



Identification du mécanisme impliqué dans la formation de délétions de l'ADN mitochondrial: cas de la "Common Deletion"

Armél Raffour-Millet

► To cite this version:

Armél Raffour-Millet. Identification du mécanisme impliqué dans la formation de délétions de l'ADN mitochondrial: cas de la "Common Deletion". Biologie moléculaire. Museum national d'histoire naturelle - MNHN PARIS, 2017. Français. NNT : 2017MNHN0017 . tel-01843610

HAL Id: tel-01843610

<https://theses.hal.science/tel-01843610v1>

Submitted on 18 Jul 2018

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.



MUSEUM NATIONAL D'HISTOIRE NATURELLE
Ecole Doctorale Sciences de la Nature et de l'Homme – ED 227

Année 2017

N°attribué par la bibliothèque

██

THESE

Pour obtenir le grade de

DOCTEUR DU MUSEUM NATIONAL D'HISTOIRE NATURELLE

Spécialité : Sciences Moléculaires et Cellulaires de la Biologie

Présentée et soutenue publiquement par

Armél RAFFOUR-MILLET

Le 11 Septembre 2017

**Identification du mécanisme impliqué dans la
formation de délétions de l'ADN mitochondrial :
cas de la « Common Deletion »**

Sous la direction de : **Mme Emmanuelle DELAGOUTTE et Mme Erika BRUNET**

JURY :

Mme Delagoutte Emmanuelle	Chargée de Recherche CNRS, Muséum national d'histoire naturelle, Paris (075)	Directeur de Thèse
Mme Brunet Erika	Chargée de Recherche INSERM, Institut Imagine, Paris (075)	Directeur de Thèse
Mme Lombès Anne	Directrice de Recherche INSERM, Institut Cochin, Paris (075)	Rapporteur
M Hyrien Olivier	Directeur de Recherche CNRS, IBENS, Paris (075)	Rapporteur
Mme Mayer Claudine	Professeure, Université Paris-Diderot, Institut Pasteur, Paris (075)	Examinateur
M Strauss François	Directeur de Recherche CNRS, Muséum national d'histoire naturelle, Paris (075)	Examinateur



MUSEUM NATIONAL D'HISTOIRE NATURELLE
Ecole Doctorale Sciences de la Nature et de l'Homme – ED 227

Thèse présentée et soutenue par
Armél RAFFOUR-MILLET

Le 11 Septembre 2017

**Identification du mécanisme impliqué dans la
formation de délétions de l'ADN mitochondrial :
cas de la « Common Deletion »
&
Développement d'un nouvel outil de caractérisation
des topoisomérases**

Sous la direction de :

Mme Emmanuelle DELAGOUTTE et Mme Erika BRUNET

Remerciements

Tout d'abord, je suis particulièrement touchée par le temps consacré par les membres de mon jury à l'appréciation de mon travail. Merci d'avoir accepté ce rôle, j'espère que mon manuscrit sera à la hauteur des attentes.

Je remercie Anne Lombès et Olivier Hyrien d'avoir accepté d'être rapporteur de cette thèse.

Je vous remercie, Claudine Mayer, pour votre suivi de mon parcours de thèse et votre participation à mes comités de thèse. J'ai particulièrement apprécié votre attention permanente et vos conseils.

Merci à François Strauss de m'avoir transmis la synthèse des minicercles et pour sa disponibilité. Merci pour tous nos échanges et nos discussions, notamment sur mes résultats et mes présentations.

Je remercie chaleureusement mes deux directrices de thèse pour leur encadrement pendant ces 4 années :

Merci Emmanuelle, pour m'avoir fait confiance dès le master M2 et pour votre formation attentive.

Je dois te remercier Erika pour avoir accepté la co-direction de ma thèse sans vraiment me connaître. J'ai toujours été très touchée de ta confiance. Grâce à ta rigueur et ton exigence, tu as su me faire énormément évoluer, me faire dépasser mes limites. Merci pour ta disponibilité et ton accompagnement de chaque instant (soir, nuit, weekend, vacances, ...). Grâce à toi, j'ai pu aller au bout de ce projet. Merci pour ton soutien et tes encouragements dans les moments difficiles. Tu as toujours eu les mots justes dans les périodes critiques pour me remotiver. Merci de m'avoir « poussée vers le haut ».

Ces quatre années, riches en événements, resteront inoubliables pour moi. J'ai été ravie de les partager avec toi.

Je remercie la « Fondation pour la Recherche Médicale » pour avoir financé ma 4^{ème} année de thèse. Je remercie Jean-François Riou de m'avoir accueillie au sein de son unité et Carine Giovannangeli pour m'avoir intégrée dans son équipe.

Merci à mes deux « super copines » de thèse, Marine et Karine. Rien que pour cette belle rencontre, cette thèse en valait la peine ! Marine, tu as été un binôme de folie !

Heureusement que tu étais là pour les arrivées matinales, les sorties tardives, les Amaxa à 1000 points, les journées western, etc... Merci pour toutes les pauses « ronchonades », « coup de gueules », et fous rires. Karine, tu es incroyable ... Qu'est-ce que tu as pu me faire rire ! Merci pour tes discours mobilisateurs, ton oreille attentive et la bonne ambiance assurée ! Encore bravo pour tes dons de voyance ! Merci à mes deux « licornasses », d'avoir été des amies incroyables et d'être maintenant des supers tatas pour Ewen. Vive les ***tch !

Merci Hind, tu resteras mon mentor de manips ! Merci d'être toujours présente, même à distance.

Merci Anne pour ton attention tout au long de ma thèse et en particulier pour toutes les discussions « non scientifiques » toujours pleines de bons conseils !

Je te remercie Anne-Laure pour ton accompagnement tout au long de ma thèse, pour ton écoute et tes conseils toujours judicieux.

Je remercie chaleureusement toutes les personnes que j'ai côtoyées pendant ces quatre années au laboratoire, notamment le « super groupe » du midi (Anthony, Gildas, François, Charlotte, Loïc) et mes « super collègues » de bureau. Un grand merci à la « team Erika » : Anna, Loelia, Marion, Hind, pour votre amitié et les précieux coups de main.

Je remercie mes amies, ma famille et ma belle-famille pour leur soutien et leur accompagnement. Un grand merci tout particulier au papa de Thibaud pour son aide précieuse pendant la rédaction de mon manuscrit.

Et pour finir, merci Thibaud....Pour ton infinie patience, ton soutien, ton écoute et ton accompagnement de chaque instant et surtout, pour n'avoir jamais cessé de croire en moi ! Merci d'avoir supporté les weekends « boulot », ma « ronchonade » légendaire, ma fatigue, le manque de temps, etc... Merci pour toutes tes attentions, d'avoir été autant investi dans cette thèse et de m'avoir permis de surmonter tous les moments de crises avec douceur et humour.

Et enfin, merci à Ewen, d'avoir été un « super bébé », qui m'a laissé un peu de temps pour écrire.

Je dédicace ma thèse à Thibaud

Avant-propos

Pendant ma thèse, j'ai travaillé sur deux projets différents :

Le premier projet, dirigé par Emmanuelle Delagoutte, a consisté à la mise au point d'un modèle d'étude *in vitro* pour les topoisomérases. La finalité de ce travail a été la publication d'un article en Août 2015. Cette première année de thèse m'a également permis de commencer à étudier la topologie de minicercles d'ADN en présence de magnésium et de calcium. Emmanuelle Delagoutte a ensuite poursuivi ce projet avec un manuscrit actuellement en préparation. La partie II de mon manuscrit traitera de cet ensemble.

Le second projet, encadrée par Erika Brunet, a été réalisée pendant les trois années suivantes, en collaboration avec le laboratoire d’Agnel Sfeir à New York. Son objectif a été d’étudier la réPLICATION et la réPARATION de l’ADN mitochondrial et plus particulièrement la formation d’une délétion appelée la « Common Deletion ». Ce travail publié en février 2017 sera présenté dans la partie I de mon manuscrit.

Partie I	p 1
Partie II	p137

PARTIE I

Abréviations

ADN : acide désoxyribonucléique
ADNc : ADN complémentaire
ADP : adénosine diphosphate
AFM : microscopie à force atomique
Alt-NHEJ : NHEJ alternatif
AMPc : adénosine monophosphate cyclique
APAF-1 : Apoptotic peptidase activating factor 1
APE1 : Apurinic/apyrimidinic endonuclease
ARN : acide ribonucléique
ARNm : ARN messager
ARNt : ARN de transfert
ATP : adénosine triphosphate
BER : réparation par excision de base
BET : bromure d'éthidium
bp : paires de bases
C terminal : Carboxy terminal
CBD : cassure double-brin
CD : Common Deletion
COX : cytochrome c oxidase
CSB : conserved sequence block
CtiP : C-terminal-binding protein interacting protein
Cy5 : cyanine 5
DNA-PK : DNA-Dependent Protein Kinase
DNMT3 : DNA methyl-transferase 3
EF-G : elongation factor G
EF-TU : Elongation Factor Thermo Unstable
EF-TS : elongation factor thermo stable
FISH : Hybridation in situ en fluorescence
gARN : guide ARN
gARNmito : guide ARN version mitochondriale
GTP : guanosine triphosphate

HR : recombinaison homologue
HSP : promoteur du brin H
IF2mt : mitochondrial initiation factor 2
IF3mt : mitochondrial initiation factor 3
kDa : kiloDaltons
LSP : promoteur du brin L
MCU : mitochondrial calcium uniporter
MDH : malate deshydrogénase
MEF : mouse embryonic fibroblast
MGME1 : Mitochondrial Genome Maintenance Exonuclease 1
MIA : Mitochondrial Intermembrane Space Assembly
mtADN : ADN mitochondrial
MLS : signal d'import à la mitochondrie
MMR : réparation des mésappariements
MPP : mitochondrial processing peptidase
Mre11 : Meitoic Recombination 11 homolog
MRN : Mre11 Rad50 Nbs1
mtADN : ADN mitochondrial
mtPAP : poly A polymerase, mitochondriale
mtRF1 : Release Factor 1, mitochondriale
mtSSB : Single-stranded DNA-binding protein, mitochondriale
NER : réparation par excision de base
N terminal : terminaison amine
nucléotides/min : nucléotides par minute
 O_H : origine de réplicatoio du brin H
 O_L : origine de replication du brin L
O/N : Over/Night
OPE : ophtalmoplégie progressive externe
PAM : Presequence translocase-Associated import Motor
PARP-1 : poly(ADP-ribose) polymérase 1
PCR : Polymerase Chain Reaction
Pi : phosphate inorganique
PNPase : Polynucleotide Phosphorylase

POLMRT : polymérase ARN mitochondriale

qPCR : PCR quantitative

RAD51 : Radiation sensitive 51

RAD52 : Radiation sensitive 51

RPA : Protein Replication A

ROS : Reactive Oxygen Species

RVD : Repeat-variable diresidue

SAM : Sorting and Assembly Machinery

siARN : Small interfering ARN

SOD2 : SuperOxyde Dismutase 2

TALEN : Transcription activator-like effector nucleases

TAS : séquence associée à la terminaison

TFAM : Mitochondrial transcription factor A

TFB1M : Transcription Factor B1, mitochondriale

TFB2M : Transcription Factor B2, mitochondriale

TIM : Translocase of the Inner Membrane

TOM : Translocase of the Outer Membrane

UTP : uridine triphosphate

VDAC : Voltage-dependent anion channel

XLF : XRCC4 Like Factor

XRCC4 : X-ray repair Complementing defective repair in Chinese hamster cells 4

2D : deux dimensions

Sommaire

Introduction	1
Quelques généralités sur la mitochondrie	1
I. La mitochondrie : une organelle cellulaire.....	1
II. Les fonctions de la mitochondrie	2
L'ADN mitochondrial	7
I. L'ADN mitochondrial humain.....	7
II. Les modèles proposés de réPLICATION.....	10
III. Transcription	14
IV. Traduction.....	15
V. Import des protéines	16
VI. Quelques protéines essentielles de la mitochondrie.....	19
VII. L'import des ARNs	22
VIII. La réparation de l'ADN mitochondrial.....	23
Les pathologies de la mitochondrie	36
I. Les délétions dans l'ADN mitochondrial	36
II. Les pathologies liées à des mutations des protéines du réplisome	39
Comment modifier le génome mitochondrial ?	43
I. La modification du génome mitochondrial dans les cellules cybrides.....	43
II. La modification du génome mitochondrial par l'utilisation des nucléases.....	43
Projet de thèse	52
Résultats	54
I. Article : « Single-Molecule Analysis of mtDNA Replication uncovers the basis of the Common Deletion ».....	54
II. Etude du rôle des protéines PARP-1 et RAD52 dans la formation de la « Common Deletion ».....	94
III. Etude de la quantité en ADN mitochondrial au cours du temps après coupure par les mitoTALENS	96
IV. Etude de la quantité en ADN mitochondrial au cours du temps après coupure par les mitoTALENS lors de l'inhibition de POLG et d'EXOG	97

V. Augmenter la proportion de « Common Deletion » dans une lignée cellulaire humaine.....	99
VI. Adaptation du système CRISPR/Cas9 pour la mitochondrie.....	102
Discussion-Perspectives	108
Conclusion.....	119
Annexes	121
Bibliographie.....	123

Table des figures

<u>FIGURE 1 : LA MITOCHONDRIE EST UNE ORGANELLE CELLULAIRE</u>	1
<u>FIGURE 2 : SCHEMA REPRESENTANT LA CHAINE RESPIRATOIRE PRESENTE AU SEIN DE LA MEMBRANE INTERNE DE LA MITOCHONDRIE....</u>	4
<u>FIGURE 3 : REPRESENTATION DE L'ADN MITOCHONDRIAL HUMAIN.....</u>	8
<u>FIGURE 4 : SCHEMA DE LA STRUCTURE D-LOOP RETROUVEE AU NIVEAU DE L'ORIGINE DE REPLICATION DU BRIN H DE L'ADN MITOCHONDRIAL HUMAIN</u>	9
<u>FIGURE 5 : REPRESENTATION DE LA STRUCTURE TIGE/BOUCLE DE L'ORIGINE DE REPLICATION O_L AU NIVEAU D'UN CLUSTER DE GENES CODANTS POUR 5 ARNs DE TRANSFERT.</u>	10
<u>FIGURE 6 : REPRESENTATION DU MODELE DE REPLICATION PAR DEPLACEMENT DE BRIN</u>	11
<u>FIGURE 7 : REPRESENTATION DU MODELE SYMETRIQUE DE REPLICATION</u>	12
<u>FIGURE 8 : REPRESENTATION DU MODELE RITOLS.....</u>	13
<u>FIGURE 9 : SCHEMA REPRESENTANT LES DIFFERENTS TRANSCRITS ISSUS DE L'ADN MITOCHONDRIAL.....</u>	15
<u>FIGURE 10 : SCHEMA REPRESENTANT LES PRINCIPAUX COMPLEXES PARTICIPANT A L'ADRESSAGE DES PROTEINES DANS LES DIFFERENTS COMPARTIMENTS DE LA MITOCHONDRIE</u>	18
<u>FIGURE 11 : SCHEMA REPRESENTANT LES ARN MODIFIES AFIN DE PERMETTRE LEUR IMPORT DANS LA MITOCHONDRIE</u>	23
<u>FIGURE 12 : LES DIFFERENTES SOURCES DE DOMMAGES A L'ADN ET LEURS CONSEQUENCES</u>	24
<u>FIGURE 13 : SCHEMA RECAPITULATIF DE LA VOIE DE REPARATION PAR EXCISION DE BASE (BER)</u>	26
<u>FIGURE 14 : SCHEMA REPRESENTANT LA VOIE DU NHEJ CLASSIQUE</u>	29
<u>FIGURE 15 : SCHEMA PRESENTANT LA VOIE ALTERNATIVE DU NHEJ</u>	31
<u>FIGURE 16 : SCHEMA PRESENTANT LES PRINCIPALES ETAPES DE LA RECOMBINAISON HOMOLOGUE.....</u>	34
<u>FIGURE 17 : LES DELETIONS LES PLUS SOUVENT RETROUVEES ET LEURS PATHOLOGIES ASSOCIEES</u>	37
<u>FIGURE 18 : REPRESENTATION DE LA « COMMON DELETION »AU SEIN DE L'ADN MITOCHONDRIAL</u>	38
<u>FIGURE 19 : SCHEMA D'UNE NUCLEASE A DOIGT DE ZINC.....</u>	45
<u>FIGURE 20 : SCHEMA REPRESENTANT L'AUGMENTATION DE LA SPECIFICITE DE COUPURE PAR AMELIORATION DU DESIGN DES NUCLEASES</u>	47
<u>FIGURE 21 : SCHEMA D'UNE TALEN</u>	48
<u>FIGURE 22 : SCHEMA REPRESENTANT LE SYSTEME CRISPR/CAS9, MECANISME DU SYSTEME IMMUNITAIRE DES BACTERIES</u>	49
<u>FIGURE 23 : SCHEMA REPRESENTANT LE SYSTEME CRISPR/CAS9.....</u>	50
<u>FIGURE 24 : EFFET DE L'INHIBITION DE RAD52 SUR LA FORMATION DE LA « COMMON DELETION ».....</u>	94
<u>FIGURE 25 : EFFET DE L'INHIBITION DE PARP-1 SUR LA FORMATION DE LA « COMMON DELETION »</u>	95
<u>FIGURE 26 : RESULTATS DE LA QUANTITE D'ADN MITOCHONDRIAL AU COURS DU TEMPS APRES TRANSFECTION PAR LES MITOTALENS DANS DES HCT116</u>	97
<u>FIGURE 27 : QUANTITE D'ADN MITOCHONDRIAL AU COURS DU TEMPS APRES TRANSFECTION PAR LES MITOTALENS LORS DE L'INHIBITION DE POLG OU ExoG DANS DES HCT116.....</u>	98
<u>FIGURE 28 : ENRICHISSEMENT EN MOLECULES DELETEES APRES DES DOUBLES TRANSFECTIONS PAR LES MITOTALENS DANS DES U2OS</u>	101

<u>FIGURE 29 : SCHEMA REPRESENTANT LES ARN MODIFIES AFIN DE PERMETTRE LEUR IMPORT DANS LA MITOCHONDRIE</u>	103
<u>FIGURE 30 : TEST D'INDUCTION DE LA « COMMON DELETION » PAR LA MITOCAS9</u>	103
<u>FIGURE 31 : TEST DE CLIVAGE IN VITRO AVEC LA CAS9 ET LES GARNS AVEC OU SANS SIGNAUX D'ADRESSAGE A LA MITOCHONDRIE .</u>	104
<u>FIGURE 32 : IMAGES D'EXPERIENCES IMMUNOFLUORESCENCE APRES TRANSFECTION AVEC LA CAS9 NUCLEAIRE OU MITOCHONDRALE EN PRESENCE D'UN GARN MITOCHONDRIAL OU NUCLEAIRE DANS DES U2OS</u>	105
<u>FIGURE 33 : RESULTATS DE LA SYNTHESE DES GARNS FLUORESCENTS PAR TRANSCRIPTION <i>IN VITRO</i></u>	106
<u>FIGURE 34 : LOCALISATION DES GARNS AVEC OU SANS SIGNAUX D'ADRESSAGE POUR LA MITOCHONDRIE DANS DES U2OS.....</u>	106
<u>FIGURE 35 : QUANTITE DE « COMMON DELETION » AU COURS DU TEMPS APRES COUPURE PAR LA MITOTALEN CD 5' DANS DES U2OS</u>	115

Introduction

Quelques généralités sur la mitochondrie

I. La mitochondrie : une organelle cellulaire

L'étymologie du mot mitochondrie nous vient du grec ancien, « mitos » le filament et « chondros » la graine, vocable qui lui a été attribué en raison de sa forme. La mitochondrie est une organelle de la cellule située dans le cytoplasme. Elle est considérée comme la source d'énergie des cellules, car elle leur fournit l'adénosine triphosphate (ATP) nécessaire. Elle joue aussi un rôle central dans le métabolisme cellulaire puisqu'elle assure, entre autres, le métabolisme des acides gras, le métabolisme du calcium et celui du fer, et se trouve aussi impliquée dans la mort cellulaire.

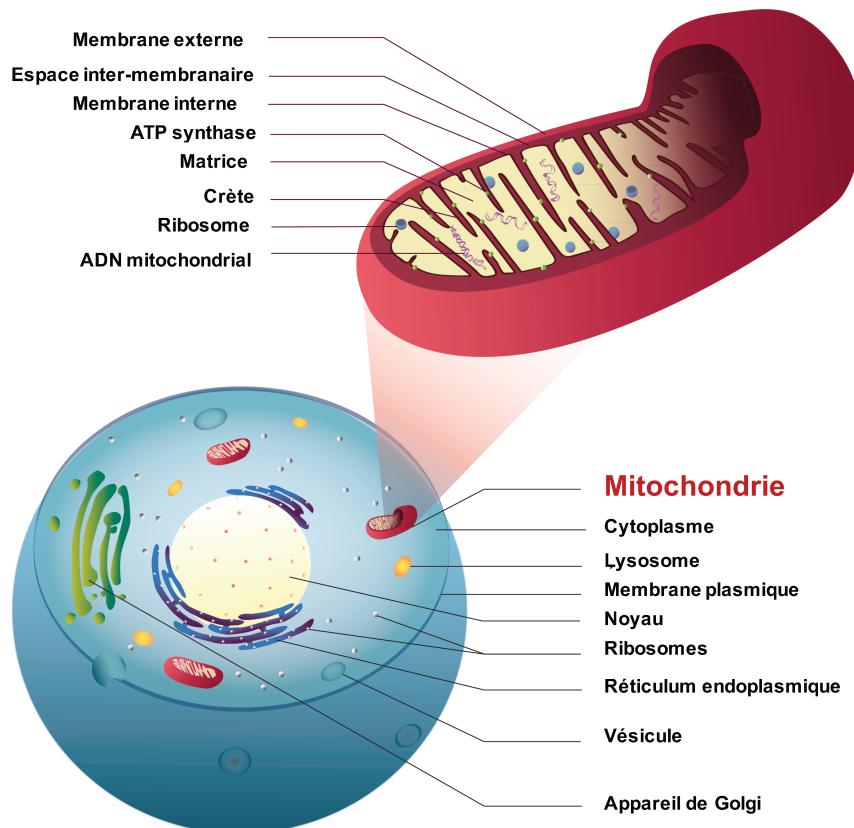


Figure 1 : La mitochondrie est une organelle cellulaire

La mitochondrie est une organelle présente dans le cytoplasme de la cellule. Elle est constituée de deux membranes externe et interne, séparées par un espace inter-membranaire. L'intérieur de la mitochondrie est appelé matrice. L'ADN mitochondrial, les ribosomes sont présents au sein de la matrice.

L'origine des mitochondries au sein des cellules eucaryotes est expliquée par la théorie endosymbiotique. De par ses caractéristiques morphologiques, il a été suggéré dès l'année 1926 que les mitochondries auraient une origine bactérienne (Wallin 1926). En 1985, les alpha-protéobactéries sont considérées comme étant les meilleurs candidats pour être les ancêtres de la mitochondrie (Yang et al. 1985). Cette hypothèse a été depuis confirmée et il est admis que la mitochondrie est issue d'une alpha-protéobactérie intégrée dans les cellules par endosymbiose (Thrash et al. 2011).

Concernant sa morphologie, la taille d'une mitochondrie est en moyenne d'1 μm . Elle est entourée de deux membranes : la membrane externe et la membrane interne. L'espace entre ces deux membranes est appelé espace intermembranaire. A l'intérieur de la mitochondrie, se situe la matrice, dans laquelle se trouvent plusieurs copies d'ADN circulaire (Figure 1).

Les mitochondries ne peuvent pas être créées *de novo*. Ainsi, la régulation du nombre de mitochondries se fait par des mécanismes de fission et de fusion. Ces deux mécanismes sont présents en permanence dans la cellule et fonctionnent en équilibre selon ses besoins. Lorsque la cellule requiert davantage de mitochondries, notamment pour assurer son métabolisme ou lors des divisions cellulaires, les mitochondries fissionnent alors afin de fournir de nouvelles mitochondries. La fusion des mitochondries, quant à elle, permet de maintenir une hétérogénéité de l'ADN mitochondrial et de complémenter des mitochondries endommagées avec de l'ADN intact provenant d'une autre mitochondrie. Avant la division cellulaire, il a été observé que les mitochondries fusionnent puis fissionnent préalablement, permettant la transmission d'un patrimoine hétérogène de la cellule mère aux cellules filles. La fission est aussi observée lors du mécanisme de mort cellulaire appelé apoptose. Les protéines responsables de la fusion et de la fission sont des protéines de la famille de la Dynamine, GTP dépendantes (Scott & Youle 2010) (van der Bliek et al. 2013).

II. Les fonctions de la mitochondrie

1) La production d'ATP

La mitochondrie permet la production d'adénosine triphosphate (ATP) à partir des sucres, des lipides et des protéines. Les sucres rentrent dans la mitochondrie sous forme de pyruvate après la glycolyse et sont transformés en acétyl-CoA. La β -oxydation qui permet le métabolisme des lipides a lieu dans la mitochondrie, aboutissant elle aussi à la production

d'acétyl-CoA. Concernant les protéines, selon les acides aminés, elles peuvent être converties, soit directement en acétyl-CoA, soit en passant par une étape intermédiaire donner lieu à la formation de pyruvate. L'acétyl-CoA obtenu par les différentes voies métaboliques est converti en ATP via le cycle de Krebs. Les électrons obtenus lors du cycle de Krebs sont pris en charge par le coenzyme NADH. Ces électrons vont servir d'énergie à la chaîne respiratoire qui se situe au niveau des crêtes de la membrane interne de la mitochondrie. La coordination des différents complexes de la chaîne respiratoire via des réactions d'oxydo-réduction aboutit à la synthèse d'ATP (Figure 2). Cet ATP, source d'énergie, sera ensuite redistribué à la cellule.

Les acteurs principaux de la chaîne respiratoire sont :

- Le complexe 1 : -la NADH-coenzyme Q-oxydoréductase- est constitué de 45 sous-unités. Il utilise les électrons provenant du NADH et transporte quatre ions H⁺ de la matrice vers l'espace intermembranaire.
- Le complexe 2 : -la succinate-coenzyme Q-oxydoréductase- est entièrement codé par les gènes nucléaires. Ce complexe utilise les électrons provenant du FADH₂ et les transfère au coenzyme Q.
- Le coenzyme Q est une molécule qui diffuse dans la membrane interne. Il transfère sous forme réduite les électrons du complexe I et II vers le complexe III.
- Le complexe 3 : -la cytochrome c réductase- oxyde le coenzyme Q et transmet les électrons libérés au cytochrome c et transporte deux ions H⁺ de la matrice vers l'espace intermembranaire.
- Le cytochrome c, une petite molécule de 13 kDa, assure le transport des électrons entre le complexe III et le complexe IV.
- Le complexe 4 : l'apport des électrons par le cytochrome c permet la formation de molécules d'H₂O à partir d'O₂ et transporte quatre ions H⁺ de la matrice vers l'espace intermembranaire

Le transfert des électrons réalisé par ces complexes crée une force protomotrice, composée d'une part d'un gradient de pH (plus acide dans l'espace intermembranaire) et d'autre part d'une différence de charge électrostatique (plus positive dans l'espace intermembranaire).

Le complexe V est l'ATP synthase. La découverte du mécanisme de ce complexe a donné lieu au prix Nobel de chimie en 1997 pour Paul Boyer et John Walker (Boyer 1975). Il fonctionne comme un « moteur moléculaire » grâce à la force protomotrice. Il est composé de deux domaines F₀ et F₁. Les ions H⁺ sont transportés par le canal F₀ de l'espace

intermembranaire vers la matrice. Le domaine F_1 , lié au domaine F_0 du côté de la matrice, synthétise de l'ATP à partir de l'ADP et du phosphate inorganique (P_i).

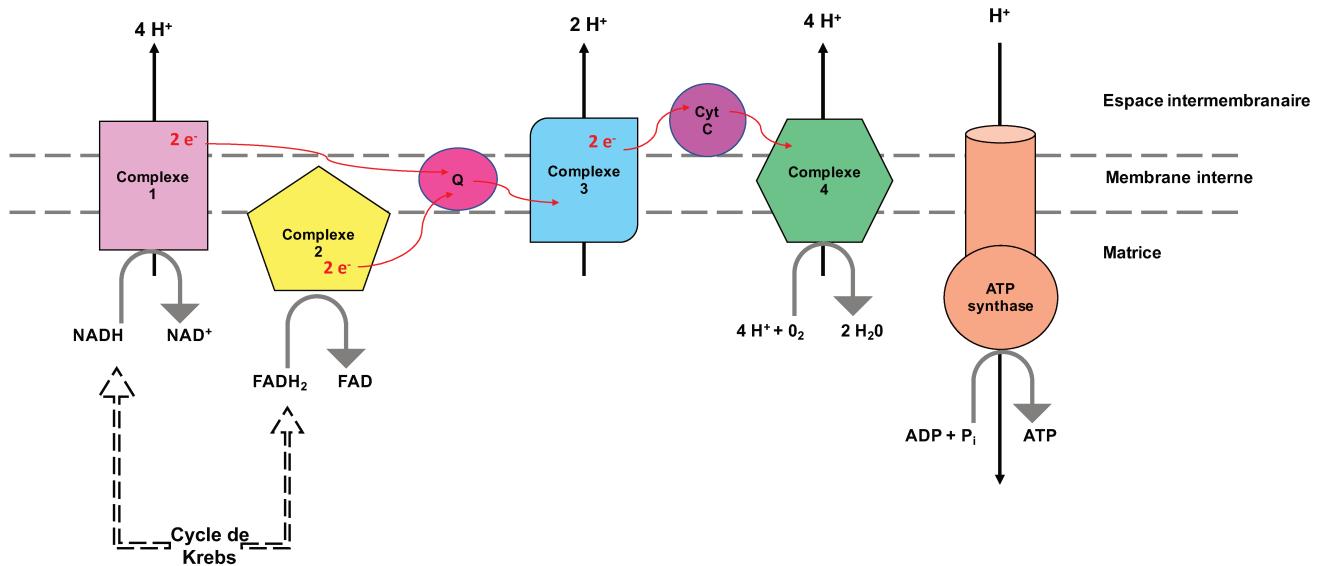


Figure 2 : Schéma représentant la chaîne respiratoire présente au sein de la membrane interne de la mitochondrie

Le cycle de Krebs produit du NADH et FADH₂. Ils vont être oxydés et fournir deux électrons au niveau du complexe 1 et 2. Le complexe 1 transfère 4 protons dans l'espace intermembranaire. Les électrons récupérés au sein des complexes 1 et 2 sont transférés au complexe 3 par le coenzyme Q (Q). Le coenzyme Q est oxydé au niveau du complexe 3 et transmet les deux électrons au cytochrome c. Le complexe 4 récupère les deux électrons transmis par le cytochrome c et permet la réduction de l'O₂ afin de former de l'H₂O. Au sein de ce complexe 4, quatre protons sont transférés dans l'espace intermembranaire. L'ATP synthase utilise le gradient de protons créé par les quatre complexes de la chaîne respiratoire pour fournir l'énergie nécessaire à la synthèse d'ATP.

2) L'apoptose

L'apoptose correspond à un processus de mort cellulaire qui se produit notamment lors du développement embryonnaire, permettant par exemple la réduction de l'appendice caudal. Elle est aussi impliquée dans l'élimination de cellules présentant des dommages irréparables de leur ADN. L'apoptose est déclenchée par des facteurs tels que le stress ou des lésions dans l'ADN. Ces signaux vont inhiber la protéine BCl-2, protéine anti-apoptotique et activer les protéines pro-apoptotiques, Bax et Bak. Ces protéines vont s'oligomériser et provoquer une perméabilisation de la membrane externe de la mitochondrie. Cela entraîne un relargage du cytochrome c. Le cytochrome c se lie à APAF-1 (Apoptotic Peptidase Activating Factor). Ce complexe active la voie des caspases et entrainera la formation de l'apoptosome. La finalité sera une condensation de l'ADN nucléaire puis une fragmentation de la cellule. Les débris cellulaires seront ensuite phagocytés (Ichim & Tait 2016).

3) Le métabolisme du Fer

La mitochondrie participe à l'homéostasie du fer. Le fer rentre dans la mitochondrie grâce à la mitoferrine. Le fer peut ainsi former l'hème qui sera ensuite utilisé pour la production d'hémoglobine. Les premières et les dernières étapes de la synthèse sont réalisées dans la mitochondrie. Cette dernière joue aussi un rôle dans la synthèse des groupements fer-soufre des protéines, comme par exemple celles du complexe de la chaîne respiratoire. La mitochondrie participe aussi à l'homéostasie du fer en participant à son stockage grâce à la mitoferritine (Chifman et al. 2014) (Barupala et al. 2016).

4) L'homéostasie du calcium

La mitochondrie joue un rôle dans l'homéostasie du calcium. Elle a la capacité de stocker le calcium, agissant tel un réservoir pour les ions calcium afin qu'ils soient rapidement disponibles selon les besoins de la cellule. Le calcium rentre dans la mitochondrie via le canal VDAC (Voltage-Dependent Anion Channel) pour passer la membrane externe puis la membrane interne via le MCU (Mitochondrial Calcium Uniporter). Le calcium dans la mitochondrie est aussi impliqué dans l'apoptose, et participe donc à la morphologie des mitochondries, car le mécanisme de fission est associé à l'apoptose. Enfin, le calcium est un régulateur du cycle de Krebs, trois enzymes étant dépendantes du calcium : la pyruvate déshydrogénase, l'isocitrate déshydrogénase et l'α-cétoglutarate déshydrogénase. Ainsi, il y a une régulation entre l'activité des enzymes du cycle de Krebs et le calcium disponible (Bianchi et al. 2004) (Contreras et al. 2010).

5) Le métabolisme des lipides

La mitochondrie est impliquée dans la dégradation des acides gras. La dégradation fournit de l'acétyl-CoA qui sera utilisé par le cycle de Krebs. Seuls les acides gras à chaîne courte (moins de quatorze carbones) peuvent diffuser à travers les membranes mitochondrielles. Pour les acides gras à chaîne longue (plus de quatorze carbones), le système de la carnitine translocase permet aux acides gras de passer la membrane externe sous forme d'acyl-carnitine grâce aux acyl-carnitine transférases puis le passage de la membrane interne de la mitochondrie se fait par une translocase, la carnitine-acylcarnitine translocase. Une fois dans la matrice, au terme de quatre étapes de dégradation des acides gras, l'acétyl-CoA finalement libéré rentrera dans le cycle de Krebs. Les mitochondries du foie peuvent réaliser la cétogénèse lors d'un excès d'acétyl-CoA ou lors d'un jeûne (Kastaniotis et al. 2017).

6) Les interactions mitochondrie-noyau

La mitochondrie établit des liens étroits avec le noyau. En effet, son rôle dans le métabolisme nécessite une bonne communication pour adapter ce dernier au besoin de la cellule. Le noyau code pour la majorité des protéines mitochondrielles. De ce fait, la régulation de leurs expressions est une façon de réguler les différentes fonctions de la mitochondrie. Les petites molécules comme le calcium, l'AMPc, les ROS (Reactive Oxygen Species), sont impliquées dans de nombreuses voies de signalisation cellulaires, et participent donc aux interactions noyau-mitochondrie. Un autre exemple est celui de la régulation de la morphologie des mitochondries, fission et fusion ayant lieu lors de la division cellulaire. (Dominy & Puigserver, 2013).

L'ADN mitochondrial

La mitochondrie est la seule organelle dans la cellule qui possède son propre ADN (mtADN). En 1991, l'équipe de Kuroiwa observe par microscopie à épifluorescence que chaque mitochondrie possède de 1 à 15 copies de son ADN, regroupées au sein de structure de type nucléoïdes (Satoh & Kuroiwa 1991). Le nombre de copies peut être très différent selon les tissus. Ainsi, il est d'environ 3700 molécules pour les cellules du muscle squeletique tandis qu'il est de 7000 pour les cardiomyocytes (Miller et al. 2003).

I. L'ADN mitochondrial humain

Chez l'homme, l'ADN mitochondrial est un ADN double-brin circulaire de 16 569 paires de bases (bp). Cet ADN est essentiellement codant. L'ADN mitochondrial possède 37 gènes dont 13 codent pour 13 des 90 protéines du complexe de la chaîne respiratoire, 2 pour des ARNs ribosomiaux et 22 pour des ARNs de transfert (Anderson et al. 1981). Les deux brins de l'ADN mitochondrial s'appellent respectivement brin H pour « heavy » (lourd) et brin L pour « light » (léger). Leur nom vient de leurs différences de densité observées par gradient de chlorure de césium. Le brin H est en effet plus riche en guanine comparé au brin L, 31% pour le brin H au lieu de 13% pour le brin L pour l'ADN mitochondrial humain (Figure 3).

1) Les origines de réPLICATION

Le génome mitochondrial est essentiellement codant. Les parties non codantes serviraient essentiellement pour l'initiation de la réPLICATION et de la transcription.

a) L'origine de réPLICATION du brin H : La D-loop

La principale partie non codante de l'ADN mitochondrial est retrouvée au niveau d'une structure D-Loop. Cette structure particulière servirait pour l'initiation de la transcription et de la réPLICATION. En effet, on y retrouve l'origine de réPLICATION du brin H (O_H) et deux promoteurs, un pour chaque brin. Le LSP (Light-Strand Promoter) est le promoteur pour le brin L, tandis que le HSP (Heavy-Strand Promoter) est le promoteur du brin H qui peut utiliser deux sites pour l'initiation de la transcription : HSP1 et HSP2.

Cette structure en D-loop a une taille de 1124 bp (positions 16024-576), elle est délimitée de part et d'autre par la séquence codante de deux ARNs de transfert. Des alignements de séquence réalisés chez le rat, la souris et l'homme montrent que la séquence de la D-loop est

conservée notamment au niveau de 3 blocs de séquences appelés « Conserved Sequence Block » (CSB) (Walberg & Clayton 1981).

Cette région prend le nom de D-loop car elle présente un troisième brin d'ADN entre les deux brins de l'ADN mitochondrial. Il est appelé ADN 7S étant donné son coefficient de sédimentation, pourtant c'est un abus de langage car il existerait certainement quatre ADNs 7S qui diffèrent légèrement en taille. Les ADNs 7S font en moyenne 650 nucléotides. Il semblerait qu'ils se terminent tous au même endroit en 3', au niveau des séquences TAS (Termination Associated Sequences) de la D-loop. La variabilité de ces ADNs 7S viendrait donc de leur site d'initiation en 5'. A l'heure actuelle, on ignore si ces ADNs sont élongués lors de la réPLICATION ou si un nouveau brin est re-synthétisé à partir de l'origine de réPLICATION O_H.

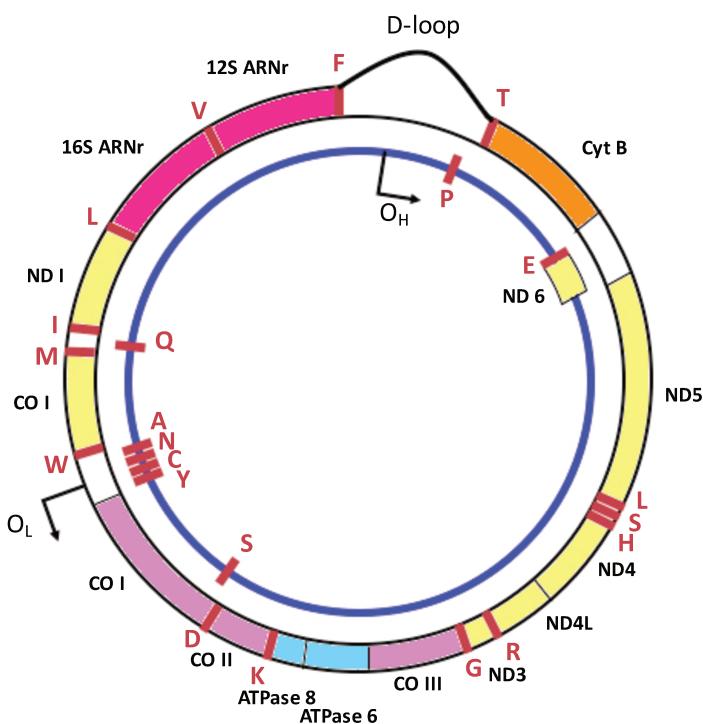


Figure 3 : Représentation de l'ADN mitochondrial humain

Le brin épais coloré correspond au brin H tandis que le brin bleu correspond au brin L de l'ADN mitochondrial. O_H est l'origine de réPLICATION du brin lourd et est située dans la région D Loop. O_L est l'origine de réPLICATION pour le brin léger. HSP1 et HSP2 sont les deux promoteurs présents sur le brin H. LSP est le promoteur du brin L. Les deux ARNs ribosomiaux sont le 12S et le 16S. Les ARNs de transferts sont annotés par une lettre majuscule uniquement. La D-loop est la plus grande partie non codante. Les treize gènes codant pour 13 des protéines de la chaîne respiratoire sont : CO I-III : sous-unités 1 à 3 du cytochrome c oxydase, ND1-6 : sous-unités 1 à 6 de la NADH déshydrogénase, ATP6 et 8 sous unités 6 et 8 de l'ATP synthase.

En amont de l'ADN 7S, un ARN 7S a été identifié. Cette fois encore, plusieurs ARNs 7S ont été observés. Ces ARNs 7S débutent au niveau du promoteur LSP et se terminent dans la région des CSB. Il a été suggéré que l'ARN 7S servirait de primer pour l'initiation de la réPLICATION au niveau de O_H en servant de régulateur de la réPLICATION (Chang & Clayton 1985) (Figure 4).

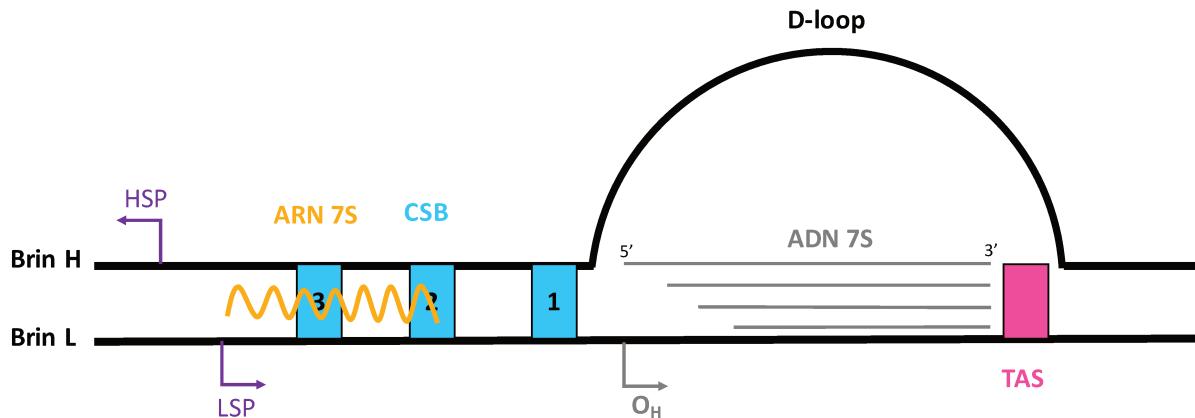


Figure 4 : Schéma de la structure D-Loop retrouvée au niveau de l'origine de réPLICATION du brin H de l'ADN mitochondrial humain

L'ADN 7S (pouvant être de taille différente) forme une structure D-loop est représenté par les traits gris. La D-loop s'étend de l'origine de réPLICATION O_H et se finit vers les séquences TAS. TAS signifie « séquence associé à la terminaison » (en rose). En amont, l'ARN 7S débute au niveau du promoteur du brin L, (LSP) et se termine au niveau du CSB 2. CSB 1,2,3 représentent les blocs de séquences conservés (en bleu). HSP est le promoteur du brin H.

b) L'origine de réPLICATION du brin L : O_L

L'origine supposée de réPLICATION du brin L se trouve au 2/3 de l'ADN mitochondrial en partant de l' O_H (Figure 4). Elle est située dans une région codante pour un cluster de 5 ARNs de transfert (Trp, Ala, Asn, Cys, Tyr). Il est suggéré qu'à cet endroit de l'ADN, une structure tige/boucle est adoptée et permet la formation de l'origine de réPLICATION O_L (Tapper & Clayton 1981) (Figure 5). L'initiation au niveau de l' O_L est réalisée par la protéine POLRMT (Fusté et al. 2010).

Des origines de réPLICATION alternatives du brin L pourraient aussi être utilisées. En effet, l'équipe de Clayton (Brown et al. 2005) observe des intermédiaires de réPLICATION par microscopie à force atomique (AFM) qui correspondent à l'utilisation d'origines alternatives. De plus, Seligmann (Seligmann, 2010) montre par une étude bioinformatique que les ARNs de transfert présents sur le brin H peuvent former des structures OL-like qui pourraient servir d'origine de réPLICATION alternative.

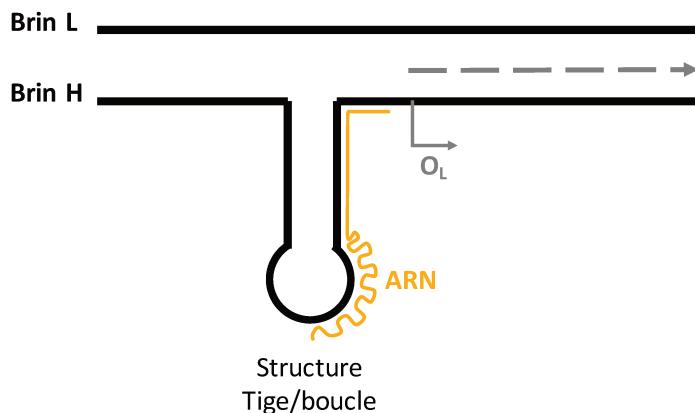


Figure 5 : Représentation de la structure tige/boucle de l'origine de réPLICATION O_L au niveau d'un cluster de gènes codants pour cinq ARNs de transfert.

Une structure tige/boucle va permettre la formation d'une origine de réPLICATION, appelée O_L . Les ARNs transcrits vont pouvoir servir de primers d'initiation de la réPLICATION.

II. Les modèles proposés de réPLICATION

Comme au sein du noyau, l'ADN mitochondrial doit être répliqué pour être maintenu et transmis lors des divisions cellulaires et/ou lors des événements de fission mitochondriale. Les mécanismes de régulation de la réPLICATION ne sont pas encore entièrement compris. La régulation de la réPLICATION mitochondriale est notamment importante lorsque les besoins de la cellule sont modifiés et pour les cellules en division.

Depuis la découverte de l'ADN mitochondrial, les chercheurs tentent d'élucider le mécanisme de réPLICATION. Les stratégies utilisées ont été : l'extraction du mtADN puis l'observation des intermédiaires de réPLICATION par séparation par gradient (Robberson & Clayton 1972), électrophorèse en 2 dimensions (2D) (Yang et al. 2002) ou la microscopie par force atomique (Brown *et al.*, 2005). A l'heure actuelle, plusieurs modèles sont proposés et font encore débat. Concernant la vitesse de réPLICATION de l'ADN mitochondrial, elle est considérée comme étant plutôt lente (270 nucléotides/min) et il faut donc plus d'une heure pour répliquer la totalité de cet ADN circulaire (Clayton 1982).

1) Modèle de déplacement de brin

Historiquement, il a été proposé que l'ADN mitochondrial se réplique selon un modèle de déplacement de brin. Ce modèle est dit asymétrique et asynchrone. Les deux brins ne sont pas répliqués exactement au même moment. Ce modèle utilise les deux origines de réPLICATION (Figure 6).

L'initiation de la réPLICATION du brin H commence au niveau de l' O_H présente dans la D-loop. Lorsque le brin H est répliqué au 2/3, la fourche de réPLICATION arrive au niveau de l'origine de réPLICATION du brin L. L'initiation du brin L commence alors au niveau de l' O_L . La réPLICATION du brin L progresse en sens inverse du brin H et remonte vers l' O_H . Il a été d'abord proposé que l'ADN simple brin exposé pendant la réPLICATION du brin H est stabilisé par la protéine mitochondriale de liaison à l'ADN simple brin, mtSSB.

David Clayton était le référent du domaine. Ses travaux depuis les années 80 ont étayé et validé ce modèle (Clayton 1982). En 1972, par exemple, son laboratoire réalise des expériences chez la souris (Robberson & Clayton 1972). Ils utilisent des expériences de centrifugation en présence d'un gradient de bromure d'éthidium avec du chlorure de césium et la microscopie électronique pour séparer des intermédiaires de réPLICATION de l'ADN mitochondrial. Leurs résultats montrent la présence d'intermédiaires présentant deux zones ouvertes (suggérant l'utilisation des deux origines de réPLICATION) et des intermédiaires de réPLICATION présentant une D-loop étendue. Ces résultats valident le modèle de déplacement de brins. Des travaux plus récents, réalisés par l'équipe de Falkenberg (Fusté et al. 2014), montrent par des expériences *in vitro* et du Chip-seq que la protéine mtSSB protégerait bien le brin H lors de la progression de la fourche de réPLICATION entre l' O_H et l' O_L . Ces expériences participent à valider le modèle de déplacement des brins.

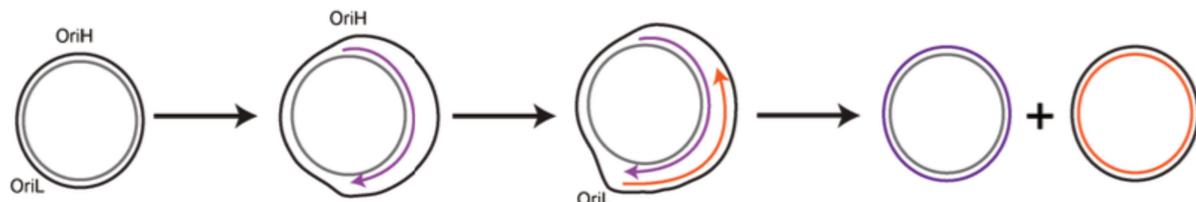


Figure 6 : Représentation du modèle de réPLICATION par déplacement de brin

L'initiation de la réPLICATION du brin H commence au niveau de l' O_H puis la fourche de réPLICATION (brin violet) arrive au niveau de l'origine de réPLICATION du brin L : l' O_L . L'initiation du brin L commence alors au niveau de l' O_L . La réPLICATION du brin L progresse en sens inverse du brin H et remonte vers l' O_H (brin orange).

2) Modèle symétrique

Dans un second temps, le modèle symétrique a été proposé. Ce modèle utilise une seule origine de réPLICATION qui serait située un peu en aval de l'origine de réPLICATION O_H . La réPLICATION des deux brins se ferait au sein d'une seule fourche de réPLICATION comme au sein du noyau. Le brin principal répliqué est le brin H et le brin retardé est le brin L (Figure 7).

Ce modèle a été proposé pour la première fois en 2000 par le laboratoire de Holt (Holt et al. 2000). Ses travaux reposent sur l'interprétation des intermédiaires de réPLICATION par gels d'électrophorèse 2D (Brewer & Fangman 1967). Les résultats de ces gels en électrophorèse 2D leur permettent de mettre en évidence des arcs Y « classiques », qui suggèrent un mode de réPLICATION traditionnel avec un brin principal et un brin retardé. Pourtant, ils observent aussi des intermédiaires de réPLICATION qui correspondraient au modèle de déplacement de brins. Les expériences réalisées dans des cellules humaines de placenta et dans des cellules murines de foie montrent des résultats similaires. De plus, les auteurs testent si les cellules utilisent le même mode de réPLICATION selon qu'elles reconstituent leur population en mtADN ou qu'elles sont en quiescence (quand elles maintiennent un nombre constant de copies). Pour cela, après traitement au bromure d'éthidium (BET) provoquant une perte drastique en mtADN, ils étudient la repopulation en mtADN. Ils observent ainsi des intermédiaires de réPLICATION correspondant au modèle symétrique pour les cellules en repopulation alors que les cellules quiescentes qui maintiennent uniquement une quantité constante de mtADN révèlent des intermédiaires de réPLICATION suggérant le modèle de déplacement des brins.

Les auteurs proposent donc que les deux modèles pourraient coexister et que la réPLICATION ne se ferait pas selon le même modèle selon que les cellules maintiennent ou se repeuplent en ADN mitochondrial.

En 2002, les mêmes auteurs (Yang et al. 2002) reviennent sur leurs travaux publiés en 2000 et avancent que seul le modèle symétrique est présent (Holt et al. 2000). Dans ces travaux, les auteurs utilisent l'électrophorèse 2D comme précédemment pour conclure sur les intermédiaires de réPLICATION. Leurs précédentes conclusions sur le modèle de déplacement de brins viendraient d'un artéfact dû à la méthode de purification de l'ADN mitochondrial. Avec leur nouvelle méthode de préparation d'ADN mitochondrial en gradient de sucrose, ils ne retrouvent plus les intermédiaires de réPLICATION du modèle de déplacement des brins. Ils concluent donc à l'utilisation du modèle symétrique par les cellules.

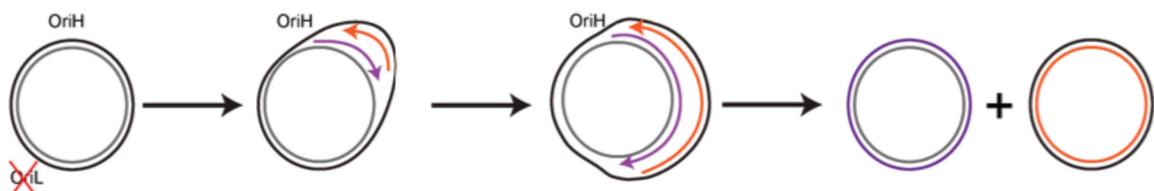


Figure 7 : Représentation du modèle symétrique de réPLICATION

Une seule origine de réPLICATION est utilisée un peu en aval de l'origine de réPLICATION O_H . La réPLICATION des deux brins se ferait au sein d'une seule fourche de réPLICATION. Le brin principal répliqué est le brin H et le brin retardé est le brin L.

3) Le modèle RITOLs (RNA incorporation throughout the lagging strand)

En plus, des deux modèles actuellement proposés, le modèle RITOLs est défini comme une sorte de modèle alternatif entre les deux précédents. L'équipe de Holt, eux-mêmes qui avaient proposé le modèle symétrique, (Yasukawa et al. 2006) envisage ce modèle après avoir visualisé la présence d'ARN lors de la synthèse du brin retardé par électrophorèse 2D. Ce modèle s'oppose au rôle de mtSSB dans le modèle de déplacement des brins, et ce serait des ARNs qui protégeraient le brin L lors de la réPLICATION du brin H. De plus, le laboratoire de Holt (Holt & Reyes 2012a) renforce cette idée en suggérant que des longs ARNs issus de la transcription pourrait servir de matrice de réparation lors des erreurs de réPLICATION en plus de protéger le brin L. L'origine de réPLICATION utilisée avec ce modèle n'est pas bien définie, mais il va dans le sens du modèle de déplacement de brins concernant la synthèse asynchrone des deux brins (Figure 8).

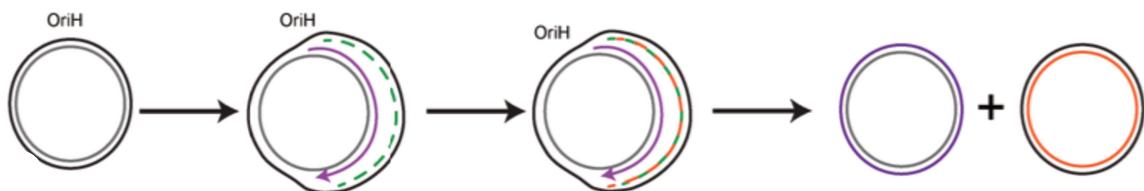


Figure 8 : Représentation du modèle RITOLs

Ce modèle est un mélange entre les deux modèles. Les origines de réPLICATION ne sont pas bien définies. La différence majeure repose sur la présence d'ARN au niveau du brin retardé et non pas par une protection par la protéine mtSSB.

4) Nécessité d'approfondir la réPLICATION de l'ADN mitochondrial

La réPLICATION de l'ADN mitochondrial est étudiée depuis de nombreuses années mais fait encore débat. En effet, jusqu'à présent, aucun modèle ne concilie complètement les résultats de toutes les expériences réalisées jusqu'à maintenant.

Une grande partie du débat vient des techniques utilisées. En effet, en 2002, l'équipe de Holt (Yang et al. 2002) revient sur ses résultats obtenus en 2000 (Holt et al. 2000) en remettant en cause la technique d'extraction de l'ADN mitochondrial. De surcroît, en 2005, l'équipe de Clayton (Brown et al., 2005) utilise l'AFM pour observer les intermédiaires de réPLICATION. Leurs travaux montrent des intermédiaires avec des D-loops élongées et d'autres qui correspondent à une initiation au niveau de l' O_L . Ils réalisent alors des gels en électrophorèse 2D sur les échantillons observés en AFM et retrouvent des structures en arc Y « classiques » comme observées par Holt et ses collègues. Les auteurs associent ces différences entre les deux techniques aux limites de l'électrophorèse 2D à séparer distinctement des

intermédiaires de réPLICATION DIFFéRENTS. Il est par ailleurs très souvent reconnu que l'analyse des gels 2D est très délicate. Concernant l'incorporation d'ARN lors de la synthèse du brin L, lorsque des résultats montrent l'absence d'ARN le long de ce brin, certains auteurs parlent de présence de RNase lors des préparations d'ADN mitochondrial.

Les modèles de réPLICATION FONT donc encore l'OBJET DE déBAT ET LES DIFFéRENTES PUBLICATIONS DE CES DERNIÈRES ANNÉES DONNENT à PENSER QUE POUR L'HEURE LES TECHNIQUES UTILISÉES NE SONT PAS OPTIMALES POUR POUVOIR CONCLURE AVEC CERTITUDE. Par conséquent, il paraît nécessaire de trouver des techniques plus directes afin de mieux appréhender la réPLICATION de l'ADN mitochondrial.

III. Transcription

L'ADN mitochondrial contient un promoteur pour chaque brin, le LSP (Light Strand Promotor) et le HSP (Hight Strand Promotor) mais il y a en fait trois sites d'initiation de la transcription car le HSP contient deux sites d'initiation : HSP1 et HSP2. Ces sites d'initiation ont été identifiés dans des cellules Hela par l'équipe de Attardi (Montoya et al. 1982). Le HSP1 transcrit un ARN polycistronique pour les deux ARNs ribosomiaux : le 12S et le 16S. La terminaison de ce transcrit se fait au niveau du 3' de l'ARN 16S et sa transcription est vingt fois celle du transcrit initié au niveau du HSP2. Le transcrit issu de l'HSP2 code pour le brin H entier et forme un unique ARN polycistronique qui sera ensuite maturé afin de donner les différents ARNm et ARNt. Le LSP permet aussi la synthèse d'un ARN polycistronique et de l'ARN 7S présent dans la D-loop. Les sites de terminaison pour les deux ARNs polycistroniques provenant du HSP2 et LSP sont méconnus actuellement (Figure 9). Les ARNs polycistroniques doivent être ensuite clivés et maturés. Il est proposé que les ARNt servent de signaux de clivage, ce modèle est appelé « tRNA punctuation model » (Ojala et al. 1981). Concernant les ARNm, une queue polyA leur est ajoutée par la Poly(A) Polymérase mitochondriale (mtPAP) (Nagaike et al. 2005).

Les protéines essentielles pour la transcription sont TFAM (mitochondrial transcription factor A), TFB1M et TFB2M (mitochondrial transcription factor B1 et B2) (Falkenberg et al. 2002) ainsi que la POLMRT (mitochondrial DNA-directed RNA polymerase) (Morozov and Temiakov, 2016).

Une étude présente pour la première fois des résultats de séquençage du transcriptome mitochondrial (Mercer et al., 2011). Les auteurs obtiennent des transcrits représentant 99,9%

du brin H et 97,6% sur le brin L, chaque brin étant donc entièrement transcrit. De façon intéressante, ils démontrent que bien que les ARNs soient polycistroniques, les proportions d'ARNs maturés provenant d'un même ARN polycistronique ne sont pas équivalentes, ce qui suggère des stabilités différentes lors de la maturation. Leur étude apporte aussi certaines réponses concernant les sites de terminaison de la transcription, puisqu'ils observent un potentiel site de terminaison pour le brin L environ à 160-185 nucléotides en amont du LSP.

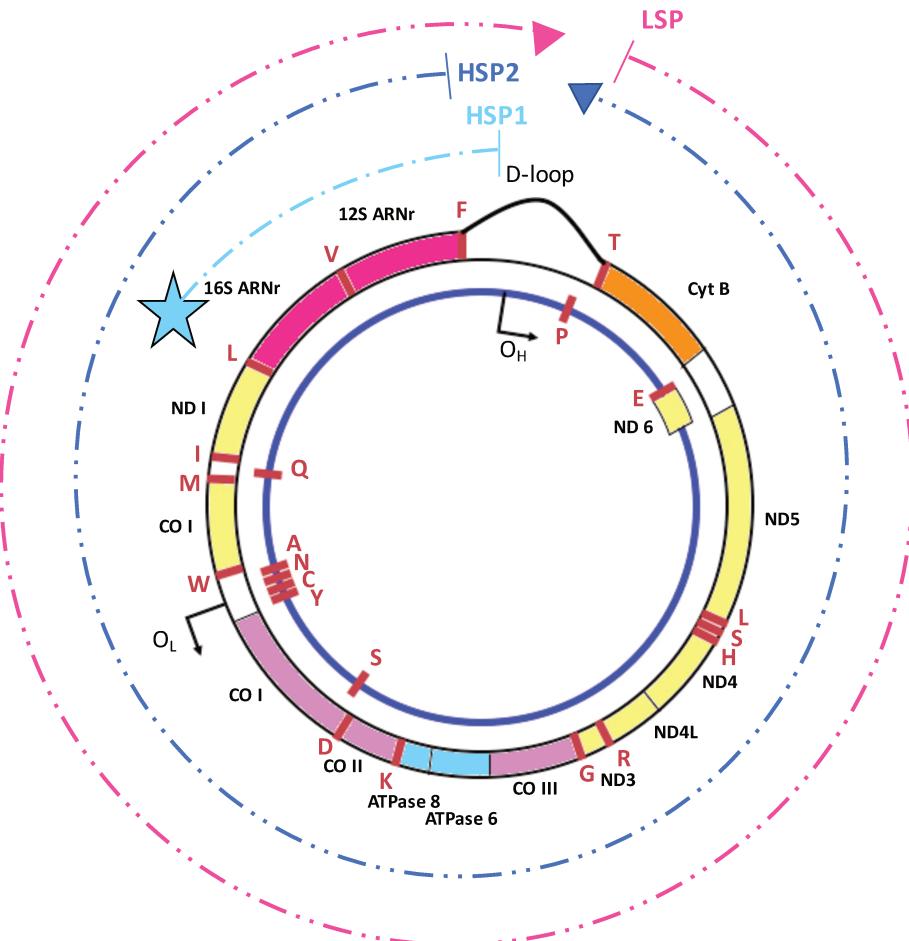


Figure 9 : Schéma représentant les différents transcrits issus de l'ADN mitochondrial
Trois transcrits polycistroniques sont issus de l'ADN mitochondrial à partir de trois promoteurs présents dans la D-loop. Le LSP est le promoteur pour le transcrit du brin L. Pour le brin H, 2 promoteurs sont présents. HSP1 est le promoteur qui initie la transcription des ARNs ribosomiaux 12S et 16S. Sa terminaison a lieu en 3' du 16S (représentée par l'étoile bleue). Le promoteur HSP2 initie un ARN polycistronique qui permet la transcription du brin H. L'endroit de la terminaison pour les transcrits issus de HSP2 et LSP n'est pas clairement établi.

IV. Traduction

Sur les 1500 protéines présentes dans la mitochondrie, seulement 1% seront traduites dans la matrice mitochondriale. Même si cela représente peu, la mitochondrie doit posséder une

machinerie de traduction. De façon originale, il a été observé dans la mitochondrie qu'il existait quelques différences de codons par rapport au code génétique universel (Tableau 1).

Codon	Code génétique universel	Code génétique de la mitochondrie
AUA	Isoleucine	Codon start
UGA	Codon stop	Tryptophane
AGA	Arginine	Codon stop
AGG	Arginine	Codon stop

Tableau 1 : différences de codons entre le code génétique universel et celui retrouvé au sein de la mitochondrie

La traduction se déroule comme au sein du noyau en trois grandes étapes : l'initiation, l'elongation et la terminaison de la traduction. La traduction mitochondriale se fait grâce à un ribosome mitochondrial aidé de nombreux facteurs. Dans la mitochondrie, le ribosome mitochondrial a un coefficient de sédimentation de 55S composé d'une petite sous-unité 28S et d'une grande sous-unité 39S (O'Brien 1971). Pour l'initiation de la traduction, deux facteurs mitochondriaux sont nécessaires, IF2mt et IF3mt (mitochondrial initiation factor 2 ou 3) qui vont aider à charger le premier tARN-Met au niveau de l'ARNm au sein du ribosome (Spencer & Spremulli 2004). L'étape suivante est l'elongation. Trois facteurs semblent y participer : EF-TU, EF-TS, EF-G. La terminaison s'accomplice grâce au facteur de terminaison : le mtRF1 (release factor) qui reconnaît le codon stop et participe à la dissociation du ribosome avec la chaîne peptidique (Cristodero et al. 2013).

Les mécanismes exacts ne sont pas encore complètement décryptés pour chacune des étapes. Concernant la traduction humaine, l'identification des différents facteurs participant à la traduction n'a notamment été faite qu'à partir d'extraits bovins et nous n'avons donc pas une idée précise sur les facteurs présents chez l'homme.

V. Import des protéines

1% des protéines étant traduites dans la matrice mitochondriale, les 99% de protéines restantes présentes dans la mitochondrie, sont codées par l'ADN nucléaire et doivent être importées dans la mitochondrie. Les protéines importées iront soit dans la membrane externe ou bien dans la membrane interne, soit dans l'espace inter-membranaire ou dans la matrice mitochondriale.

Des travaux chez la levure ont montré grâce à des expériences de FISH (Hybridation in situ en fluorescence) que des ARNm codant pour des protéines à importer sont retrouvés à la surface de la mitochondrie, ce qui suggère un mécanisme de rapprochement spatial pour faciliter l'import (Garcia et al. 2007). Chez l'homme il a ensuite été observé le même phénomène que chez la levure, les ARNm codant pour les futures protéines mitochondrielles étant retrouvés au sein de polysomes proches de la mitochondrie (Sylvestre et al. 2003).

Les protéines seront importées sous forme de chaines polypeptidiques non structurées. L'adressage à la mitochondrie des protéines repose sur la reconnaissance d'un signal d'import. Dans la très grande majorité des cas, cette séquence d'import se situe en N terminal de la protéine. Elle possède une charge globale positive grâce à la présence d'acides aminés comme l'arginine et présente une fréquence élevée d'acides aminés hydrophobes tels que la leucine et l'alanine. Ce signal mitochondrial sera clivé à la dernière étape de l'import. Ces propriétés permettent l'identification par bioinformatique du signal d'import mitochondrial potentiel au sein d'une séquence protéique (Negi et al. 2015). Cette séquence d'import peut néanmoins ne pas se trouver en N terminal, et rend l'identification par les outils bioinformatiques plus ardue. C'est par exemple le cas de la protéine APE-1 qui possède son signal d'import en C terminal (Li et al. 2010).

Après traduction proche de la mitochondrie, les protéines non structurées sont amenées près de la membrane externe grâce à des protéines chaperonnes, notamment Hsp70 et Hsp90 (Young et al. 2003). La première étape de l'import est le passage de la membrane externe de la mitochondrie par les protéines via le complexe oligomérique TOM (Translocase of the Outer Membrane). L'élément central du canal TOM est la protéine TOM40 qui permet de former deux canaux au sein de ce complexe (Ahting et al. 1999). Quant aux protéines qui sont adressées dans la matrice, elles empruntent le canal TIM (Translocase of the Inner Membrane) à la sortie du canal TOM (Horst et al. 1995). Le complexe TIM coopère avec celui de la membrane externe, son élément central étant la protéine TIM23 (Berthold et al. 1995) et il utilise le potentiel de membrane créé par la chaîne respiratoire (Bauer et al. 1996). Les protéines sont donc prises en charge via la reconnaissance de leur séquence d'import par le complexe TIM qui est associé au PAM (Presequence translocase-Associated import Motor). Le PAM se lie à la chaîne peptidique à importer et la dirige vers la matrice mitochondriale. Le PAM nécessite un apport d'énergie obtenu grâce à l'hydrolyse de l'ATP en ADP (Ting et al. 2014). La séquence d'import est ensuite clivée lorsque la chaîne polypeptidique arrive dans la matrice par la MPP (Mitochondrial Processing Peptidase) (Taylor et al. 2001). Une fois dans la matrice, les

protéines sont prises en charge par la mtHsp70 pour permettre leur structuration (Berthold et al. 1995) (Figure 10).

Pour les protéines allant dans l'espace intermembranaire, elles sont prises en charge par le complexe MIA (Mitochondrial Intermembrane Space Assembly) à la sortie du TOM (Mordas & Tokatlidis 2015). Concernant les protéines adressées à la membrane interne ou à la membrane externe, elles sont prises en charge à la sortie du TOM respectivement par le complexe TIM22 (Rehling et al. 2003) ou le complexe SAM (Sorting and Assembly Machinery) (Kutik et al. 2008).

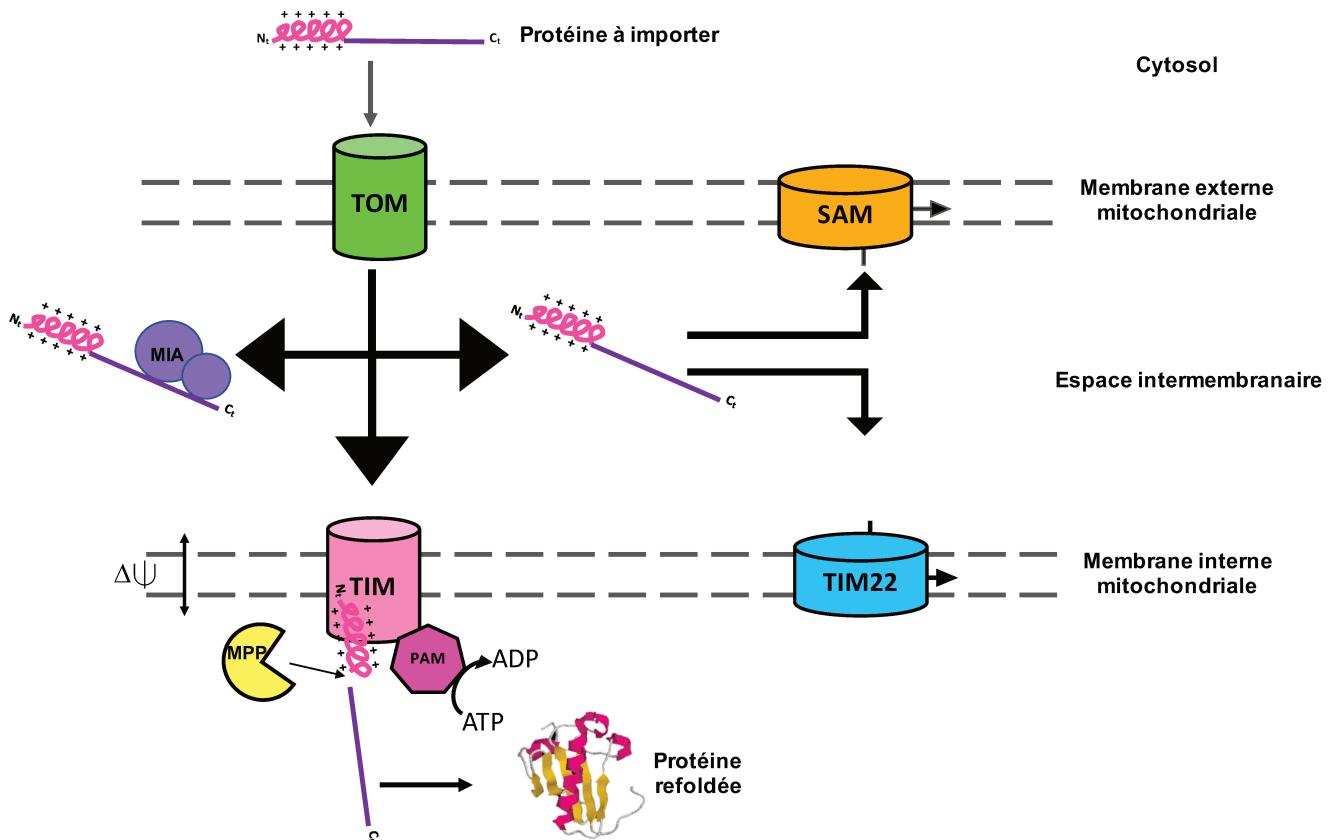


Figure 10 : Schéma représentant les principaux complexes participant à l'adressage des protéines dans les différents compartiments de la mitochondrie

Les protéines à importer sont reconnues sous forme de chaîne polypeptidique grâce à leur signal d'adressage (en rose). Après reconnaissance et passage par le canal TOM (Translocase of the Outer Membrane), le canal TOM va coopérer avec le complexe TIM (Translocase of the Inner Membrane) pour faire passer les protéines vers la matrice mitochondriale. Contrairement au canal TOM, le complexe TIM aidé du PAM nécessite de l'énergie. Pour les protéines qui sont adressées dans l'espace intermembranaire, dans la membrane externe ou interne, elles sont prises en charge à la sortie du TOM soit par le complexe MIA (pour l'espace intermembranaire) soit grâce au complexe SAM (pour la membrane externe) soit avec l'aide du complexe TIM22 (pour la membrane interne).

VI. Quelques protéines essentielles de la mitochondrie

1) La polymérase POLG

La polymérase POLG est, à ce jour, la seule polymérase identifiée dans la mitochondrie. Elle a été identifiée pour la première fois en 1977 dans des cellules Hela (Bolden et al. 1977). Elle est codée par le gène *POLG* et composée de deux sous-unités, α et β . La sous-unité α de 140 kDa comporte une activité exonucléase 3' vers 5' (en N terminal), une activité 5'dRP lyase et une activité polymérase (en C terminal) (Graves et al. 1998). La sous-unité β de 55 kDa participe en effet à la liaison à l'ADN mais aussi à la processivité de l'enzyme (Lim et al. 1999). La cristallisation de la sous-unité β a révélé qu'elle interagit en homodimère et participe à la liaison à l'ADN (Fan et al. 2006). POLG est impliquée dans la réPLICATION de l'ADN mitochondrial, en faisant partie du réplisome minimal (Korhonen et al. 2004) et dans la réparation notamment dans la voie de la réparation par excision de base (BER) (Longley et al. 1998).

2) L'hélicase Twinkle

La protéine Twinkle a été identifiée pour la première fois en 2001 comme étant une hélicase présente dans la mitochondrie (Spelbrink et al. 2001). Cette hélicase fonctionne de 5' vers 3' et a besoin d'ATP. Elle possède un domaine primase en N terminal et un domaine hélicase en C terminal. Elle fonctionne en hexamère (Korhonen et al. 2003). Elle fait partie du réplisome minimal (Korhonen et al. 2004). En effet, *in vitro*, Twinkle agit sur les intermédiaires de réPLICATION ainsi que sur la structure de la D-loop (Khan et al., 2016). Des travaux (Milenkovic et al. 2013) démontrent que Twinkle est nécessaire dans l'embryogénèse chez la souris, les embryons KO pour Twinkle n'étant pas viables. Il est alors utilisé des souris avec un KO inductible uniquement dans les cardiomyocytes. Ces expériences mettent en évidence que chez une souris avec un KO Twinkle induit uniquement dans ses cellules cardiaques, l'ADN mitochondrial restant n'est plus que de 6% à 16 semaines de vie. Comme observé *in vitro*, ces résultats tendent aussi à prouver que Twinkle joue un rôle au niveau de la D-loop, dans la régulation de la réPLICATION et de la transcription car le KO de Twinkle perturbe la quantité d'ADN 7S et de transcrits.

3) La protéine mitochondriale de liaison à l'ADN simple brin : mtSSB

La protéine mitochondriale SSB fonctionne en tétramère et se lie à l'ADN simple brin (Tiranti et al. 1993) (Yang et al. 1997) (Oliveira & Kaguni 2010). Elle fait partie du réplisome

minimal en stimulant la polymérase POLG et l'hélicase Twinkle (Korhonen et al. 2004). Une étude en cellules humaines (cellules Hela) utilisant un siARN contre mtSSB constate une diminution de 40% de l'ADN mitochondrial (comme pour les autres membres du réplisome) reflétant une inhibition de la réPLICATION. De surcroit, comme pour les autres protéines de la réPLICATION, après trois jours d'inhibition de mtSSB, il n'y a seulement que 20% d'ADN 7S restant (Ruhanen et al. 2010).

4) La protéine endo/exonucléase MGME1

La protéine MGME1 (mitochondrial genome maintenance exonuclease1) est codée à partir du gène *C20orf72* (Szczesny et al. 2013). Cette équipe identifie chez cette protéine une séquence d'adressage potentielle à la mitochondrie et transfecTE le cDNA afin de valider la localisation de cette protéine à la mitochondrie. Les tests *in vitro* qui ont été réalisés témoignent que cette protéine est une exonucléase qui dégrade spécifiquement un substrat d'ADN simple brin de 3' en 5' et qui n'est pas capable de dégrader de l'ARN ou de l'ADN circulaire. Elle semble aussi posséder une activité endonucléase car elle peut cliver l'ADN d'un substrat ARN-ADN-ARN. Concernant leurs résultats en cellules, l'inhibition par siARN de la protéine provoque une augmentation de la quantité d'ADN 7S alors qu'une surexpression conduit à une diminution en ADN 7S. Ces travaux identifient donc cette nouvelle nucléase de la mitochondrie avec un potentiel rôle dans la régulation de l'ADN 7S (Szczesny et al. 2013). Des travaux publiés au même moment réalisent les mêmes observations sur les fonctions de la protéine MGME1 (Kornblum et al. 2013). De plus, des protéines MGME1 possédant des mutations non-sens identifiées chez des patients entraînent une diminution du nombre de copies d'ADN mitochondrial et une augmentation de délétions comparables à celles observées avec des protéines POLG mutées (Kornblum et al. 2013). Les travaux de l'équipe de Falkenberg (Uhler et al. 2016) montrent que MGME1 peut agir sur les structures Flap. Les auteurs reconstituent un système minimal qui mime la réPLICATION. Cela leur permet de constater le rôle primordial de MGME1 pour résoudre la structure Flap présente lors de la réPLICATION.

5) L'hélicase/nucléase DNA2

L'hélicase/nucléase DNA2 est une protéine de 120 kDa qui possèDE un domaine nucléase et ATPase ainsi qu'un domaine hélicase. Une première publication (Zheng et al. 2008) conclut à la présence de la protéine DNA2 uniquement dans la mitochondrie. Ces résultats montrent que la protéine est retrouvée au niveau de l'ADN mitochondrial (colocalisation avec le marquage de l'ADN mitochondrial). Peu après, d'autres travaux contredisent cette

conclusion affirmant que DNA2 est présente aussi au sein du noyau (Duxin et al. 2009). Ces deux études s'accordent sur le rôle de DNA2 agissant sur les structures Flap qui sont présentes lors de la réPLICATION et de la réPARATION par excision de base du « long patch BER ». Leurs expériences montrent que DNA2 clive les structures Flap, en tant qu'endonucléase. Il est aussi proposé que DNA2 en interagissant avec POLG stimule l'activité de cette dernière d'après des expériences de coimmunoprecipitation (Zheng et al. 2008). Lors de dommages induits par un traitement avec de l'H₂O₂, quand DNA2 est inhibée, il est observé une baisse concomitante de la réparation et de la réPLICATION de l'ADN mitochondrial (Zheng et al. 2008). L'inhibition de DNA2 fait aussi état d'une diminution de la réPLICATION par analyse en électrophorèse 2D (Zheng et al. 2008). Ces différents travaux attribuent donc une fonction dans la réPLICATION et dans la réPARATION à la protéine DNA2 en tant qu'endonucléase sur les structures Flap.

6) Le facteur de transcription A mitochondrial : TFAM

La protéine TFAM (mitochondrial transcription factor A) d'environ 25 kDa a été purifiée et identifiée pour la première fois comme un facteur de transcription en 1988. Elle peut se lier aux deux promoteurs, le LSP et le HSP mais aussi à l'ADN de façon non spécifique (Fisher & Clayton 1988) (Litton et al. 2010). La présence de TFAM permet d'augmenter l'efficacité de transcription (Fisher & Clayton 1988) . Des années plus tard, une autre fonction pour la protéine TFAM est suggérée. En effet, elle pourrait former les structures nucléoïdes de l'ADN mitochondrial, car il est démontré que dans les cellules Hela, 900 molécules de TFAM se lient à une molécule d'ADN et que ces foyers interagiraient avec les membranes internes de la mitochondrie (Alam et al. 2003). Quelques années plus tard, l'équipe de Shoubridge (Kaufman et al. 2007) identifie TFAM comme un dimère qui a les capacités de compacter l'ADN mitochondrial, et qui serait à la base de la formation de structures nucléoïdes contenant plusieurs copies d'ADN mitochondrial.

7) La polymérase ARN mitochondriale : POLMRT

La polymérase POLMRT est la polymérase ARN de la mitochondrie. Sa taille est d'environ 126 kDa et son domaine polymérase se trouve dans sa partie C terminale (Tiranti et al. 1997) (Ringel et al. 2011). La polymérase POLMRT est impliquée à la fois dans la transcription de l'ADN mitochondrial mais aussi dans la réPLICATION du brin L.

Concernant la transcription, une interaction a été observée entre POLMRT et TFAM, et la structure du complexe de transcription a été obtenue avec les facteurs TFM1 et 2 au niveau du promoteur (Yakubovskaya et al. 2014). Il a été montré que POLMRT reconnaît

spécifiquement les séquences promotrices (la séquence murine n'étant pas reconnue par la protéine humaine) (Gaspari et al. 2004).

Au niveau de sa fonction dans la réPLICATION, une étude montre en cellules et *in vitro* que POLMRT permet l'initiation de la réPLICATION au niveau de l'origine O_L grâce à son rôle de primase. Cette étude met en évidence que la POLMRT reconnaît une série de T au niveau de l' O_L et qu'après 25 nucléotides en cellules et 25 à 75 nucléotides *in vitro*, POLG prend le relai de POLMRT (Fusté et al. 2010).

VII. L'import des ARNs

Il semblerait que chez l'Homme plusieurs types d'ARN soient importés dans la mitochondrie (Kim et al. 2017). Il a été observé que des ARNs non codants (Mercer et al. 2011), des ARNt provenant du noyau (Rubio et al. 2008), l'ARNr 5S et l'ARN MRP (Wang et al. 2010) seraient importés au sein de la mitochondrie. Cependant, à l'heure actuelle, chez l'Homme, il n'existe pas de voie d'import des ARNs qui soit décrite.

L'équipe de Tarassov (Kolesnikova et al. 2000) réussit pour la première fois à importer un ARNt dans des mitochondries humaines. L'équipe de Alfonzo (Rubio et al. 2008) démontre *in vitro* l'import du ARNt^{Gln} dans des mitochondries purifiées. Ses expériences mettent en évidence que le mécanisme d'import est ATP dépendant mais n'est pas lié au potentiel de membrane. Ces résultats suggèrent une voie d'import différente de celle des protéines. L'équipe de Tarassov (Kolesnikova et al. 2000) avait pourtant conclu à l'utilisation du potentiel de membrane pour permettre l'import. L'équipe de Teitell (Wang et al. 2010) évoque que la polynucleotide phosphorylase (PNPase), retrouvée au sein de l'espace intermembranaire, participerait à l'import des ARNs, notamment concernant l'ARN de la RNase P. La voie d'import reste donc encore très mal comprise.

Quelques caractéristiques des ARNs importés ont pu être décrites dans les travaux de Wang (Wang et al. 2012). Ses expériences ont permis d'identifier des séquences d'adressage, dont une structure en loupe en 5' de l'ARN semblant faciliter l'import, ainsi qu'une séquence en 3'UTR permettant d'approcher l'ARN à la surface de la mitochondrie. Des études utilisent des ARNs comme outil thérapeutique afin d'éliminer des copies délétées (Comte et al. 2013) ou encore afin de compenser un ARNt muté (Karicheva et al. 2011). Par exemple, afin de bloquer la réPLICATION de copies d'ADN mitochondrial délétées impliquées dans le syndrome de Kearns Sayre, l'import d'un ARN modifié qui reconnaît l'ADN mitochondrial muté a été utilisé. Deux

structures boucles en 5' et en 3' provenant de l'ARN^{lys} de la levure ont été ajoutées sur cet ARN afin de permettre l'import (Comte et al. 2013) (Figure 11).

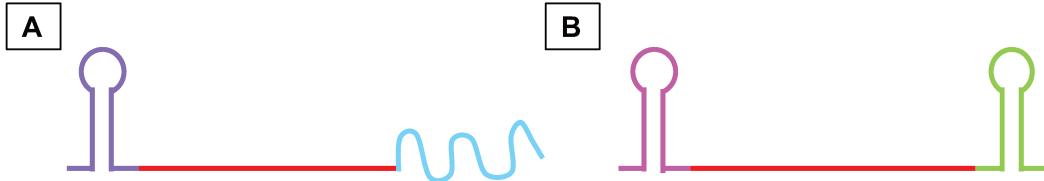


Figure 11 : Schéma représentant les ARN modifiés afin de permettre leur import dans la mitochondrie

A) l'ARN (en rouge) possède une loop en 5' (violet), identifiée sur l'ARN de la RNase P et une séquence en 3' (bleue), identifiée sur l'ARN de la MRP, pour permettre son import à la mitochondrie (Wang et al., 2012). B) l'ARN (en rouge) possède une loop en 5' (rose) et une loop en 3' (vert) identifiée sur l'ARN^{lys} de la levure (Comte et al. 2013).

VIII. La réparation de l'ADN mitochondrial

1) Les dommages de l'ADN mitochondrial

Comme au sein du noyau, l'ADN mitochondrial est exposé à différents facteurs endogènes ou exogènes à la cellule pouvant induire différents types de dommages. Les sources exogènes de dommages sont par exemple les ultraviolets (UV), les rayonnements ionisants et certains produits chimiques ou médicaments. Les sources de dommage endogènes à la cellule sont les espèces réactives de l'oxygène (ROS), certains produits secondaires du métabolisme et des erreurs lors de la réPLICATION. Les conséquences de ces dommages sur l'ADN vont entraîner des modifications des bases de l'ADN, des sites abasiques, des adduits, des cassures sur un brin de l'ADN, des cassures double-brin ou des mésappariements de bases (Figure 12). Ces dommages doivent être réparés le plus fidèlement possible pour maintenir l'intégrité de l'ADN et la survie cellulaire. Ainsi, comme l'ADN nucléaire, l'ADN mitochondrial doit être réparé afin de limiter l'apparition de mutations délétères. Cependant, la multiplicité du nombre de copies de mtADN permet d'envisager la possibilité que chaque copie endommagée n'est pas forcément toujours réparée.

Du fait de la proximité de l'ADN mitochondrial avec la chaîne respiratoire, il est majoritairement admis que la principale source de dommage pour l'ADN mitochondrial serait due aux ROS, produits du métabolisme respiratoire. Le stress oxydatif qui en découle induirait des modifications de base (Richter et al. 1988). La base modifiée le plus souvent incorporée dans l'ADN est la 8-oxoguanine. Le principal problème de cette base est qu'elle peut entraîner

des transversions des bases G-C en A-T. Les modifications de bases ont donc un fort pouvoir mutagène (comme revue voir (Mandavilli et al. 2002)).

Pour limiter l'apparition de mutations au sein de son génome mitochondrial, la cellule doit posséder des voies de réparation. Actuellement, les mécanismes de réparation ne sont pas entièrement élucidés. Seule la voie par excision de base (BER) est bien décrite. Les autres voies de réparation proposées sont : la réparation des mésappariements (MMR), la recombinaison homologue, le Non Homologous End Joining (NHEJ). Une autre stratégie potentiellement utilisée par la cellule serait de dégrader les copies endommagées au lieu de les réparer.

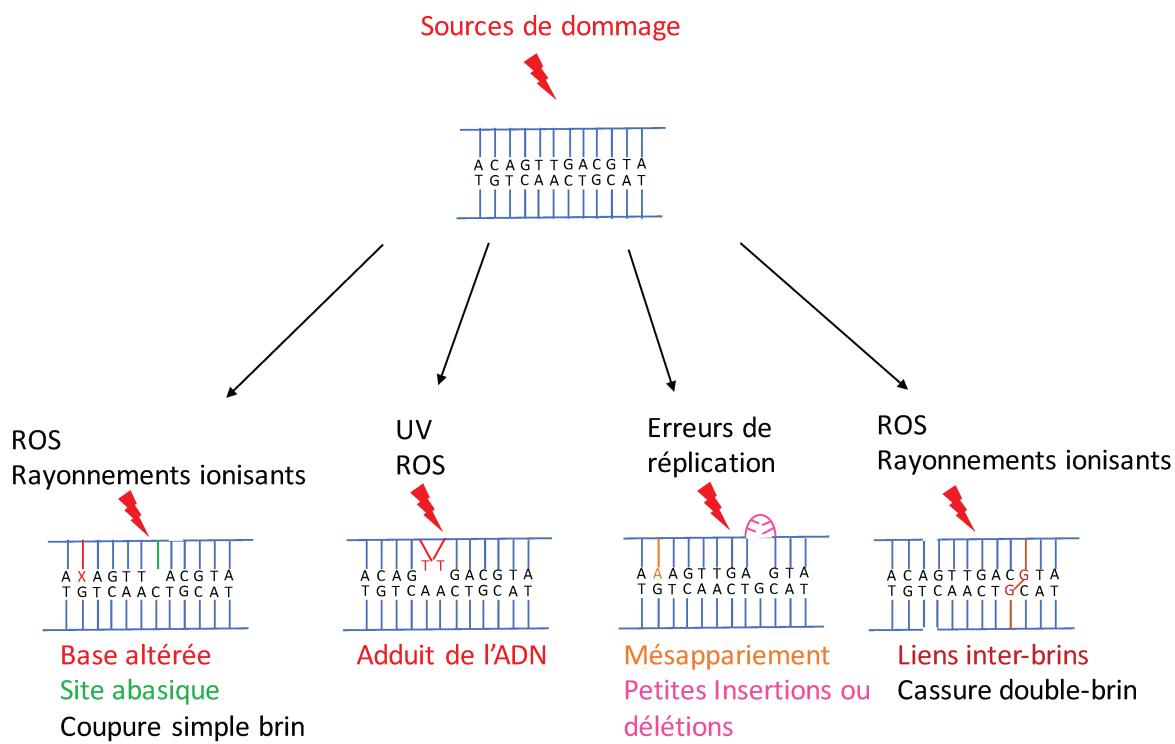


Figure 12 : Les différentes sources de dommages à l'ADN et leurs conséquences

Les espèces réactives de l'oxygène (ROS), les rayonnements ultra-violets (UV), les rayonnements ionisants ainsi que les erreurs de réPLICATION peuvent entraîner des modifications au sein de l'ADN qui doivent être réparées. Ces modifications sont des altérations de bases (en orange), des sites abasiques (en vert), des adduits de l'ADN (en rouge), des mésappariements (en marron), des délétions ou insertions (en rose), des liens inter-brins (en bordeaux), des cassures sur un brin ou sur les deux brins.

2) Les différentes voies de réparation au sein de la mitochondrie

a) La réparation par excision de base (BER) dans la mitochondrie

La réparation par excision de base (BER) est la voie de réparation de l'ADN mitochondrial la mieux décrite. Cette voie est utilisée pour enlever des bases modifiées au sein de l'ADN. Cette voie fonctionne comme le BER présent dans le noyau.

Le BER se déroule en quatre grandes étapes. La première étape est la reconnaissance et l'élimination de la base modifiée. La seconde étape est l'élimination du site abasique (AP) engendré par le retrait de la base endommagée. L'étape suivante est de combler la brèche formée puis la dernière étape est celle de la religation au niveau des bases ajoutées. Il existe deux sous-voies dans le BER au niveau de la troisième étape, soit le « short patch » dans le cas de l'élimination d'une seule base, soit le « long patch » si plusieurs bases consécutives sont éliminées (de 2 à 6 nucléotides).

L'étape 1 est réalisée par des glycosylases qui peuvent être monofonctionnelles ou bifonctionnelles. Les glycosylases monofonctionnelles enlèvent la base endommagée en laissant un site abasique (AP) au sein de l'ADN. Les glycosylases bifonctionnelles enlèvent aussi la base endommagée mais possèdent en plus leur propre activité lyase éliminant le site abasique, elles réalisent alors l'étape 1 et 2. Dans la mitochondrie, trois glycosylases monofonctionnelles ont été identifiées : l'uracil DNA glycosylase (UNG1), l'alkyladenine DNA glycosylase (AAG), la «MutY homolog» DNA glycosylase (MUTYH) (Ohtsubo et al. 2000) et 4 glycosylases bifonctionnelles : la 8-oxoguanine DNA glycosylase (OGG1) (Takao et al. 1998), la «NTH endonucleaseIII-like 1» glycosylase (NTH1) (Takao et al. 1998), la «NEIL like 1» glycosylase (NEIL1) (Holt & Reyes 2012b) (Hu et al. 2005) et la «NEIL like 2» glycosylase (NEIL2) (Mandal et al. 2012). L'étape 2 est réalisée par l'apurinic/apyrimidinic (AP) endonucléase (APE-1). Elle a été identifiée dans la mitochondrie (Chattopadhyay et al. 2006). Elle reconnaît le site abasique et laisse le 3'OH et le 5'P de part et d'autre de la base manquante. L'étape 3 est réalisée par la polymérase POLG. Dans le cas du « short patch », après l'action de POLG, la ligase 3 réalise la ligation de la nouvelle base. Dans le cas du « long patch », après le remplissage des plusieurs bases manquantes, une structure Flap est formée et est éliminée par l'exonucléase EXOG ou par l'action concomitante des protéines DNA2 et

FEN1. Comme dans le « short patch », l'étape de ligation est réalisée par la ligase 3 (Figure 13).

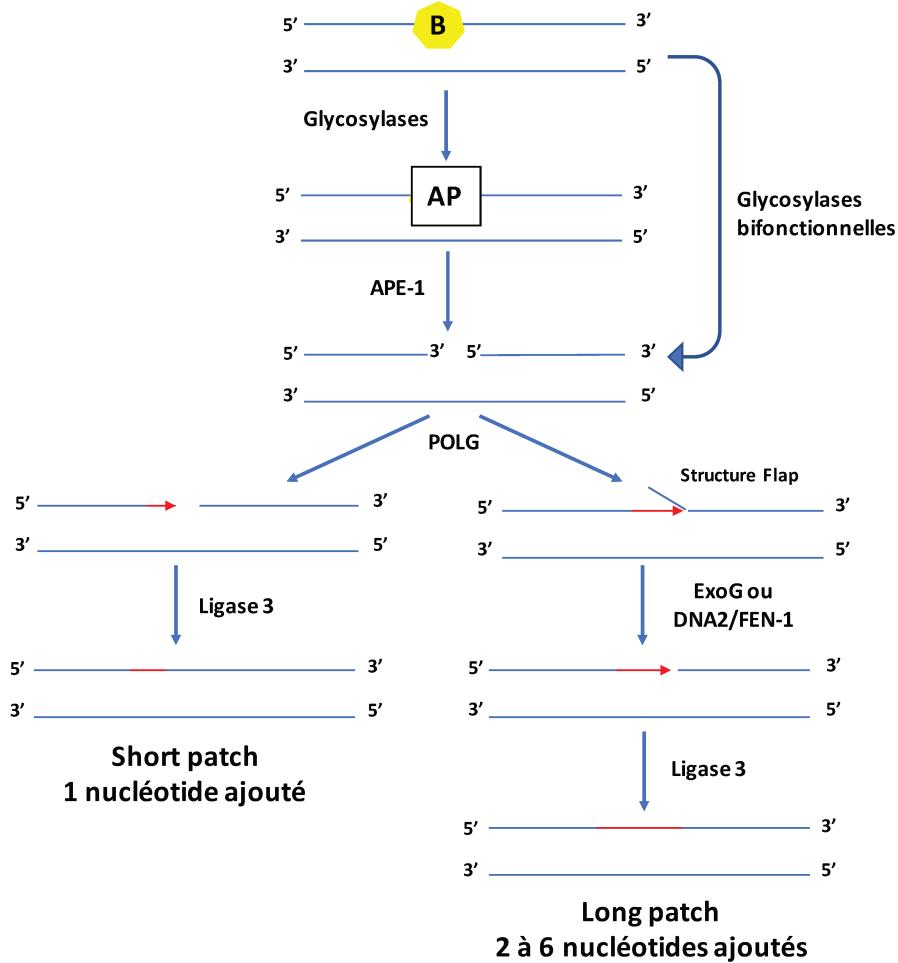


Figure 13 : Schéma récapitulatif de la voie de réparation par excision de base (BER)

Après élimination de la base endommagée par les glycosylases, le site abasique (AP) créé est éliminé par APE-1. La polymérase POLG va combler la brèche. Le short patch BER est utilisé lorsqu'une seule base est remplacée tandis que le long patch est utilisé pour le remplacement de 2 à 6 nucléotides. Après l'ajout du(des nucléotide(s), la ligation est effectuée par la ligase 3.

Il est considéré que le BER est la voie la plus utilisée pour la réparation car l'ADN mitochondrial est plus exposé aux dommages oxydatifs en raison de sa proximité avec la chaîne respiratoire. Ceci est remis en question par certaines études qui s'opposent à ce dogme et suggèrent que la proportion plus importante de 8 oxo-G par rapport au noyau a été surestimée (Beckman & Ames 1999). Ces études alertent sur les méthodes d'extraction et d'analyse de l'ADN mitochondrial qui peuvent induire un biais. Par exemple, lorsqu'il y a préalablement une purification de mitochondrie les taux de mutations sont plus élevées que sur une extraction totale (Anson et al. 2000). De plus, l'équipe de Epe (Trapp et al. 2007) montre que des souris déficientes pour la glycosylase OGG1 ne présentent pas d'augmentation significative de bases

endommagées. Il faut aussi rester critique sur la présence de certaines protéines de réparation à la mitochondrie. En effet, la majorité des publications démontrant la présence des protéines du BER au sein de la mitochondrie, notamment celles concernant les glycosylases, reposent sur des analyses bioinformatiques recherchant un potentiel signal d'adressage à la mitochondrie et/ou sur leur présence dans des extraits mitochondriaux. Or la préparation d'extraits mitochondriaux reste délicate et est facilement contaminée par d'autres compartiments cellulaires. La réparation des dommages oxydatifs dans la mitochondrie, bien qu'elle soit très étudiée, nécessite encore une meilleure compréhension et doit être discutée avec prudence étant donné le nombre d'études présentant des résultats différents.

b) La réparation par Excision de Nucléotides (NER) dans la mitochondrie

Dans le noyau, la réparation par excision de nucléotides (NER) est le système utilisé pour la réparation des adduits, comme les dimères de thymines, provoqués notamment par les rayonnements ultraviolets (UV). Dans la mitochondrie, il ne semble pas que cette voie de réparation soit utilisée. Dès les années 70, l'équipe de Friedberg (Clayton et al. 1974) n'observe pas la réparation des adduits provoqués par les UV au sein de l'ADN mitochondrial et conclut à l'absence de NER au sein de la mitochondrie. Après exposition aux UV puis action de la T4 endonuclease V, les auteurs analysent la proportion d'ADN présentant des brèches (nicks). En effet, en présence d'un adduit, la T4 endonuclease V va convertir cette lésion en nick. Les auteurs ont testé différentes lignées cellulaires humaines et murines et conviennent de l'absence de NER dans tous ces types cellulaires. L'équipe de Bohr étudie l'impact des rayonnements UV sur l'ADN mitochondrial dans les cellules de hamster CHO (Ledoux et al. 1992). Après exposition aux UV, l'ADN mitochondrial a été purifié et traité comme précédemment par la T4 endonucléase V. Leurs résultats confirment, de la même façon que les travaux précédents, l'absence de NER au sein de la mitochondrie.

Il est maintenant largement accepté que la mitochondrie ne possède pas la voie de réparation du NER.

c) La réparation des mésappariements (MMR) dans la mitochondrie

La présence de la voie de réparation des mésappariements (MMR) au sein de la mitochondrie fait encore débat. Dans le noyau, cette voie de réparation est utilisée pour corriger une erreur d'incorporation en face de la base complémentaire, par exemple quand un T est incorporé face à un G. La réparation des mésappariements est aussi utilisée dans le cas où des

bases supplémentaires ont été intégrées sur un des brins créant une petite boucle sur un des deux brins. Seule l'équipe de Lightowers (Mason et al. 2003) s'accorde sur l'existence d'une voie de réparation MMR dans la mitochondrie dans des extraits mitochondriaux de cellules de foie de rat. Toutefois, ces travaux reposent uniquement sur des tests *in vitro* à partir de préparation d'extraits mitochondriaux. La pureté du fractionnement n'est pas présentée, et cette méthode de préparation reste extrêmement délicate pour être parfaite au point de n'avoir absolument aucune contamination par des protéines nucléaires de réparation (dans ce cas). De plus, les auteurs ne détectent pas la protéine Msh2, protéine essentielle du MMR au sein du noyau. Si la réparation par MMR est utilisée au sein de la mitochondrie, cela serait donc grâce à d'autres protéines que celles identifiées dans le noyau. Les protéines de la voie du MMR dans la mitochondrie n'ont toujours pas été identifiées chez l'Homme à ce jour.

d) Le Non Homologous End Joining (NHEJ) dans la mitochondrie

i. Le NHEJ classique

Dans le noyau, la voie du Non Homologous End joining (NHEJ) peut être utilisée pour la réparation des cassures double-brin tout au long du cycle cellulaire. Cette voie relie les deux extrémités cassées avec ou sans modifications de séquences. Cette voie se décline en trois grandes étapes : 1-reconnaissance de la cassure double-brin (CBD), 2-maintien à proximité et nettoyage/maturation si besoin des deux extrémités, 3-religation (Lieber 2008).

Ku est la protéine qui reconnaît la cassure double-brin. Cette protéine de deux sous-unités (Ku70 et Ku86) forme un anneau autour de l'ADN cassé. Ku permet de garder les deux extrémités de la cassure à proximité et va interagir et stimuler les protéines nécessaires à la réparation.

Après la fixation de Ku, la DNA-PKcs (DNA-dependent protein kinase) arrive au niveau de la cassure. Sa liaison à l'ADN lui permet de s'activer par autophosphorylation. Il semblerait que la DNA-PKcs puisse se lier indépendamment de Ku ou interagir avec Ku et former le complexe Ku/DNA-PK. Ku glisse alors vers l'intérieur de l'ADN pour laisser la DNA-PKcs au niveau de la cassure (Yaneva et al. 1997). La DNA-PKcs s'autoactive en se liant à l'ADN par phosphorylation. Elle peut activer de très nombreuses protéines par phosphorylation. L'ensemble de ses protéines cibles n'est pas encore démontré *in vivo*.

Ku va ensuite permettre de recruter le complexe XRCC4-DNA LIGASE 4 (X-ray repair cross-complementing protein 4) et de stimuler l'activité de ligation du complexe (Nick McElhinny et al. 2000). Plus récemment, il a été démontré que la protéine XLF/Cernunnos

(XRCC4 like factor) participe au complexe XRCC4-DNA LIGASE 4 (Ahnesorg et al. 2006) (Buck et al. 2006).

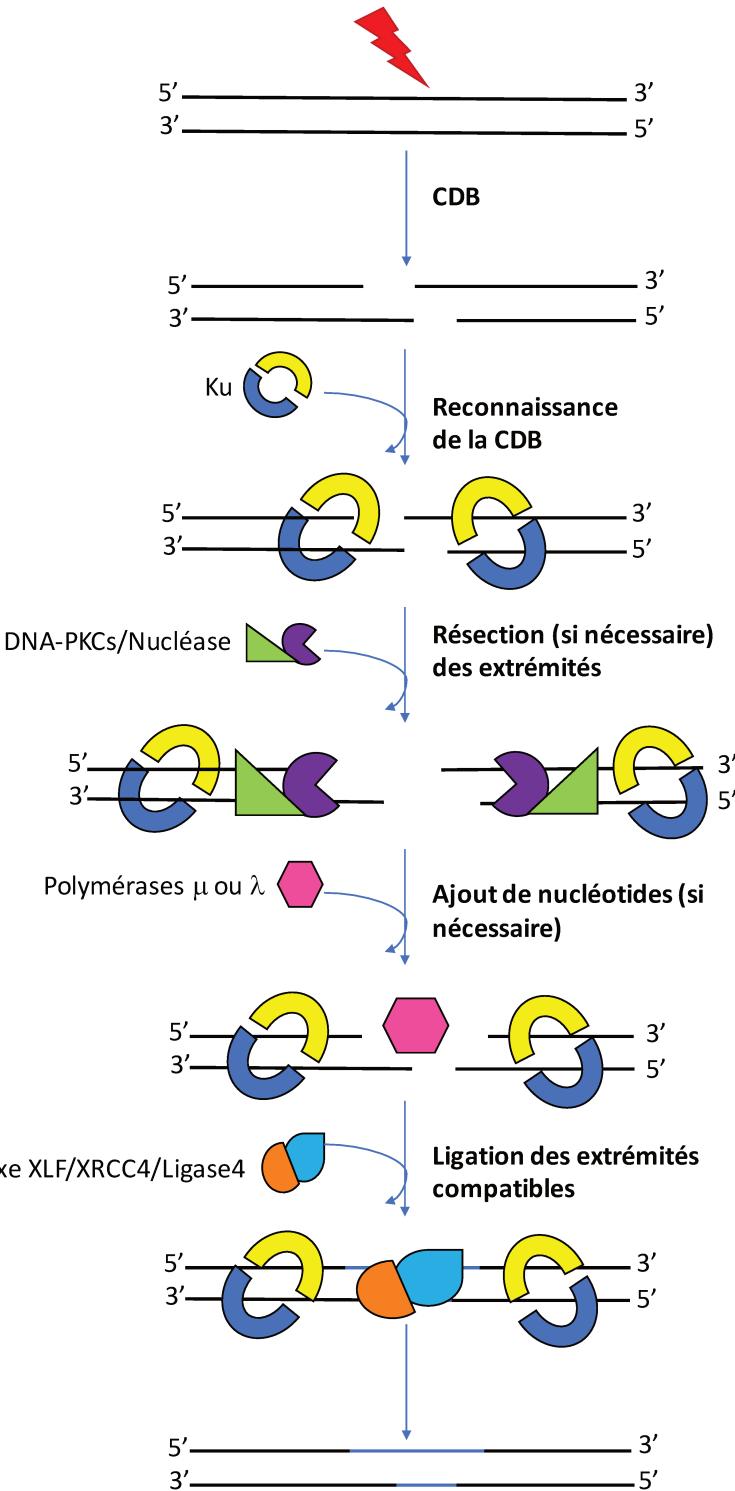


Figure 14 : Schéma représentant la voie du NHEJ classique

Après un dommage entraînant une cassure double-brin, Ku se lie aux extrémités cassées. La DNA-PKcs s'autophosphoryle et active des nucléases telle que Artémis si les extrémités ne sont pas compatibles directement pour la religation. Si nécessaire, les polymérasases μ et λ combinent les brèches. Une fois, les extrémités rendues compatibles, le complexe XRCC4/XLF/Ligase4 réalise la ligation finale.

Lorsque les extrémités ne sont pas directement religables, des nucléases ou des polymérases vont être nécessaires pour rendre les extrémités compatibles. Une des nucléases identifiées est Artémis (Moshous et al. 2001). Une étude *in vitro* (Ma et al. 2004) purifie les principaux acteurs du NHEJ et montre que le complexe DNA-PKcs/Artemis peut réséquer les extrémités après cassures et que les polymérases Pol μ et Pol λ ajoutent si nécessaire des nucléotides aux extrémités (Figure 14).

L'activité du NHEJ classique dans la mitochondrie n'est pas clairement démontrée et semble peu probable. En effet, au sein de la mitochondrie, une seule DNA ligase, la DNA LIGASE 3, est présente (Simsek et al. 2011). La DNA LIGASE 4, une des protéines clés du NHEJ, est donc manquante. De plus, la présence des autres protéines du NHEJ n'a pas été clairement démontrée, même si certaines ont été suggérées comme ayant une localisation mitochondriale, comme par exemple XRCC4 (Wisnovsky et al. 2016). La réparation des cassures double-brin au sein de la mitochondrie via le NHEJ classique fait donc encore débat.

ii. Le NHEJ alternatif

Au sein du noyau, il a été détecté une activité résiduelle de type NHEJ lorsque le NHEJ classique est inhibé. Cette voie a été décrite comme la voie alternative du NHEJ (alt-NHEJ) (Boulton & Jackson 1996) (Kabotyanski et al. 1998). Elle est dite moins fidèle que le NHEJ car au niveau du site de cassure plus de délétions sont retrouvées. Des microhomologies sont fréquemment utilisées. Une microhomologie est une séquence d'au moins 2 nucléotides homologues. Dans la voie du alt-NHEJ, l'utilisation de microhomologies de part et d'autre de la CBD, peut participer à stabiliser les extrémités cassées en s'appariant. Cette voie est maintenant considérée comme une voie présente en permanence dans les cellules et pas seulement en voie de secours lorsque le NHEJ classique n'est plus fonctionnel. Toutes les protéines de cette voie n'ont pas été identifiées clairement. L'inhibition de Ku favorise l'utilisation du alt-NHEJ alors que sa présence l'inhibe (Bennardo et al. 2008). L'initiation de cette voie repose sur une compétition entre les protéines Ku et PARP-1 (Poly (ADP-ribose) polymerase 1) (Wang et al. 2006). L'initiation du alt-NHEJ repose donc sur la liaison de PARP-1 aux extrémités cassées. Comme Ku, PARP-1 se lie directement aux extrémités cassées mais avec deux fois moins d'affinité. Le complexe MRN (Mre11-Rad50-Nbs1) et la nucléase CtIP participent ensuite à la résection (Truong et al. 2013).

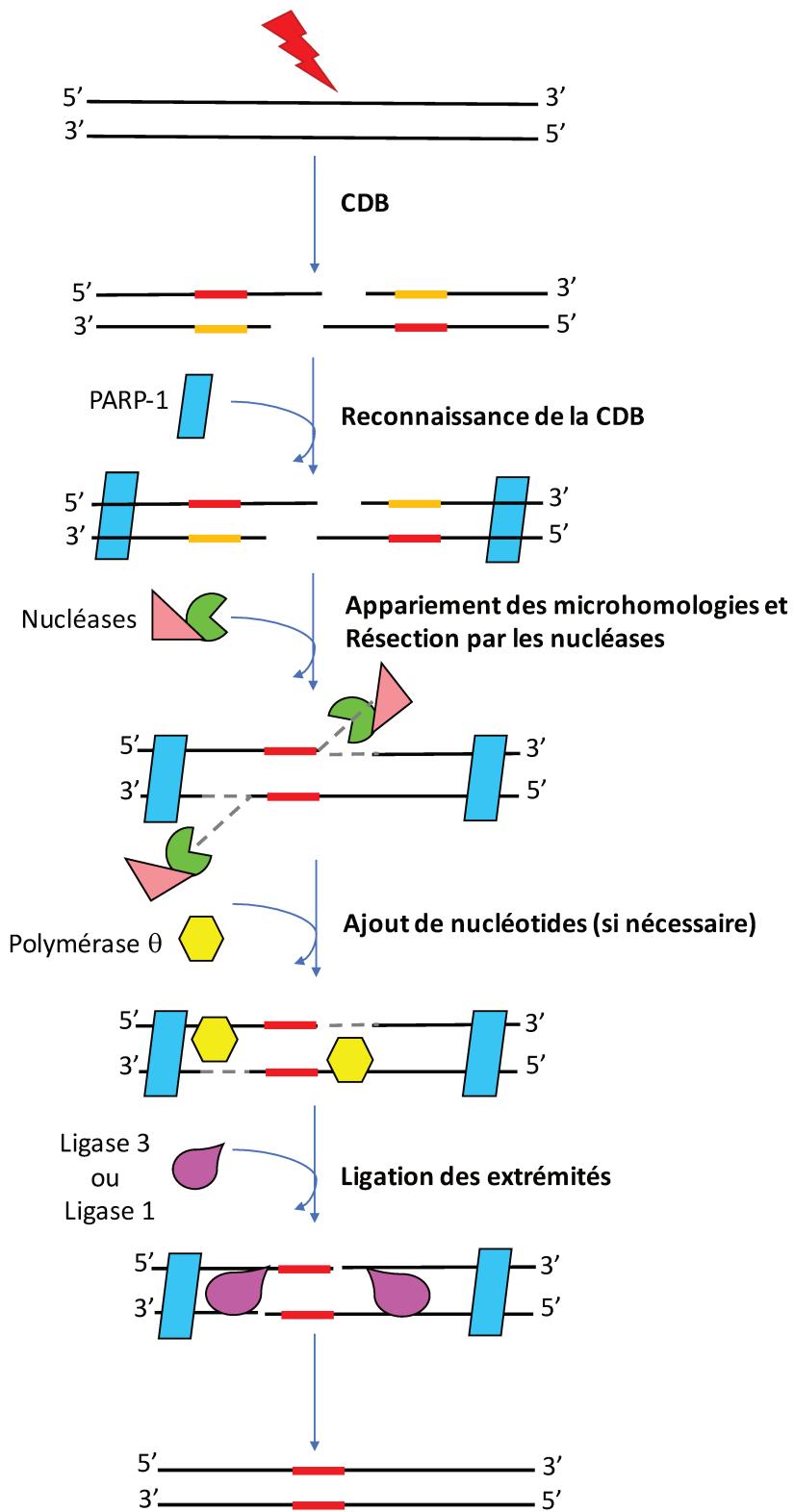


Figure 15 : Schéma présentant la voie alternative du NHEJ

La protéine PARP-1 reconnaît les brins cassés. Les brins sont réséqués par des nucléases telles que le complexe MRN et la protéine CtIP afin de permettre l'appariement des microhomologies (en rouge et orange). L'ajout de nucléotides est réalisé par la polymérase θ. L'étape finale est la ligation par la Ligase 3 ou la Ligase 1.

Pour combler les brèches engendrées par la résection, la polymérase θ pourrait être la polymérase utilisée (Kent et al. 2015). L'étape finale passe par la religation des deux brins de l'ADN. Elle est effectuée par la DNA LIGASE 3 voire la DNA LIGASE 1 (dans le cas où on ne retrouve pas de microhomologies) (Audebert et al. 2004) (Figure 15).

Dans la mitochondrie, seule la DNA LIGASE 3 est présente et son isoforme mitochondriale est essentielle à la survie cellulaire (Simsek et al. 2011). De plus, 85% des délétions identifiées dans l'ADN mitochondrial présentent des microhomologies de part et d'autre de leur séquence. Ces découvertes pourraient suggérer l'utilisation du alt-NHEJ par la mitochondrie dans le cas des cassures double-brin.

L'équipe de Raghavan (Tadi et al. 2016) a donc cherché l'activité du alt-NHEJ au sein de la mitochondrie et conclut à son utilisation à partir d'expériences sur des extraits mitochondriaux de rat. Les auteurs ont obtenu les mêmes résultats sur les cellules humaines Hela (données non publiées). Ses travaux sont réalisés grâce à des expériences *in vitro* qui mettent en présence des substrats de réparation présentant ou non des microhomologies avec des extraits mitochondriaux. Ils observent que contrairement aux extraits totaux qui montrent une activité de réparation par NHEJ, les extraits mitochondriaux présentent une activité de réparation utilisant des microhomologies de 5 nucléotides minimum, ce qui suggèrent une réparation par alt-NHEJ. Par déplétion et par l'utilisation d'inhibiteur, les auteurs identifient les protéines FEN1, CtIP, Mre11, PARP1 et la DNA LIGASE 3 comme étant impliquées dans la voie du alt-NHEJ au sein de la mitochondrie.

Ses résultats sont à considérer avec prudence car il s'agit d'expériences *in vitro* faites à partir de préparation d'extraits mitochondriaux. A ce jour, la présence des protéines CtIP, Mre11 et PARP1 à la mitochondrie n'est toujours pas encore bien établie. De surcroit, l'étude bioinformatique réalisée par les auteurs n'a pas permis de détecter un signal d'adressage à la mitochondrie au sein des séquences de ces protéines.

e) La recombinaison homologue (HR) dans la mitochondrie

Le noyau utilise la recombinaison homologue (HR) pour réparer ses cassures double-brin ou lors de fourches de réPLICATION bloquée. Elle fonctionne majoritairement en phase S/G2 du cycle cellulaire. Le principe de cette voie repose sur l'utilisation par la cellule d'une matrice de réparation homologue. Cette matrice de réparation peut être la chromatide sœur ou rarement le chromosome homologue.

Après cassure double-brin, la première étape est de réséquer les brins cassés afin d'obtenir des brins avec des extrémités 3' sortantes. Cette étape est réalisée par les mêmes partenaires que dans la voie du alt-NHEJ : le complexe MRN et CtIP. La recombinaison homologue nécessite une résection importante. Pour cela, la cellule fait appel à Exo1/BLM (Truong et al. 2013). La protéine RPA (Protein Replication A) se lie à l'ADN simple-brin et empêche la formation de structures secondaires. L'étape suivante est l'invasion du brin cassé au sein du chromosome homologue ou la chromatide sœur qui va lui servir de matrice de réparation. Pour cette étape, RPA est chassée puis remplacée par Rad51. La protéine Brca2 interagit avec Rad51, ce qui la stimule pour chasser RPA (Moynahan et al. 2001). L'ADN simple-brin protégé par Rad51 est appelé nucléofilament. Il y a ensuite formation d'une D-loop lors l'invasion du nucléofilament au sein de l'ADN homologue. Après formation de la D-loop, la restauration de l'ADN peut être effectuée via le « Synthesis Dependent Strand Annealing » (SDSA) ou par le « Double Strand Break Repair » (DSBR) via la résolution de la jonction de Holiday. Cette jonction de Holiday formée pourra être résolue par l'hélicase Bloom aidée par la topoisomérase 3 α (Figure 16).

Etant données les copies multiples d'ADN présentes dans une mitochondrie, il a été proposé que la recombinaison homologue pourrait être un bon candidat de voie de réparation. En effet, lors de cassures double-brin sur une copie d'ADN mitochondrial, la mitochondrie pourrait utiliser une autre copie comme matrice de réparation. Une étude réalisée dans la levure suggère l'utilisation par la mitochondrie d'une voie proche de la recombinaison homologue (Stein et al. 2015). Chez le rat, une étude *in vitro* réalisée à partir d'extraits mitochondriaux conclut à la possibilité de la présence d'une activité de réparation par recombinaison homologue au sein des mitochondries (Thyagarajan et al. 1996).

Il n'y a pour autant aucune démonstration pertinente de l'utilisation de la recombinaison homologue dans les mitochondries de cellules humaines, la présence des acteurs de la recombinaison homologue étant notamment loin d'être statuée.

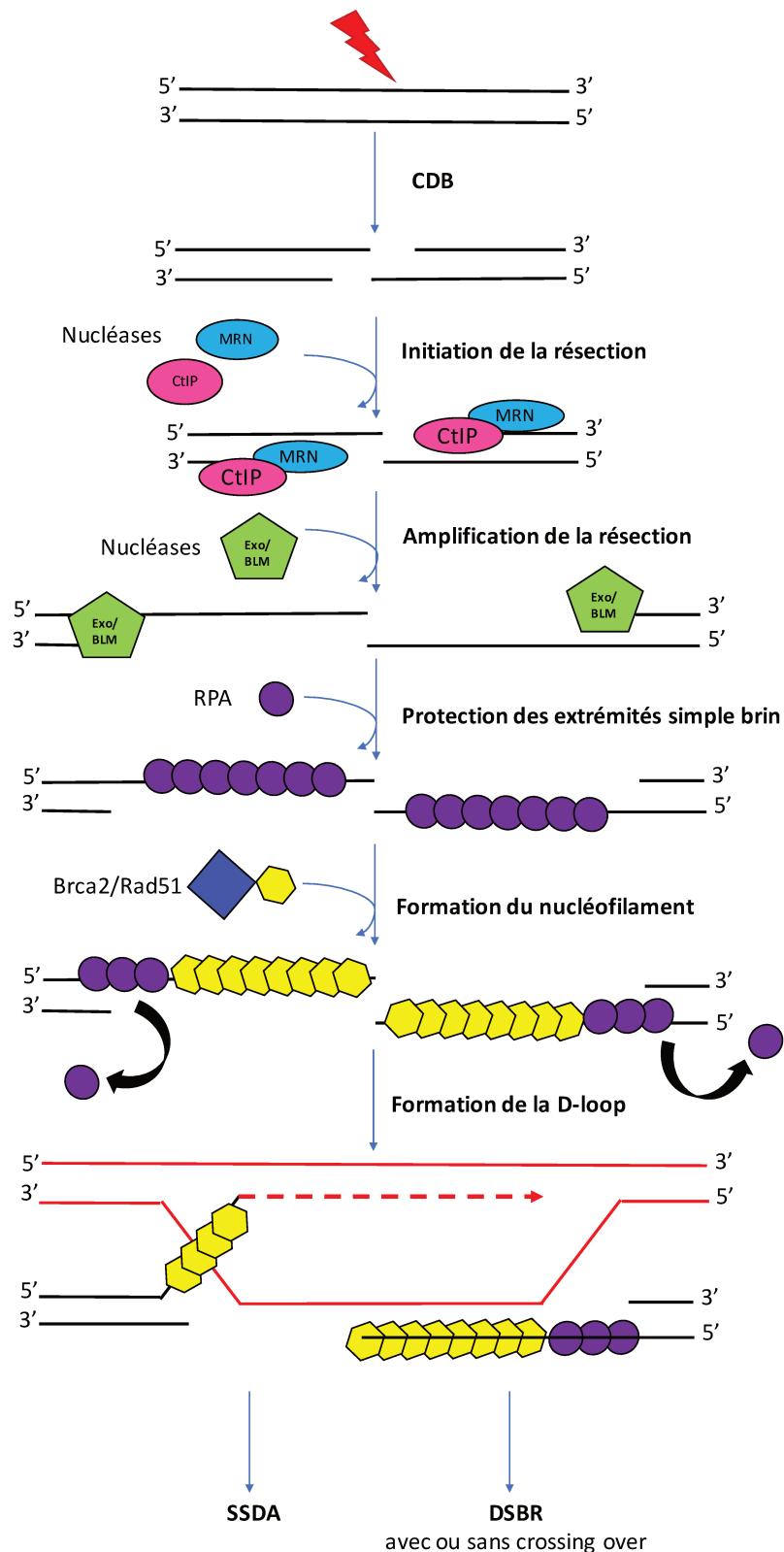


Figure 16 : Schéma présentant les principales étapes de la recombinaison homologue
 Après cassure double-brin, l'ADN nécessite une résection importante par le complexe MRN et CtIP puis par Exo1/BLM. RPA protège l'ADN simple-brin puis est remplacée par Rad51 stimulée par Brca2. Le nucléofilament formé envahit l'ADN homologue qui va servir de matrice de réparation. L'intégrité de l'ADN cassé est restaurée soit par le SSDA soit par le DSBR avec ou sans crossing over.

3) Maintien de l'intégrité de l'ADN mitochondrial par dégradation

Il semblerait que l'ADN mitochondrial possède une façon originale de ne conserver que les molécules d'ADN mitochondrial intactes contrairement à l'ADN nucléaire. Cette méthode consiste à dégrader son ADN au lieu de le réparer. Cette propriété a même été utilisée pour créer une lignée sans ADN mitochondrial. L'équipe de Seibel (Kukat et al. 2008) utilise une enzyme de restriction EcoRI modifiée pour aller à la mitochondrie. Les cassures double-brin provoquées par cette enzyme provoquent la dégradation de l'ADN mitochondrial. En utilisant cette propriété particulière de non réparation des cassures double-brin, ils ont ainsi obtenu une lignée cellulaire sans ADN mitochondrial. D'autres travaux menés par l'équipe d'Alexeyev (Shokolenko et al. 2009) ont permis de mettre en évidence que la dégradation n'avait pas lieu uniquement dans le cas des cassures double-brin. Après exposition à des ROS, l'équipe d'Alexeyev n'observe pas d'accumulation significative de mutations au sein de l'ADN mitochondrial, contrairement à ce qui était attendu, mais une augmentation de sites abasiques et de coupures intra-brin. Cette équipe se rend alors compte par Southern Blot que l'ADN mitochondrial a été dégradé rapidement (une heure après exposition à un stress oxydatif) en passant par une étape de linéarisation de l'ADN mitochondrial (50% d'ADN linéaire supplémentaires est détecté comparé au control). Ces différents travaux illustrent que la mitochondrie semble, du fait de ses multiples copies d'ADN au sein de chaque organite, ne pas privilégier la réparation de ses copies endommagées mais plutôt leur dégradation.

Pour résumé, les mécanismes de réparation de l'ADN mitochondrial sont encore mal compris. Il manque notamment d'études réalisées en cellules pour étoffer les études *in vitro*. De surcroît, les études actuelles tendent peut-être à trop rechercher l'équivalent des voies de réparation du noyau sur la mitochondrie. Du fait de ses nombreuses copies d'ADN, la mitochondrie pourrait simplement dégrader ses molécules trop endommagées et stimuler ensuite la réPLICATION. Par ailleurs, le protéosome réduit de la mitochondrie est différent de celui du noyau. L'absence de certaines protéines n'exclut peut-être pas l'absence de certaines voies de réparation qui pourrait être prises en charge par la polyvalence de certaines protéines.

Les pathologies de la mitochondrie

De nombreuses mutations et délétions sont retrouvées au sein de l'ADN mitochondrial. Ces modifications du génome mitochondrial ne sont pas nécessairement associées à des pathologies. En effet, selon le nombre de copies modifiées présentes dans la cellule ou dans un tissu, les conséquences sont très différentes (non pathologiques à pathologiques, plus ou moins sévères). Le rapport du nombre de copies mitochondrielles mutées/délétées et sauvages est appelé degré d'hétéroplasmie. Il est considéré que les pathologies neurologiques, cardiaques et musculaires surviennent lorsque le degré d'hétéroplasmie atteint 50 à 90% de copies modifiées (Wallace 2016). Les tissus qui sont le plus sensibles à l'apparition de copies modifiées de leur ADN mitochondrial sont en général les tissus qui nécessitent le plus d'énergie tels que les muscles et le cerveau.

L'apparition de délétions/mutations de l'ADN mitochondrial peut survenir soit (i) à la suite de dommages, par exemple lors de l'exposition aux ROS soit (ii) à cause de mutations dans des protéines essentielles pour la réPLICATION, comme par exemple PolG et Twinkle. Les protéines mutées possèdent alors des fonctions altérées, engendrant une modification du nombre de copies de l'ADN mitochondrial car elles sont moins efficaces pour la réPLICATION, et/ou l'apparition de délétions et/ou de mutations par un manque de fidélité. Ces deux circonstances auront pour conséquences une perturbation des besoins énergétiques de la cellule.

I. Les délétions dans l'ADN mitochondrial

Les délétions au sein de l'ADN mitochondrial sont associées à de nombreuses pathologies : vieillissement, cancers, maladies mitochondrielles comme le syndrome de Kearns-Sayre, des myopathies, l'ophtalmoplégie externe progressive, surdité, etc... (Schapira 2006) (Chen, et al. 2011). A ce jour, il est considéré que l'apparition des délétions a pour origine soit une perturbation de la réPLICATION soit un mécanisme de réparation utilisant des séquences répétées. La délétion la plus souvent retrouvée est une délétion de 4977 bp appelée « Common Deletion » (Figure 17).

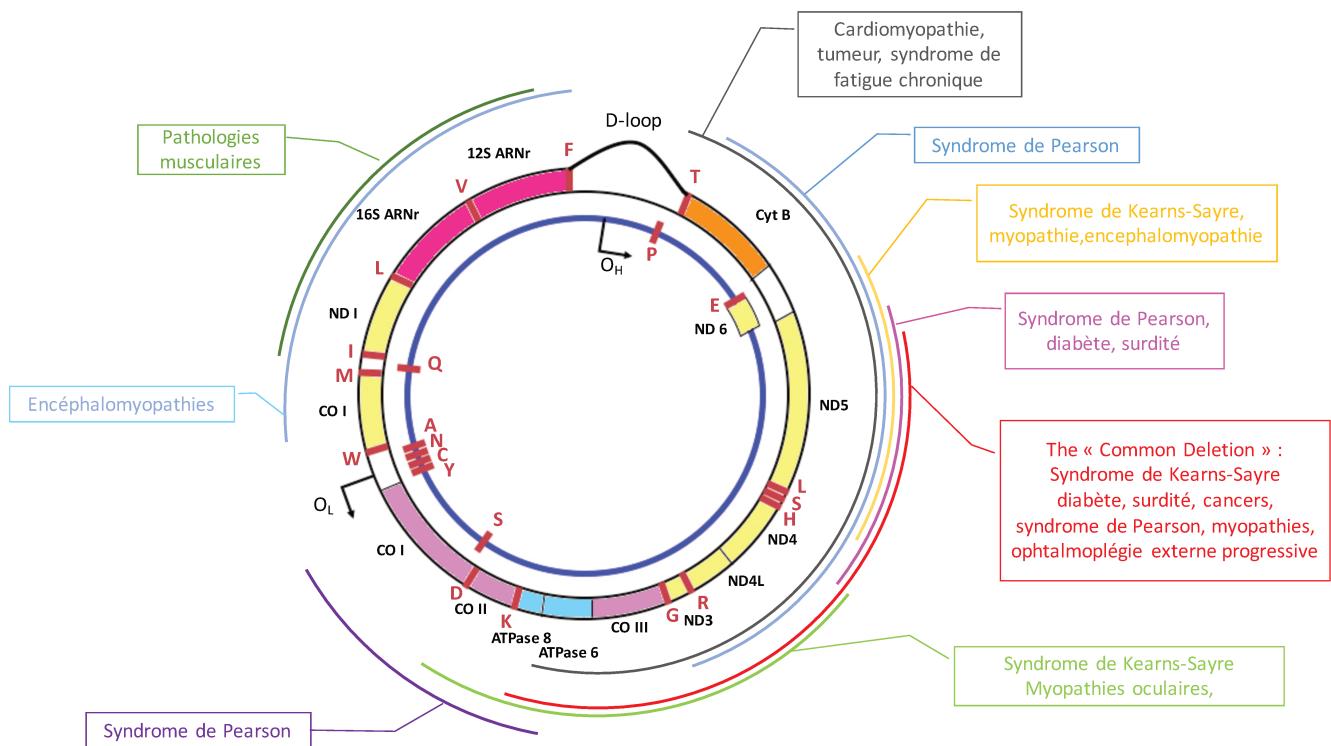


Figure 17 : Les délétions les plus souvent retrouvées et leurs pathologies associées (schéma inspiré de (Chen, He, Huang, et al. 2011))

Les délétions de l'ADN mitochondrial sont retrouvées tout au long de son génome mais majoritairement dans l'arc majeur. L'arc majeur correspond au 2/3 de l'ADN entre l' O_H et l' O_L . Les délétions sont impliquées dans différentes pathologies : syndrome de Kearns-Sayre, des myopathies, syndrome de Pearson, l'ophtalmoplégie externe progressive, des diabètes et dans des tumeurs. La « Common Deletion » (en rouge) est la délétion la plus fréquemment retrouvée.

1) La « Common Deletion »

Cette délétion est appelée « Common Deletion » car c'est la délétion la plus souvent retrouvée dans l'ADN mitochondrial. Elle est à l'origine de nombreuses pathologies. En particulier elle a été retrouvée dans les cellules musculaires chez des patients atteints du syndrome de Kearns-Sayre et d'ophtalmoplégie externe progressive (Shoffner et al. 1989), des myopathies (Holt et al. 1989) et de nombreux cancers, carcinome des cellules basales et carcinome des cellules épithéliales (Kamenisch et al. 2007), cancer du sein (Dimberg et al. 2011), cancer colorectal (Chen et al. 2011). Il faut néanmoins signaler que la « Common Deletion » est aussi présente de façon asymptomatique chez presque la moitié des sujets testés (27/60) dans l'étude de Min (Wang et al. 2013)) mais à des taux très faible de l'ordre du 10^{-5} selon le ratio Common Deletion/ mtADN total.

Cette délétion représente une perte de 4977 bp. Ce fragment contient 2 sous-unités du complexe V, une sous-unité du complexe IV, quatre sous-unités du complexe I et 5 ARNr

(Figure 17). Elle est entourée de part et d'autre par la même séquence, une séquence de 13 bp (ACCTCCCTCACCA), l'une se trouve dans le gène ND5 (position 8470-8482) et l'autre dans le gène de l'ATPase 8 (position 13447-13459). Lorsque cette délétion se forme, une seule séquence de 13 bp est conservée (Figure 18).

Des études sur les caractéristiques cellulaires montrent que l'accumulation de « Common Deletion » n'est pas anodine. En effet, plus le pourcentage de « Common Deletion » est important dans les cellules, plus la quantité de ROS augmente. Il existe aussi des modifications du potentiel de membrane et de la production en ATP. En condition de stress oxydatif, les perturbations morphologiques et l'apoptose sont plus importantes dès que les cellules possèdent environ 50% de « Common Deletion » (Peng et al. 2006), (Porteous et al. 1998)

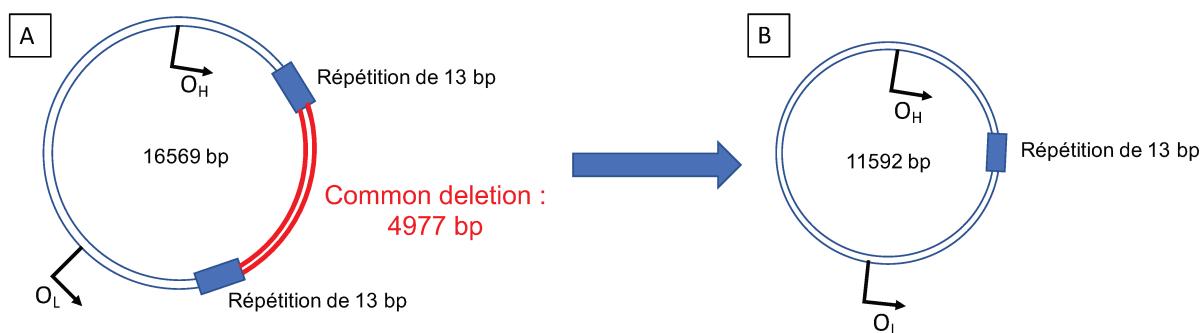


Figure 18 : Représentation de la « Common Deletion » au sein de l'ADN mitochondrial
A) La common deletion fait 4977 bp et est entourée par une séquence répétée de 13 bp (ACCTCCCTCACCA). B) lors de la formation de la « Common Deletion », une seule répétition de 13 bp est conservée

Puisque la « Common Deletion » est retrouvée dans de très nombreux cancers, des équipes ont analysé la proportion de « Common Deletion » dans la tumeur et ses tissus périphériques. L'équipe de Bai (Nie et al. 2013) observe notamment la présence de la délétion dans de très nombreux tissus cancéreux et constate que le tissu périphérique à la tumeur à une quantité plus importante de « Common Deletion » que la tumeur. La même observation a été faite dans le cas d'une étude sur des patients atteints de cancer colorectal (Chen et al. 2011). Ces résultats montrent la complexité de l'apparition et de la régulation de la délétion, évoquant une sélection négative par les cellules cancéreuses dans un environnement cellulaire enrichi en délétion. Cela suggère des mécanismes de régulation encore non décrits à l'heure actuelle.

Il est communément répandu que les mutations et/ou délétions vont s'accumuler au fur et à mesure du temps. Certains résultats paraissent conforter cette affirmation, l'accumulation étant, en effet, liée à l'âge mais avec un taux de délétions très variable selon les tissus (Meissner

et al. 2008). Pourtant, dans d'autres études, les patients de plus de 65 ans ne présentent pas de quantité plus importante en « Common Deletion » au sein de leur tumeur (Wang et al. 2013) (Chen et al. 2011). Les résultats contradictoires de plusieurs études sont souvent dus à la taille des échantillons. C'est pourquoi, l'équipe de Nebel (von Wurmb-Schwarz et al. 2010) quantifie la proportion de « Common Deletion » dans le sang de plus de 400 individus âgés (61–75 ans) ou très âgés (95–109 ans) et constate que le groupe « âgé » est porteur d'une quantité plus importante de « Common Deletion » que le groupe « très âgé ».

Contrairement à l'idée reçue, l'accumulation des délétions ne semble donc pas s'accumuler au cours du temps sans régulation mais pourrait être contrôlée par différents mécanismes.

II. Les pathologies liées à des mutations des protéines du réplisome

Les maladies liées à une perturbation de la mitochondrie peuvent aussi être la conséquence de mutations dans les protéines essentielles à la réPLICATION, telles que les protéines POLG et Twinkle. De la même façon, ces pathologies touchent des organes jouant un rôle dans le métabolisme et des tissus ayant des besoins énergétiques importants.

1) Les maladies dues à des mutations de POLG

Actuellement, 147 mutations dans la polymérase POLG ont été recensées, ces mutations pouvant être dominantes ou récessives. De nombreuses pathologies sont associées à ces mutations, notamment le syndrome d'Alpers, le syndrome MCHS (myocerebrohépatopathie pédiatrique, ataxie, syndrome neuropathique), le syndrome MEMSA (épilepsie myoclonique, myopathie et ataxie sensorielle), et l'ophtalmoplégie externe progressive. Les mutations sont retrouvées tout au long de la protéine, et touchent aussi bien le domaine exonucléase, le linker, que le domaine polymérase. Les conséquences de ces mutations dans la protéine sont soit des délétions, soit une diminution du nombre de copies dans l'ADN mitochondrial, notamment dans le syndrome d'Alpers (Stumpf & Copeland 2011).

Etant donné le nombre de pathologies et leur gravité, beaucoup d'études se sont intéressées à mieux comprendre les effets des mutations de POLG. Un nombre important d'études ont été publiées comprenant des expériences *in vitro*. Par exemple, la mutation A467T située dans la partie linker de la protéine entre le domaine exonucléase et le domaine polymérase, est la mutation la plus retrouvée dans le syndrome d'Alpers. Cette mutation provoque une activité polymérase inférieure à 4% et une activité exonucléase diminuée de deux

fois comparée à la protéine WT. De plus, il y a été observé une perte d'interaction avec la sous-unité POLG β , (Chan et al. 2005). Deux autres mutations G848S and R852C présentes cette fois dans le domaine polymérase de POLG entraînent une activité polymérase de seulement 1% comparée à la protéine WT, ces mutations semblant perturber la liaison à l'ADN de la protéine (Kasiviswanathan et al. 2009). De façon intéressante, un mutant de POLG D198A dans le domaine exonucléase révèle une augmentation du pouvoir mutagène (à savoir une mutation toutes les 1700 bp soit environ 10 mutations par molécule d'ADN au lieu d'un taux entre 0 et 1 mutation par génome mitochondrial pour la protéine WT) (Spelbrink et al. 2000).

Des modèles de souris transgéniques ont aussi été mis au point afin d'étudier les effets des mutations de POLG. Un modèle souris exprimant la POLG mutée D181A dans ses cardiomycocytes montre qu'à un mois de vie, 1,4 mutations par 10 000 bp sont présentes comparées à 0,06 par 10 000 bp dans les autres tissus n'exprimant pas la protéine mutée. Ces souris ne présentent pas de modification du nombre de copies, ni de perturbation de la chaîne respiratoire pourtant elles développent une cardiomyopathie dès 4 à 5 semaines (Zhang et al. 2000). Une souris transgénique possédant la mutation Y955C connaît une mortalité prématuée associée à une cardiomyopathie. L'analyse de son ADN mitochondrial montre une diminution d'un facteur 2 dans le nombre de copies, une augmentation d'un facteur 3 en 8 oxoguanine illustrant l'augmentation du stress oxydatif (Lewis et al. 2007). Une autre étude compare des souris WT, POLG^{mut/WT} et POLG^{mut/mut} (Vermulst et al. 2008). Les souris transgéniques POLG^{mut/mut} font état d'un vieillissement prématucléaire associé à la présence de délétions augmentée d'un facteur de 7 à 11 fois. Les souris POLG^{mut/WT} n'affichent pas d'augmentation en délétions mais une augmentation en mutations d'un facteur 100 comparée aux souris WT.

2) Les maladies dues à des mutations de Twinkle

La principale pathologie liée à des mutations de la protéine Twinkle est l'ophtalmoplégie externe progressive (OPE) dans des formes autosomales dominantes ou récessives. L'équipe de Larsson analyse onze mutations retrouvées dans la protéine Twinkle chez des patients atteints de OPE (Spelbrink et al. 2001). Ces mutations n'empêchent pas la localisation mitochondriale ou la multimérisation de la protéine *in vitro*, bien que ces mutations soient présentes dans le domaine de la protéine permettant la multimérisation (partie linker de la protéine). Ces patients présentent un nombre plus important de molécules d'ADN mitochondrial délétées associées à leurs différentes mutations dans la protéine Twinkle.

L'équipe de Spelbrink testent sept mutations présentes dans des formes d'ophtalmoplégie externe progressive (Goffart et al. 2009). Ces sept mutations touchent la partie N terminale, le linker ou le domaine hélicase de la protéine. L'expression en cellules (HEK293) des différents mutants montre une diminution d'ADN mitochondrial dès trois jours d'expression. Par exemple, la surexpression du mutant Twinkle R374Q (mutation présente dans le linker de la protéine) provoque une diminution de 25% d'ADN mitochondrial à trois jours. Ce mutant possède aussi une activité hélicase réduite *in vitro* bien que la mutation ne soit pas dans le domaine hélicase. L'étude de la réPLICATION par électrophorèse 2D montre l'accumulation d'intermédiaires de réPLICATION. Ce mutant présente les mêmes caractéristiques que les six autres mutants étudiés. Les auteurs concluent donc que les mutations dans la protéine Twinkle entraînent des pauses dans la réPLICATION ce qui provoque des perturbations dans l'ADN mitochondrial (délétions et baisse du nombre de copies). Comme pour POLG, des modèles souris ont été établis. Des souris transgéniques avec une duplication d'un acide aminé en position 352 et 364 dans Twinkle ont été créées (Tyynismaa et al. 2005). Cette mutation est retrouvée dans l'ophtalmoplégie externe progressive et ce modèle de souris recrée les mêmes symptômes que ceux observés chez les patients. En effet, ces souris accusent une restriction de 50% d'ADN mitochondrial dans le cerveau, et montrent diverses délétions (environ 5%) dans le muscle et le cerveau ainsi que des modifications morphologiques de leurs fibres musculaires.

L'apparition des pathologies étant rattachée à des effets seuil selon l'hétéroplasmie, maîtriser le degré d'hétéroplasmie constitue un objectif majeur en tant qu'outil thérapeutique face aux pathologies liées à l'ADN mitochondrial. Il reste donc encore beaucoup à comprendre sur les changements d'hétéroplasmie. Il est par exemple surprenant que certains patients développent des maladies avec des degrés d'hétéroplasmie très différents (Rossignol et al. 2003). Il a été suggéré que les molécules délétées auraient un avantage réPLICatif (Fukui & Moraes 2009) (Diaz et al. 2002). Pourtant il semblerait que ces hypothèses soient remises en cause. En effet, les souris transgéniques ont par exemple permis d'observer une repopulation en molécules d'ADN WT lors de la pratique d'exercices physiques (Safdar et al. 2011). Un autre argument réside dans le fait que l'étude concernant des personnes très âgées démontre l'absence d'augmentation de molécules délétées liée au vieillissement (von Wurmb-Schwark et al. 2010). Les mécanismes de régulation de l'hétéroplasmie semblent donc mal compris. Une autre raison importante de pouvoir agir sur l'hétéroplasmie est que la transmission des mitochondries est uniquement maternelle (Gyllensten et al. 1985). Les copies mutées ou

délétées seront transmis à la descendance. Des outils sont donc nécessaires pour la modification du génome mitochondrial.

Comment modifier le génome mitochondrial ?

Le nombre de pathologies touchant l'ADN mitochondrial est très important. Les outils thérapeutiques étant à ce jour très insuffisants, il est donc nécessaire de mettre au point de nouvelles stratégies. L'une d'entre elles consiste à éliminer et/ou corriger les copies d'ADN mitochondrial endommagées. Afin d'y parvenir, les nucléases artificielles semblent être une piste prometteuse. Cette technologie est une révolution en thérapie génique qui permet de modifier « à façon » l'ADN nucléaire. De ce fait, depuis quelques années, les chercheurs tentent d'adapter cette technologie à la mitochondrie.

I. La modification du génome mitochondrial dans les cellules cybrides

Les cybrides sont la fusion de l'ADN mitochondrial de cellules donneuses enucléées avec des cellules acceptrices sans ADN mitochondrial. Leur ADN mitochondrial a été éliminé à l'aide d'un traitement au bromure d'éthydium (BET). Cette méthode est notamment utilisée pour créer des lignées cellulaires modèles de pathologies. Selon l'hétéoplasmie des mitochondries donneuses, des clones avec des hétéoplasmies variables seront obtenus. Ce seront ensuite de bons outils pour comprendre les effets de seuils dans l'apparition des pathologies. Des modèles cellulaires ont ainsi pu être obtenus. Par exemple, un modèle a été conçu à partir de cellules Hela avec l'intégration d'ADN mitochondrial portant une délétion de 5196 bp entre les nucléotides 8563 et 13758 impliquée dans une forme d'ophtalmoplégie progressive externe (Hayashi et al. 1991). Des lignées cellulaires présentant des hétéoplasmies variables pour la « Common Deletion » ont aussi été établies en utilisant des fibroblastes provenant de patients atteints d'ophtalmoplégie progressive externe fusionnés avec des cellules 143B sans ADN mitochondrial (Peng et al. 2006).

II. La modification du génome mitochondrial par l'utilisation des nucléases

La stratégie générale afin de modifier le génome mitochondrial a été d'adapter la technologie des nucléases artificielles utilisée dans la modification du génome nucléaire. Pour appliquer cette technologie à la mitochondrie, il convient d'ajouter aux nucléases artificielles des signaux d'import afin qu'elles empruntent la voie d'import des protéines mitochondrielles pour pénétrer dans la matrice mitochondriale (cf partie voie d'import).

1) Les enzymes de restriction

L'utilisation d'enzymes de restriction importées à la mitochondrie a été la première stratégie pour éliminer des copies de l'ADN mutées (Srivastava & Moraes 2001). Plusieurs études montrent que cette méthode réussit à modifier le degré d'hétéoplasmie (Bacman et al. 2007) (Fukui & Moraes 2009). L'équipe de Seibel (Kukat et al. 2008), quant à elle, a utilisé la transfection de l'enzyme EcoRI pour obtenir des cellules sans ADN mitochondrial. Dans certains cas, l'apparition d'une mutation crée un site de reconnaissance pour l'enzyme de restriction. La transfection avec l'enzyme de restriction pourra donc permettre l'élimination des copies mutées. Le problème de cette méthode reste la limitation du nombre de sites de restriction existants ainsi que le fait que la majorité des mutations ne créent pas un site de restriction. Cette technologie a même été mise au point pour faire des souris transgéniques (Srivastava & Moraes 2005) (Fukui & Moraes 2009) (Reddy et al. 2015). Par exemple, la transfection d'un ARNm codant pour l'enzyme de restriction ApaLI dans des oocytes de souris possédant 2 formes asymptomatiques d'ADN mitochondrial, BALB ou NZB (seule BALB possède un site ApaLI) conduit à une diminution drastique de la forme BALB (de 80% à 15%) (Reddy et al. 2015). Ce succès de modification de l'ADN mitochondrial d'oocytes de souris est très prometteur. En effet, l'ADN mitochondrial étant hérité par la mère, cette stratégie peut être un vrai outil thérapeutique pour limiter la transmission de maladies génétiques liées à l'ADN mitochondrial à sa descendance.

L'utilisation des enzymes de restriction est donc un outil très intéressant pour l'élimination de copies mutées mais reste limitée par la nécessité que les mutations forment un site de restriction.

2) Les nucléases à doigt de zinc

a) Description

Les nucléases à doigts de zinc sont composées d'un domaine de liaison à l'ADN spécifique d'une séquence choisie séparé par un linker du domaine endonucléase non spécifique provenant de la protéine FokI (Kim et al. 1996) (Figure 19). Ce domaine de liaison à l'ADN est un enchainement de doigts de zinc qui permet une reconnaissance spécifique d'une séquence d'ADN. Le domaine de liaison à l'ADN peut enchaîner 3 à 6 doigts de zinc. Le motif doigt de zinc est appelé de forme Cys₂-His₂ car il est composé de la séquence suivante : X₃-Cys-X₂₋₄-Cys-X₁₂-His-X₃₋₄-His-X₄ (X correspond à n'importe quel acide aminé). Un doigt de zinc forme

une hélice alpha, est stabilisé par un ion Zn^{2+} et lie le grand sillon de l'ADN. Il reconnaît un triplet de base de l'ADN (Pavletich & Pabo 1991). Selon la séquence du doigt de zinc, le triplet de l'ADN reconnu varie. Cependant, il existe des exceptions dans la reconnaissance du triplet de bases par le doigt de zinc. En effet, il a été observé qu'un doigt de zinc ne reconnaît pas forcément qu'un seul triplet de bases d'ADN selon les doigts de zinc qui l'entourent (Fairall et al. 1993).

Le domaine endonucléase couplé au domaine doigt de zinc, qui va induire une cassure double-brin, provient de la partie C terminale de l'endonucléase FokI. Il ne possède pas de spécificité de séquence. Pour couper l'ADN, le domaine FokI doit se dimériser (Bitinaite et al. 1998). De ce fait, pour une coupure double-brin, les nucléases à doigt de zinc doivent être utilisées en dimère (Figure 19). Une modification par mutation du domaine FokI a été effectuée pour forcer l'hétérodimérisation entre les monomères. Il en résulte une diminution de plus de 40% des coupures par homodimérisation (Doyon et al. 2011). Des nucléases à doigt de zinc en version nickase (dont un des domaines FokI est inactivé) ont aussi été mises au point (Ramirez et al. 2012).

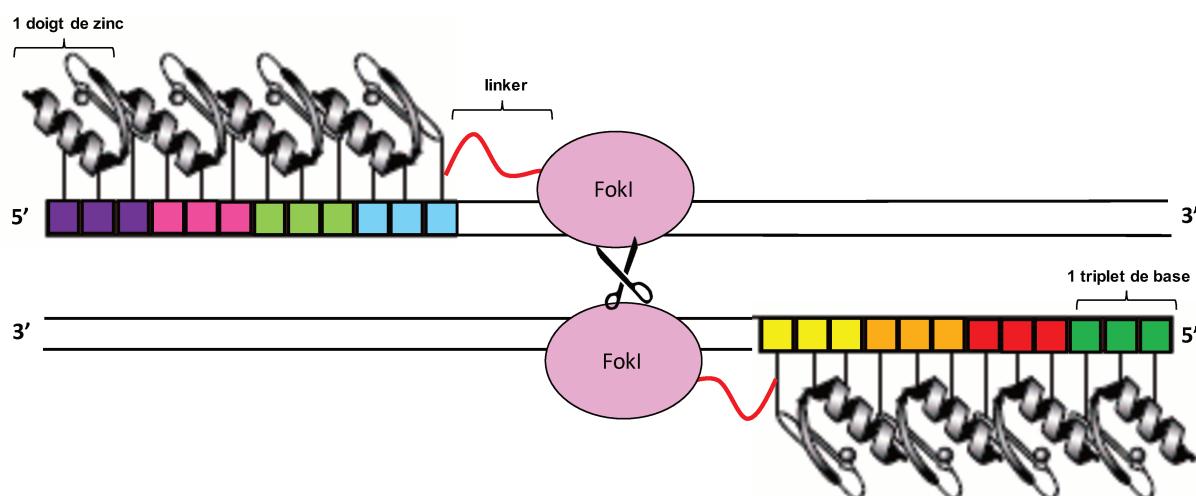


Figure 19 : Schéma d'une nucléase à doigt de zinc

La nucléase à doigt de zinc est composée d'un enchainement de doigt de zinc associé grâce à un linker (en rouge) au domaine endonucléase FokI (en rose). Un doigt de zinc reconnaît 3 bases de l'ADN. Deux hétérodimères doivent se faire face afin que les deux domaines FokI puissent se dimériser pour permettre la coupure double brin.

b) Les nucléases à doigts de zinc dans la mitochondrie

L'équipe de Klug (Minczuk et al. 2006) utilise un domaine de liaison à l'ADN, grâce à des motifs doigts de zinc, couplé au domaine méthylase de la DNMT3- α humaine pour cibler

la mutation T8993G impliquée dans plusieurs pathologies (le syndrome NARP (neuropathie, ataxie, rétinite pigmentaire) et syndrome de Leigh). Pour permettre l'import à la mitochondrie de l'enzyme, les auteurs testent différents signaux d'import à la mitochondrie. Ils observent que selon les signaux d'import, l'adressage dans la mitochondrie est plus au moins efficace et qu'une partie des protéines à doigt de zinc se retrouvent exprimées dans le noyau. L'ajout d'un signal d'exclusion nucléaire en C terminal résout ce problème. L'ajout de la séquence 3' UTR de l'ARNm de l'ATP synthase participe aussi à un meilleur import des protéines à doigt de zinc (cf partie import). Après transfection, leurs résultats montrent une augmentation de la méthylation sur les cytosines adjacentes à la base G mutée d'un facteur 6 au niveau de la séquence mutée comparée à la séquence WT (le domaine méthylase méthyle les CpG donc seule la séquence mutée sera méthylée).

Deux ans plus tard, la même équipe (Minczuk et al. 2008) utilise à nouveau la stratégie des nucléases à doigt de zinc pour cette fois éliminer des copies mutées. Ils s'intéressent à la même mutation. Ils utilisent les mêmes signaux d'adressage pour importer une nucléase à doigt avec quatre domaines doigt de zinc et deux domaines FokI. Cette stratégie permet la transfection d'un seul monomère de nucléase à doigt de zinc. Les deux domaines FokI de cette nucléase sont séparés par un linker de 40 acides aminés. Cette structure empêche la coupure des molécules non mutées par des homodimères et augmente la spécificité de coupure pour les molécules mutées (Figure 20). Les résultats montrent que 2 et 30 jours après transfection, la lignée cellulaire qui possédait 85% de molécules mutées et 15% de molécules WT récupère environ 40% de molécules WT, ce qui a permis une diminution de 25% de forme mutée. Concernant les clones exprimant stablement la nucléase, les auteurs constatent que l'expression de la nucléase est perdue au fur et à mesure du temps mais que les proportions d'ADN WT obtenues après transfection sont maintenues.

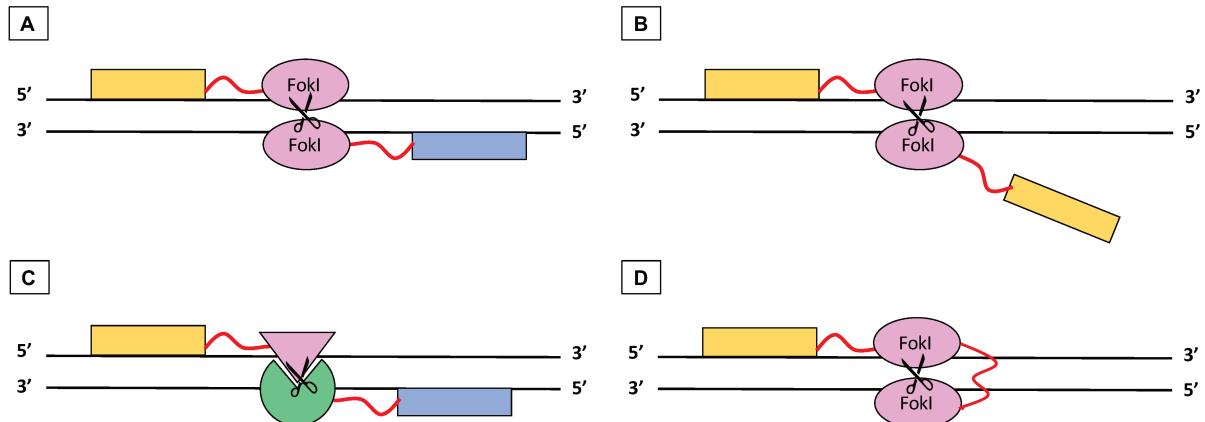


Figure 20 : Schéma représentant l'augmentation de la spécificité de coupure par amélioration du design des nucléases

A) Pour permettre une coupure spécifique au niveau de la séquence cible d'ADN, deux hétérodimères doivent s'associer. B) Dans le cas où la coupure doit être uniquement dans un ADN qui diffère par une mutation comparée à un autre ADN, une coupure non spécifique peut avoir lieu si deux homodimères s'apparentent par le domaine FokI. C) Pour augmenter la spécificité de coupure, des mutations dans le domaine FokI augmentent la spécificité en rendant l'hétéromimérisation obligatoire pour la coupure (Doyon et al. 2011). D) Pour augmenter la spécificité de coupure, une nucléase à doigt de zinc avec deux domaines FokI a été mise au point (Minczuk et al. 2008)

3) Les TALE nucléases (TALEN)

a) Description

Les TALENs (Transcription Activator-Like Effector Nuclease) sont des nucléases artificielles qui possèdent un domaine de liaison à l'ADN en N terminal et un domaine endonucléase, le domaine FokI. Ce dernier est le même domaine nucléase utilisé que pour les nucléases à doigt de zinc. Le domaine de liaison à l'ADN est appelé domaine TALE. Il a été identifié chez le phytopathogène de *Xanthomonas*. Il est composé de répétitions en tandem reconnaissant chacune une base de l'ADN, un domaine d'activation de la transcription et un signal d'adressage au noyau. Une répétition est composée de 34 acides aminés. Ce sont les acides aminés en position 12 et 13 appelés « Repeat-variable diresidue » (RVD) qui permettent la reconnaissance et la discrimination des quatre différentes bases de l'ADN. Le code de reconnaissance a été complètement élucidé, NG=T ; NI=A ; NS=A, C, G ou T ; NN=A ou G (Boch et al. 2009) (Deng et al. 2012). Les propriétés de ce domaine de liaison ont été utilisées pour cibler une séquence d'ADN choisie. En général, les nucléases artificielles sont composées de 18 répétitions (la première répétition pour le T +16 répétitions + une demi-répétition toujours présente naturellement). La seule limitation pour le choix de la séquence cible est d'avoir une base T en 5' de la séquence reconnue par le monomère (Mak et al. 2012) (Figure 21). Le linker doit être entre 12 à 21 bp entre le domaine TALE et le domaine FokI. Le domaine FokI étant le

même que pour les nucléases à doigt zinc, ce domaine peut être utilisé en version ELD/KKR pour forcer l'hétérodimérisation (Doyon et al. 2011) et en version nickase lorsque qu'un domaine FokI est inactivé (Wu et al. 2014).

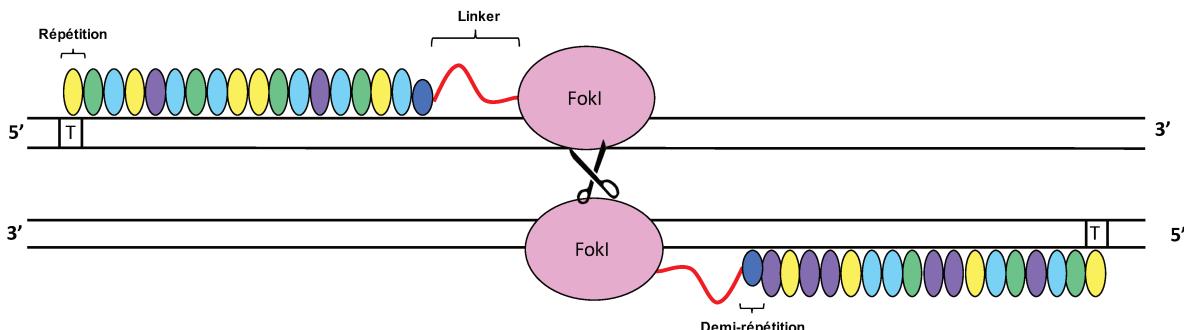


Figure 21 : Schéma d'une TALEN

La TALEN est composée d'un enchainement de répétitions associé grâce à un linker (en rouge) au domaine endonucléase FokI (en rose). Une répétition reconnaît une base de l'ADN. La première répétition doit reconnaître la base T de l'ADN en 5' de la séquence cible. La dernière répétition est une demi-répétition (en bleu foncé). Deux hétérodimères doivent se faire face afin que les deux domaines FokI puissent se dimériser et permettre la coupure double brin.

b) Les TALENs dans la mitochondrie

L'équipe de Moraes (Bacman et al. 2013) adapte la technologie TALEN en ajoutant un signal d'import à la mitochondrie en N terminal pour cibler l'ADN mitochondrial. Ces TALENs adaptées à la mitochondrie seront appelées mitoTALENs. Leur but est d'éliminer spécifiquement les copies d'ADN mitochondriales porteuses de la « Common Deletion » (cf partie 4). Le design de la TALEN permet la coupure spécifique des ADN portant la délétion. La TALEN a été transfectée dans des clones de cellules d'ostéosarcomes humains portant 90% ou 55 % de « Common Deletion ». Les résultats obtenus montrent que dans les cellules traitées un taux de 80% d'ADN sauvage se maintient au moins pendant quatorze jours après transfection. La récupération d'un pourcentage plus élevé de molécules non mutées est corrélée à une récupération totale de l'activité du complexe I de la chaîne respiratoire qui était affectée (30 à 40% de perte d'activité) dans le clone porteur de 90% de délétions.

Les mitoTALENs ont été testées également sur des oocytes de souris qui possèdent des copies mutées de leur ADN mitochondrial retrouvées dans 2 pathologies humaines, le syndrome NARP (neuropathie, ataxie, rétinite pigmentaire) (mutation 9176T>C) et la neuropathie optique de Leber (mutation 14459G>A). 72h après injection de l'ARNm codant pour les mitoTALENs, moins de formes mutées d'ADN ciblée est détectée. Cependant ces travaux n'ont pas permis l'obtention de souris modifiées viables (Reddy et al. 2015).

4) Le système CRISPR/Cas9

a) Description

Le système CRISPR/Cas9 (Clustered Regulazy Interspaced Short Palindromic Repeats, CRISPR associated protein 9) est issu du système immunitaire des bactéries. Ce système repose sur deux composants : l'ARN qui reconnaît l'ADN cible et une nucléase, la protéine Cas9 qui vont permettre la coupure double-brin de l'ADN lors de la reconnaissance du complexe ARN/ADN (Jinek et al. 2012).

Le système CRISPR a été identifié chez les bactéries. A ce jour, trois systèmes bactériens ont été décrits. Chez la bactérie, ce système est un mécanisme de défense adaptative. Lors de l'infection par un agent pathogène, l'ADN de ce dernier est coupé, des portions étant intégrées dans le génome bactérien entre des séquences répétées appelées locus CRISPR (Barrangou et al. 2007). Ces locus CRISPR sont transcrits et maturés afin d'obtenir des crARN (CRISPR ARN). Un crARN possède une séquence complémentaire au trans-activating crARN (tracrARN) qui permet le recrutement de la protéine Cas9. Lors d'une seconde infection, le complexe ARN (crARN et tracrARN) va reconnaître une séquence de 20 bp de l'ADN pathogène. Si cette hybridation a lieu en aval d'un protospacer adjacent motif (PAM), la protéine Cas9 reconnaîtra ce complexe ARN-ADN et coupera les deux brins de l'ADN pathogène (Mali et al. 2013) (Figure 22).

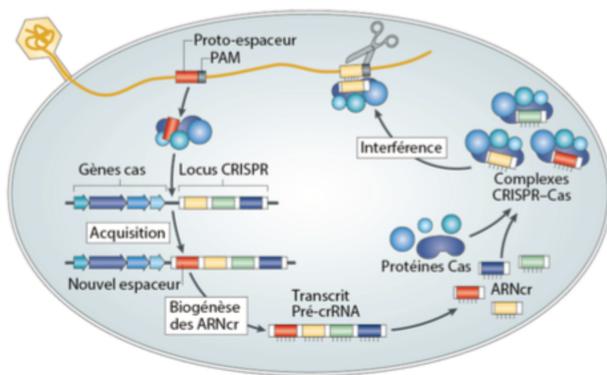


Figure 22 : Schéma représentant le système CRISPR/Cas9, mécanisme du système immunitaire des bactéries. Adapté d'après (Samson et al. 2013)

Lors de l'infection par un pathogène, son ADN est clivé et certaines courtes séquences d'environ 20 bp sont intégrées au niveau de locus CRISPR entre de courtes séquences répétées. Ces locus CRISPR sont transcrits en un long ARN, qui sera ensuite mûr pour donner des crARN. Lors d'une seconde infection, le complexe Cas9, crARN associé avec le tracrARN reconnaîtra l'ADN pathogène et s'il s'hybride en aval d'un protospacer associated motif (PAM) entraînera une coupure double-brin de l'ADN pathogène grâce à l'activité nucléase de la Cas9.

Le système CRISPR II de la bactérie a été adapté pour en faire un système de nucléase artificielle. Bien que le système bactérien utilise deux ARNs (le crARN et le tracrARN), un design simplifié a été mis au point permettant l'utilisation de la protéine Cas9 avec un seul ARN long pour permettre la coupure double-brin de l'ADN. Ce qui permet la reconnaissance est une séquence ARN de 20 nucléotides appelée guide ARN qui se trouve sur la séquence cible en aval d'un signal spécifique, le PAM. La protéine Cas9 vient alors reconnaître l'association de l'ARN à la séquence d'ADN cible et coupe les deux brins de l'ADN à 3 nucléotides en amont du PAM. Le plus souvent, la protéine Cas9 utilisée est la protéine Cas9 provenant de *Streptococcus pyogenes* et le PAM correspond à une séquence NGG, où N correspond à n'importe quel nucléotide (Jinek et al. 2012) (Figure 23).

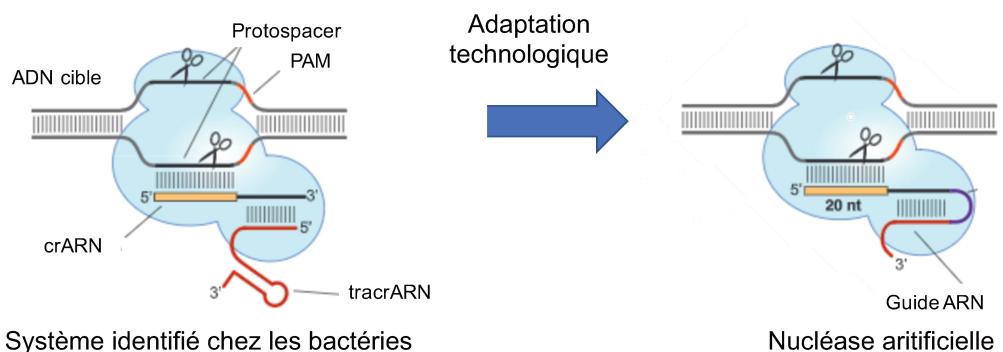


Figure 23 : Schéma représentant le système CRISPR/Cas9. Adapté d'après (Jinek et al. 2012)

Le système identifié chez la bactérie pour se protéger des infections par des phages ou des plasmides est un système à trois acteurs : la Cas9, le crRNA qui s'hybride à la séquence d'ADN cible, et le tracARN qui reconnaît le crARN (partie de gauche). Pour transformer ce système immunitaire bactérien en nucléase artificielle, le système a été simplifié et permet l'utilisation d'un système faisant appel à seulement 2 acteurs : la protéine Cas9 et un ARN guide qui s'autohybride (partie droite)

b) Le système CRIPR/Cas9 dans la mitochondrie

Actuellement, la nouvelle génération de nucléases artificielles CRISP/Cas9 n'a pas encore été adaptée à la mitochondrie. Ce système qui possède une grande flexibilité de séquence cible, car la présence du PAM dans la séquence est peu restrictive, est une alternative très encourageante pour augmenter les possibilités de modifications du génome mitochondrial. Contrairement aux nucléases précédentes qui étaient seulement constituées de protéines, les CRISPR/Cas9 possèdent une composante ARN. Or, faire rentrer l'ARN dans la mitochondrie semble être le challenge à relever pour faire marcher cette technologie dans cette organelle...

L'hétéroplasmie de l'ADN mitochondrial est un vrai problème dans l'émergence des maladies mitochondrielles. Les nucléases artificielles semblent donc apporter des solutions pour la correction et/ou l'élimination des copies mutées ou délétées. L'amélioration de ces technologies pourrait à terme être une solution thérapeutique pour les pathologies mitochondrielles.

Projet de thèse

Mon projet de thèse s'intéresse à une délétion de 4,9 kilobases située dans l'ADN mitochondrial appelée « Common Deletion ». Cette délétion est impliquée dans de nombreuses maladies notamment dans 2 syndromes : le syndrome de Pearson et le syndrome de Kearns Sayre et a été retrouvée dans différentes tumeurs. Elle est appelée « Common Deletion » car 30% des personnes présentant une délétion dans leur ADN mitochondrial possèdent cette délétion en particulier. D'un point de vue mécanistique, notre hypothèse de travail était que la « Common Deletion » apparaissait soit suite à une erreur de réPLICATION soit suite à la mauvaise réparation de l'ADN mitochondrial après une cassure double-brin.

Les enjeux de ce travail étaient donc de créer dans un premier temps un modèle d'étude de la « Common Deletion » en induisant cette délétion dans l'ADN mitochondrial puis en identifiant les protéines responsables de l'apparition de cette délétion afin de mieux comprendre les mécanismes de réparation et de réPLICATION dans la mitochondrie. Ce projet a été réalisé en étroite collaboration avec le laboratoire d'Agnel Sfeir à New York où ils ont particulièrement étudié le mode de réPLICATION actif au sein de la mitochondrie par peignage moléculaire. De notre côté, nous avons mis au point l'induction de la « Common Deletion » en induisant des cassures spécifiques sur l'ADN mitochondrial pour étudier les protéines impliquées dans sa formation.

Les objectifs de ma thèse se composaient donc en deux grandes parties :

(i) L'induction de la « Common Deletion »

Notre stratégie a été d'utiliser des nucléases artificielles adaptées pour être importées à la mitochondrie. Nous avons donc transfecté des TALENs (transcription activator-like effector nucleases) en leur rajoutant des signaux d'adressage à la mitochondrie (Bacman et al., Nat med, 2013) dans des cellules humaines. Nous avons ainsi obtenu des TALENs mitochondriaux appelés mitoTALENs. Les premières étapes étaient de vérifier l'adressage des mitoTALENs à la mitochondrie en fractionnant les cellules en un extrait cytoplasmique et un extrait mitochondrial puis de vérifier et de quantifier la formation de la « Common Deletion » par PCR et qPCR.

(ii) L'identification de protéines impliquées dans l'apparition de la « Common Deletion »

Notre objectif était donc d'induire la formation de cette délétion afin de comprendre le mécanisme qui conduisait à son apparition. Pour cela nous voulions identifier les protéines impliquées, ce qui nous a également permis de mieux caractériser les mécanismes de réPLICATION et de réPARATION dans la mitochondrie. Notre stratégie a été d'inhiber des protéines de réPLICATION et de réPARATION de l'ADN mitochondrial en utilisant des siARNs tout en induisant la « Common Deletion » avec les mitoTALENs dans des lignées cellulaires humaines. Ainsi, nous avons pu observer les conséquences de l'inhibition de la protéine ciblée sur la quantité de « Common Deletion ».

Résultats

I. Article : « Single-Molecule Analysis of mtDNA Replication uncovers the basis of the Common Deletion »

Ce travail a été réalisé et publié en collaboration avec le laboratoire d’Agnel Sfeir à New York. J’ai signé cette publication en tant que co-premier auteur.

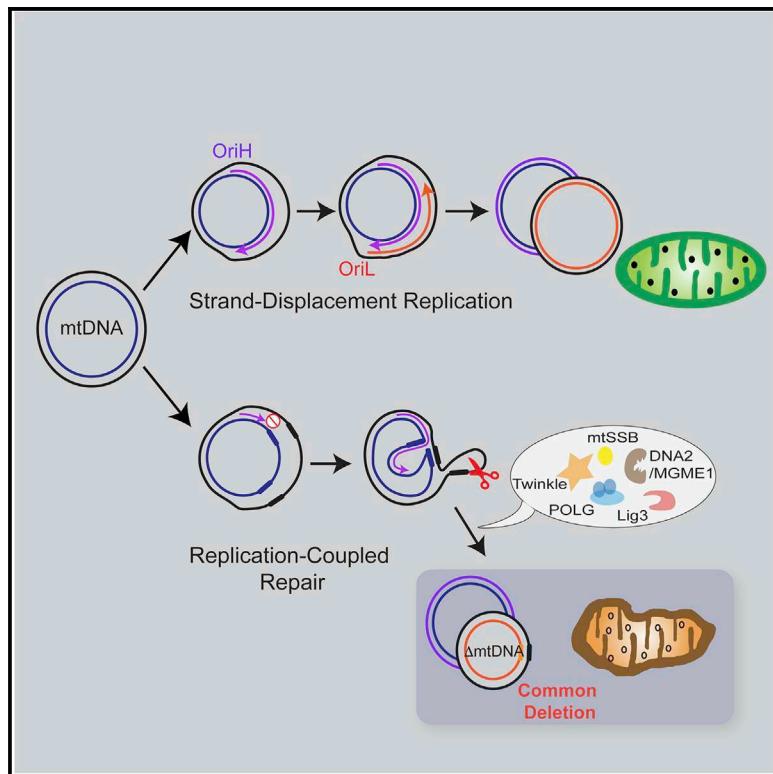
Nous avons étudié la réplication de l’ADN mitochondrial en adaptant la technique de peignage moléculaire sur l’ADN mitochondrial (laboratoire Agnel Sfeir). Les profils visualisés par peignage moléculaire montrent l’utilisation de deux origines de réplication, au niveau de l’ O_H et de l’ O_L . De plus, la protéine POLMRT est nécessaire à l’utilisation de l’origine de réplication O_L . L’ensemble des résultats nous ont permis de conclure à l’utilisation du modèle de déplacement de brin proposé dès les années 80. Le peignage moléculaire nous a aussi permis de conclure que dans le cas où la protéine Twinkle était mutée, nous observions un arrêt de la réplication dans la région de la « Common Deletion ».

En parallèle au laboratoire, j’ai étudié la formation de la « Common Deletion ». L’objectif était d’apporter des réponses sur les mécanismes qui mènent à la formation de délétions au sein de l’ADN mitochondrial. Pour mieux saisir ces mécanismes, nous avons tout d’abord mis au point l’induction de la « Common Deletion » en utilisant des nucléases artificielles adaptées à la mitochondrie, les mitoTALENs. Il s’est avéré qu’une seule mitoTALEN permettait d’induire la « Common Deletion ». Ensuite, notre approche par siARN a montré l’implication des protéines POLG, mtSSB, Twinkle, DNA2, MGME1 et PIF-1 dans la formation de la « Common Deletion ». Les protéines identifiées se sont révélées être principalement impliquées dans la réplication de l’ADN mitochondrial. Les protéines de réparation testées (CtIP, MRE11, RAD51) n’ont pas montré d’effet sur la formation de la délétion. Ceci implique que la « Common Deletion » induite par une cassure double-brin est le fruit d’une réparation illégitime induite lors de la réplication. Ainsi, nos travaux apportent-ils une meilleure compréhension de la réplication et de la réparation au sein de l’ADN mitochondrial.

Molecular Cell

Single-Molecule Analysis of mtDNA Replication Uncovers the Basis of the Common Deletion

Graphical Abstract



Authors

Aaron F. Phillips, Armel R. Millet,
Marco Tigano, ..., Marion Piganeau,
Erika Brunet, Agnel Sfeir

Correspondence

erika.brunet@inserm.fr (E.B.),
agnel.sfeir@med.nyu.edu (A.S.)

In Brief

Mutations in mtDNA are associated with human diseases and aging. Phillips et al. developed a single-molecule mtDNA combing approach to elucidate the molecular basis for the most frequent mtDNA aberration, the “common deletion,” and provide insights into mtDNA replication and repair.

Highlights

- Development of mito-SMARD, a single-molecule approach to study mtDNA replication
- *In vivo* analysis validates the strand displacement mode of mtDNA replication
- Twinkle helicase mutations promote fork stalling near the common deletion
- Repair of mito-TALEN-induced breaks triggers the common deletion

Single-Molecule Analysis of mtDNA Replication Uncovers the Basis of the Common Deletion

Aaron F. Phillips,^{1,4} Armel R. Millet,^{2,3,4} Marco Tigano,¹ Sonia M. Dubois,² Hannah Crimmins,¹ Loelia Babin,^{2,3} Marine Charpentier,² Marion Piganeau,² Erika Brunet,^{2,3,*} and Agnel Sfeir^{1,5,*}

¹ Skirball Institute of Biomolecular Medicine, Department of Cell Biology, NYU School of Medicine, New York, NY 10016, USA

² Structure et Instabilité des Génomes, Muséum National d'Histoire Naturelle, INSERM U 1154, CNRS UMR 7196, 75005 Paris, France

³ Genome Dynamics in the Immune System Laboratory, INSERM, UMR 1163, Institut Imagine, 75015 Paris, France

⁴ Co-first author

⁵ Lead Contact

*Correspondence: erika.brunet@inserm.fr (E.B.), agnel.sfeir@med.nyu.edu (A.S.)

<http://dx.doi.org/10.1016/j.molcel.2016.12.014>

SUMMARY

Mutations in mtDNA lead to muscular and neurological diseases and are linked to aging. The most frequent aberrancy is the “common deletion” that involves a 4,977-bp region flanked by 13-bp repeats. To investigate the basis of this deletion, we developed a single-molecule mtDNA combing method. The analysis of replicating mtDNA molecules provided *in vivo* evidence in support of the asymmetric mode of replication. Furthermore, we observed frequent fork stalling at the junction of the common deletion, suggesting that impaired replication triggers the formation of this toxic lesion. In parallel experiments, we employed mito-TALENs to induce breaks in distinct loci of the mitochondrial genome and found that breaks adjacent to the 5' repeat trigger the common deletion. Interestingly, this process was mediated by the mitochondrial replisome independent of canonical DSB repair. Altogether, our data underscore a unique replication-dependent repair pathway that leads to the mitochondrial common deletion.

INTRODUCTION

Human mtDNA consists of a 16.6-kb double-stranded circular genome and encodes 37 genes that are essential for oxidative phosphorylation (Anderson *et al.*, 1981). In addition to point mutations, mtDNA abnormalities comprise several large-scale deletions. The most common aberrancy is a 4,977-bp deletion spanning nucleotides 8,483–13,459 of the mitochondrial genome. The common deletion (also known as mtDNA⁴⁹⁷⁷) was first detected in neuromuscular diseases—Kearns-Sayre syndrome (KSS) and progressive external ophthalmoplegia (PEO) (Shoffner *et al.*, 1989)—and later found to accumulate in aging tissues (Corral-Debrinski *et al.*, 1992; Pikó *et al.*, 1988; Yang *et al.*,

1994). The loss of 12 essential mitochondrial genes within the deleted region cripples mitochondrial biogenesis. As a result, the common deletion is heteroplasmic in nature, manifesting clinically when mtDNA⁴⁹⁷⁷ molecules exceed a critical threshold (Wallace and Chalkia, 2013). The molecular mechanism underlying the formation of the common deletion remains unknown. The deleted region is flanked by two 13-bp repeats, with only one repeat being retained in mtDNA⁴⁹⁷⁷ (Schon *et al.*, 1989). This characteristic feature, shared by the majority of mtDNA deletions, implicates replication slippage or aberrant double-strand break (DSB) repair during the process of common deletion formation (Krishnan *et al.*, 2008; Srivastava and Moraes, 2005; Figure S1A, available online).

Over the years, the mode of mtDNA replication has been a matter of controversy (Figure S1B). The strand-displacement model was put forth in the 1970s based on the visualization of replicating mtDNA intermediates by electron microscopy (EM) (Robberson and Clayton, 1972). According to this model, mtDNA synthesis employs two distinct replication origins marked by unidirectional fork progression. Heavy-strand synthesis initiates from the origin of heavy-strand replication (OriH), located in the displacement loop, and advances for ~11 kb before the synthesis of the light chain initiates from the origin of light-strand replication (OriL) in the opposite direction. Later, Holt *et al.* (2000) proposed the strand-coupled model based on 2D gel electrophoresis analyses of mtDNA replication intermediates. According to this model, heavy- and light-strand replication is coupled and proceeds bi-directionally from a broad initiation zone downstream of OriH (Holt *et al.*, 2000). The strand-coupled model was later amended to include extensive RNA priming of the lagging strand, which is then replaced by DNA in subsequent maturation steps (Yang *et al.*, 2002; Yasukawa *et al.*, 2006).

mtDNA replication is accomplished with the help of a unique set of enzymes, distinct from the nuclear replisome. The minimal mitochondrial replisome was reconstituted *in vitro* and consists of polymerase gamma (POLG), the hexameric Twinkle helicase, and the mitochondrial single-stranded DNA binding protein (mtSSB) (Korhonen *et al.*, 2004). Additional factors were found to assist mtDNA synthesis, including the transcription and

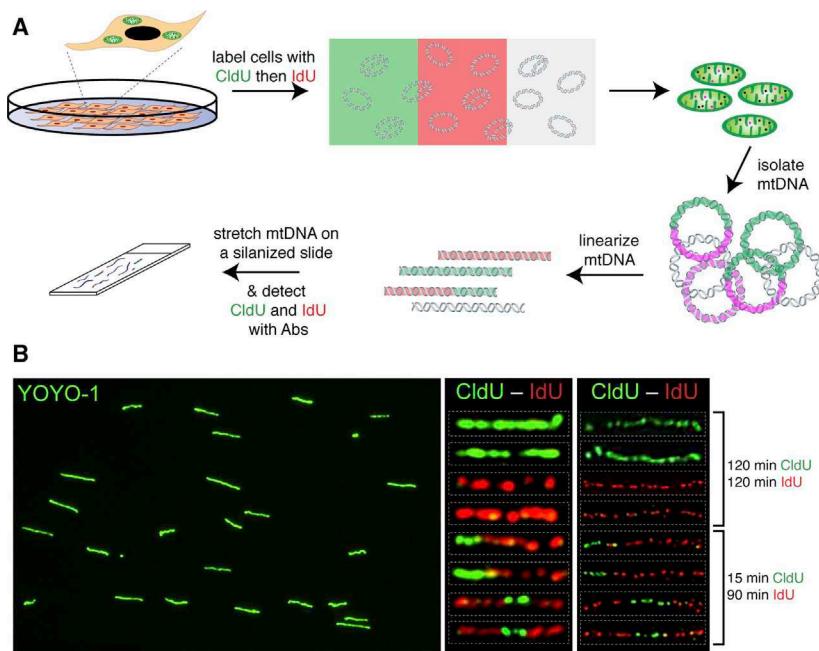


Figure 1. A Single-Molecule DNA Combing Approach to Visualize mtDNA Replication

(A) Schematic of mito-SMARD. Cells were sequentially pulse-labeled with media containing CldU and IdU, followed by a chase period. mtDNA was purified and digested with the Alel restriction enzyme, which cleaves upstream of OriH. Linearized DNA was then stretched on a silanized glass slide and stained with antibodies that distinguish CldU (green) from IdU (red).

(B) Left: stretched mtDNA molecules marked with a fluorescent DNA intercalator (YOYO-1). Center: examples of mtDNA molecules from HT1080 cells labeled with mito-SMARD and visualized with conventional microscopy (Nikon Eclipse Ti-V) using a $100\times/1.45$ NA oil objective. Shown are mtDNA molecules replicating in the presence of CldU (green molecules) and IdU (red molecules) as well as those replicating during the switch between CldU and IdU (dually labeled molecules). Right: mtDNA molecules detected by mito-SMARD and imaged with an OMX 3D-SIM super-resolution microscope (Applied Precision).

packaging protein TFAM (Falkenberg et al., 2007), the mitochondrial RNA polymerase (POLRMT) (Wanrooij et al., 2008), and the mitochondrial exonuclease MGME1 (Kornblum et al., 2013). Mutations in POLG and Twinkle have been implicated in mitochondrial disease and linked to large-scale mtDNA deletions (Spelbrink et al., 2001; Tyynismaa et al., 2005; Van Goethem et al., 2001). 2D gel analysis of mtDNA in cells expressing disease-associated Twinkle mutations highlights the presence of replication intermediates (Goffart et al., 2009; Wanrooij et al., 2007), indicative of replication fork pausing along the major arc of the mitochondrial genome. These observations provide circumstantial evidence in support of replication slippage as the basis of the common deletion (Figure S1A).

Originally, it was assumed that mitochondria lack the basic machinery to fix DSBs. However, recent reports have described DSB repair activities in this organelle (Bacman et al., 2009; Kryatsberg et al., 2004; Tadi et al., 2016), and several nuclear DSB repair proteins appear to localize to mitochondria (Dmitrieva et al., 2011; Sage et al., 2010; Simsek et al., 2011). Furthermore, introducing DSBs with restriction endonucleases triggers the accumulation of deletions across the mitochondrial genome (Srivastava and Moraes, 2005), thus implicating aberrant DSB repair during the process of deletion formation. The 13-bp repeats flanking the junction of the common deletion are predicted to promote microhomology-mediated end joining (MMEJ) of broken DNA (Samuels, 2004; Shoffner et al., 1989; Srivastava and Moraes, 2005; Figure S1A). MMEJ is an error-prone pathway of DSB repair that is associated with both deletions and the usage of microhomology and is dependent on LIG3, MRE11/CtIP, and PARP1 (Sfeir and Symington, 2015). Although mitochondrial protein extracts are known to stimulate MMEJ ex vivo (Tadi et al., 2016), in vivo evidence in support of a role for this erroneous pathway in common deletion formation is lacking.

Here we define the underlying molecular mechanism of mtDNA⁴⁹⁷⁷ formation in mammalian cells and distinguish between the contribution of replication slippage and DSB repair. We developed a single-molecule assay based on DNA combing technology to examine key features (mode, origin, and directionality) of mtDNA replication and investigate stalling of the replisome in the vicinity of the common deletion. In parallel, using mitochondrially targeted transcription activator-like effector nucleases (mito-TALENs) to induce breaks in specific loci of the mitochondrial genome, we investigated the underlying genetic basis of mitochondrial DSB repair. Our results identify a unique replication-dependent repair pathway that triggers the formation of the mitochondrial common deletion. Understanding the basis of this widespread aberration is an essential first step to prevent its accumulation in mitochondrial diseases.

RESULTS

A Single-Molecule Approach to Study mtDNA Replication

We aimed to investigate the dynamics of mtDNA replication at the single-molecule level and with high resolution. To do so, we adapted the single-molecule analysis of replicating DNA (SMARD) technique (Norio and Schildkraut, 2001), a DNA combing method that has been widely applied to study the replication of specific loci within the nuclear genome. This method, which we named mito-SMARD (Figure 1A), is based on labeling cells sequentially with two halogenated thymidine derivatives (5-Iodo-2'-deoxyuridine [IdU] and 5-Chloro-2'-deoxyuridine [CldU]) that are incorporated into replicating DNA molecules. After a short recovery period, during which cells are incubated with media not supplemented with any thymidine derivatives, circular mtDNA was isolated from lysed cell extracts using a miniprep kit (Quispe-Tintaya et al., 2013). The DNA was linearized with a restriction enzyme that cleaves upstream of OriH and resolved by gel electrophoresis. The isolated 16.6-kb mtDNA fragment

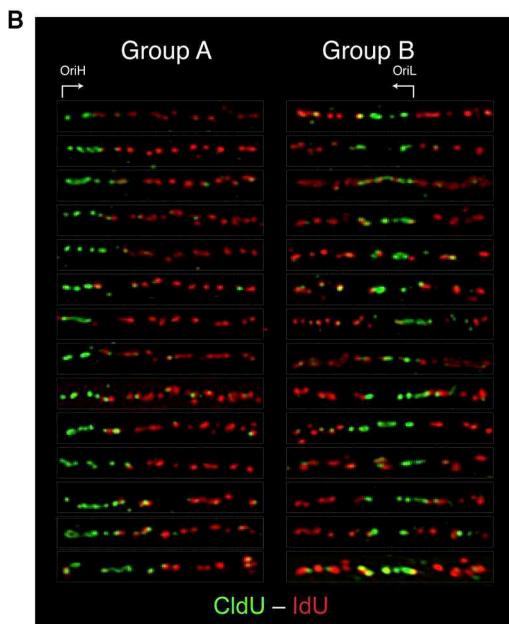
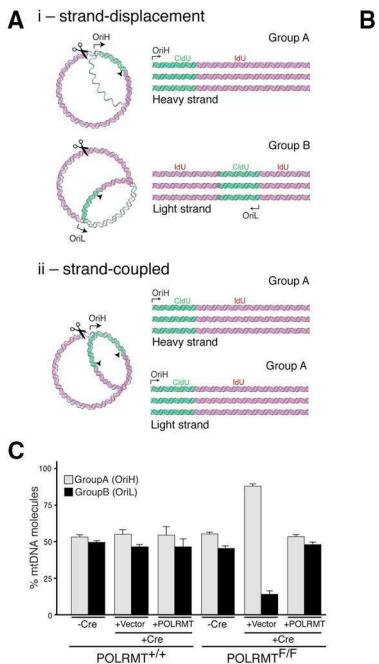


Figure 2. Mito-SMARD Highlights the Mode of mtDNA Replication In Vivo

(A) Predicted pattern of CldU and IdU incorporation based on the strand-coupled or strand displacement models. Taking into consideration dual-labeled molecules, the strand-coupled model is expected to yield one replication pattern in which green is followed by red. In contrast, according to the strand displacement model, two independent groups of molecules will be visualized: one labeled with green followed by red (group A) and an equivalent fraction with red flanked by green on either side (group B).

(B) Mito-SMARD analysis of mtDNA in HT1080 cells yields molecules that are categorized in two distinct groups: group A molecules show CldU (green) followed by IdU (red) labeling, whereas group B molecules show an IdU-CldU-IdU labeling pattern.

(C) Quantification of the frequency of group A versus group B molecules in *Polrmt^{F/F}* and *Polrmt^{+/+}* cells with the indicated treatment. Cells were analyzed 48 hr after induction of a 4-hydroxytamoxifen-inducible Cre-ER^{T2}. Average frequencies were calculated from three independent experiments ($n \geq 100$ molecules in each experiment) with SD.

was then stretched on silanized glass slides, and the two analogs were detected with antibodies that distinguish CldU (green) from IdU (red). The mito-SMARD procedure yielded mtDNA molecules that were substituted with either of the two analogs and, therefore, replicated fully in the presence of either CldU or IdU. In addition, a significant fraction of mtDNA molecules were labeled with both dyes, indicating that such molecules were synthesized during the switch between CldU to IdU (Figures 1B, S2A, and S2B). This class of molecules was especially informative to determine the speed, direction, and origin of the replication fork. Super-resolution microscopy using a DeltaVision optical microscope experimental (OMX) enabled us to more accurately delineate the boundaries between CldU and IdU (Figure 1B). A key optimization step when developing mito-SMARD involved the identification of the appropriate pulse-labeling periods. Because it takes more than 1 hr to completely duplicate an mtDNA molecule (Clayton, 1982; Korhonen et al., 2004), the optimized procedure consisted of 15-min pulses of CldU (30 min in the case of mouse cells) and 90-min pulses of IdU, followed by a 2 hr chase. Isolation of mtDNA using a miniprep kit was shown previously to result in ~2,000-fold enrichment of mtDNA molecules (Quispe-Tintaya et al., 2013). To confirm the purity of the isolated DNA, we monitored the levels of ND5 mtDNA relative to actin. Indeed, the results validate the purification method and confirm a significant enrichment for mtDNA in samples used in mito-SMARD (Figure S2C).

mito-SMARD Supports the Strand Displacement Model of mtDNA Replication

We first exploited mito-SMARD to uncover key features of mtDNA replication, including mode, origin, and directionality. To that end, we examined the pattern of CldU/IdU incorporation in the dual-labeled mtDNA molecules. We anticipated

that mito-SMARD analysis would allow us to distinguish between the two opposing models for mtDNA replication (Figure 2A). Should the mtDNA be replicated asymmetrically—according to the strand-displacement model—mito-SMARD would yield molecules with two distinct labeling patterns (groups A and B) (Figure 2A). Heavy-strand synthesis from OriH will appear as molecules labeled with CldU in green at the 5' end, followed by IdU in red (group A), whereas replication of light-strand through OriL will result in a set of molecules in which CldU is flanked with IdU on both sides (group B). In contrast, the strand-coupled model, which involves firing of a single origin, is expected to yield one labeling pattern. Specifically, synchronous replication of mtDNA molecules from OriH would result in a labeling pattern consisting of green (CldU) followed by red (IdU, group A) (Figure 2A). Inspection of dual-labeled mtDNA molecules in human and mouse cells revealed two distinct labeling patterns representing group A and group B (Figures 2B and S2B). The visualization of these two configurations with similar frequency (Figures 2B, 2C, and S2B) underscores efficient firing of both OriH and OriL and is, therefore, consistent with the strand-displacement model of mtDNA replication (Clayton, 1982; Robberson and Clayton, 1972).

The initiation of light-strand synthesis has been proposed to be mediated by POLRMT, an RNA polymerase that acts as an OriL-specific primase upon recognizing a stem-like loop within the lagging-strand origin (Fusté et al., 2010; Wanrooij et al., 2008). To validate the mito-SMARD results and confirm that group B molecules are products of OriL firing, we tested the effect of depleting POLRMT on the efficiency of lagging-strand synthesis in vivo. To do so, we treated *Polrmt^{F/F}* and *Polrmt^{+/+}* mouse embryonic fibroblasts (MEFs) (Kühl et al., 2014) with Cre recombinase (Figure S3A). POLRMT-deficient and -proficient cells were pulse-labeled with CldU and IdU 2 days after Cre treatment and analyzed by mito-SMARD. We chose to investigate an early

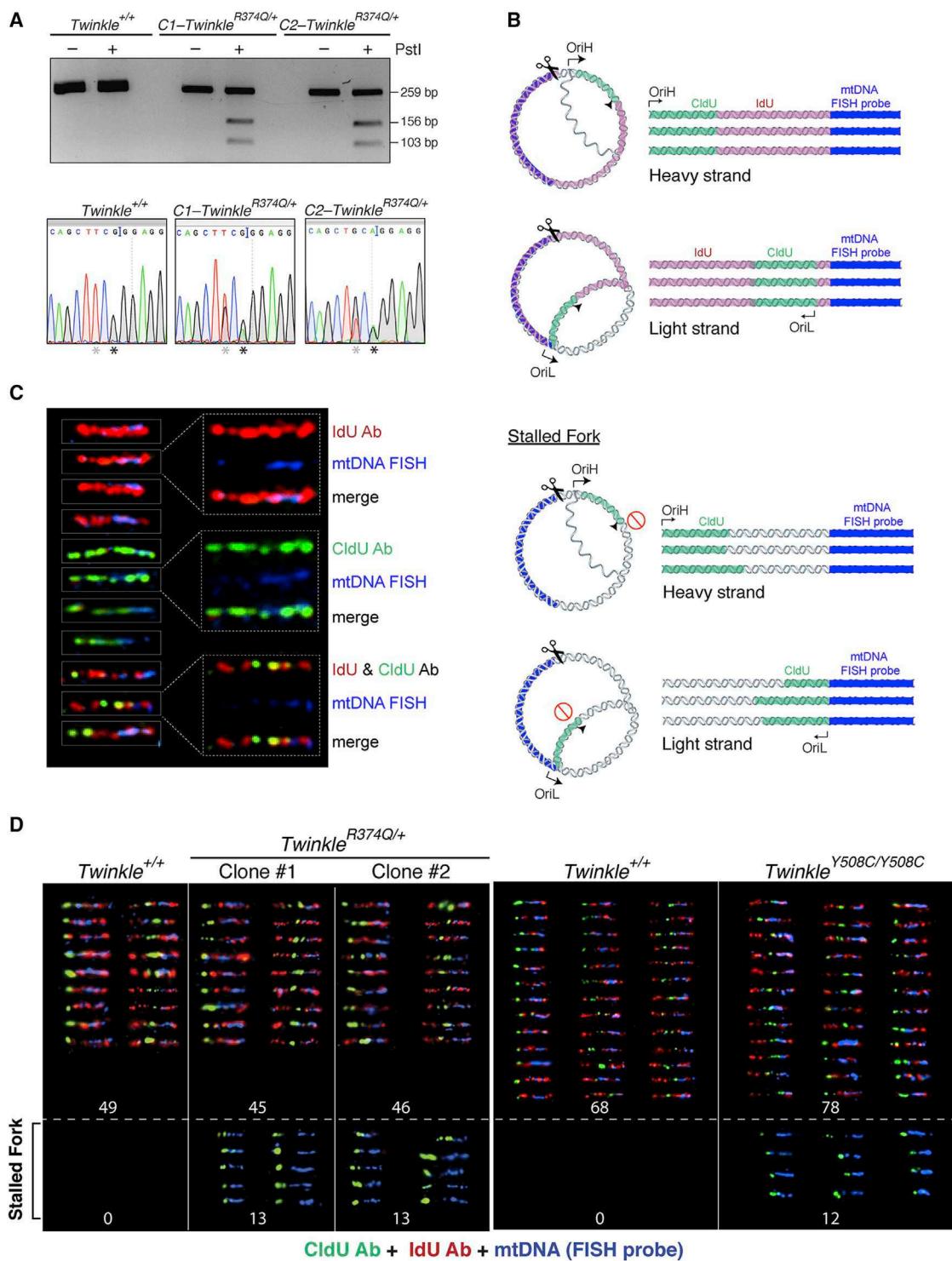


Figure 3. Evidence for Replication Fork Stalling in Cells Expressing Mutant Twinkle

(A) Genotyping PCR to detect heterozygous Twinkle^{R374Q/+} generated by CRISPR/Cas9 targeting of RPE-1. The PCR product was treated with a restriction enzyme that specifically cleaves the targeted allele. Two independently derived clonal cell lines were generated. Targeting was confirmed by Sanger sequencing of cDNA from the indicated clones. The asterisks mark mutated residues, including the R374Q mutation and a silent mutation in the protospacer adjacent motif (PAM) sequence.

(legend continued on next page)

time point to avoid the confounding effect of POLRMT inhibition on mtDNA transcription and mtDNA content (Figures S3A–S3E). Depletion of POLRMT impaired firing of OriL, as depicted by a significant reduction in the frequency of group B molecules (Figure 2C). As a control, stable expression of exogenous POLRMT in cells deficient for this RNA polymerase rescued OriL firing (Figures 2C and S3F). The demonstration that POLRMT-mediated priming initiates mtDNA synthesis from OriL provides further evidence in support of the strand-displacement model of mtDNA replication (Figure S1B, top; Clayton, 1982; Robberson and Clayton, 1972) and argues against both strand-coupled mtDNA replication (Figure S1B, center; Holt et al., 2000) and discontinuous lagging-strand synthesis via the RNA incorporation throughout the lagging strand (RITOLS) model (Figure S1B, bottom; Yang et al., 2002; Yasukawa et al., 2006).

Using CRISPR/Cas9 to Establish Cells with Disease-Associated Twinkle Mutations

Having developed mito-SMARD to explore the link between common deletion and replication infidelity, we next tested whether disease-associated Twinkle mutations, previously linked to the formation of the mitochondrial common deletion (Spelbrink et al., 2001; Tyynismaa et al., 2005), induce replication fork stalling in the vicinity of the 4,977-bp region. Pausing of the fork is expected to foster the mis-annealing of the flanking repeats, a key step during the genesis of the common deletion. In the first instance, we utilized the previously established Flp-In T-Rex-inducible cell lines that express wild-type and mutant variants of Twinkle (Goffart et al., 2009; Figure S4A; data not shown). Induction of Twinkle in this heterologous setting was achieved by treating cells with doxycycline. Upon examining the mtDNA replication profile, we observed that the expression of either wild-type or mutant Twinkle led to similar impairment of mtDNA replication. Specifically, analysis of the ratio of group A to group B molecules revealed a significant bias toward mtDNA replication through OriH in cells treated with doxycycline (Figure S4B; data not shown). These results underscore the limitation of the Flp-In system when examining the function of mitochondrial replication factors. It is possible that overexpression of a core component of the replisome sequesters the remaining subunits of the complex, leading to dysregulated mitochondrial replication. Alternatively, tetracycline was recently shown to impair mitochondrial function (Moullan et al., 2015) and could indirectly affect mtDNA replication.

To avoid artifacts emanating from Twinkle overexpression and/or doxycycline treatment and to address the function of Twinkle mutations *in vivo*, we employed the clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated

protein 9 (Cas9) gene-editing technology and derived independent Twinkle^{R374Q/+} and Twinkle^{Y508C/Y508C} sub-clones from normal human somatic RPE-1 cells. R374Q is a dominant mutation in Twinkle that has been previously identified in patients with mitochondrial disorders and lies in the linker region that separates the helicase domain from the N-terminal primase-like domain (Martin-Negrier et al., 2011). Y508C is a recessive disease-associated Twinkle mutation (Nikali et al., 2005) within the helicase domain. Genotyping PCR and sequence analysis revealed mono-allelic editing for the dominant R374Q mutations and bi-allelic editing for the Y508C mutant (Figures 3A, S4C, and S4D). The mtDNA content in Twinkle-mutant cells, as determined by qPCR, was similar to that of wild-type cells, confirming that the presence of the disease-linked mutations did not completely abolish mtDNA replication (Figure S4E).

Evidence for Replication Fork Stalling in Cells Expressing Mutant Twinkle

We next investigated whether disease-associated mutations in Twinkle lead to replication fork stalling. We anticipated that transient pausing of the replisome during the 10-min CldU pulse or shortly after the switch to IdU would lead to a class of molecules labeled with green because of CldU incorporation but little or no red denoting IdU incorporation (Figure 3B). To distinguish fork stalling events from partially labeled fragments emanating from broken mtDNA molecules, we further amended mito-SMARD and included a library of fluorescence *in situ* hybridization (FISH) probes to mark the end of the stretched mtDNA molecules (spanning nucleotides 10,000 and 16,400 of the mitochondrial genome) (Figures 3C, S4F, and S4G).

Upon analyzing molecules obtained from independent clonally derived cells comprising Twinkle^{R374Q/+} and Twinkle^{Y508C/Y508C} alleles, we observed fibers that contained CldU but no IdU labeling (Figures 3D and S4H). Importantly, this pattern, which is indicative of fork stalling, was not identified in cells carrying wild-type Twinkle. Although the number of stalled molecules was small, they were solely found in cells with mutant Twinkle and observed in independent experiments (Figures 3D and S4H), providing qualitative evidence in support of fork stalling in the vicinity of the common deletion. In conclusion, based on the analysis of mtDNA replication *in vivo*, we conclude that the onset of the common deletion in human cells is potentially coupled to replication fork stalling adjacent to the 4,977-bp deleted region.

Breaks at the 5' End of the Deletion Breakpoint Trigger mtDNA⁴⁹⁷⁷

So far, our analysis of mtDNA replication using mito-SMARD provided *in vivo* evidence in support of the replication model of

(B) Schematic for mito-SMARD to assess replication fork stalling. FISH probes that label the control region of the mitochondrial genome were used to mark the end of the linearized mtDNA molecules stretched on glass slides. Top: schematic of molecules representative of fully replicated mtDNA. Bottom: schematic of partially labeled molecules indicative of replication fork stalling.

(C) Representative mtDNA molecules labeled with CldU (green), IdU (red), and CldU-IdU (green-red) and stained with mito-FISH probes (blue).

(D) Evidence of replication fork stalling in the major mtDNA arc in Twinkle^{R374Q/+} and Twinkle^{Y508C/Y508C} cells. Shown are mtDNAs obtained from cells expressing mutant and wild-type Twinkle. Top: examples of fully replicating mtDNA molecules with IdU/CldU incorporation patterns consistent with replication initiation from OriL and OriH. Bottom: molecules represent labeling patterns consistent with fork stalling. It is important to note that the duration of the IdU pulse (90 min) allows mtDNA molecules that have initiated replication during the CldU pulse to fully complete synthesis (rate of mtDNA synthesis, ~190 nt/min) (data not shown; Clayton, 1982; Korhonen et al., 2004).

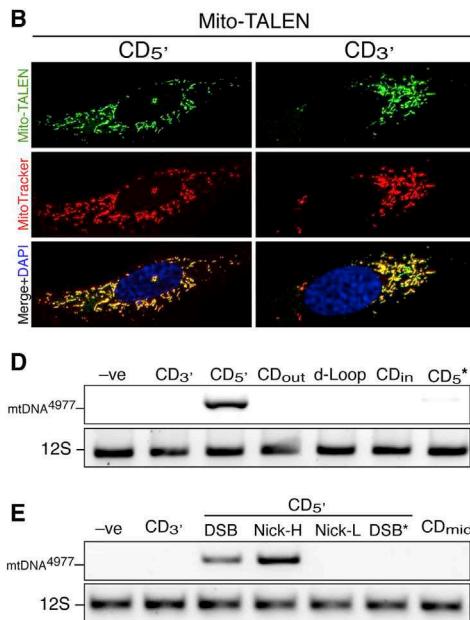
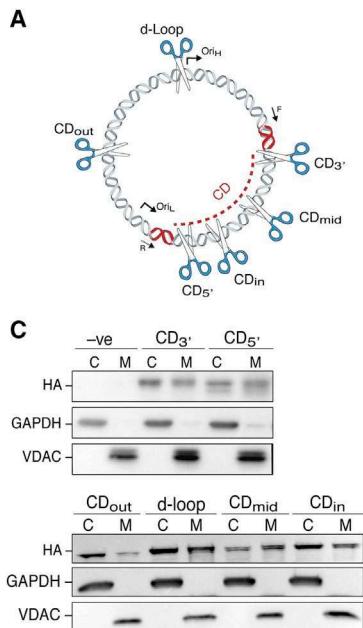


Figure 4. Engineering TALENs to Induce the Mitochondrial Common Deletion

(A) Schematic of mtDNA with cleavage sites for each mito-TALEN (D-loop, CD_{3'}, CD_{mid}, CD_{5'}, CD_{out}). Indicated in red are the two 13-bp repeats surrounding the common deletion (CD). The primers used for PCR amplification of the common deletion are also shown.

(B) Immunofluorescence microscopy of U2OS cells transfected with mito-TALENs that target CD_{5'} (left) and CD_{3'} (right). Hemagglutinin (HA)-tagged mito-TALEN was visualized 48 hr after transfection (top), MitoTracker marks mitochondria (center), and DAPI marks nuclei.

(C) Fractionation experiments followed by western blot analysis to detect mito-TALENs (HA tag) in the cytoplasm (C) and mitochondria (M). Glyceraldehyde 3-phosphate dehydrogenase (GAPDH) marks the cytosolic fraction, and voltage-dependent anion-selective channel (VDAC) denotes the mitochondrial fraction. (-ve), no mito-TALEN.

(D) PCR detection of mtDNA⁴⁹⁷⁷ in U2OS cells treated with the indicated mito-TALEN and harvested 7 days after transfection. The asterisk marks a catalytically inactive mito-TALEN-CD_{5'}.

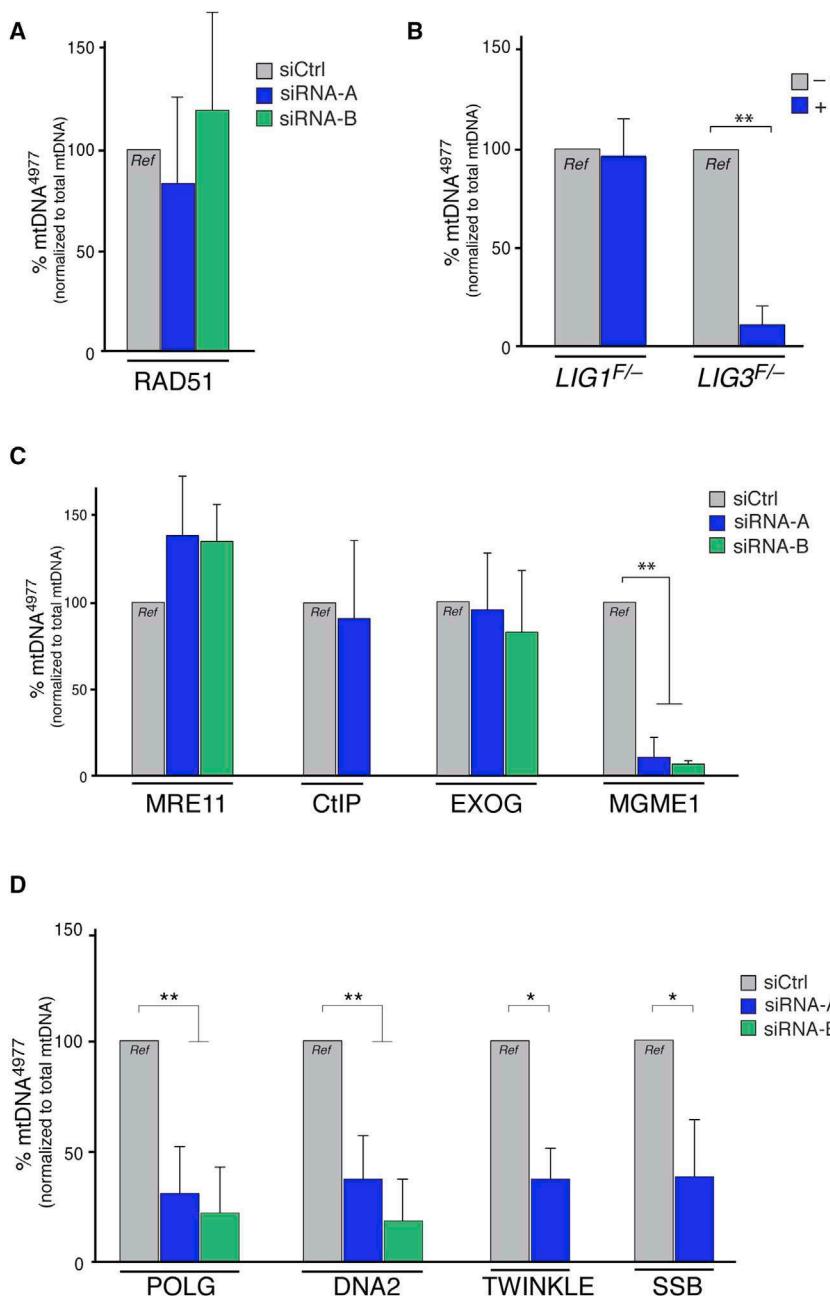
(E) PCR detection of the common deletion in cells treated with the indicated mito-TALEN nickase (Nick-H induces a break in the heavy strand, Nick-L induces a break in the light strand) and nuclease (DSB).

common deletion formation. In an independent set of experiments, we aimed to investigate the role of DNA break repair by MMEJ during common deletion formation. To do so, we engineered mito-TALENs capable of inducing DSBs in distinct loci within the mitochondrial genome, including D-loop, common deletion (CD_{3'}), CD_{5'}, CD_{mid}, CD_{in}, and CD_{out} (Figure 4A). Immunofluorescence analysis as well as fractionation experiments confirmed the import of mito-TALENs to the mitochondria (Figures 4B, 4C, and S5A–S5C). In addition, using a T7 assay, we verified that the different TALENs cleave the mtDNA target sequence with similar efficiency *in vivo* (Figures S5C–S5F). To monitor the effect of the different mito-TALENs on the common deletion, we used a nested PCR designed to detect mtDNA⁴⁹⁷⁷ molecules. We observed that generation of a DSB internal to the 5' repeat (CD_{5'}) was sufficient to trigger the common deletion. In contrast, breaks incurred within the D-loop, CD_{out}, CD_{mid}, CD_{in}, and CD_{3'} did not trigger the accumulation of mtDNA⁴⁹⁷⁷ (Figures 4D and S5D). Furthermore, no deletion was formed when a DSB was induced within or external to the 5' repeat (CD_{5'} repeat or CD_{5'} external) or when using an additional alternative CD_{3'}bis mito-TALEN (Figures S5A and S5E–S5H). Analysis of the junction by next generation (Next-Gen) sequencing confirmed the loss of one of the two 13-bp repeats and the intervening sequences (Figures S5I and S5J). We next assessed whether single-stranded breaks (strand-specific) are sufficient to result in the common deletion. To do so, we used a combination of catalytically inactive and wild-type FokI domains (Wu et al., 2014) and designed mito-TALEN pairs capable of inducing nicks in the heavy or light strands. Our results demonstrate that introducing a single-strand break on the heavy strand of CD_{5'} led to the accumulation of mtDNA⁴⁹⁷⁷ molecules. However, nicks at CD_{3'} and

on the light strand of CD_{5'} failed to stimulate the mitochondrial common deletion (Figures 4E, S5K, and S5L). In conclusion, by monitoring the outcome of breaks induced at distinct loci within the mitochondrial genome, we show that mtDNA⁴⁹⁷⁷ molecules specifically accumulate as a result of breaks adjacent to the 5'-end repeat.

A Critical Role for LIG3 during the Accumulation of mtDNA⁴⁹⁷⁷

The induction of the common deletion in response to mito-TALEN-CD_{5'} provided us with a robust tool to decipher the genetic requirement for mtDNA⁴⁹⁷⁷ formation. Candidate genes included nuclear DSB repair factors that localize to the mitochondrial matrix as well as mitochondrial-specific factors implicated in mtDNA metabolism. RAD51-mediated recombination has been shown to stimulate mtDNA rearrangements in yeast (Stein et al., 2015), and human RAD51 is found in the mitochondrial matrix (Sage et al., 2010). We observed that the levels of mtDNA⁴⁹⁷⁷ accruing in response to mito-TALEN-CD_{5'} were not reduced upon depletion of RAD51 (Figures 5A and S6A). This observation potentially ruled out inappropriate recombination as the underlying basis of the common deletion and prompted us to ask whether the deletion is contingent on end-joining activity. Of the three DNA ligases involved in replication and repair in mammalian cells, LIG3 is the only enzyme that localizes to the mitochondria and is crucial for mtDNA integrity (Gao et al., 2011; Simsek et al., 2011). To investigate the role of LIG3 during common deletion formation, we introduced a plasmid expressing mito-TALEN-CD_{5'} into *LIG3*^{F/F} HCT116 cells. Following depletion of LIG3 with Cre recombinase, we noted a marked decrease in the frequency of mtDNA⁴⁹⁷⁷. As a control, we



showed that the frequency of mtDNA⁴⁹⁷⁷ formation was not affected by the loss of LIG1 (Figures 5B and S6A). It is important to note that the levels of mtDNA⁴⁹⁷⁷ were normalized to the total mtDNA content in each setting (Figure S6B). In summary, these results invoke a LIG3-mediated end joining step during the process of common deletion formation.

Induction of the Common Deletion Is Dependent on mtDNA Replication

Prior to joining of broken DNA ends, end resection is an essential step to degrade the intervening sequence, resulting in mtDNA⁴⁹⁷⁷. To decipher this key processing step, we targeted

Figure 5. Replication-Mediated Repair Leads to the Common Deletion

(A) qPCR analysis of mtDNA⁴⁹⁷⁷ in cells treated with mito-TALEN-CD_{5'}. The amount of mtDNA⁴⁹⁷⁷ was determined 7 days after mito-TALEN transfection and normalized to total mtDNA (12S). The frequency of mtDNA⁴⁹⁷⁷ in cells treated with siRAD51 is presented relative to siControl. (B) Relative percentage of mtDNA⁴⁹⁷⁷ accumulation in LIG1^{F/-} and LIG3^{F/-} cells after Cre induction and mito-TALEN expression. (C) qPCR analysis of mtDNA⁴⁹⁷⁷ induced by mito-TALEN-CD_{5'} in cells depleted for the indicated nucleases and analyzed 7 days after transfection. (D) Induction of mtDNA⁴⁹⁷⁷ upon inhibition of replication in cells treated with the indicated siRNA and analyzed 7 days after transfection. The frequency of mtDNA⁴⁹⁷⁷ in all genetic perturbations was normalized to mtDNA content (12S) and is presented relative to siControl. Values represent the mean of three independent experiments with SD.

a number of mitochondrion-residing nucleases, including MRE11 and CtIP (Dmitrieva et al., 2011; Tadi et al., 2016)—key components of the MMEJ pathway. In addition, we inhibited EXOG, which is implicated in mitochondrial single-strand break repair (Cymerman et al., 2008), and MGME1, a mitochondrion-specific nuclease that acts together with POLG to process DNA flap intermediates during replication (Kornblum et al., 2013). To identify the resecting nuclease responsible for common deletion formation, we used a small interfering RNA (siRNA)-based approach and depleted MRE11, CtIP, and MGME1 from HCT116 cells (Figures 5C, S6A, and S6B). Concomitantly, we introduced plasmids encoding mito-TALEN-CD_{5'} and monitored the frequency of mtDNA⁴⁹⁷⁷. We observed no differences in the levels of mtDNA⁴⁹⁷⁷ upon inhibition of repair-related nucleases—EXOG, CtIP, and MRE11. Instead, upon depletion of the replicative nuclease

MGME1, the levels of common deletion were significantly reduced (Figure 5C). To account for the variation in mtDNA content because of these genetic perturbations, the common deletion was normalized to total mtDNA (Figure S6B).

Given the role of MGME1 during mtDNA replication, we wondered whether other components of the mitochondrial replisome are also implicated in the formation of mtDNA⁴⁹⁷⁷. To that end, we treated HCT116 cells with siRNA to deplete POLG, DNA2, Twinkle, and mtSSB independently. We then expressed mito-TALEN-CD_{5'} and monitored the frequency of the common deletion by qPCR. Our results reveal an attenuation in mtDNA⁴⁹⁷⁷ upon inhibition of various members of the

mitochondrial replisome (Figures 5D and S6A–S6C). Similarly, treatment of cells with ethidium bromide (EtBr), known to block mtDNA replication, led to a significant decrease in the frequency of TALEN-induced common deletion (Figures S6D–S6F). In contrast, depletion of PIF1, an ATP-dependent helicase that unwinds secondary structures and ensures progression of the replication fork (Paeschke et al., 2011), led to a 2-fold increase in the common deletion in response to TALEN cleavage of the mtDNA (Figures S6G and S6H). Altogether, our results highlight a replication-dependent form of DNA repair acting in the mitochondria and capable of stimulating the common deletion.

DISCUSSION

The mitochondrial common deletion was among the first identified mitochondrial mutations causative of human diseases and is currently one of 200 documented pathogenic mtDNA deletions (<http://www.mitomap.org/>). Deletions in mtDNA are predominantly located on the major arc of the genome and are typically flanked by direct repeats (Samuels et al., 2004), potentially implicating a common underlying mechanism that triggers their formation. The experiments described here define the mechanism that leads to mitochondrial common deletion *in vivo*. By deciphering mtDNA replication and repair, we highlight a replication-dependent mode of DNA repair that is unique to mitochondria and leads to accumulation of the toxic common deletion (Figure S1A, bottom).

The Mechanism of Common Deletion Formation

It has been proposed that the mtDNA common deletion is triggered by replication slippage, MMEJ-mediated repair of DSBs, or the combination of defective replication coupled with breaks in the DNA (Figure S1A). According to the DSB repair hypothesis, any break internal to the two 13-bp repeats is expected to undergo resection and re-joining, ultimately yielding mtDNA⁴⁹⁷⁷ molecules. On the contrary, our data demonstrate that only breaks adjacent and internal to the 5'-end breakpoint trigger formation of the common deletion (Figures 4D, 4E, and S5D–S5H). Furthermore, we provide genetic evidence indicating that formation of the common deletion in response to TALEN-induced breaks is independent of the core MMEJ factors MRE11 and CtIP (Figure 5C). Altogether, these results argue against a role for canonical MMEJ during common deletion formation. Instead, our data demonstrate that mtDNA break repair is executed by components of the mitochondrial replisome and coupled to replication (Figures 5D and S6D). According to the replication-dependent repair model (Figure S1A, bottom), frequent stalling of the replication fork facilitates the mispairing of the two 13-bp repeats, thereby exposing a long single-stranded loop on the heavy strand. Damage incurred to the looped DNA triggers loss of the intervening sequence, resulting in an mtDNA⁴⁹⁷⁷ genome. In effect, impaired mtDNA replication combined with aberrant break repair lead to the common deletion.

G4 DNA-forming sequences, typically unwound by PIF1 (Duan et al., 2015), are prevalent in regions proximal to the deletion breakpoints (Bharti et al., 2014) and could, therefore, contribute to stalling of the replisome. Similarly, Twinkle mutations impede fork progression to foster mispairing of the 13-bp repeats and

stimulate common deletion formation. The mis-annealing of homologous repeats flanking the deleted region represents a crucial step during common deletion formation. The exposed heavy strand is typically covered with mtSSB, but the factor(s) that promote annealing of the complementary junctions remain(s) unknown. Twinkle possesses strand-annealing activity that could potentially promote the annealing of flanking repeats (Sen et al., 2012), although we cannot exclude the involvement of additional factors.

According to the replication-dependent repair mode of the common deletion, nicks anywhere within the ssDNA loop are expected to induce the common deletion. On the contrary, our data indicate that TALEN-induced nicks adjacent to the 3' end of the common deletion are refractory to common deletion formation (Figures 4E and S5L). This contradiction is inherent to the source of the damage in our experimental setup. TALENs recognize double-stranded DNA and will, therefore, induce breaks prior to the unwinding of the double helix during replication. As a result, breaks at the 3' end of the deleted region are expected to obstruct the progression of the fork and subsequent unwinding of the DNA and, therefore, hinder formation of the common deletion.

A Method to Address a Fundamental Question Related to the Mode of mtDNA Replication

Replication of mtDNA has been classically studied using 2D gels and EM analysis (Clayton, 1982; Robberson and Clayton, 1972), with both methods leading to contradictory results as to what is the exact mode of mtDNA replication. Here we developed a method based on DNA combing to examine mtDNA replication at the single-molecule level and with high resolution. DNA combing has long been used to address key questions related to nuclear DNA replication. By adapting this powerful technology to the specifics of DNA synthesis in mitochondria, we observed efficient firing of two replication origins (Figure 2), thus confirming the strand displacement model of mtDNA replication (Clayton, 1982; Robberson and Clayton, 1972; Figure S1B, top). In addition, we provided genetic evidence in support of a role for POLRMT in the initiation of OriL *in vivo* (Figure 2C). These data are consistent with previous experiments demonstrating that POLRMT primes lagging strand synthesis *in vitro* by recognizing a specific loop in the unwound OriL (Fusté et al., 2010; Wanrooij et al., 2008) and argues against strand-coupled mtDNA replication (Figure S1B, center; Holt et al., 2000) as well as discontinuous lagging-strand synthesis via the RITOLS model (Figure S1B, bottom; Yang et al., 2002; Yasukawa et al., 2006). Whether a protein factor (such as mtSSB) or RNA coats the lagging strand prior to OriL firing remains to be settled.

It is worth noting that the analysis of mtDNA replication was carried out in cells in which DNA synthesis served to maintain steady-state levels of mtDNA molecules following normal turnover. In the future, it will be informative to investigate whether the mode of mtDNA replication is altered in cells undergoing mtDNA repopulation following a sudden drop in mtDNA content or in response to increased energy demand.

The Endogenous Source of DNA Damage

Our study raises an obvious question related to the physiological source of damage that induces breaks in mtDNA. The

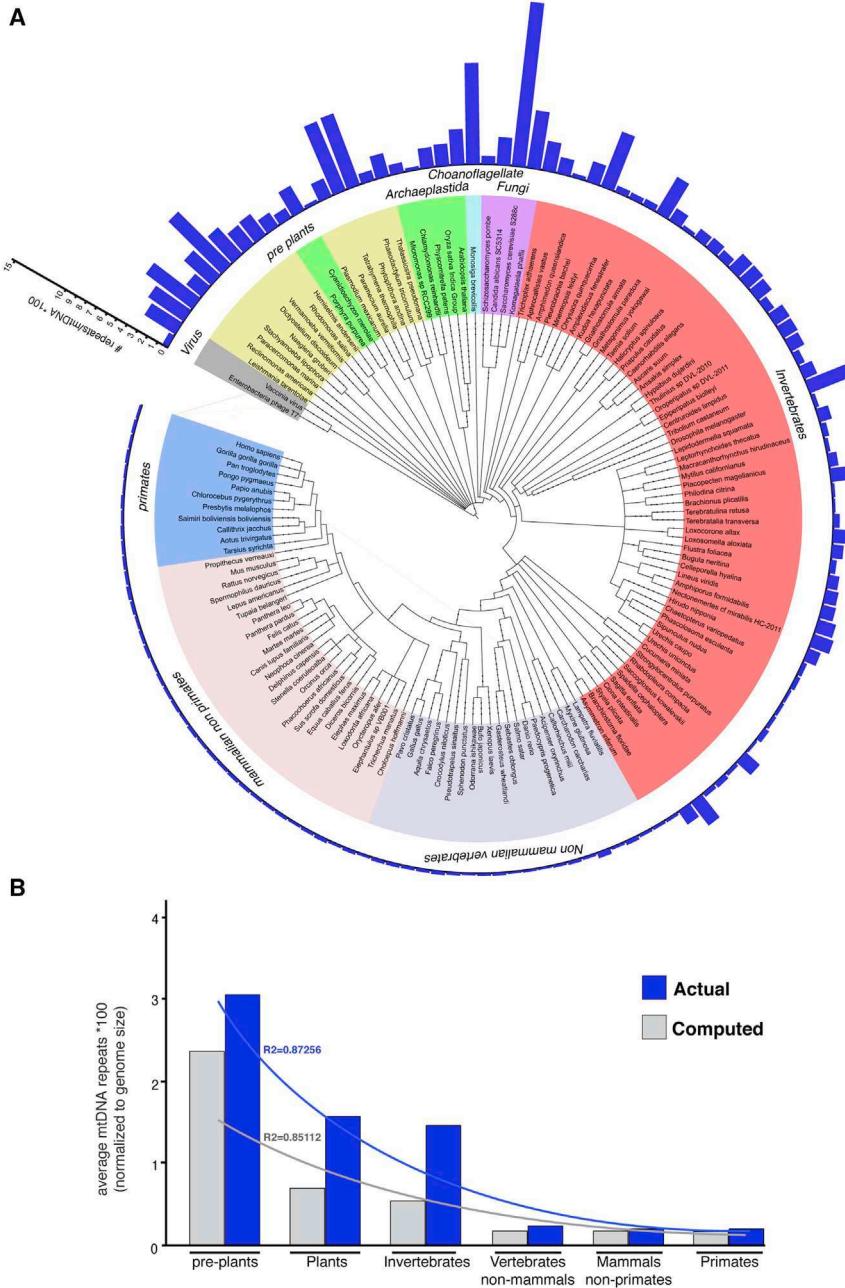


Figure 6. The Evolution of Repeats in the Mitochondrial Genome

(A) Phylogenetic distribution of mtDNA repeats of 142 species. Phylogenetic relationships were analyzed through the NCBI Taxonomy Database, and organisms are divided into evolutionary groups. Bars represent the total number of direct and inverse repeats (>12 bp) normalized to the size of the genome.

(B) Bar graph representing the average number of repeats compared with a theoretical number computed for each evolutionary group, taking into account the average size of the mitochondrial genome and the percentage of GC content.

increased mtDNA deletions (Kokoszka et al., 2001). Alternatively, mtDNA breaks could result from the uncontrolled activity of DNA metabolic enzymes. For example, cleavage of abasic sites by apurinic/apyrimidinic endonuclease (APEX1) (Tomkinson et al., 1988) would induce mtDNA breaks, particularly when subsequent base excision repair steps are ineffective. Similarly, abortive topoisomerase I (TOP1) activity (Zhang et al., 2001) could generate DNA nicks that would then trigger formation of the common deletion. Lastly, the 5'-end breakpoint of the common deletion is flanked by AT-rich sequences that assume a distorted structure in vitro (Hou and Wei, 1998), potentially creating a hotspot for DNA damage.

DSB Repair in Mitochondria

Cells evolved three independent pathways (homologous recombination [HR], non-homologous end joining [NHEJ], and MMEJ) to repair nuclear DSBs. Despite the early notion that mitochondria lack the imperative for DSB repair, indirect evidence suggests that breaks in this extra-nuclear genome can be processed by DSB repair pathways (Bacman et al., 2009; Srivastava and Moraes, 2005).

Furthermore, it has been hypothesized that misrepair of DSBs results in deletions, including the common deletion (Krishnan et al., 2008). In yeast, members of the RAD52 epistasis groups, which are involved in HR, were recently shown to promote DSB repair of a reporter construct in a way that triggers the deletion of sequences flanked by direct repeats (Phadnis et al., 2005; Stein et al., 2015). Illegitimate recombination has also been proposed to underlie deletion formation in human mitochondria (Mita et al., 1990). However, the 13-bp repeats flanking the common deletion are not sufficiently long for an HR-mediated homology search (Hua et al., 1997). Instead, it was recently suggested that the common deletion results from

mitochondrial genome is in proximity to the electron transport chain and, therefore, permanently exposed to reactive oxygen species (ROSs) that can damage the DNA and trigger formation of deletions (Shokolenko et al., 2009). There is extensive evidence in the literature to support this idea. For example, disruption of ANT1 (adenine nucleotide translocator isoform 1) inhibits the exchange of ADP and ATP across the inner mitochondrial membrane, leading to increased production of ROS and subsequent accumulation of mtDNA deletions (Kaukonen et al., 2000; Subramaniam et al., 2008). Similarly, mitochondrial oxidative stress and increased ROS levels in response to partial loss of SOD2 (manganese superoxide dismutase) correlate with

the activity of MMEJ, which typically requires less than 12 bp of homology (Sfeir and Symington, 2015). Our *in vivo* data indicate that the formation of the common deletion is not mediated by HR and MMEJ. In contrast, we report that the mitochondrial common deletion results from a unique mode of repair that is coupled to DNA replication and mediated by the replisome. Interestingly, the described mode of replication-dependent DNA repair is highly reminiscent of DNA repair in T7 bacteriophages (Kong and Masker, 1994; Yu and Masker, 2001). Components of the mitochondrial replisome, notably Twinkle and mtSSB, have a striking similarity to their counterparts in T7 phages (Spelbrink et al., 2001). It has been proposed that a bacteriophage ancestor might have contributed to the early “mitochondrial” endosymbiosis event (Shutt and Gray, 2006). As such, it appears as if mitochondria have retained an autonomous mode of replication-dependent repair independent of nuclear DNA repair.

The Evolution of mtDNA Repeats throughout Eukaryotes

The human mitochondrial genome contains a total of 31 repeats (direct and inverted) that are more than 12 bp in length. The presence of these sequences predisposes our cells to mitochondrial genome instability and raises the question of why such repeats have been retained throughout evolution. It would be informative to address this question by altering the sequence of the 13-bp repeats and monitoring mitochondrial fitness. Unfortunately, it is currently impossible to manipulate the human mitochondrial genome, largely because of the limitations of current gene-targeting approaches. Instead, we decided to carry out a comparative mtDNA analysis by examining the frequency of repeats in 142 eukaryotes and viruses (Figure 6). Our phylogenetic analysis depicts a trend toward reduced repetitive DNA content in the mitochondrial genome of higher eukaryotes (Figures 6A and 6B). For example, genomes of primates, mammals, and vertebrates are less frequently littered with repeats than those of invertebrates and lower eukaryotes. In addition, the demarcation of vertebrates from invertebrates was marked by a significant drop in mtDNA repeats, with frequencies that are equivalent to what would have been predicted by chance alone (Figure 6B). These observations potentially reflect a selective pressure that drives the loss of repeats in mitochondrial genomes of higher organisms. The question of whether the residual repeats in human mtDNA are critical for optimal mitochondrial function awaits mitochondrial gene-editing technology to be tested experimentally.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT AND RESOURCE SHARING
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS
 - Cell culture
 - Mito-SMAR
 - Generation of mito-FISH probes
 - CRISPR/Cas9 targeting to generate Twinkle mutants

- Mito-TALEN
- Detection of the common deletion by nested PCR or QPCR
- RT-q-PCR
- Western blotting
- Fluorescence *in situ* hybridization
- Analysis of repeats content in mtDNA sequences
- T7 endonuclease assay to determine the efficiency of mito-TALEN cutting
- Next-Gen sequencing

● QUANTIFICATION AND STATISTICAL ANALYSIS

● DATA AND SOFTWARE AVAILABILITY

● ADDITIONAL RESOURCES

SUPPLEMENTAL INFORMATION

Supplemental Information includes six figures and five tables and can be found with this article online at <http://dx.doi.org/10.1016/j.molcel.2016.12.014>.

AUTHOR CONTRIBUTIONS

The project was conceived by A.S. A.S. and E.B. designed the experiments and were responsible for data interpretation. The analysis of mtDNA replication was carried out by A.F.P. Experiments related to mito-TALENs were performed by A.R.M. with the help of M.P., M.C., M.T., S.M.D., E.B., and L.B. M.T. performed the analysis of repeats in the mitochondrial genome. H.C. generated cell lines with Twinkle mutations. A.S. wrote the manuscript, and all authors provided feedback.

ACKNOWLEDGMENTS

We are grateful to Lillian Walton Masters and Khadija Lamribet for technical support. Eric Hendrickson, Gregory David, Johannes Spelbrink, and Nils-Göran Larsson are thanked for providing key reagents. We thank Eros Lazzerini-Denchi, Ruth Lehmann, Thomas Hurd, Flavia Fontanesi, and Antonio Barrientos for commenting on the manuscript and Loïc Ponger, Carine Giovannangeli, and Jean-Paul Concordet for helpful discussions. This work was supported by a joint grant (to A.S. and E.B.) from the Human Frontier Science Program (HFSP RGY0088) and by grants (to A.S.) from the David and Lucile Packard Foundation and a Damon Runyon-Rachleff innovator award. A.F.P. was supported by NIH Training Grant 2T32GM007238-39. M.P. was supported by a fellowship from Le Cancéropôle Idf and La Ligue Contre le Cancer and A.R.M. by a fellowship from the Fondation pour la Recherche Médicale (FDT20160435361). We acknowledge the support of the Genome Technology Center (GTC) and the Perlmutter Cancer Center at NYU Langone Medical Center. The Bio-Imaging Resource Center at the Rockefeller University is thanked for assistance with OMX imaging.

Received: June 19, 2016

Revised: October 4, 2016

Accepted: December 16, 2016

Published: January 19, 2017

REFERENCES

- Ali, S.A., Zaidi, S.K., Dacwag, C.S., Salma, N., Young, D.W., Shakoori, A.R., Montecino, M.A., Lian, J.B., van Wijnen, A.J., Imbalzano, A.N., et al. (2008). Phenotypic transcription factors epigenetically mediate cell growth control. *Proc. Natl. Acad. Sci. USA* 105, 6632–6637.
- Anderson, S., Bankier, A.T., Barrell, B.G., de Brujin, M.H., Coulson, A.R., Drouin, J., Eperon, I.C., Nierlich, D.P., Roe, B.A., Sanger, F., et al. (1981). Sequence and organization of the human mitochondrial genome. *Nature* 290, 457–465.

- Bacman, S.R., Williams, S.L., and Moraes, C.T. (2009). Intra- and inter-molecular recombination of mitochondrial DNA after *in vivo* induction of multiple double-strand breaks. *Nucleic Acids Res.* 37, 4218–4226.
- Bacman, S.R., Williams, S.L., Pinto, M., Peralta, S., and Moraes, C.T. (2013). Specific elimination of mutant mitochondrial genomes in patient-derived cells by mitoTALENs. *Nat. Med.* 19, 1111–1113.
- Bharti, S.K., Sommers, J.A., Zhou, J., Kaplan, D.L., Spelbrink, J.N., Mergny, J.L., and Brosh, R.M., Jr. (2014). DNA sequences proximal to human mitochondrial DNA deletion breakpoints prevalent in human disease form G-quadruplexes, a class of DNA structures inefficiently unwound by the mitochondrial replicative Twinkle helicase. *J. Biol. Chem.* 289, 29975–29993.
- Clayton, D.A. (1982). Replication of animal mitochondrial DNA. *Cell* 28, 693–705.
- Corral-Debrinski, M., Horton, T., Lott, M.T., Shoffner, J.M., Beal, M.F., and Wallace, D.C. (1992). Mitochondrial DNA deletions in human brain: regional variability and increase with advanced age. *Nat. Genet.* 2, 324–329.
- Cymerman, I.A., Chung, I., Beckmann, B.M., Bujnicki, J.M., and Meiss, G. (2008). EXOG, a novel paralog of Endonuclease G in higher eukaryotes. *Nucleic Acids Res.* 36, 1369–1379.
- Dmitrieva, N.I., Malide, D., and Burg, M.B. (2011). Mre11 is expressed in mammalian mitochondria where it binds to mitochondrial DNA. *Am. J. Physiol. Regul. Integr. Comp. Physiol.* 301, R632–R640.
- Doyon, Y., Vo, T.D., Mendel, M.C., Greenberg, S.G., Wang, J., Xia, D.F., Miller, J.C., Urnov, F.D., Gregory, P.D., and Holmes, M.C. (2011). Enhancing zinc-finger-nuclease activity with improved obligate heterodimeric architectures. *Nat. Methods* 8, 74–79.
- Duan, X.L., Liu, N.N., Yang, Y.T., Li, H.H., Li, M., Dou, S.X., and Xi, X.G. (2015). G-quadruplexes significantly stimulate Pif1 helicase-catalyzed duplex DNA unwinding. *J. Biol. Chem.* 290, 7722–7735.
- Falkenberg, M., Larsson, N.G., and Gustafsson, C.M. (2007). DNA replication and transcription in mammalian mitochondria. *Annu. Rev. Biochem.* 76, 679–699.
- Fusté, J.M., Wanrooij, S., Jemt, E., Granycome, C.E., Cluett, T.J., Shi, Y., Atanassova, N., Holt, I.J., Gustafsson, C.M., and Falkenberg, M. (2010). Mitochondrial RNA polymerase is needed for activation of the origin of light-strand DNA replication. *Mol. Cell* 37, 67–78.
- Gao, Y., Katyal, S., Lee, Y., Zhao, J., Rehg, J.E., Russell, H.R., and McKinnon, P.J. (2011). DNA ligase III is critical for mtDNA integrity but not Xrcc1-mediated nuclear DNA repair. *Nature* 471, 240–244.
- Goffart, S., Cooper, H.M., Tyynismaa, H., Wanrooij, S., Suomalainen, A., and Spelbrink, J.N. (2009). Twinkle mutations associated with autosomal dominant progressive external ophthalmoplegia lead to impaired helicase function and *in vivo* mtDNA replication stalling. *Hum. Mol. Genet.* 18, 328–340.
- Guschin, D.Y., Waite, A.J., Katibah, G.E., Miller, J.C., Holmes, M.C., and Rebar, E.J. (2010). A rapid and general assay for monitoring endogenous gene modification. *Methods Mol. Biol.* 649, 247–256.
- Holt, I.J., Lorimer, H.E., and Jacobs, H.T. (2000). Coupled leading- and lagging-strand synthesis of mammalian mitochondrial DNA. *Cell* 100, 515–524.
- Hou, J.H., and Wei, Y.H. (1998). AT-rich sequences flanking the 5'-end breakpoint of the 4977-bp deletion of human mitochondrial DNA are located between two bent-inducing DNA sequences that assume distorted structure in organello. *Mutat. Res.* 403, 75–84.
- Hua, S.B., Qiu, M., Chan, E., Zhu, L., and Luo, Y. (1997). Minimum length of sequence homology required for *in vivo* cloning by homologous recombination in yeast. *Plasmid* 38, 91–96.
- Kaukonen, J., Juselius, J.K., Tiranti, V., Kyttälä, A., Zeviani, M., Comi, G.P., Keränen, S., Peltonen, L., and Suomalainen, A. (2000). Role of adenine nucleotide translocator 1 in mtDNA maintenance. *Science* 289, 782–785.
- Kim, E., Kim, S., Kim, D.H., Choi, B.S., Choi, I.Y., and Kim, J.S. (2012). Precision genome engineering with programmable DNA-nicking enzymes. *Genome Res.* 22, 1327–1333.
- Kokoszka, J.E., Coskun, P., Esposito, L.A., and Wallace, D.C. (2001). Increased mitochondrial oxidative stress in the Sod2 (+/-) mouse results in the age-related decline of mitochondrial function culminating in increased apoptosis. *Proc. Natl. Acad. Sci. USA* 98, 2278–2283.
- Kong, D., and Masker, W. (1994). Deletion between direct repeats in T7 DNA stimulated by double-strand breaks. *J. Bacteriol.* 176, 5904–5911.
- Korhonen, J.A., Pham, X.H., Pellegrini, M., and Falkenberg, M. (2004). Reconstitution of a minimal mtDNA replisome in vitro. *EMBO J.* 23, 2423–2429.
- Kornblum, C., Nicholls, T.J., Haack, T.B., Schöler, S., Peeva, V., Danhauser, K., Hallmann, K., Zsurka, G., Rorbach, J., Iuso, A., et al. (2013). Loss-of-function mutations in MGME1 impair mtDNA replication and cause multisystemic mitochondrial disease. *Nat. Genet.* 45, 214–219.
- Kraytsberg, Y., Schwartz, M., Brown, T.A., Ebralidse, K., Kunz, W.S., Clayton, D.A., Vissing, J., and Khrapko, K. (2004). Recombination of human mitochondrial DNA. *Science* 304, 981.
- Krishnan, K.J., Reeve, A.K., Samuels, D.C., Chinnery, P.F., Blackwood, J.K., Taylor, R.W., Wanrooij, S., Spelbrink, J.N., Lightowers, R.N., and Turnbull, D.M. (2008). What causes mitochondrial DNA deletions in human cells? *Nat. Genet.* 40, 275–279.
- Kühl, I., Kukat, C., Ruzzenente, B., Milenkovic, D., Mourier, A., Miranda, M., Koolmeister, C., Falkenberg, M., and Larsson, N.G. (2014). POLRMT does not transcribe nuclear genes. *Nature* 514, E7–E11.
- Letunic, I., and Bork, P. (2016). Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res.* 44 (W1), W242–W245.
- Martin-Negrier, M.L., Sole, G., Jardel, C., Vital, C., Ferrer, X., and Vital, A. (2011). TWINKLE gene mutation: report of a French family with an autosomal dominant progressive external ophthalmoplegia and literature review. *Eur. J. Neurol.* 18, 436–441.
- Mita, S., Rizzuto, R., Moraes, C.T., Shanske, S., Arnaudo, E., Fabrizi, G.M., Koga, Y., DiMauro, S., and Schon, E.A. (1990). Recombination via flanking direct repeats is a major cause of large-scale deletions of human mitochondrial DNA. *Nucleic Acids Res.* 18, 561–567.
- Moullan, N., Mouchiroud, L., Wang, X., Ryu, D., Williams, E.G., Mottis, A., Jovaisaita, V., Frochaux, M.V., Quiros, P.M., Deplancke, B., et al. (2015). Tetracyclines disturb mitochondrial function across eukaryotic models: a call for caution in biomedical research. *Cell Rep.* 10, 1681–1691.
- Murgha, Y.E., Rouillard, J.M., and Gulari, E. (2014). Methods for the preparation of large quantities of complex single-stranded oligonucleotide libraries. *PLoS ONE* 9, e94752.
- Nikali, K., Suomalainen, A., Saharinen, J., Kuokkanen, M., Spelbrink, J.N., Lönnqvist, T., and Peltonen, L. (2005). Infantile onset spinocerebellar ataxia is caused by recessive mutations in mitochondrial proteins Twinkle and Twinky. *Hum. Mol. Genet.* 14, 2981–2990.
- Norio, P., and Schildkraut, C.L. (2001). Visualization of DNA replication on individual Epstein-Barr virus episomes. *Science* 294, 2361–2364.
- Oh, S., Harvey, A., Zimbric, J., Wang, Y., Nguyen, T., Jackson, P.J., and Hendrickson, E.A. (2014). DNA ligase III and DNA ligase IV carry out genetically distinct forms of end joining in human somatic cells. *DNA Repair (Amst.)* 21, 97–110.
- Okonechnikov, K., Golosova, O., and Fursov, M.; UGENE team (2012). Unipro UGENE: a unified bioinformatics toolkit. *Bioinformatics* 28, 1166–1167.
- Paeschke, K., Capra, J.A., and Zakian, V.A. (2011). DNA replication through G-quadruplex motifs is promoted by the *Saccharomyces cerevisiae* Pif1 DNA helicase. *Cell* 145, 678–691.
- Phadnis, N., Sia, R.A., and Sia, E.A. (2005). Analysis of repeat-mediated deletions in the mitochondrial genome of *Saccharomyces cerevisiae*. *Genetics* 171, 1549–1559.
- Pikó, L., Hougham, A.J., and Bulpitt, K.J. (1988). Studies of sequence heterogeneity of mitochondrial DNA from rat and mouse tissues: evidence for an increased frequency of deletions/additions with aging. *Mech. Ageing Dev.* 43, 279–293.
- Quispe-Tintaya, W., White, R.R., Popov, V.N., Vijg, J., and Maslov, A.Y. (2013). Fast mitochondrial DNA isolation from mammalian cells for next-generation sequencing. *Biotechniques* 55, 133–136.

- Renaud, J.B., Boix, C., Charpentier, M., De Cian, A., Cochenne, J., Duvernois-Berthet, E., Perrouault, L., Tesson, L., Edouard, J., Thinard, R., et al. (2016). Improved genome editing efficiency and flexibility using modified oligonucleotides with TALEN and CRISPR-Cas9 nucleases. *Cell Rep.* **14**, 2263–2272.
- Robberson, D.L., and Clayton, D.A. (1972). Replication of mitochondrial DNA in mouse L cells and their thymidine kinase-derivatives: displacement replication on a covalently-closed circular template. *Proc. Natl. Acad. Sci. USA* **69**, 3810–3814.
- Sage, J.M., Gildemeister, O.S., and Knight, K.L. (2010). Discovery of a novel function for human Rad51: maintenance of the mitochondrial genome. *J. Biol. Chem.* **285**, 18984–18990.
- Samuels, D.C. (2004). Mitochondrial DNA repeats constrain the life span of mammals. *Trends Genet.* **20**, 226–229.
- Samuels, D.C., Schon, E.A., and Chinnery, P.F. (2004). Two direct repeats cause most human mtDNA deletions. *Trends Genet.* **20**, 393–398.
- Schon, E.A., Rizzuto, R., Moraes, C.T., Nakase, H., Zeviani, M., and DiMauro, S. (1989). A direct repeat is a hotspot for large-scale deletion of human mitochondrial DNA. *Science* **244**, 346–349.
- Sen, D., Nandakumar, D., Tang, G.Q., and Patel, S.S. (2012). Human mitochondrial DNA helicase TWINKLE is both an unwinding and annealing helicase. *J. Biol. Chem.* **287**, 14545–14556.
- Sfeir, A., and Symington, L.S. (2015). Microhomology-mediated end joining: a back-up survival mechanism or dedicated pathway? *Trends Biochem. Sci.* **40**, 701–714.
- Shoffner, J.M., Lott, M.T., Voljavec, A.S., Soueidan, S.A., Costigan, D.A., and Wallace, D.C. (1989). Spontaneous Kearns-Sayre/chronic external ophthalmoplegia plus syndrome associated with a mitochondrial DNA deletion: a slip-replication model and metabolic therapy. *Proc. Natl. Acad. Sci. USA* **86**, 7952–7956.
- Shokolenko, I., Venediktova, N., Bochkareva, A., Wilson, G.L., and Alexeyev, M.F. (2009). Oxidative stress induces degradation of mitochondrial DNA. *Nucleic Acids Res.* **37**, 2539–2548.
- Shutt, T.E., and Gray, M.W. (2006). Bacteriophage origins of mitochondrial replication and transcription proteins. *Trends Genet.* **22**, 90–95.
- Simsek, D., Furda, A., Gao, Y., Artus, J., Brunet, E., Hadjantonakis, A.K., Van Houten, B., Shuman, S., McKinnon, P.J., and Jasins, M. (2011). Crucial role for DNA ligase III in mitochondria but not in Xrcc1-dependent repair. *Nature* **471**, 245–248.
- Spelbrink, J.N., Li, F.Y., Tiranti, V., Nikali, K., Yuan, Q.P., Tariq, M., Wanrooij, S., Garrido, N., Comi, G., Morandi, L., et al. (2001). Human mitochondrial DNA deletions associated with mutations in the gene encoding Twinkle, a phage T7 gene 4-like protein localized in mitochondria. *Nat. Genet.* **28**, 223–231.
- Srivastava, S., and Moraes, C.T. (2005). Double-strand breaks of mouse muscle mtDNA promote large deletions similar to multiple mtDNA deletions in humans. *Hum. Mol. Genet.* **14**, 893–902.
- Stein, A., Kalifa, L., and Sia, E.A. (2015). Members of the RAD52 epistasis group contribute to mitochondrial homologous recombination and double-strand break repair in *Saccharomyces cerevisiae*. *PLoS Genet.* **11**, e1005664.
- Subramaniam, V., Golik, P., Murdock, D.G., Levy, S., Kerstann, K.W., Coskun, P.E., Melkonian, G.A., and Wallace, D.C. (2008). MITOCHIP assessment of differential gene expression in the skeletal muscle of *Ant1* knockout mice: coordinate regulation of OXPHOS, antioxidant, and apoptotic genes. *Biochim. Biophys. Acta* **1777**, 666–675.
- Tadi, S.K., Sebastian, R., Dahal, S., Babu, R.K., Choudhary, B., and Raghavan, S.C. (2016). Microhomology-mediated end joining is the principal mediator of double-strand break repair during mitochondrial DNA lesions. *Mol. Biol. Cell* **27**, 223–235.
- Tomkinson, A.E., Bonk, R.T., and Linn, S. (1988). Mitochondrial endonuclease activities specific for apurinic/apyrimidinic sites in DNA from mouse cells. *J. Biol. Chem.* **263**, 12532–12537.
- Tynnismäa, H., Mjosund, K.P., Wanrooij, S., Lappalainen, I., Ylikallio, E., Jalanko, A., Spelbrink, J.N., Paetau, A., and Suomalainen, A. (2005). Mutant mitochondrial helicase Twinkle causes multiple mtDNA deletions and a late-onset mitochondrial disease in mice. *Proc. Natl. Acad. Sci. USA* **102**, 17687–17692.
- Van Goethem, G., Dermaut, B., Löfgren, A., Martin, J.J., and Van Broeckhoven, C. (2001). Mutation of POLG is associated with progressive external ophthalmoplegia characterized by mtDNA deletions. *Nat. Genet.* **28**, 211–212.
- Wallace, D.C., and Chalkia, D. (2013). Mitochondrial DNA genetics and the heteroplasmy conundrum in evolution and disease. *Cold Spring Harb. Perspect. Biol.* **5**, a021220.
- Wanrooij, S., Goffart, S., Pohjoismäki, J.L., Yasukawa, T., and Spelbrink, J.N. (2007). Expression of catalytic mutants of the mtDNA helicase Twinkle and polymerase POLG causes distinct replication stalling phenotypes. *Nucleic Acids Res.* **35**, 3238–3251.
- Wanrooij, S., Fusté, J.M., Farge, G., Shi, Y., Gustafsson, C.M., and Falkenberg, M. (2008). Human mitochondrial RNA polymerase primes lagging-strand DNA synthesis in vitro. *Proc. Natl. Acad. Sci. USA* **105**, 11122–11127.
- Wu, Y., Gao, T., Wang, X., Hu, Y., Hu, X., Hu, Z., Pang, J., Li, Z., Xue, J., Feng, M., et al. (2014). TALE nickase mediates high efficient targeted transgene integration at the human multi-copy ribosomal DNA locus. *Biochem. Biophys. Res. Commun.* **446**, 261–266.
- Yang, J.H., Lee, H.C., Lin, K.J., and Wei, Y.H. (1994). A specific 4977-bp deletion of mitochondrial DNA in human ageing skin. *Arch. Dermatol. Res.* **286**, 386–390.
- Yang, M.Y., Bowmaker, M., Reyes, A., Vergani, L., Angeli, P., Gringeri, E., Jacobs, H.T., and Holt, I.J. (2002). Biased incorporation of ribonucleotides on the mitochondrial L-strand accounts for apparent strand-asymmetric DNA replication. *Cell* **111**, 495–505.
- Yasukawa, T., Reyes, A., Cluett, T.J., Yang, M.Y., Bowmaker, M., Jacobs, H.T., and Holt, I.J. (2006). Replication of vertebrate mitochondrial DNA entails transient ribonucleotide incorporation throughout the lagging strand. *EMBO J.* **25**, 5358–5371.
- Yu, M., and Masker, W. (2001). T7 single strand DNA binding protein but not T7 helicase is required for DNA double strand break repair. *J. Bacteriol.* **183**, 1862–1869.
- Zhang, H., Barceló, J.M., Lee, B., Kohlhagen, G., Zimonjic, D.B., Popescu, N.C., and Pommier, Y. (2001). Human mitochondrial topoisomerase I. *Proc. Natl. Acad. Sci. USA* **98**, 10608–10613.

STAR★METHODS**KEY RESOURCES TABLE**

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
anti-MGME1 polyclonal antibody	Sigma	Cat# HPA040913; RRID: AB_10962835
anti-CTIP monoclonal antibody (clone 14-1)	Active motif	Cat#61141
anti-human LIG3 monoclonal antibody (clone 7)	Becton Dickinson	Cat# 611877
anti-LIG1 monoclonal antibody (clone 10H5)	Abcam	Cat# ab615; RRID: AB_2136117
anti-HA monoclonal antibody (clone 12CA5)	Roche	Cat# 11-867-423-001
anti-PIF1 polyclonal antibody	Sigma	Cat# SAB4301117
anti-VDAC1 polyclonal antibody	Abcam	Cat# ab15895
anti-GAPDH monoclonal antibody (clone 6C5)	Abcam	Cat# ab8245; RRID: AB_2107448
anti-RAD51 monoclonal antibody (clone 14BH)	Novus Biologicals	Cat# NB100-148; RRID: AB_10002131
anti-MRE11 polyclonal antibody	Novus	Cat# NB100-142
anti-POLG polyclonal antibody	Novus	Cat# NBP1-52300; RRID: AB_11027699
anti-DNA2 polyclonal antibody	Abcam	Cat# ab194939
anti-mtSSB polyclonal antibody	Proteintech	Cat# 12212-1-AP; RRID: AB_2195320
anti-c-MYC monoclonal antibody (clone 9E10)	Millipore	Cat# 05-419; RRID: AB_309725
anti-FLAG monoclonal antibody (clone M2)	Sigma Aldrich	Cat#F1804; RRID: AB_262044
anti- γ -Tubulin monoclonal antibody (GTU-88)	Abcam	Cat#ab11316
anti- α -Tubulin monoclonal antibody (clone DM1A)	Sigma	Cat#T9026
mouse anti-BrdU monoclonal antibody	BD Biosciences	Cat#347580
rat anti-BrdU monoclonal antibody	Abd Serotec	Cat#OBT0030CX
Alexa Fluor 488-conjugated goat anti-rat	Molecular Probes	Cat#A-11006
Alexa Fluor 568-conjugated goat anti-mouse	Molecular Probes	Cat#A-11004
Biotinylated anti-streptavidin	Vector Labs	Cat#BA-0500
Chemicals, Peptides, and Recombinant Proteins		
Protease Inhibitor Cocktail	Roche	Cat#04693159001
T7 Endonuclease I	New England Biolabs	Cat# M0302S
Proteinase K	New England Biolabs	Cat# P8107S
(Z)-4-Hydroxytamoxifen	Sigma	Cat#H7904-5MG
5-Chloro-2'-deoxyuridine	Sigma	Cat#C6891
5-Iodo-2'-deoxyuridine	Sigma	Cat#I7125
Alexa Fluor 405-conjugated streptavidin	Molecular Probes	Cat#S32351
(3-Aminopropyl)triethoxysilane	Sigma	Cat#440140
UltraPure low melting point agarose gel	Thermo Fisher	Cat#16520050
Critical Commercial Assays		
QIAprep Spin Miniprep Kit	QIAGEN	Cat# 27106
RNeasy Mini Kit	QIAGEN	Cat# 74106
iScript gDNA Clear cDNA Synthesis Kit	Bio-Rad	Cat# 1725034
High Capacity cDNA Reverse Transcription Kit	Thermo Fisher	Cat# 4368814
SYBR Green PCR Master Mix	Thermo Fisher	Cat# 4309155
Cell Fractionation Kit	Abcam	Cat# Ab109719
Brilliant III Ultra-Fast SYBR QPCR MasterMix	Agilent	Cat#600882
Lab-Tek Chamber Slide system	Sigma	Cat#C7182
Mito-Tracker Red CMXRos	Invitrogen	Cat#M7512
Hoechst 33342	Thermo Fisher	Cat#H1399

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
FastStart Taq DNA Polymerase, dNTPack	Roche	Cat#04738420001
Lipofectamine RNAiMAX	Thermo Fisher	Cat#13778150
Gel extraction kit	EZNA	Cat#D2500-02
Experimental Models: Cell Lines		
U2OS	ATCC	HTB-96
HCT116	ATCC	CCL-247
HT1080	ATCC	CCL-121
BJ-hTERT	ATCC	CRL-4001
C2C12 myoblasts	ATCC	CRL-1772
RPE-1-hTERT WT	ATCC	CRL-4000
MEF	This paper	N/A
Twinkle ^{R374Q/WT} and RPE-1-hTERT	This paper	N/A
Twinkle ^{Y508C/Y508C} RPE-1-hTERT		
POLRMT ^{+/+} and POLRMT ^{F/F} mouse embryonic fibroblasts	Laboratory of N.G. Larsson	Kühl et al. (2014)
HCT116 Lig1 F/-	E. Hendrickson	Oh et al. (2014)
HCT116 Lig3 F/-	E. Hendrickson	Oh et al. (2014)
Recombinant DNA		
pVAX-mitoTALENs, see Table S3	This paper	N/A
Plasmid: PLJM1-Flag-POLRMT-puro	This paper	N/A
Plasmid: pBabe-Cre-ER ^{T2}	Laboratory of G. David	N/A
Plasmid: pSpCas9(Bbs)(Bsa)-2A-GFP	This paper	N/A
Sequence-Based Reagents		
siRNA sequences, see Table S1	This paper	N/A
Custom mito-FISH probe library, see Table S2	This paper	MYcroarray custom synthesis
Mito-TALEN target sequences, see Table S3	This paper	N/A
Primers for common deletion detection and quantification, see Table S4	This paper	N/A
Primers for muscle cell differentiation q-PCR, see Table S5		Ali et al. (2008)
Primers for q-PCR, see Table S5	PrimerBank	https://pga.mgh.harvard.edu/primerbank/
Primers for POLRMT genotyping: 3F: 5'-CCACTAACCATGGCTGTGC-3', 3R: 5'-GGCACATACTTGATACAGGCTTGG-3', and Int3_Rv3: 5'-CTTGCAGGCCAGGCATATG-3'	This paper	N/A
Primers for semiquantitative PCR of POLRMT:	This paper	N/A
2F: 5'-GCAACATGTCCTGAGGGAGTGG-3'		
4R: 5'- GCCAGCATCCTTCAGCATGAGG-3'		
Twinkle R374Q CRISPR guides: exon1 gRNA #1: 5'-CCATCGTATCTTCCGGCAGCT-3', Twinkle exon1 gRNA #2: 5'-GCTAGGAGAACTGTCAAATGTGG-3'	This paper	N/A
Twinkle R374Q genotyping primers: Fwd: 5'- GCACG AAAACTGAACCCCAAACG-3' and Rev: 5'- CCTTCG ATGTCCTTCAAGATACG-3'	This paper	N/A
Twinkle Y508C CRISPR guides: Twinkle exon3 gRNA #1: 5'-GACAACCTGCAGTT CATGAT -3', Twinkle exon3 gRNA #2: 5'-ATTGTATCTATTACAGTCCT-3'	This paper	N/A
Twinkle Y508C genotyping primers: Fwd: 5'- ACTCACAC TCAGATCCCTGTCC-3' and Rev: 5'-ATCGCTTGAGCCT AGTAGGTTGG-3'	This paper	N/A

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Software and Algorithms		
NCBI Taxonomy database		https://www.ncbi.nlm.nih.gov/taxonomy
NCBI Organelle database		https://www.ncbi.nlm.nih.gov/genome/organelle/
UGENE bioinformatics suite		http://www.ugene.net
Random DNA sequence generator		http://www.faculty.ucr.edu/~mmaduro/random.htm
Custom python pipeline	Renaud et al., 2016	N/A
Other		
Next-Gen Sequencing data	This paper	NCBI Bioproject : PRJNA357534
Unprocessed image data	This paper	http://dx.doi.org/10.17632/hctwmmpj9r.1

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for reagents may be directed to, and will be fulfilled by, the lead contact, Dr. Agnel Sfeir, agnel.sfeir@med.nyu.edu.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Source of cell lines used in the study is reported in the reagent and resource table.

METHOD DETAILS**Cell culture**

HT1080, BJ-hTERT, C2C12 myoblasts, RPE-1, and MEF cell lines were grown in DMEM supplemented with 10% FBS (fetal bovine serum), 2 mM L-glutamine (Sigma), 100 U ml⁻¹ penicillin (Sigma), 0.1 µg ml⁻¹ streptomycin (Sigma), 0.1 mM non-essential amino acids (Invitrogen) and 200 µM uridine (Sigma) as needed. C2C12 myoblasts were differentiated into myotubes by replacing the media with DMEM supplemented with 2% horse serum, 2 mM L-glutamine (Sigma), 100 U ml⁻¹ penicillin (Sigma), 0.1 µg ml⁻¹ streptomycin (Sigma), and 0.1 mM non-essential amino acids (Invitrogen) at 90% confluency and cultured for 5-7 days. *Polmt*^{F/F} and *Polmt*^{+/-} MEFs were a gift from N.G. Larsson (Kühl et al., 2014). U2OS and HCT116 cells were cultured in McCoy's 5A medium containing 10% FCS (Fetal Calf serum). *L/G3*^{F/-} and *L/G1*^{F/-} HCT116 cells were a gift from E. Hendrickson (Oh et al., 2014). Deletion of *L/G3* and *L/G1* was induced by treating cells with 10nM 4-hydroxy-tamoxifen for 24 hr. To delete POLRMT, Cre-ER^{T2} was stably expressed in *Polmt*^{F/F} and *Polmt*^{+/-} MEFs, and cells were treated with 10 µM 4-Hydroxy-tamoxifen (Sigma) for 24 hr. Cells were pulsed with CldU and IdU for mito-SMARD, 48 hr post Cre induction. The primers used for *Polmt* genotyping are: 3F: 5'-CCACTAACCAT GGCTGTCTGC-3', 3R: 5'-GGCACATACTTGATACAGCTGG-3', and Int3_Rv3: 5'-CTTGCAGGCCAGGCATATG-3'.

U2OS and HCT116 cells were transfected with 3.5 µg of mito-TALEN using Amaxa Nucleofector technology (Lonza). Typically, 1-2*10⁶ cells were transfected with the Kit V and program D-032 (Amaxa 2b device) or SE. Solution and CM-104 program (Amaxa 4D). For siRNA transfection, cells were plated at a density of 3*10⁵ cells per well in a 6- well and transfected 24 hr later using LipofectamineTM RNAiMax (Invitrogen) following the manufacturer's protocol. The siRNA sequences are listed in Table S1. Cells were then incubated for 24 hr before transfection with Mito-TALEN. mtDNA depletion was achieved by treating cells with 50 ng/mL EtBr in presence of uridine for a total of 5 days, followed by 2 days of chasing in regular media supplemented with 200 µM uridine.

Mito-SMARD

Cells were initially pulsed with CldU for 2 hr, followed by IdU for 2 hr, yielding mostly all green or all red molecules. Pulse-labeling was experimentally optimized in order to capture a greater number of dual-labeled molecules. Human cells by sequentially labeling with 15 min of CldU (30 µM) and 90 min of IdU (30 µM), followed by 2 hr chase in standard media. Mouse cells were pulsed with 30 min CldU, followed by 90 min with IdU. After repeating the procedure twice, cells were collected by trypsinization and mtDNA was purified using the QIAprep Spin Miniprep Kit (QIAGEN) (Quispe-Tintaya et al., 2013). mtDNA was eluted from the column with 55 µL of H2O, mixed with an equal volume of 2% agarose and loaded into a plug mold (Bio-Rad). Plugs were incubated with 10 U of Ael (New England BioLabs) overnight at 37°C. The plugs were run on a 0.7% UltraPure low melting point agarose gel (Invitrogen) at 30 V for 2 hr. The linearized mtDNA band was excised from the gel and melted in 100 µL melting buffer (1xTE, 100mM NaCl, and 0.1% beta-mercaptoethanol) at 68°C for 20 min before being moved to 45°C and incubated for 2 hr with the addition of 3 U Gelase. After melting, an additional 10 µL of beta-mercaptoethanol and 0.2 µL of YoYo-1 (Invitrogen) were added to the tube and incubated overnight. The mtDNA was combed by capillary action across microscope cover glasses coated with 3-aminopropyltriethoxysilane (Sigma). The stretched DNA was denatured in an alkali-denaturing buffer (0.1 N NaOH in 70% ethanol and 0.1% beta-mercaptoethanol) for

9 min and fixed by adding 0.5% glutaraldehyde for 5 min. mtDNA FISH probes were hybridized overnight in buffer containing 2x SSC, 40% formamide, and 10% dextran sulfate. The probe was labeled with Alexa Fluor 405-conjugated streptavidin (Molecular Probes). The signal was amplified by two rounds of labeling with biotinylated anti-streptavidin antibody (Vector) and additional Alexa Fluor 405-conjugated streptavidin. The thymidine analogs were detected with a mouse anti-BrdU monoclonal antibody recognizing IdU (Becton Dickinson) and a rat anti-BrdU monoclonal antibody recognizing CldU (AbD Serotec). Secondary antibodies used in these studies were Alexa Fluor 488-conjugated goat anti-rat (Molecular Probes) and Alexa Fluor 568-conjugated goat anti-mouse (Molecular Probes). Slides were imaged using an Eclipse Ti2 with a 100x Apo TIRF oil objective (Nikon) or on a DeltaVision OMX (GE Healthcare Life Sciences).

Generation of mito-FISH probes

A library of mtDNA probes for FISH analysis was generated using MYtag technology. A library of probes labeling nucleotides 10000-16400 of the mtDNA genome were custom synthesized by MYcroarray. The library was prepared according to manufacturer's recommendations ([Murgha et al., 2014](#)). Briefly, the oligonucleotide library was transcribed in vitro to generate RNA. The RNA fragments were then reverse transcribed using a common biotinylated primer. Degradation of the RNA molecules yielded single stranded oligonucleotides. The sequences for the probes are found in [Table S2](#).

CRISPR/Cas9 targeting to generate Twinkle mutants

Twinkle heterozygote R374Q mutation was introduced into human, retina pigmented epithelium cells (RPE-1) using a CRISPR/Cas9 double nickase strategy. Briefly the plasmid pSpCas9(BB)-2A-GFP (Addgene #48138), harboring a BbsI cloning site for gRNA expression and the Cas9-2A-GFP nickase was reengineered with an additional BsaI gRNA cloning cassette giving rise to pSpCas9(Bbs)(Bsa)-2A-GFP. Two gRNAs targeting Twinkle Exon 1 (Twinkle exon1 gRNA #1: 5'-CCATCGTATCTTCGGCAGCT-3', Twinkle exon1 gRNA #2: 5'-GCTAGGAGAACTGTCAAATGTGG-3') were cloned sequentially as dsDNA oligos into BbsI and BsaI sites and sequence confirmed. The construct was co-transfected with a repair template consisting of an 180bp ssDNA oligo (IDT) introducing the desired point mutation, a silent point mutation introducing a PstI restriction site for genotyping purposes and silent mutations in the PAM sites. 48 hr post-transfection, GFP-positive cells were sorted as single cells into 96-well plates. Clonal cell lines with heterozygote targeting were identified by genotypic PCR, followed by restriction digest using PstI (NEB). Positive clones were confirmed sequencing both DNA and cDNA. The genotyping PCR primers consist of Fwd: 5'- GCACGAAA ACTGAACCCCAAACG-3' and Rev: 5'- CCTTTCGATG TCCCTTCAAGATACG-3'. A similar strategy was used to generate the homozygote Y508C mutation in exon 3. Two gRNAs (Twinkle exon3 gRNA #1: 5'-GACAACCTGCAGTT CATGAT -3', Twinkle exon3 gRNA #2: 5'-ATTGTATCTATTACAGTCCT-3') were cloned sequentially into pSpCas9(Bbs)(Bsa)-2A-GFP and co-transfected with a repair template consisting of an 180bp ssDNA oligo (IDT) introducing the desired point mutation, a silent point mutation introducing a ClaI (NEB) restriction site for genotyping purposes and silent mutations in the PAM sites. Single clones were screened by genotypic PCR followed by ClaI digestion. The genotyping PCR primers consisted of Fwd: 5'- ACTCACACTCAGATCCCTGTCC'-3' and Rev: 5'-ATCGCTTGAGCCTAGTAGGTTGG-3'.

Mito-TALEN

For each TALEN subunit, the fragment containing the 16 RVD segment was obtained from single unit vectors: A (NI), T (NG), G (NN) and C (HD), kindly provided by B. Zhang. Mito-TALENs were assembled into pVax vector and included the 5'UTR and mitochondrial localization signal (MLS) derived from SOD2. 3'UTR of SOD2 was also introduced at the C-term end ([Bacman et al., 2013](#)). The target sequences for each TAL domain is described in [Table S3](#). To decrease mitoTALEN off-target effects, we used two obligate heterodimeric domains of FokI: FokI-ELD and FokI-KKR (Enhancing zinc-finger-nuclease activity with improved obligate heterodimeric architectures ([Doyon et al., 2011](#)). The T2A 'self-cleaving' peptide sequence and GFP reporter were fused to FokI-ELD heterodimeric domain. Similarly, the E2A and RFP reporter coding sequences (kindly provided by T. Auer) were fused to KKR domain. Mito-TALEN nickases were generated using a catalytically inactive FokI domain (D450A) on one of the monomers ([Kim et al., 2012](#)).

Detection of the common deletion by nested PCR or QPCR

Cells were trypsinized and were re-suspended in 200 µL of PBS buffer containing Proteinase K (0.2mg/ml), SDS (0.2%), EDTA (5mM) for 3 hr at 56°C. After isopropanol precipitation, DNA was amplified by PCR. For direct detection, 100 ng of DNA was amplified by nested PCR (PCR1 - 15 cycles; PCR2 - 30 cycles). For quantification of the common deletion formation, 50 ng of DNA is amplified by Q-PCR (Brilliant III Ultra-Fast SYBR QPCR Master Mix (Agilent)). A list of primers is provided in [Table S4](#).

RT-q-PCR

Cells were collected by trypsinization and RNA was purified using the RNeasy kit (QIAGEN) according to the manufacturer's protocol. cDNA was generated using 1 µg of RNA and the High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems). Q-PCR was run using SYBR Green PCR Master Mix (Applied Biosystems). Primers listed in [Table S5](#).

Western blotting

Whole-cell extracts were prepared in protein lysis buffer (50 mM Tris-HCl pH 7.4, 1% Triton X-100, 0.1% SDS, 150 mM NaCl, 1 mM EDTA, and 1 mM DTT), with addition of Protease Inhibitor Cocktail (Complete, Roche). Typically 30 µg of protein extracts

from cells were resolved on an 8% or 10% SDS-page gel, blotted, and probed with the following antibody: anti-MGME1 polyclonal antibody (Sigma, HPA040913), anti-CTIP monoclonal antibody (Active motif, 61141), anti-human LIG3 monoclonal antibody (Becton Dickinson, 611877), anti-LIG1 monoclonal antibody (Abcam, ab615), anti-HA monoclonal antibody (Roche, 11-867-423-001), anti-PIF1 polyclonal antibody (Sigma, SAB4301117), anti-VDAC1 polyclonal antibody (Abcam, ab15895), anti-GAPDH monoclonal antibody (Abcam, ab8245), anti-RAD51 monoclonal antibody (Novus Biologicals, NB100-148), anti-MRE11 polyclonal antibody (Novus, NB100-142), anti-POLG polyclonal antibody (Novus, NBP1-52300), anti-DNA2 polyclonal antibody (Abcam, ab194939), anti-mtSSB polyclonal antibody (Proteintech 12212-1-AP), anti-c-MYC antibody clone 9E10 (Millipore, 05-419), anti-FLAG antibody clone M2 (Sigma Aldrich, F1804), anti- γ -TUBULIN antibody clone GTU-88 (Abcam, ab11316) and anti- α -TUBULIN antibody (Sigma, T9026). Cellular Fractionation was performed using an Abcam fractionation kit (Ab109719) according to the manufacturer protocol.

Fluorescence in situ hybridization

U2OS cells were cultured in LabTek chambers (Sigma). Mitochondria were stained with 62.5 μ M Mito-Tracker Red CMXRos (Invitrogen) for 25 min at 37°C. Cells were then washed with PBS and fixed for 15 min at room temperature in 4% paraformaldehyde. After two washes in PBS, cells were permeabilized with 1% triton for 5 min. Pre-blocking was performed for 1 hr in 10% FBS (fetal bovine serum). After an additional PBS wash, the slides are incubated with primary HA antibody (Roche) for 1 hr, washed, and incubated with Alexa Fluor A488 (Jackson Immuno Research) for an additional hour. After 3 more washes, nuclei were stained with Hoechst 33342 (Thermo Fisher). Images were captured with the Zeiss Axio Observer Z1 system.

Analysis of repeats content in mtDNA sequences

142 organisms with sequenced mtDNA were chosen to cover the whole eukaryotic reign. Two viruses (*Enterobacteria phage T7* and *Vaccinia virus*) were also included in the analysis due to homology of DNA metabolism enzymes with mitochondrial replication/repair machinery. Phylogenetic relationships were inferred from NCBI Taxonomy database and plotted as circular tree with iTOL interactive tools v3.1 (Letunic and Bork, 2016). The tree was further annotated with repeat content computed as follow. mtDNA sequences were retrieved at NCBI Organelle database and repeats content and GC% were analyzed with UGENE bioinformatics suite (Okonechnikov et al., 2012). Unique direct and inverse repeats > 12bp with 100% identity were computed for each sequence, normalized to the size of the respective mtDNA (or genome in case of viruses) and multiplied by 100. Additional control sequences with defined size and GC% were randomly generated (<http://www.faculty.ucr.edu/~mmaduro/random.htm>) and analyzed in the same way.

T7 endonuclease assay to determine the efficiency of mito-TALEN cutting

U2OS cells were transfected with 2 μ g mito-TALEN target sequence (PCR amplified from U2OS cells (Table S4) and cloned using TOPO-TA cloning, Invitrogen) and 3.5 μ g of the corresponding nuclear TALEN expression plasmids. DNA was extracted after 36 hr and indels formation was monitored using a T7 endonuclease as previously described (Guschin et al., 2010) using the primers T3: GCAATTAACCCTCACTAAAGG and T7: TAATACGACTCACTATAGGG.

Next-Gen sequencing

Cells treated with mito-TALENs were trypsinized and resuspended in lysis buffer (10 mM Tris-MOPS pH 7.4, 1 mM EGTA/Tris pH 7.4, 0.2 M sucrose) for 10 min on ice. After douncing, lysates were centrifuged at 700 g for 10 min at 4°C. Supernatants were incubated with 1,000 Units of DNase I at 37°C for 30-60 min and spin at 10000 g for 30 min at 4°C. Pellets were washed with 1ML of lysis buffer, centrifuged at 10,000 g for an additional 30 min at 4°C, and resuspended in PBS buffer containing 0.2% SDS, 5mM EDTA and 0.2mg/ml Proteinase K. After incubation at 56°C for 3 hr, mtDNA was precipitated and PCR amplification was performed using primers from Table S4 to detect the common deletion. PCR fragments were purified using Gel extraction kit (EZNA). Libraries were prepared with NEBNext DNA Library Prep Master Mix (IDT) and NEBNext Multiplex Oligos Index Primers (set 1 and set 2, IDT) following manufacturer's instructions. A custom python pipeline was used to characterize the indels (Renaud et al., 2016).

QUANTIFICATION AND STATISTICAL ANALYSIS

For Next-Gen sequencing analysis, a custom python pipeline was used to characterize the indels (Renaud et al., 2016).

Quantification of the semiquantitative PCR shown in Figure S3B. Band intensity was measured using Image Lab and the intensity is normalized to the untreated sample.

Quantification of common deletion: The frequency of common deletion in all genetic perturbations was calculated by RT PCR. Values were normalized to mtDNA content (12S) and presented relative to siControl. Values are mean of three independent experiments with SDs.

Quantification of group A and group B mtDNA molecules: Average frequencies were calculated from three independent experiments ($n \geq 100$ molecules in each experiment) with SDs.

DATA AND SOFTWARE AVAILABILITY

Original imaging data have been deposited to Mendeley Data and are available at <http://dx.doi.org/10.17632/hctwmmpj9r.1>. The accession number for the Next-Gen sequencing data reported in this paper is NCBI BioProject: PRJNA357534 (<https://www.ncbi.nlm.nih.gov/bioproject/357534>).

ADDITIONAL RESOURCES

Random DNA sequence generator: <http://www.faculty.ucr.edu/~mmaduro/random.htm>

NCBI Organelle database: <https://www.ncbi.nlm.nih.gov/genome/organelle/>

NCBI Taxonomy database: <https://www.ncbi.nlm.nih.gov/taxonomy>

Python pipeline for the analysis of indels: Renaud et al., 2016

Supplemental Information

Single-Molecule Analysis of mtDNA Replication Uncovers the Basis of the Common Deletion

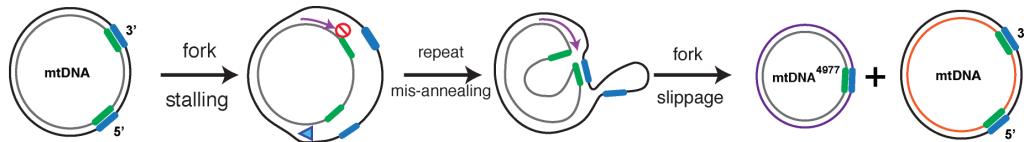
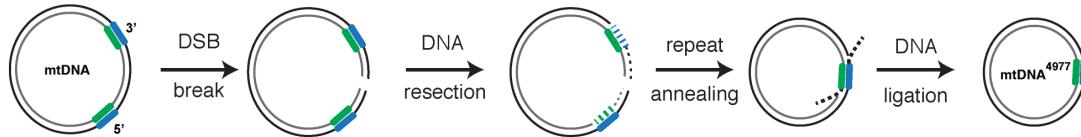
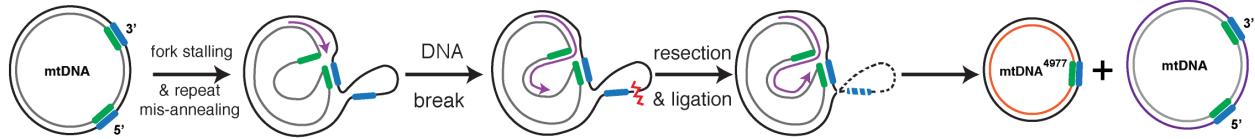
Aaron F. Phillips, Armél R. Millet, Marco Tigano, Sonia M. Dubois, Hannah Crimmins, Loelia Babin, Marine Charpentier, Marion Piganeau, Erika Brunet, and Agnel Sfeir

Single-molecule analysis of mtDNA replication uncovers the basis of the common deletion

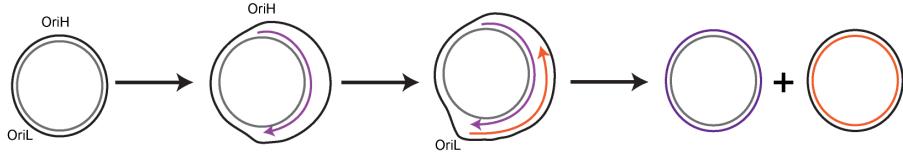
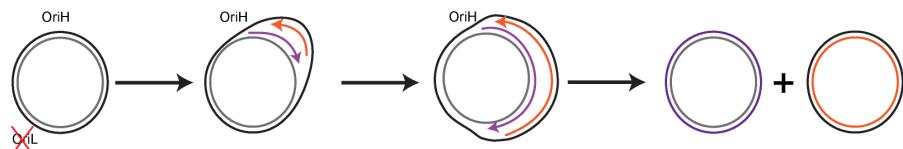
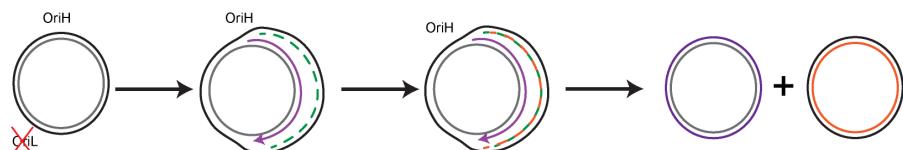
Aaron F. Phillips^{1*}, Armêl R. Millet^{2*}, Marco Tigano¹, Sonia M. Dubois², Hannah Crimmins¹, Loelia Babin², Marine Charpentier², Marion Piganeau², Erika Brunet^{2**}, and Agnel Sfeir^{1,4**}

SUPPLEMENTAL FIGURES

Figure. S1

A**Replication Slippage****DSB repair by MMEJ****Replication-mediated repair**

H strand **newly synthesized H strand**
 L strand **newly synthesized L strand**

B**Strand-displacement (asymmetric – Clayton)****Strand-coupled (synchronous – Holt)****RITOLS (based on strand-coupled – Holt)**

H strand **newly synthesized H strand**
 L strand **newly synthesized L strand**

RNA

Figure S1. Related to introduction. Schematic of the proposed models for the formation of the mtDNA common deletion. (A) Top panel, Replication slippage can lead to the common deletion. Replication fork stalling in the vicinity of the common deletion exposes long stretches of single-stranded DNA, allowing the mis-annealing of the flanking repeats. In the event of replication slippage, the common deletion would result. Middle panel, MMEJ can lead to the common deletion. A DSB (double-strand break) incurred anywhere within the common deletion region is subject to exonuclease resection, exposing the 5'- and 3'- complementary repeat sequences, which can anneal onto each other. Subsequent degradation of the unbound single-stranded DNA and ligation of the ends results in a deleted mtDNA. A third hypothetical model suggest that replication and repair converge to stimulate the common deletion. According to the replication-mediated repair model, mis-annealing of repeats in response to replication fork stalling generates a single stranded loop that is susceptible to breakage. The damaged DNA is subject to nuclease resection until the free ends ligate to each other, resulting in the loss of intervening sequence and formation of mtDNA⁴⁹⁷⁷. (B) Schematic representing models that explain the mechanism by which mtDNA duplicates. The strand-displacement model (also termed asymmetric mode) was put forth based on the visualization of replicating mtDNA intermediates by electron microscopy (EM). According to this model, mtDNA synthesis employs two distinct origins of replication and progresses uni-directionally for each strand. Heavy-strand synthesis initiates from the OriH (origin of heavy-strand replication) located in the displacement loop and advances for ~11kb before the synthesis of the light chain initiates from the OriL (origin of light strand replication) in the opposite direction. Holt and colleagues later proposed the contrasting strand-coupled model (also referred to as synchronous mode), which is based on 2-dimensional (2D) gel electrophoresis analyses of mtDNA replication. According to this model, replication of the heavy- and light- strands is coupled, and proceeds in a bi-directional manner from a broad initiation zone downstream of OriH. The strand-coupled model was later amended to include an RNA that first coats the lagging-strand and is replaced by DNA in subsequent maturation step. According to the RITOLS (RNA incorporation throughout the lagging strand) model, RNA intermediates are processed into several small RNA primers that then initiate lagging strand DNA synthesis discontinuously.

Figure S2

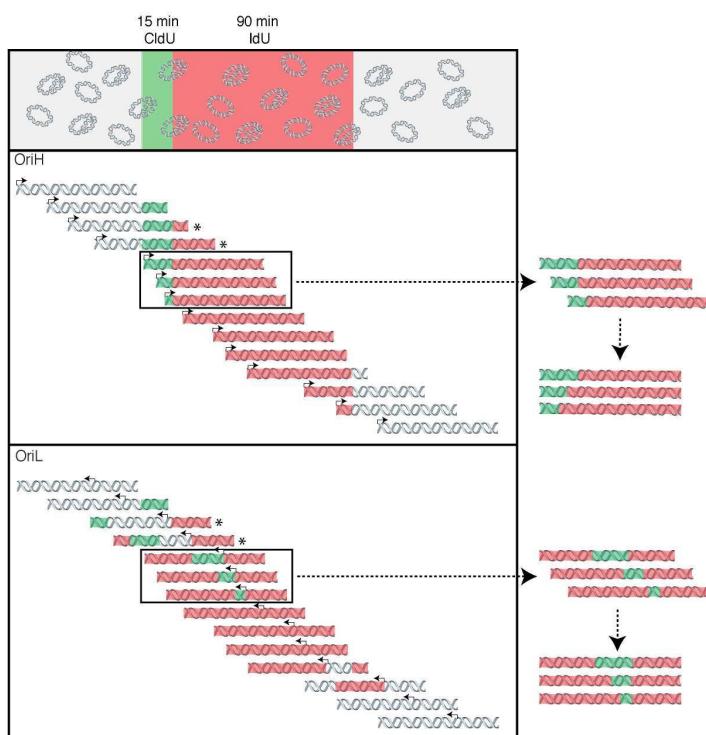
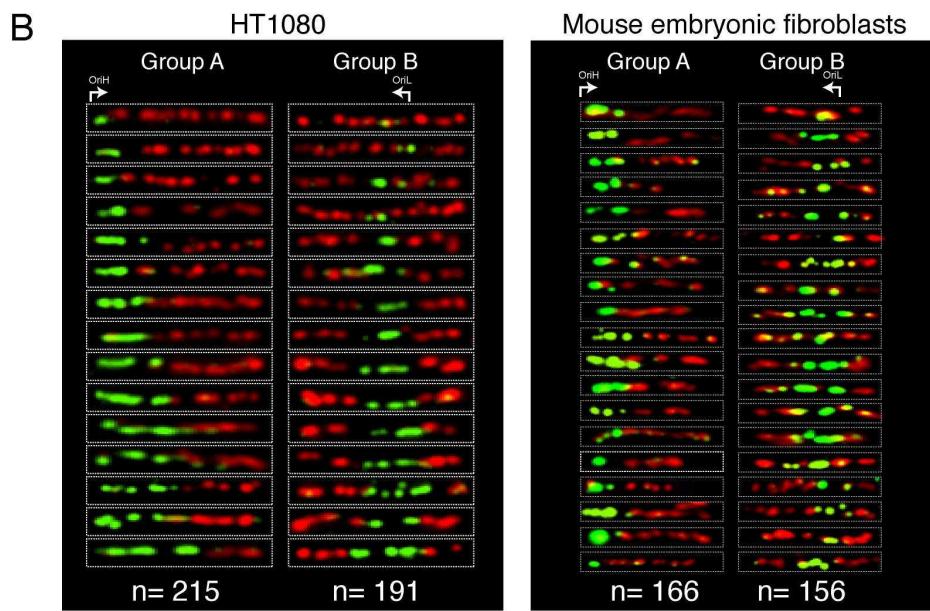
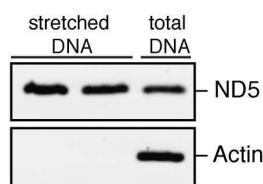
A**B****C**

Figure S2. Related to Figure 1. Mito-SMARD: a novel approach to study mtDNA replication. (A) Criteria for identifying dual-labeled molecules. Cells were sequentially pulse labeled with media containing CldU for 15 minutes and IdU for 90 minutes, followed by a chase period. Schematic depiction of the predicted pattern of labeling shown for heavy- and light-strand replication. The fragments are organized based on the fork progressing from OriH (Top) and OriL (Bottom). Replication termination during the CldU pulse and replication initiation during the IdU pulse yield single labeled molecules that were not considered in the analysis. In addition, incomplete labeling due to firing of OriH prior to the CldU pulse generate partially labeled molecules (shorter tracts and gaps; highlighted with asterisks) that were excluded from the analysis. In order to determine key features of mtDNA synthesis, including replication origin and speed, we only examined full-length and dual-labeled molecules (highlighted in a box). These fragments were organized assuming the fork progresses from OriH copying the leading strand and OriL copying the lagging strand. (B) Examples of mtDNA molecules from cells labeled with mito-SMARD. Mito-SMARD yields mtDNA molecules that have been labelled with red or green, and a third, and very important set of molecules that are dually labelled. When investigating DNA replication with combing methods, molecules labelled with one analogue (red or green) are not informative to deducing replication speed, direction, and origin, and therefore, these molecules were disregarded from our analysis. Instead, key to examining replication parameters are dual-labeled molecules, which we carefully document throughout the study. (C) To confirm the purity of the isolated DNA, we performed PCR for mtDNA (ND5) and nuclear DNA (actin). The results validate the purification method and confirm a significant enrichment for mtDNA in samples stretched on glass slides.

Figure S3

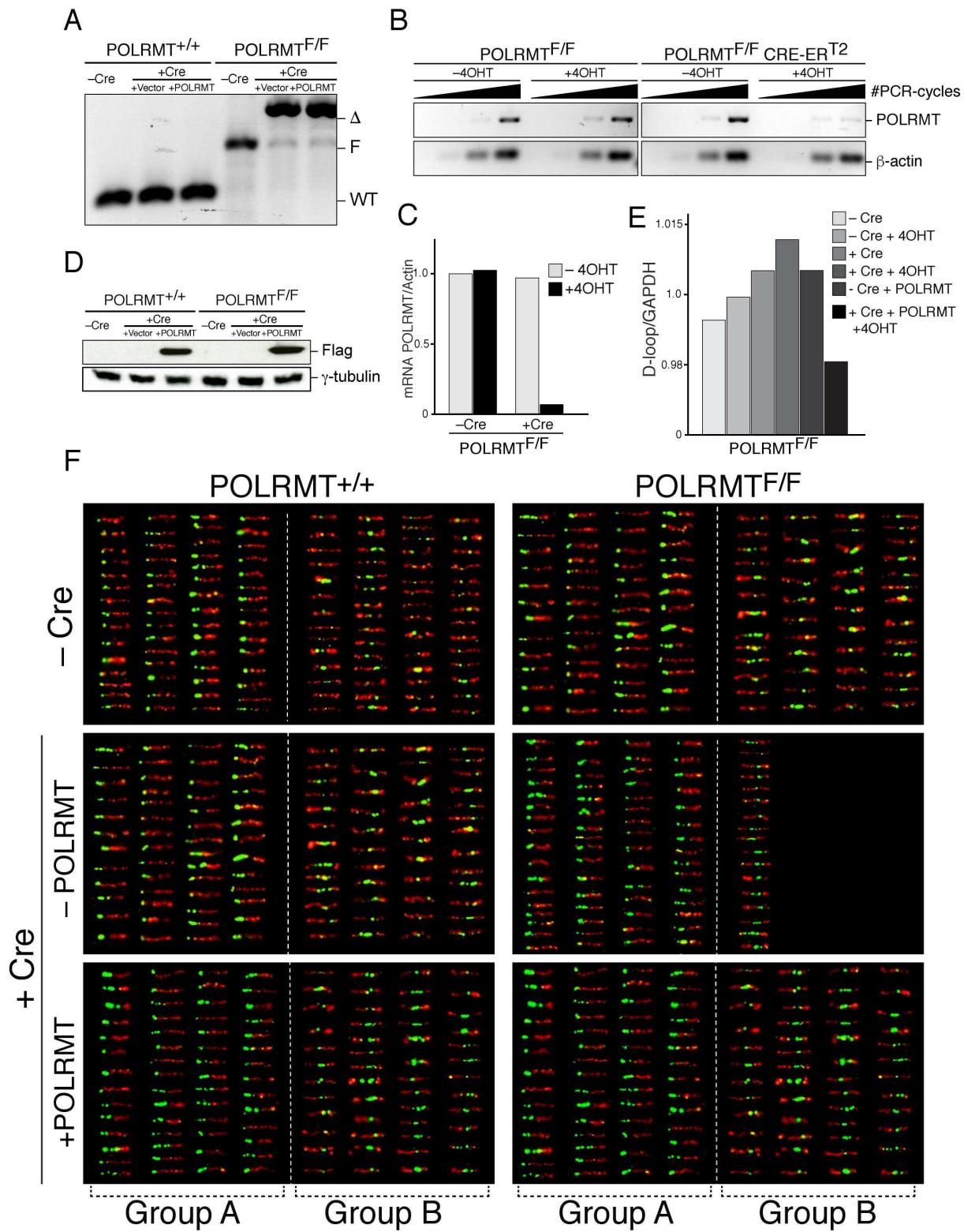


Figure S3. Related to Figure 2. POLRMT promotes the initiation of OriL. (A) Genotyping PCR for POLRMT in *Polrmt^{+/+}* and *Polrmt^{F/F}* MEFs following the expression of a tamoxifen inducible Cre-ER^{T2}. Cells were analyzed 48 hours following Cre expression. (B) Semi-quantitative PCR for *Polrmt^{F/F}* MEFs following the expression of a tamoxifen inducible Cre-ER^{T2}. Cells were analyzed 48 hours following Cre expression. Samples were collected after 25, 30, 35, and 40 rounds of amplification. (C) Quantification of the semi-quantitative PCR shown in B. Band intensity was measured using Image Lab and the intensity is normalized to the untreated sample. (D) Western blot displaying the expression of a flag-tagged POLRMT in *Polrmt^{+/+}* and *Polrmt^{F/F}* MEFs with the indicated treatment. (E) Quantification of mtDNA content (D-loop) in MEFs with the indicated treatment and analyzed 48 hours post-Cre treatment. Values were normalized to GAPDH. It is important to note that while long-term depletion of POLRMT impacts replication from both OriL and OriH (OriH effect is due to impaired transcription), the analysis of POLRMT depleted cells shortly after Cre depletion (48 hours) did not affect mtDNA content. As such, the effect of POLRMT depletion on OriL firing is unlikely to be due to pleiotropic effect. (F) Representative images of mtDNA molecules obtained from one of three experiments using MEFs with the indicated genotype and treatment, and subject to mito-SMARD.

Figure S4

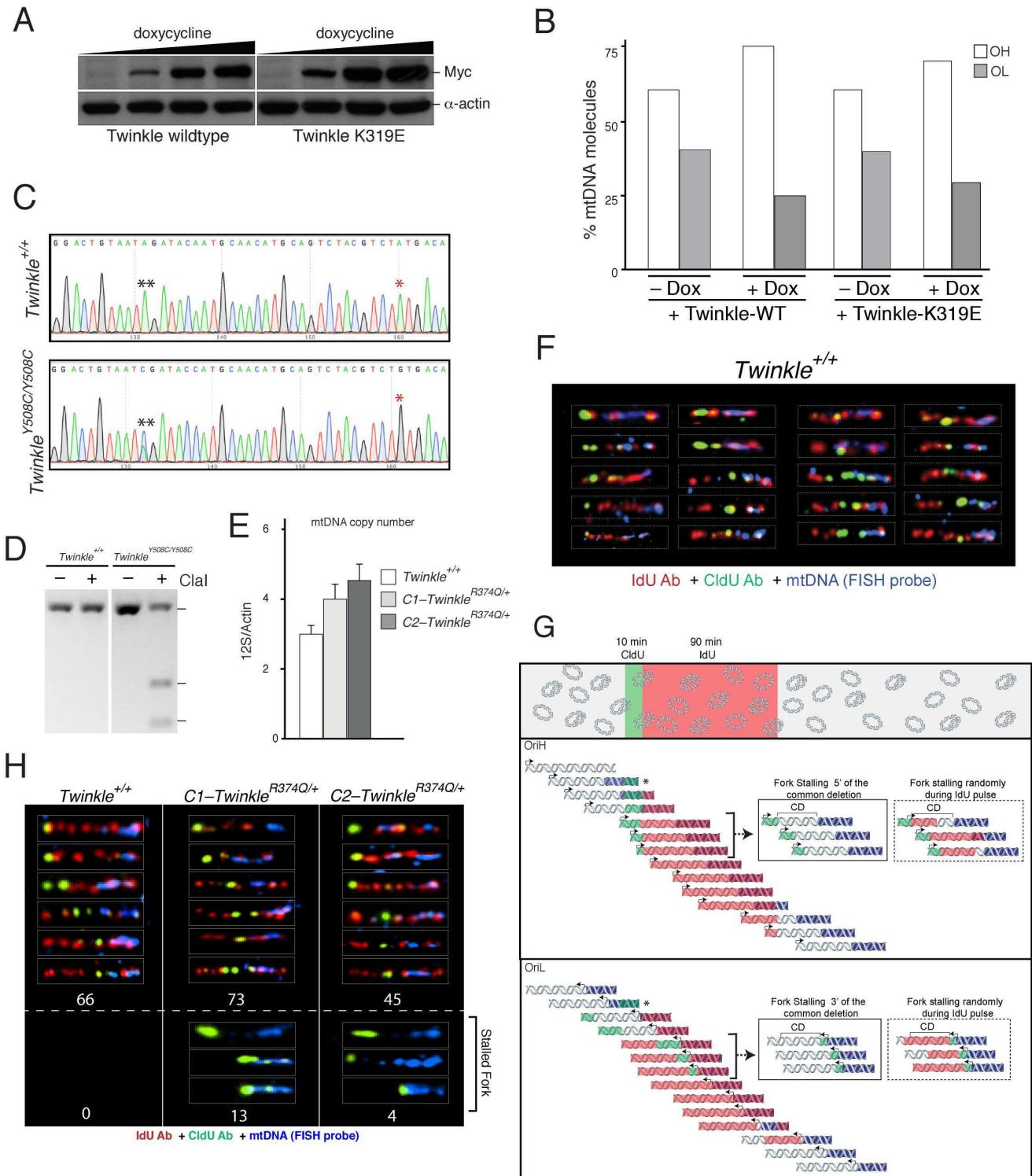
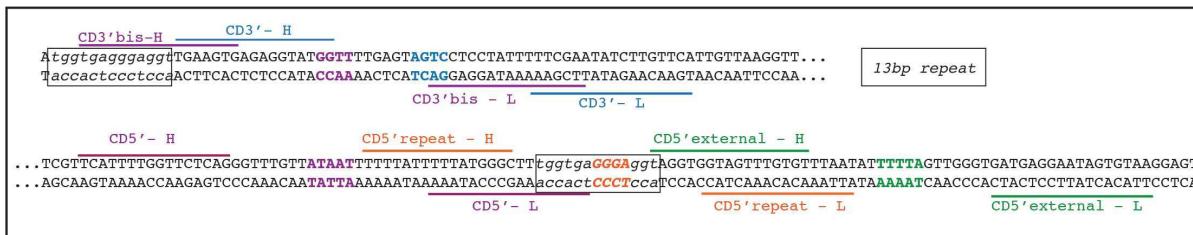


Figure S4. Related to Figure 3. Evidence for replication fork stalling in cells expressing mutant Twinkle. (A) Western blot for the expression of myc-tagged Twinkle in doxycycline-inducible HeLa Flip-In cells. Cells were treated with doxycycline in a concentration dependent manner (0, 1, 10, or 100 ng/mL) for 24 hours. (B) Frequency of Group A (OH) and Group B (OL) molecules in HeLa Flip-In cells following the induction of wild-type and K319E Twinkle. (C) Targeting was confirmed by Sanger Sequencing of cDNA generated from the indicated clones. The single asterisk marks the Y508C mutated allele. Silent mutations were introduced in the donor sequence to prevent repeated CRISPR cleavage after incorporation of the donor. An additional silent mutation was introduced upstream of the targeted residue to create a recognition site for Clal restriction enzyme (marked with two asterisks.). To note, while the Y508C targeting was homozygote, the Clal site is heterozygote). (D) Genotyping PCR to detect homozygous Twinkle^{Y508C/Y508C} alleles generated by CRISPR/Cas9 targeting of RPE-1. The PCR product was treated with a restriction enzyme that specifically cleaves that targeted allele. As mentioned in (C), incomplete digestion of the cDNA amplicon by Clal is due to incorporation of this silent mutation on only one locus. (E) mtDNA copy number in RPE-1 cells with the indicated genotype. Cells were analyzed shortly (<20 population doubling) after generating the CRISPR/CAS9 targeted clones. We avoided long term culture of the cells as it is expected to result in mtDNA depletion. Values are mean of three independent experiments with SDs. (F) Examples of dual-labeled mtDNA molecules from Twinkle^{+/+} cells subject to mito-SMARD and stained with mito-FISH probes. It is worth noting that the mtDNA probes were designed to be sufficiently distant from one another as to allow for the anti-ClDU and anti-IdU antibodies to detect halogenated nucleotides in this region. (G) Criteria for identifying stalling events in the vicinity of the common deletion. Cells were sequentially pulse-labeled with media containing ClDU for 10 minutes and IdU for 90 minutes, followed by a chase period. Presented is the predicted pattern of labeling for heavy- and light-strand replication, marking mtDNA ends with a FISH probe. The fragments are organized based on the fork progressing from OriH (Top) and OriL (Bottom). Replication fork stalling during the 90 minute pulses will yield partially labeled molecules that contain both green and red (highlighted in dotted boxes). On the other hand, replication fork stalling in the proximity of the 5' and 3' junctions of the common deletion will yield molecules that are labeled in green (highlighted in solid boxes). It is important to note that such molecules can be distinguished from single-labeled molecules resulting from replication termination during the ClDU pulse. In the later molecules (marked with asterisk), the green label is expected to exclusively overlap with the FISH probes. (H) Representative mtDNA molecules labeled with ClDU (green), IdU (red), and ClDU-IdU (green-red) and stained with mito-FISH probes (blue). Presented mtDNA molecules

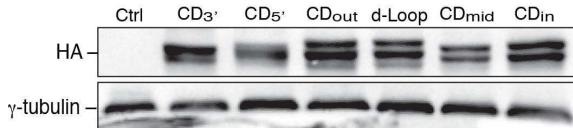
were obtained from Twinkle^{+/+} and Twinkle^{R374Q/+} cells. Top: Examples of fully replicating mtDNA molecules with IdU/CldU incorporation patterns consistent with replication initiation from OriL and OriH. Bottom: Molecules shown represent labeling patterns consistent with fork stalling near the end of a CldU pulse. These molecules were only present in cells expressing mutant Twinkle.

Figure S5

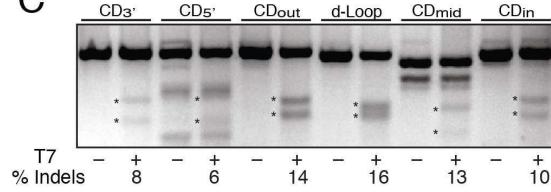
A



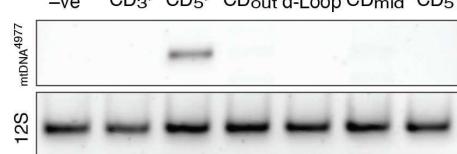
B



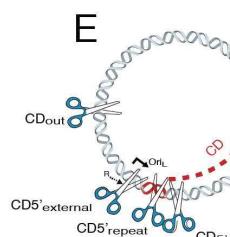
C



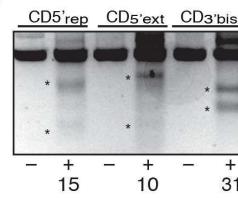
D



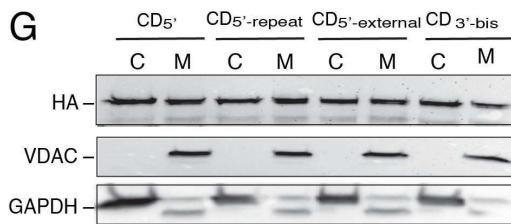
E



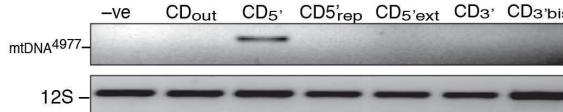
F



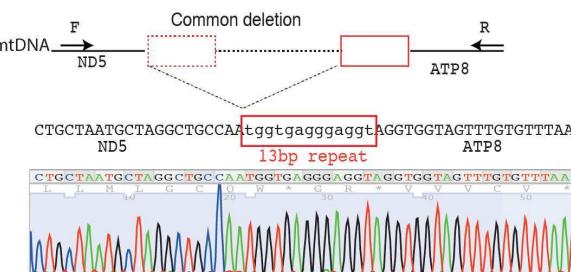
G



H



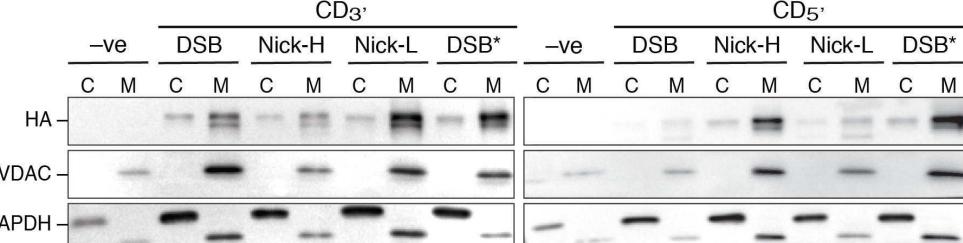
I



J

Deletions	0,39%
Insertions	0,04%
Mutations(>1 bp)	0,24%
SNP	15,51%
Common deletion	83,78%

K



L

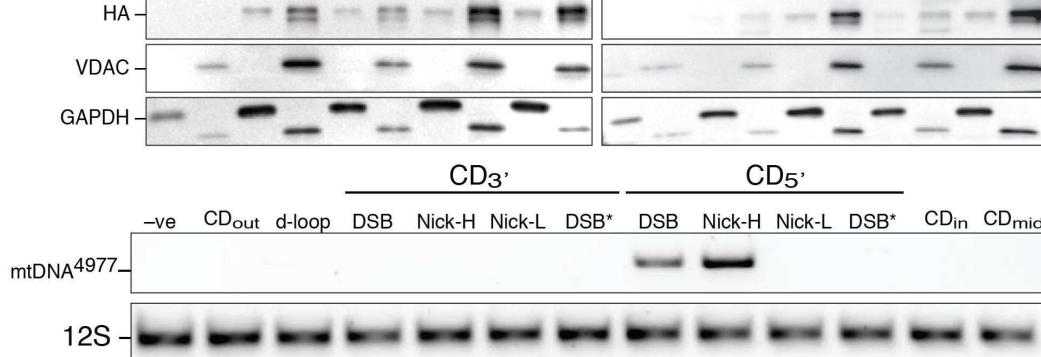


Figure S5. Related to Figure 4. Mito-TALENs induce common deletion formation. (A) Sequence of mtDNA located downstream of the 3' repeat with the binding sites for the mito-TALEN CD_{3'} and CD_{3'bis} and around the 5' repeat with the binding sites for the mito-TALEN CD_{5'}, CD_{5' external}, and CD_{5' repeat}. Theoretical DSB sites are in bold. TALEN cleavage results primarily in 5' overhangs and the cleaved region lies in the center of the spacer region (TAL domain H: recognition of the Heavy strand and TAL domain L: recognition of the Light strand) (B) Western Blot analysis for HA tagged mito-TALEN in cells with the indicated treatment. Whole cell lysate was analyzed. (C) Efficiency of TALEN cleavage monitored by T7 endonuclease assay in U2OS cells co-transfected with the indicated nuclear TALENs and a plasmid containing the corresponding targeting sequence (36h after transfection). The asterisks indicate T7 endonuclease cleaved bands. Non-specific bands were detected for CD_{5'} and CD_{mid}. (D) PCR detection of mtDNA⁴⁹⁷⁷ in U2OS cells analysed 7 days post-transfection with the indicated mito-TALENs nuclease. 12S represented total levels of mtDNA. The asterisk marks a catalytically inactive version of the mito-TALEN. The mito-TALENs were generated using two obligatory heterodimeric FokI endonucleases to reduce potential off-target cleavage due to homodimerization of TALENs. (E) Schematic of mtDNA with cleavage sites for mito-TALEN CD_{3'}, CD_{3'bis} CD_{5'}, CD_{5' external}, CD_{5' repeat}, and CD_{out}. (F) T7 endonuclease assay to test the efficiency of CD_{5' repeat}, CD_{5' external} and CD_{3'bis} as described in (C). (G) Fractionation experiments in U2OS cells followed by western blot analysis to detect CD_{5'}, CD_{5' external}, CD_{5' repeat} and CD_{3'bis}, in cytoplasm (C) and mitochondria (M) 48h after transfection. GAPDH is a control for cytoplasmic fraction and VDAC is a mitochondrial marker (H) PCR detection of mtDNA⁴⁹⁷⁷ in U2OS cells analysed 7 days post-transfection with the indicated TALENs nuclease (I) Sanger sequencing of the PCR fragment across the common deletion junction in cells treated with mito-TALEN-CD_{5'}. (J) Frequency of the different types of indels (deletions, insertions, and mutations) estimated by deep sequencing of the mitochondrial locus targeted by the mito-TALEN-CD_{5'}. (K) Fractionation experiments in U2OS cells followed by western blot analysis detecting mito-TALEN-nickase in cytoplasm (C) and mitochondria (M) (L) PCR detection of the common deletion in cells treated with the indicated mito-TALEN nickase (Nick-H induces a break in the Heavy-strand, Nick-L induces a break in the Light-strand) and nuclease (DSB). The asterisk marks a catalytically inactive control TALEN.

Figure S6

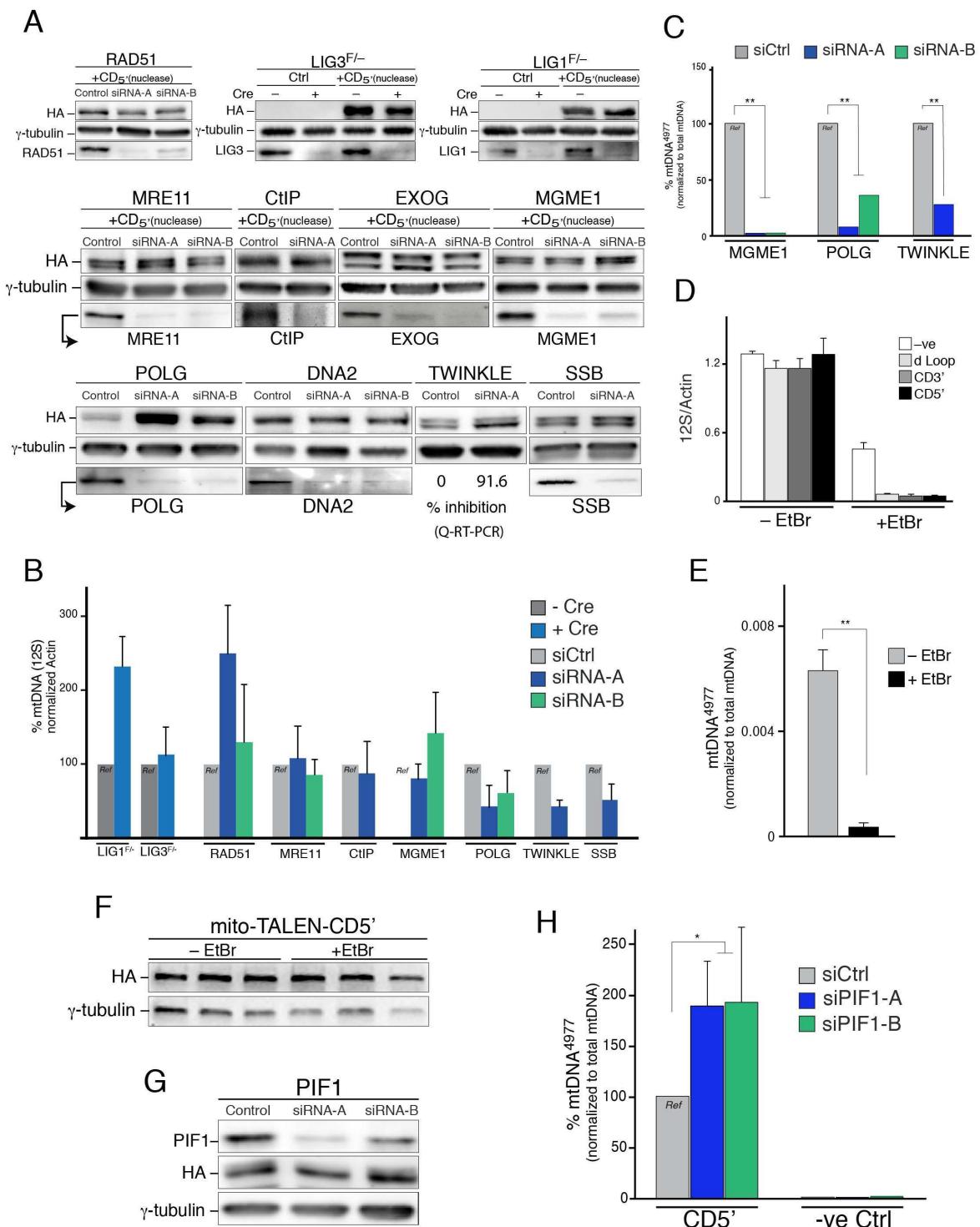


Figure S6. Related to Figure 5. The role of the mitochondrial replisome during common deletion formation. (A) Western blot analysis to confirm the expression of mito-TALEN-CD_{5'} (HA tag) in HCT116 cells with indicated treatment. (B) mtDNA content (12S/Actin) in cells with the indicated siRNA treatment. It is important to note that in order to account for the variation in mtDNA content due to various genetic perturbations (especially ones in which replisome was inhibited), the common deletion was normalized to mtDNA content in each sample. Values are mean of three independent experiments with SDs. (C) Q-PCR analysis to determine the levels of mtDNA⁴⁹⁷⁷ induced by mito-TALEN-CD_{5'} Nickase in cells treated with the indicated siRNA. The frequency of mtDNA⁴⁹⁷⁷ was normalized to total mtDNA (12S) and presented relative to siControl. Values are mean of three independent experiments with SDs. (D) Graph representing mtDNA content in cells treated with EtBr and the indicated mito-TALENs. Values are mean of three independent experiments with SDs. (E) Q-PCR analysis to determine the levels of mtDNA⁴⁹⁷⁷ induced by mito-TALEN-CD_{5'} in cells treated with EtBr. Values are mean of three independent experiments with SDs. (F) Western blot analyses for the expression HA-tagged mito-TALEN in U2OS cells treated with EtBr. γ -tubulin serves as a loading control. (G) Western blot analysis for mito-TALEN-CD_{5'} (HA tag), PIF1, and γ -tubulin in HCT116 cells with the indicated treatment. (H) Quantification of mtDNA⁴⁹⁷⁷ in HCT116 cells with the indicated siRNA treatment, analysed 7 days post-transfection with mito-TALEN-CD_{5'} or – ve Ctrl (no TALEN treatment). Values are mean of three independent experiments with SDs. Values were normalized to total mtDNA content (12S).

SUPPLEMENTAL TABLES**Table S1. Related to STAR methods. siRNA sequences used in this study.**

Si Control	5'UGUGACUUUAUCGGUGUGAC3'
Si CtIP	5'GCUAAAACAGGAACGAAUC3'
Si MRE11 A	5'GAUGCCAUUGAGGAAUUAG3'
Si MRE11 B	5'GCUAAUGACUCUGAUGAUA3'
Si RAD51 A	5'GAGCUUGACAAACUACUUC3'
Si RAD51 B	5'UGUAGCAUAUGCUCGAGCG3'
Si MGME1 A	5'GGGAGAAGAUGGUUUAAA3'
Si MGME1 B	5'GUCUGGUUACAUUGAAAGU3'
Si PIF A	5'GCCAGAGCAUCUUCUCA3'
Si PIF B	5'GGGAAGUCAUAUCUGCUAA3'
Si SSB	5'GUACUUCGUCAGUUUGUAA3'
Si DNA2 A	5'AGACAAGGUUCCAGCGCCA3'
Si DNA2 B	5'CCACACGCCAAUUCUAAU3'
Si Twinkle	invitrogen 10620318
Si POIG A	5'GGAUGGUAAUAGCUGUAAU3'
Si POLG B	Ambion S10787
Si EXOG A	5'GACCUUACCUCAGACUAGA3'
Si EXOG B	Ambion S19298

Table S2. Related to STAR methods. Mito-FISH probes. Related to STAR methods.

5'-GTATAATAGTACCGTTAACCTCCAATTAACTAGTTTGACAACATT-3'
5'-TTCAAAAAAGAGTAATAAACCTCGCCTTAATTTAATAATCACACC-3'
5'-CCCTCCTAGCCTACTACTAATAATTATTACATTGACTACCAC-3'
5'-AACTCAACGGCTACATAGAAAAATCCACCCCTACGAGTCGGCT-3'
5'-TCGACCCTATATCCCCCGCCGCCTCTCCATAAAATT-3'
5'-TCTTCTTAGCTATTACCTTCTTATTATTGATCTAGAAATTGCC-3'
5'-CCCTCCTTTACCCCTACCATGAGCCTACAAACAACCTGC-3'
5'-CACTAATAGTTATGTCATCCCTCTTATTAAATCATCATCTAGCCC-3'
5'-TAAGTCTGGCCTATGAGTGAATACAAAAAGGATTAGACTGAACCG-3'
5'-AATTGGTATATAGTTAAACAAAACGAATGATTGACTCATTAAT-3'
5'-ATTATGATAATCATATTACCAAATGCCCTCATTTACATAAATATT-3'
5'-TTAACTACTGCAATTACCATCTCACTCTAGGAATACTAGTATATCGC-3'
5'-GCTCACACCTCATATCCTCCCTACTATGCCCTAGAAGGAATAATAC-3'
5'-TATCGCTGTCATTATAGCTACTCTCATACCCCTAACACCCACT-3'
5'-CCCTCTAGCCAATTGTGCCCTATTGCCATACTAGTCTTGCCG-3'
5'-CCTGCGAAGCAGCGGTGGGCCTAGCCCTACTAGTCTCAATCTC-3'
5'-CCAACACATATGGCCTAGACTACGTACATAACCTAAACCTACTCC-3'
5'-AATGCTAAAACATACTGTCACAAATTATATTACTACCACTGAC-3'
5'-ATGACTTTCCAAAAAACACATAATTGAATCAACACAAACCAACCA-3'
5'-CAGCCTAATTATTAGCATCATCCCTCTACTATTTTAACCAAAT-3'

5'-CAACAACAACCTATTAGCTGTTCCCAACCTTCCGACCC-3'
5'-CCTAACAACCCCCTCTAATACTAACTACCTGACTCCTACCCCT-3'
5'-CACAAATCATGGCAAGCCAACGCCACTTATCCAGTGAACCAACTATC-3'
5'-ACGAAAAAAACTCTACCTCTCTATACTAATCTCCCTACAAATCTC-3'
5'-CTTAATTATAACATTACAGGCCACAGAACTAATCATATTATATCT-3'
5'-CTTCTCGAAACCACACTTATCCCCACCTGGCTATCATCACCCG-3'
5'-ATGAGGCAACCAGCCAGAACGCCTGAACGCAGGCACATACTTC-3'
5'-CCTATTCTACACCCTAGTAGGCTCCCTCCACTCATCGCACT-3'
5'-AATTTACACTCACAAACACCCCTAGGCTCACTAAACATTCTACTACT-3'
5'-CACTCTCACTGCCAAGAACTATCAAACCTTGAGCCAACAACCTT-3'
5'-AATATGACTAGCTTACACAATAGCTTTATAGTAAAGATACTCTT-3'
5'-TTACGGACTCCACTTATGACTCCCTAAAGCCCAGTCGAAGCCC-3'
5'-CCCCATCGCTGGGTCAATAGTACTTGCAGTACTCTAAAACCT-3'
5'-AGGCGGCTATGGTATAATACGCCACTCATTCTAACCCCCC-3'
5'-CCTGACAAAACACATAGCCTACCCCTCTGTACTATCCCTATG-3'
5'-AGGCATAATTATAACAAGCTCCATCTGCCTACGACAAACAGACCT-3'
5'-AAAATCGCTCATTGCATACTCTCAATCAGCCACATAGCCCTCGT-3'
5'-AGTAACAGCCATTCTCATCCAAACCCCTGAAGCTTCACCGGC-3'
5'-CGCAGTCATTCTCATATGCCAACGGGCTACATCCTCATTACT-3'
5'-ATTCTGCCTAGCAAACACTCAAACACTACGAACGCACTCACAGTCGCAT-3'
5'-CATAATCCTCTCTAAGGACTCAAACACTACTCCACTAATAGC-3'
5'-TTTTGATGACTCTAGCAAGCCTCGCTAACCTCGCCTACCC-3'
5'-CCCCACTATTAACCTACTGGAGAACTCTGTGCTAGTAACCAC-3'
5'-GTTCTCCTGATCAAATATCACTCTCCTACTTACAGGACTCAACAT-3'
5'-ACTAGTCACAGCCCTATACTCCCTACATATTACCAACACA-3'
5'-ATGGGGCTCACTCACCCACCACATTAACAACATAAAACCCCTCATT-3'
5'-CACACGAGAAAACACCCCTCATGTTCATACACCTATCCCCCATTCT-3'
5'-CCTCCTATCCCTAACCCCGACATCATTACCGGGTTTCCTCTTG-3'
5'-TAAATATAGTTAACAAAACATCAGATTGTGAATCTGACAACAGA-3'
5'-AGGCTTACGACCCCTTATTACCGAGAAAGCTCACAAGAACTGCT-3'
5'-AACTCATGCCCATGTCTAACAAACATGGCTTCTCAACTTTAA-3'
5'-AGGATAACAGCTATCCATTGGCTTAGGCCAAAAATTGGT-3'
5'-CAACTCCAAATAAAAGTAATAACCAGTCACACTATAACCACC-3'
5'-CTAACCCCTGACTTCCCTAACCCCCCATCCTTACCAACCTCG-3'
5'-GTTAACCCCTAACAAAAAAACTCATACCCCCATTATGTAAAATCC-3'
5'-ATTGTCGCATCCACCTTATTATCAGTCTTCCCCACAACAATA-3'
5'-TTCATGTGCCTAGACCAAGAAGTTATTATCTGAACTGACACTGA-3'
5'-GCCACAAACCCAAACAACCCAGCTCTCCCTAACGCTCAAACACTAGAC-3'
5'-TACTTCTCCATAATATTACCCCTGTAGCATTGCTTACATGG-3'
5'-TCCATCATAGAATTCTCACTGTGATATATAAACTCAGACCCAAAC-3'
5'-ATTAATCAGTTCTCAAATATCTACTCATCTTCTAATTACCATACT-3'
5'-CTAATCTTAGTTACCGCTAACAAACCTATTCCAACGTTCATCGGC-3'
5'-TGAGAGGGCGTAGGAATTATCCTTCTGCTCATCAGTTGATGA-3'
5'-TACGCCGAGCAGATGCCAACACAGCAGCCATTCAAGCAATCC-3'
5'-CTATACAACCGTATCGGCATATCGGTTCATCCTCGCCTAGCA-3'
5'-TGATTTATCCTACACTCCAACCTATGAGACCCACAACAAATGCC-3'
5'-CTTCTAAACGCTAATCCAAGCCTACCCCACTACTAGGCCTCCTC-3'
5'-CTAGCAGCAGCAGGCAAATCAGCCCAATTAGGTCTCCACCCCT-3'
5'-TGACTCCCTCAGCCATAGAAGGCCCCACCCAGTCTCAGCCC-3'
5'-CTACTCCACTCAAGCACTATAGTTGTAGCAGGAATCTTCTACTC-3'
5'-ATCCGCTCCACCCCTAGCAGAAAATAGCCCACAAATCCAAACT-3'
5'-CTAACACTATGCTTAGGCGCTATCACCACCTGTTCGCAGCAGTC-3'
5'-TGCGCCCTTACACAAAATGACATCAAAAAATCGTAGCCTCTCC-3'

5'-ACTTCAAGTCAACTAGGACTCATAATAGTTACAATCGGCATCAAC-3'
5'-CAACCACACCTAGCATTCCTGCACATCTGTACCCACGCCCTCT-3'
5'-TTCAAAGCCATACTATTTATGTGCTCCGGTCCATCATCCACAAC-3'
5'-CTTAACAATGAACAAAGATATTGAAAAAATAGGAGGACTACTAAA-3'
5'-ACCATACTCTCACTCAACCTCCCTCACCATGGCAGCCTAG-3'
5'-GCATTAGCAGGAATACCTTCCTCACAGGTTCTACTCCAAAGAC-3'
5'-CACATCATCGAAACCGCAAACATATCATAACACAAACGCCCTGAGCC-3'
5'-CTATCTATTACTCTCATCGTACCTCCCTGACAAGCGCCTATAGC-3'
5'-ACTCGAATAATTCTCTCACCTAACAGGTCAACCTCGCTCCCC-3'
5'-ACCCCTACTAACATTAACGAAAATAACCCCACCTACTAAACCCC-3'
5'-ATTAAACGCCCTGGCAGCCGGAAGCCTATTGCGAGGATTCTCA-3'
5'-ATTACTAACACATTCCCCCGATCCCCCTTCAAACAAACAATC-3'
5'-CCCCTCTACCTAAAACACAGCCCTAGACCTAACCTAACCAACAAACTT-3'
5'-GGACTTCTAACAGCCCTAGACCTAACCTAACCAACAAACTT-3'
5'-AAAATAAAATCCCCACTATGCACATTATTCTCCAACATACTC-3'
5'-GGATTCTACCCCTAGCATCACACACCCGACAATCCCCTATCTAGGC-3'
5'-CTTCTTACGAGCCAAAACCTGCCCCACTCCTCTAGACCTAACCC-3'
5'-TGACTAGAAAAGCTATTACCTAAAACAATTTCACAGCACCAAATC-3'
5'-TCCACCTCCATCATCACCTAACCCAAAAGGCATAATTAAACTT-3'
5'-TACTCCTCTTTCTTCCACTCATCCTAACCCCTACTCCTA-3'
5'-ATCACATAACCTATTCCCCGAGCAATCTCAATTACAATATAC-3'
5'-ACCAACAAACAATGTTAACCGAGTAACACTACTAACACGCC-3'
5'-ATAATCATACAAAGCCCCCGACCAATAGGATCCTCCGAATCAA-3'
5'-CCCTGACCCCTCTCTTCAAATTATTAGCCTTACACTATT-3'
5'-AAAGTTACCAACACCACCCATCATACTCTTCAACCCACAG-3'
5'-CACCAATCCTACCTCCATCGCTAACCCACTAAACACTCACCAA-3'
5'-GACCTAACCCCTGACCCCCATGCCTCAGGAACTCCTCAATA-3'
5'-AGCCATCGCTGTAGTATATCAAAGACAACCATCATTCCCCCTAA-3'
5'-ATAAATTAAAAAAACTATTAAACCCATATAACCTCCCCAAAATTCA-3'
5'-CAGAATAATAACACACCCGACCACACCGCTAACATCAACTAA-3'
5'-ACCCCCATAAAATAGGAGAAGGCTTAGAAGAAAACCCCACAAACCC-3'
5'-CATTACTAAACCCACACTAACAGAAACAAAGCATAACATCATTAT-3'
5'-TCTCGCACGGACTACAACCAACGACCAATGATATGAAAAACCATCG-3'
5'-TTGTATTCAACTACAAGAACACCAATGACCCAATACGCAAAC-3'
5'-TAACCCCCATAATAAAATTAAACCACTCATTATCGACCTCCC-3'
5'-CACCCCATCCAACACATCTCCGATGATGAAACTCCTGGCTCACTC-3'
5'-CCTGGCGCCTGCCTGATCCTCAAATCACCACAGGACTATTC-3'
5'-CCTAGCCATGCACTACTCACCAAGACGCCCTAACCGCCTTTCA-3'
5'-ATCAATGCCAACATCACTCGAGACGTAATTATGGCTGAATCAT-3'
5'-CCGCTACCTTCAGGCCATGGCGCCTCAATTATTCTTATCTGC-3'
5'-CCTCTTCTACACATGGCGAGGCCTATATTACGGATCATTCT-3'
5'-CTACTCAGAAACCTGAAACATCGGCATTATCCTCTGCTGCAAC-3'
5'-TATAGCAACAGCCTCATAGGCTATGCTCTCCGTGAGGCCAAAT-3'
5'-ATCATTCTGAGGGGCCACAGTAATTACAAACTTACTATCCGCCAT-3'
5'-CCCATACATTGGGACAGACCTAGTTCAATGAATCTGAGGAGGCTA-3'
5'-CTCAGTAGACAGTCCCACCCCTCACACGATTCTTACCTTCACTT-3'
5'-CATCTTGCCTTCATTATTGCAAGCCCTAGCAACACTCCACCTCC-3'
5'-CCTATTCTGCACGAAACGGGATCAAACAAACCCCTAGGAATCAC-3'
5'-CTCCCATCCGATAAAATCACCTCCACCCCTACTACACAATCAA-3'
5'-AGACGCCCTGGCTTACTTCTCTTCTCCTTAATGACATT-3'
5'-AACACTATTCTCACCAGACCTCCTAGGCGACCCAGACAATTATAC-3'
5'-CCTAGCCAACCCCTAAACACCCCTCCCCACATCAAGCCCGAA-3'
5'-ATGATATTCTTCTATTGCGCTACACAATTCTCGATCGTCCCTAA-3'

5'-CAAACTAGGAGGCGTCCTGCCCTATTACTATCCATCCTCATCCT-3'
5'-AGCAATAATCCCCATCCTCCATATATCCAAACAAACAAGCATAAT-3'
5'-ATTCGCCCACTAAGCCAATCACTTATTGACTCCTAGCCGCAGA-3'
5'-CCTCCTCATTCTAACCTGAATCGGAGGACAACCAGTAAGCTACCC-3'
5'-TTTACCATCATTGGACAAGTAGCATCCGTACTATACTTCACAAC-3'
5'-AATCCTAATCCTAACCAACTATCTCCCTAATTGAAAACAAAT-3'
5'-ACTCAAATGGGCCTGTCCTGTAGTATAAAACTAACACCCAGTCT-3'
5'-TGTAAACCGGAGATGAAAACCTTTCCAAGGACAAATCAGAGAA-3'
5'-AAAGTCTTAACTCCACCATTAGCACCCAAAGCTAAGATTCTAAT-3'
5'-TTAAACTATTCTCTGTTCTTCATGGGGAAAGCAGATTGGGTACC-3'
5'-ACCCAAGTATTGACTCACCCATCAACAACCGCTATGTATTCGTA-3'
5'-CATTACTGCCAGCCACCATGAATATTGTACGGTACCATAAACT-3'
5'-TGACCACCTGTAGTACATAAAACCAATCCACATCAAAACCCC-3'
5'-TCCCCATGCTTACAAGCAAGTACAGCAATCAACCCTCAACTATCA-3'
5'-CACATCAACTGCAACTCCAAAGCCACCCCTCACCCACTAGGAT-3'
5'-TACCAACAAACCTACCCACCCCTAACAGTACATAGTACATAAAGC-3'
5'-CATTTACCGTACATAGCACATTACAGTCAAATCCCTCTCGTCCC-3'

Table S3. Related to STAR methods. mito-TALEN target sequences.

TALEN	Target sequences (T-)	FOK1
d-Loop Light strand	GGTCCTACTTCAGGG	ELD
d-Loop Heavy strand	TATTTAAGGGAAACGT	KKR
CDin Light strand	TCAACCAATAGCCCTG	ELD
CDin Heavy strand	GGCCTGCAGTAATGTT	KKR
CDout Light strand	CATATACCAAATCTCT	ELD
CDout Heavy strand	GAGAGAGTGAGGAGAA	KKR
CD5' Light strand	CACCAAAGCCCATAAA	ELD
CD5' Heavy strand	TCATTTGGTTCTCAG	KKR
CD3' Light strand	GAACAAGATATTGAA	ELD
CD3' Heavy strand	GAAGTGAGAGGTATGG	KKR
CDmid Light strand	AATTATTAGCATCATC	ELD
CDmid Heavy strand	AGGTTGTTGTTGATT	KKR
CD5'repeat Light strand	ATTAAACACAAACTAC	ELD
CD5'repeat Heavy strand	TTTTATTTTATGGGC	KKR
CD5'external Light strand	TACACTATTCCCTCATC	ELD
CD5'external Heavy strand	AGGTGGTAGTTGTGT	KKR
CD3'bis Light strand	TCGAAAATAGGAGGA	ELD
CD3'bis Heavy strand	GAGGGAGGTTGAAGTG	KKR

Table S4. Related to STAR methods. Primers for detection amplification of mito-TALEN region (T7 assay) and quantification of the common deletion.

T7-CD ₅ FWD	TTACAGTGAAATGCCCAACT
------------------------	----------------------

T7-CD _{5'} REV	TCAGGTTCGTCCTTAGTGTG
T7-CD _{3'} FWD	CGCCCTTACACAAAATGACA
T7-CD _{3'} REV	CAGGGAGGTAGCGATGAGAG
T7-CD _{OUT} FWD	CAATATACTCTCCGGACAATGA
T7-CD _{OUT} REV	ATGGTTATGTTAGGGTTGACGG
T7-CD _{IN} FWD	TATGAGCGGGCACAGTGAT
T7-CD _{IN} REV	TGTTGTCGTGCAGGTAGAGG
T7-CD _{dLoop} FWD	CATTACAGTCAAATCCCTCTCG
T7-CD _{dLoop} REV	TACTGCGACATAGGGTGCTC
T7-CD _{MID} FWD	AGCCCTACTAGTCTCAATCTCCA
T7-CD _{MID} REV	CGTTGGCTTGCCATGATTGT
Common deletion PCR1 FWD	5'AACCACAGTTCATGCCCATC3'
Common deletion PCR1 REV	5'TGTTAGTAAGGGTGGGAAGC3'
Common deletion PCR2 FWD	5'ACCTATAGCACCCCCCTAC3'
Common deletion PCR2 REV	5'CTTGTCAAGGAGGTAGCGATG3'
12S FWD	5'TAGCCCTAACCTCAACAGT3'
12S REV	5'TGCGCTTACTTGTAGCCTTCAT3'
Common deletion Q-PCR FWD	5'TTCCTCATACCCAACTAAAAA3'
Common deletion Q-PCR REV	5'TTCGATGATGTGGTCTTGG3'
Actin Q-PCR FWD	5'TCACCCACACTGTGCCATCTACGA3'
Actin Q-PCR REV	5'CAGCGGAACCGCTCATTGCCATGG3'

Table S5. Related to STAR methods. Primers for Q-PCR.

TFAM FWD	5'-ATTCCGAAGTGTTCAGCA-3'
TFAM REV	5'-TCTGAAAGTTGCATCTGGGT-3'
TWINKLE FWD	5'-TTCTCCGACGTGCATATCCC-3'
TWINKLE REV	5'-GCGCTTCTTCTGTACCTTCT-3'
POLG FWD	5'-GCAGGAACAGTTAGTGGTGGG-3'
POLG REV	5'-GCTCATAGTATCGAGAAAACGCA-3'
POLRMT FWD	5'-TGGCGCAAAAGCTAGAGG-3'
POLRMT REV	5'-GTGAAGGTCCAGAACCTCTG-3'
SSB FWD	5'-TGGCTGACAAGGAGAACACG-3'
SSB REV	5'-CGCCACATCTCATTGTTGCTA-3'
TK1 FWD	5'-AAGTGCCTGGTCATCAAGTATG-3'
TK1 REV	5'-GCTGCCACAATTACTGTCTTGC-3'
TK2 FWD	5'-AGCAGTGGTTGTATTGAGGG-3'
TK2 REV	5'-ACATGAGGCTCAGAGGGTTATG-3'
MFN1 FWD	5'-ATGGCAGAAACGGTATCTCCA-3'

MFN1 REV	5'-GCCCTCAGTAACAAACTCCAGT-3'
MFN2 FWD	5'-AGAACTGGACCCGGTTACCA-3'
MFN2 REV	5'-CACTCGCTGATAACCCCTGA-3'
DESMIN FWD	5'-GTGGAGCGTGACAACCTGAT-3'
DESMIN REV	5'-GATGGTCTCATACTGAGCCG-3'
CREATINE KINASE FWD	5'-GC GGAGGCAGTGTAACCCTTG-3'
CREATINE KINASE REV	5'-CAGACCTCAGCAAGCACAAATCAC-3'
ALPHA-ACTIN FWD	5'-CAGAGCAAGCGAGGTATCC-3'
ALPHA-ACTIN REV	5'-GTCCCCAGAATCCAACACG-3'
TWINK FWD	5'-AATATGATCACTGGGCTGACC-3'
TWINK REV	5'-CGTAGACTGCATGTTGCATTG-3'
GAPDH FWD	5'-GAGTCCACTGGCGTCTTCAC-3'
GAPDH REV	5'-TTCACACCCATGACGAACAT-3'

II. Etude du rôle des protéines PARP-1 et RAD52 dans la formation de la « Common Deletion »

Afin de compléter l'étude des protéines impliquées dans l'apparition de la « Common Deletion », nous avons souhaité étudier deux autres protéines, la protéine PARP-1 et la protéine RAD52 afin de confirmer que le alt-NHEJ et la recombinaison homologue ne sont pas impliqués dans la formation des délétions de l'ADN mitochondrial.

1) Le rôle de RAD52 dans la formation de la « Common Deletion »

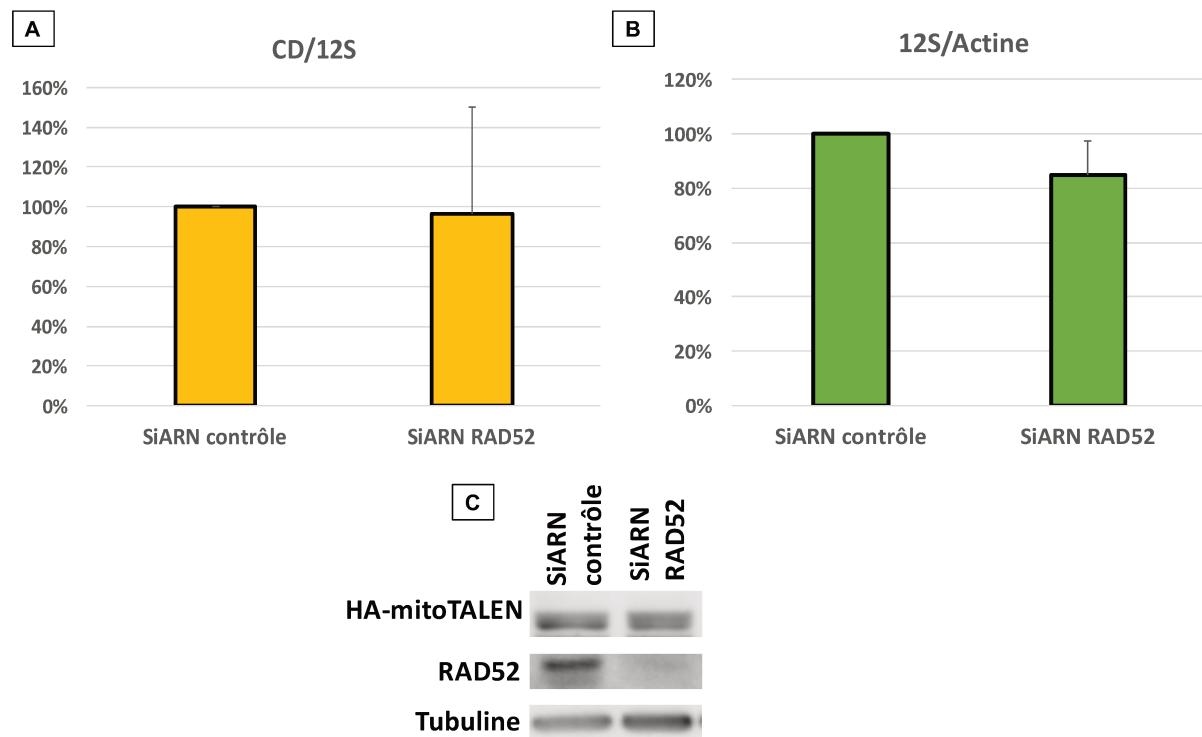


Figure 24 : Effet de l'inhibition de RAD52 sur la formation de la « Common Deletion »
A) Quantification par qPCR de la proportion de « Common Deletion » après transfection de la mitoTALEN CD 5', avec (siARN RAD52) ou sans inhibition (siARN contrôle) de la protéine RAD52. Le rapport CD/12S (« Common Deletion » /ADN mitochondrial) représente la quantité de « Common Deletion » au sein de l'ADN mitochondrial. **B)** Résultats par qPCR de la quantité d'ADN mitochondrial après transfection avec la mitoTALEN CD 5' avec (siARN RAD52) ou sans inhibition (siARN contrôle) de la protéine RAD52. Le rapport ADN mitochondrial/ADN nucléaire (12S/Actine) correspond à la quantité d'ADN mitochondrial au sein des cellules. **C)** Western permettant de visualiser l'expression de la mitoTALEN CD 5' et l'inhibition de la protéine RAD52.

Après avoir vérifié l'expression de la mitoTALEN CD 5' et l'inhibition de RAD52 par le siARN (Figure 24C), les résultats qPCR n'ont pas montré de variations significatives sur la quantité de « Common Deletion » lors de l'inhibition de la protéine RAD52 (Figure 24A). De plus, nos résultats montrent que l'inhibition de RAD52 n'influence pas la quantité d'ADN mitochondrial (Figure 24B). En conclusion, la formation de la « Common Deletion » ne semble pas dépendante de RAD52.

2) Le rôle de PARP-1 dans la formation de la « Common Deletion »

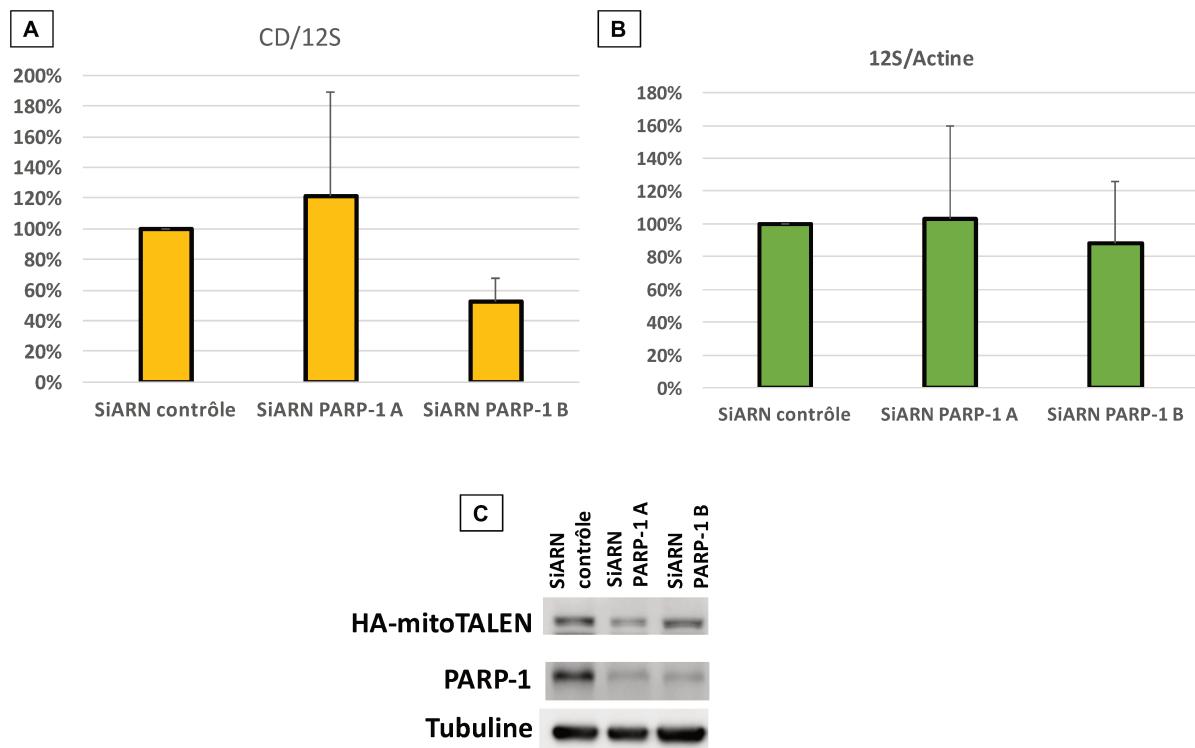


Figure 25 : Effet de l'inhibition de PARP-1 sur la formation de la « Common Deletion »

A) Quantification par qPCR de la proportion de « Common Deletion » après transfection par la mitoTALEN CD 5' et avec (siARN PARP-1 A ou B) ou sans inhibition (siARN contrôle) de la protéine PARP-1. Le rapport CD/12S (« Common Deletion » /ADN mitochondrial) représente la quantité de « Common Deletion » au sein de l'ADN mitochondrial. **B)** Résultats par qPCR de la quantité d'ADN mitochondrial après transfection avec la mitoTALEN CD 5' avec (siARN PARP-1 A ou B) ou sans inhibition (siARN contrôle) de la protéine PARP-1. Le rapport ADN mitochondrial /ADN nucléaire (12S/Actine) correspond à la quantité d'ADN mitochondrial au sein des cellules. **C)** Western permettant de visualiser l'expression de la mitoTALEN CD 5' et l'inhibition de la protéine PARP-1.

La transfection de deux siARNs différents contre PARP-1 (siARNs PARP-1 A et B) inhibent bien l'expression de la protéine PARP-1 (Figure 25C). Cependant, les résultats de qPCR ne montrent pas d'effet significatif de cette inhibition sur la formation de la délétion (même si le siARN PARP-1 B semble induire une inhibition, le siARN PARP-1 A n'a aucun effet) (Figure 25A). L'inhibition de PARP-1 n'a pas d'effet sur la quantité d'ADN mitochondrial (Figure 25B).

Concernant, les résultats pour RAD52, ils renforcent nos observations préalables faites lors de l'inhibition de la protéine RAD51 : la recombinaison homologue ne semble pas participer directement à la formation des délétions. Pour PARP-1, les résultats vont aussi dans le même sens que ceux obtenus pour MRE11 et CtIP, ils renforcent donc l'idée que le alt-NHEJ n'est pas majoritairement impliqué dans la formation de la délétion.

III. Etude de la quantité en ADN mitochondrial au cours du temps après coupure par les mitoTALENs

Lors de nos différentes expériences, nous avons observé par qPCR une diminution en ADN mitochondrial (rapport 12S/Actine) après transfection avec les mitoTALENs. Nous voulions donc mieux comprendre ce phénomène. Il est proposé dans la littérature, que les molécules endommagées peuvent être dégradées. Nous avons donc testé si nos résultats concordaient avec cette hypothèse. Pour cela, nous avons suivi l'évolution de la quantité d'ADN mitochondrial après transfection par les mitoTALENs.

Après transfection avec la mitoTALENs CD 3', CD 5', CD OUT, CD 5'ou la CD 5' Dead (qui ne coupe pas), nous avons récupéré l'ADN pendant une semaine puis quantifié la proportion d'ADN mitochondrial par qPCR. Les résultats sont normalisés en présentant le rapport 12S/Actine, qui représente la quantité d'ADN mitochondrial présente dans les cellules par rapport à la quantité de cellules (corrélée à la quantité d'ADN nucléaire).

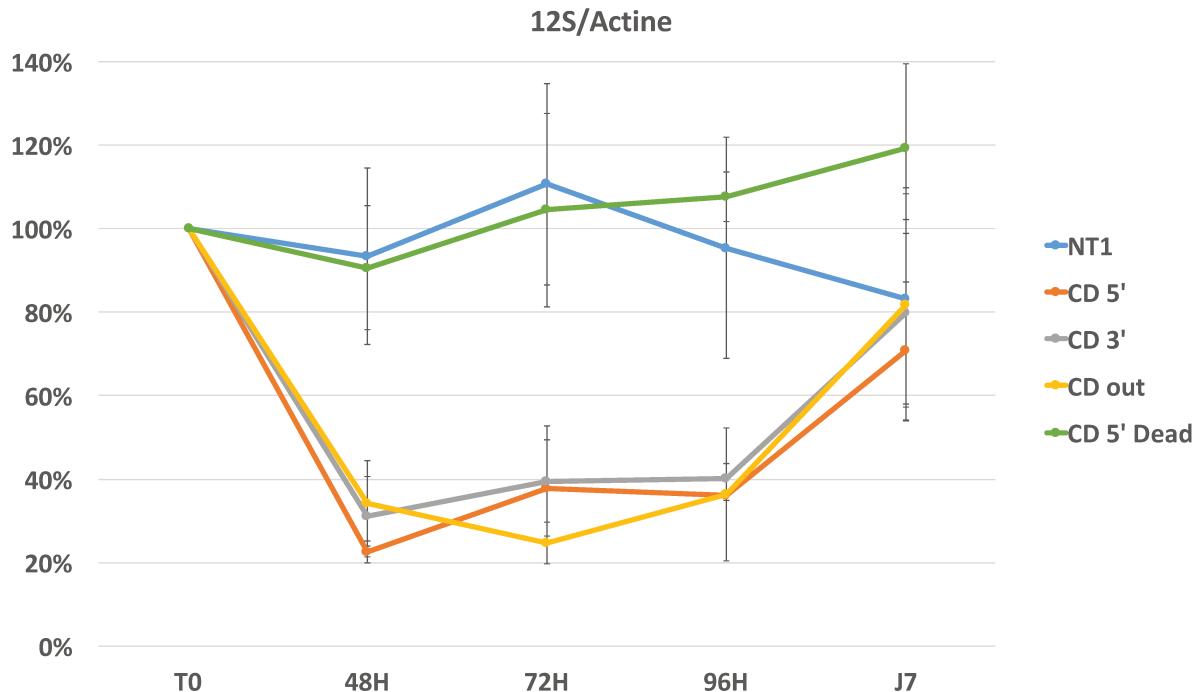


Figure 26 : Résultats de la quantité d'ADN mitochondrial au cours du temps après transfection par les mitoTALENs dans des HCT116

Résultats qPCR de la quantité d'ADN mitochondrial présente à différents temps (48h, 72h, 96h ou 7 jours (J7)) après transfection avec différentes mitoTALENs (CD 3', CD 5', CD OUT, CD 5' Dead (mitoTALEN mutée qui ne peut pas couper l'ADN)) ou seulement avec le choc électrique de la nucléofection (NT). Le rapport 12S/Actine correspond à la quantité d'ADN mitochondrial au sein des cellules. Les résultats obtenus présentent la moyenne de deux expériences indépendantes réalisées chacune en dupliquat.

Nos résultats montrent que lorsqu'il y a coupure par une mitoTALEN, une diminution significative en ADN mitochondrial est observée. Il faut environ cinq jours pour que la quantité d'ADN mitochondrial augmente à nouveau (Figure 26). Cette dégradation a lieu uniquement dans le cas d'une cassure double-brin car la transfection avec la mitoTALEN CD 5' Dead ne provoque pas de variation dans la quantité d'ADN mitochondrial (l'effet observé n'est donc pas le résultat de la liaison de la mitoTALEN sur l'ADN).

IV. Etude de la quantité en ADN mitochondrial au cours du temps après coupure par les mitoTALENs lors de l'inhibition de POLG et d'EXOG

Nous avons voulu étudier si la dégradation des molécules d'ADN mitochondrial après cassure double-brin pouvait avoir un lien avec l'activité exonucléase de la protéine ExoG. Nous avons donc voulu tester si son inhibition empêchait la dégradation des molécules d'ADN après coupure par les mitoTALENs. Nous avons aussi voulu comprendre si la réaugmentation en ADN mitochondrial environ 5 jours après la cassure double-brin pouvait impliquer la protéine

POLG, ce qui suggérerait le rôle de la réPLICATION après cassure pour restaurer la quantité initiale en ADN mitochondrial.

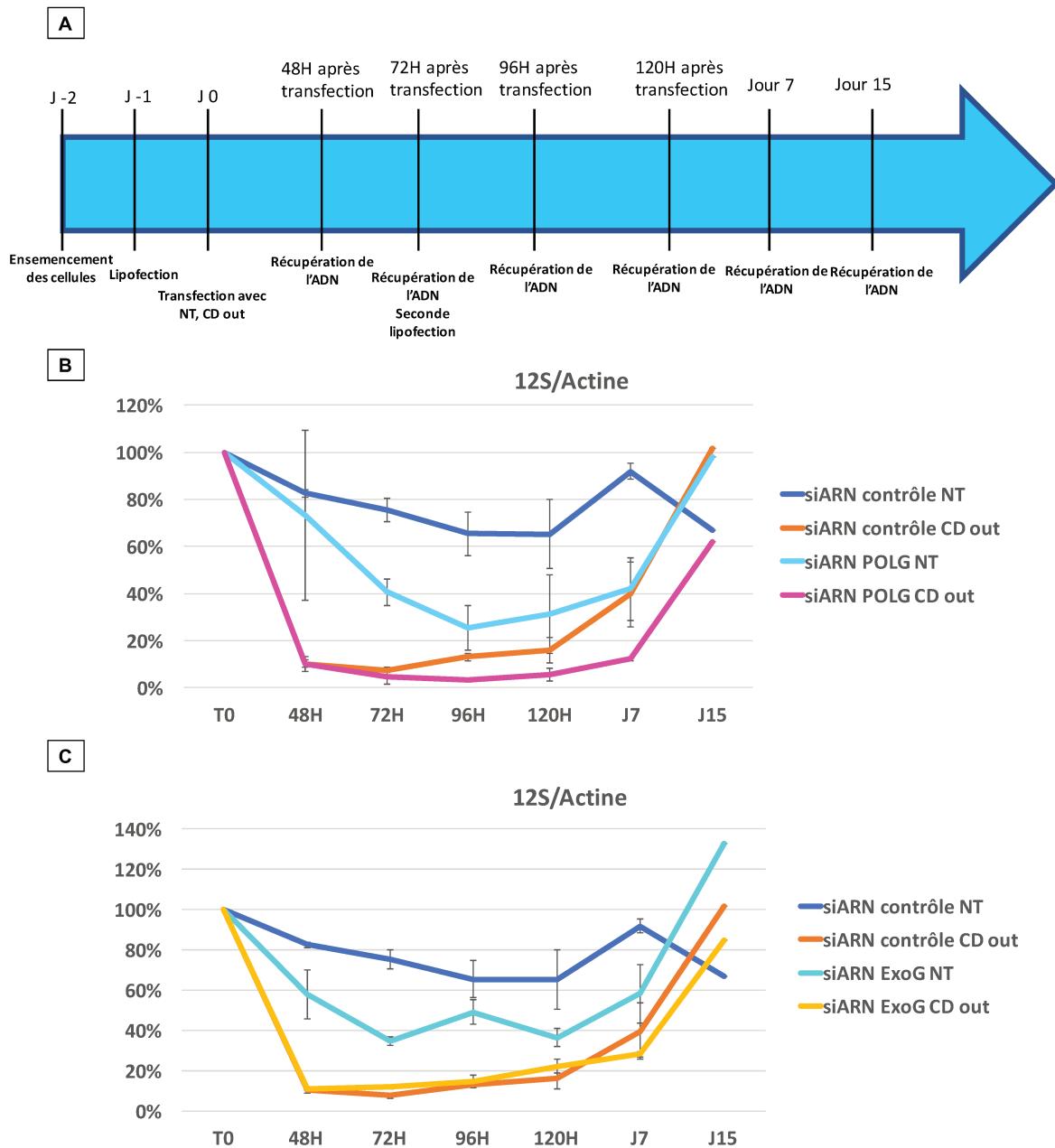


Figure 27 : Quantité d'ADN mitochondrial au cours du temps après transfction par les mitoTALENs lors de l'inhibition de POLG ou ExoG dans des HCT116

A) Schéma du déroulement de l'expérience. B) Résultats qPCR de la cinétique après transfction avec la mitoTALEN CD OUT ou sans (NT) lors de l'inhibition de la protéine POLG. L'ADN est récupéré à 48h, 72h, 96h, 120h, 7 jours (J7) et à 15 jours (J15). Le rapport 12S/Actine correspond à la quantité d'ADN mitochondrial au sein des cellules. Les résultats obtenus présentent la moyenne de 2 expériences indépendantes réalisées chacune en duplique. C) Résultats qPCR de la cinétique après transfction avec la mitoTALEN CD OUT ou sans (NT) lors de l'inhibition de la protéine ExoG. L'ADN est récupéré à 48h, 72h, 96h, 120h, 7 jours (J7) et à 15 jours (J15). Le rapport 12S/Actine correspond à la quantité d'ADN mitochondrial au sein des cellules. Les résultats obtenus présentent la moyenne de deux expériences indépendantes réalisées chacune en duplique.

Avant la transfection par la mitoTALEN CD OUT, les cellules sont préalablement transfectées par lipofection 24h auparavant avec les différents siARNs (siARN contrôle, siARN POLG, siARN ExoG). L'inhibition par les siARNs durant trois jours, une deuxième transfection avec les siARNs est réalisée afin de maintenir l'inhibition des protéines pendant sept jours. L'ADN est récupéré pendant sept jours consécutifs puis à quinze jours. Nous avons ensuite quantifié par qPCR la quantité d'ADN mitochondrial (Figure 27A).

Lors de l'inhibition de la protéine POLG, la transfection avec la mitoTALEN CD OUT montre une diminution de la quantité en ADN mitochondrial identique à la condition contrôle (siARN contrôle avec la mitoTALEN CD OUT). Cependant, il semble que la quantité d'ADN mitochondrial réaugmente plus tardivement, et qu'elle a lieu après sept jours une fois la levée de l'inhibition (courbe rose et orange, Figure 27B). La baisse d'ADN observée avec la condition siARN POLG sans mitoTALEN (NT) s'explique par l'inhibition de la réPLICATION provoquée par l'inhibition de POLG (courbe bleu et bleu ciel, Figure 27B). Ces expériences devront être reproduites afin de confirmer cet effet.

Les résultats obtenus lors de l'inhibition de la protéine ExoG avec la mitoTALEN CD OUT semblent identiques à ceux obtenus avec le siARN contrôle (courbe jaune et orange, Figure 27C). L'inhibition d'ExoG n'empêche donc pas la dégradation des molécules d'ADN mitochondrielles. Il est difficile de conclure sur les profils des conditions siARN ExoG et siARN contrôle sans mitoTALEN (NT). Les légères variations peuvent être dues à des variations expérimentales ou à une perturbation du nombre de copies d'ADN à cause de l'inhibition de la protéine ExoG. Les expériences devront elles-aussi être reproduites afin de confirmer que la protéine ExoG n'est pas impliquée dans la dégradation des molécules endommagées.

V. Augmenter la proportion de « Common Deletion » dans une lignée cellulaire humaine

Grâce aux mitoTALENs, nous avons pu induire la « Common Deletion » (cf résultats I). Nos valeurs en qPCR de la « Common Deletion » par rapport à l'ADN mitochondrial total (rapport CD/12S) est de l'ordre de 10^{-4} . Ces valeurs sont faibles comparées à la quantité de délétions retrouvées chez les malades (de l'ordre d'au moins 20%). Nous avons donc essayé d'augmenter la quantité de « Common Deletion ».

Pour ce faire, notre stratégie a été de procéder à des doubles transfections dans des cellules U20S. La première transfection est réalisée avec la mitoTALEN CD 5' afin d'induire la « Common Deletion » ou seulement avec le choc électrique de la nucléofection (NT). Puis sept jours après la première transfection, la seconde est réalisée soit pour la deuxième fois avec la mitoTALEN CD 5' soit avec la mitoTALEN CD IN qui coupe dans la région de la Common deletion, ou bien encore avec le choc électrique de la nucléofection (NT). Sept jours après la deuxième transfection, l'ADN est récupéré et la proportion de « Common Deletion » est quantifiée par qPCR.

Le principe de cette stratégie de double transfection est d'induire plus de molécules déletées en transfectant deux fois la mitoTALEN CD 5' ou après avoir induit la « Common Deletion » lors de la première transfection puis de couper spécifiquement les molécules non déletées grâce à la mitoTALEN CD IN. L'utilisation de la mitoTALEN CD IN repose sur l'observation que la coupure double-brin par les mitoTALENs entraîne une dégradation d'une partie de l'ADN mitochondrial (Figure 26).

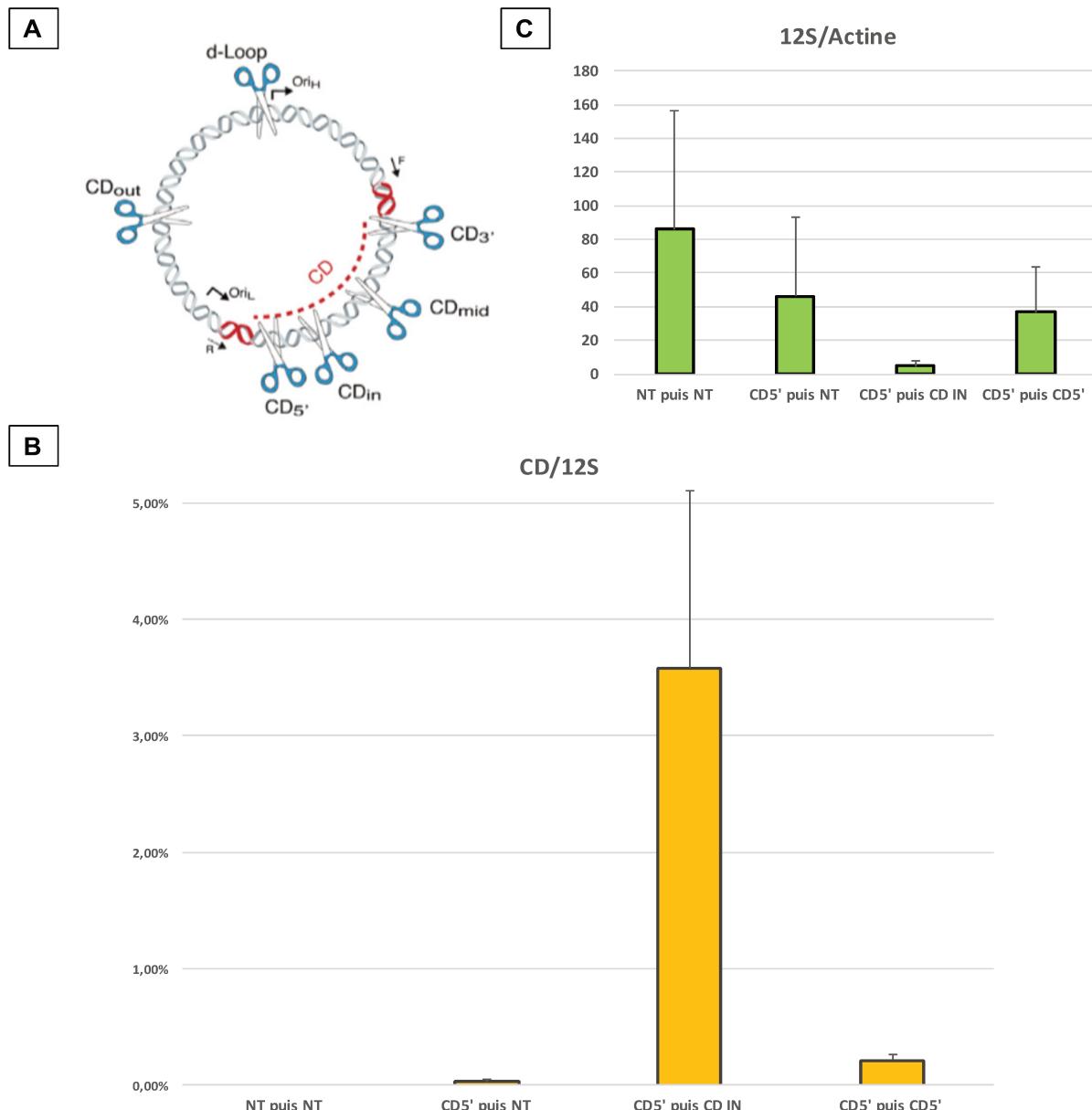


Figure 28 : Enrichissement en molécules déletées après des doubles transfections par les mitoTALENs dans des U2OS

A) Schéma de l'ADN mitochondrial avec les sites de coupure des mitoTALENs : dans cette expérience, nous utilisons les mitoTALENs CD 5' et CD IN B) Résultats qPCR de la quantité de « Common Deletion » formée après double transfection avec la mitoTALEN CD 5' puis la mitoTALEN CD 5' ou CD IN. NT signifie que les cellules ont uniquement reçu le choc électrique de la nucléofection sans plasmide. Les résultats présentent le rapport CD/12S (« Common Deletion »/ADN mitochondrial), qui représentent la quantité de délétions au sein de l'ADN mitochondrial. L'ADN est récupéré 7 jours après la seconde transfection, soit 15 jours après le début de l'expérience. C) Résultats qPCR de la quantité d'ADN mitochondrial après double transfection avec la mitoTALEN CD 5' puis la mitoTALEN CD 5' ou CD IN. NT signifie que les cellules ont uniquement reçu le choc électrique de la nucléofection sans plasmide. Le rapport ADN mitochondrial (12S)/ADN nucléaire (Actine) correspond à la quantité d'ADN mitochondrial au sein des cellules. L'ADN est récupéré 7 jours après la seconde transfection, soit 15 jours après le début de l'expérience. Les résultats présentés sont la moyenne de deux expériences indépendantes réalisées chacune en duplicat.

Nos résultats montrent que la double transfection permet une augmentation significative en « Common Deletion ». De façon intéressante, c'est la stratégie avec la mitoTALEN CD IN utilisée lors de la seconde transfection qui montre la plus grande augmentation en molécules délétées. Nous avons ainsi réussi à obtenir des cellules qui possèdent environ 5% de molécules d'ADN mitochondrial possédant la « Common Deletion » (Figure 28). On observe aussi une diminution importante de la quantité d'ADN mitochondrial après la double transfection avec CD5' et CD IN reflétant l'efficacité de CD IN à dégrader l'ADN mitochondrial intact. Nous avons tenté de transfecter une nouvelle fois les cellules afin d'atteindre un pourcentage encore plus élevé de « Common Deletion » mais cette stratégie n'a pas fonctionné, la « Common Deletion » a été perdue par les cellules lors de cette nouvelle transfection. Il serait peut-être intéressant de laisser plus de temps avant de réaliser une troisième transfection pour laisser les cellules restaurer leur quantité d'ADN mitochondrial et de peut-être mieux tolérer la délétion.

VI. Adaptation du système CRISPR/Cas9 pour la mitochondrie

Nous avons réussi à induire la « Common Deletion » avec des mitoTALENs. Quelques années après les TALENs, le système de nucléase artificielle CRISPR/Cas9 est apparu. Nous avons donc cherché à adapter ce système pour cibler l'ADN mitochondrial et induire la « Common Deletion ».

Contrairement aux TALENs, ce système est plus flexible concernant le choix de la séquence cible. En effet, la séquence cible des TALENs doit avoir un T en 5' reconnue par la première répétition de chaque domaine TALE (Figure 21) et l'assemblage des monomères des TALENs est plus complexe que la synthèse des gARNs. Un autre avantage du système CRISPR/Cas9 est de pouvoir construire plusieurs gARNs sur une séquence cible assez courte (une centaine de bp) car le PAM est en général assez souvent retrouvé. Nous avons utilisé la Cas9 *Neisseria meningitidis* dont le PAM est NNNNGATT, car cette protéine était plus petite que la Cas9 *Streptococcus pyogenes* donc potentiellement plus facile à importer. Cependant, la difficulté de ce système est d'importer deux acteurs de nature différente, le gARN et la mitoCas9 au sein de la mitochondrie.

Pour adresser la protéine Cas9 à la mitochondrie, nous avons utilisé les mêmes signaux d'adressage que pour les mitoTALENs. Concernant l'import du gARN, la voie d'import des ARNs étant encore mal connue, nous avons utilisé des séquences identifiées comme permettant

l'import d'ARN dans la mitochondrie (Wang et al. 2012). Nous avons donc ajouté au gARN, une loop en 5' identifiée sur l'ARN de la RNase P et une séquence en 3' provenant de l'ARN de la MRP (Figure 29).



Figure 29 : Schéma représentant les ARN modifiés afin de permettre leur import dans la mitochondrie

Une loop en 5' (violet), identifiée sur l'ARN de la RNase P et une séquence en 3' (bleue), identifiée sur l'ARN de la MRP sont ajoutée au gARN (en rouge) pour permettre son import à la mitochondrie (Wang et al., 2012).

Nous avons transfecté les plasmides codant pour la protéine mitoCas9 (2,5 µg) et pour le gARN (5 µg) par nucléofection. Malgré de nombreux essais avec des conditions différentes (transfection du gARN et de la mitoCas9 à 24h d'écart, transfaction en lipofectamine), nous n'avons pas réussi à détecter la « Common Deletion » par PCR (Figure 30).

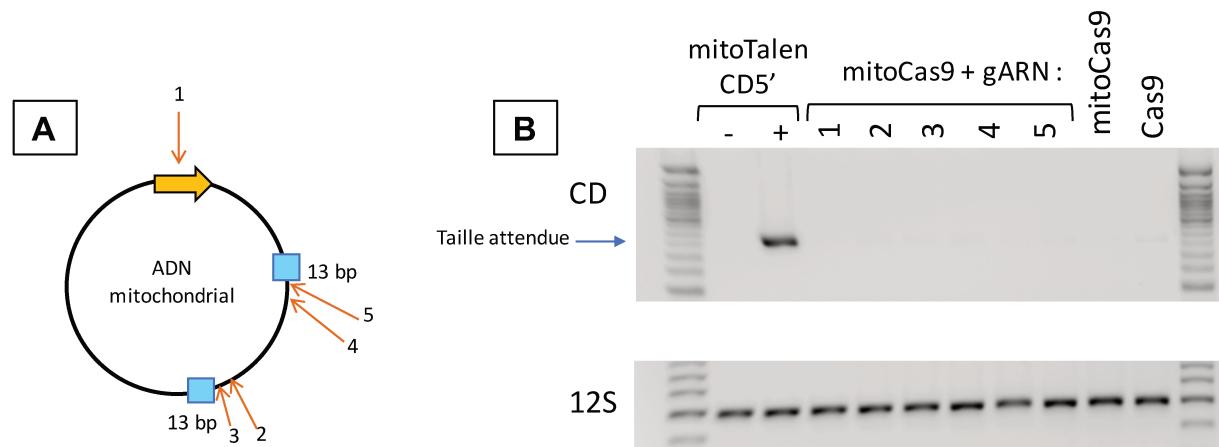


Figure 30 : Test d'induction de la « Common Deletion » par la mitoCas9

A) Schéma représentant l'ADN mitochondrial. La flèche jaune représente la D-loop, les carrés bleus représentent les séquences de 13 bp qui encadrent la « Common Deletion ». Les flèches oranges représentent l'endroit du génome mitochondrial ciblé par les gARNs. B) CD représente les résultats de la PCR nichée pour la détection de la « Common Deletion ». mitoCas9 correspond à la transfaction uniquement avec la mitoCas9 sans gARN, Cas9 correspond à la transfaction uniquement avec la Cas9 (version nucléaire) sans gARN. 12S correspond au standard interne de PCR.

Afin de mieux comprendre pourquoi nous n'arrivions pas à détecter la formation de la « Common Deletion », nous avons réalisé des expériences *in vitro* pour vérifier que les signaux

d'adressage ajoutés au gARN n'empêchaient pas la reconnaissance et/ou la coupure par la mitoCas9. Pour réaliser ces expériences *in vitro*, nous avons incubé à 37°C la Cas9 purifiée, les gARNs avec ou sans les signaux d'adressage (obtenus par transcription *in vitro*), et la cible ADN préalablement amplifiée par PCR. Après incubation, les échantillons sont analysés par électrophorèse sur gel BET. Nous avons ainsi pu montrer que la loop en 5' et la séquence en 3'UTR n'empêchent pas la coupure de la cible par la Cas9 *in vitro*, car les pourcentages de coupure sont assez similaires selon les différents gARNs (Figure 31).

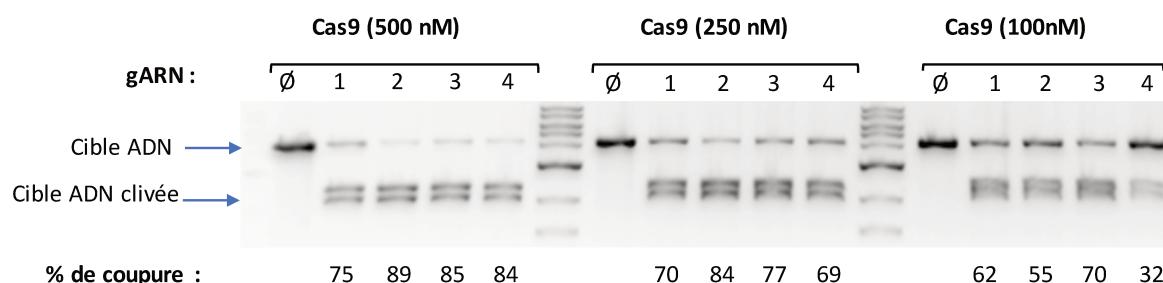


Figure 31 : Test de clivage *in vitro* avec la Cas9 et les gARNs avec ou sans signaux d'adressage à la mitochondrie

Ø correspond à la Cas9 seule sans gARN, 1 correspond à un gARN avec une loop en 5', 2 correspond à un gARN sans signaux d'adressage pour la mitochondrie, 3 correspond à un gARN avec une loop en 5' et la séquence en 3'UTR, 4 correspond à un gARN avec la séquence en 3'UTR. La Cas9 est incubée à différentes concentrations pendant 30 minutes à 37°C avec les différents gARNs. La visualisation des produits de coupure est faite par électrophorèse sur gel BET.

Nous avons ensuite cherché à savoir si dans les cellules, la mitoCas9 et le gARN étaient efficacement adressés à la mitochondrie après transfection. Nous avons donc réalisé des expériences d'immunofluorescences pour regarder la localisation de la mitoCas9 48h après transfection. Nos résultats ont montré que de façon surprenante, la mitoCas9 complexée avec un gARNmito montrait un signal diffus au sein de la cellule, non spécifique de la mitochondrie (Figure 32A). Afin de mieux comprendre cette observation, nous avons transfété la mitoCas9 avec un gARN nucléaire. La mitoCas9 est alors bien visualisée dans la mitochondrie lors de la transfection avec un gARN nucléaire (Figure 32B). Le dernier contrôle a été de transférer la Cas9 nucléaire avec le gARNmito. Dans ces conditions, la Cas9 nucléaire est bien présente dans le noyau, le gARNmito ne perturbe pas la localisation de la Cas9 nucléaire (Figure 32C). Les transfections simultanées ou à 24h d'intervalle des plasmides ont montré des résultats identiques. Ces résultats suggèrent qu'il y a un problème d'import dans la mitochondrie de la protéine mitoCas9 en présence d'un gARNmito, cela semble bloquer l'import de la mitoCas9.

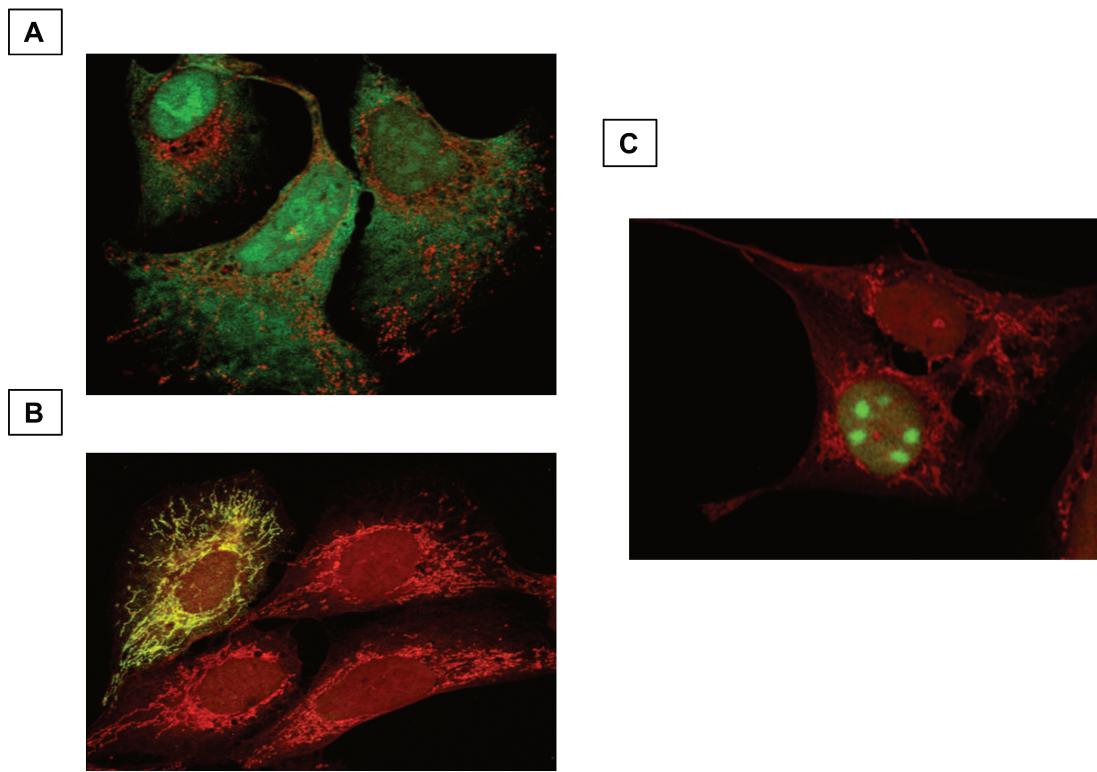


Figure 32 : Images d'expériences immunofluorescence après transfection avec la Cas9 nucléaire ou mitochondriale en présence d'un gARN mitochondrial ou nucléaire dans des U2OS

A) Visualisation de la mitoCas9 (en vert) 48h après transfection avec la mitoCas9 et un gARNmito. B) Visualisation de la mitoCas9 (en vert) 48h après transfection avec la mitoCas9 et un gARN nucléaire. C) Visualisation de la Cas9 (en vert) 48h après transfection avec Cas9 et un gARNmito. La Cas9 et la mitoCas9 sont observées en vert grâce à un anticorps secondaire anti HA couplé à l'Alexa Fluor A488, les mitochondries marquées au mitotracker sont visualisées en rouge.

Afin de vérifier l'import du gARNmito, nous avons produit les gARNmito *in vitro* (marqué avec de l'UTP-Cy5), avec les deux signaux d'adressage, avec un seul ou sans signal d'adressage par transcription (Figure 33).

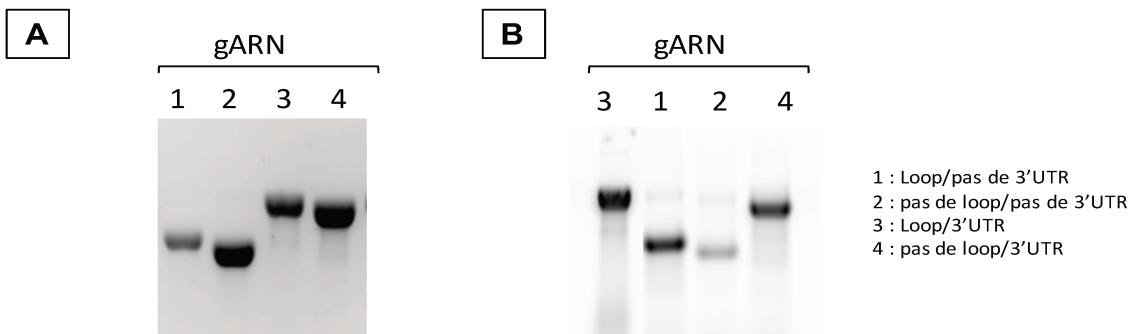


Figure 33 : Résultats de la synthèse des gARNs fluorescents par transcription *in vitro*

A) Visualisation des gARNs marqués avec de l'UTP-Cy5 sur gel BET après transcription *in vitro*. B) Visualisation par fluorescence des gARNs marqués avec de l'UTP-Cy5. 1- correspond à un gARN avec une loop en 5', 2- correspond à un gARN sans signaux d'adressage pour la mitochondrie, 3- correspond à un gARN avec une loop en 5' et la séquence en 3'UTR, 4- correspond à un gARN avec la séquence en 3'UTR.

Nous avons ensuite transfété ces gARNs en lipofectamine et 48h après transfection, la localisation des gARNs marqués est observée en microscopie. Les résultats montrent que quelque soit le gARN, ils sont tous retrouvés dans des vésicules et ne semblent pas être présents en quantité suffisante pour être détectés dans la mitochondrie (Figure 34).

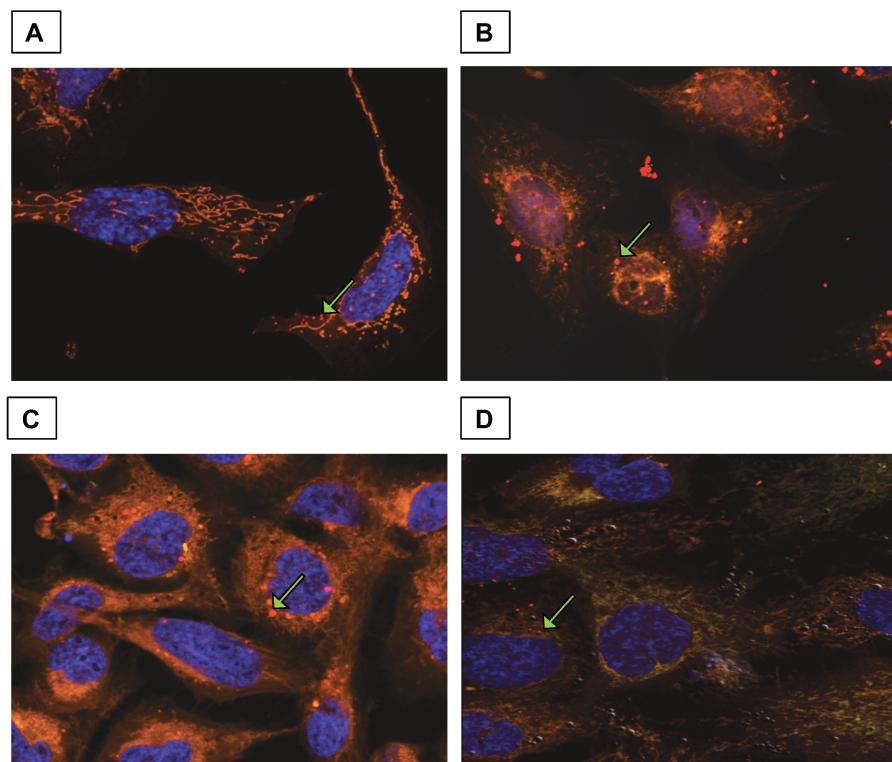


Figure 34 : Localisation des gARNs avec ou sans signaux d'adressage pour la mitochondrie dans des U2OS

A) gARN avec une loop en 5' (en rouge) B) gARN sans signaux d'adressage pour la mitochondrie (en rouge) C) gARN avec une loop en 5' et la séquence en 3'UTR (en rouge) D) gARN avec la séquence en 3'UTR (en rouge). Les mitochondries sont visualisées grâce à du mitotracker (en orange), les noyaux sont visualisés grâce à du Hoechst (bleu). Les flèches vertes montrent un exemple de la présence des gARNs.

Nos différentes conditions expérimentales n'ont pas permis d'observer l'induction de la « Common Deletion » par la mitoCas9. Nos différents tests semblent montrer que le problème est certainement l'adressage du gARN à la mitochondrie. Le système d'import des gARNs nécessite d'être optimisé.

Discussion-Perspectives

Ma thèse a porté sur l'étude de la formation d'une délétion, appelée « Common Deletion » la plus fréquemment retrouvée dans les maladies associées à l'ADN mitochondrial. Nous avons pour cela utilisé des TALENs, appelées mitoTALENs, adaptées pour être exprimées au sein de la mitochondrie afin d'induire des cassures double-brin sur l'ADN mitochondrial. Nous avons vérifié leur adressage en réalisant des fractionnements cellulaires et par immunofluorescence. Nous avons ensuite cherché si la « Common Deletion » était induite par la coupure de certaines mitoTALENs. Il s'est avéré que seule la mitoTALEN coupant proche de la répétition située du côté de l' O_L induit la « Common Deletion » dans des lignées cellulaires humaines. Nous avons alors montré que la formation de la « Common Deletion » est liée à une mauvaise réparation lors de la réPLICATION et ainsi identifié les protéines du réplisome minimal (POLG, Twinkle, mtSSB) ainsi que l'exonucléase MGME1, la DNA LIGASE 3 et les hélicases DNA2 et PIF-1 comme étant impliquées dans la formation de la « Common Deletion ». Ces protéines sont connues à l'heure actuelle pour participer à la réPLICATION mais aussi pour leur implication dans la voie de réparation par excision de bases (BER) pour les protéines POLG et DNA2.

Ce projet mené en collaboration avec le laboratoire d'Agnel Sfeir a aussi permis d'apporter des informations sur le mode de réPLICATION utilisé lors de la réPLICATION de l'ADN mitochondrial.

Le peignage moléculaire adapté à la mitochondrie confirme le modèle de réPLICATION par déplacement de brin

Nos collaborateurs ont mis au point la technique de peignage de l'ADN sur l'ADN mitochondrial, technique qui jusque-là était utilisée pour étudier la réPLICATION dans le noyau. Comme présenté dans l'introduction de ce manuscrit, les modèles de réPLICATION pour l'ADN mitochondrial sont encore débattus. Les deux grands modèles, celui du déplacement de brin et le modèle synchrone ont été étudiés majoritairement par des expériences d'électrophorèse 2D. Cette technique s'est avérée présenter un inconvénient : la purification préalable de l'ADN mitochondrial, qui peut modifier les structures des intermédiaires de réPLICATION (Yang et al. 2002). Nous avons par conséquent opté pour la méthode de peignage moléculaire ne comportant

pas cet aléa. Elle possède l'avantage de marquer l'ADN directement en court de la réPLICATION dans les cellules pour ensuite l'analyser par microscopie sous forme de fibres. L'utilisation de cette technique a permis de montrer l'utilisation prépondérante du modèle de déplacement brin pour la réPLICATION mitochondriale dans un certain nombre de lignées cellulaires (HT1080, MEF, RPE1, BJ hTERT, C2C12 myoblastes et myotubes). En effet, d'une part le profil de peignage moléculaire obtenu suggère l'utilisation des deux origines de réPLICATION (O_H et O_L) et d'autre part, les résultats montrent un rôle essentiel de la protéine POLMRT au niveau de la seconde origine de réPLICATION. Ces différentes observations s'accordent parfaitement avec les différents travaux publiés sur le modèle de déplacement de brins (Chang & Clayton 1985) (Tapper & Clayton 1981). L'étude de la vitesse de réPLICATION a elle-aussi été en accord avec les données publiées de 270 nucléotides/min (Clayton 1982), puisque nos collaborateurs ont calculé une vitesse de 220 nucléotides/min. La méthode de peignage moléculaire s'est donc révélée être un bon outil pour étudier l'ADN mitochondrial. Nos résultats ne permettent cependant pas de conclure sur le troisième modèle, intermédiaire aux précédents, appelé le modèle « RITOLS ». En effet, dans ce modèle, le brin retardé serait protégé par des ARNs et non par la protéine mtSSB. Notre outil possède donc une limite car il ne permet pas de discriminer entre ces deux alternatives car seule l'incorporation des nucléotides marqués est observée.

Le peignage moléculaire a aussi permis d'étudier l'effet d'un mutant de Twinkle, Twinkle R374Q notamment impliqué dans l'ophtalmoplégie externe progressive (Martin-Negrier et al. 2011). Nous avons ainsi pu observer que l'expression de ce mutant dans les cellules entraînait des pauses de la réPLICATION dans la région de la « Common Deletion ». Ces pauses dans la réPLICATION semblent faciliter la formation de délétions en provoquant des sauts de réPLICATION. Le peignage moléculaire pourra donc servir d'outil pour mieux comprendre les conséquences de protéines mutées sur la réPLICATION de l'ADN mitochondrial.

Enfin nous aurions voulu analyser directement la formation de « Common Deletion » par peignage moléculaire après son induction avec les mito-TALENs mais cela n'a pas été possible car le taux de délétion dans les cellules traitées est trop bas pour observer un nombre significatif de molécules délétées.

Y-a-t-il un système de réparation efficace lors de cassures double-brin dans la mitochondrie ?

Dans cette étude, nous avons étudié la formation de la délétion la plus fréquemment retrouvée au sein de l'ADN mitochondrial pour tenter de comprendre pourquoi et comment de si longues délétions peuvent se former. En partant de l'hypothèse que les délétions soient la conséquence de cassures double-brin, nous avions envisagé l'implication possible des mêmes voies de réparation que celles actives dans le noyau : le NHEJ, le alt-NHEJ ou la recombinaison homologue (HR). Les délétions au sein de l'ADN mitochondrial présentent la caractéristique d'être bordées de courtes séquences répétées dans 85% des cas (<http://www.mitomap.org/>) (Eshaghian et al. 2006). Une fréquence aussi élevée de séquences répétées entourant les délétions suggère l'utilisation de la voie alt-NHEJ. En effet, dans le noyau, elle est la voie de réparation qui s'appuie sur la présence de courtes séquences homologues avoisinant les cassures double-brin pour les réparer. Par ailleurs, la présence de plusieurs copies d'ADN mitochondrial proches les unes des autres évoque la possible action de la recombinaison homologue pour réparer les cassures, les copies intactes pouvant servir de matrice de réparation.

Nous avons donc testé le rôle des protéines CtIP (alt-NHEJ et HR), MRE11 (alt-NHEJ et HR), RAD51 (HR), RAD52 (HR), PARP-1 (alt-NHEJ) et de la DNA LIGASE 3 (alt-NHEJ). L'inhibition des protéines CtIP, MRE11, RAD51, RAD52 et PARP-1 n'a pas montré d'effet sur la formation de la délétion, ce qui suggère que le NHEJ, le alt-NHEJ et la recombinaison homologue ne participent pas directement à la formation des délétions dans l'ADN mitochondrial. Concernant la DNA LIGASE 3, son KO entraîne une diminution de la formation de la délétion, ce qui valide le rôle de cette protéine dans la formation de la « Common Deletion ». Cependant, comme la DNA LIGASE 3 est la seule ligase présente dans la mitochondrie, elle est aussi directement impliquée dans la réPLICATION de l'ADN mitochondrial. De plus, l'absence de la forme mitochondriale de la DNA LIGASE 3 induit en effet la mort cellulaire probablement par perte des copies d'ADN mitochondrial (Simsek et al. 2011).

Nos résultats s'opposent aux résultats publiés récemment montrant l'activité d'un mécanisme alt-NHEJ au sein de la mitochondrie (Tadi et al. 2016). En effet, nous ne voyons pas d'effet direct des protéines du alt-NHEJ sur la formation de la « Common Deletion » et seule une de nos mito-TALENs (la CD 5') induit la formation de la délétion. En supposant que le alt-NHEJ soit actif au sein de la mitochondrie, toutes les mitoTALENs coupant dans la délétion devraient pouvoir entraîner la formation de la délétion lors de la réparation de la cassure après résection et reconnaissance des séquences de 13 bp, en particulier celles qui coupent proches des extrémités. Ainsi, l'analyse des séquences au niveau du site de coupure

des différentes mitoTALENs par séquençage haut débit et test T7 endonucléase ne montre pas de mutations (insertions ou délétions), ce qui s'oppose à une réparation par alt-NHEJ.

La différence de résultat peut s'expliquer par le fait que nous avons utilisé une stratégie très différente de celle utilisée par l'équipe de Raghavan (Tadi et al. 2016). En effet, dans leur étude, les auteurs utilisent des extraits mitochondriaux et testent leur activité sur différents types de substrats *in vitro*. Les capacités de réparation de certaines protéines présentes dans les extraits mimant un mécanisme de alt-NHEJ ne signifient pas forcément que la cellule utilise cette voie de réparation.

Nous n'avons donc pas constaté l'utilisation du NHEJ, du alt-NHEJ ou de la recombinaison homologue dans la formation de la « Common Deletion » tels qu'ils sont décrits dans le noyau mais il serait intéressant de réaliser la même étude avec une autre délétion présentant des séquences répétées d'une taille différente pour confirmer nos résultats.

Il a été publié que lors de dommages, telles que les cassures double-brin, cela entraîne une dégradation de l'ADN mitochondrial (Kukat et al. 2008) (Moretton et al. 2017). L'utilisation des différentes mitoTALENs qui induisent des cassures double-brin nous ont permis de faire la même observation. Nous avons pu quantifier par qPCR une baisse significative en molécules d'ADN mitochondrielles après transfection par les mitoTALENs. Cet effet est spécifique à une cassure double-brin car la mitoTALEN CD 5' Dead (domaine Fok inactivé, ce qui empêche la coupure) n'induit pas de diminution en quantité d'ADN mitochondrial. Les cellules mettent ensuite cinq jours à récupérer leur quantité d'ADN mitochondrial initiale (Figure 26). Il est donc concevable que la prise en charge des cassures double-brin soit régie par la mitochondrie en dégradant les molécules cassées puis en restaurant sa quantité d'ADN via la réPLICATION.

Nous avons cherché à identifier des protéines responsables de ce mécanisme de dégradation. Nous avons choisi de tester le rôle de l'exonucléase ExoG, car sa processivité semblait plus compatible avec la capacité de dégrader entièrement les molécules d'ADN endommagées comparée à la protéine MGME1. Nous avons donc induit une inhibition durant sept jours de cette protéine et réalisé en parallèle des cassures double-brin par une de nos mitoTALENs (la CD OUT). Nos résultats ne nous ont pas permis d'observer de différences significatives avec les conditions contrôles. Une étude récemment publiée n'identifie pas non plus la protéine ExoG comme étant responsable de la dégradation des molécules endommagées (Moretton et al. 2017). De façon surprenante, cette étude qui teste les 5 exonucléases connues

à ce jour (ExoG, MGME1, EndoG, DNA2, FEN1) ne détecte pas leurs implications sur la dégradation de l'ADN mitochondrial. Est-il possible que nous ne connaissons pas encore toutes les exonucléases présentes à la mitochondrie ou que nous n'ayons pas encore identifié une fonction exonucléase d'une protéine connue de la mitochondrie ?

L'absence de certaines protéines de réparation signifie-t-elle l'absence de certaines voies de réparation ?

Dans la mitochondrie, la présence de nombreuses protéines de réparation participant à différentes voies n'a pas été validée. Il en est ainsi pour les protéines de la voie du MMR, NER, NHEJ, alt-NHEJ, ou la recombinaison homologue. Il est actuellement considéré que la mitochondrie ne répare pas via la voie du MMR en raison de l'absence d'identification des protéines clées telles qu'elles sont décrites dans la voie MMR, présente dans le noyau. Concernant les voies du NHEJ, du alt-NHEJ, ou de la recombinaison homologue, la présence des protéines n'est pas clairement établie dans la mitochondrie. Des études concluent par exemple à la présence de XRCC4 (Wisnovsky et al. 2016) et d'une forme de KU tronqué en C terminal (Coffey & Campbell 2000) à la mitochondrie. La présence de XRCC4 sans celle de la DNA LIGASE 4 est cependant étrange car aucun rôle n'a été identifié jusqu'à présent pour XRCC4 non complexé à la DNA LIGASE 4.

On peut alors se demander s'il existe un biais dans la façon d'identifier une voie de réparation propre à la mitochondrie. En effet, nous cherchons à identifier les mêmes protéines et les mêmes mécanismes que ceux actuellement connus dans le noyau. Cette stratégie est peut-être une erreur. Les mécanismes diffèrent peut-être légèrement ou ils sont réalisés par d'autres protéines. De plus, le nombre de protéines présentes dans la mitochondrie étant réduit comparé au noyau, des protéines déjà identifiées pourraient jouer plusieurs rôles. POLG, DNA2 assurent déjà des fonctions au sein de la réPLICATION et du BER, on peut imaginer qu'il en est de même pour d'autres protéines.

De quelle façon une cassure double-brin est-elle détectée au sein de l'ADN mitochondrial ?

Aussi, au sein de la mitochondrie nous ignorons également comment se déroule la prise en charge des cassures double-brin. Dans le noyau, ces dommages sont reconnus et activent des voies de signalisation qui permettront la réparation. La reconnaissance et la prise en charge des cassures dans le noyau mettent en œuvre le remodelage de la chromatine. Dans la mitochondrie,

la reconnaissance des cassures est forcément différente car l'ADN mitochondrial n'est pas entouré d'histones. Il est pour autant probable qu'il existe un système de reconnaissance des cassures. Toute cette signalisation est donc à l'heure actuelle complètement inconnue. Néanmoins, la mitochondrie reconnaît a priori rapidement les lésions, car comme en témoigne une étude, 30 minutes après induction de dommages par un traitement à l' H_2O_2 , des molécules linéaires sont déjà identifiées (Shokolenko et al. 2009).

Pourquoi seule la mitoTALEN CD 5' coupant près de la répétition proche de l' O_L induit-elle la « Common Deletion » ?

Nous avons construit différents couples de mitoTALENs dont plusieurs coupent près des répétitions de 13 bp qui entourent la « Common Deletion ». Pourtant, seules la mitoTALEN CD 5' et son équivalent en TALEN nickase qui coupe uniquement le brin H, provoquent la formation de la « Common Deletion ». Est-il nécessaire que la réPLICATION soit suffisamment avancée le long de la molécule pour qu'il y ait mésappariement des séquences de 13 bp et que la délétion se forme (Figure S1 de l'article) ? La première répétition de 13 bp étant proche de l' O_H , la cellule décide-t-elle de stopper la réPLICATION lors d'une cassure et de dégrader les molécules lorsque la réPLICATION est seulement à peine initiée ? Ceci pourrait expliquer pourquoi uniquement la cassure proche de la répétition 5' et non celle en 3' peut permettre la formation de la délétion (Figure 4A de l'article). L'avancement de la réPLICATION peut en effet être un argument car nos résultats montrent le rôle des protéines de réPLICATION dans la formation de la délétion.

Malgré cette idée, des mitoTALENs très proches de la mitoTALEN CD 5' n'induisent pas la délétion, ce qui suggère qu'il y a vraiment une particularité au niveau de la séquence reconnue par la mitoTALEN CD 5'. La topologie de l'ADN mitochondrial joue peut-être un rôle, une structure particulière (courbure) ayant été relevée *in vitro* (Hou & Wei 1998). Les G-quadruplexes pourraient aussi être des candidats de structure particulière qui favorisent la formation de délétions. En effet, l'inhibition de la protéine PIF-1, hélicase qui peut agir sur les G-quadruplexes, entraîne une augmentation de la formation de la « Common Deletion ».

De plus, nous savons que l'ADN mitochondrial ne possède pas de chromatine et se retrouve au niveau de structures nucléoïdes regroupant plusieurs copies de cet ADN très proches dans l'espace. Cependant nous ne disposons que de peu d'informations sur l'exacte organisation spatiale des molécules. En effet, on peut envisager l'existence de certains

rapprochements de séquences intra ou intermoléculaires susceptibles de favoriser et/ou stabiliser des structures particulières. Ces dernières pourraient faciliter l'appariement des séquences répétées 5' et 3' de 13 bp, induisant alors un saut de réPLICATION avec pour conséquence la formation de la « Common Deletion ». L'absence de chromatine au sein de l'ADN mitochondrial pourrait influencer l'apparition des délétions. En effet, l'ADN mitochondrial n'étant pas protégé par des histones certaines régions sont peut-être plus exposées ou mieux protégées en cas de dommages. L'hypothèse suivante pourrait être avancée : la région ciblée par la mitoTALEN CD 5' pourrait mieux tolérer la cassure et ne pas être immédiatement réparée ou dégradée, ce qui laisserait la possibilité d'induire un saut de réPLICATION. Ainsi serait favorisée la formation de la délétion, contrairement aux autres séquences ciblées par les mitoTALENs. Selon la topologie d'une molécule d'ADN mitochondrial et de l'organisation des molécules d'ADN en nucléoïdes, des dommages dans certaines régions pourraient être plus facilement détectés/réparés contrairement à d'autres. Pour vérifier cette hypothèse, il est impératif de mieux appréhender les mécanismes de détection et de réparation des dommages à la mitochondrie et la topologie de son ADN.

Pourquoi les taux de « Common Deletion » induite par la mitoTALEN CD 5' restent-ils faibles comparés à ceux des patients ?

Nos résultats ont permis d'induire la « Common Deletion » dans différentes lignées cellulaires humaines. Cependant nos données de quantification de la délétion par qPCR révèlent une fréquence d'environ 10^{-4} à 10^{-3} (« Common Deletion »/ADN mitochondrial (CD/12S)). Ces valeurs sont très faibles en comparaison des valeurs d'hétéroplasmie retrouvées chez les patients, les symptômes apparaissant à partir de 20% d'hétéroplasmie (Rossignol et al. 2003). Des cellules transfectées avec la mitoTALEN CD 5' ont été gardées en culture durant un mois. Nous avons alors constaté que la fréquence de la « Common Deletion » augmentait au cours des premières semaines puis les cellules conservaient leurs molécules délétées à un taux constant (Figure 35). Ces résultats s'opposent à l'idée de plus en plus controversée que les molécules délétées possèdent un avantage réPLICATIF. En effet, si ces molécules de plus petites tailles possèdent un avantage réPLICATIF nous aurions dû constater une augmentation constante. Dans l'hypothèse où ces molécules présenteraient bien un avantage réPLICATIF, un mécanisme de régulation négative doit alors exister pour contrecarrer cet effet, ce qui permettrait à la cellule de tolérer une certaine quantité de délétions et de maintenir une valeur constante en « Common Deletion ».

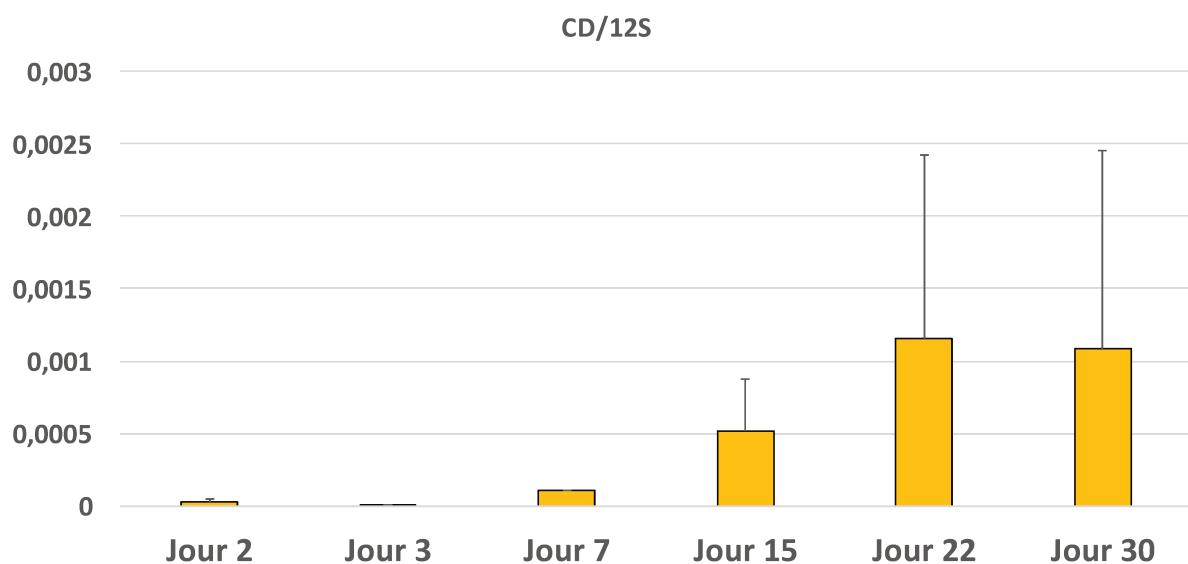


Figure 35 : Quantité de « Common Deletion » au cours du temps après coupure par la mitoTALEN CD 5' dans des U2OS

La quantité de « Common Deletion » est quantifiée par qPCR pendant un mois après transfection avec la mitoTALEN CD 5'. Le rapport CD sur 12S signifie «Common Deletion»/ADN mitochondrial, ce qui représente la quantité de « Common Deletion » au sein de l'ADN mitochondrial.

L'équipe de Nebel (von Wurmb-Schwark et al. 2010) et celle de Turnbull (Campbell et al. 2014) par exemple ne semblent pas d'accord avec l'idée que les molécules délétées présentent un avantage réplicatif. De fait, l'accumulation de molécules délétées au cours du vieillissement n'a pas été vérifiée au sein d'une cohorte de patients (von Wurmb-Schwark et al. 2010). Par ailleurs, l'équipe de Turnbull n'observe pas un envahissement plus rapide dans le cas de molécules avec des grandes délétions comparé à la vitesse d'envahissement des molécules avec de petites délétions au sein de fibres musculaires. Enfin, l'étude réalisée avec des souris transgéniques mutées sur POLG et soumises à un exercice physique apporte des résultats d'un réel intérêt (Safdar et al., 2011). En effet, ces souris diminuent leur taux de molécules mutées et restaurent une quantité d'ADN mitochondrial équivalente à celle des souris WT. Ces résultats renforcent l'idée d'un mécanisme de régulation, puisque chez ces souris, il semble que les cellules aient la capacité d'opter pour une réamplification des molécules WT afin qu'elles puissent répondre aux besoins énergétiques exigés par l'activité physique malgré une polymérase POLG défective.

Nos taux de « Common Deletion » étant relativement bas comparés à ceux observés chez les patients, nous avons cherché à augmenter le degré d'hétéoplasmie. Nos observations

sur la dégradation des molécules d'ADN mitochondrial après coupure double-brin nous ont conduit à élaborer une stratégie de double transfection. Nous avons donc après une première transfection avec la mitoTALEN CD 5' pour induire la « Common Deletion » soit transfété à nouveau avec la mitoTALEN CD 5' afin de former la délétion sur les molécules encore intactes, soit transfété avec la mitoTALEN CD IN afin d'éliminer les molécules WT. De façon intéressante, la double transfection qui donne les meilleurs résultats est celle effectuée avec les mitoTALENs CD 5' puis CD IN. Nous avons ainsi pu obtenir des taux d'environ 5% de « Common Deletion ». Pour continuer d'augmenter l'hétéroplasmie, nous avons tenté de transfecter une troisième fois avec la mitoTALEN CD IN. Une semaine après transfection, nous avons quantifié la proportion de « Common Deletion ». Nos résultats de qPCR se sont révélés négatifs pour la détection de la délétion. Cette troisième transfection a donc entraîné la perte des copies délétées. Bien que les trois transfections aient été réalisées à des intervalles de 7 jours, il est peu probable que la délétion soit perdue au cours du temps. De fait, nos précédentes expériences de cinétique avaient révélé que la délétion était maintenue au cours du temps et que les cellules doublement transfectées avaient elles aussi conservé leur degré d'hétéroplasmie (Figure 35). Une autre stratégie réalisée par nos collaborateurs a été de diminuer la quantité d'ADN mitochondrial avec un traitement au BET. Lorsque la quantité d'ADN mitochondrial avait diminué drastiquement, la « Common Deletion » était induite par la mitoTALEN CD 5'. Le but était d'obtenir la délétion sur la faible quantité d'ADN mitochondrial restante, puis que les molécules délétées soient répliquées lors de la levée du traitement au BET. Les résultats n'ont pas montré d'augmentation significative en « Common Deletion », les molécules WT sont redevenues majoritaires. L'ensemble de ces résultats suggèrent la présence d'un mécanisme de régulation des délétions par lequel les cellules semblent tolérer jusqu'à un certain seuil la présence de copies délétées puis les éliminent.

Une question reste donc inexorable, que se passe-t-il alors chez les patients présentant des degrés d'hétéroplasmie beaucoup plus important que ceux obtenus par nos méthodes ? Certes, des protéines mutées telles que POLG et Twinkle, peuvent entraîner la formation de délétions ou de mutations, cependant, la mitochondrie semble posséder des acteurs régulant l'hétéroplasmie encore méconnus à ce jour. Ces mécanismes paraissent ne plus fonctionner chez ces patients, permettant ainsi l'émergence des copies endommagées et le début des symptômes. La connaissance de ces mécanismes constitue un enjeu majeur pour la recherche de thérapies adaptées aux maladies liées à des mutations/délétions de l'ADN mitochondrial.

Pourquoi le système CRISPR n'induit-il pas la délétion ?

Nous avons tout d'abord utilisé les mitoTALENs pour cibler l'ADN mitochondrial. Peu de temps après les TALENs, le système de CRISPR/Cas9 est apparu. Une partie de ma thèse a été consacrée à tenter d'adapter le système CRISPR/Cas9 à la mitochondrie. Nous avons donc tenté d'induire la « Common Deletion » avec ce nouveau système. La difficulté de celui-ci était d'importer au sein de la mitochondrie à la fois la protéine et l'ARN.

Nos différentes expériences ne nous ont pas permis d'induire la « Common Deletion » grâce à ce système. Nous avons donc essayé de mieux comprendre pourquoi il ne semblait pas fonctionner dans nos conditions. Notre hypothèse a été de vérifier si les deux acteurs, la mitoCas9 et le gARN étaient importés assez efficacement au sein de la mitochondrie pour obtenir un taux de coupure suffisant pour induire la « Common Deletion ». Nos résultats nous ont permis de conclure que le gARNmito perturbait l'adressage dans la mitochondrie de la mitoCas9. De plus, nos expériences de localisation du gARNmito ne nous ont pas permis de détecter avec certitude la présence de ce dernier dans la mitochondrie.

La perturbation de l'import de la mitoCas9 en présence du gARNmito est intéressante, il pourrait y avoir une compétition dans l'import de la mitoCas9 et du gARNmito. Cela s'oppose à l'hypothèse recensée dans les publications, qui suppose que les voies d'import sont séparées pour les ARNs et les protéines.

Le manque d'efficacité du système semble être la conséquence de la difficulté à importer les gARNs dans la mitochondrie. Bien que l'import des ARNs dans la mitochondrie ait déjà été couronné de succès dans plusieurs études, la méthode ne semble pas convenir pour les gARNs. Les publications faisant état de réussite pour l'import des ARNs étaient majoritairement des ARNt. La structure du gARN diffère peut-être trop de ces ARNt pour être efficacement importé (notamment le trackRNA présente une boucle ou les séquences sont appariées). Pour permettre son import, une stratégie pourrait être de s'inspirer des travaux de l'équipe de Numata (Chuah et al. 2016) et d'utiliser des peptides. Cette étude a permis l'import d'un ADN en utilisant un peptide de 30 acides aminés composé d'une partie MLS de 12 acides aminés provenant de la protéine COX4 pour l'import dans la mitochondrie puis d'une partie de 18 acides aminés composée de lysine et d'histidine pour la complexation avec l'ADN. Cette méthode pourrait peut-être aussi fonctionner avec de l'ARN. Globalement, une meilleure connaissance de la voie d'import des ARNs pourrait permettre d'optimiser les séquences d'import des ARNs.

Enfin pour augmenter encore l'efficacité d'import de la mitoCas9, il a été mis en évidence que certains MLS pouvait fonctionner pour certaines protéines et pas pour d'autres (Bacman et al. 2013). Pour augmenter l'import de la mitoCas9 à la mitochondrie, nous avons donc testé d'autres MLS que celui de SOD2 initialement utilisé : ceux de la malate déshydrogénase (MDH), de POLG et de TFAM, pour la mitoCas9 (laboratoire Agnel Sfeir). Cependant il n'a pu être mis en évidence une réelle augmentation de l'efficacité d'import de la mitoCas9 par rapport à notre construction initiale.

Conclusion

Mon projet de thèse m'a permis de contribuer à une meilleure compréhension du mécanisme impliqué dans la formation de délétions retrouvées au sein de l'ADN mitochondrial mais également d'éclaircir le mode de réPLICATION de ce dernier.

La mise au point de la technique de peignage moléculaire pour l'ADN mitochondrial nous a permis de conclure à l'utilisation du modèle de déplacement de brin proposé par Clayton lors de la réPLICATION de l'ADN mitochondrial. Lorsque des protéines mutées sont identifiées chez des patients souffrant le plus souvent de pathologies sévères telles que des neuromyopathies, cet outil permettra d'étudier les conséquences sur la réPLICATION.

Afin de mieux appréhender la formation fréquente de délétions au sein de l'ADN mitochondrial, nos travaux se sont appuyés sur l'étude de la « Common Deletion ». Pour induire la délétion, nous avons utilisé des mitoTALENs, TALENs qui ciblent l'ADN mitochondrial. Nous avons également tenté d'adapter le système CRISP/Cas9 à la mitochondrie. Cet outil moins restrictif que les TALENs dans le choix des séquences cibles aurait pu donner l'opportunité de compléter nos travaux afin de mieux comprendre la prise en charge des cassures double-brin. Malheureusement, la tentative d'adaptation de ce système à la mitochondrie a échoué.

Grâce à une des mitoTALENs, nous sommes ainsi parvenus à induire la délétion puis à étudier des protéines participant à sa formation. Notre stratégie a consisté à inhiber des protéines candidates de la réparation et de la réPLICATION. Seules les protéines de réPLICATION se sont avérées modifier la quantité de « Common Deletion ». Nous avons ainsi identifié les protéines POLG, mtSSB, Twinkle, DNA2, PIF-1, MGME1 et la DNA LIGASE 3. Ces résultats semblent suggérer d'une part que cette délétion se forme après une réparation défective lorsque la molécule d'ADN est en cours de réPLICATION et d'autre part que la mitochondrie ne répare pas les cassures double-brin grâce à l'utilisation des voies du NHEJ, alt-NHEJ et de la recombinaison homologue, contrairement à ce qui se passe dans le noyau. De plus, nous avons observé que l'induction de cassures double-brin entraînait une diminution drastique en ADN mitochondrial, renforçant comme préalablement évoqué, un mécanisme de dégradation en cas de dommages. Ma thèse a donc conduit à une meilleure compréhension des mécanismes de réparation et de réPLICATION présents dans la mitochondrie. Néanmoins, il paraît essentiel de

compléter ses observations car les modifications de l'ADN mitochondrial sont impliquées dans un nombre important de pathologies sévères dont des cancers. A l'heure actuelle, nous ne possédons pas chez l'Homme de stratégie thérapeutique efficace pour modifier les degrés d'hétéroplasmie responsables de leur survenue. Néanmoins l'élimination *in vitro* de mutations comme celle à l'origine du syndrome NARP ou l'arrêt spécifique de réPLICATION DES molécules mutées responsable du syndrome MERF suscitent beaucoup d'espoir pour l'élaboration de nouvelles thérapies. La proposition de nouveaux traitements passera par une meilleure compréhension de la réparation et de la réPLICATION ET DES mécanismes contrôlant la régulation de l'hétéroplasmie.

Annexes

I. Matériel et méthodes

1) Préparation des gARNs par transcription *in vitro*

Les différents gARNs (avec ou sans signaux d'adressage pour la mitochondrie) sont préparés par PCR à partir du plasmide codant pour le gARNmito (Loop et 3'UTR). La PCR est réalisée avec la polymérase Phusion (New England Biolabs) avec le programme suivant de 35 cycles : 98°C 30s, 98°C 10s, 60°C 30s, 72°C 15s, 72°C 10min, RT ∞. Les produits de PCR sont ensuite précipités et resuspendus dans de l'H₂O (RNase free). Les primers utilisés sont les suivants :

Primer	Séquence
Loop Forward	GAAATTAATACGACTCACTATAAGGTCTCCCTGAGCTTCAGGGAG
Pas de loop Forward	GAAATTAATACGACTCACTATAAGGATTGTGGGGCAATGAATGA AGCGGTT
3'UTR Reverse	AAAAAAGTGGCCTGATGGAAGAG
Pas de 3'UTR Reverse	AAAAAACCTCGAGAAGCTTAATAAACG

Les différentes séquences des gARNs obtenus par PCR sont présentées ci-dessus : le promoteur T7 est en rouge, la loop en rose, la séquence du guide en vert, le trackARN en violet et le 3'UTR en orange.

Nom du gARN	Séquence
Pas de loop-gARN-pas de 3'UTR	TAATACGACTCACTATAAGGGATTGTGGGGCAATGAATGAAGCGGTTGTAG CTCCCTTCGAAAGAACCGTTGCTACAATAAGGCCGTGAAAAGATGTGC CGCAACGCTCTGCCCTAAAGCTTCTGCTTAAGGGGCATCGTTAAG CTTCTCGAGGTTTTTTT
Loop-gARN-pas de 3'UTR	TAATACGACTCACTATAAGGGCTCCCTGAGCTTCAGGGAGCGTCTCGACAC ATTGTGGGGCAATGAATGAAGCGGTTGTAGCTCCCTTCGAAAGAACCGT TGCTACAATAAGGCCGTGAAAAGATGTGCCGCAACGCTCTGCCCTAA AGCTTCTGCTTAAGGGGCATCGTTATTAAGCTTCTGAGGTTTTTTT
Pas de loop-gARN- 3'UTR	TAATACGACTCACTATAAGGGCGTCTGACACATTGTGGGGCAATGAATGA AGCGGTTGTAGCTCCCTTCGAAAGAACCGTTGCTACAATAAGGCCGTG AAAAGATGTGCCGCAACGCTCTGCCCTAAAGCTTCTGCTTAAGGGCA TCGTTATTAAGCTTCTCGAGG CAGAAGAAGTGA CGGGCTGGGGCACAGTG GGCTGGCGCCCTGCAGAACATGAACCTCCGCTCTGGCTGCCACAGGG TCCTCCGATGCTGGCCTTGCAGCTAGAGGGCAGCCACTCATGGATTCAAG TCCTGGCTCCGCCTTCCATCAGGACCACTTTTTTT

Loop-gARN-3'UTR	TAATACGACTCACTATA GGG TCTCCCTGAGCTTCAGGGAG CGTCTCGACAC ATTGTGGGGCAATGAATGAAGCGGTAGCTCCCTTCGAAAGAACCGT TGCTACAATAAGGCCGTCTGAAAAGATGTGCCGCAACGCTCTGCCCTAA AGCTTCTGCTTAAGGGCATCGTTATTAGCTTCTGAGG CAGAAGAAGT GACGGCTGGGGCACAGTGGGCTGGCGCCCTGCAGAACATGAACCTCC GCTCCTGGCTGCCACAGGGCCTCCGATGCTGGCCTTGCGCCTCTAGAGGC AGCCACTCATGGATTCAAGTCCTGGCTCCGCCTTCATCAGGACCACTTT
-----------------	---

La transcription *in vitro* est réalisée O/N à 37°C à partir des différents produits de PCR en utilisant le kit Hscribe T7 High Yield kit (New England Biolabs). Le lendemain, les transcrits sont traités pendant 30 min at 37°C avec la Murine RNase Inhibitor (NEB) et la DNase I (NEB). Les ARNs sont ensuite purifiés grâce au kit Microelute RNA clean UP kit (Omega Biotek). Avant d'être chargés sur gel, les ARNs sont mis dans du bleu de charge contenant du Formaldehyde/formamide et chauffé 2 minutes à 90°C puis laissés une minute dans la glace. Ils sont ensuite chargés sur un gel agarose 3%, MAE 1X.

2) Expérience de coupure *in vitro*

La Cas9 et l'ARN du gARN sont incubés 30 minutes à 37°C dans un tampon A (Hepes-NaOH 20mM pH 7.5, KCl 150mM, BSA 0,1 mg/mL, DTT 1mM). La stoechiométrie utilisée pour la Cas9/gARN est 1, 5 : 1 en présence de 5 nM d'ADN cible. La réaction est stoppée en ajoutant un tampon B (Protéinase K 55U/mL, RNase A (Dnase free) (1,4µg/µL), sucre 50%), le mélange est incubé pendant 30 minutes à 37°C. Les échantillons sont ensuite chargés sur un gel agarose 1%.

3) Expériences avec les siARN RAD52 et PARP-1

Les expériences sont réalisées dans les conditions publiées.

Les siARNS utilisés sont :

siARN	Référence ou séquence
siRAD52	Dharmacon, référence : L-011760-00-0005
siPARP-1 A	GAUAGAGCGUGAAGGCGAATT
siPARP-1 B	GCCUCCGCUCCUGAACAAUTT

Les anticorps utilisés sont :

Protéine	Référence de l'anticorps
RAD52	Rad52 (F-7) Santa Cruz Biotechnology
PARP-1	PARP Antibody #9542, cell signaling

Bibliographie

- Ahnesorg, P., Smith, P. & Jackson, S.P., 2006. XLF Interacts with the XRCC4-DNA Ligase IV Complex to Promote DNA Nonhomologous End-Joining. *Cell*, 124(2), pp.301–313.
- Ahting, U. et al., 1999. The TOM core complex: the general protein import pore of the outer membrane of mitochondria. *The Journal of cell biology*, 147(5), pp.959–68.
- Alam, T.I. et al., 2003. Human mitochondrial DNA is packaged with TFAM. *Nucleic acids research*, 31(6), pp.1640–5.
- Anderson, S. et al., 1981. Sequence and organization of the human mitochondrial genome. *Nature*, 290(5806), pp.457–65.
- Anson, R.M., Hudson, E. & Bohr, V.A., 2000. Mitochondrial endogenous oxidative damage has been overestimated. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology*, 14(2), pp.355–60.
- Audebert, M., Salles, B. & Calsou, P., 2004. Involvement of Poly(ADP-ribose) Polymerase-1 and XRCC1/DNA Ligase III in an Alternative Route for DNA Double-strand Breaks Rejoining. *Journal of Biological Chemistry*, 279(53), pp.55117–55126.
- Bacman, S.R. et al., 2007. Modulating mtDNA heteroplasmy by mitochondria-targeted restriction endonucleases in a “differential multiple cleavage-site” model. *Gene therapy*, 14(18), pp.1309–18.
- Bacman, S.R. et al., 2013. Specific elimination of mutant mitochondrial genomes in patient-derived cells by mitoTALENs. *Nature Medicine*, 19(9), pp.1111–1113.
- Barrangou, R. et al., 2007. CRISPR Provides Acquired Resistance Against Viruses in Prokaryotes. *Science*, 315(5819), pp.1709–1712.
- Barupala, D.P. et al., 2016. Synthesis, delivery and regulation of eukaryotic heme and Fe–S cluster cofactors. *Archives of Biochemistry and Biophysics*, 592, pp.60–75.
- Bauer, M.F. et al., 1996. Role of Tim23 as Voltage Sensor and Presequence Receptor in Protein Import into Mitochondria. *Cell*, 87(1), pp.33–41.
- Beckman, K.B. & Ames, B.N., 1999. Endogenous oxidative damage of mtDNA. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, 424(1), pp.51–58.
- Bennardo, N. et al., 2008. Alternative-NHEJ is a mechanistically distinct pathway of mammalian chromosome break repair. *PLoS genetics*, 4(6), p.e1000110.
- Berthold, J. et al., 1995. The MIM Complex Mediates Preprotein Translocation across the Mitochondrial Inner Membrane and Couples It to the mt-Hsp70/ATP Driving System. *Cell*, 81, pp.1085–1093.

- Bianchi, K. et al., 2004. Calcium and mitochondria: mechanisms and functions of a troubled relationship. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, 1742(1–3), pp.119–131.
- Bitinaite, J. et al., 1998. FokI dimerization is required for DNA cleavage. *Proceedings of the National Academy of Sciences of the United States of America*, 95(18), pp.10570–5.
- van der Bliek, A.M., Shen, Q. & Kawajiri, S., 2013. Mechanisms of mitochondrial fission and fusion. *Cold Spring Harbor perspectives in biology*, 5(6), p.a011072.
- Boch, J. et al., 2009. Breaking the Code of DNA Binding Specificity of TAL-Type III Effectors. *Science*, 326(5959), pp.1509–1512.
- Bolden, A., Noy, G.P. & Weissbach, A., 1977. DNA polymerase of mitochondria is a gamma-polymerase. *The Journal of biological chemistry*, 252(10), pp.3351–6.
- Boulton, S.J. & Jackson, S.P., 1996. Saccharomyces cerevisiae Ku70 potentiates illegitimate DNA double-strand break repair and serves as a barrier to error-prone DNA repair pathways. *The EMBO journal*, 15(18), pp.5093–103.
- Boyer, P.D., 1975. A model for conformational coupling of membrane potential and proton translocation to ATP synthesis and to active transport. *FEBS letters*, 58(1), pp.1–6.
- Brewer, B.J. & Fangman, W.L., 1967. The Localization of Replication Origins on ARS Plasmids in *S. cerevisiae*. *Cell*, 51, pp.463–471.
- Brown, T.A. et al., 2005. Replication of mitochondrial DNA occurs by strand displacement with alternative light-strand origins, not via a strand-coupled mechanism. *Genes & Development*, 19(20), pp.2466–2476.
- Brown, T.A. et al., Replication of mitochondrial DNA occurs by strand displacement with alternative light-strand origins, not via a strand-coupled mechanism.
- Buck, D. et al., 2006. Cernunnos, a novel nonhomologous end-joining factor, is mutated in human immunodeficiency with microcephaly. *Cell*, 124(2), pp.287–99.
- Campbell, G. et al., 2014. Dissecting the mechanisms underlying the accumulation of mitochondrial DNA deletions in human skeletal muscle. *Human molecular genetics*, 23(17), pp.4612–20.
- Chan, S.S.L., Longley, M.J. & Copeland, W.C., 2005. The common A467T mutation in the human mitochondrial DNA polymerase (POLG) compromises catalytic efficiency and interaction with the accessory subunit. *The Journal of biological chemistry*, 280(36), pp.31341–6.
- Chang, D.D. & Clayton, D.A., 1985. Priming of human mitochondrial DNA replication occurs at the light-strand promoter (displacement loop/regulatory sequences/template signals/transcription). *Biochemistry*, 82, pp.351–355.
- Chattopadhyay, R. et al., 2006. Identification and characterization of mitochondrial abasic (AP)-endonuclease in mammalian cells. *Nucleic acids research*, 34(7), pp.2067–76.

- Chen, T., He, J., Huang, Y., et al., 2011. The generation of mitochondrial DNA large-scale deletions in human cells. *Journal of Human Genetics*, 56(10), pp.689–694.
- Chen, T., He, J., Shen, L., et al., 2011. The mitochondrial DNA 4,977-bp deletion and its implication in copy number alteration in colorectal cancer. *BMC Medical Genetics*, 12(1), p.8.
- Chen, T. et al., The mitochondrial DNA 4,977-bp deletion and its implication in copy number alteration in colorectal cancer.
- Chifman, J., Laubenbacher, R. & Torti, S. V, 2014. A systems biology approach to iron metabolism. *Advances in experimental medicine and biology*, 844, pp.201–25.
- Chuah, J.-A. et al., 2016. Self-Assembled Peptide-Based System for Mitochondrial-Targeted Gene Delivery: Functional and Structural Insights. *Biomacromolecules*, 17(11), pp.3547–3557.
- Clayton, D.A., 1982. Replication of animal mitochondrial DNA. *Cell*, 28(4), pp.693–705.
- Clayton, D.A., Doda, J.N. & Friedberg, E.C., 1974. The absence of a pyrimidine dimer repair mechanism in mammalian mitochondria. *Proceedings of the National Academy of Sciences of the United States of America*, 71(7), pp.2777–81.
- Coffey, G. & Campbell, C., 2000. An alternate form of Ku80 is required for DNA end-binding activity in mammalian mitochondria. *Nucleic acids research*, 28(19), pp.3793–800.
- Comte, C. et al., 2013. Mitochondrial targeting of recombinant RNAs modulates the level of a heteroplasmic mutation in human mitochondrial DNA associated with Kearns Sayre Syndrome. *Nucleic acids research*, 41(1), pp.418–33.
- Contreras, L. et al., 2010. Mitochondria: The calcium connection. *Biochimica et Biophysica Acta (BBA) - Bioenergetics*, 1797(6–7), pp.607–618.
- Cristodero, M. et al., 2013. Mitochondrial translation factors of *Trypanosoma brucei*: elongation factor-Tu has a unique subdomain that is essential for its function. *Molecular Microbiology*, 90(4), pp.744–755.
- Deng, D. et al., 2012. Structural Basis for Sequence-Specific Recognition of DNA by TAL Effectors. *Science*, 335(6069), pp.720–723.
- Diaz, F. et al., 2002. Human mitochondrial DNA with large deletions repopulates organelles faster than full-length genomes under relaxed copy number control. *Nucleic acids research*, 30(21), pp.4626–33.
- Dimberg, J. et al., 2011. Common 4977 bp deletion and novel alterations in mitochondrial DNA in Vietnamese patients with breast cancer.
- Dominy, J.E. & Puigserver, P., 2013. Mitochondrial Biogenesis through Activation of Nuclear Signaling Proteins. *Cold Spring Harbor Perspectives in Biology*, 5(7), pp.a015008–a015008.

- Doyon, Y. et al., 2011. Enhancing zinc-finger-nuclease activity with improved obligate heterodimeric architectures. *Nature methods*, 8(1), pp.74–9.
- Duxin, J.P. et al., 2009. Human Dna2 Is a Nuclear and Mitochondrial DNA Maintenance Protein. *Molecular and Cellular Biology*, 29(15), pp.4274–4282.
- Eshaghian, A. et al., 2006. Mitochondrial DNA deletions serve as biomarkers of aging in the skin, but are typically absent in nonmelanoma skin cancers. *The Journal of investigative dermatology*, 126(2), pp.336–44.
- Fairall, L. et al., 1993. The crystal structure of a two zinc-finger peptide reveals an extension to the rules for zinc-finger/DNA recognition. *Nature*, 366(6454), pp.483–7.
- Falkenberg, M. et al., 2002. Mitochondrial transcription factors B1 and B2 activate transcription of human mtDNA. *Nature genetics*, 31(3), pp.289–94.
- Fan, L. et al., 2006. A Novel Processive Mechanism for DNA Synthesis Revealed by Structure, Modeling and Mutagenesis of the Accessory Subunit of Human Mitochondrial DNA Polymerase. *Journal of Molecular Biology*, 358(5), pp.1229–1243.
- Fisher, R.P. & Clayton, D.A., 1988. Purification and characterization of human mitochondrial transcription factor 1. *Molecular and cellular biology*, 8(8), pp.3496–509.
- Fukui, H. & Moraes, C.T., 2009. Mechanisms of formation and accumulation of mitochondrial DNA deletions in aging neurons. *Human Molecular Genetics*, 18(6), pp.1028–1036.
- Fusté, J.M. et al., 2014. In Vivo Occupancy of Mitochondrial Single-Stranded DNA Binding Protein Supports the Strand Displacement Mode of DNA Replication.
- Fusté, J.M. et al., 2010. Mitochondrial RNA Polymerase Is Needed for Activation of the Origin of Light-Strand DNA Replication. *Molecular Cell*, 37(1), pp.67–78.
- Garcia, M. et al., 2007. Mitochondria-associated yeast mRNAs and the biogenesis of molecular complexes. *Molecular biology of the cell*, 18(2), pp.362–8.
- Gaspari, M. et al., 2004. The mitochondrial RNA polymerase contributes critically to promoter specificity in mammalian cells. *The EMBO journal*, 23(23), pp.4606–14.
- Goffart, S. et al., 2009. Twinkle mutations associated with autosomal dominant progressive external ophthalmoplegia lead to impaired helicase function and in vivo mtDNA replication stalling. *Human molecular genetics*, 18(2), pp.328–40.
- Graves, S.W., Johnson, A.A. & Johnson, K.A., 1998. Expression, purification, and initial kinetic characterization of the large subunit of the human mitochondrial DNA polymerase. *Biochemistry*, 37(17), pp.6050–8.
- Gyllensten, U., Wharton, D. & Wilson, A.C., 1985. Maternal inheritance of mitochondrial DNA during backcrossing of two species of mice. *Journal of Heredity*, 76(5), pp.321–324.
- Hayashi, J. et al., 1991. Introduction of disease-related mitochondrial DNA deletions into HeLa cells lacking mitochondrial DNA results in mitochondrial dysfunction. *Proceedings of the National Academy of Sciences of the United States of America*, 88(23), pp.10614–8.

- Holt, I.J., Harding, A.E. & Morgan-Hughes, J.A., 1989. Deletions of muscle mitochondrial DNA in mitochondrial myopathies: sequence analysis and possible mechanisms. *Nucleic acids research*, 17(12), pp.4465–9.
- Holt, I.J., Lorimer, H.E. & Jacobs, H.T., 2000. Coupled Leading- and Lagging-Strand Synthesis of Mammalian Mitochondrial DNA. *Cell*, 100(5), pp.515–524.
- Holt, I.J. & Reyes, A., 2012a. Human mitochondrial DNA replication. *Cold Spring Harbor perspectives in biology*, 4(12).
- Horst, M. et al., 1995. Dynamic interaction of the protein translocation systems in the inner and outer membranes of yeast mitochondria. *The EMBO journal*, 14(10), pp.2293–7.
- Hou, J.H. & Wei, Y.H., 1998. AT-rich sequences flanking the 5'-end breakpoint of the 4977-bp deletion of human mitochondrial DNA are located between two bent-inducing DNA sequences that assume distorted structure in organello. *Mutation research*, 403(1–2), pp.75–84.
- Hu, J. et al., 2005. Repair of Formamidopyrimidines in DNA Involves Different Glycosylases ROLE OF THE OGG1, NTH1, AND NEIL1 ENZYMES * □ S.
- Ichim, G. & Tait, S.W.G., 2016. A fate worse than death: apoptosis as an oncogenic process. *Nature reviews. Cancer*, 16(8), pp.539–48.
- Jinek, M. et al., 2012. A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science*, 337(6096), pp.816–821.
- Kabotyanski, E.B. et al., 1998. Double-strand break repair in Ku86- and XRCC4-deficient cells. *Nucleic Acids Research*, 26(23), pp.5333–5342.
- Kamenisch, Y. et al., 2007. The Mitochondrial DNA Common Deletion Is Present in Most Basal and Squamous Cell Carcinoma Samples Isolated by Laser Capture Microdissection but Generally at Reduced Rather than Increased Levels. *Journal of Investigative Dermatology*, 127(2), pp.486–490.
- Karicheva, O.Z. et al., 2011. Correction of the consequences of mitochondrial 3243A>G mutation in the MT-TL1 gene causing the MELAS syndrome by tRNA import into mitochondria. *Nucleic acids research*, 39(18), pp.8173–86.
- Kasiviswanathan, R. et al., 2009. Disease mutations in the human mitochondrial DNA polymerase thumb subdomain impart severe defects in mitochondrial DNA replication. *The Journal of biological chemistry*, 284(29), pp.19501–10.
- Kastaniotis, A.J. et al., 2017. Mitochondrial fatty acid synthesis, fatty acids and mitochondrial physiology. *Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids*, 1862(1), pp.39–48.
- Kaufman, B.A. et al., 2007. The mitochondrial transcription factor TFAM coordinates the assembly of multiple DNA molecules into nucleoid-like structures. *Molecular biology of the cell*, 18(9), pp.3225–36.

- Kent, T. et al., 2015. Mechanism of microhomology-mediated end-joining promoted by human DNA polymerase θ. *Nature structural & molecular biology*, 22(3), pp.230–7.
- Khan, I. et al., 2016. Biochemical Characterization of the Human Mitochondrial Replicative Twinkle Helicase: substrate specificity, dna branch migration, and ability to overcome blockades to dna unwinding. *The Journal of biological chemistry*, 291(27), pp.14324–39.
- Kim, K.M. et al., 2017. Mitochondrial noncoding RNA transport. *BMB reports*, 50(4), pp.164–174.
- Kim, Y.G., Cha, J. & Chandrasegaran, S., 1996. Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. *Proceedings of the National Academy of Sciences of the United States of America*, 93(3), pp.1156–60.
- Kolesnikova, O.A. et al., 2000. Suppression of Mutations in Mitochondrial DNA by tRNAs Imported from the Cytoplasm. *Science*, 289(5486).
- Korhonen, J.A. et al., 2004. Reconstitution of a minimal mtDNA replisome in vitro. *The EMBO journal*, 23(12), pp.2423–9.
- Korhonen, J.A., Gaspari, M. & Falkenberg, M., 2003. TWINKLE Has 5' -> 3' DNA helicase activity and is specifically stimulated by mitochondrial single-stranded DNA-binding protein. *The Journal of biological chemistry*, 278(49), pp.48627–32.
- Kornblum, C. et al., 2013. Loss-of-function mutations in MGME1 impair mtDNA replication and cause multisystemic mitochondrial disease. *Nature Genetics*, 45(2), pp.214–219.
- Kukat, A. et al., 2008. Generation of Δ0 cells utilizing a mitochondrially targeted restriction endonuclease and comparative analyses. *Nucleic Acids Research*, 36(7), pp.e44–e44.
- Kutik, S. et al., 2008. Dissecting Membrane Insertion of Mitochondrial β-Barrel Proteins. *Cell*, 132(6), pp.1011–1024.
- Ledoux, S.P. et al., 1992. Repair of mitochondrial DNA after various types of DNA damage in Chinese hamster ovary cells. , 13(11), pp.1967–1973.
- Lewis, W. et al., 2007. Decreased mtDNA, oxidative stress, cardiomyopathy, and death from transgenic cardiac targeted human mutant polymerase gamma. *Laboratory investigation; a journal of technical methods and pathology*, 87(4), pp.326–35.
- Li, M. et al., 2010. Identification and characterization of mitochondrial targeting sequence of human apurinic/apyrimidinic endonuclease 1. *The Journal of biological chemistry*, 285(20), pp.14871–81.
- Lieber, M.R., 2008. The mechanism of human nonhomologous DNA end joining. *The Journal of biological chemistry*, 283(1), pp.1–5.

- Lim, S.E., Longley, M.J. & Copeland, W.C., 1999. The mitochondrial p55 accessory subunit of human DNA polymerase gamma enhances DNA binding, promotes processive DNA synthesis, and confers N-ethylmaleimide resistance. *The Journal of biological chemistry*, 274(53), pp.38197–203.
- Litonin, D. et al., 2010. Human mitochondrial transcription revisited: only TFAM and TFB2M are required for transcription of the mitochondrial genes in vitro. *The Journal of biological chemistry*, 285(24), pp.18129–33.
- Longley, M.J. et al., 1998. Identification of 5'-deoxyribose phosphate lyase activity in human DNA polymerase gamma and its role in mitochondrial base excision repair in vitro. *Proceedings of the National Academy of Sciences of the United States of America*, 95(21), pp.12244–8.
- Ma, Y. et al., 2004. A Biochemically Defined System for Mammalian Nonhomologous DNA End Joining. *Molecular Cell*, 16(5), pp.701–713.
- Mak, A.N.-S. et al., 2012. The crystal structure of TAL effector PthXo1 bound to its DNA target. *Science (New York, N.Y.)*, 335(6069), pp.716–9.
- Mali, P., Esvelt, K.M. & Church, G.M., 2013. Cas9 as a versatile tool for engineering biology. *Nature methods*, 10(10), pp.957–63.
- Mandal, S.M. et al., 2012. Role of Human DNA Glycosylase Nei-like 2 (NEIL2) and Single Strand Break Repair Protein Polynucleotide Kinase 3'-Phosphatase in Maintenance of Mitochondrial Genome. *Journal of Biological Chemistry*, 287(4), pp.2819–2829.
- Mandavilli, B.S., Santos, J.H. & Van Houten, B., 2002. Mitochondrial DNA repair and aging. *Mutation research*, 509(1–2), pp.127–51.
- Martin-Negrier, M.-L. et al., 2011. TWINKLE gene mutation: report of a French family with an autosomal dominant progressive external ophthalmoplegia and literature review. *European journal of neurology*, 18(3), pp.436–41.
- Mason, P.A. et al., 2003. Mismatch repair activity in mammalian mitochondria. *Nucleic Acids Research*, 31(3), pp.1052–1058.
- Meissner, C. et al., 2008. The 4977bp deletion of mitochondrial DNA in human skeletal muscle, heart and different areas of the brain: A useful biomarker or more? *Experimental Gerontology*, 43(7), pp.645–652.
- Mercer, T.R. et al., 2011. The Human Mitochondrial Transcriptome. *Cell*, 146, pp.645–658.
- Milenkovic, D. et al., 2013. TWINKLE is an essential mitochondrial helicase required for synthesis of nascent D-loop strands and complete mtDNA replication. *Human Molecular Genetics*, 22(10), pp.1983–1993.
- Miller, F.J. et al., 2003. Precise determination of mitochondrial DNA copy number in human skeletal and cardiac muscle by a PCR-based assay: lack of change of copy number with age. *Nucleic acids research*, 31(11), p.e61.

- Minczuk, M. et al., 2008. Development of a single-chain, quasi-dimeric zinc-finger nuclease for the selective degradation of mutated human mitochondrial DNA. *Nucleic Acids Research*, 36(12), pp.3926–3938.
- Minczuk, M. et al., 2006. Sequence-specific modification of mitochondrial DNA using a chimeric zinc finger methylase. *Proceedings of the National Academy of Sciences*, 103(52), pp.19689–19694.
- Montoya, J. et al., 1982. Identification of initiation sites for heavy-strand and light-strand transcription in human mitochondrial DNA. *Proceedings of the National Academy of Sciences of the United States of America*, 79(23), pp.7195–9.
- Mordas, A. & Tokatlidis, K., 2015. The MIA Pathway: A Key Regulator of Mitochondrial Oxidative Protein Folding and Biogenesis. *Accounts of Chemical Research*, 48(8), pp.2191–2199.
- Moretton, A. et al., 2017. Selective mitochondrial DNA degradation following double-strand breaks M. Sola, ed. *PLOS ONE*, 12(4), p.e0176795.
- Morozov, Y.I. & Temiakov, D., Human mitochondrial transcription initiation complexes have similar topology on the light and heavy strand promoters.
- Moshous, D. et al., 2001. Artemis, a novel DNA double-strand break repair/V(D)J recombination protein, is mutated in human severe combined immune deficiency. *Cell*, 105(2), pp.177–86.
- Moynahan, M.E., Pierce, A.J. & Jasin, M., 2001. BRCA2 is required for homology-directed repair of chromosomal breaks. *Molecular cell*, 7(2), pp.263–72.
- Nagaike, T. et al., 2005. Human Mitochondrial mRNAs Are Stabilized with Polyadenylation Regulated by Mitochondria-specific Poly(A) Polymerase and Polynucleotide Phosphorylase. *Journal of Biological Chemistry*, 280(20), pp.19721–19727.
- Negi, S. et al., 2015. LocSigDB: a database of protein localization signals. *Database : the journal of biological databases and curation*, 2015.
- Nick McElhinny, S.A. et al., 2000. Ku recruits the XRCC4-ligase IV complex to DNA ends. *Molecular and cellular biology*, 20(9), pp.2996–3003.
- Nie, H. et al., 2013. Mitochondrial common deletion, a potential biomarker for cancer occurrence, is selected against in cancer background: a meta-analysis of 38 studies. J. Santos, ed. *Plos one*, 8(7), p.e67953.
- O'Brien, T.W., 1971. The general occurrence of 55 S ribosomes in mammalian liver mitochondria. *The Journal of biological chemistry*, 246(10), pp.3409–17.
- Ohtsubo, T. et al., 2000. Identification of human MutY homolog (hMYH) as a repair enzyme for 2-hydroxyadenine in DNA and detection of multiple forms of hMYH located in nuclei and mitochondria. *Nucleic Acids Research*, 28(6), pp.1355–1364.
- Ojala, D., Montoya, J. & Attardi, G., 1981. tRNA punctuation model of RNA processing in human mitochondria. *Nature*, 290(5806), pp.470–4.

- Oliveira, M.T. & Kaguni, L.S., 2010. Functional Roles of the N- and C-Terminal Regions of the Human Mitochondrial Single-Stranded DNA-Binding Protein J. Santos, ed. *PLoS ONE*, 5(10), p.e15379.
- Pavletich, N.P. & Pabo, C.O., 1991. Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 Å. *Science (New York, N.Y.)*, 252(5007), pp.809–17.
- Peng, T.-I. et al., 2006. Visualizing common deletion of mitochondrial DNA-augmented mitochondrial reactive oxygen species generation and apoptosis upon oxidative stress. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease*, 1762(2), pp.241–255.
- Porteous, W.K. et al., 1998. Bioenergetic consequences of accumulating the common 4977-bp mitochondrial DNA deletion. *European Journal of Biochemistry*, 257(1), pp.192–201.
- Ramirez, C.L. et al., 2012. Engineered zinc finger nickases induce homology-directed repair with reduced mutagenic effects. *Nucleic Acids Research*, 40(12), pp.5560–5568.
- Reddy, P. et al., 2015. Selective Elimination of Mitochondrial Mutations in the Germline by Genome Editing. *Cell*, 161(3), pp.459–469.
- Rehling, P. et al., 2003. Protein insertion into the mitochondrial inner membrane by a twin-pore translocase. *Science (New York, N.Y.)*, 299(5613), pp.1747–51.
- Richter, C., Park, J.-W. & Amest, B.N., 1988. Normal oxidative damage to mitochondrial and nuclear DNA is extensive. *Genetics*, 85, pp.6465–6467.
- Ringel, R. et al., 2011. Structure of human mitochondrial RNA polymerase. *Nature*, 478(7368), pp.269–273.
- Robberson, D.L. & Clayton, D.A., 1972. Replication of Mitochondrial DNA in Mouse L Cells and Their Thymidine Kinase-Derivatives: Displacement Replication on a Covalently-Closed Circular Template (electron microscopy/mouse DNA/density gradient centrifugation). , 69(12), pp.3810–3814.
- Rossignol, R. et al., 2003. Mitochondrial threshold effects. *Biochem. J*, 370, pp.751–762.
- Rubio, M.A.T. et al., 2008. Mammalian mitochondria have the innate ability to import tRNAs by a mechanism distinct from protein import. *Proceedings of the National Academy of Sciences*, 105(27), pp.9186–9191.
- Ruhanen, H. et al., 2010. Mitochondrial single-stranded DNA binding protein is required for maintenance of mitochondrial DNA and 7S DNA but is not required for mitochondrial nucleoid organisation. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, 1803(8), pp.931–939.
- Safdar, A. et al., 2011. Endurance exercise rescues progeroid aging and induces systemic mitochondrial rejuvenation in mtDNA mutator mice. *Proceedings of the National Academy of Sciences of the United States of America*, 108(10), pp.4135–40.
- Samson, J.E. et al., 2013. Revenge of the phages: defeating bacterial defences. *Nature reviews. Microbiology*, 11(10), pp.675–87.

- Satoh, M. & Kuroiwa, T., 1991. Organization of Multiple Nucleoids and DNA Molecules in Mitochondria of a Human Cell. *EXPERIMENTAL CELL RESEARCH*, 196, pp.137–140.
- Schapira, A.H., 2006. Mitochondrial disease. *The Lancet*, 368(9529), pp.70–82.
- Scott, I. & Youle, R.J., 2010. Mitochondrial fission and fusion. *Essays in biochemistry*, 47, pp.85–98.
- Seligmann, H., 2010. Mitochondrial tRNAs as light strand replication origins: Similarity between anticodon loops and the loop of the light strand replication origin predicts initiation of DNA replication. *Biosystems*, 99(2), pp.85–93.
- Shoffner, J.M. et al., 1989. Spontaneous Kearns-Sayre/chronic external ophthalmoplegia plus syndrome associated with a mitochondrial DNA deletion: a slip-replication model and metabolic therapy. *Proceedings of the National Academy of Sciences of the United States of America*, 86(20), pp.7952–6.
- Shokolenko, I. et al., 2009. Oxidative stress induces degradation of mitochondrial DNA. *Nucleic Acids Research*, 37(8), pp.2539–2548.
- Simsek, D. et al., 2011. Crucial role for DNA ligase III in mitochondria but not in Xrcc1-dependent repair. *Nature*, 471(7337), pp.245–8.
- Spelbrink, J.N. et al., 2001. Human mitochondrial DNA deletions associated with mutations in the gene encoding Twinkle, a phage T7 gene 4-like protein localized in mitochondria. *Nature genetics*, 28(3), pp.223–31.
- Spelbrink, J.N. et al., 2000. In Vivo Functional Analysis of the Human Mitochondrial DNA Polymerase POLG Expressed in Cultured Human Cells. *Journal of Biological Chemistry*, 275(32), pp.24818–24828.
- Spencer, A.C. & Spremulli, L.L., 2004. Interaction of mitochondrial initiation factor 2 with mitochondrial fMet-tRNA. *Nucleic acids research*, 32(18), pp.5464–70.
- Srivastava, S. & Moraes, C.T., 2005. Double-strand breaks of mouse muscle mtDNA promote large deletions similar to multiple mtDNA deletions in humans. *Human Molecular Genetics*, 14(7), pp.893–902.
- Srivastava, S. & Moraes, C.T., 2001. Manipulating mitochondrial DNA heteroplasmy by a mitochondrially targeted restriction endonuclease. *Human molecular genetics*, 10(26), pp.3093–9.
- Stein, A., Kalifa, L. & Sia, E.A., 2015. Members of the RAD52 Epistasis Group Contribute to Mitochondrial Homologous Recombination and Double-Strand Break Repair in *Saccharomyces cerevisiae*. *PLoS genetics*, 11(11), p.e1005664.
- Stumpf, J.D. & Copeland, W.C., 2011. Mitochondrial DNA replication and disease: insights from DNA polymerase γ mutations. *Cellular and Molecular Life Sciences*, 68(2), pp.219–233.

- Sylvestre, J. et al., 2003. The Role of the 3' Untranslated Region in mRNA Sorting to the Vicinity of Mitochondria Is Conserved from Yeast to Human Cells. *Molecular Biology of the Cell*, 14, pp.3848–3856.
- Szczesny, R.J. et al., 2013. Identification of a novel human mitochondrial endo-/exonuclease Ddk1/c20orf72 necessary for maintenance of proper 7S DNA levels. *Nucleic Acids Research*, 41(5), pp.3144–3161.
- Tadi, S.K. et al., 2016. Microhomology-mediated end joining is the principal mediator of double-strand break repair during mitochondrial DNA lesions. *Molecular Biology of the Cell*, 27(2), pp.223–235.
- Takao, M. et al., 1998. Mitochondrial targeting of human DNA glycosylases for repair of oxidative DNA damage. *Nucleic Acids Research*, 26(12).
- Tapper, D.P. & Clayton, D.A., 1981. Mechanism of replication of human mitochondrial DNA. Localization of the 5' ends of nascent daughter strands. *The Journal of biological chemistry*, 256(10), pp.5109–15.
- Taylor, A.B. et al., 2001. Crystal Structures of Mitochondrial Processing Peptidase Reveal the Mode for Specific Cleavage of Import Signal Sequences. *Structure*, 9(7), pp.615–625.
- Thrash, J.C. et al., 2011. Phylogenomic evidence for a common ancestor of mitochondria and the SAR11 clade. *Scientific Reports*, 1(1), p.13.
- Thyagarajan, B., Padua, R.A. & Campbell, C., 1996. Mammalian Mitochondria Possess Homologous DNA Recombination Activity. *Journal of Biological Chemistry*, 271(44), pp.27536–27543.
- Ting, S.-Y. et al., 2014. Architecture of the TIM23 inner mitochondrial translocon and interactions with the matrix import motor. *The Journal of biological chemistry*, 289(41), pp.28689–96.
- Tiranti, V. et al., 1993. Cloning of human and rat cDNAs encoding the mitochondrial single-stranded DNA-binding protein (SSB). *Gene*, 126(2), pp.219–25.
- Tiranti, V. et al., 1997. Identification of the gene encoding the human mitochondrial RNA polymerase (h-mtRPOL) by cyberscreening of the Expressed Sequence Tags database. *Human molecular genetics*, 6(4), pp.615–25.
- Trapp, C., McCullough, A.K. & Epe, B., 2007. The basal levels of 8-oxoG and other oxidative modifications in intact mitochondrial DNA are low even in repair-deficient (Ogg1^{-/-}/Csb^{-/-}) mice. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, 625(1), pp.155–163.
- Truong, L.N. et al., 2013. Microhomology-mediated End Joining and Homologous Recombination share the initial end resection step to repair DNA double-strand breaks in mammalian cells. *Proceedings of the National Academy of Sciences of the United States of America*, 110(19), pp.7720–5.

- Tyynismaa, H. et al., 2005. Mutant mitochondrial helicase Twinkle causes multiple mtDNA deletions and a late-onset mitochondrial disease in mice. *Proceedings of the National Academy of Sciences of the United States of America*, 102(49), pp.17687–92.
- Uhler, J.P. et al., 2016. MGME1 processes flaps into ligatable nicks in concert with DNA polymerase γ during mtDNA replication. *Nucleic Acids Research*, 44(12), pp.5861–5871.
- Vermulst, M. et al., 2008. DNA deletions and clonal mutations drive premature aging in mitochondrial mutator mice. *Nature Genetics*, 40(4), pp.392–394.
- Walberg, M.W. & Clayton, D.A., 1981. Sequence and properties of the human KB cell and mouse L cell D-loop regions of mitochondrial DNA. *Nucleic acids research*, 9(20), pp.5411–21.
- Wallace, D.C., 2016. Genetics: Mitochondrial DNA in evolution and disease. *Nature*, 535(7613), pp.498–500.
- Wallin, I.E., 1926. bacteria and the origin of species. *Science (New York, N.Y.)*, 64(1651), pp.173–5.
- Wang, G. et al., 2012. Correcting human mitochondrial mutations with targeted RNA import. *Proceedings of the National Academy of Sciences of the United States of America*, 109(13), pp.4840–5.
- Wang, G. et al., 2010. PNPASE Regulates RNA Import into Mitochondria. *Cell*, 142(3), pp.456–467.
- Wang, M. et al., 2006. PARP-1 and Ku compete for repair of DNA double strand breaks by distinct NHEJ pathways. *Nucleic acids research*, 34(21), pp.6170–82.
- Wang, P. et al., 2013. Mitochondria DNA 4977 bp common deletion in peripheral whole blood from healthy donors. *Biomedical and environmental sciences : BES*, 26(12), pp.990–3.
- Wisnovsky, S., Jean, S.R. & Kelley, S.O., 2016. Mitochondrial DNA repair and replication proteins revealed by targeted chemical probes. *Nature chemical biology*, 12(7), pp.567–73.
- Wu, Y. et al., 2014. TALE nickase mediates high efficient targeted transgene integration at the human multi-copy ribosomal DNA locus. *Biochemical and biophysical research communications*, 446(1), pp.261–6.
- von Wurmb-Schwarz, N. et al., 2010. Low level of the mtDNA4977 deletion in blood of exceptionally old individuals. *Mechanisms of Ageing and Development*, 131(3), pp.179–184.
- Yakubovskaya, E. et al., 2014. Organization of the human mitochondrial transcription initiation complex. *Nucleic acids research*, 42(6), pp.4100–12.

- Yaneva, M., Kowalewski, T. & Lieber, M.R., 1997. Interaction of DNA-dependent protein kinase with DNA and with Ku: biochemical and atomic-force microscopy studies. *The EMBO Journal*, 16(16), pp.5098–5112.
- Yang, C. et al., 1997. Crystal structure of human mitochondrial single-stranded DNA binding protein at 2.4 Å resolution. *Nature structural biology*, 4(2), pp.153–7.
- Yang, D. et al., 1985. Mitochondrial origins. *Proceedings of the National Academy of Sciences of the United States of America*, 82(13), pp.4443–7.
- Yang, M.Y. et al., 2002. Biased Incorporation of Ribonucleotides on the Mitochondrial L-Strand Accounts for Apparent Strand-Asymmetric DNA Replication. *Cell*, 111(4), pp.495–505.
- Yasukawa, T. et al., 2006. Replication of vertebrate mitochondrial DNA entails transient ribonucleotide incorporation throughout the lagging strand. *The EMBO journal*, 25(22), pp.5358–71.
- Young, J.C., Hoogenraad, N.J. & Hartl, F.U., 2003. Molecular Chaperones Hsp90 and Hsp70 Deliver Preproteins to the Mitochondrial Import Receptor Tom70. *Cell*, 112(1), pp.41–50.
- Zhang, D. et al., 2000. Construction of transgenic mice with tissue-specific acceleration of mitochondrial DNA mutagenesis. *Genomics*, 69(2), pp.151–61.
- Zheng, L. et al., 2008. Human DNA2 Is a Mitochondrial Nuclease/Helicase for Efficient Processing of DNA Replication and Repair Intermediates. *Molecular Cell*, 32(3), pp.325–336.

PARTIE III

Abréviations

ADN : acide désoxyribonucléique

ARN : acide ribonucléique

AFM : microscopie à force atomique

ATP : adénosine triphosphate

bp : paire de bases

C terminal : carboxy terminal

EcTopoA : topoisomérase A d'*Escherichia coli*

hTopoI : topoisomérase IB humaine

hTopoImito : topoisomérase IB humaine forme mitochondriale

hTopo2 α : topoisomérase humaine 2 α

hTopo2 β : topoisomérase humaine 2 β

hTopo3 α : topoisomérase humaine 3 α

hTopo3 β : topoisomérase humaine 3 β

Lk : indice d'enlacement

RMI1 : RecQ-mediated genome instability 1

RMI2 : RecQ-mediated genome instability 2

SDS-PAGE : sodium dodecyl sulfate polyacrylamide gel electrophoresis

Tw : twist

Wr : writhe

WT : Wild Type

Sommaire

Introduction	141
I. La topologie de l'ADN	141
1) Présentation des différentes topologies de l'ADN.....	141
2) Etude de la topologie de l'ADN grâce à des protéines ou molécules chimiques ..	142
II. Les protéines qui agissent sur la topologie de l'ADN : les topoisomérases.....	143
1) Présentation des différentes familles de topoisomérases.....	145
2) Les topoisomérases comme cible thérapeutique	148
III. Les tests <i>in vitro</i> et leurs lacunes pour l'étude des topoisomérases	148
Le projet de thèse	150
Résultats	151
I. Article 1 : « Use of double-stranded DNA mini-circles to characterize the covalent topoisomerase-DNA complex »	151
II. Article 2 : “Probing hyper-negatively supercoiled mini-circles with nucleases and DNA binding proteins”	175
Discussion	208
Conclusion.....	210
Bibliographie.....	211

Table des figures

<u>FIGURE 1 : REPRESENTATION DE LA DOUBLE HELICE DE L'ADN</u>	141
<u>FIGURE 2 : SCHEMA REPRESENTANT DIFFERENTES TOPOLOGIES QUE L'ADN PEUT ADOPTER</u>	142
<u>FIGURE 3 : CONTRAINTES TOPOLOGIQUES SUBIES PAR L'ADN LORS DE LA TRANSCRIPTION</u>	144
<u>FIGURE 4 : SCHEMA PRESENTANT L'ACTIVITE DE RELAXATION OU D'INTRODUCTION DE SUPERTOURS ET DE CATENATION/DECATENATION DE DEUX MOLECULES D'ADN SŒURS</u>	144
<u>FIGURE 5 : SCHEMA PRESENTANT LE MECANISME DE RELAXATION DE L'ADN GRACE A UNE REACTION DE TRANS-ESTERIFICATION PAR LES TOPOISOMERASES DE LA FAMILLE IA</u>	146
<u>FIGURE 6 : SCHEMA REPRESENTANT LE MECANISME DE ROTATION CONTROLEE UTILISE PAR LES TOPOISOMERASES DE TYPE IB</u>	147

Introduction

I. La topologie de l'ADN

1) Présentation des différentes topologies de l'ADN

L'acide désoxyribonucléique (ADN) est le support de l'information génétique. Sa structure en double hélice a été découverte par Watson et Crick en 1953. Elle est composée d'un squelette de sucres et de phosphates et de 4 bases différentes : adénine (A), cytosine (C), guanine (G) ou thymine (T) (Watson & Crick 1953)

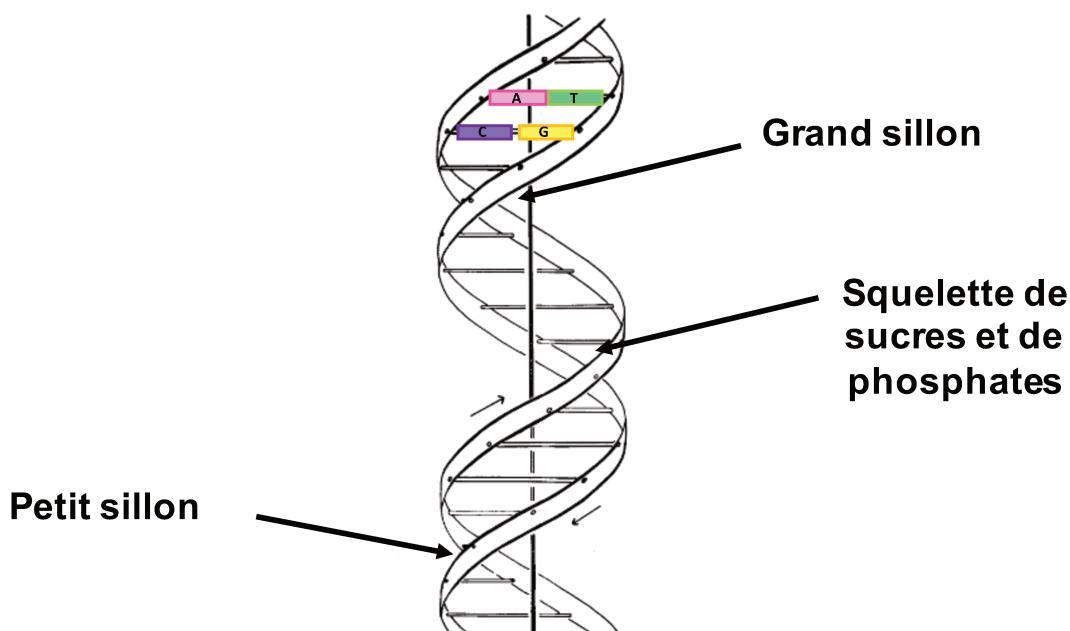


Figure 1 : Représentation de la double hélice de l'ADN

L'ADN est composé de 4 bases adénine (A), cytosine (C), guanine (G) ou thymine (T) liées à une chaîne de sucres et de phosphates qui forme le squelette sucré de l'ADN. La structure en double hélice s'articule en un grand et un petit sillon. (D'après (Watson & Crick 1953))

L'ADN peut prendre différentes conformations selon qu'il est linéaire, circulaire, et en fonction des contraintes qui agissent sur lui. Ces différents états sont appelés topologie de l'ADN. Au sein de la cellule, l'ADN change de topologies notamment pendant la majorité des processus cellulaires : la réPLICATION, la transcription, la division cellulaire, la réparation, etc... (Espeli & Marians 2004).

Cette topologie est définie par trois paramètres : le nombre de liens topologiques ou indice d'enlacement Lk, la torsion Tw et le surenroulement ou vrillage Wr. L'indice d'enlacement Lk correspond au nombre de fois où un brin de la double hélice tourne autour de l'autre. La torsion Tw est égale au nombre total de paires de bases divisé par le nombre de paires de bases par tour d'hélice. Le surenroulement Wr exprime le nombre de croisements de l'axe de la double hélice d'ADN autour d'elle-même pour former un supertour. Les supertours peuvent être négatifs ou positifs selon la polarité du croisement de brins. Ces trois paramètres sont reliés par l'équation : $Lk = Tw + Wr$ (Deweese et al. 2008) (Figure 2).

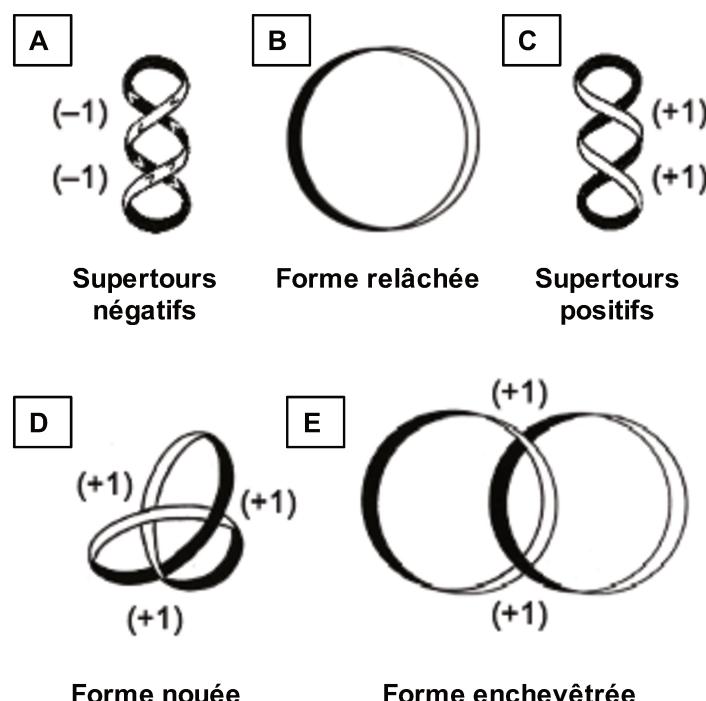


Figure 2 : Schéma représentant différentes topologies que l'ADN peut adopter

A) Supertours négatifs, le signe – correspond à une polarité droite dans le croisement des brins. B) La forme relâchée correspond à la structure sans contrainte de l'ADN. C) Supertours positifs, le signe + correspond à une polarité gauche dans le croisement des brins. D) Forme nouée, le sens des croisements entre les brins peut être positif ou négatif. E) Forme enchevêtrée ou caténée entre deux molécules d'ADN, le sens des croisements entre les brins peut être positif ou négatif.

2) Etude de la topologie de l'ADN grâce à des protéines ou molécules chimiques

L'étude de la topologie de l'ADN est un vrai enjeu dans la compréhension des mécanismes cellulaires. Dans la cellule, les acteurs agissant sur les contraintes topologiques sont les topoïsomérases (cf partie II). La majorité des interactions ADN-protéines est influencée

par la topologie de l'ADN. Par exemple, dans le cas des nucléosomes, ils révèlent une sensibilité à la torsion de l'ADN (Seol & Neuman 2016). Il est donc important d'avoir des outils, sondes chimiques ou protéiques, pour détecter des changements de structure au sein de l'ADN.

Afin de mieux appréhender les modifications de structure lors du changement de topologie, *in vitro*, certaines molécules chimiques ou protéines permettent la détection de zones de simple ou de double-brin. C'est notamment, le cas pour la DNaseI qui coupe préférentiellement l'ADN double-brin et les zones flexibles de l'ADN (Heddi et al. 2010) ou encore la nucléase SI qui reconnaît en particulier des zones simple-brin (Kohwi-Shigematsu et al. 1983). Des sondes chimiques peuvent aussi être utilisées, tels que le bromoacétaldehyde reconnaissant des bases A et C désappariées (Kohwi-Shigematsu et al. 1983) ou encore le diméthyl sulfate et le permanganate de potassium (Borowiec et al. 1987).

L'étude de la topologie des acides nucléiques en présence ou non de protéines ou de molécules qui l'influencent se fait grâce à des techniques telles que l'électrophorèse, la microscopie électronique (Stros et al. 2007), la microscopie à force atomique (AFM) (Qian et al. 2015) ou encore la cristallographie (Parkinson et al. 2002) (Chenoweth & Dervan 2009). Des études en « molécule unique » utilisent des substrats linéaires avec des billes aimantées à leurs extrémités. Les aimants permettent d'exercer des contraintes plus au moins fortes. Les déformations de l'ADN par exemple en présence de topoisomérases sont alors quantifiées (Charvin et al. 2005) (Seol et al. 2015).

II. Les protéines qui agissent sur la topologie de l'ADN : les topoisomérases

Les topoisomérases ont été identifiées comme étant les acteurs clés pour résoudre les problèmes de topologie. Ces enzymes furent découvertes pour la première fois découvertes chez *E. coli* dans les années 1970 (Wang, 1970). Il a été démontré que les topoisomérases sont impliquées dans la réPLICATION de l'ADN, la transcription, la recombinaison, la ségrégation des chromosomes et l'instabilité génétique (Pommier et al. 2016). Les fonctions des topoisomérases ne seront pas détaillées plus amplement dans ma thèse.

Chez l'Homme, six topoisomérases sont recensées : la hTopoI (famille IB), la hTopoImit (famille IB), la hTopo3 α (famille IA), la hTopo3 β (famille IA), la hTopo2 α (famille 2A), la hTopo2 β (famille 2A).

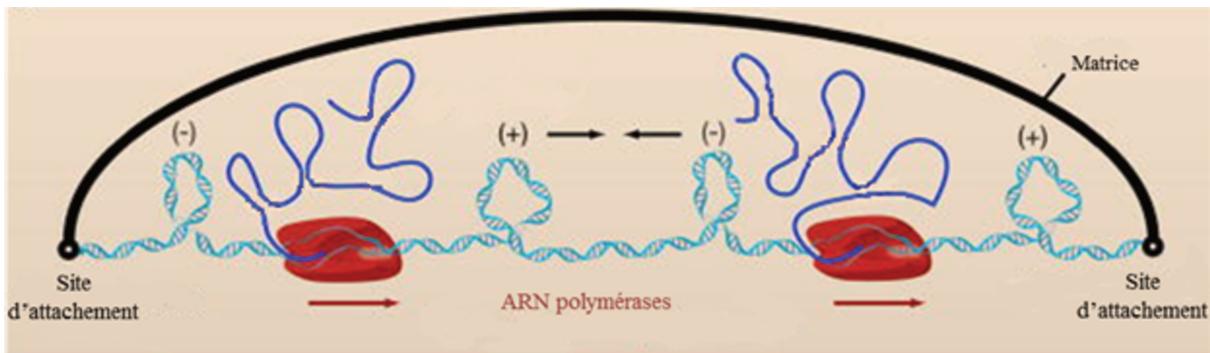


Figure 3 : Contraintes topologiques subies par l'ADN lors de la transcription

Lorsque l'ARN polymérase transcrit un ADN, des supertours négatifs et positifs s'établissent de part et d'autre de l'enzyme. L'intervention des topoisomérasées est nécessaire pour les éliminer. Le signe associé à un supertour est fonction de la polarité du croisement des brins. Pour un supertour négatif, le croisement de l'axe de la double hélice est de polarité droite. Pour un supertour positif, il est de polarité gauche. (D'après (Koster et al. 2010)).

Par exemple, au cours de la transcription, des superenroulements positifs de l'ADN s'accumulent en aval de l'ARN polymérase et des superenroulements négatifs en amont (Figure 3). Les topoisomérasées sont nécessaires pour enlever ces contraintes et permettre au complexe transcriptionnel de progresser le long de l'ADN. De même, lors de la réplication, outre les superenroulements positifs qui s'accumulent à l'avant de la fourche comme dans le cas de la transcription, les brins fils nouvellement synthétisés peuvent s'entremêler. Les topoisomérasées peuvent donc agir sur différents types de substrats. De façon générale, elles peuvent relaxer l'ADN, c'est à dire éliminer les supertours ou décatener des molécules d'ADN (Figure 4).

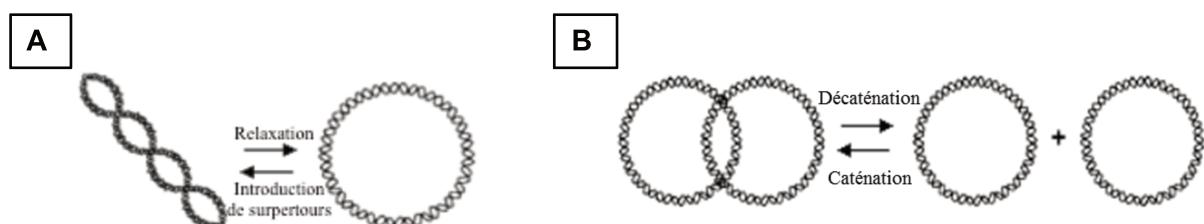


Figure 4 : Schéma présentant l'activité de relaxation ou d'introduction de supertours et de caténation/décaténation de deux molécules d'ADN sœurs

A) La forme superenroulée (à gauche) peut passer dans sa forme relâchée (à droite) grâce à l'action des topoisomérasées, ce qui correspond à l'activité de relaxation des topoisomérasées. La réaction inverse de la réaction de relaxation correspond à l'introduction de supertours dans une molécule relâchée. B) Les deux molécules d'ADN double-brin caténées (à gauche) sont séparées (à droite) grâce à l'activité de décaténation des topoisomérasées. (D'après (Nitiss, 2001)).

1) Présentation des différentes familles de topoisomérases

Pour résoudre ces contraintes topologiques au sein de l'ADN, les topoisomérases possèdent une activité enzymatique qui se décompose en trois grandes étapes : 1- Clivage transitoire de l'ADN avec établissement d'un lien covalent entre l'enzyme et une des extrémités de l'ADN, 2- Changement de topologie, 3- Religation de l'ADN (Champoux 2001). Les topoisomérases coupent l'ADN grâce à une attaque nucléophile. Celle-ci est réalisée par l'oxygène de la tyrosine catalytique de l'enzyme sur un groupement phosphate du squelette de l'ADN. Il se forme ainsi une liaison phosphotyrosyle entre l'enzyme et l'ADN (Figure 5). Après le changement de topologie et afin de refermer la coupure dans l'ADN, une attaque nucléophile se fait à nouveau. Elle est réalisée, cette fois-ci, par l'oxygène de l'extrémité du brin clivé qui attaque la liaison phosphotyrosyle du complexe covalent topoisomérase-ADN (Figure 5B).

Bien que toutes les topoisomérases fonctionnent selon un mécanisme en 3 étapes : (incision, changement de topologie, religation), elles sont séparées en familles et en sous-familles selon plusieurs critères :

1. le nombre de brins d'ADN qu'elles clivent (un ou deux),
2. le mécanisme de changement de topologie (mécanisme de passage de brin au travers de la cassure ou de rotation contrôlée d'une partie du duplexe d'ADN autour de la liaison phosphodiester intacte opposée à la cassure),
3. l'extrémité de l'ADN qui établit le lien covalent avec la tyrosine catalytique (3' ou 5'),
4. la dépendance en Mg^{2+} et en ATP des enzymes
5. les homologies de séquences.

Les topoisomérases de classe I coupent un seul brin de l'ADN et n'utilisent pas d'ATP pour modifier la topologie de l'ADN. Grâce à des comparaisons de mécanismes d'action et de séquence, elles ont été réparties en 3 sous-familles : IA, IB et IC. Les topoisomérases de la famille IA sont dépendantes du magnésium (Mg^{2+}). Elles se lient à des régions simple-brin de l'ADN double-brin (elles ont en effet peu d'affinité pour l'ADN double-brin). Après coupure du brin d'ADN par l'attaque nucléophile de la tyrosine catalytique de l'enzyme, elles établissent un lien covalent avec l'extrémité 5' du brin clivé (Figure 5). Le second brin de l'ADN peut alors passer à travers la brèche préalablement créée. C'est donc par un mécanisme de passage de brin que la topologie de l'ADN est modifiée. Après le passage de brin, la religation met en

jeu l'extrémité 3'OH libre du brin coupé, qui, grâce à une attaque nucléophile, rompt la liaison covalente enzyme-ADN et permet la fermeture de l'ADN (Figure 5).

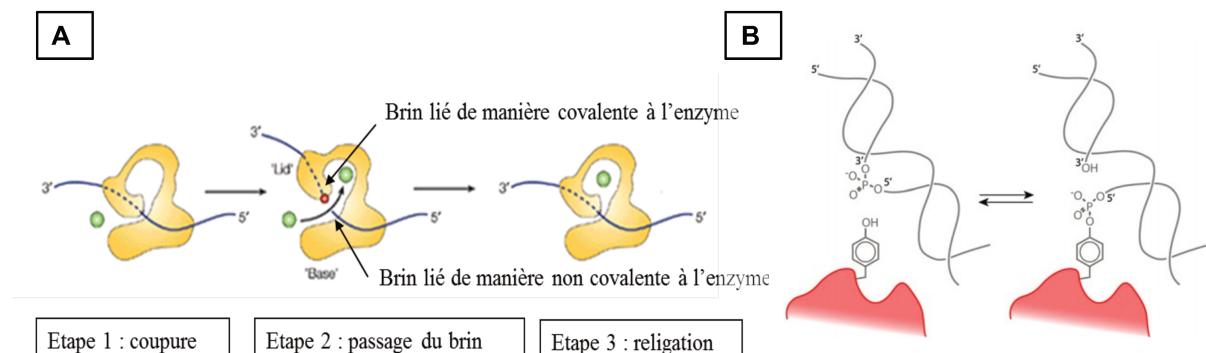


Figure 5 : Schéma présentant le mécanisme de relaxation de l'ADN grâce à une réaction de trans-estérification par les topoisomérases de la famille IA

A) Après rupture transitoire d'un brin d'ADN (trait bleu), l'extrémité 5' du brin d'ADN clivé se lie de manière covalente au groupe tyrosyl (cercle rouge) situé dans un sous-domaine protéique appelé couvercle ou 'lid', et l'extrémité 3' interagit de manière non covalente avec le socle ou 'base' de l'enzyme. En soulevant le 'lid' loin de la 'base', l'enzyme ouvre une brèche permettant le passage d'un autre brin (cercle vert). Enfin, le brin d'ADN clivé est religué et le complexe est dissocié. (Adaptée d'après (Wang 2002)).
B) La partie coloriée en rouge représente la protéine. La chaîne latérale de la tyrosine catalytique ainsi qu'un groupement phosphate de la molécule d'ADN sont montrés. Dans cet exemple, un lien covalent est établi entre l'extrémité 5' de l'ADN et l'oxygène de la tyrosine catalytique. (D'après (Viard & de la Tour 2007)).

Ce mécanisme permet aux topoisomérases de la famille IA d'éliminer les supertours négatifs (réaction de relaxation, Figure 4A) avec un pas égal à 1 ($\Delta Lk = 1$) et de décatener deux molécules d'ADN (Figure 4B) si l'une d'elles contient une région d'ADN simple-brin (présence d'une brèche ou d'une cassure simple-brin).

Plus récemment, il a été montré par des tests *in vitro* que les topoisomérases IA possèdent une activité topoisomérase ARN (Ahmad et al. 2016). Par exemple, chez l'Homme, la hTopo3 β peut agir sur les noeuds ARN et sur les R-Loop, ce qui lui donne un rôle lors de la transcription et de la traduction.

Les topoisomérases de type IB sont des enzymes qui modifient la topologie d'un ADN par un mécanisme de rotation contrôlée indépendant du Mg $^{2+}$ (Figure 6). Ces enzymes se lient préférentiellement à l'ADN double-brin. Elles clivent un des deux brins de l'ADN et maintiennent le brin clivé de manière covalente par l'extrémité 3'. La partie du duplex d'ADN non liée de manière covalente à l'enzyme peut alors tourner une ou plusieurs fois autour de la liaison phosphodiester intacte opposée à la cassure. Ce mécanisme est appelé rotation contrôlée.

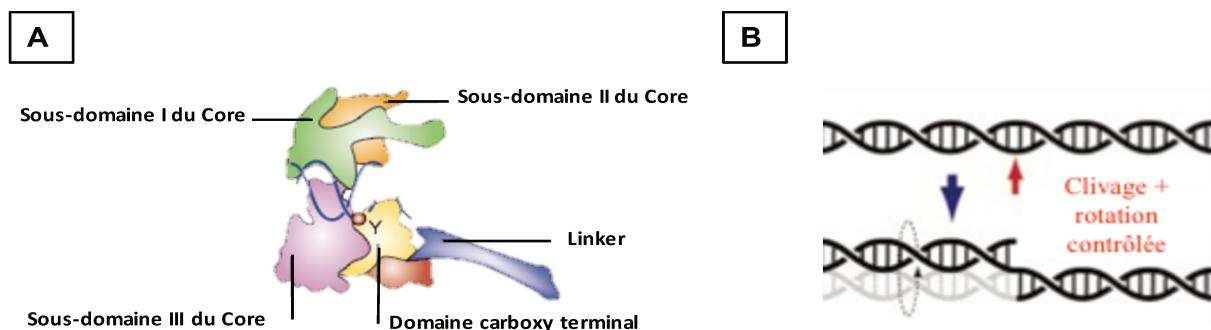


Figure 6 : Schéma représentant le mécanisme de rotation contrôlée utilisé par les topoisomérases de type IB

A) Structure schématique de la topoisomérase IB. Elle est composée d'un domaine core qui contient 3 sous-domaines, un linker et un domaine C terminal. D'après (Wang 2002). B) Le schéma résume le mécanisme de rotation contrôlée utilisé par les topoisomérases de la famille IB. Un des deux brins de l'ADN est clivé par l'enzyme. Le brin clivé est maintenu par l'enzyme. Le duplex d'ADN non maintenu de manière covalente par l'enzyme peut ainsi faire plusieurs rotations sur lui-même afin d'éliminer des supertours négatifs et positifs. (D'après (Corbett & Berger 2004)).

Les topoisomérases IB peuvent éliminer les supertours négatifs et positifs (Figure 4A) avec un pas égal à 1 ($\Delta Lk = 1$). L'Homme possède deux enzymes de cette famille, appelées topoisomérase I humaine (hTopoI) et une forme mitochondriale (hTopoImito).

La famille IC ne compte à ce jour qu'un seul représentant, la topoisomérase V, appartenant au genre *Methanopyrus* des Archea. Cette enzyme a été classée dans une nouvelle famille car elle diffère par sa séquence. D'un point de vue mécanistique, elle est plus proche des enzymes de la famille IB, car elle change de topologie par un mécanisme de rotation contrôlée, crée un lien covalent transitoire avec l'extrémité 3' de l'ADN, et présente une activité indépendante du Mg^{2+} .

À la différence des topoisomérases de la famille I, les topoisomérases de classe II coupent chaque brin de l'ADN double-brin créant une cassure double-brin transitoire dans l'ADN. Leur activité est dépendante de l'ATP et du Mg^{2+} . Elles utilisent un mécanisme de passage de brin pour changer la topologie d'un ADN. La coupure double-brin de l'ADN créée par la topoisomérase permet le passage d'une seconde molécule d'ADN double-brin à travers cette brèche. Travailant sur de l'ADN double-brin, ces topoisomérases modifient l'indice d'enlacement d'un ADN avec un pas égal à 2 ($\Delta Lk = 2$). Ces enzymes sont spécialisées dans la décaténation des molécules d'ADN (Figure 4B). La sous-famille IIA des topoisomérases est

composée des topoisomérases II bactériennes et humaines alors que la sous-famille IIB regroupe uniquement une topoisomérase archéenne.

2) Les topoisomérases comme cible thérapeutique

Les êtres vivants nécessitent la présence d'au moins une topoisomérase pour être viables. Par conséquent, cette protéine est une cible thérapeutique antibactérienne et antivirale, une topoisomérase étant aussi présente dans la famille des poxvirus. De plus, chez l'Homme, les topoisomérases étant impliquées dans tous les mécanismes cellulaires, elles ont été rapidement ciblées par les stratégies anticancéreuses (Ashour et al. 2015).

Dans la catégorie des anticancéreux, la camptothécine et ses dérivés sont utilisés comme anticancéreux, ils ciblent la topoisomérase IB tandis que les étoposides, anthracyclines (doxorubicine, daunorubicine), et le mitoxantrone ciblent les topoisomérases IIA. Concernant les antibiotiques, les molécules de la famille des quinolones et aminocoumarines inhibent les topoisomérases bactériennes de type II (Pommier et al. 2010).

III. Les tests *in vitro* et leurs lacunes pour l'étude des topoisomérases

Depuis plusieurs années, des partenaires des topoisomérases ont été identifiés. C'est par exemple le cas pour la hTopo3 α qui interagit avec les protéines RMI1 et RMI2, mais leurs fonctions ne sont pas encore bien élucidées (Mankouri & Hickson 2007). Les tests *in vitro* peuvent donc apporter des réponses afin de mieux comprendre l'influence des protéines partenaires sur l'activité de la topoisomérase et/ou permettre de caractériser le mécanisme d'action de nouvelles topoisomérases (le type de topologie sur laquelle elle peut agir (supertours positifs, négatifs), capacité à décatainer, mécanisme de rotation libre ou non, etc...).

De surcroit, les topoisomérases étant des cibles thérapeutiques, de nouvelles molécules peuvent être testées dans un premier temps par des tests *in vitro*. Ces tests doivent permettre de décortiquer les mécanismes d'action de ces molécules et d'identifier par quels moyens elles inhibent la relaxation, par exemple en stabilisant le complexe covalent. En effet, bloquer le complexe covalent va créer des cassures dans l'ADN, et/ou bloquer des fourches de réPLICATION qui pourront entraîner la mort cellulaire.

Initialement, les substrats utilisés pour les études *in vitro* sont des plasmides ou des oligonucléotides afin de tester la relaxation, l'incision, la présence du complexe covalent (Depew et al. 1978). Les tests de relaxation sont réalisés sur des plasmides tandis que les tests

d'incision et de religation utilisent les oligonucléotides. La visualisation du complexe covalent a pu être obtenue grâce à des plasmides et des oligonucléotides (Tse et al. 1980). Ces études permettent, par exemple, de comparer une topoisomérase WT avec des protéines mutantes (Cheng et al. 2005).

Ces différentes méthodes ont néanmoins quelques limites. La principale lacune réside dans le fait que ces tests ne permettent pas de quantifier toutes les étapes de la relaxation en une expérience. En effet, lors de la relaxation sur plasmides, seuls les topoisomères obtenus sont observables. Concernant le complexe covalent, il a pu être visualisé par centrifugation en chlorure de césium en utilisant des plasmides mais les conditions expérimentales ne permettaient pas l'étude de la relaxation en parallèle (Champoux 1977). Dans le cas où les oligonucléotides sont utilisés, seule les étapes d'incision et/ou de religation sont analysables. Le complexe covalent peut aussi être visualisé lors de ces étapes par SDS-PAGE (Tse-Dinh 1986) (Cheng et al. 2005).

Bien qu'ils puissent être marqués et ainsi permettre la quantification de réactions biochimiques, les oligonucléotides manquent cruellement de ressemblance avec un substrat physiologique de topoisomérases car ils ne présentent aucune contrainte topologique. La relaxation est donc impossible sur ce genre de substrats, ce qui les rend contestable pour l'étude des topoisomérases. Concernant les plasmides, ils ont l'avantage de posséder une topologie, mais leurs grandes tailles présentent un manque de sensibilité. En effet, tous les topoisomères obtenus lors de la relaxation ne sont pas séparables et la visualisation du complexe covalent n'est pas possible dans les conditions physiologiques de relaxation. Le complexe covalent sur plasmides a pu être observé mais les conditions utilisées dans ces études n'étaient pas des conditions physiologiques de relaxation car une proportion très importante de forme d'ADN possèdant une brèche et de forme relâchée était présente.

À l'heure actuelle, les études *in vitro* ne disposent pas de substrats qui permettent d'observer globalement toutes les étapes du mécanisme d'action des topoisomérases et de les quantifier.

Le projet de thèse

Mon projet de thèse a eu pour finalité de mettre au point et de valider un nouveau modèle d'étude *in vitro* pour les topoisomérases.

Les objectifs de ce projet étaient :

- i) De proposer un système qui permet d'étudier et de quantifier la relaxation de topoisomérases notamment de la famille I et en même temps de pouvoir visualiser leur incision, le complexe covalent et leur religation. Pour cela, des petits cercles d'ADN de 235 bp radioactivement marqués, appelés « minicercles » ont été utilisés. La petite taille de ces cercles permet de discriminer des substrats libres ou liés à des protéines et la radioactivité permet une grande sensibilité de détection des différents espèces produites pendant la relaxation.
- ii) De fabriquer des substrats de topologie différentes qui permettent d'observer des différences d'affinités pour les topoisomérases et d'apporter des réponses sur la topologie de ces minicercles.

La première partie de ce travail proposant un nouveau modèle d'étude *in vitro* pour les topoisomérases 1 a été publiée dans la revue Scientific Report en Août 2015. Concernant la seconde partie du projet qui consistait à mieux comprendre la topologie des minicercles, j'ai initié l'étude en testant l'effet du calcium et du magnésium sur les minicercles. Le travail a été poursuivi par Emmanuelle Delagoutte et le manuscrit est actuellement en cours de préparation.

Résultats

I. Article 1 : « Use of double-stranded DNA mini-circles to characterize the covalent topoisomerase-DNA complex »

Ce travail avait pour but de proposer un modèle d'étude *in vitro* pour les topoisomérases à l'aide de minicercles d'ADN.

Nous avons utilisé des minicercles de 235 bp avec un Lk relatif de 0 à -6. Ils nous ont permis d'étudier l'activité de relaxation de deux topoisomérases commerciales, la topoisomérase IA d'*E. coli*, la EcTopoI et la topoisomérase IB de thymus de veau, la ctTopoI sur les 7 topoisomères obtenus. Ces substrats se sont révélés être bien adaptés aux topoisomérases puisque nous avons pu observer et quantifier des différences de produits de relaxation finale pour ces deux enzymes. Ces différences sont corrélées au fait que ces deux topoisomérases appartiennent à des familles différentes et ne possèdent pas le même mécanisme d'action. Nos différentes conditions de gels et d'électrophorèse nous ont aidé à visualiser l'étape de complexe covalent lors de la réaction de relaxation. L'utilisation de camptothécine, inhibiteur de topoisomérase IB, a montré un enrichissement en complexe covalent pour la ctTopoI, ce qui renforce l'intérêt de notre modèle pour tester des molécules de type « poison », c'est à dire qui stabilise le complexe covalent.

SCIENTIFIC REPORTS



OPEN

Use of double-stranded DNA mini-circles to characterize the covalent topoisomerase-DNA complex

Received: 26 February 2015

Accepted: 24 June 2015

Published: 24 August 2015

Armél Millet, François Strauss & Emmanuelle Delagoutte

The enzymatic DNA relaxation requires the DNA to be transiently nicked and rejoined, the covalent topoisomerase-DNA complex being a key intermediate of the nicking-joining reaction. Practically, this reaction is most often characterized by oligonucleotides. However, the incision-religation of an oligonucleotide does not fully recapitulate the incision-religation occurring during relaxation and the preferred substrate for such reaction characterization is supercoiled DNA. We therefore developed a method that used radiolabeled supercoiled DNA mini-circles to characterize the covalent enzyme-DNA complex formed during a relaxation reaction. Resolution of the relaxation products under different conditions permitted to quantify the proportion of covalent complex formed during the relaxation catalyzed by two topoisomerase models, the *Escherichia coli* topoisomerase I and the calf thymus topoisomerase I. As expected, the covalent complex formed with the calf thymus topoisomerase I was significantly enriched by camptothecin, a widely-used inhibitor of this topoisomerase, and a salt jump permitted the multiple topoisomerases trapped per mini-circle to complete the reaction cycle. The identified positions of the camptothecin-induced incision sites were shown to be independent of the linking number and the substrate circular nature. Overall, our results demonstrate that supercoiled mini-circles constitute a powerful and polyvalent substrate to characterize the mechanism of action of novel topoisomerases and inhibitors, including the incision-religation reaction.

Topoisomerases are essential enzymes whose function is to regulate cellular DNA supercoiling and to solve topological problems associated with the manipulation of DNA (for reviews on DNA topoisomerases see^{1–6}). The replication, transcription, repair, recombination or chromatin remodeling machineries indeed modify the DNA topology locally inducing positive or negative supercoils and create intermediates of complex topology, such as knots, catenanes or hemicatenanes that can become toxic if left unresolved. To modulate the topology of the DNA, topoisomerases start by nicking the sugar-phosphate backbone of the DNA, establishing a covalent bond with the DNA. From this nicked intermediate, either a strand passage takes place with a DNA strand passing through the break or a controlled rotation occurs with the rotation of the DNA duplex not covalently attached to the enzyme around the phosphodiester bond opposite the nick. At the end of these processes, the DNA backbone is resealed, the enzyme released, and the DNA global topology has been modified.

The classification of topoisomerases into families is based on the number of strands of the duplex DNA that are broken during the enzymatic reaction. Topoisomerases of family I break one strand of the duplex DNA whereas topoisomerases of family II introduce two breaks, one on each strand of the duplex DNA. Both families have been further divided into subfamilies (IA, IB, IC and IIA and IIB), depending on the polarity of the DNA end that is linked to the topoisomerase, the mechanism used to relax DNA, and the primary sequence of the protein. Topoisomerases of family IA, IIA and IIB change the topology by a strand passage mechanism whereas those of family IB and IC use a controlled rotation mechanism.

Structure et Instabilité des Génomes, Sorbonne Universités, Muséum national d'Histoire naturelle, Inserm U 1154, CNRS UMR 7196; CP26, 57 rue Cuvier 75005 Paris, France. Correspondence and requests for materials should be addressed to E.D. (email: emmanuelle.delagoutte@mnhn.fr)

The covalent bond between the topoisomerase and the DNA is established between the lateral chain of the catalytic tyrosine in the protein and the 3'-phosphate (in the case of topoisomerases IB and IC) or the 5'-phosphate (in the case of topoisomerases IA and II) of the DNA. The resulting covalent complex is a short-lived species but its half-life time, which depends directly on the relative rates of the nicking-closing reactions, can be altered by drugs, such as camptothecin that decreases the joining rate^{7,8} or by protein partners of topoisomerases as suggested for the RecQ Mediated Instability 1 (RMI1) protein partner of the human topoisomerase III alpha⁹ and the yeast RMI1 homolog¹⁰. For the cell, the consequences of altering the relative rates of the nicking and joining reactions of a topoisomerase can be dramatic as exemplified by the cytotoxic effect of camptothecin or its derivatives (for reviews on the IB family of topoisomerase inhibitors see^{11,12}).

Several types of substrates have been used to trap and characterize the covalent complex. The method based on plasmids relied on the fact that a nucleoprotein complex exhibited a lower buoyant density than free DNA when sedimented at equilibrium on an alkaline cesium chloride gradient^{13,14}. However, the conditions of relaxation were not optimal for relaxation since they promoted the accumulation of nicked DNA instead of relaxed plasmids as final products^{13,14}. Other pioneer assays using ss DNA of various lengths permitted to characterize the covalent complex formed with type IA topoisomerases^{15–17}. Recently, procedures relying on the use of oligonucleotides have been developed to elucidate the mechanism of action of topoisomerase drugs and to facilitate comparison of the efficiency of different drugs in stabilizing the covalent complex¹⁸. Over the past decade, mutations of the *Yersinia pestis* and *Escherichia coli* Topoisomerase I have been isolated¹⁹, and the stabilization of the covalent complex by these point mutations has been inferred from the high level of cleaved oligonucleotides generated by these mutants^{20–22}.

Whereas all these approaches using oligonucleotides are undoubtedly very valuable and have been very helpful in elucidating the mechanism of action of topoisomerases and their inhibitors, it is nevertheless worth noticing that the incision and religation of oligonucleotides do not fully recapitulate the nicking and closing reactions that occur during a relaxation reaction since no strand passage or controlled rotation takes place with such substrates. It may indeed well be that the relative kinetic of the incision and religation reactions be modulated by the strand passage step in the case of the topoisomerases of family IA and II or the swiveling step in the case of the topoisomerases of family IB. Our goal was therefore to develop a quantitative method that used the preferred and natural substrate of the nicking-joining reaction, i.e. supercoiled DNA, to characterize these two coupled steps of the topoisomerase-catalyzed relaxation reaction. Such method should also permit to estimate the proportion of covalent complex during a relaxation reaction, since topoisomerase mutations, topoisomerase protein partners and topoisomerase inhibitors may alter this proportion. As an alternative to plasmids, double-stranded DNA mini-circles (dsMCs) might represent the ideal relaxation substrate for such a study due first, to their small size permitting the resolution of all relaxation intermediates, including the covalent complex, and second, to the possibility of radiolabelling them, thus conferring a high detection sensitivity and a reliable quantification aspect to the method. We therefore prepared dsMCs of 235 base pairs (bps) and used them to characterize the covalent complex of two topoisomerase models, the *Escherichia coli* TopoI (*Ec* TopoI) of the IA family of topoisomerases and the calf thymus TopoI (*Ct* TopoI) of the IB family. Altogether, our results demonstrate that dsMCs are usable to resolve and isolate the covalent topoisomerase-DNA complex and therefore suitable to characterize the nicking-joining reaction of any DNA topoisomerase under various conditions (in the presence of protein partners or inhibitors, with mutant topoisomerases, etc.) Our method represents an improvement in the field of DNA topoisomerases since the coupled nicking-joining reaction is examined with the preferred and natural substrate of relaxation.

Results

Relaxation activity of the Ct TopoI and Ec TopoI on dsMCs. Seven topoisomers with relative linking numbers ranging from 0 to -6 (labeled in what follows T_i where $i =$ relative linking number) were prepared (Supplementary Fig. 1 and Online Methods) and used to accurately assay the relaxation activity of two topoisomerase models, the Ct TopoI and the Ec TopoI (Supplementary Figs. 2, 3). The Ct TopoI and Ec TopoI efficiently relaxed the topoisomers with relative linking numbers ≤ -2 . The final reaction products were T_0 and T_{+1} for Ct TopoI and T_{-1} for Ec TopoI. The Ec TopoI had no activity on T_{-1} and T_0 whereas the Ct TopoI had a strong and moderate activity on T_{-1} and T_0 , respectively.

Visualization of a covalent complex during the relaxation of dsMCs by the Ct TopoI. The relaxation reaction involves the formation of a covalent topoisomerase-DNA complex. This intermediate is a short-lived species but during the incision of an oligonucleotide or longer ss DNA it can nevertheless be irreversibly trapped by adding SDS^{15,16}. Since the incision of an oligonucleotide is not a perfect mimic of the incision reaction taking place during plasmid relaxation, we investigated whether the covalent complex could also be trapped during the relaxation of dsMCs. The relaxation of dsMCs by the Ct TopoI was then performed at 37°C over a 10 min period, after which SDS was added to trap the covalent complex and stop the reaction. The sample was next split into two, with one half being treated with proteinase K, the other half not. Both samples were then analyzed by gel electrophoresis under several gel conditions.

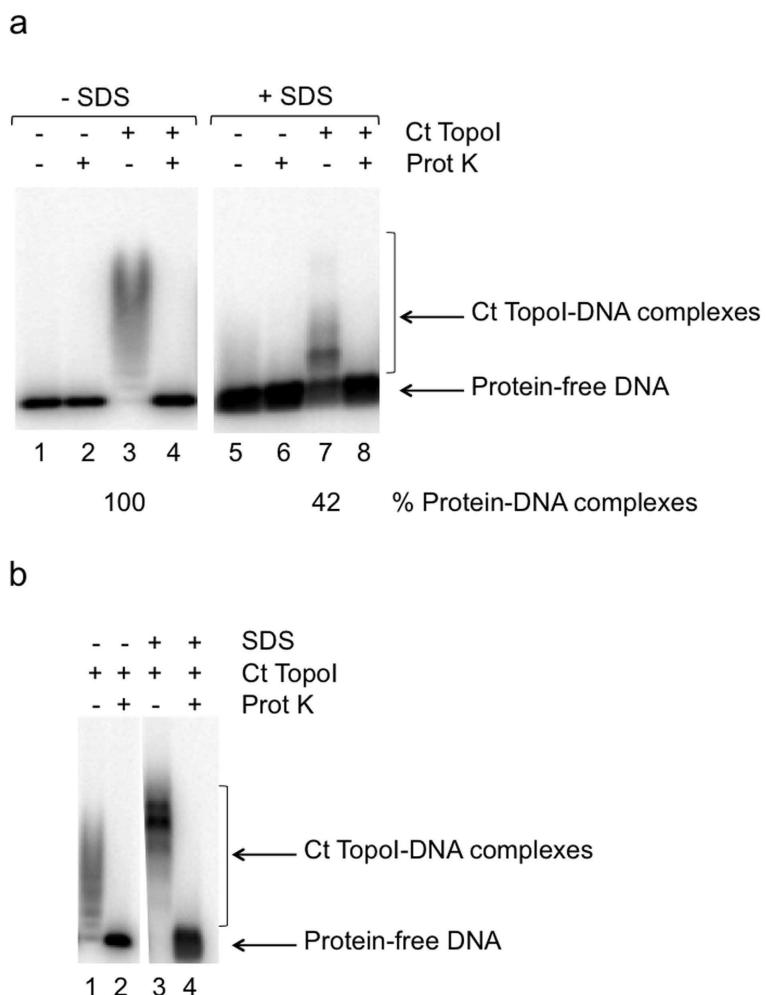


Figure 1. Relaxation of a mixture of the topoisomers T₋₄ and T₋₅ analyzed on an agarose gel under native conditions. (a) Relaxation was performed with the Ct TopoI at a concentration of $18.75 \text{ mU} \mu\text{L}^{-1}$ ($\approx 0.2 \text{ nM}$) as described in the methods section. After 10 minutes of incubation at 37°C , SDS was added to the reaction mixtures to irreversibly trap the Ct TopoI covalently bound to DNA (lanes 5–8). After the SDS treatment, the samples were treated with proteinase K (Prot K) to confirm the presence of proteins in the retarded bands (lanes 2, 4, 6 and 8). The % of proteins engaged in a complex with dsMCs is indicated; (b) Relaxation was performed with the Ct TopoI at a concentration of $18.75 \text{ mU} \mu\text{L}^{-1}$ ($\approx 0.2 \text{ nM}$) in the presence of camptothecin (final concentration, $10 \mu\text{M}$) as described in the methods section. After 10 minutes of incubation at 37°C , SDS was added to the reaction mixtures to irreversibly trap the Ct TopoI covalently bound to DNA (lanes 3 and 4). After the SDS treatment, the samples were treated with proteinase K (Prot K) to confirm the presence of proteins in the retarded bands (lanes 2 and 4).

Complexes with Ct TopoI covalently linked to DNA could be detected on an agarose gel under native conditions (Fig. 1a, lane 7). The efficiency of the SDS treatment to trap the protein covalently linked to dsMCs was estimated by comparing the percentage of protein-DNA complexes formed in the presence or absence of SDS. The results indicate that SDS could convert nearly 40% of the Ct TopoI bound to DNA into a covalent complex (Fig. 1a, compare lanes 3 and 7).

As the agarose gel did not resolve the different protein-free DNA species (topoisomers and nicked dsMCs), the reaction products were analyzed on polyacrylamide gels. In theory, the proteinase K treatment should make the radiolabeled species corresponding to the covalent complex disappear. A concomitant enrichment of either (i) linear 235 nucleotide (nt)-long DNA if the sample is fully denatured before electrophoresis, or (ii) nicked plasmid if the sample is not denatured, should be measured. SDS-PAGE revealed radiolabeled material that was sensitive to proteinase K treatment (Fig. 2, compare lanes 8 and 9, arrows 1–2) and a noticeable enrichment of linear DNA upon proteinase K treatment (Fig. 2, compare lanes 8 and 9, arrow labeled “lin.”). Among the two radiolabeled bands sensitive to proteinase K treatment, one of them had an apparent molecular mass close to that of the Ct TopoI-DNA covalent complex

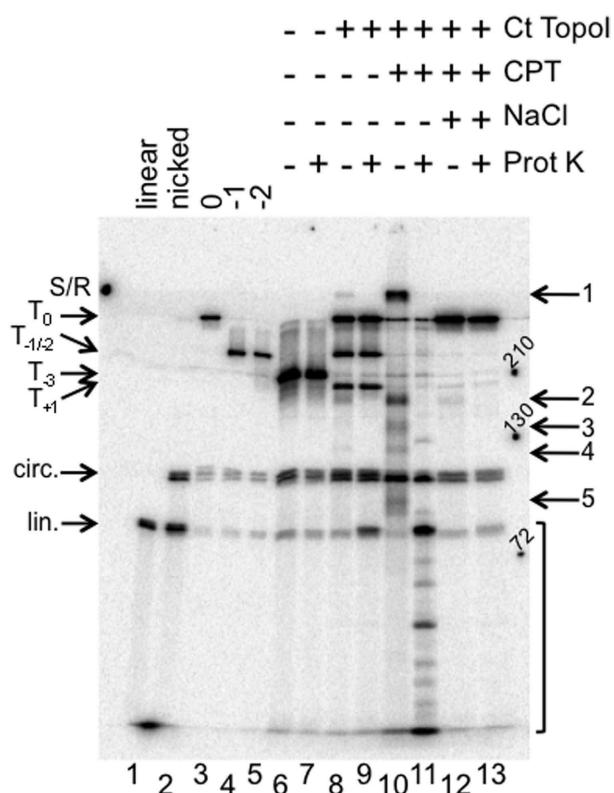


Figure 2. Visualization of the covalent Ct TopoI-DNA complex. Relaxation assay analyzed by SDS-PAGE. The relaxation of topoisomer T_{-3} by the Ct TopoI (at a final concentration, $18.75 \text{ mU} \mu\text{L}^{-1}$ ($\approx 0.2 \text{ nM}$)) was performed in the absence (lanes 8 and 9) or in the presence of camptothecin (CPT) at a final concentration of $10 \mu\text{M}$ (lanes 10–13) as described in the methods section. After 10 minutes of incubation at 37°C , SDS was added to the reaction mixtures to irreversibly trap the Ct TopoI covalently bound to DNA (lanes 8–13). When indicated, the samples were next treated with proteinase K (Prot K) (lanes 9, 11 and 13). Lanes 12 and 13, NaCl was added to the samples prior to SDS and proteinase K treatments. Lane 1, linear 235 nt-long fragment; lane 2, denatured nicked dsMCs; lane 3, topoisomer T_0 ; lane 4, topoisomer T_{-1} ; lane 5, topoisomer T_{-2} ; lanes 6 and 7, samples were treated as in lanes 8 and 9, respectively, except that no enzyme was added. S/R, limit between the resolving and the stacking gel. lin., linear. cir., circular. Arrows 1, 2, 3, 4 and 5 on the right side of the gel point to proteinase K-sensitive material. The bracket on the right side of the gel represents DNA species that migrate faster than the linear 235 nt-long fragment. Numbers 210, 130 and 72 on the right side of the gel correspond to the molecular mass of proteins of 210, 130 and 72 kDa, respectively.

(Fig. 2, lane 8, arrow 2, $\approx 170 \text{ kDa}$). The other species (Fig. 2, lane 8, arrow 1) was trapped at the interface between the resolving and the stacking gel (S/R). When samples were analyzed on chloroquine-containing polyacrylamide gels under native (Supplementary Fig. 4a) or denaturing (Supplementary Fig. 4b) conditions, a similar enrichment of nicked DNA (Supplementary Fig. 4a, compare lanes 7 and 8) or linear 235 nt-long DNA (Supplementary Fig. 4b, compare lanes 7 and 8) was observed. Under such gel conditions, the high molecular weight species that was sensitive to proteinase K treatment (Supplementary Fig. 4a, b, arrow 1) could not penetrate into the gel. The results of these experiments demonstrate that a small amount of the Ct TopoI covalently linked to DNA could be detected during the relaxation of dsMCs.

Enrichment in the covalent complex in the presence of camptothecin. To confirm that the covalent complex isolated during the relaxation of dsMCs was indeed the covalent complex involved in the reaction cycle of topoisomerases, we characterized the proteinase K-sensitive species in the presence of camptothecin, a topoisomerase drug known to inhibit relaxation by slowing down the rejoicing reaction and thus stabilizing the covalent complex^{11,12}. Trapping of the Ct TopoI in a covalent complex with dsMCs in the presence of camptothecin could be visualized on an agarose gel under native conditions (Fig. 1b). This trapping reaction was highly efficient as all DNA became engaged in a covalent complex (Fig. 1b, lane 3). Furthermore, SDS-PAGE revealed first, that the relaxation was efficiently inhibited by camptothecin as T_0 and T_{+1} represented less than 10% of the radiolabeled species (Fig. 2, lane 10).

Second, among the radiolabeled material sensitive to proteinase K treatment (Fig. 2, compare lanes 10 and 11, arrows 1–5) two of them had the same mobility as those detected in the absence of camptothecin and were clearly enriched by the presence of camptothecin (Fig. 2, lanes 8 and 10, arrows 1 and 2). Third, both linear 235 nt-long DNA and faster migrating species were significantly enriched upon proteinase K treatment (Fig. 2, lane 11, bracket). Fourth, when samples were analyzed on chloroquine-containing polyacrylamide gels, a high molecular weight species sensitive to proteinase K was also detected (Supplementary Fig. 4a, b, lane 9, arrow 1). After proteinase K treatment this species gave rise to either nicked DNA when the sample was not denatured prior to electrophoresis (Supplementary Fig. 4a, lane 10) or DNA \leq 235 nts when the sample was denatured (Supplementary Fig. 4b, lane 10, bracket). These results demonstrate that the proteinase K-sensitive material isolated during our experimental conditions corresponds well to the functional covalent topoisomerase-DNA complex.

Effect of salt on the covalent complex trapped on dsMCs by camptothecin. It is well established that a salt jump can promote DNA religation within the covalent complex before complete dissociation of the protein from the DNA. To establish that the intermediate covalently stabilized on the dsMCs by camptothecin was competent for religation following an increase in salt concentration, relaxation of dsMCs by the Ct TopoI was performed in the presence of camptothecin before successively adding NaCl and finally SDS. The results indicate that the covalent complex stabilized on the dsMCs by camptothecin was indeed competent for religation since T_0 was the major product of the reaction (Fig. 2, lane 12; Supplementary Fig. 4a, b, lane 11). Its almost complete disappearance agreed well with its capability to complete the reaction cycle.

Covalent linking of multiple Ct TopoI per strand of dsMC by camptothecin. Electrophoresis on a sequencing gel was used to investigate the size of the material that migrated faster than the 235 nt-long linear DNA fragment when relaxation was performed in the presence of camptothecin (Fig. 2, lane 11, bracket; Supplementary Fig. 4b, lane 10, bracket). This analysis revealed that DNA fragments of \approx 45 to 235 nts were produced (Fig. 3a, lane 3). Such fragments might originate from the covalent binding of several Ct TopoI molecules per strand of dsMC. To test this hypothesis, we titrated camptothecin during a relaxation reaction. Analysis on an agarose gel under native conditions demonstrated that up to five distinct complexes with Ct TopoI covalently bound to DNA could be resolved (Fig. 3b), confirming the possibility of covalently linking multiple topoisomerases per dsMC. The increase of the level of incision of the dsMCs as the concentration of the camptothecin increased (Fig. 3c) was also consistent with this possibility. The result that more than one Ct TopoI could be covalently attached per dsMC explains the formation of covalent complexes with an apparent molecular mass smaller than the predicted molecular mass of this complex formed with a 235 nt-long DNA fragment (Fig. 2, arrows 3–5).

Specificity of the pattern of the camptothecin-induced incision of the dsMCs. The radio-labeled nature of the dsMCs made it possible to identify the DNA sequences where the Ct TopoI had established a covalent bond with the DNA during relaxation in the presence of camptothecin. As both strands of the dsMCs were radiolabeled, after the relaxation reaction, the reaction products needed to be submitted to a double restriction enzyme digestion (BamHI + BglII or BamHI + HindIII) to visualize the nicks on one of the two strands (Supplementary Fig. 5). Furthermore, to unambiguously identify the strand that was incised, the pattern of incision obtained after the double digestion was compared with that obtained after the single BamHI digestion. The resistance or sensitivity of a specific band to digestion with the second enzyme (BglII or HindIII) indeed indicated the strand that did or did not carry the nick (Supplementary Fig. 5). Once the strand that carried the nick had been identified, the site of incision was positioned on the sequence by comparison with the sequencing lanes. We first analysed the incision profile of the Ct TopoI in the presence of camptothecin with the BamHI-linearized form of dsMCs. 4 bands along the BamHI digestion profile (Fig. 4a, lanes 5 and 10) were resistant to the additional HindIII digestion and sensitive to the additional BglII digestion (Fig. 4a, compare lanes (5–7), and (8–10); scans comparing intensity of the bands S_1 , S_5 , and S_7 of lanes 8 and 9), indicating that these 4 incision sites were located on strand 1 of the DNA (Supplementary Fig. 5). Conversely, 8 bands along the BamHI digestion profile (Fig. 4a, lanes 5 and 10) were resistant to the additional BglII digestion and sensitive to the additional HindIII digestion (Fig. 4a, compare lanes (5–7), and (8–10); scans comparing intensity of bands S_2 , S_3 , S_4 , S_6 and S_8 of lanes 8 and 9). These 8 incision sites were thus located on strand 2 of the DNA (Supplementary Fig. 5). The disappearance of these 8 bands by HindIII digestion was not complete due to the partial HindIII enzyme digestion. However a difference in intensity of these bands between the BamHI-BglII and BamHI-HindIII double digestions was clearly measurable (Fig. 4b). The bands S_{13} and S_{14} were resistant to BglII and HindIII (Fig. 4a). As the “5'-TTC” sequence that partially overlapped the HindIII restriction site (“5'-AAGCTT”) represented a potential attachment site of the Ct TopoI²³, it is possible that the nick created by covalent binding of Ct TopoI inhibits the restriction enzyme reaction, thus making these two bands resistant to HindIII. Overall, our results indicate that in the presence of camptothecin, Ct TopoI could covalently bind the 235 bp-long DNA at 12 different positions with 8 incision sites located on strand 2, and 4 on strand 1 of the linear DNA (Fig. 4a,b; Supplementary Fig. 5). The three strong cleavage sites (sites S_4 , S_5 and S_8) matched quite well the published consensus sequence²³ since (i) two of them had a thymine at position –1 (sites S_4 and S_5), (ii)

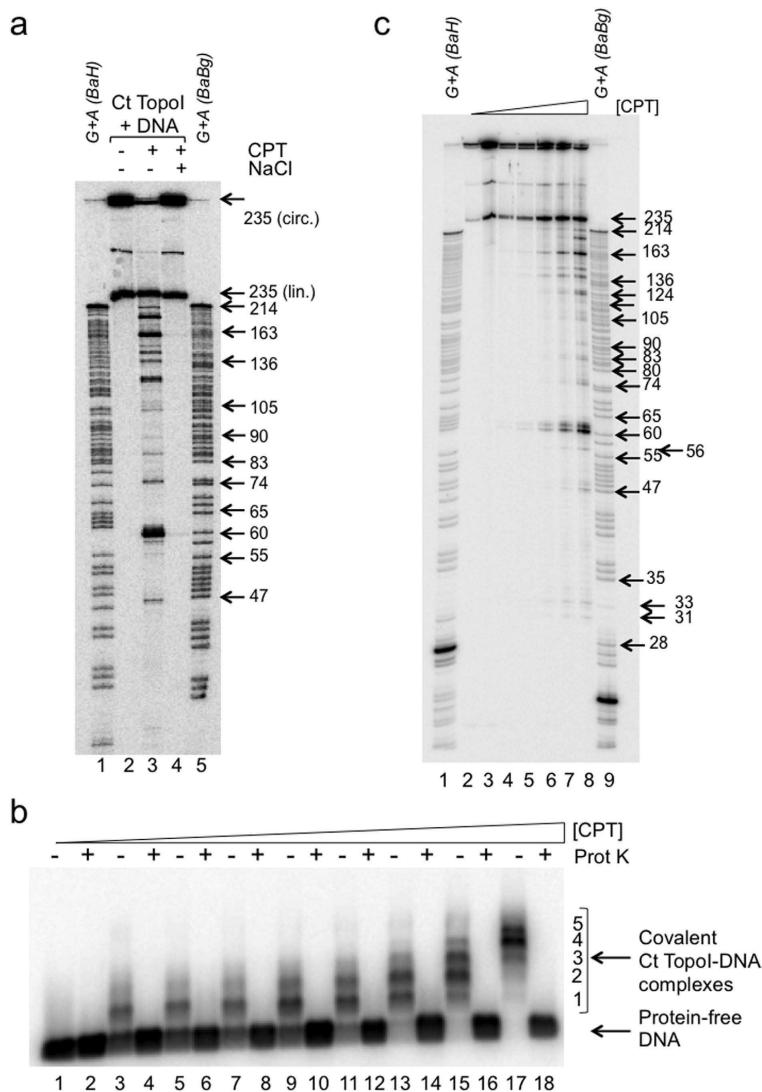


Figure 3. Multiple Ct TopoI can be covalently linked to the dsMCs in the presence of camptothecin.
(a) The relaxation of a mixture of the topoisomers T₋₄ and T₋₅ by the Ct TopoI (18.75 mU μ L⁻¹ (\approx 0.2 nM)) was performed in the absence (lane 2) or in the presence of 10 μ M of camptothecin (CPT, lanes 3 and 4). After 10 minutes of incubation at 37°C, SDS was added to the reaction mixtures to irreversibly trap the Ct TopoI covalently bound to DNA, and the samples were next treated with proteinase K. In lane 4, NaCl was added to the sample prior to SDS and proteinase K treatments. Samples were precipitated and analyzed on 12% sequencing gel. dsMCs were digested with BamHI and HindIII (lane 1) or with BamHI and BglII (lane 5) and sequenced to determine the positions of the guanines and adenines in the sequence³⁶. The size of single-stranded DNA fragments is indicated on the right side of the gel. lin., linear; circ., circular;
(b) As in (a) except for the concentrations of camptothecin (CPT; 0 (lanes 3 and 4); 0.05 (lanes 5 and 6); 0.1 (lanes 7 and 8); 0.25 (lanes 9 and 10); 0.5 (lanes 11 and 12); 1 (lanes 13 and 14); 5 (lanes 15 and 16); 10 μ M (lanes 17 and 18)) and the gel analysis (1% agarose gel, native conditions). When indicated, the samples were next treated with proteinase K (Prot K). Protein-free DNA and covalent Ct TopoI-DNA complexes are indicated. Lanes 1 and 2 contained 4% DMSO (no Ct TopoI no camptothecin); **(c)** As in (a) except for the concentrations of camptothecin (CPT; 0 (lane 3); 0.1 (lane 4); 0.25 (lane 5); 0.5 (lane 6); 1 (lane 7); 10 μ M (lane 8)). dsMCs were digested with BamHI and HindIII (lane 1) or BamHI and BglII (lane 9) and sequenced to determine the positions of the guanines and adenines in the sequence³⁶. Lane 2, mixture of T₋₄ and T₋₅. The size of single-stranded DNA fragments is indicated on the right side of the gel.

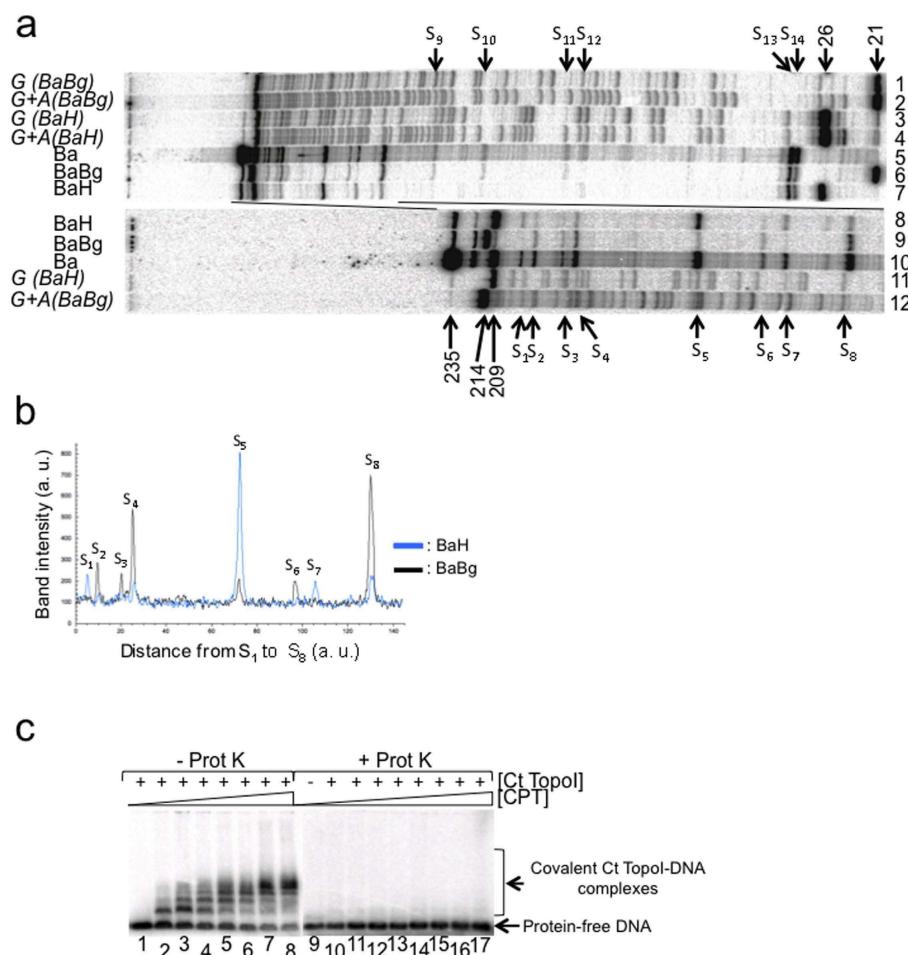


Figure 4. Determination of the sequence of the camptothecin-induced cleavage sites. (a) Identification of the camptothecin-induced cleavage sites on the BamHI-digested dsMCs. The dsMC linearized by BamHI was treated with Ct TopoI ($18.75 \text{ mU } \mu\text{L}^{-1}$ ($\approx 0.2 \text{ nM}$)) in the presence of $10 \mu\text{M}$ camptothecin as described in the methods section. After 10 minutes of incubation at 37°C SDS was added to irreversibly trap the Ct TopoI covalently bound to DNA. The sample was next treated with proteinase K and precipitated. Lanes 5 and 10, sample directly loaded on a sequencing gel after precipitation. Sample further digested with BglII (lanes 6 and 9) or HindIII (lanes 7 and 8). dsMCs were digested with BamHI and BglII (lanes 1, 2 and 12) or BamHI and HindIII (lanes 3, 4 and 11) and sequenced to determine the positions of the guanines (lanes 1, 3 and 11) or the guanines and adenines (lanes 2, 4 and 12) in the sequence³⁶. Samples loaded onto a 10% (lanes 1–7) or 7% (lanes 8–12) acrylamide sequencing gel. The camptothecin-induced cleavage sites are labelled from S₁ to S₁₄; (b) Overlayed scans of lanes 8 and 9 from site S₁ to S₈. Black line, scan of lane 8 (BamHI-BglII double digestion). Blue line, scan of lane 9 (BamHI-HindIII double digestion); a. u., arbitrary unit; (c) The linearized dsMCs were incubated with Ct TopoI ($18.75 \text{ mU } \mu\text{L}^{-1}$ ($\approx 0.2 \text{ nM}$)) and increasing concentrations of camptothecin [CPT; 0.05 (lanes 2 and 11); 0.1 (lanes 3 and 12); 0.2 (lanes 4 and 13); 0.5 (lanes 5 and 14); 1 (lanes 6 and 15); 5 (lanes 7 and 16); 10 μM (lanes 8 and 17)] as described in the methods section. After 10 minutes of incubation at 37°C , SDS was added to the reaction mixtures to irreversibly trap the Ct TopoI covalently bound to DNA. When indicated, the samples were next treated with proteinase K (Prot K). The products of the reactions were finally analyzed on a 1% agarose gel under native conditions. Lane 9 contained 4% DMSO (no Ct TopoI no camptothecin). Lanes 1 and 10 contained 4% DMSO instead of camptothecin.

all three sites had a guanine at position +1 and (iii) neither a guanine and nor a thymine was found at positions –4 and –3, respectively (Table 1).

The patterns of incision of the topoisomers T₀, T₋₃ and T₋₆ (shown only for T₋₃) were similar to that of the BamHI-linearized dsMC, and the 12 cleavage sites observed with the linear form of the dsMCs were detectable with the dsMCs (Supplementary Fig. 6, only S₁ to S₉ cleavage sites indicated by * are shown). This result is consistent with the observation that the profile of appearance of the covalent

site position (5' to 3')	-4	-3	-2	-1 ↓	+1	%
4	A	G	G	T	G	23.5%
5	T	C	G	T	G	50%
8	A	C	C	T	G	26.5%
base position (5' to 3')	-4	-3	-2	-1 ↓	+1	
A	50%	0%	0%	26.5%	0%	
C	0%	76.5%	26.5%	0%	0%	
G	0%	23.5%	73.5%	0%	100%	
T	50%	0%	0%	73.5%	0%	

Table 1. Sequences of the strong cleavage sites and percentage of observed bases at cleavage sites.

The upper part of the table shows the sequence at the designated site. The vertical arrow (↓) indicates the cleavage site. The right column (%) represents the relative percentage of incision at the designated site. The lower part of the table represents the percentage of each of the four bases taking into account the relative percentage of incision at each site.

complexes trapped by SDS on linear DNA as a function of camptothecin (Fig. 4c) was very similar to that obtained with a mixture of topoisomers T₋₄ and T₋₅ (Fig. 3b). However the dsMC cleavage profiles either singly digested by BamHI (Supplementary Fig. 6, lane 5) or doubly digested by BamHI-BglII (Supplementary Fig. 6, lane 4) or BamHI-HindIII (Supplementary Fig. 6, lane 3) were more complex than their equivalents obtained with the BamHI-linearized dsMCs (Fig. 4a, lanes 7–9). This higher complexity stems from the partial digestion of BamHI giving rise to additional bands detectable on the incision profile of (i) the undigested dsMC (Supplementary Fig. 6, indicated by ↑ on lanes 5 and 8), (ii) the singly BglII-digested dsMC (Supplementary Fig. 6, indicated by ▽ on lanes 4 and 6) and (iii) the singly HindIII-digested dsMC (Supplementary Fig. 6, indicated by ▲ on lanes 1 and 3). The results of these experiments demonstrate that in the presence of camptothecin, Ct TopoI incised the dsMCs in a linking number independent manner and that the circular nature of the dsMCs did not influence the positioning of the Ct TopoI on the DNA. Therefore the use of the linearized form of dsMCs to determine the consensus sequence of incision of a novel type IB topoisomerase is more straightforward.

Visualization of the covalent complex during the relaxation of dsMCs by the Ec TopoI. Our method made it possible to visualize the covalent complex Ct TopoI-DNA during the relaxation of dsMCs. We wondered whether the covalent complex could also be detected with another topoisomerase, the *Ec* TopoI, since this topoisomerase was able to efficiently relax most dsMCs (Supplementary Figs. 2b, 3b). Complexes of *Ec* TopoI covalently linked to DNA could be detected on an agarose gel under native conditions (Fig. 5a, lane 2). Whereas SDS could convert nearly 40% of the Ct TopoI bound to DNA into a covalent complex (Fig. 1a, compare lanes 3 and 7), less than 20% of the *Ec* TopoI bound to DNA could be irreversibly linked to DNA by the SDS treatment (Fig. 5a, compare lanes 2 and 4), indicating that the efficiency of the SDS treatment was topoisomerase-dependent.

A covalent *Ec* TopoI-DNA complex could also be detected by SDS-PAGE as a proteinase K-sensitive species (Fig. 5b, arrow 1) and a slight enrichment of linear DNA were clearly observed (Fig. 5b, compare lanes 8 and 9). When samples were analyzed on chloroquine-containing polyacrylamide gels under native (Supplementary Fig. 7a) or denaturing (Supplementary Fig. 7b) conditions, a high molecular weight species (Supplementary Fig. 7a, b, lane 7, arrow 1) that failed to enter into the gel and that was sensitive to proteinase K treatment (Supplementary Fig. 7a, b, compare lanes 7 and 8) was detected. As with Ct TopoI, a small amount of *Ec* TopoI was covalently bound to DNA during the relaxation of dsMCs. In agreement with the results obtained by the analysis on an agarose gel, a smaller proportion of *Ec* TopoI than of Ct TopoI was covalently trapped on DNA.

Discussion

DsMCs have been described as powerful substrates to investigate the structure of specific DNA sequences since the influence of adjacent random sequences is minimized²⁴. They have also been widely used in the past by Prunell's laboratory to investigate multiple aspects of the chromatin assembly process^{25–30}. In

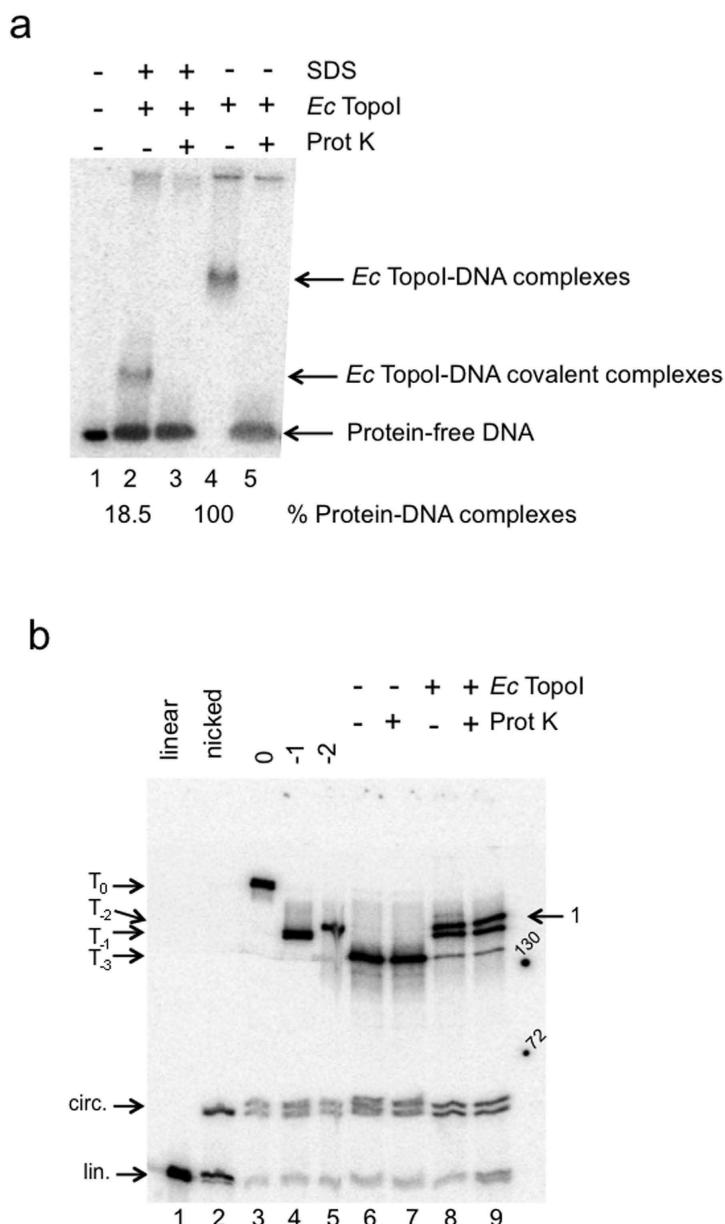


Figure 5. Visualization of the covalent *Ec TopoI*-DNA complex. (a) Relaxation assay analysed on an agarose gel under native conditions. Relaxation of topoisomers T₋₃ was performed with the *Ec TopoI* at a concentration of 500 mU μ L⁻¹ (\approx 30 nM) (lanes 2–5) as described in the methods section. After an incubation of 10 minutes at 37°C, SDS was added in the reaction mixtures (lanes 2 and 3) to irreversibly trap the *Ec TopoI* covalently bound to DNA. The samples were next treated with proteinase K (Prot K) to confirm the presence of proteins in the retarded bands (lanes 3 and 5). The % of proteins engaged in a complex with dsMCs is indicated. Protein-free DNA and complexes with *Ec TopoI* covalently linked or not to the DNA are indicated; (b) Relaxation assay analyzed by SDS-PAGE. The relaxation of topoisomer T₋₃ was performed by the *Ec TopoI* at a concentration of 500 mU μ L⁻¹ (\approx 30 nM) as described in the methods section. After 10 minutes at 37°C, SDS was added to the reaction mixtures to irreversibly trap the *Ec TopoI* covalently bound to DNA (lane 8). Lane 9, the sample was next treated with proteinase K (Prot K) to detect radiolabeled bands that were sensitive to the proteinase K treatment. Lane 1, linear 235 nt-long fragment; lane 2, denatured nicked dsMCs; lane 3, topoisomer T₀; lane 4, topoisomer T₋₁; lane 5, topoisomer T₋₂; lanes 6 and 7, samples were treated as in lanes 8 and 9, respectively, except that no enzyme was added. Arrow 1 on the right side of the gel points to the proteinase K-sensitive material. Numbers 130 and 72 on the right side of the gel correspond to the molecular mass of proteins of 130 and 72 kDa, respectively. lin., linear. cir., circular.

addition to allowing to measure the change of linking number per relaxation step of the topoisomerase reaction cycle³¹ and specific kinetic and partitioning parameters of relaxation³², they offer the possibility to dissect other key biochemical features of the relaxation reactions performed by topoisomerases. As reported here, the quantitative analysis of relaxation reactions in terms of the proportion of each topoisomeric as a function of enzyme concentration is indeed made possible by the radioactive nature of the DNA substrates and the possibility of resolving the topoisomers by one dimensional PAGE (Supplementary Fig. 1b). We validated the use of radiolabeled dsMCs of variable linking numbers to characterize topoisomerases by first testing the relaxation activity of two topoisomerase models, the *Ec* TopoI and the *Ct* TopoI, on these different topoisomers (Supplementary Figs. 2, 3). Both topoisomerases indeed relaxed the dsMCs efficiently, although with differences not only in the reaction yield but also in the final product. The *Ec* TopoI was unable to fully relax the dsMCs since the major final product was T_{-1} , a result which is in agreement with the early observation made by Wang in 1971 that the relaxation of negatively supercoiled plasmid does not go to completion, probably because of the slow rate of the reaction when the degree of superhelicity is low³³. In contrast, the *Ct* TopoI fully relaxed all substrates.

Second, using a method that relied on the use of radiolabeled supercoiled dsMCs and a combination of gels to analyze the products of a relaxation reaction (agarose gels, polyacrylamide gels with SDS or chloroquine), we succeeded in visualizing the covalent topoisomerase-DNA intermediate during a relaxation reaction (Figs. 1–3, 5, Supplementary Figs. 4, 7). The covalent topoisomerase-DNA complex is a transient species of the relaxation reaction with a very short half-life time. As a consequence, its detection and isolation during a relaxation reaction has been very difficult. Our method has filled this gap since the proportion of covalent complexes formed during a relaxation reaction and trapped by SDS has been estimated for both topoisomerases ($\approx 40\%$ and $\approx 20\%$ for the *Ct* TopoI and *Ec* TopoI, respectively; Figs. 1a, 5a). Its dependence with respect to the topoisomerase possibly reflects mechanistic differences between the type IA and type IB families of topoisomerases, especially in the mechanism conferring processivity. The molecular mass of the covalent complex measured on the SDS-PAGE, although close to the expected molecular mass in the case of the *Ct* TopoI covalent complex (165 kDa versus $\approx 170\text{ kDa}$ for the expected versus the measured molecular mass), greatly differs from the expected molecular mass in the case of the *Ec* TopoI covalent complex (170 kDa versus $\approx 260\text{ kDa}$ for the expected versus the measured molecular mass). The difference between expected and apparent molecular mass might stem from the difference in the nature of the species that are monitored (protein for the protein molecular weight marker and protein-DNA species for our samples). Discrepancies between predicted and observed buoyant densities have been reported in the case of the protein-DNA covalent complex formed with the DNA untwisting enzyme¹³.

Third, as expected, the covalent *Ct* TopoI-DNA intermediate was greatly enriched when the relaxation was performed in the presence of camptothecin (Figs. 1–3, Supplementary Fig. 4), consistent with the covalent linking of multiple topoisomerases per dsMC (Fig. 3). This species was competent for religation following an increase of salt (Fig. 2, Supplementary Fig. 4), and interestingly, T_0 was the major product of the reaction under these experimental conditions. Other products (such as T_{+1} and T_{-1} , clearly detectable during a relaxation reaction without camptothecin) were surprisingly barely detectable, suggesting that camptothecin targets the covalent complex trapped with relaxed DNA but cannot insert into the DNA of the covalent *Ct* TopoI-DNA complex assembled on the topoisomers T_{+1} or T_{-1} .

Fourth, our analysis of the incision profiles as a function of supercoiling revealed that in the presence of camptothecin, *Ct* TopoI established covalent bonds at positions on dsMCs that were influenced neither by the linking number nor by the circular nature of the substrate (Fig. 4, Supplementary Figs. 5, 6). Although the number of incision sites is small (12), the sequence of the three strongest sites agrees quite well with the published consensus sequence²³ with thymine at position -1 and guanine at position $+1$ in 73.5% and 100% of the cases, respectively (Fig. 4 and Table 1). The frequency of detectable cleavage sites ranged from $\approx 1/30$ to $\approx 1/60$ for the strand that carries 8 and 4 sites, respectively. This percentage is low compared to the frequency published previously²³ and may stem from the reaction conditions that differ between the studies.

All together our data show that radiolabeled dsMCs constitute a powerful alternative substrate to plasmids and oligonucleotides to biochemically characterize the relaxation mechanism of any new topoisomerase, either wild type or mutant, alone or with its binding partners. In addition, our method makes it possible to detect the covalent complex with the preferred and natural substrate of the topoisomerase-catalyzed relaxation reaction. It therefore constitutes a novelty and an improvement in the field of topoisomerases since most of the *in vitro* experiments investigating the nicking-joining reaction and the mechanism of action of topoisomerase drugs use DNA substrates (such as oligonucleotides of various configurations; ss; ds-ss junctions etc) that do not position the topoisomerase in its functional state e.g. ready to change the topology. Consequently, relationships between the proportion of covalent complex trapped and the efficiency of relaxation will possibly be established in the future. Finally, our study demonstrates that dsMCs are powerful tools to investigate the mechanism of action of topoisomerase drugs during a relaxation reaction, especially molecules that act by stabilizing the covalent topoisomerase-DNA complex.

Methods

Material. The calf thymus DNA topoisomerase I (Ct TopoI) was from Invitrogen. The *Escherichia coli* Topoisomerase I (*Ec* TopoI), T4 polynucleotide kinase, calf intestinal phosphatase, T4 DNA ligase, High Fidelity BamHI (HF-BamHI), BglII, High Fidelity HindIII (HF-HindIII) Nb., and BbvCI were from New England Biolabs. [γ -³²P]-ATP was from Perkin Elmer. The PageRuler Plus prestained protein ladder was from Thermoscientific. The plasmid pHCO3 has been described elsewhere³⁴.

Preparation of ds DNA mini-circles (dsMCs) with a specific relative linking number. To prepare dsMCs of 235 bps, the pHCO3 plasmid was digested by HF-BamHI [1U μ L⁻¹, 4 h, 37°C in CutSmart buffer (50 mM KOAc, 20 mM Tris-OAc, 10 mM Mg(OAc)₂, pH 7.9)] and dephosphorylated by calf intestinal phosphatase (33 mU μ L⁻¹, 2 h, 37°C in CutSmart buffer), leading to DNA fragments of 235 and 2765 bps. The calf intestinal phosphatase was inactivated by phenol, phenol:chloroform and chloroform extraction. The dephosphorylated DNA fragment of 235 bps was gel-purified, radiolabeled with polynucleotide kinase [2 μ M [γ -³²P]-ATP, 0.6 U μ L⁻¹, 30 min, 37°C in polynucleotide kinase buffer (70 mM Tris-HCl, 10 mM MgCl₂, 5 mM DTT, pH 7.6)] and ligated with T4 DNA ligase (2 U μ L⁻¹, 4 h, 16°C in T4 DNA ligase buffer (50 mM Tris-HCl, 10 mM MgCl₂, 1 mM ATP, 10 mM DTT, pH 7.5) in the presence of various concentrations of ethidium bromide (from 0 to 20 mg mL⁻¹). After ethidium bromide extraction, the covalently closed dsMCs were gel-purified on a 4% acrylamide (29:1 = acrylamide:bisacrylamide mass ratio) gel without chloroquine and recovered by electro-elution (Supplementary Fig. 1a). A specific relative linking number has been attributed to each topoisomer using the band counting method³⁵. The relative linking number of 0 has been arbitrarily attributed to the dsMC obtained by ligation without ethidium bromide. When starting with the purified linear 235 bp-long fragment, the yield of preparation of dsMCs ranges from 25% to 2%.

Relaxation assay. Unless indicated otherwise, the relaxation assays were performed as follows and in 20 μ L of Ct buffer (10 mM Tris-HCl, 26 mM KCl, 8 mM MgCl₂, 500 μ M DTT, pH 7.5) for the Ct TopoI or 20 μ L of *Ec* buffer (20 mM Tris-OAc, 50 mM KOAc, 10 mM Mg(OAc)₂, 1 mM DTT, pH 7.9) for the *Ec* TopoI. 1 nM of dsMCs of indicated topology or of the 235 bp-long fragment linearized by BamHI were pre-incubated for 5 min at 37°C in relaxation buffer before adding the enzyme at the indicated concentration. After 10 min of incubation at 37°C, the relaxation reactions were stopped by adding SDS (final concentration, 1%) to trap some or all of the topoisomerase covalently linked to DNA. The samples were next treated with proteinase K (final concentration, 1.2 mg mL⁻¹; 10 min at 37°C followed by 10 min at 56°C) and precipitated with ethanol. When used, camptothecin was diluted in DMSO, added to a final concentration of up to 10 μ M (leading to 4% DMSO in the reaction mixture) and pre-incubated with the DNA for 5 min at 37°C prior to enzyme addition. When NaCl was used to promote DNA religation within the covalent complex, its final concentration was 0.5 M and it was added at the end of the incubation period between the topoisomerase and the dsMCs.

Conditions of gel analysis. All gels were made with TBE 0.5x (44.5 mM Tris-Base, 44.5 mM boric acid, 1 mM EDTA). When the samples were analyzed under native conditions on an agarose gel, at the end of the relaxation reaction the samples were not precipitated and glycerol was added to a concentration of 4% before loading of the samples onto the 1% agarose gel.

When the samples were analyzed under native conditions on a polyacrylamide gel, after ethanol precipitation the pellets were redissolved in TE supplemented with 4% glycerol and loaded onto a 6% polyacrylamide (29:1 = acrylamide:bisacrylamide mass ratio) gel supplemented with chloroquine (final concentration, 30 μ g mL⁻¹).

When the samples were analyzed under denaturing conditions on a polyacrylamide gel, after ethanol precipitation the pellets were redissolved in a denaturing loading buffer (TE supplemented with 4% glycerol and 0.1 M NaOH) before being analyzed on a 6% polyacrylamide (29:1 = acrylamide:bisacrylamide mass ratio) gel supplemented with chloroquine (final concentration, 30 μ g mL⁻¹). [Note that to prevent the loss of the ss DNA (either circular or linear) during electrophoresis and to maintain a good resolution of the topoisomers, the 6% polyacrylamide gel (38 mL) was poured above a 12% polyacrylamide gel (12 mL; 29:1 = acrylamide:bisacrylamide mass ratio)].

When the samples were analyzed by SDS-PAGE, at the end of the relaxation reaction the samples were not precipitated but denatured with SDS (final concentration, 3.25%), formamide (final concentration, 3%) and heat (5 min, 95°C); their pH was adjusted to 6.8. [Note that the reticulation of the stacking part of the protein gel was 80:1 (acrylamide: bisacrylamide mass ratio)].

Conditions of digestion with restriction and nicking enzymes. The digestion of dsMCs by BamHI, BglII and HindIII were performed for 18 h at 37°C in 200 μ L in the buffer recommended by the manufacturer (20 mM Tris-OAc, 50 mM KOAc, 10 mM Mg(OAc)₂, pH 7.9 for BamHI and HindIII, and 50 mM Tris-HCl, 100 mM NaCl, 10 mM MgCl₂, pH 7.9 for BglII) with the restriction enzyme at a concentration of 1.5 U μ L⁻¹. To prepare nicked dsMCs, dsMCs were treated with the Nb.BbvCI nicking enzyme (50 mU μ L⁻¹, 1 h, 37°C in 20 mM Tris-OAc, 50 mM KOAc, 10 mM Mg(OAc)₂, 1 mM DTT, pH 7.9 buffer).

Sequencing reactions. To sequence the 235 nt-long DNA, the 235 bp-long fragment linearized by BamHI was completely digested by BglII or HindIII according to the manufacturer's recommendations and sequenced according to the Maxam and Gilbert procedures specific for either guanines or adenines³⁶.

References

- Chen, S. H., Chan, N.-L. & Hsieh, T. New mechanistic and functional insights into DNA topoisomerases. *Annu. Rev. Biochem.* **82**, 139–170 (2013).
- Baker, N. M., Rajan, R. & Mondragon, A. Structural studies of type I topoisomerases. *Nucleic Acids Res.* **37**, 693–701 (2009).
- Schoeffler, A. J. & Berger, J. M. DNA topoisomerases: harnessing and constraining energy to govern chromosome topology. *Q. Rev. Biophys.* **41**, 41–101 (2008).
- Wang, J. C. *Untangling the double helix: DNA entanglement and the action of the DNA topoisomerases*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, USA, 2009.
- Wang, J. C. Cellular roles of DNA topoisomerases: a molecular perspective. *Nat. Rev. Mol. Cell. Biol.* **3**, 430–440 (2002).
- Champoux, J. J. DNA topoisomerases: structure, function, and mechanism. *Annu. Rev. Biochem.* **70**, 369–413 (2001).
- Staker, B. L. *et al.* Structures of three classes of anticancer agents bound to the human topoisomerase I-DNA covalent complex. *J. Med. Chem.* **48**, 2336–2345 (2005).
- Staker, B. L. *et al.* The mechanism of topoisomerase I poisoning by a camptothecin analog. *Proc. Natl. Acad. Sci. USA* **99**, 15387–15392 (2002).
- Bocquet, N. *et al.* Structural and mechanistic insight into Holliday-junction dissolution by Topoisomerase III α and RMI1. *Nat. Struct. Mol. Biol.* **21**, 261–268 (2014).
- Cejka, P., Plank, J. L., Dombrowski, C. C. & Kowalczykowski, S. C. Decatenation of DNA by the *S. cerevisiae* Sgs1-Top3-Rmi1 and RPA complex: a mechanism for disentangling chromosomes. *Mol. Cell* **47**, 886–896 (2012).
- Pommier, Y. Drugging topoisomerases: lessons and challenges. *ACS Chem. Biol.* **8**, 82–95 (2013).
- Pommier, Y. DNA topoisomerase I inhibitors: chemistry, biology, and interfacial inhibition. *Chem. Rev.* **109**, 2894–2902 (2009).
- Champoux, J. J. Strand breakage by the DNA untwisting enzyme results in covalent attachment of the enzyme to DNA. *Proc. Natl. Acad. Sci. USA* **74**, 3800–3804 (1977).
- Liu, L. F., Rowe, T. C., Yang, L., Tewey, K. M. & Chen, G. L. Cleavage of DNA by mammalian DNA topoisomerase II. *J. Biol. Chem.* **258**, 15365–15370 (1983).
- Depew, R. E., Liu, L. F. & Wang, J. C. Interaction between DNA and *Escherichia coli* protein omega. Formation of a complex between single-stranded DNA and omega protein. *J. Biol. Chem.* **253**, 511–518 (1978).
- Tse-Dinh, Y. C. Uncoupling of the DNA breaking and rejoining steps of *Escherichia coli* type I DNA topoisomerase. Demonstration of an active covalent protein-DNA complex. *J. Biol. Chem.* **261**, 10931–10935 (1986).
- Tse, Y. C., Kirkegaard, K. & Wang, J. C. Covalent bonds between protein and DNA. Formation of phosphotyrosine linkage between certain DNA topoisomerases and DNA. *J. Biol. Chem.* **255**, 5560–5565 (1980).
- Dexheimer, T. S. & Pommier, Y. DNA cleavage assay for the identification of topoisomerase I inhibitors. *Nat. Protoc.* **3**, 1736–1750 (2008).
- Tse-Dinh, Y.- C. Bacterial topoisomerase I as a target for discovery of antibacterial compounds. *Nucleic Acids Res.* **37**, 731–737 (2009).
- Cheng, B., Sorokin, E. P. & Tse-Dinh, Y. C. Mutation adjacent to the active site tyrosine can enhance DNA cleavage and cell killing by the TOPRIM Gly to Ser mutant of bacterial topoisomerase I. *Nucleic Acids Res.* **36**, 1017–1025 (2008).
- Cheng, B. *et al.* Asp-to-Asn substitution at the first position of the DxD TOPRIM motif of recombinant bacterial topoisomerase I is extremely lethal to *E. coli*. *J. Mol. Biol.* **385**, 558–567 (2009).
- Cheng, B., Shukla, S., Vasunilashorn, S., Mukhopadhyay, S. & Tse-Dinh, Y.- C. Bacterial cell killing mediated by topoisomerase I DNA cleavage activity. *J. Biol. Chem.* **280**, 38489–38495 (2005).
- Tanizawa, A., Kohn, K. W. & Pommier, Y. Induction of cleavage in topoisomerase I c-DNA by topoisomerase I enzymes from calf thymus and wheat germ in the presence and absence of camptothecin. *Nucleic Acids Res.* **21**, 5157–5166 (1993).
- Camilloni, G., Negri, R., Caserta, M. & Di Mauro, E. Purification and Use of DNA minicircles with different linking numbers. *Methods in Molecular Biology*, DNA topoisomerase protocols, Humana Press Inc. Totowa, New Jersey, USA, 1999, 94, 51–60.
- Hamiche, A. & Prunell, A. Chromatin reconstitution on small DNA rings. V. DNA thermal flexibility of single nucleosomes. *J. Mol. Biol.* **228**, 327–337 (1992).
- Douc-Rasy, S., Kolb, A. & Prunell, A. Protein-induced unwinding of DNA: measurement by gel electrophoresis of complexes with DNA minicircles. Application to restriction endonuclease EcoRI, catabolite gene activator protein and lac repressor. *Nucleic Acids Res.* **17**, 5173–5189 (1989).
- Hamiche, A. *et al.* Interaction of the histone (H3-H4)2 tetramer of the nucleosome with positively supercoiled DNA minicircles: Potential flipping of the protein from a left- to a right-handed superhelical form. *Proc. Natl. Acad. Sci. USA* **93**, 7588–7593 (1996).
- Sivolob, A., De Lucia, F., Révet, B. & Prunell, A. Nucleosome dynamics. II. High flexibility of nucleosome entering and exiting DNAs to positive crossing. An ethidium bromide fluorescence study of mononucleosomes on DNA minicircles. *J. Mol. Biol.* **285**, 1081–1099 (1999).
- Sivolob, A., De Lucia, F., Alilat, M. & Prunell, A. Nucleosome dynamics. VI. Histone tail regulation of tetrasome chiral transition. A relaxation study of tetrasomes on DNA minicircles. *J. Mol. Biol.* **295**, 55–69 (2000).
- Sivolob, A. & Prunell, A. Nucleosome dynamics V. Ethidium bromide versus histone tails in modulating ethidium bromide-driven tetrasome chiral transition. A fluorescence study of tetrasomes on DNA minicircles. *J. Mol. Biol.* **295**, 41–53 (2000).
- Douc-Rasy, S., Kayser, A., Riou, J. F. & Riou, G. ATP-independent type II topoisomerase from trypanosomes. *Proc. Natl. Acad. Sci. USA* **83**, 7152–7156 (1986).
- Anderson, B. G. & Stivers, J. T. Variola type IB DNA topoisomerase: DNA binding and supercoil unwinding using engineered DNA minicircles. *Biochemistry* **53**, 4302–4315 (2014).
- Wang, J. C. Interaction between DNA and an *Escherichia coli* protein omega. *J. Mol. Biol.* **55**, 523–533 (1971).
- Gaillard, C. & Strauss, F. Construction of DNA hemicatenanes from two small circular DNA molecules. *PLoS One* **10**(3): e0119368 (2015).
- Shure, M. & Vinograd, J. The number of superhelical turns in native virion SV40 DNA and minicol DNA determined by the band counting method. *Cell* **8**, 215–226 (1976).
- Maxam, A. M. & Gilbert, W. A new method for sequencing DNA. *Proc. Natl. Acad. Sci. USA* **74**, 560–564 (1977).

Acknowledgments

This work was supported by INSERM, CNRS and LNCC. A. Millet was supported by a fellowship from the Muséum national d'Histoire naturelle.

Author Contributions

E.D. conceived the project. A.M. and E.D. performed the experiments. E.D., A.M. and F.S. analyzed the data. E.D. wrote the paper with input from F.S. and A.M.

Additional Information

Supplementary information accompanies this paper at <http://www.nature.com/srep>

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Millet, A. *et al.* Use of double-stranded DNA mini-circles to characterize the covalent topoisomerase-DNA complex. *Sci. Rep.* 5, 13154; doi: 10.1038/srep13154 (2015).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

SUPPLEMENTARY DATA

**Use of double-stranded DNA mini-circles to
characterize the covalent topoisomerase-DNA complex**

Armél Millet¹, François Strauss¹ and Emmanuelle Delagoutte¹

¹ Structure et Instabilité des Génomes, Sorbonne Universités, Muséum national d'Histoire naturelle, Inserm U 1154, CNRS UMR 7196; CP26, 57 rue Cuvier 75005 Paris, France

Correspondence should be addressed to E.D. (emmanuelle.delagoutte@mnhn.fr)

FIGURE LEGENDS

Supplementary Figure 1 Procedure of preparation of dsMCs. **(a)** The preparation of dsMCs includes several steps. (1) Radiolabeling of the linear 235 bp-long ds DNA with [γ -³²P]-ATP. (2) Ligation at various concentrations of EtBr (from 0 to 20 mg mL⁻¹). (3) EtBr extraction, DNA precipitation. (4) PAGE to separate the different species. (5) Cutting of the bands of the gel corresponding to the dsMCs, DNA electro-elution and precipitation. The dsMCs are eluted from the gel that does not contain chloroquine (left), where the dsMCs, the single-stranded linear and double-stranded linear fragments are indicated. The gel with chloroquine at 30 μ g mL⁻¹ (right) shows the sequential formation of the topoisomers as a function of [EtBr] present in the ligation step and their relative migration. The timing proposed on the left side of the figure with step (4) taking placing overnight (ON) makes the preparation of dsMCs a two-day procedure; **(b)** Resolution of the 7 topoisomers by electrophoresis on a chloroquine-containing gel. The chloroquine in the gel and running buffer was at 30 μ g mL⁻¹. Lane 1 contains topoisomer T₀; lane 2, topoisomer T₋₁; lane 3, topoisomer T₋₂; lane 4, topoisomer T₋₃; lane 5, a mixture of topoisomers T₋₄ and T₋₅; lane 6, topoisomer T₋₅; lane 7, topoisomer T₋₆.

Supplementary Figure 2 Relaxation of the topoisomer T₋₂ by the *Ct* TopoI and the topoisomer T₋₆ by the *Ec* TopoI. Electrophoresis was performed on a chloroquine-containing gel. The chloroquine in the gel and running buffer was at 30 μ g mL⁻¹. **(a)** Relaxation by *Ct* TopoI. Lane 1, topoisomer T₀; lane 2, topoisomer T₋₁; lane 3, topoisomer T₋₂; lanes 4 to 11 contain 37.5, 18.75, 9.38, 4.69, 2.34, 1.17, 0.59, 0.29 mU μ L⁻¹ of *Ct* TopoI, respectively. Comparison of protein titration curves made it possible to estimate that 18.75 mU μ L⁻¹ of *Ct* TopoI corresponded to \approx 0.2 nM. **(b)** Relaxation by *Ec* TopoI. Lane 1, topoisomer T₀; lane 2, a mixture of topoisomers T₋₁ and T₋₂; lane 3, a mixture of topoisomers T₋₃ and T₋₄; lane 4, topoisomer T₋₅; lane 5, topoisomer T₋₆; lanes 6 to 13 contain 0.05, 0.15, 0.46, 1.39, 4.17, 12.5, 37.5, 75 mU μ L⁻¹ of *Ec* TopoI, respectively. Comparison of protein titration curves made it possible to estimate that 18.75 mU μ L⁻¹ of *Ec* TopoI corresponded to \approx 1 nM.

Supplementary Figure 3 Quantification of the relaxation of the dsMCs by the *Ct* TopoI and the *Ec* TopoI. Comparison of protein titration curves made it possible to estimate that 18.75 mU μ L⁻¹ of *Ct* TopoI or *Ec* TopoI corresponded to \approx 0.2 or \approx 1 nM, respectively.

(a) The percentage of product formed after 10 min as a function of the concentration of Ct TopoI for topoisomer T₀ (black line and \times), topoisomer T₋₁ (black line and \diamond), topoisomer T₋₂ (black line and \square) and a mixture of topoisomers T₋₄ and T₋₅ (black line and \circ); **(b)** The percentage of product formed after 10 min of incubation at 37°C as a function of the concentration of Ec TopoI for topoisomer T₀ (black line and \times), topoisomer T₋₁ (black line and \diamond), topoisomer T₋₂ (black line and \square), a mixture of topoisomers T₋₄ and T₋₅ (black line and \circ), topoisomer T₋₅ (black line and \blacksquare) and topoisomer T₋₆ (black line and \bullet). Errors bars correspond to the standard deviation calculated from at least two independent experiments.

Supplementary Figure 4 Visualization of the covalent Ct TopoI-DNA complex. The relaxation of topoisomer T₋₃ by the Ct TopoI (at a final concentration, 18.75 mU μ L⁻¹ (\approx 0.2 nM)) was performed in the absence (lanes 7 and 8) or in the presence of camptothecin (CPT) at a final concentration of 10 μ M (lanes 9-12) as described in the methods section. After 10 minutes of incubation at 37°C, SDS was added to the reaction mixtures to irreversibly trap the Ct TopoI covalently bound to DNA (lanes 7-12). When indicated, the samples were next treated with proteinase K (Prot K) to remove the proteins (lanes 8, 10 and 12). Lanes 11 and 12, NaCl was added to the samples prior to SDS and proteinase K treatments. Lanes 5 and 6, samples were treated as in lanes 7 and 8, respectively, except that no enzyme was added. Arrow 1 on the right side of the gel points to the proteinase K-sensitive material. **(a)** Relaxation assay analyzed on a chloroquine-containing 6% polyacrylamide gel under native conditions. The chloroquine in the gel and running buffer was at 30 μ g mL⁻¹. Lane 1, nicked dsMCs; lane 2, topoisomer T₀; lane 3, topoisomer T₋₁; lane 4, topoisomer T₋₂; **(b)** As in (a) except that the samples were analyzed under denaturing conditions on a chloroquine-containing 6% polyacrylamide gel. The bracket on the right side of the gel represents DNA species that migrate faster than the linear 235 nt-long fragment. lin., linear. cir., circular.

Supplementary Figure 5 Principle of the analysis of the camptothecin-induced cleavage sites of the dsMCs by the Ct TopoI. The relative position of the BamHI, BglII and HindIII restriction sites is shown and the size of the smallest fragment generated by the double BamHI-BglII and BamHI-HindIII digestions is indicated. The dsMC has a green (strand 1) and a black (strand 2) strand. The radioactive phosphorus is represented by a circle filled in green or black, according to the color of the strand. The nick (indicated by

an arrow) is present either on the green (left) or black (right) strand. To identify the strand that is nicked, the profile of the single BamHI digestion must be compared with that of the double BamHI-BglII and BamHI-HindIII digestions. Left: The nick on the dsMC is located on the green strand. The single BamHI digestion of the dsMCs generates fragments of 235 and n nts. The additional BglII digestion generates a 214 and a 21 nt-long fragment (from the 235 and n nt-long fragments, respectively), whereas the supplementary HindIII digestion leaves the n nt-long fragment intact but shortens the 235 nt-long fragment into a 26 nt-long fragment. Right: The nick on the dsMC is located on the black strand. The single BamHI digestion of the dsMCs generates fragments of 235 and m nts. The supplementary BglII digestion leaves the m nt-long fragment intact but shortens the 235 nt-long fragment into a 21 nt-long fragment, whereas the additional HindIII digestion generates a 209 and a 26 nt-long fragment (from the 235 and m nt-long fragments, respectively).

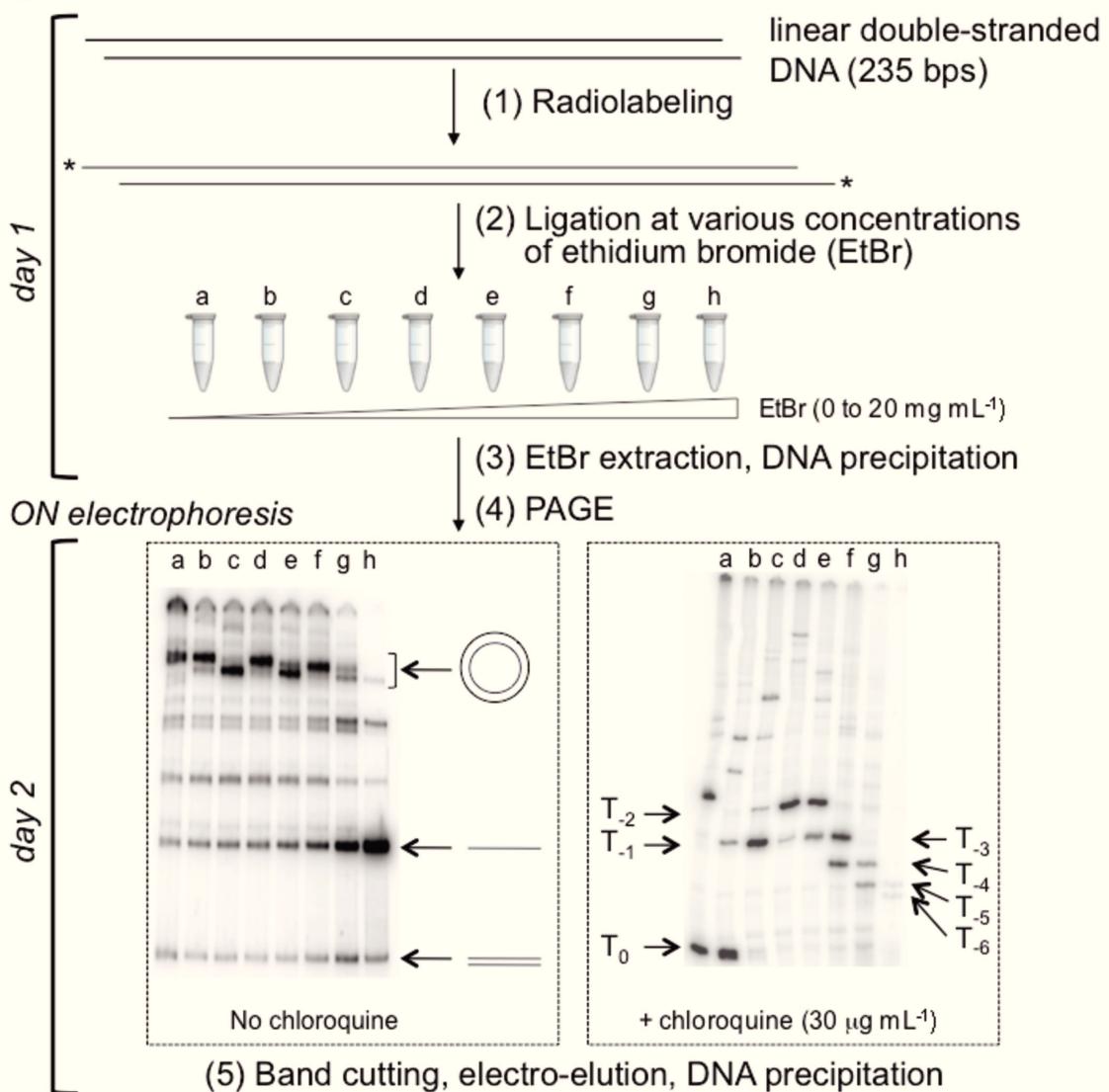
Supplementary Figure 6 Identification of the camptothecin-induced cleavage sites of the dsMCs by the Ct TopoI. All samples were loaded on a 7% acrylamide sequencing gel. The indicated dsMCs (T_0 or T_{-3}) were incubated with Ct TopoI (final concentration, 18.75 mU μL^{-1} (≈ 0.2 nM)) in the presence of camptothecin (final concentration, 10 μM), and after 10 minutes of incubation at 37°C SDS was added to irreversibly trap the Ct TopoI covalently bound to DNA. The samples were next treated with proteinase K and precipitated. The samples were either loaded directly onto the sequencing gel (lane 8), or singly digested with HindIII (lane 1), BamHI (lane 5) or BglII (lane 6), or doubly digested with BamHI and HindIII (lane 3) or BamHI and BglII (lane 4) before being loaded directly onto the sequencing gel. Lanes 2 and 12, dsMCs were digested with BamHI and HindIII and sequenced according to the Maxam and Gilbert procedure to determine the positions of the guanines and adenines in the sequence (1). Lanes 7 and 11, dsMCs were digested with BamHI and HindIII and sequenced according to the Maxam and Gilbert procedure to determine the positions of the guanines in the sequence (1). Lane 9, dsMCs were digested with BamHI and BglII and sequenced according to the Maxam and Gilbert procedure to determine the positions of the guanines in the sequence (1). Lane 10, dsMCs were digested with BamHI and BglII and sequenced according to the Maxam and Gilbert procedure to determine the positions of the adenines and guanines in the sequence (1). The camptothecin-induced cleavage sites

are designated by a *. ↑, main material when the reaction products have not been digested by any restriction enzyme. ∇, main material when the reaction products have only been digested by BglII. ▲, main material when the reaction products have only been digested by HindIII. (1) Maxam, A. M. & Gilbert, W. A new method for sequencing DNA. *Proc. Natl. Acad. Sci. U. S. A.* **74**, 560–4 (1977).

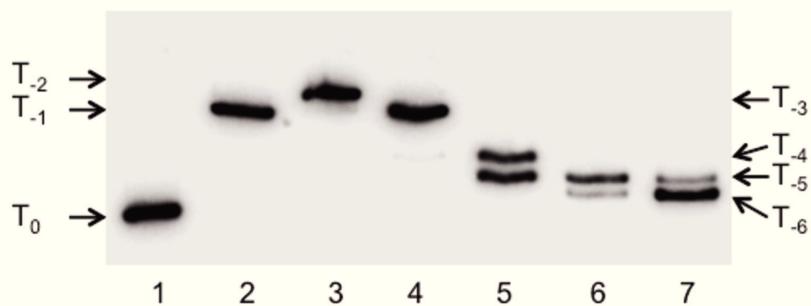
Supplementary Figure 7 Visualization of the covalent *Ec* TopoI-DNA complex. The relaxation of topoisomer T₋₃ was performed by the *Ec* TopoI at a concentration of 500 mM μ L⁻¹ (\approx 30 nM) (lanes 7 and 8). After 10 min at 37°C, SDS was added to the reaction mixtures to irreversibly trap the *Ec* TopoI covalently bound to DNA (lanes 7 and 8). Lane 8, the sample was next treated with proteinase K (Prot K) to identify proteinase K-sensitive radiolabeled material. Lanes 5 and 6, samples were treated as in lanes 7 and 8, respectively, except that no enzyme was added. Lane 1, nicked dsMCs; lane 2, topoisomer T₀; lane 3, topoisomer T₋₁; lane 4, topoisomer T₋₂. Arrow 1 on the right side of the gel points to the proteinase K-sensitive material. (a) Relaxation assay analyzed under native conditions on a chloroquine-containing 6% polyacrylamide gel. The chloroquine in the gel and in the running buffer was at 30 μ g mL⁻¹; (b) Relaxation assay analyzed on a polyacrylamide gel under denaturing conditions. As in (a) except that the samples were analyzed under denaturing conditions on a chloroquine-containing 6% polyacrylamide gel. lin., linear. cir., circular.

Supplementary Figure 1 (Delagoutte)

a

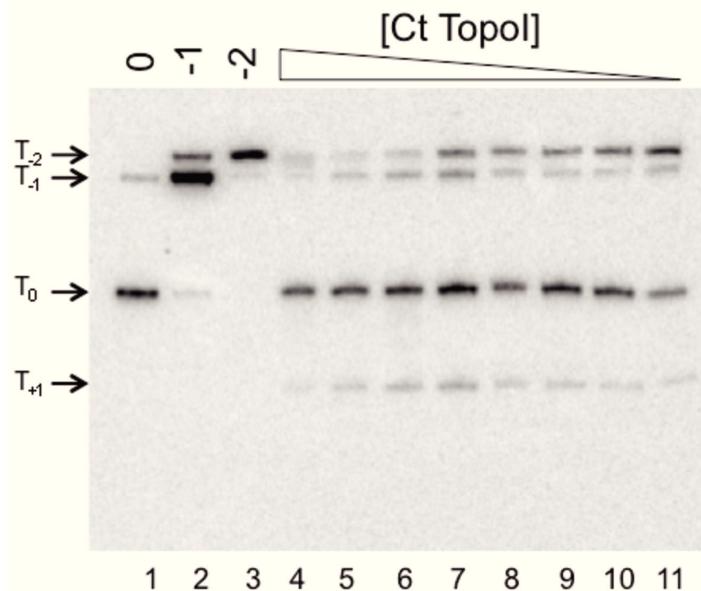


b

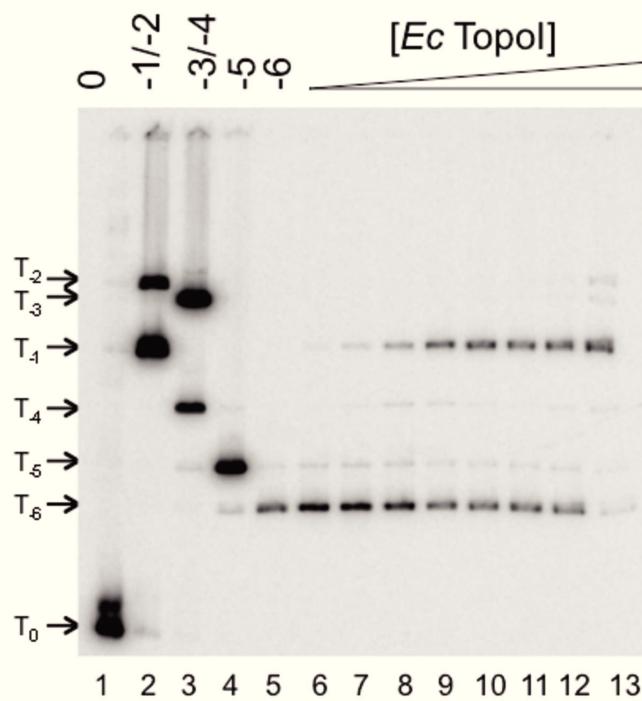


Supplementary Figure 2 (Delagoutte)

a

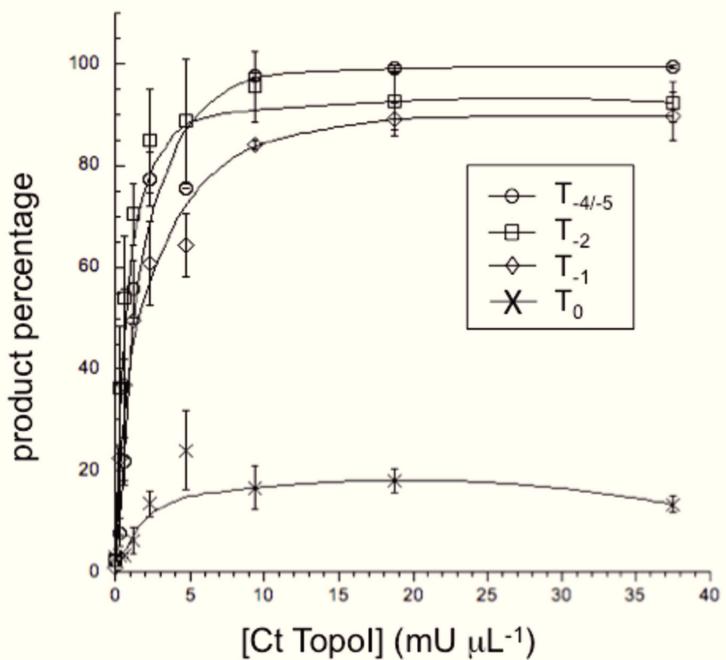


b

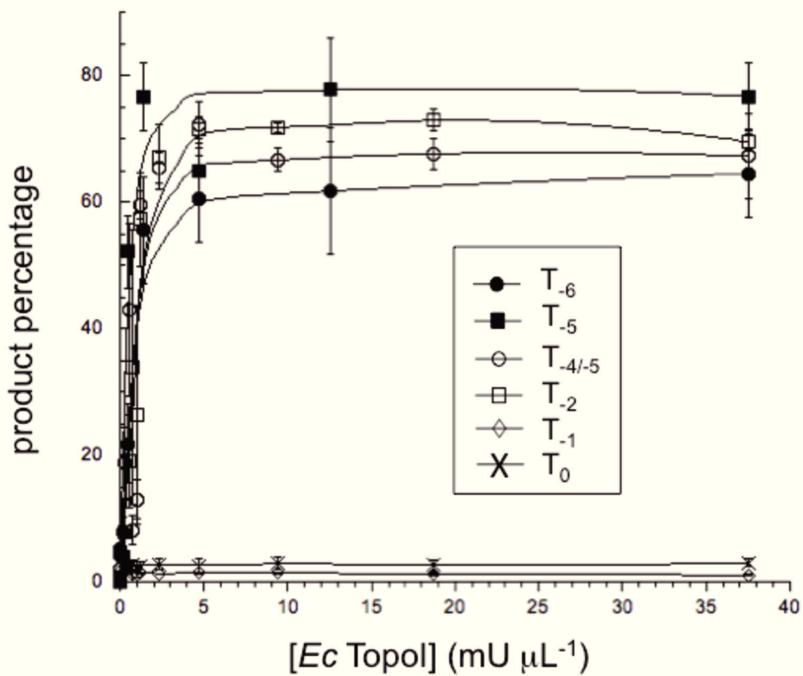


Supplementary Figure 3 (Delagoutte)

a

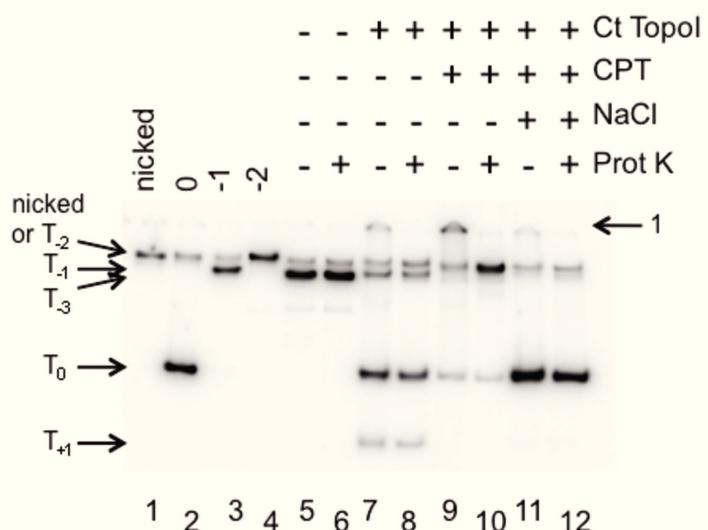


b

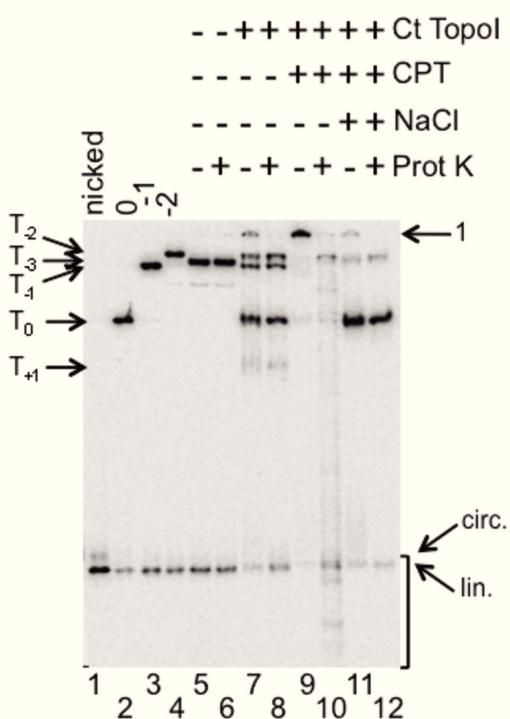


Supplementary Figure 4 (Delagoutte)

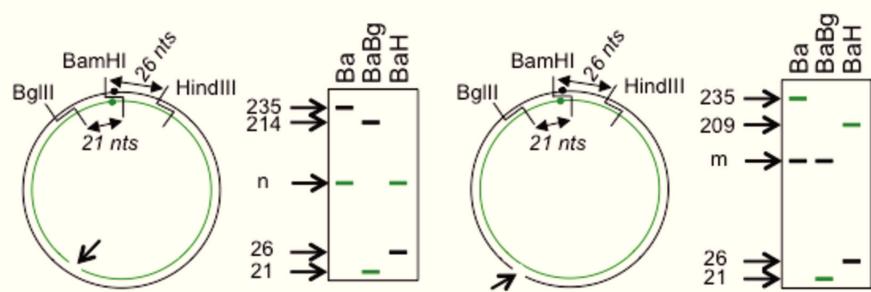
a



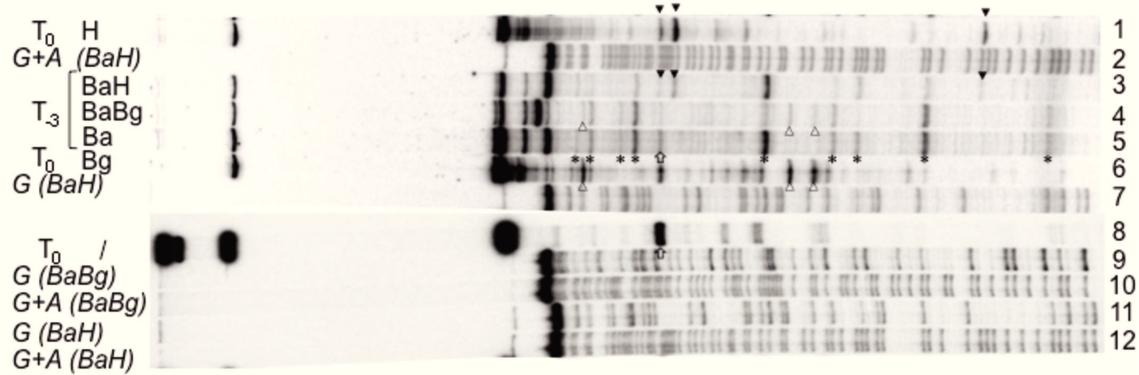
b



Supplementary Figure 5 (Delagoutte)

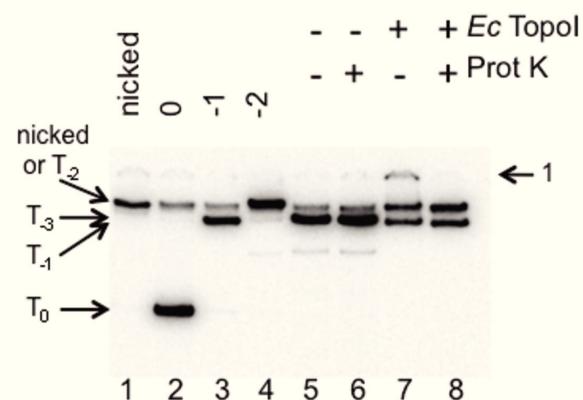


Supplementary Figure 6 (Delagoutte)

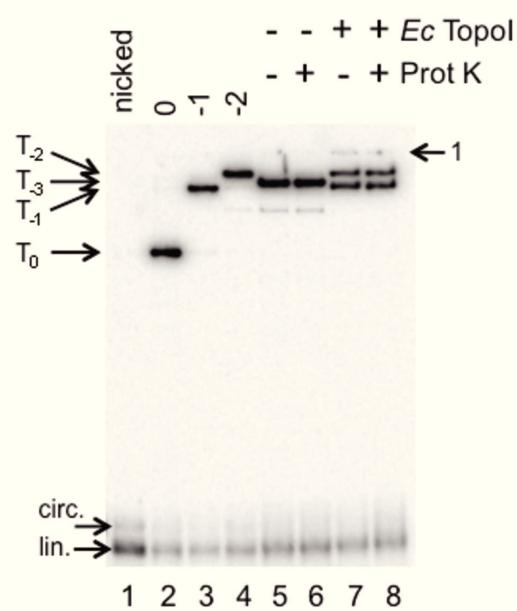


Supplementary Figure 7 (Delagoutte)

a



b



II. Article 2 : "Probing hyper-negatively supercoiled mini-circles with nucleases and DNA binding proteins"

Ce projet vise à mieux comprendre la topologie des minicercles d'ADN utilisés dans l'étude précédente (cf article 1) afin de mieux appréhender l'effet du surenroulement sur leurs structures.

Nous avons utilisé 7 topoisomères d'indice d'enlacement de 0 à -6. L'objectif du projet est de mieux comprendre les différences de topologie entre ces topoisomères. En effet, il est communément répandu que plus l'ADN est hypernégativement surenroulé, plus il présente des zones simple-brin. Grâce à la nucléase SI et à la DNaseI, nous avons pu explorer les régions plus sensibles à ces sondes enzymatiques. Pour compléter l'étude, deux protéines ayant une affinité pour l'ADN simple-brin, la protéine RPA humaine, et l'*EcTopoI* ont été utilisées pour rechercher des différences d'affinités, ce qui illustrerait des proportions différentes en région simple-brin selon les différents topoisomères.

**Probing hyper-negatively supercoiled mini-circles with nucleases and DNA
binding proteins**

Armêl R. Millet¹, Carole Saintomé^{1, 2} and Emmanuelle Delagoutte^{1*}

¹ Structure et Instabilité des Génomes, Sorbonne Universités, Muséum national
d'Histoire naturelle, Inserm U 1154, CNRS UMR 7196; CP26, 57 rue Cuvier 75005 Paris,
France

² Sorbonne Universités, UPMC University Paris 06, F-75005, Paris, France

*Correspondence should be addressed to E.D. (emmanuelle.delagoutte@mnhn.fr)

ABSTRACT

It is well accepted that the introduction of negative supercoils locally unwinds the DNA double helix, influencing thus the DNA binding properties of proteins. However, the extent and location of unpaired bases induced by the negative supercoiling have never been investigated at the nucleotide level. Radiolabeled double-stranded DNA mini-circles (dsMCs) represent useful tools to examine the effect of DNA supercoiling on (i) the base-pairing state of the DNA at the nucleotide level and (ii) the DNA binding properties of proteins. Our goals were to determine the location of unpaired bases on dsMCs whose topology ranged from relaxed to hyper-negatively supercoiled states, and to characterize the binding of two proteins involved in the DNA metabolism. Our results show that Nuclease SI and DNase I detect subtle structural changes among the hyper-negatively supercoiled, negatively supercoiled or relaxed topoisomers. As divalent cations modify the DNA topology, our binding studies were conducted with or without magnesium. As expected, without magnesium, the more negatively supercoiled is the DNA, the better the binding of the human replication protein A (hRPA) and *Escherichia coli* topoisomerase I. Unexpectedly, the magnesium-induced conformation of the topoisomers renders the topoisomer with the lowest superhelicity the best substrate of hRPA.

INTRODUCTION

The activity of proteins involved in the DNA metabolism such as specific DNA binding proteins, nucleases or RNA polymerases, is influenced by the topology of the DNA and more specifically by the degree of superhelicity (1–12). For instance, it was early recognized that the single-stranded DNA binding (SSB) protein of bacteriophage T4 invades less efficiently covalently closed circular than linear double-stranded DNA, due to the strain induced by the strand separation in closed circular DNA (13). In the mid seventies, Wang, by monitoring the transcriptional activity of the RNA polymerase on DNA samples containing varying numbers of superhelical turns, reported the influence of the level of supercoiling on transcription (1). The inhibition of transcription at high negative supercoiling levels stems from the accumulation of hyper-negative supercoiled DNA behind the RNA polymerase. This degree of supercoiling hinders the progression of the transcription complex and its accumulation can result in genome instability (14). In this context, DNA topoisomerases of family IA perform an essential function in preventing the accumulation of hyper-negative supercoiled DNA, and their preferred substrate is often hyper-negatively supercoiled DNA. For example, *Drosophila melanogaster* topoisomerase III α relaxes hyper-negatively supercoiled DNA into supercoiled DNA and cannot relax supercoiled DNA to fully relaxed DNA (15). Structural data have suggested that even if *Escherichia coli* DNA topoisomerase I (*Ec* TopoI) can relax negatively supercoiled DNA, its preference for hyper-negatively supercoiled DNA stems from the nature and the number of interactions between single-stranded DNA and its carboxy-terminal extremity that carries four cysteine zinc ribbon and zinc ribbon-like domains (16).

At the DNA molecular level and from his biochemical assays, Wang suggested that for negatively supercoiled DNA, tightly wound regions were prominent. Sharp bends exposing few bases and possibly becoming preferred substrates for nucleases could appear (1). Thermodynamic considerations led him predict that, in contrast, hyper-negatively supercoiled DNA would expose denatured segments of significant length (1). Over the past decades, experiments of electron microscopy, cryo-tomography and atomic force microscopy, various methods of simulation and 3D reconstruction, and biochemical assays have permitted to investigate the structural features (in terms of bend, kink, unwound regions, base flipping, asphericity, prolateness for example (17, 18))

of supercoiled DNA of various (i) sizes (ranging from mini-circles to plasmids (19)), (ii) sequences (*i.e.* cruciform (20), direct or inverted sequences (21, 22), TATA sequence (23)), and (iii) supercoiling degrees (from positive to hyper-negative (24, 25)) under various experimental conditions (18, 26). However, despite these essential and relevant studies conducted from structural, biophysical, biochemical and biological points of view, at least two questions remained unanswered: (i) where on the sequence and at the nucleotide level, are these structural features located, and (ii) how these structural features influence the binding properties of proteins involved in DNA metabolism?

Our goal in the present study was therefore to investigate the specific structural changes induced by negative supercoiling and more specifically by hyper-negative supercoiling. To this end, we performed two types of experiments, first, experiments that identify at the nucleotide level, the sequence(s) that become(s) hyper-sensitive to or protected by enzymatic probes upon introduction of hyper-negative supercoils, and second, gel shift assays to evaluate the influence of hyper-negative supercoiling on the DNA binding properties of proteins. As DNA substrates, we used double-stranded DNA mini-circles (dsMCs) for several reasons (reviewed in (27)). First, several methods (including our method (28)) of preparation of dsMCs make it possible to produce topoisomers whose topology ranges from relaxed to hyper-negatively or positively supercoiled forms of DNA (25, 29–31). Second, dsMCs have proved to be suitable substrates to characterize the biochemical activity of topoisomerases (30, 32, 34) and the structural properties of DNA, either naked (17, 18, 23, 25, 26, 33) or assembled into chromatin (29, 34–37). Third, they have also been used as models of small topological domains such as DNA loops created by the packaging of the DNA or chromatin remodeling (38, 39). As probes to estimate the proportion of unpaired bases and locate them, we used the methylating agent, dimethylsulfate (DMS), Nuclease SI, and DNase I that detect and modify non-B DNA conformations and react with either single-stranded DNA (DMS and Nuclease SI), or double-stranded DNA (DNase I). As proteins involved in the DNA metabolism, we used proteins able to bind single-stranded DNA, namely the human replication protein A (hRPA) and topoisomerase I from *Escherichia coli* (*Ec* TopoI). These biochemical assays were performed either in the absence or in the presence of divalent cations since specific (*e.g.* writhe) and non canonical structures (*e.g.* base flipping) may be favored under specific salt conditions (40). Such cations indeed

screen the negative charge of the DNA phosphates. As a result they (i) decrease DNA-DNA repulsion, (ii) induce a transition from a loosely to a tightly inter-wound conformation that correlates with a decrease of the effective DNA diameter and (iii) increase the writhe of the DNA (26, 41).

Our results indicate that the structural changes induced by the hyper-negative supercoiling and detected by Nuclease SI and DNase I are subtle, and are not randomly distributed all over the sequence but over a double-stranded region of ~ 40 base pairs. In addition, as expected, an increased sensitivity of Nuclease SI and a decreased sensitivity of DNase I with increasing negative supercoiling is observed. Surprisingly, our binding experiments performed with hRPA suggest that the conformation of the topoisomers induced by magnesium renders the topoisomers with a relative linking number of -1 the best substrate of hRPA.

MATERIAL AND METHODS

Material

Escherichia coli topoisomerase I (*Ec* TopoI), T4 polynucleotide kinase (PNK), calf intestinal phosphatase, T4 DNA ligase, DNase I, BamHI, BglII and HindIII were from New England Biolabs. [γ -³²P]-ATP was from Perkin Elmer. The human replication protein A (hRPA) was prepared in our laboratory (42, 43). Wheat germ Topoisomerase was from Inspiralis. Nuclease SI was from Promega. Proteinase K was from Sigma-Aldrich.

Preparation of ds DNA mini-circles (dsMCs) with a specific relative linking number

The procedure to prepare dsMCs of 235 base pairs with a specific linking number has been published elsewhere (28). Briefly, the dephosphorylated DNA fragment of 235 base pairs was radiolabeled with T4 polynucleotide kinase [2 μ M [γ -³²P]-ATP, 0.6 U μ L⁻¹ PNK, 30 min, 37°C in polynucleotide kinase buffer (70 mM Tris-HCl, 10 mM MgCl₂, 5 mM DTT, pH 7.6)] and ligated with T4 DNA ligase (2 U μ L⁻¹, 4 h, 16°C in T4 DNA ligase buffer (50 mM Tris-HCl, 10 mM MgCl₂, 1 mM ATP, 10 mM DTT, pH 7.5) in the presence of various concentrations of ethidium bromide (from 0 to 20 mg mL⁻¹). After ethidium bromide extraction, the covalently closed dsMCs were gel-purified on a 4% acrylamide (29:1 = acrylamide:bisacrylamide mass ratio) gel without chloroquine and recovered by electro-elution. A specific relative linking number has been attributed to each

topoisomer using the band counting method (44). The relative linking number of 0 has been arbitrarily attributed to the dsMC obtained by ligation without ethidium bromide.

Analysis of dsMCs in divalent cation-containing buffer

dsMCs of indicated topology (final concentration, 0.5 nM) were diluted in 10 mM Tris-HCl pH 7.5 buffer supplemented either with MgCl₂ (final concentration, 10 mM) or CaCl₂ (final concentration, 10 mM) and incubated for 1 h at 37°C. The samples were then analyzed by electrophoresis on a polyacrylamide gel.

Treatment of dsMCs with DMS

Dilutions of DMS were freshly made in 100% EtOH. A 200 µL mixture of 0.5 nM dsMCs of indicated topology and 20 ng µL⁻¹ of sonicated salmon sperm DNA was incubated with DMS at the indicated concentration in a DMS buffer (50 mM Na-cocadylate pH 8, 10 mM MgCl₂, 1 mM EDTA) for 15 min at 20°C. The reaction was stopped by adding 50 µL of DMS stop buffer (50 mM Na-acetate pH 7, 1 M β-mercaptoethanol, 0.1 mg mL⁻¹ tRNA). The sample was then precipitated with EtOH, treated with piperidine, lyophilized twice before being analyzed by electrophoresis on a polyacrylamide gel.

Nuclease SI and DNase I digestion

In the case of digestions of dsMCs with Nuclease SI, 0.5 nM of dsMC of indicated topology was incubated in 20 µL of Nuclease SI buffer (50 mM Tris-HCl, 50 mM NaCl, 20 mM MgCl₂, 5% glycerol, pH 7.5) for 5 min at 25°C before adding the enzyme at the indicated concentration. After 5 min incubation at 25°C, the reactions were stopped by adding EDTA (final concentration, 25 mM) and the samples were kept on ice. The digestions of dsMCs with DNase I were performed in 20 µL of DNase I buffer (40 mM Tris-HCl, 10 mM NaCl, 6 mM MgCl₂, 1 mM CaCl₂, 2 mM DTT, pH 7.9). 0.5 nM of dsMC of indicated topology were then incubated for 5 min at 37°C in DNase I buffer before adding the enzyme at the indicated concentration. After 15 min at 37°C, the reactions were stopped by adding EDTA (final concentration, 25 mM) and the samples were kept on ice. DNase I and Nuclease SI were removed from the samples by adding SDS (final concentration, 0.5 %), NaCl (final concentration, 0.5 M) and 1 volume of chloroform-isoamyl alcohol (24:1 v/v). The samples were vortexed at top speed for 2 min and centrifuged at 4°C for 10 min at 15,000 g. The aqueous upper phases were recovered and their DNA contents were

precipitated with linear polyacrylamide carrier and 2.5 volumes of EtOH. The tubes were left on ice for at least 10 min and centrifuged for 20 min at 4°C at 15,000 g. After discarding the supernatant, the DNA pellets were rinsed with EtOH, dried before being analyzed by electrophoresis on a polyacrylamide gel.

Interaction between dsMCs and hRPA

The tests of interaction between dsMCs and hRPA were performed at 20°C for 10 min in 10 µL of hRPA binding buffer (20 mM Tris-HCl, 50 mM KCl, 1 mM DTT, pH 7.5). When indicated, MgCl₂ was added at a final concentration of 10 mM. Each dsMC of indicated topology was at 0.5 nM and the hRPA at the indicated concentration. When required, proteinase K was added after 10 min interaction between dsMCs and hRPA at a final concentration of 2 mg mL⁻¹ for 10 min at 20°C. At the end of the interaction, glycerol was added at a final concentration of 5% and the samples were kept on ice before being analyzed by electrophoresis on a polyacrylamide gel.

Interaction between dsMCs and Ec TopoI

The tests of interaction between dsMCs and *Ec TopoI* were performed at 37°C for 10 min in 10 µL of *Ec TopoI* binding buffer (20 mM Tris-HCl, 50 mM KCl, 25 mM EDTA, 1 mM DTT, pH 7.5). Each dsMC of the indicated topology was at 0.5 nM and the *Ec TopoI* at the indicated concentration. At the end of the interaction, glycerol was added to the samples at a final concentration of 5% and samples were kept on ice before being analyzed by electrophoresis on a polyacrylamide gel.

Conditions of gel analysis

When the effect of calcium or magnesium on the electrophoretic migration of dsMC was investigated, after 1 h at 37°C, glycerol and NaOH were added to the samples at a final concentration of 4% and 0.1M, respectively. The reactions products were then loaded on a 6% polyacrylamide (29:1 = acrylamide:bisacrylamide mass ratio) gel made in 40 mM Tris-acetate pH 8.5, supplemented with either 10 mM MgCl₂ or 10 mM CaCl₂. Electrophoresis was performed at room temperature for 6 hours and at 80 V. The running buffer composed of 40 mM Tris-acetate pH 8.5, 10 mM MgCl₂ or CaCl₂ was continuously recycled during electrophoresis.

After the lyophylisation that follows the piperidine treatment, the pellets were redissolved in a denaturing loading buffer (TE (10 mM Tris-HCl pH 8, 1 mM EDTA) supplemented with 4% glycerol and 0.1M NaOH) and loaded onto a 4% polyacrylamide (29:1 = acrylamide:bisacrylamide mass ratio) gel made in TBE 0.5x (44.5 mM Tris-Base, 44.5 mM boric acid, 1 mM EDTA) supplemented with 30 µg mL⁻¹ chloroquine. Electrophoresis was performed at room temperature for 6 hours at 100 V.

When the extent of digestion of dsMCs by DNase I or Nuclease SI as a function of enzyme concentration was investigated, after ethanol precipitation the pellets were redissolved in denaturing loading buffer (TE supplemented with 4% glycerol and 0.1M NaOH) and loaded onto a 6% polyacrylamide (29:1 = acrylamide:bisacrylamide mass ratio) gel made in TBE 0.5x (44.5 mM Tris-Base, 44.5 mM boric acid, 1 mM EDTA) supplemented with 30 µg mL⁻¹ chloroquine. Electrophoresis was performed at room temperature for 15 hours at 15 V.

When the profiles of digestion of dsMCs by DNase I or Nuclease SI were investigated, after ethanol precipitation the pellets were redissolved in a formamide-containing loading buffer and loaded onto a 8% or 12% sequencing gel.

When the interaction between dsMCs and hRPA or *Ec* TopoI was investigated, the samples were analyzed under native conditions on a 4% polyacrylamide (29:1 = acrylamide:bisacrylamide mass ratio) gel made in TBE 0.5x (44.5 mM Tris-Base, 44.5 mM boric acid, 1 mM EDTA). When present, MgCl₂ was added at a final concentration of 10 mM both in the gel and in the running buffer. Electrophoresis was performed at 4°C, for 4 hours and at 150 V.

Sequencing reactions

To sequence the single-stranded 235 nucleotide-long DNA, the 235 base pair-long fragment linearized by BamHI was completely digested by BglII or HindIII according to the manufacturer's recommendations and sequenced according to the Maxam and Gilbert procedures specific for either guanines or guanines and adenines (45).

RESULTS

Native electrophoresis under various buffer conditions made it possible to establish that negative supercoiling in the absence of divalent cations, partitioned almost exclusively into untwisting of the DNA helix and that the presence of salt

modified the relative proportion of twist, writhe and compaction in the DNA (26, 31, 40, 46). These findings were clearly reproduced with our seven topoisomers of 235 base pairs whose relative linking number (ΔLk) ranged from 0 to -6 (labeled in what follows T_i where $i = \text{relative linking number}$). Indeed, the comparison of the electrophoresis patterns of the topoisomers in the absence (Supplementary Figure 1a) or in the presence of divalent cations (Supplementary Figures 1b and c), is in agreement with the analysis of topoisomers of dsMCs of similar sizes by other laboratories using cryo-electromicroscopy (26), atomic force microscopy (31) or cryo-electrotomography (25) and suggests the existence of at least two writhe sites and a high degree of compaction for hyper-negatively supercoiled dsMCs in the presence of divalent cations. The magnesium-induced conformational change of the topoisomers was also detected by the methylating agent, DMS, whose activity is independent of divalent cation and which has been widely used in the past to probe DNA conformations in particular, the local unwinding of DNA (47–49). Indeed, our results indicate that a given topoisomer is slightly more sensitive to DMS in the absence of magnesium than in the presence of magnesium (Supplementary Figure 2a).

1. Location of the nucleotides sensitive to hyper-negative supercoiling

Conformational changes of the dsMCs associated with the introduction of various levels of writhe and compaction might be detected by enzymatic and chemical probes. Nuclease SI is specific of single-stranded DNA, whereas DNase I is specific of double-stranded DNA and both nucleases can detect non canonical structures into double-stranded DNA. To locate the nucleotides whose conformation changed in response to the introduction of negative and hyper-negative supercoiling, we first checked the sensitivity of the radiolabeled topoisomers of dsMCs with these two protein probes, Nuclease SI and DNase I. It should be noted that Nuclease SI digestion was performed at neutral pH by adjusted experimental conditions as recommended (50). Each topoisomer was incubated with increasing amounts of nuclease for a given period of time and at a given temperature. Nucleases were removed from the reaction products and DNA was precipitated before being denatured by NaOH and analyzed by electrophoresis on a native polyacrylamide gel. Under these conditions, dsMCs were resolved from single-stranded DNA (either circular or linear). The results showed a decrease of the amount of substrate and an accumulation of ≤ 235 nucleotide-long DNA fragments as a function of

Nuclease SI concentration (Figure 1a, left panel). In addition, the higher was the negative supercoiling, the more active was Nuclease SI (Figure 1a, right panel). In contrast, the results obtained with DNase I showed that the topoisomer T_0 was two times more sensitive to DNase I than a mixture of topoisomers T_{-5} and T_{-6} (Figure 1b). For both experiments, the reaction products consisted of fragments whose length was ≤ 235 nucleotides, revealing multiple cuts by the nucleases per dsMC. Similar results were obtained by others (19). In contrast to the results obtained with the enzymatic probes, the topoisomers, among themselves, exhibited the same reactivity towards DMS, independently of the presence or not of divalent cations (Supplementary Figures 2b and c) revealing a difference of sensitivity between chemical and enzymatic probes. Altogether, these results indicate that hyper-negatively supercoiled dsMCs are more accessible to the Nuclease SI digestion than the relaxed topoisomer, possibly exposing additional and/or better substrates. In contrast, relaxed dsMCs are slightly better substrates for DNase I than hyper-negatively supercoiled dsMCs, in agreement with the Nuclease SI results, given the opposite specificities of the two nucleases.

The difference of reactivity of our topoisomers with Nuclease SI and DNase I having been established and shown to be significantly different, we next investigated which nucleotide(s) on the dsMCs sequence became hyper-sensitive or hypo-sensitive to Nuclease SI or DNase I upon introduction of hyper-negative supercoiling in the dsMCs. To this end, the concentration of nuclease that permitted to digest less than 20% of the initial substrate (corresponding to on average of one nick per DNA molecule, at most) was determined, and this nuclease concentration was used to compare the profiles of digestion of two topoisomers of dsMCs, T_{-2} and T_{-6} for the Nuclease SI digestion and T_0 and T_{-6} for the DNase I digestion. After the nuclease digestions, the reaction products were digested by either the BamHI/BglII or the BamHI/HindIII restriction enzyme couple (Supplementary Figure 3) in order to analyze only one of the two radiolabeled strands of the dsMCs on a sequencing gel (our procedure to prepare dsMCs indeed leads to dsMCs radiolabeled on both strands (see Material and Methods)).

Nuclease SI digestion permitted to identify two regions whose sensitivity to this nuclease was influenced by the supercoiling. One region was single-stranded and consisted of a seven nucleotide-long sequence for which the introduction of hyper-negative supercoiling in the dsMCs made it more accessible to Nuclease SI (region

shown in red, Figures 2a and c). The second region consisted of the two complementary strands of a six nucleotide-long region that, in contrast to the first region, became less sensitive to Nuclease SI upon introduction of negative supercoiling in the dsMCs (region shown in green, Figures 2a and c). Finally our experiments identified a single nucleotide (shown in blue, Figures 2a-c) that became hyper-sensitive to both Nuclease SI and DNase I upon introduction of hyper-negative supercoiling.

2. Interaction between topoisomers of dsMCs and single-stranded DNA binding proteins in the absence of divalent cations

Our probing experiments performed with Nuclease SI and DNase I identified the nucleotide(s) whose conformation was influenced by the negative supercoiling (Figure 2). We next wondered whether a protein that binds single-stranded DNA could detect the structural differences between the topoisomers of dsMCs and more specifically in the extent of locally unwound regions. In the past, electron microscopy (13), AFM (51) and single-molecule (52) experiments used various single-stranded DNA binding proteins as a reporter of single-stranded DNA. Consequently, the interaction between each topoisomer T_i and two proteins known to bind single-stranded DNA, the hRPA and *Ec* TopoI, was characterized. Since the binding studies were performed in the absence of divalent cations, relaxation of the negatively supercoiled dsMCs by the *Ec* TopoI was prevented. The topoisomers T_0 , T_{-2} , mixtures of topoisomers T_{-3}/T_{-4} , T_{-4}/T_{-5} and T_{-5}/T_{-6} were incubated with increasing amounts of hRPA or *Ec* TopoI, and the resulting products were resolved on a polyacrylamide gel under native conditions (Figure 3). The results indicated that with concentrations of hRPA ranging from 0 to 300 nM, up to five different protein-DNA complexes (called C1 to C5) could be resolved on a polyacrylamide gel (Figure 3a). It is possible that these complexes differ by the number of hRPA molecules bound per DNA molecule, the C1 complex migrating the fastest containing less hRPA molecules per dsMC than the C5 complex migrating the slowest. No protein-DNA complex was formed with the topoisomer T_0 , and a single species (C1) was assembled with the topoisomer T_{-2} (Figure 3a and Table 1). The C2 and C3 complexes started forming with the mixture of topoisomers T_{-3}/T_{-4} and the C3 complex was the most abundant complex formed with the mixture of topoisomers T_{-4}/T_{-5} at the highest hRPA concentration (300 nM) (Figure 3a). The C4 and C5 complexes only formed with the mixture of topoisomers T_{-5}/T_{-6} . Similar results were obtained with *Ec* TopoI, except

that (i) within the range of concentrations used, up to 9 protein-DNA complexes could be resolved on a polyacrylamide gel, and (ii) *Ec* TopoI can bind to the topoisomer T₀ (Figure 3b and Table 2). Altogether, our results show that without magnesium hRPA and *Ec* TopoI are sensitive to the negative supercoiling of the DNA and that the higher is the negative supercoiling, the greater the number of proteins per dsMC.

3. Interaction between topoisomers of dsMCs and hRPA in the presence of divalent cations

Since in the absence of magnesium, the extent of binding of hRPA depended on supercoiling (Figure 3a, Table 1), we investigated whether this dependency persisted in a magnesium-containing buffer. To this end, nicked, T₀, T₋₁, T₋₂, T₋₃, T₋₄ and a mixture of T₋₄ and T₋₅ DNA were incubated with increasing amounts of hRPA and the products were resolved by electrophoresis on a polyacrylamide gel (Figure 4a). In these experiments, magnesium was not only present in the reaction buffer but also in the polyacrylamide gels and in the electrophoresis buffer. The results indicated that hRPA interacted with nicked dsMC, giving rise to a complex that migrated more slowly than naked DNA (hRPA-nicked dsMC complex, Figure 4a). Surprisingly, with the T₋₁ topoisomer, in addition to the hRPA-nicked dsMC complex that was caused by the presence of nicked dsMCs in our preparation, two radiolabeled species (C_a and C_b, Figure 4a) that migrated faster than naked dsMC could be isolated. The same species were also formed with the other topoisomers of dsMCs. The appearance of T₋₁ topoisomers in the hRPA titration performed with T_i (i ≤ -2) topoisomers was probably due to a small contamination of our preparation of hRPA by *Escherichia coli* topoisomerases.

To characterize the C_a and C_b species, we first checked that they were sensitive to proteinase K treatment. The C₁ and C₂ species were indeed sensitive to the proteinase K, as was the hRPA-nicked dsMC complex, (Figure 4b), indicating that the C_a and C_b species were nucleoprotein complexes.

The topological state of the DNA in the C_a and C_b complexes formed with T₋₂ and a mixture of T_{-4/-5} topoisomers was next examined (Figures 5a-d). For this, the DNA was electro-eluted from the gel pieces corresponding to the C_a and C_b complexes and its topological state was analyzed by electrophoresis. The results showed that the C_a complex contained more than 90% of the T₋₁ dsMC and the C_b complex a mixture of the T₋₁ dsMC and the initial substrate(s) of the reaction (Figures 5b and d). Interestingly for

the experiment performed with the mixture of T_{-4/-5} topoisomers, the naked T_{-4/-5} topoisomers were clearly more abundant than the naked T₋₁ topoisomer, representing ~ 70% of the unbound species (Figure 5c). In the C_b nucleoprotein species, the T_{-4/-5} topoisomers represented ~ 30% of the topoisomers (Figure 5d). This result suggests unexpectedly that the hyper-negatively supercoiled substrates (T_{-4/-5}) are not the preferred substrates of hRPA.

Finally, to investigate whether the topology of the dsMCs changed upon binding of hRPA, an “in gel” relaxation assay was carried out. We (1) incubated the T₋₁ topoisomer with hRPA, (2) separated the C_a and C_b complexes by gel electrophoresis, (3) cut the gel bands corresponding to the C_a and C_b species, (4) incubated them with either *Ec* TopoI or wheat germ Topoisomerase, (5) electro-eluted the dsMCs from the gel pieces, and finally (6) measured the relative linking numbers of the topoisomerase products by electrophoresis on a polyacrylamide gel. The results indicate that binding of hRPA on T₋₁ topoisomer does not induce a topological change in the dsMC since the products of the “in gel” relaxation reaction are T₋₁ with *Ec* TopoI and T_{0/-1} with wheat germ Topoisomerase (Figures 5e and f), as expected (28).

DISCUSSION

Our goals in this study were first to locate at the nucleotide level, any structural changes induced by the hyper-negative supercoiling, by using two enzymatic probes, Nuclease SI and DNase I, and second to characterize the effect of hyper-negative supercoiling on the binding of two proteins known to bind single-stranded DNA, the hRPA heterotrimer and *Ec* TopoI. These two effects of hyper-negative supercoiling were investigated in the absence and in the presence of magnesium since divalent cations influence the structure of DNA, and more specifically its topology. Topoisomers of dsMCs of 235 base pairs were used as DNA substrates.

The seven topoisomers prepared span differences of topology ranging from relaxed (T₀) to hyper-negatively supercoiled DNA (T₋₃, T₋₄, T₋₅ and T₋₆). Under native conditions the results of electrophoresis (Supplementary Figure 1) are in agreement with previous structural studies that showed the introduction of writhe by the magnesium and a transition towards a more compact and highly writhed structure (≥ 3 writhe sites) when negative supercoiling was introduced into the dsMCs (25, 31). In

addition and as expected, our results show that the activity of Nuclease SI and DNase I depends on negative supercoiling: for Nuclease SI, the higher is the negative supercoiling, the higher is the nuclease activity (Figure 1a). An opposite relationship between negative supercoiling and activity is observed with DNase I (Figure 1b). The identification of the sites incised by Nuclease SI and DNase I has made it possible to locate the structural changes induced by the hyper-negative supercoiling on the topoisomer DNA sequence. Our results indicate that the structural changes are first subtle and second are not spread all along the sequence but cover a region of ~ 40 base pairs (Figure 2b). The equal reactivity of the topoisomers with DMS (Supplementary Figures 2b and c) is coherent with the subtle changes revealed by the nuclease probes. The results obtained with both types of probes are also consistent with the identification of a limited number (≤ 2) of regions cleaved by Nuclease SI on SV40 DNA (53). The nature of the hyper- and hypo-sensitive regions and the absence of a defined double-stranded region becoming hyper-sensitive to Nuclease SI upon the introduction of negative supercoils (Figure 2b) do not support the presence of denatured DNA bubbles, that in theory, should lead to a stronger and weaker reactivity of Nuclease SI and DNase I, respectively, on the two complementary strands. Similar studies performed in the absence of divalent cations used other DNA substrates, probes and methods to quantify the non-base-paired regions in supercoiled DNAs (53–56). They suggested the presence of localized regions of denaturation, that in the case of the supercoiled DNA of the bacteriophage PM2 overlapped with the earliest melting regions of the relaxed DNA (56). Similarly, in the case of the dsMCs produced by the HMGB2 (high mobility group of family B2)-induced ligation of a 75 base pair fragment, a denatured region of ~ 15 nucleotides has been mapped in a magnesium-free environment with chemical probes (33). Nucleases digestions performed on dsMCs of various sizes (from 64 to 100 base pairs) and topologies combined with theoretical analyses gave a global picture of the DNA with local unwound structures, which, depending on the size of the circle, may take the form of a small region with opened base pairs or of a sharp kink of the double helix that preserves base pairing (19). A recent electro cryo-tomography study reported a conformational heterogeneity of topoisomers of dsMCs, with a distribution of conformations specific of each topoisomer (25). Molecular dynamics simulations suggested, for a population of topoisomers with a given superhelicity, large fluctuations in the extent of writhe and in the level of compaction. This structural feature might also

exist with our smaller topoisomers, and the half-time life of the population of topoisomers with denatured regions and unpaired bases of significant length might be too short to be trapped by the Nuclease SI or the DNase I.

Whereas the structural changes associated with the introduction of negative supercoiling in a magnesium-containing buffer confer a higher sensitivity to Nuclease SI (Figure 1a), they surprisingly do not seem to stimulate the binding of hRPA (Figures 4 and 5). The determination of the topological state of the dsMCs present in the C_a and C_b nucleoprotein complexes (Figures 5a-d) strongly suggests that the best substrate for hRPA is the T₋₁ topoisomer, the DNA topoisomer that has the weakest reactivity towards Nuclease SI (Figure 1a). A similar behavior has been observed in the case of the interaction of histone f1 with supercoiled DNA. In addition to being influenced by negative supercoiling, this interaction is indeed modulated by the concentration of magnesium, with a decrease of the extent of the interaction as the concentration of magnesium increases (57). hRPA binds single-stranded DNA in a sequential manner, adopting first a compact (8-10 nucleotides occluded), next an elongated compacted (13-22 nucleotides occluded) and finally an elongated extended conformation (30 nucleotides occluded) (58). It is possible that the hyper-negative supercoiling, by inducing the formation of a compact structure with a high degree of writhe prevents the sequential binding of hRPA along the DNA and the formation of the stable elongated extended conformation occluding 30 nucleotides. The negatively supercoiled T₋₁ topoisomer may have a conformational dynamics that allows hRPA to go through its three sequential binding stages. In contrast, in the absence of magnesium, the higher is the negative supercoiling, the more hRPA bound per dsMC (Figure 3a). Under these reaction conditions, bases are more accessible to DMS and possibly not well paired with their complementary bases on the opposite strand (Supplementary Figure 3a). Furthermore, the stabilization of the double helix by counter ions does not exist and the negative supercoiling is mostly absorbed by twist changes (Supplementary Figure 1a). These ionic and conformational conditions may then favor the assembly of several hRPA with an elongated extended conformation. This hypothesis is consistent with binding behavior of hRPA on G-quadruplex structures (59), double-stranded DNA containing damaged bases, mismatches or short bubbles (60-62).

Topoisomerases of family IA have a better relaxing activity on hyper-negatively supercoiled DNA. As a matter of fact, the topoisomerase III α from *Drosophila melanogaster* can only relax hyper-negatively supercoiled DNA to negatively supercoiled DNA (15). Our data strongly suggest that this preference does not rely on the presence of a higher proportion of topoisomerase initial substrate (single-stranded DNA) on the hyper-negatively supercoiled DNA. Instead, we propose that the preference of topoisomerases of family IA for hyper-negatively supercoiled DNA relies on the presence of specific secondary structures that constitute better substrates for these enzymes. These specific structures may take the form of a flipped base within a strongly bent region. For instance, solvated molecular dynamic simulations performed with a highly writhed dsMC led to a model of topoisomer containing a sharp bend with a disrupted base pair. Alternatively, the high sensitivity of dsMCs of 64-65 base pairs towards the BAL31 nuclease has been attributed to a kink resulting from the inherent bending strain in the DNA (19). On our highly writhed dsMCs a sharp bend or kink might be located at the site becoming hyper-sensitive to both Nuclease SI and DNase I (Figure 2).

As expected, the hRPA-dsMCs complexes observed in the absence of magnesium have a lower mobility than naked DNA, and their mobility decreases as the concentration of hRPA increases (Figure 3a). In contrast, the migration profile of the hRPA-dsMCs complexes assembled in the presence of magnesium is unusual from two perspectives. First, our results indeed show that for all topoisomers, the C_a complex that forms at a weaker concentration of hRPA than the C_b complex, has a reduced mobility compared to the C_b complex, and second, both complexes migrate faster than the T₋₁ and T₋₂ topoisomers (Figure 4a). It is known that the wrapping of the DNA double helix around histones changes the topology of the DNA (35). In the case of the hRPA-dsMCs complexes, it is possible that the binding of hRPA to the dsMCs induces a topology change leading to complexes that have an unusual behavior in terms of mobility in an electric field. However, the results of the “in gel” relaxation assay (Figures 5e and f) exclude the possibility that the binding of hRPA changes the topology of the DNA. It is therefore possible that the aberrant migration of the C_a and C_b hRPA-dsMC complexes with respect to each other and with respect to the T₋₁ and T₋₂ topoisomers stems from the structure itself of the complexes.

CONCLUSION

The structural features of hyper-negatively supercoiled dsMCs in a magnesium-containing buffer and highlighted by our experiments are the subtle, very short (over a few nucleotides or base pairs) and clustered deformations, and probably not the exposure of long segments of single-stranded DNA. In addition our probing experiments demonstrate that such a conformation can either favor binding of proteins as exemplified by Nuclease SI, or in contrast inhibit protein binding as exemplified by the hRPA heterotrimer.

FUNDING

This work was supported by INSERM, CNRS MNHN and LNCC. A. Millet was supported by a fellowship from the Muséum national d'Histoire naturelle and the Fondation pour la Recherche Médicale (FDT 20160435361).

ACKNOWLEDGEMENTS

The authors thank their respective team members for their support and fruitful discussions.

AUTHOR CONTRIBUTIONS

E.D. conceived the project. A.M., C.S. and E.D. performed the experiments and analyzed the data. E.D. wrote the paper with input from C.S. and A.M. All authors revised the manuscripts.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

REFERENCES

1. Wang,J.C. (1974) Interactions between twisted DNAs and enzymes: the effects of superhelical turns. *J. Mol. Biol.*, **87**, 797–816.
2. Falco,S.C., Zivin,R. and Rothman-Denes,L.B. (1978) Novel template requirements of N4 virion RNA polymerase. *Proc. Natl. Acad. Sci. U. S. A.*, **75**, 3220–4.
3. Sumida-Yasumoto,C. and Hurwitz,J. (1977) Synthesis of phiX174 viral DNA in vitro depends on phiX replicative form DNA. *Proc. Natl. Acad. Sci. U. S. A.*, **74**, 4195–9.
4. Marians,K.J., Ikeda,J.E., Schlagman,S. and Hurwitz,J. (1977) Role of DNA gyrase in phiX replicative-form replication in vitro. *Proc. Natl. Acad. Sci. U. S. A.*, **74**, 1965–8.
5. Mizuuchi,K., Gellert,M. and Nash,H.A. (1978) Involvement of supertwisted DNA in

- integrative recombination of bacteriophage lambda. *J. Mol. Biol.*, **121**, 375–92.
- 6. Hays,J.B. and Boehmer,S. (1978) Antagonists of DNA gyrase inhibit repair and recombination of UV-irradiated phage lambda. *Proc. Natl. Acad. Sci. U. S. A.*, **75**, 4125–9.
 - 7. Eisenberg,S., Scott,J.F. and Kornberg,A. (1976) An enzyme system for replication of duplex circular DNA: the replicative form of phage phi X174. *Proc. Natl. Acad. Sci. U. S. A.*, **73**, 1594–7.
 - 8. Botchan,P., Wang,J.C. and Echols,H. (1973) Effect of circularity and superhelicity on transcription from bacteriophage lambda DNA. *Proc. Natl. Acad. Sci. U. S. A.*, **70**, 3077–81.
 - 9. Mizuuchi,K. and Nash,H.A. (1976) Restriction assay for integrative recombination of bacteriophage lambda DNA in vitro: requirement for closed circular DNA substrate. *Proc. Natl. Acad. Sci. U. S. A.*, **73**, 3524–8.
 - 10. Landy,A. and Ross,W. (1977) Viral integration and excision: structure of the lambda att sites. *Science*, **197**, 1147–60.
 - 11. Holloman,W.K., Wiegand,R., Hoessli,C. and Radding,C.M. (1975) Uptake of homologous single-stranded fragments by superhelical DNA: a possible mechanism for initiation of genetic recombination. *Proc. Natl. Acad. Sci. U. S. A.*, **72**, 2394–8.
 - 12. Lim,H.M., Lewis,D.E.A., Lee,H.J., Liu,M. and Adhya,S. (2003) Effect of varying the supercoiling of DNA on transcription and its regulation. *Biochemistry*, **42**, 10718–25.
 - 13. Delius,H., Mantell,N.J. and Alberts,B. (1972) Characterization by electron microscopy of the complex formed between T4 bacteriophage gene 32-protein and DNA. *J. Mol. Biol.*, **67**, 341–50.
 - 14. Santos-Pereira,J.M. and Aguilera,A. (2015) R loops: new modulators of genome dynamics and function. *Nat. Rev. Genet.*, **16**, 583–97.
 - 15. Chen,S.H., Wu,C.-H., Plank,J.L. and Hsieh,T. (2012) Essential functions of C terminus of Drosophila Topoisomerase III α in double holliday junction dissolution. *J. Biol. Chem.*, **287**, 19346–53.
 - 16. Tan,K., Zhou,Q., Cheng,B., Zhang,Z., Joachimiak,A. and Tse-Dinh,Y.-C. (2015) Structural basis for suppression of hypernegative DNA supercoiling by E. coli topoisomerase I. *Nucleic Acids Res.*, **43**, 11031–46.
 - 17. Demurtas,D., Amzallag,A., Rawdon,E.J., Maddocks,J.H., Dubochet,J. and Stasiak,A. (2009) Bending modes of DNA directly addressed by cryo-electron microscopy of

- DNA minicircles. *Nucleic Acids Res.*, **37**, 2882–93.
- 18. Benedetti,F., Japaridze,A., Dorier,J., Racko,D., Kwapich,R., Burnier,Y., Dietler,G. and Stasiak,A. (2015) Effects of physiological self-crowding of DNA on shape and biological properties of DNA molecules with various levels of supercoiling. *Nucleic Acids Res.*, **43**, 2390–9.
 - 19. Du,Q., Kotlyar,A. and Vologodskii,A. (2008) Kinking the double helix by bending deformation. *Nucleic Acids Res.*, **36**, 1120–8.
 - 20. Panayotatos,N. and Wells,R.D. (1981) Cruciform structures in supercoiled DNA. *Nature*, **289**, 466–70.
 - 21. Lilley,D.M. (1983) Structural perturbation in supercoiled DNA: hypersensitivity to modification by a single-strand-selective chemical reagent conferred by inverted repeat sequences. *Nucleic Acids Res.*, **11**, 3097–112.
 - 22. Lilley,D.M. (1981) In vivo consequences of plasmid topology. *Nature*, **292**, 380–2.
 - 23. Amzallag,A., Vaillant,C., Jacob,M., Unser,M., Bednar,J., Kahn,J.D., Dubochet,J., Stasiak,A. and Maddocks,J.H. (2006) 3D reconstruction and comparison of shapes of DNA minicircles observed by cryo-electron microscopy. *Nucleic Acids Res.*, **34**, e125.
 - 24. Harris,S.A., Laughton,C.A. and Liverpool,T.B. (2008) Mapping the phase diagram of the writhe of DNA nanocircles using atomistic molecular dynamics simulations. *Nucleic Acids Res.*, **36**, 21–9.
 - 25. Irobalieva,R.N., Fogg,J.M., Catanese,D.J., Catanese,D.J., Sutthibutpong,T., Chen,M., Barker,A.K., Ludtke,S.J., Harris,S.A., Schmid,M.F., *et al.* (2015) Structural diversity of supercoiled DNA. *Nat. Commun.*, **6**, 8440.
 - 26. Bednar,J., Furrer,P., Stasiak,A., Dubochet,J., Egelman,E.H. and Bates,A.D. (1994) The twist, writhe and overall shape of supercoiled DNA change during counterion-induced transition from a loosely to a tightly interwound superhelix. Possible implications for DNA structure in vivo. *J. Mol. Biol.*, **235**, 825–47.
 - 27. Bates,A.D., Noy,A., Piperakis,M.M., Harris,S.A. and Maxwell,A. (2013) Small DNA circles as probes of DNA topology. *Biochem. Soc. Trans.*, **41**, 565–70.
 - 28. Millet,A., Strauss,F. and Delagoutte,E. (2015) Use of double-stranded DNA mini-circles to characterize the covalent topoisomerase-DNA complex. *Sci. Rep.*, **5**, 13154.
 - 29. Hamiche,A. and Prunell,A. (1992) Chromatin reconstitution on small DNA rings. V. DNA thermal flexibility of single nucleosomes. *J. Mol. Biol.*, **228**, 327–337.
 - 30. Anderson,B.G. and Stivers,J.T. (2014) Variola type IB DNA topoisomerase: DNA

binding and supercoil unwinding using engineered DNA minicircles. *Biochemistry*, **53**, 4302–15.

31. Fogg,J.M., Kolmakova,N., Rees,I., Magonov,S., Hansma,H., Perona,J.J. and Zechiedrich,E.L. (2006) Exploring writhe in supercoiled minicircle DNA. *J. Phys. Condens. Matter*, **18**, S145–S159.
32. Douc-Rasy,S., Kayser,A., Riou,J.F. and Riou,G. (1986) ATP-independent type II topoisomerase from trypanosomes. *Proc. Natl. Acad. Sci. U. S. A.*, **83**, 7152–7156.
33. Nakamura,Y., Shimizu,M. and Yoshida,M. (2002) Distorted DNA structures induced by HMGB2 possess a high affinity for HMGB2. *J. Biochem.*, **131**, 153–60.
34. Douc-Rasy,S., Kolb,A. and Prunell,A. (1989) Protein-induced unwinding of DNA: measurement by gel electrophoresis of complexes with DNA minicircles. Application to restriction endonuclease EcoRI, catabolite gene activator protein and lac repressor. *Nucleic Acids Res.*, **17**, 5173–5189.
35. Hamiche,A., Carot,V., Alilat,M., De Lucia,F., O'Donohue,M.F., Revet,B. and Prunell,A. (1996) Interaction of the histone (H3-H4)₂ tetramer of the nucleosome with positively supercoiled DNA minicircles: Potential flipping of the protein from a left-to a right-handed superhelical form. *Proc. Natl. Acad. Sci. U. S. A.*, **93**, 7588–93.
36. Sivolob,A., De Lucia,F., Révet,B. and Prunell,A. (1999) Nucleosome dynamics. II. High flexibility of nucleosome entering and exiting DNAs to positive crossing. An ethidium bromide fluorescence study of mononucleosomes on DNA minicircles. *J. Mol. Biol.*, **285**, 1081–1099.
37. Sivolob,A. and Prunell,A. (2000) Nucleosome dynamics V. Ethidium bromide versus histone tails in modulating ethidium bromide-driven tetrasome chiral transition. A fluorescence study of tetrasomes on DNA minicircles. *J. Mol. Biol.*, **295**, 41–53.
38. Pasi,M. and Lavery,R. (2016) Structure and dynamics of DNA loops on nucleosomes studied with atomistic, microsecond-scale molecular dynamics. *Nucleic Acids Res.*, **44**, 5450–6.
39. Pasi,M., Mornico,D., Volant,S., Juchet,A., Batisse,J., Bouchier,C., Parissi,V., Ruff,M., Lavery,R. and Lavigne,M. (2016) DNA minicircles clarify the specific role of DNA structure on retroviral integration. *Nucleic Acids Res.*, **44**, 7830–47.
40. Randall,G.L., Zechiedrich,L. and Pettitt,B.M. (2009) In the absence of writhe, DNA relieves torsional stress with localized, sequence-dependent structural failure to preserve B-form. *Nucleic Acids Res.*, **37**, 5568–77.

41. Adrian,M., ten Heggeler-Bordier,B., Wahli,W., Stasiak,A.Z., Stasiak,A. and Dubochet,J. (1990) Direct visualization of supercoiled DNA molecules in solution. *EMBO J.*, **9**, 4551–4.
42. Audry,J., Maestroni,L., Delagoutte,E., Gauthier,T., Nakamura,T.M., Gachet,Y., Saintomé,C., Géli,V. and Coulon,S. (2015) RPA prevents G-rich structure formation at lagging-strand telomeres to allow maintenance of chromosome ends. *EMBO J.*, **34**, 1942–1958.
43. Safa,L., Gueddouda,N.M., Thiébaut,F., Delagoutte,E., Petrusova,I., Lavrik,O., Mendoza,O., Bourdoncle,A., Alberti,P., Riou,J.-F., *et al.* (2016) 5' to 3' Unfolding Directionality of DNA Secondary Structures by Replication Protein A: G-QUADRUPLEXES AND DUPLEXES. *J. Biol. Chem.*, **291**, 21246–21256.
44. Shure,M. and Vinograd,J. (1976) The number of superhelical turns in native virion SV40 DNA and minicol DNA determined by the band counting method. *Cell*, **8**, 215–26.
45. Maxam,A.M. and Gilbert,W. (1977) A new method for sequencing DNA. *Proc. Natl. Acad. Sci. U. S. A.*, **74**, 560–4.
46. Horowitz,D.S. and Wang,J.C. (1984) Torsional rigidity of DNA and length dependence of the free energy of DNA supercoiling. *J. Mol. Biol.*, **173**, 75–91.
47. Dalby,A.B., Goodrich,K.J., Pfingsten,J.S. and Cech,T.R. (2013) RNA recognition by the DNA end-binding Ku heterodimer. *RNA*, **19**, 841–851.
48. Bandhu,A., Ganguly,T., Jana,B., Mondal,R. and Sau,S. (2010) Regions and Residues of an Asymmetric Operator DNA Interacting with the Monomeric Repressor of Temperate Mycobacteriophage L1. *Biochemistry*, **49**, 4235–4243.
49. Shaw,P.E. and Stewart,A.F. (2009) Identification of Protein/DNA Contacts with Dimethyl Sulfate: Methylation Protection and Methylation Interference. In:pp. 97–104.
50. Esteban,J.A., Salas,M. and Blanco,L. (1992) Activation of S1 nuclease at neutral pH. *Nucleic Acids Res.*, **20**, 4932.
51. Jeon,J.-H., Adamcik,J., Dietler,G. and Metzler,R. (2010) Supercoiling Induces Denaturation Bubbles in Circular DNA. *Phys. Rev. Lett.*, **105**, 208101.
52. Takahashi,S., Motooka,S., Usui,T., Kawasaki,S., Miyata,H., Kurita,H., Mizuno,T., Matsuura,S., Mizuno,A., Oshige,M., *et al.* (2015) Direct Single-Molecule Observations of Local Denaturation of a DNA Double Helix under a Negative Supercoil State. *Anal.*

Chem., **87**, 3490–3497.

53. Beard,P., Morrow,J.F. and Berg,P. (1973) Cleavage of circular, superhelical simian virus 40 DNA to a linear duplex by S1 nuclease. *J. Virol.*, **12**, 1303–13.
54. Dean,W.W. and Lebowitz,J. (1971) Partial alteration of secondary structure in native superhelical DNA. *Nat. New Biol.*, **231**, 5–8.
55. Beerman,T.A. and Lebowitz,J. (1973) Further analysis of the altered secondary structure of superhelical DNA. Sensitivity to methylmercuric hydroxide a chemical probe for unpaired bases. *J. Mol. Biol.*, **79**, 451–70.
56. Brack,C., Bickle,T.A. and Yuan,R. (1975) The relation of single-stranded regions in bacteriophage PM2 supercoiled DNA to the early melting sequences. *J. Mol. Biol.*, **96**, 693–702.
57. Vogel,T. and Singer,M.F. (1975) Interaction of f1 histone with superhelical DNA. *Proc. Natl. Acad. Sci. U. S. A.*, **72**, 2597–600.
58. Chen,R. and Wold,M.S. (2014) Replication protein A: single-stranded DNA's first responder: dynamic DNA-interactions allow replication protein A to direct single-strand DNA intermediates into different pathways for synthesis or repair. *Bioessays*, **36**, 1156–61.
59. Safa,L., Delagoutte,E., Petrusseva,I., Alberti,P., Lavrik,O., Riou,J.-F. and Saintomé,C. (2014) Binding polarity of RPA to telomeric sequences and influence of G-quadruplex stability. *Biochimie*, **103**, 80–8.
60. Lao,Y., Gomes,X. V., Ren,Y., Taylor,J.-S. and Wold,M.S. (2000) Replication Protein A Interactions with DNA. III. Molecular Basis of Recognition of Damaged DNA †. *Biochemistry*, **39**, 850–859.
61. Iftode,C. and Borowiec,J.A. (2000) 5' → 3' Molecular Polarity of Human Replication Protein A (hRPA) Binding to Pseudo-Origin DNA Substrates †. *Biochemistry*, **39**, 11970–11981.
62. Wang,M., Mahrenholz,A. and Lee,S.H. (2000) RPA stabilizes the XPA-damaged DNA complex through protein-protein interaction. *Biochemistry*, **39**, 6433–9.

TABLES

Table 1: % of DNA either naked or engaged in complexes with hRPA

$[hRPA] = 100 \text{ nM}$

	Naked DNA	Bound DNA			
		C1	C2	C3	C3+C4+C5
T ₀	100	/	/	/	/
T ₋₂	64	36	/	/	/
T _{-3/-4}	43	48	5	4	/
T _{-4/-5}	32	5	13	50	/
T _{-5/-6}	44	/	/	/	56

Table 2: % of DNA either naked or engaged in complexes with *Ec* TopoI

$[Ec \text{ TopoI}] = 28 \text{ mU}/\mu\text{L}$

	Naked DNA	Bound DNA			
		C1	C2	C3	C4
T ₀	98	2		/	/
T ₋₂	39.5	53	7.5	/	/
T _{-3/-4}	38	16	46	/	/
T _{-4/-5}	44	13.5	25	17.5	/
T ₋₆	20.5	6	20	43.5	10

FIGURE LEGENDS

Figure 1: Digestion of topoisomers of dsMCs by nucleases

The topoisomers were incubated with increasing amounts of nuclease. At the end of the reaction, nucleases were removed from the reaction products. The DNAs were precipitated before being denatured by NaOH and resolved by electrophoresis on a native polyacrylamide gel containing chloroquine ($30 \mu\text{g mL}^{-1}$). Under these conditions, dsMCs were resolved from linear and circular single-stranded DNA. 1a. Reactivity of topoisomers of dsMCs towards Nuclease SI. Left panel: picture of the gel showing the degradation of T₀ and T_{-5/-6} topoisomers at increasing concentrations of Nuclease SI (0; 0.7; 2; 6.2; 18.5; 55; 170 mU μL^{-1}). Right panel: Quantification of the degradation of the topoisomers at increasing concentrations of Nuclease SI. The % of nicked DNA is plotted as a function of Nuclease SI concentration. Error bars correspond to the standard errors calculated from three independent experiments. 1b. Reactivity of topoisomers of dsMCs

towards DNase I. Left panel: picture of the gel showing the degradation of T_0 and $T_{-5/-6}$ topoisomers at increasing concentrations of DNase I (0; 1; 2.5; 5; 10 mU μL^{-1}). Right panel: Quantification of the degradation of the T_0 and $T_{-5/-6}$ topoisomers at increasing concentrations of DNase I. The % of nicked DNA is plotted as a function of DNase I concentration.

Figure 2: Sites of structural changes induced by the hyper-negative supercoiling detected by Nuclease SI and DNase I.

2a. The enzymatic probe used to map the fine structure of the T_{-2} and T_{-6} topoisomers is Nuclease SI. Nuclease SI is at 2 mU μL^{-1} and DNA at 0.5 nM. After the Nuclease SI reaction, the samples are treated to remove the proteins. The DNAs are precipitated and submitted to the double digestion as indicated (BglII+BamHI, or BamHI+HindIII) to only visualize DNA fragments from one of the two radiolabeled strands. The reaction products are analyzed on two different sequencing gels (8% to see long DNA fragments, 12% to see short DNA fragments) as indicated. G and G+A lanes correspond to the products of the Maxam and Gilbert reactions to identify specifically the guanines (G lanes) or the guanines and adenines (G+A lanes). The color code used to indicate sites sensitive to hyper-negative supercoiling is the following: Nuclease SI hyper-sensitive sites are indicated by a red rectangle, Nuclease SI hypo-sensitive sites by a green rectangle, and the site hyper-sensitive to DNase I and Nuclease SI by a blue rectangle. 2b. Same as 2a except that the dsMCs are the T_{-2} and T_{-6} topoisomers and the enzymatic probe is DNase I. DNase I is at 0.5 mU μL^{-1} and DNA at 0.5 nM. 2c. The hyper- and hypo-sensitive sites are shown on the sequence with the same color code used in 2a and 2b. The horizontal double-headed arrow designates the 37 base pair regions where the structural changes revealed by Nuclease SI and DNase I are concentrated.

Figure 3: Binding of hRPA and *Ec* TopoI to the topoisomers of dsMCs in the absence of magnesium

Each dsMC of indicated topology is at 0.5 nM and the DNA binding protein at the indicated concentration. For both titrations, the samples are analyzed by electrophoresis on native 4% polyacrylamide gels (29:1 = acrylamide:bisacrylamide mass ratio) made in TBE 0.5x. Electrophoresis is performed at 150 V, 4°C and for 4 h.

1a. Binding of hRPA. [hRPA] ranges from 0 to 300 nM (0; 3; 10; 30; 100; 300 nM). 1b. Binding of *Ec* TopoI. [*Ec* TopoI] ranges from 0 to 250 mU μL^{-1} (0; 3; 9; 28; 81; 250 mU μL^{-1}).

Figure 4: Interaction between hRPA and topoisomers of dsMCs in the presence of magnesium

The dashed arrows indicate the direction of the electrophoretic migration. 4a. The topoisomers of indicated topology are incubated with increasing amounts of hRPA. [hRPA] ranges from 0 to 1 μM (final concentration, 0; 30; 100; 300; 1000 nM). The reaction products are analyzed by electrophoresis on a native 4% polyacrylamide (29:1 = acrylamide:bisacrylamide mass ratio) gel made in TBMg (44.5 mM Tris-Base, 44.5 mM boric acid, 10 mM MgCl₂). Electrophoresis is performed at 150 V, 4°C and for 4 hours. Three complexes form upon mixing a given topoisomer with hRPA: the hRPA-nicked dsMC complex, the C_a and the C_b complexes. 4b. Same as 4a with a hRPA final concentration of 300 nM. Samples were (“+” lanes) or not (“-” lanes) treated with proteinase K (Prot K) before their analysis by electrophoresis on a native polyacrylamide gel. The three protein-DNA complexes that assemble on the nicked, T₋₁, T₋₂, T₋₄ and T_{-4/-5} DNA are sensitive to Proteinase K.

Figure 5: Characterization of the hRPA-dsMC complexes formed in the presence of magnesium

5a. DNA content of the C_a and C_b complexes. hRPA (final concentration, 300 nM) is incubated with the T₋₂ topoisomer (final concentration, 0.5 nM) and the different molecular species are resolved by electrophoresis on a native 4% polyacrylamide (29:1 = acrylamide:bisacrylamide mass ratio) gel made in TBMg (44.5 mM Tris-Base, 44.5 mM boric acid, 10 mM MgCl₂). Electrophoresis is performed at 150 V, 4°C and for 4 h. 5b. The DNAs contained in the gel bands corresponding to the C_a and C_b complexes in 5a are electro-eluted and analyzed by electrophoresis on a chloroquine-containing gel (6% acrylamide; 29:1 = acrylamide:bisacrylamide mass ratio; 30 $\mu\text{g mL}^{-1}$ of chloroquine; TBE 0.5x). 5c and d. Same as 5a and b, respectively, except that the DNA is the T_{-4/-5} topoisomer. 5e. Determination of the topological state of the dsMC in the C_a and C_b complexes assembled on the T₋₁ topoisomer. The species formed upon incubating hRPA (300 nM) with the T₋₁ topoisomer (0.5 nM) are resolved by electrophoresis on a 4%

native polyacrylamide (29:1 = acrylamide:bisacrylamide mass ratio) gel made in TBMg (44.5 mM Tris-Base, 44.5 mM boric acid, 10 mM MgCl₂). 5f. The gel pieces corresponding to the C_a and C_b complexes in 5e are treated with *Ec* TopoI (*Ec*) or wheat germ Topoisomerase (wg) before performing an electro-elution to recover the DNA. The DNA recovered is precipitated and analyzed by electrophoresis on a chloroquine-containing gel (6% acrylamide; 29:1 = acrylamide:bisacrylamide mass ratio; 30 µg mL⁻¹ of chloroquine; TBE 0.5x).

Supplementary Figure 1: Effect of magnesium or calcium on the conformations of the topoisomers of dsMCs

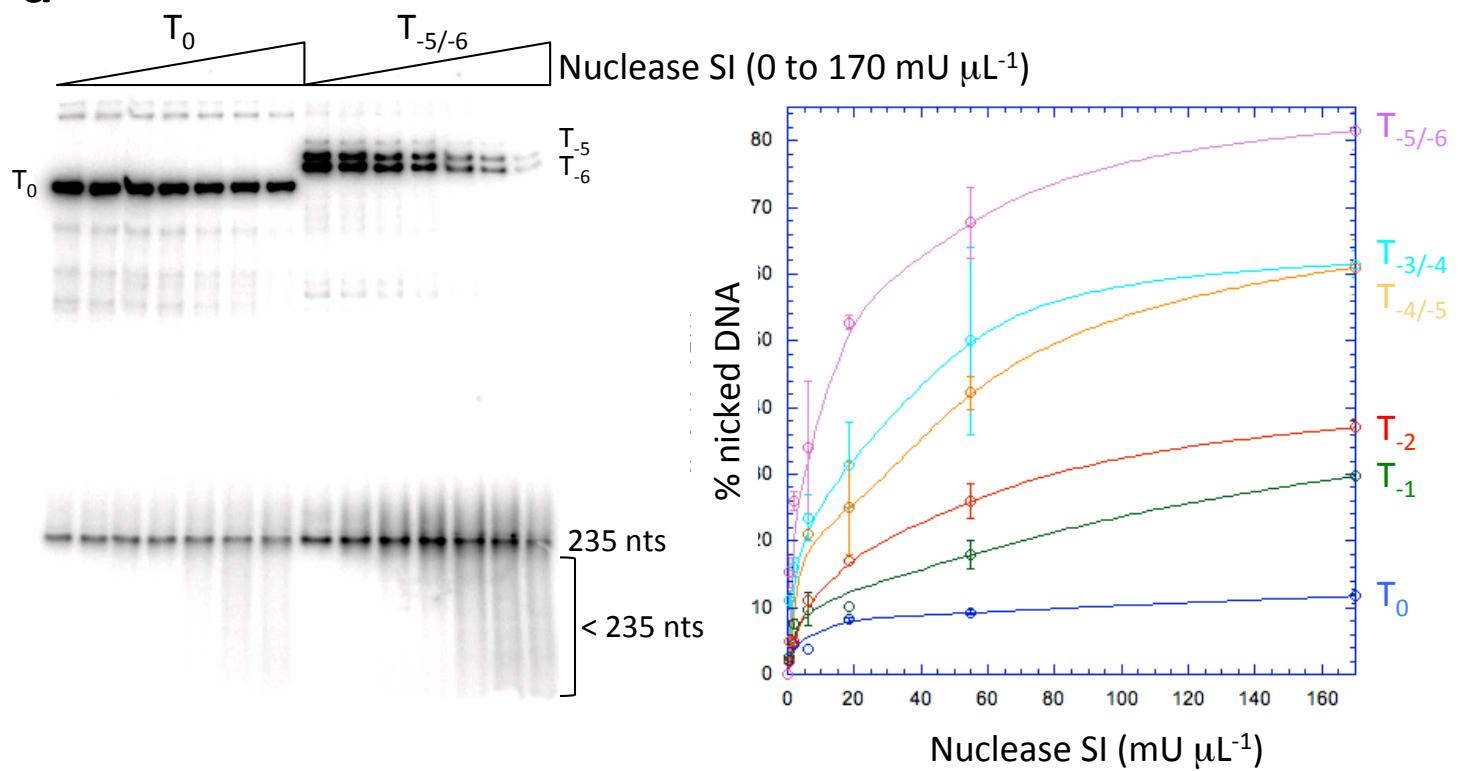
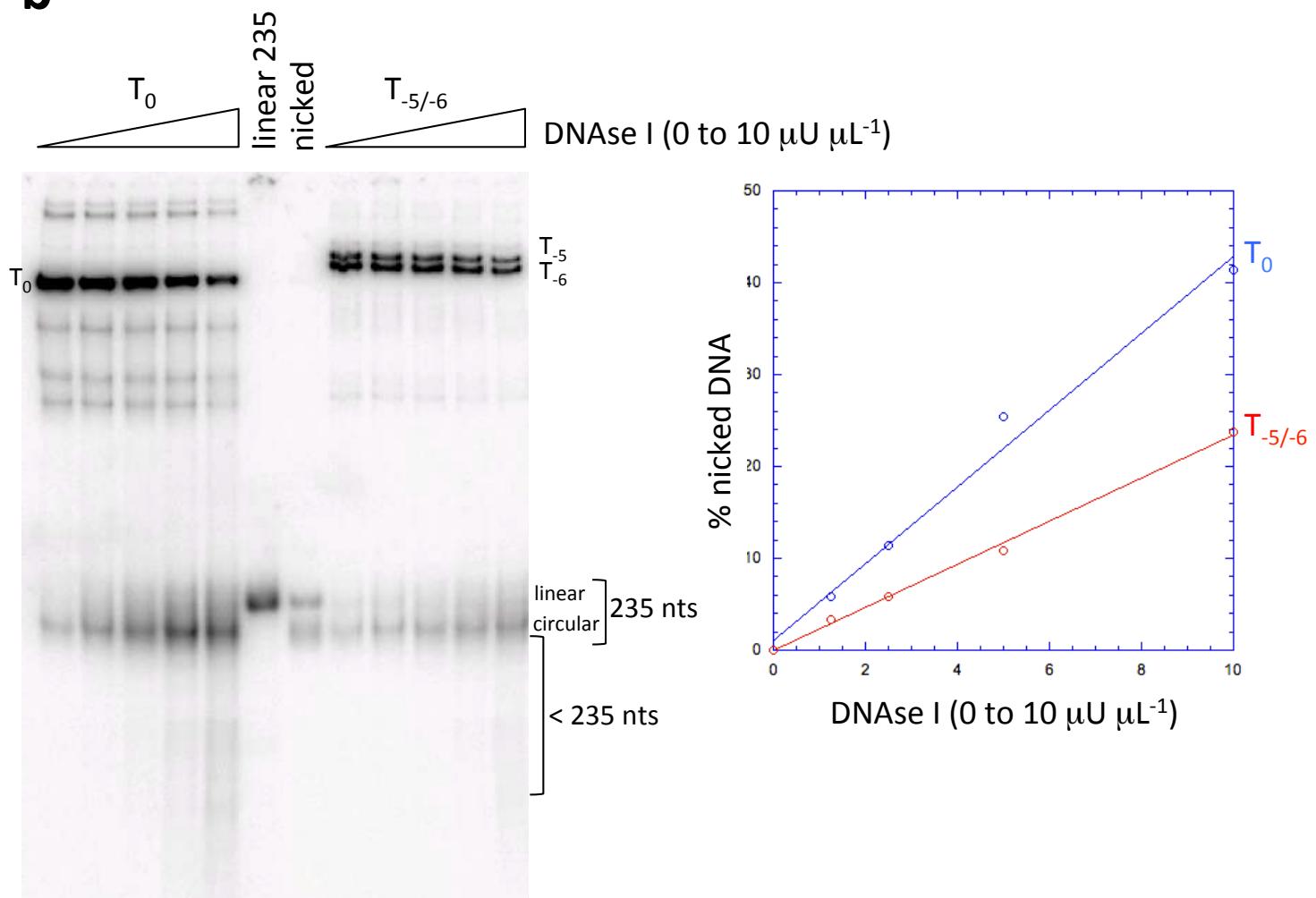
1a. Electrophoretic migration profile of the seven topoisomers on a native gel made in TBE 0.5x. 1b. Electrophoretic migration profile of the seven topoisomers in the presence of magnesium. Magnesium (10 mM) is present in the native polyacrylamide gel and in the running buffer. 1c. Electrophoretic migration profile of the seven topoisomers in the presence of calcium. Calcium (10 mM) is present in the native polyacrylamide gel and in the running buffer. In 1b and 1c, the topoisomers differ by their relative linking number ΔLk, ranging from 0 (T₀ topoisomer, lane 1) to -6 (T₋₆ topoisomer, lane 7) and electrophoreses are performed at 4°C.

Supplementary Figure 2: Reactivity of the DMS on the topoisomers of dsMCs in the absence or in the presence of magnesium

3a. The designated topoisomer (T₀, T₋₃, T_{-4/-5}) is incubated with DMS in the absence (black curve) or in the presence (blue curve) of magnesium. 3b. Comparison of the reactivity of three topoisomers (T₀, T₋₃, T_{-4/-5}) with DMS in the absence of magnesium. 3c. Comparison of the reactivity of three topoisomers (T₀, T₋₃, T_{-4/-5}) with DMS in the presence of magnesium. [DMS] ranges from 0 to 0.125%. The % of covalently closed dsMCs is represented as a function of [DMS]. Error bars correspond to the standard errors calculated from three independent experiments.

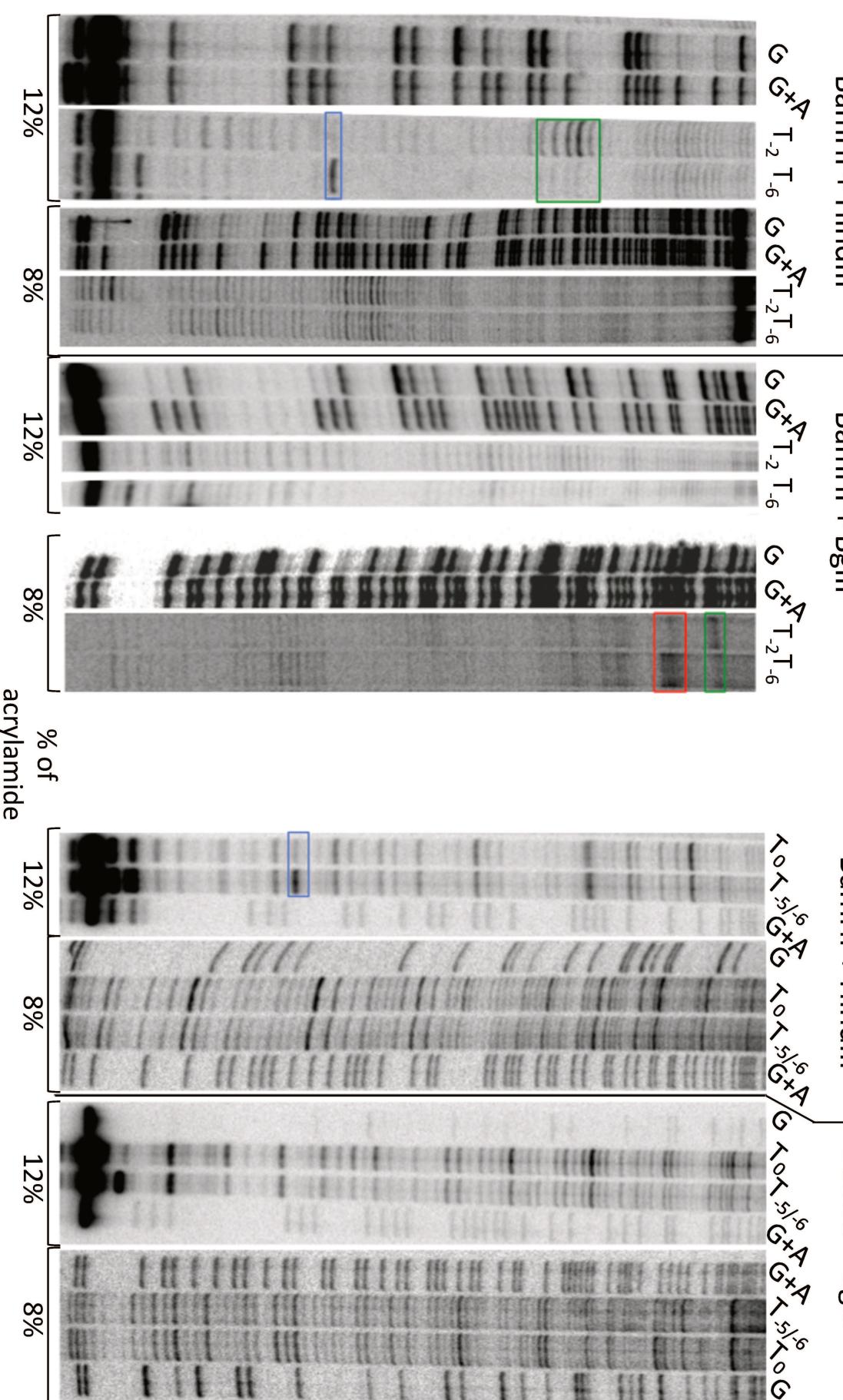
Supplementary Figure 3: Position of the BglII, BamHI and HindIII restriction sites along the dsMC sequence.

The strands that are radioactive upon the double digestion (BglII+BamHI, or BamHI+HindIII) are colored in red.

a**b****FIGURE 1**

a**Nuclease SI**

BamHI + HindIII BamHI + BgIII

 G $G^{x\Delta} T_2 T_6$ $G G^{x\Delta} T_{-2} T_6$ $G G^{x\Delta} T_2 T_{-6}$ $G G^{x\Delta} T_{-2} T_6$
 $\nwarrow \nwarrow \nwarrow G^{x\Delta} G$ $\nwarrow \nwarrow \nwarrow G^{x\Delta} / G$ $\nwarrow \nwarrow \nwarrow G^{x\Delta} G^{x\Delta} \nwarrow \nwarrow \nwarrow G$
**b****DNase I**

BamHI + HindIII BamHI + BgIII

 $\nwarrow \nwarrow \nwarrow G^{x\Delta} G$ $\nwarrow \nwarrow \nwarrow G^{x\Delta} / G$ $\nwarrow \nwarrow \nwarrow G^{x\Delta} G^{x\Delta} \nwarrow \nwarrow \nwarrow G$
a

5' -GATCCGGAGACGTGGTAGGTAGATCTGGCGCCCTCGA **G** CTTGGTCGTGACC **GGTATT** TTCGGGACCA CCTTACC TGTGGCGAGC-3'
3' -CTAGGGCTCTGCACCCATCCATCTAGACCCGGGGAGCT **C** GAACCAGCACTGG **CCATAA** AAGCCCTGGT **GGAATGG** ACAGCCGCTCG-5'

FIGURE 2

37 base pairs

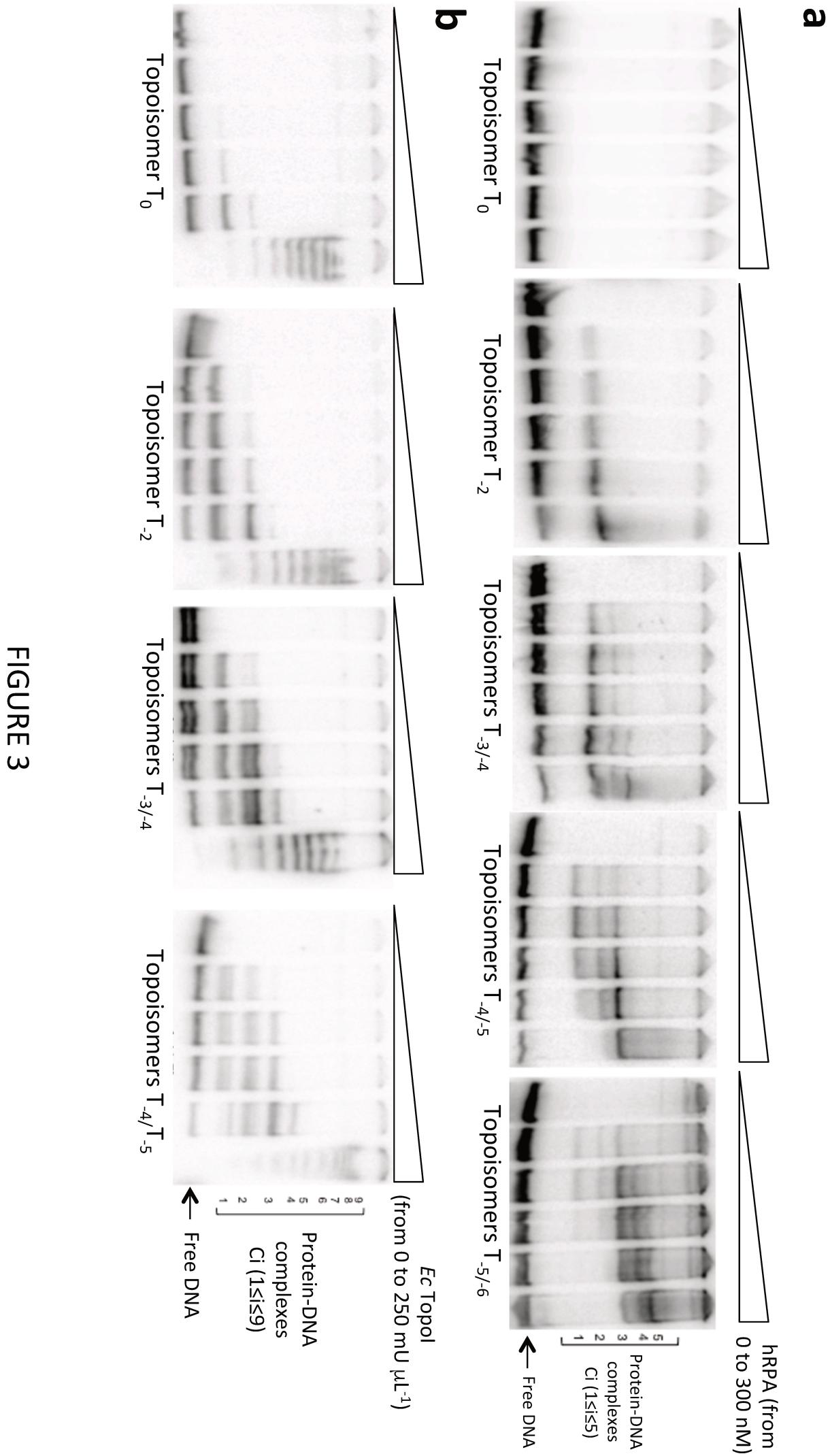
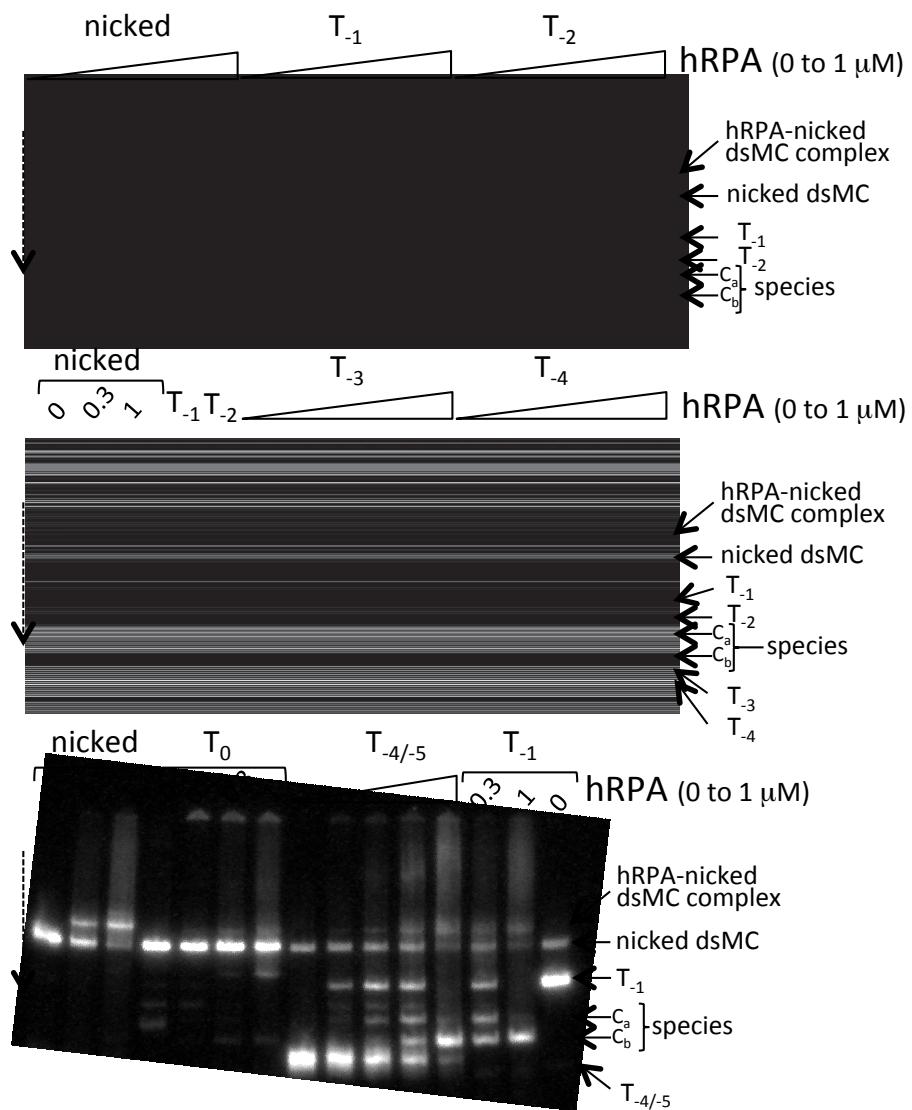
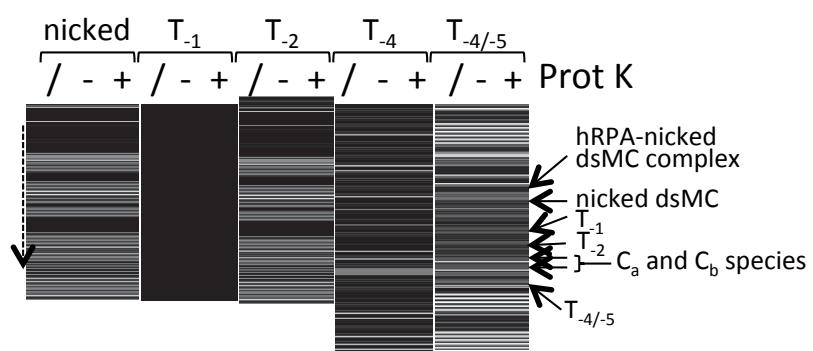


FIGURE 3

a**b****FIGURE 4**

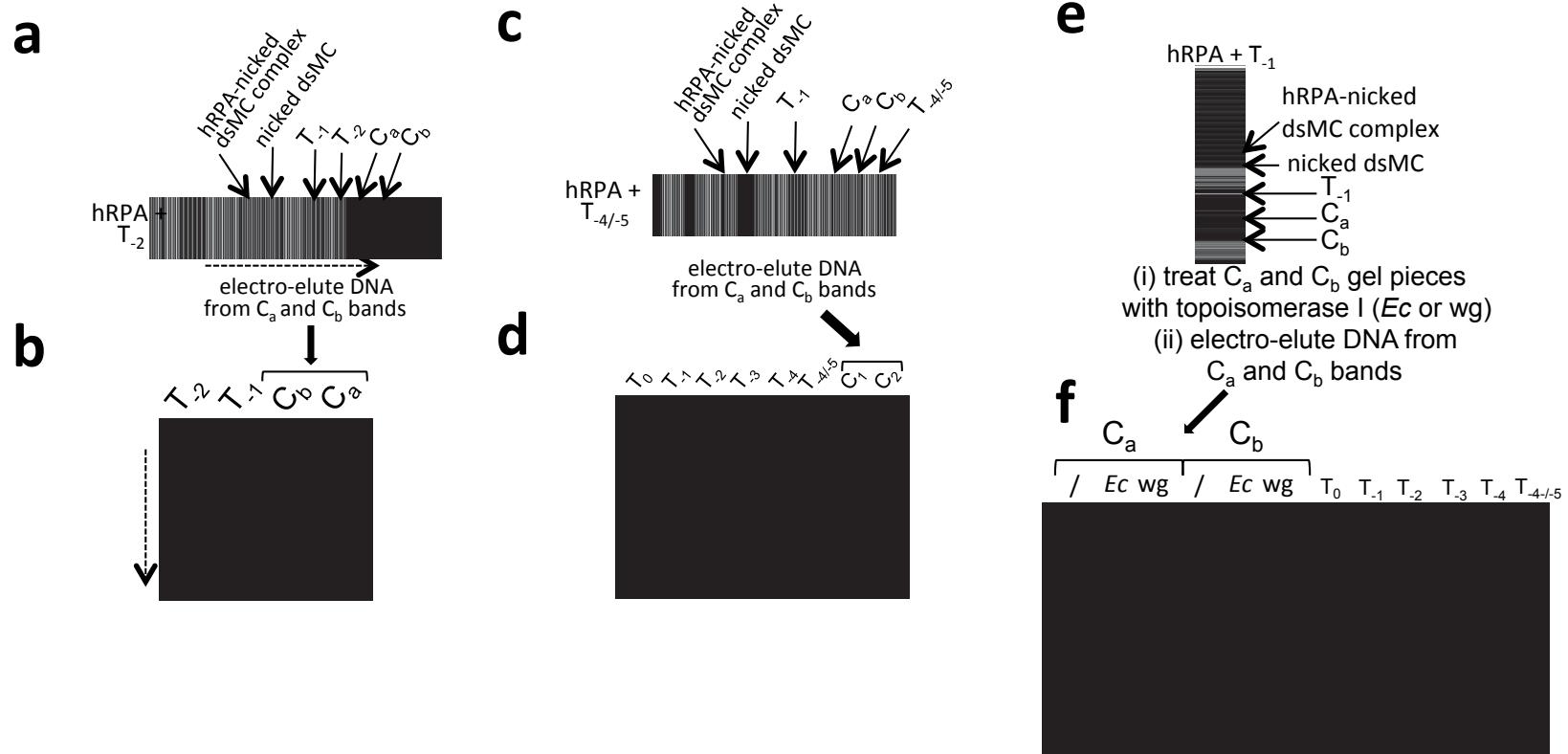
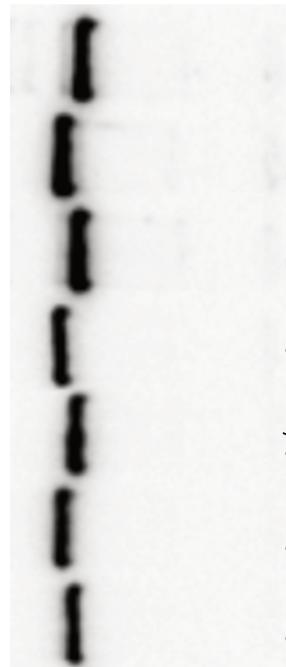


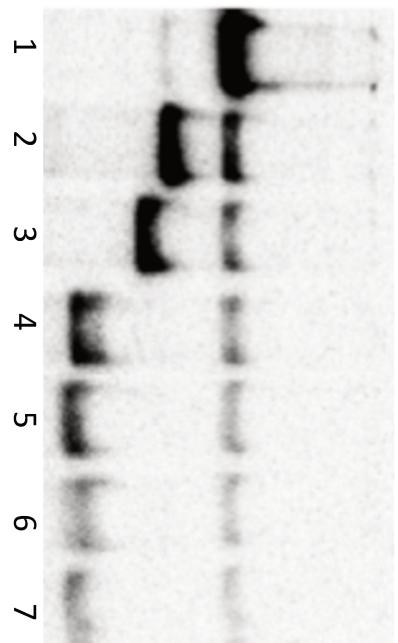
FIGURE 5

a

	0	-1	-2	-3	-4/-5	-5	-6	ΔLk
T_0	T_{-1}	T_{-2}	T_{-3}	$T_{-4/5}$	T_{-5}	T_{-6}	Topoisomer	

**b**

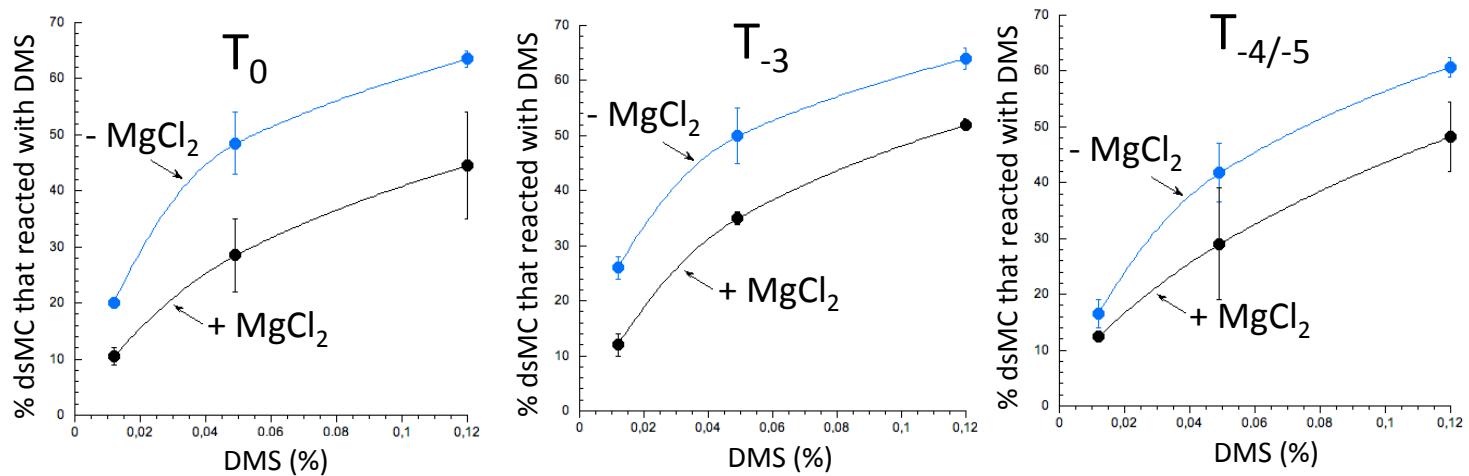
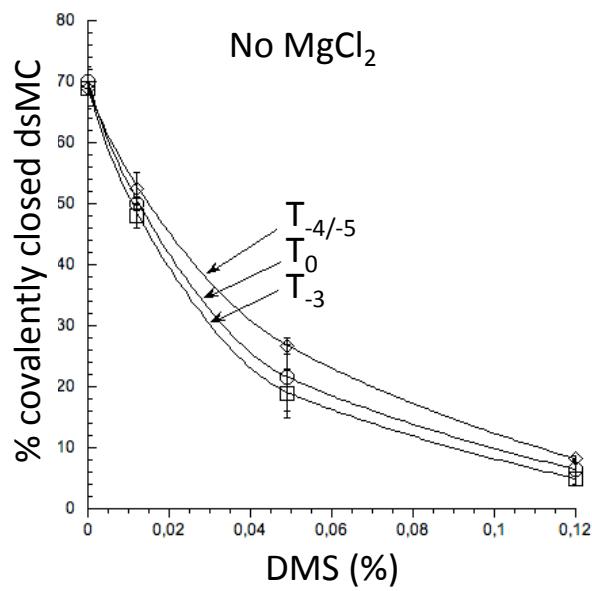
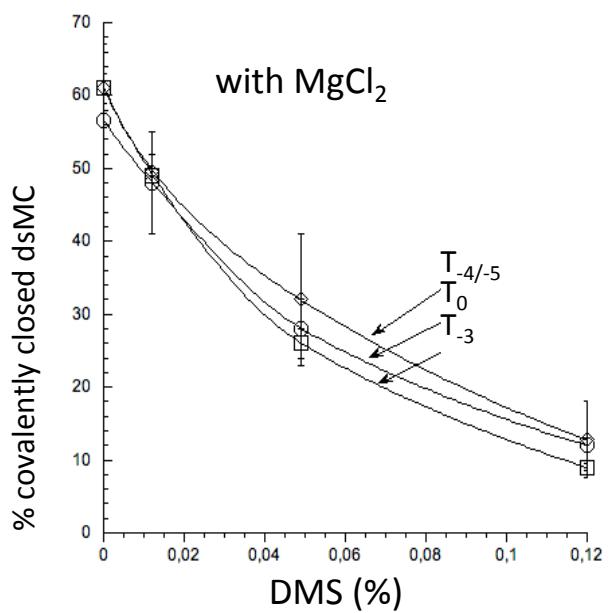
	0	-1	-2	-3	-4/-5	-5	-6	ΔLk
T_0	T_{-1}	T_{-2}	T_{-3}	$T_{-4/5}$	T_{-5}	T_{-6}	Topoisomer	

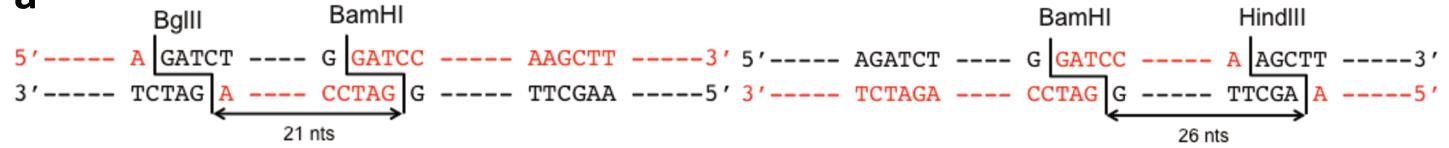
**c** CaCl_2

	0	-1	-2	-3	-4/-5	-5	-6	ΔLk
T_0	T_{-1}	T_{-2}	T_{-3}	$T_{-4/5}$	T_{-5}	T_{-6}	Topoisomer	

Supplementary Figure 1: Effect of magnesium or calcium on the conformations of the topoisomers of dsMCs

- 1a. Electrophoretic migration profile of the seven topoisomers on a native gel made in TBE 0.5x.
 - 1b. Electrophoretic migration profile of the seven topoisomers in the presence of magnesium. Magnesium (10 mM) is present in the native polyacrylamide gel and in the running buffer.
 - 1c. Electrophoretic migration profile of the seven topoisomers in the presence of calcium. Calcium (10 mM) is present in the native polyacrylamide gel and in the running buffer.
- In 1b and 1c, the topoisomers differ by their relative linking number ΔLk , ranging from 0 (T_0 topoisomer, lane 1) to -6 (T_{-6} topoisomer, lane 7) and electrophoreses are performed at 4°C.

a**b****c**

a

Discussion

Les topoisomérases sont des enzymes essentielles intervenant dans le métabolisme de l'ADN. Leur rôle est de résoudre les contraintes topologiques qui s'exercent sur l'ADN lors de différents processus cellulaires. Le caractère primordial de ces enzymes justifie leur étude sur le plan mécanistique et fonctionnel. Notre travail a proposé un nouveau modèle d'étude pour les topoisomérases utilisant des minicercles d'ADN de 235 bp. Nous avons cherché à mieux appréhender les différences de structure de ces minicercles en fonction de leur linking number (Lk).

La topologie des minicercles peut être précisément contrôlée lors de leur synthèse. Ainsi, nous avons pu construire 7 minicercles différents de 235 bp avec un nombre relatif d'enlacement variant de 0 à -6. Ces différents topoisomères peuvent être séparés et clairement identifiés par électrophorèse. Leur marquage radioactif permet une grande sensibilité de détection.

L'étude de la topologie de ces minicercles a permis de mettre en évidence qu'ils ne possèdent pas de grandes régions simple-brin, contrairement à l'idée que nous nous en faisions (plus le substrat est hypernégativement superenroulé, plus il possède de régions simple-brin). Nos résultats montrent que plus le substrat est hypernégativement surenroulé, plus il est sensible à la nucléase SI mais seulement au niveau d'une petite région de 37 bp.

L'utilisation de minicercles pour la caractérisation de l'activité de relaxation de topoisomérases présente plusieurs avantages : la détection et la quantification de chaque produit de relaxation (permettant des analyses fines de la relaxation), la visualisation et la quantification du complexe covalent pendant la relaxation et l'étude des zones d'incision par l'enzyme. Grâce à deux topoisomérases commerciales, la *EcTopoI* et la *ctTopoI* nous avons pu valider les minicercles comme étant des substrats pertinents pour l'étude de la relaxation. Pour les deux enzymes testées, nous avons pu détecter et quantifier une activité de relaxation. Nous avons démontré que leurs profils de relaxation diffèrent. Par conséquent, le modèle de minicercles d'ADN double-brin permet de caractériser l'activité de relaxation des topoisomérases en termes d'efficacité de relaxation et de nature des produits formés. Ces conclusions mettent en évidence leur différence de mécanisme de relaxation, passage de brin pour la *EcTopoI* (famille IA) et un mécanisme de rotation contrôlée pour la *ctTopoI* (famille IB).

Par l'utilisation de ces minicercles, nous avons réussi à visualiser le complexe covalent car leurs petites tailles comparées à un plasmide permettent de discriminer un minicercle libre d'un minicercle lié à une topoisomérase par électrophorèse. Ainsi, notre modèle nous a permis de visualiser pour chaque topoisomérase testée, l'étape de complexe covalent avec l'ADN, lors d'une réaction de relaxation dans des conditions physiologiques. Nos expériences de relaxation en présence de camptothécine, inhibiteur de la ctTopoI, ont bien montré d'une part une inhibition de la relaxation et d'autre part une augmentation de la proportion en complexe covalent. Ce modèle représente donc un système supplémentaire pour évaluer l'effet de petites molécules pouvant se comporter comme des poisons de topoisomérases (stabilisateurs de l'intermédiaire covalent). Ce modèle, permettant aussi bien de quantifier la relaxation et le complexe covalent que d'identifier les zones d'incision par l'enzyme, peut donc aussi participer à l'étude du rôle de protéines partenaires des topoisomérases, notamment pour observer si ces protéines modulent l'activité de l'enzyme. Par exemple, dans le cas de la hTopo3 α , la fonction de ses protéines partenaires RMI1 et RMI2 n'est pas encore établie. Un autre exemple d'utilisation de notre modèle pourrait être l'étude du rôle de la hTopo2 β dans la formation de cassure double-brin (Madabhushi et al. 2015). La réalité de ce phénomène n'est à l'heure actuelle pas clairement établie, car on ne sait si il provient d'une observation réelle ou d'un artéfact expérimental (Pommier et al. 2016). Notre modèle pourrait permettre de contrôler si ce mécanisme utilise bien une plus grande stabilité du complexe covalent.

Conclusion

Mes travaux de thèse ont permis de mettre au point un nouveau modèle d'étude des topoisomérases. Ce modèle repose sur l'utilisation de minicercles d'ADN double-brin de 235 bp. Avec nos conditions expérimentales, nous avons pu obtenir 7 topoisomères distincts. Ces différents minicercles d'indice d'enlacement relatif de 0 à -6 nous ont permis de tester la relaxation pour deux topoisomérases commerciales, la *EcTopoI* et la *ctTopoI*. Mes résultats ont ainsi pu montrer que les minicercles sont un outil simple et pertinent pour l'étude des topoisomérases. Ces minicercles présentent l'avantage de pouvoir quantifier la relaxation et le complexe covalent ADN-protéine des topoisomérases dans la même expérience, ce qui apporte un bénéfice par rapport aux études précédentes. L'utilisation des minicercles pourra être étendue à l'identification de nouvelles molécules stabilisant le complexe covalent, l'étude du rôle de protéines partenaires aux topoisomérases ou encore la caractérisation de nouvelles topoisomérases.

Bibliographie

- Ahmad, M. et al., 2016. RNA topoisomerase is prevalent in all domains of life and associates with polyribosomes in animals. *Nucleic acids research*, 44(13), pp.6335–49.
- Ashour, M.E., Atteya, R. & El-Khamisy, S.F., 2015. Topoisomerase-mediated chromosomal break repair: an emerging player in many games. *Nature reviews. Cancer*, 15(3), pp.137–51.
- Borowiec, J.A. et al., 1987. DNA supercoiling promotes formation of a bent repression loop in lac DNA. *Journal of molecular biology*, 196(1), pp.101–11.
- Champoux, J.J., 2001. DNA topoisomerases: structure, function, and mechanism. *Annual review of biochemistry*, 70, pp.369–413.
- Champoux, J.J., 1977. Strand breakage by the DNA untwisting enzyme results in covalent attachment of the enzyme to DNA. *Proceedings of the National Academy of Sciences of the United States of America*, 74(9), pp.3800–4.
- Charvin, G. et al., 2005. Tracking topoisomerase activity at the single-molecule level. *Annual review of biophysics and biomolecular structure*, 34(1), pp.201–19.
- Cheng, B. et al., 2005. Bacterial cell killing mediated by topoisomerase I DNA cleavage activity. *The Journal of biological chemistry*, 280(46), pp.38489–95.
- Chenoweth, D.M. & Dervan, P.B., 2009. Allosteric modulation of DNA by small molecules. *Proceedings of the National Academy of Sciences*, 106(32), pp.13175–13179.
- Corbett, K.D. & Berger, J.M., 2004. Structure, molecular mechanisms, and evolutionary relationships in DNA topoisomerases. *Annual review of biophysics and biomolecular structure*, 33, pp.95–118.
- Depew, R.E., Liu, L.F. & Wang, J.C., 1978. Interaction between DNA and Escherichia coli protein omega. Formation of a complex between single-stranded DNA and omega protein. *The Journal of biological chemistry*, 253(2), pp.511–8.
- Deweese, J.E., Osheroff, M.A. & Osheroff, N., 2008. DNA Topology and Topoisomerases: Teaching a "Knotty" Subject. *Biochemistry and molecular biology education : a bimonthly publication of the International Union of Biochemistry and Molecular Biology*, 37(1), pp.2–10.
- Espeli, O. & Marians, K.J., 2004. Untangling intracellular DNA topology. *Molecular microbiology*, 52(4), pp.925–31.
- Heddi, B. et al., 2010. Sequence-dependent DNA flexibility mediates DNase I cleavage. *Journal of molecular biology*, 395(1), pp.123–33.

- Kohwi-Shigematsu, T., Gelinas, R. & Weintraub, H., 1983. Detection of an altered DNA conformation at specific sites in chromatin and supercoiled DNA. *Proceedings of the National Academy of Sciences of the United States of America*, 80(14), pp.4389–93.
- Koster, D.A. et al., 2010. Cellular strategies for regulating DNA supercoiling: a single-molecule perspective. *Cell*, 142(4), pp.519–30.
- Madabhushi, R. et al., 2015. Activity-Induced DNA Breaks Govern the Expression of Neuronal Early-Response Genes. *Cell*, 161(7), pp.1592–605.
- Mankouri, H.W. & Hickson, I.D., 2007. The RecQ helicase-topoisomerase III-Rmi1 complex: a DNA structure-specific “dissolvosome”? *Trends in biochemical sciences*, 32(12), pp.538–46.
- Nitiss, J.L., 2001. Topoisomerase assays. *Current protocols in pharmacology / editorial board, S.J. Enna (editor-in-chief) ... [et al.]*, Chapter 3, p.Unit3.3.
- Parkinson, G.N., Lee, M.P.H. & Neidle, S., 2002. Crystal structure of parallel quadruplexes from human telomeric DNA. *Nature*, 417(6891), pp.876–880.
- Pommier, Y. et al., 2010. DNA Topoisomerases and Their Poisoning by Anticancer and Antibacterial Drugs. *Chemistry & Biology*, 17(5), pp.421–433.
- Pommier, Y. et al., 2016. Roles of eukaryotic topoisomerases in transcription, replication and genomic stability. *Nature Reviews Molecular Cell Biology*, 17(11), pp.703–721.
- Qian, Z. et al., 2015. A New Noncoding RNA Arranges Bacterial Chromosome Organization. *mBio*, 6(4).
- Seol, Y. et al., 2015. Single-Molecule Supercoil Relaxation Assay as a Screening Tool to Determine the Mechanism and Efficacy of Human Topoisomerase IB Inhibitors. *Molecular Cancer Therapeutics*, 14(11), pp.2552–2559.
- Seol, Y. & Neuman, K.C., 2016. The dynamic interplay between DNA topoisomerases and DNA topology. *Biophysical Reviews*, 8(S1), pp.101–111.
- Stros, M. et al., 2007. HMGB1 interacts with human topoisomerase II and stimulates its catalytic activity. *Nucleic Acids Research*, 35(15), pp.5001–5013.
- Tse-Dinh, Y.C., 1986. Uncoupling of the DNA breaking and rejoining steps of Escherichia coli type I DNA topoisomerase. Demonstration of an active covalent protein-DNA complex. *The Journal of biological chemistry*, 261(23), pp.10931–5.
- Tse, Y.C., Kirkegaard, K. & Wang, J.C., 1980. Covalent bonds between protein and DNA. Formation of phosphotyrosine linkage between certain DNA topoisomerases and DNA. *The Journal of biological chemistry*, 255(12), pp.5560–5.
- Viard, T. & de la Tour, C.B., 2007. Type IA topoisomerases: a simple puzzle? *Biochimie*, 89(4), pp.456–67.
- Wang, J.C., 2002. Cellular roles of DNA topoisomerases: a molecular perspective. *Nature reviews. Molecular cell biology*, 3(6), pp.430–40.

Watson, J.D. & Crick, F.H.C., 1953. the structure of dna. *Cold Spring Harb Symp Quant Biol*, 18, pp.123–131.

Résumé :

Mots clefs : mitochondries, délétion de l'ADN mitochondrial, réPLICATION, réPARATION, topologie de l'ADN, topoisomérases, complexe covalent

La mitochondrie est une organelle essentielle possédant son propre ADN circulaire. Cet ADN peut présenter des mutations et/ou des délétions, consécutives à l'exposition à différents types de dommages ou en raison de protéines mutées. Ces mutations ou délétions sont impliquées dans de nombreuses pathologies, dont les cancers, et le vieillissement. Leur apparition peut survenir notamment lors de la réPLICATION ou de la réPARATION. A ce jour, la réPLICATION et la réPARATION mitochondrielles ne sont pas encore bien élucidées. L'objectif de ce projet est donc de mieux en appréhender les mécanismes et de mieux comprendre l'émergence d'anomalies en nous intéressant plus particulièrement à une délétion appelée « Common Deletion ». Ce travail reposait sur l'hypothèse que cette délétion put résulter d'une mauvaise réPARATION de cassure(s) double-brin et/ou d'une erreur durant la réPLICATION de l'ADN mitochondrial. L'analyse de ces résultats révèle que la formation de la « Common Deletion » ne nécessite qu'une seule cassure double-brin proche des séquences répétées entourant cette dernière et implique les protéines de la réPLICATION de l'ADN mitochondrial. Ainsi, ce travail permet de mieux saisir les mécanismes de réPLICATION et de réPARATION assurant la stabilité de l'ADN mitochondrial. Un second projet a été de proposer un modèle d'étude *in vitro* des topoisomérases en utilisant des minicercles d'ADN permettant la visualisation du complexe covalent, étape clef de la réACTION de relaxation de ces enzymes.

Summary :

Keywords: mitochondria, deletion of mitochondrial DNA, replication, repair, DNA topology, topoisomerases, covalent complex

Mitochondria is an essential organelle with its own circular DNA. This DNA may exhibit mutations and/or deletions, as a result of exposure to different types of damage or due to mutated proteins. These mutations or deletions are involved in many pathologies, including cancers, and aging. They may occur during replication or repair. For now, mitochondrial replication and repair have not yet been fully elucidated. The objective of this project is therefore to better understand the mechanisms and the emergence of anomalies by focusing on a deletion called "Common Deletion". This work was based on the assumption that this deletion could result from poor repair of double-strand break(s) and/or error during mitochondrial DNA replication. Analysis of these results reveals that the formation of the "Common Deletion" requires only a single double-strand break close to the repeated sequences surrounding the latter and involves the proteins of mitochondrial DNA replication. Thus, this work makes it possible to better understand the mechanisms of replication and repair ensuring the stability of mitochondrial DNA. A second project was to propose an *in vitro* model for topoisomerases using DNA minicircles allowing visualization of the covalent complex, a key step in the relaxation reaction of these enzymes.