A(B) et U (C) trouvés dans les queues des différents transcrits, en fonction de leur position vis-à-vis de la fin de la 3'UTR (position 0, trait rouge pointillé). La légende des courbes et mutants associés sont au-dessus de chaque figure.

Bleu : contrôle pHUN ; **Rose** : *amipap4-37*.



Herrade MEICHEL CLAUDEL

Maturation des ARN messagers mitochondriaux et système traductionnel chez l'algue verte *Chlamydomonas reinhardtii*

Résumé

Les mitochondries sont des organites semi-autonomes, possédant leur propre machinerie d'expression des gènes. Mon travail de thèse mené sur l'algue verte *Chlamydomonas reinhardtii* avait pour but de mieux comprendre le processus de maturation des ARNm mitochondriaux qui possèdent des caractéristiques particulières. En effet, il n'y a pas de séquence 5'UTR et des queues polyC-riches et polyU-riches sont ajoutées post-transcriptionnellement. Des mutants de 2 Nucleotidyl transférases mitochondriales, PAP4 et PAP6, ont été obtenus. L'analyse des transcrits mitochondriaux et de la composition de leurs queues ajoutées post-transcriptionnellement suggèrent que PAP4 soit l'enzyme ajoutant les queues polyC-riches en 3' des ARNm, permettant probablement leur stabilisation. La protéine PAP6 quant à elle ajouterait des queues polyU-riches, qui semblent être un signal de dégradation des ARNm. Par ailleurs, dans le cadre de la caractérisation de la structure et la composition du ribosome mitochondrial, l'analyse de mutants de protéines mitoribosomales spécifiques de *C. reinhardtii* supportent leur rôle dans l'assemblage et de la stabilisation des fragments d'ARNr.

Mots clés : mitochondrie, maturation ARN, PAP, polycytidylation, polyuridylation, traduction, mitoribosome

Abstract

Mitochondria are semi-autonomous organelles with their own gene expression machinery. The aim of my thesis work on the green alga *Chlamydomonas reinhardtii* was to gain a better understanding of the maturation process of mitochondrial mRNAs, which have specific characteristics. In fact, there is no 5'UTR sequence and polyC-rich and polyU-rich tails are added post-transcriptionally. Mutants of 2 mitochondrial nucleotidyl transferases, PAP4 and PAP6, were obtained. Analysis of the mitochondrial transcripts and the composition of their post-transcriptionally added tails suggest that PAP4 is the enzyme adding the polyC-rich tails to the 3' of the mRNAs, probably enabling their stabilisation. The PAP6 protein is thought to add polyU-rich tails, which appear to be a signal for mRNA degradation. Furthermore, as part of the characterisation of the structure and composition of the mitochondrial ribosome, analysis of mutants of mitoribosomal proteins specific to *C. reinhardtii* support their role in the assembly and stabilisation of rRNA fragments.

Keywords : mitochondria, RNA maturation, PAP, polycytidylation, polyuridylation, translation, mitoribosome