



HAL
open science

Fonction et régulation de l'ADN polymérase zêta au cours de la réplication de l'ADN : conséquences sur la stabilité du génome.

Barbara Ben Yamin Rosen

► To cite this version:

Barbara Ben Yamin Rosen. Fonction et régulation de l'ADN polymérase zêta au cours de la réplication de l'ADN : conséquences sur la stabilité du génome.. Biochimie, Biologie Moléculaire. Université Paris-Saclay, 2020. Français. NNT : 2020UPASS031 . tel-03868263

HAL Id: tel-03868263

<https://theses.hal.science/tel-03868263v1>

Submitted on 23 Nov 2022

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.

Fonction et régulation de l'ADN
polymérase ζ au cours de la réplication de
l'ADN : conséquences sur la stabilité du
génom

Thèse de doctorat de l'université Paris-Saclay

École doctorale n° 582 de Cancérologie : Biologie, Médecine, Santé
Spécialité de doctorat: Sciences de la vie et de la santé
Unité de recherche : UMR9019 CNRS « Genome Integrity and Cancers »
Réfèrent : Faculté de médecine

Thèse présentée et soutenue à Villejuif, le 5 février 2020, par

Barbara BEN YAMIN ROSEN

Composition du Jury

Pierre-Henry GAILLARD Directeur de Recherche CNRS – CRCN, Marseille	Président & Rapporteur
Angelos CONSTANTINO Directeur de Recherche INSERM – IGH, Montpellier	Rapporteur & Examineur
Jean-Charles CADORET Maître de Conférences – Paris Université – IJM, Paris	Examineur
Agnès CORDONNIER Chargée de Recherche INSERM – ESBS, Illkirch	Examinatrice
Gérald PEYROCHE Maître de Conférences – ENS Paris Saclay	Examineur
Patricia KANNOUCHE Directrice de Recherche CNRS – Gustave Roussy, Villejuif	Directrice de thèse

Résumé

La réplication de l'ADN est un processus cellulaire fondamental qui assure la duplication fidèle de l'information génétique et sa transmission aux cellules filles. Différentes perturbations d'origine endogène ou exogène peuvent interférer avec la progression des fourches de réplication menaçant ainsi l'intégrité du génome. Pour éviter le blocage du réplisome, les polymérases répliquatives peuvent être remplacées temporairement par des ADN polymérases translésionnelles (TLS) mutagènes, capables de franchir différents types de lésions. Parmi ces TLS polymérases, l'ADN polymérase zêta (Pol ζ) est un facteur-clé dans ce processus de franchissement de lésions. Elle a aussi une fonction essentielle dans le développement embryonnaire et la prolifération cellulaire. Avant le début de ma thèse, la sous unité catalytique de Pol ζ , REV3L, a été montrée jouer un rôle important dans le maintien des sites fragiles communs, suggérant que Pol ζ serait impliquée dans la réplication "normale" de l'ADN. Cependant sa fonction biologique restait encore à préciser.

Au cours de la première partie de ma thèse, je me suis intéressée à la caractérisation des fonctions biologiques de l'ADN polymérase ζ dans les cellules de mammifères. Ces travaux suggèrent que, dans des conditions physiologiques, Pol ζ participe à la réplication de certaines régions du génome difficiles à répliquer. En effet, l'absence de REV3L entraîne un retard de réplication de 5% du génome associée à une augmentation de la mutagénèse dans ces régions spécifiques. Nous avons également montré que Pol ζ joue un rôle critique dans le maintien épigénétique en interagissant avec des composants de l'hétérochromatine telle que les marques répressives d'histones. Finalement, des résultats préliminaires semblent suggérer un rôle de Pol ζ dans la réplication des télomères et le maintien de la stabilité télomérique. Nous proposons donc que Pol ζ a été sélectionnée au cours de l'évolution pour assister les polymérases répliquatives à répliquer des régions hétérochromatiennes compactes, afin de limiter l'instabilité génomique.

Dans une seconde partie, je me suis intéressée à la régulation cellulaire de la protéine REV3L au cours de la réplication non perturbée de l'ADN ou après exposition des cellules à un stress génotoxique. Nous avons observé que REV3L est finement régulée, avec des temps de demi-vie courte de la protéine comme de l'ARN messenger. De façon inattendue, nous avons mis en évidence un mécanisme de régulation unique de la protéine REV3L. En effet, nous avons démontré que REV3L est naturellement clivée par l'endopeptidase TASP1 au niveau de sa séquence consensus QLDGTAD en position 525. Ce clivage conduit à la formation d'un hétérodimère stable (REV3L p70/p300), composé des deux fragments résultants du clivage, suggérant que cet hétérodimère correspond à la forme active de REV3L. De manière intéressante, le fragment N-terminal est hyperphosphorylé en réponse à un stress génotoxique. Finalement, des résultats préliminaires laissent penser que le clivage de REV3L par TASP1 est nécessaire pour la prolifération cellulaire, la résistance aux UV ainsi que le maintien de la stabilité génomique dans les cellules de mammifères. Mes travaux de thèse ont permis de mettre en lumière un mécanisme unique pour contrôler la fonction d'une TLS polymérase dans les cellules de mammifères.

En conclusion, ces résultats apportent une meilleure compréhension de la fonction et de la régulation de Pol ζ dans les cellules de mammifères. De plus, d'un point de vue thérapeutique, ces résultats sont particulièrement importants étant donné que cette ADN polymérase mutagène est impliquée dans les mécanismes de chimio-résistance. Ainsi pouvoir réduire son activité dans les tumeurs lors des traitements serait une stratégie thérapeutique.

Remerciements

Tout d'abord, je souhaite remercier Angelos CONSTANTINO et Pierre-Henri GAILLARD d'avoir accepté d'être rapporteur de ma thèse. Je remercie également Gerald PEYROCHE et Agnès CORDONNIER d'avoir accepté d'être examinateurs. Un merci particulier à Jean-Charles CADORET qui m'a permis de rencontrer Patricia et qui suit mon travail depuis le début de ma thèse, pour avoir accepté aujourd'hui d'être examinateur ainsi que pour sa collaboration à mon papier de thèse.

Merci aux personnes qui m'ont accompagné chaque jour au cours de ces 5 dernières années, l'équipe TLS :

Merci Patricia pour m'avoir accueillie dans votre équipe et guidée tout au long de ma thèse. Merci d'avoir pris du temps malgré toutes vos responsabilités, pour vos précieux conseils ainsi que les nombreuses discussions sur ce projet plein de rebondissements qui nous a fait nous poser beaucoup de questions mais toujours si intéressant. Je vous souhaite encore beaucoup de belles découvertes sur tous les projets du laboratoire.

Merci Caro pour ta gentillesse et ton aide précieuse toutes ces années, tu es et resteras la reine du Western blot ;) ! Merci Emma pour tous tes précieux conseils et idées. Merci Laetitia de m'avoir supporté dans le même bureau que toi au cours de ces 2 dernières années. Je te souhaite beaucoup de réussite dans ton projet ainsi que pour la suite ! Merci à tous les stagiaires ayant travaillé sur le projet REV3L, Amélie, Léa et plus particulièrement Dorian pour leur aide technique. Last but not least, merci Alice, mon mari de la journée, simplement pour tout... Je n'aurai jamais réussi cette thèse sans toi, tant coté manip que coté mental, et même si je sais que tu n'aimes pas ça, aujourd'hui tu n'es plus qu'une simple collègue, tu es mon amie.

Un grand merci à toute l'unité UMR8200 pour votre bonne humeur ainsi que pour les discussions scientifiques ou non au détour d'un couloir ou au cours des déjeuners à la cafeteria. Ce fut un vrai plaisir de venir travailler chaque jour à vos côtés.

Un merci spécial à mes copines Houaria et Caroline Z pour ces longues discussions dans mon bureau, je vous kiffe les filles !

Ensuite, je tiens à remercier les personnes qui me soutiennent au quotidien :

Merci à mes parents d'avoir cru en moi peu importe les obstacles, pour m'avoir toujours donner les moyens de réussir dans mes projets et sans qui je n'en serai pas la aujourd'hui.

Merci à mes sœurs, Rebecca, Léa et Johanna, je vous l'ai déjà dit un million de fois mais ma vie serait si triste sans vous. Vous êtes à la fois mes sœurs et mes meilleures amies et j'ai tellement de chance de vous avoir à mes côtés !

Merci à ma nièce, ma princesse Sienna, pour avoir été mon havre de paix au cours de l'écriture de cette thèse.

Merci à mon mari, Jonathan, qui a attrapé ma thèse en vol, pour ton soutien à toutes épreuves et pour toujours faire semblant de t'intéresser à mes résultats même si tu ne comprends pas grand-chose... Merci d'être tout simplement présent à mes côtés chaque jour !

Merci à tous de m'avoir épaulé et encourager toutes ces années, je vous aime tous à la folie et cette thèse c'est aussi un peu la vôtre !

Merci à Joy, Matt et Jihad, les Originals, sans qui mes études auraient été beaucoup moins drôles. Même si on se voit moins et que vous êtes loin, je pense fort à vous maintenant que c'est finalement à moi de terminer ma thèse!

Je tiens finalement à remercier le ministère de l'enseignement supérieur, la faculté de médecine de Kremlin Bicêtre, la Fondation pour la recherche médicale et le CNRS pour avoir financé mes 5 années de thèse.

Merci encore mille fois à tous et maintenant en route pour de nouvelles aventure !



Table des matières

I. Introduction.....	12
A. Organisation du génome nucléaire	13
1. Le matériel génétique	13
2. L'information épigénétique	14
a) Méthylation de l'ADN	14
b) Les modifications d'histones.....	15
(1) Acétylation	16
(a) Effet et localisation des histones acétylées.....	16
(b) Régulation de l'acétylation des histones	16
(2) Méthylation.....	18
(a) Localisation des différentes marques de méthylation.....	18
(b) Régulation de la méthylation des histones.....	20
(3) Phosphorylation.....	21
(4) Ubiquitination.....	22
(5) Le code histone	22
(6) Les ARNs non-codants.....	23
3. Les territoires structuraux de la chromatine	24
a) L'euchromatine	24
b) L'hétérochromatine.....	25
(1) L'hétérochromatine constitutive	26
(a) Structure.....	26
(b) Mise en place et maintien de l'hétérochromatine constitutive.....	27
(2) L'hétérochromatine facultative.....	27
B. La réplication de l'ADN chez les eucaryotes supérieurs.....	28
1. Le réplisome.....	28
2. Les ADN polymérases	29
3. La réplication de l'ADN	31
a) La réplication de l'ADN en conditions physiologiques.....	31
(1) Les ADN polymérases répliquatives.....	31
(a) Pol α	31
(b) Pol δ	32
(c) Pol ϵ	32
(2) Déroulé de la réplication	32
(a) Initiation	32
(b) Élongation.....	34

(c) Terminaison.....	34
b) Cas particulier des régions centromériques et péricentromériques.....	35
(1) Structure	35
(2) Réplication	36
c) Cas particulier des télomères	36
(1) Structure.....	36
(2) Réplication	37
d) Drogues affectant la réplication de l'ADN.....	39
(1) Aphidicoline.....	39
(2) Hydroxyurée.....	39
(3) Cisplatine	40
(4) Camptothécine.....	40
e) Rôle de la réplication dans le processus d'oncogenèse	40
4. Régulation spatio-temporelle de la réplication.....	41
a) Régulation spatiale.....	41
b) Régulation temporelle.....	42
(1) Méthodes d'analyses.....	42
(a) Incorporation d'analogues de thymidine	42
(b) Timing de réplication à l'échelle du génome.....	43
(2) Corrélacion entre le programme temporel de la réplication, la séquence d'ADN et la transcription.....	44
(3) Corrélacion entre le paysage épigénétique et le programme temporel de la réplication..	45
(4) Influence du programme temporel de la réplication sur le paysage mutationnel.....	45
(5) Les régions de transition temporelle	45
C. Problèmes au cours de la réplication : causes et conséquences sur l'intégrité du génome	47
1. Les différents obstacles de la fourche de réplication.....	47
a) Les lésions endogènes.....	48
b) Les lésions exogènes	48
(1) Les agents physiques.....	48
(a) Les rayons ultraviolets.....	48
(b) Les rayonnements ionisants.....	49
(2) Les agents chimiques.....	49
c) Les barrières naturelles de la fourche de réplication et régions difficiles à répliquer.....	50
(1) Les G-quadruplexes.....	50
(2) Les R-loops.....	51
(3) Les sites fragiles communs.....	51
(4) Les conflits transcription-réplication	52
(5) Modification de la composition du réplisome au cours de la phase S.....	52

2.	Mécanismes de signalisation et prise en charge des lésions de l'ADN.....	53
a)	Signalisation des lésions au cours de la phase S	53
(1)	Mécanismes généraux des points de contrôle du cycle cellulaire.....	53
(2)	Le point de contrôle S/G2	53
b)	La synthèse translésionnelle	54
(1)	Les ADN polymérases translésionnelles.....	55
(a)	Pol η	56
(b)	Pol ι	56
(c)	Pol κ	57
(d)	Rev1	57
(e)	Primpol.....	57
(2)	Mécanisme de synthèse translésionnelle	58
c)	Prise en charge et réparation des lésions.....	58
3.	Conséquences sur l'intégrité du génome	59
a)	Instabilité génétique.....	59
b)	Instabilité chromosomique.....	59
D.	L'ADN polymérase zêta	60
1.	Le complexe ADN polymérase ζ 4.....	60
2.	Le gène et la protéine Rev3L	62
a)	<i>Saccharomyces cerevisiae</i>	62
b)	<i>Mus musculus</i>	62
c)	<i>Homo sapiens</i>	62
3.	Structure de la protéine Rev3L de mammifère.....	62
4.	Conservation au cours de l'évolution du gène Rev3L	64
5.	Rôles connus de Pol ζ	65
a)	Synthèse Translésionnelle.....	65
b)	Rôle dans la mutagenèse et signature mutationnelle.....	65
(1)	Mutagenèse	66
(2)	Signature mutationnelle.....	66
c)	Rôle dans la synthèse d'ADN après exposition à un stress réplicatif.....	67
d)	Rôle dans le maintien de la stabilité des sites fragiles communs	68
e)	Rôle dans la réparation de l'ADN	68
f)	Rôle dans la mitochondrie.....	69
6.	Rev3L in vivo.....	70
7.	REV3L et cancer.....	71
E.	Objectifs de la thèse.....	74
II.	Résultats	76

Partie I - Conséquences de l'absence de Rev3L sur la stabilité génétique et épigénétique.....	77
1. DNA polymerase zeta contributes to heterochromatin replication to prevent genome instability (Article 1, <i>submitted</i>).....	78
2. Addendum: Rev3L might contribute to telomere stability.....	135
Partie II – Caractérisation fonctionnelle et régulatrice de l'ADN Polymérase zêta.	140
1. Introduction : L'endopeptidase Threonine ASpartase 1 (TASP1).....	141
a) Le gène et la protéine TASP1.....	141
b) Structure de la protéine TASP1.....	141
c) Fonctions et cibles principales de TASP1.....	141
(1) MLL1 et MLL2.....	142
(2) TFIIA.....	142
d) Tasp1 in vivo.....	143
(1) Chez la souris.....	143
(2) Chez l'homme.....	145
e) TASP1 et cancer.....	146
2. REV3L is proteolytically processed by Taspase1 to form an active heterodimer required for polζ activity. (Article 2, <i>in preparation</i>).....	148
III. Discussion / Perspectives.....	198
IV. Conclusion.....	204
V. Références.....	207

Liste des abréviations et anglicismes

53BP1 : p53 Binding Protein 1
6-4PP : adduits 6-4 pyrimidines-pirimidiones
aa : Acide aminé
ADN : Acide Désoxyribonucléique
APH : Aphidicoline
ARN : Acide ribonucléique
ARNm : ARN messenger
ARNnc : ARN non codant
ARNsi : ARN silencing
ATM : Ataxia Telangectasia Mutated
ATR : Ataxia Telangectasia and Rad-5-related
ATRX : Alpha Thalassemia/Mental Retardation Syndrome X-Linked
BER : Réparation par excision de base
BrdU : Bromodésoxyuridine
c-Het : Hétérochromatine constitutive
CHX : Cycloheximide
CENPA : Centromere Protein A
CFS : Site fragile commun
CisPT : Cisplatine
CMG : Cdc45-MCM-GINS
CNV : Variant du nombre de copies
CPD : Cyclobutane dimères de pyrimidine
CPT : Camptothécine
CTR : Région de timing de réplication constant
CS : Site de clivage (Cleaving Site)
DAPI : 4',6-diamidino-2-phenylindole
DNMT1 : DNA methyl transferase 1
dNTP : Di-nucléotide triphosphate
DRIM : mutagénèse induite par le réplisome (Defective Replisome Induced Mutagenesis)
DSB: Cassure double brin (Double Strand Break)
EdU : 5-Ethynyl-2'-deoxyuridine
FA : Formaldéhyde
FACS : Fluorescence-Activated Cell Sorting
G4 : G quadruplex
HAT : Histone acétyltransférase
HDAC : Histone déacétylase
HDM : Histone déméthylase
HMT : Histone méthyltransférase
Hox : Gènes homéotiques (Homeobox)
HP1 : Hétérochromatin Protein 1
HU : Hydroxyurée
IdU : Iododeoxyuridine
IP : Iodure de Propidium ou Immunoprécipitation
kDa : Kilo Daltons

KI : Knock In
KO : Knock-out
LINE : Long INterspersed Elements
Mb : Mégabases
MBD : Metyl-CpG binding protein
MCM : Minichromosome Maintenance
MEF : Fibroblaste embryonnaire de souris
MLL1 : mixed-lineage leukemia 1
MLS : Mitochondrial Localization Signal
MMR : Réparation des mismatches de l'ADN
NER : Réparation par excision de nucléotide
NGS : Next Generation Sequencing
NHEJ : Jonction d'extrémités non-homologues
NLS : Nuclear Localization Signal
ORC : Origin Recognition Complex
Pb : Paire de base
PCNA : Proliferating Cell Nuclear Antigen
Pol : Polymérase
PRIM : Primase
RH : recombinaison Homologue
RI : Rayons ionisants
RNase A/H : Ribonucléase A/H
ROS : Espèces réactives de l'Oxygène
RPA : Réplication Protein A
RT : Timing de réplication ou Reverse Transcription
RTT : Régions de transitions temporelles
RTEL1 : regulator of telomere length helicase 1
SINE : Short INterspersed Elements
SNP : Polymorphisme de nucléotide simple
TASP1 : Threonine aspartase 1
TLS : Synthèse translésionelle
TFIIA : Transcription Factor II A
TRF1-2 : Telomeric Repeat binding factor 1-2
TSS : Transcription Start Site
TTR : Régions de transition temporelle
UV : Rayons ultraviolets
WGA : Whole Genome Amplification

Liste des figures

Figure 1: Répartition des types de séquences présents dans le génome humain.

Figure 2: Ensemble des modifications des queues N-terminales des Histones.

Figure 3: Acétylation des histones, condensation de la chromatine et expression génique.

Figure 4: Distribution des marques de méthylation des histones, des HMTs et des HDMs tout au long du chromosome.

Figure 5: Lien entre la méthylation des histones et la méthylation de l'ADN.

Figure 6: Les différentes modifications épigénétiques sont impliquées dans de nombreux mécanismes cellulaires et sont nécessaires au bon fonctionnement de la cellule.

Figure 7: Observation de chromosome métaphasique après coloration au Giemsa.

Figure 8: L'hétérochromatine se localise dans des territoires distincts du noyau.

Figure 9: Tableau récapitulatif des différentes familles de polymérases eucaryotes et leurs fonctions.

Figure 10: Structure par cristallographie des sous-unités catalytiques des polymérases répliquatives de la famille B.

Figure 11: Initiation de la réplication de l'ADN.

Figure 12: Représentation du réplisome en cours d'élongation au niveau d'une fourche de réplication.

Figure 13: Structure des télomères de mammifères sous forme de t-loop associés aux protéines du complexe shelterin.

Figure 14: Réplication des telomeres.

Figure 15: Modèle de développement tumoral dû au stress répliquatif.

Figure 16: Pattern de réplication de l'ADN par incorporation d'analogue de thymidine dans des cellules humaines.

Figure 17: Analyse du timing de réplication.

Figure 18: Corrélations entre timing de réplication, paysage génétique et épigénétique.

Figure 19: Structure des principales lésions de l'ADN

Figure 20: Structure d'un G-quadruplex

Figure 21: Représentation schématique de la structure d'une R-loop

Figure 22: Vue simplifiée de la réponse aux dommages de l'ADN

Figure 23: Structure par cristallographie des domaines catalytiques des polymérases translésionnelles

Figure 24: Modèle du changement de polymérase impliquant deux polymérases translésionnelles pour le franchissement d'une lésion

Figure 25: Structure de la Polymérase ζ 4

Figure 26: Rev3L au cours de l'évolution.

Figure 27: Modèle du rôle de Pol ζ dans la réplication induite de cassures médiées par micro-homologie (MMBIR).

Figure 28: L'absence de Rev3L entraîne une létalité embryonnaire précoce chez la souris

Figure 29: Pol ζ agit comme suppresseur de tumeur en préservant la stabilité chromosomique au prix d'une augmentation de la mutagenèse chez les mammifères

Figure 30: TASP1 régule divers mécanismes cellulaires à travers son clivage des facteurs de transcription MLL1 et TFIIA

Figure 31: Phénotypes associés aux souris *Tasp1*^{-/-}

Figure 32: Conclusion

I. Introduction

A. Organisation du génome nucléaire

1. *Le matériel génétique*

La cellule représente l'unité du vivant et elle est le véhicule de l'information héréditaire qui définit une espèce. La vie d'une cellule est régie par le cycle cellulaire correspondant à l'ensemble des étapes permettant sa multiplication. Il est constitué de 4 étapes : (1) Une étape de réplication de l'information génétique (phase de synthèse ou S) de la manière la plus fidèle possible afin de générer deux copies exactes. (2) Une étape de division cellulaire (phase de mitose ou M) donnant lieu à deux cellules filles identiques. (3-4) Deux étapes d'interphase ou « gaps » (phases G1 et G2).

Toute cellule vivante conserve son information génétique sous la forme d'une molécule d'ADN bicaténaire. Dans le noyau, l'ADN est enroulé autour d'un octamère de protéines appelées histones composé de deux hétérodimères H3-H4, H2a et H2b. Cette association d'ADN et de protéines est appelée la chromatine et son unité fonctionnelle est le nucléosome. L'histone H1 vient ensuite stabiliser ce complexe.

Chez les métazoaires, l'ADN est composé de plusieurs types de séquences. Environ 2% du génome est constitué de séquences codantes, ou géniques, et conduisent à la formation de protéines. Le reste du génome est non-codant et contient des séquences régulatrices de l'expression génique, des séquences répétées ainsi que des éléments transposables tels que les LINE ayant joué un rôle dans l'évolution des génomes (**Figure 1**). Cet ADN intergénique serait impliqué dans le maintien de la structure de la chromatine.

La séquence nucléotidique et les gènes forment le premier niveau d'information génétique. Or un organisme tel que le corps humain est composé de plusieurs tissus avec des fonctions différentes portant tous cette même information génétique. Il existe donc un niveau supérieur de régulation de l'expression des gènes, l'épigénétique.

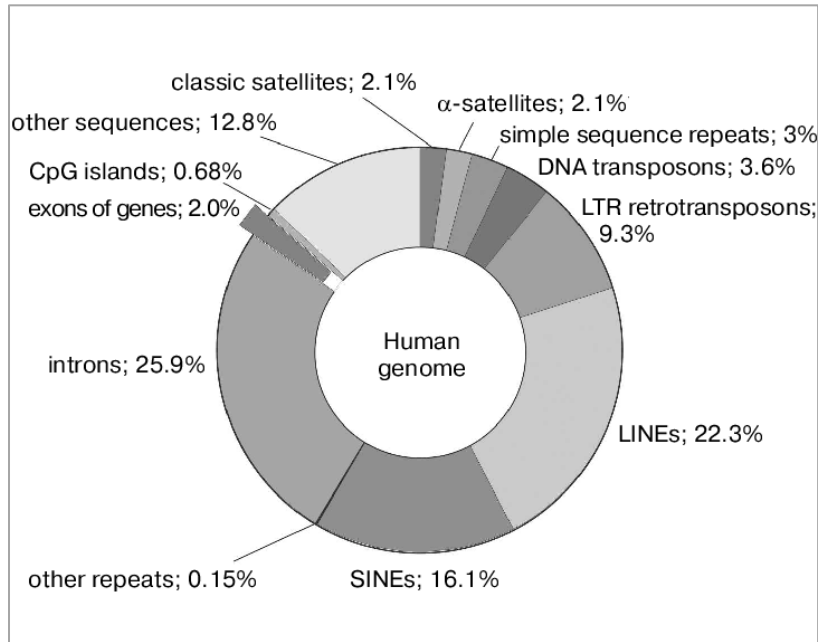


Figure 1: Répartition des types de séquences présents dans le génome humain. (Patrushev and Minkevich, 2008). Les régions codantes de l'ADN ne représentent que 2% du génome tandis que plus de la moitié du génome est composée de séquences répétées.

2. L'information épigénétique

L'épigénétique désigne des modifications transmissibles et réversibles de l'expression des gènes ne s'accompagnant pas de changements des séquences nucléotidiques. Ces modifications peuvent être faites directement sur l'ADN ou sur les protéines associées à l'ADN telles que les histones (Deans and Maggert, 2015).

a) Méthylation de l'ADN

Les modifications chimiques de l'ADN ont été découvertes chez les mammifères en 1948 par Hotchkiss (Hotchkiss, 1948). Ce n'est qu'au milieu des années 70 que la méthylation de l'ADN a été démontrée comme étant impliquée dans la régulation de l'expression des gènes et plus particulièrement dans la différenciation cellulaire (Holliday and Pugh, 1975).

La méthylation de l'ADN est un mécanisme épigénétique de transfert d'un groupement méthyle (-CH₃) sur le carbone 5 d'une cytosine donnant une 5-méthylcytosine (5mC), considérée comme la « 5^{ème} base » de l'ADN. Ce mécanisme est catalysé par une famille de protéines appelées DNA-methyl-transferases (DNMT) et peut être inhibé par la drogue 2'-deoxy-5-azacytidine. Cette famille de protéines peut être divisée en deux catégories : DNMT1 est responsable du maintien

du profil de méthylation préexistant au cours de la réplication de l'ADN tandis que DNMT3a et b sont capables de méthyler l'ADN non modifié *de novo* au cours de l'embryogenèse et de la différenciation cellulaire. Ces protéines sont très fortement impliquées dans le développement embryonnaire et leur expression est fortement réduite dès la différenciation cellulaire. En effet, des changements dynamiques du profil de méthylation ont été observés dans l'embryon au cours du développement. Par exemple, les zygotes de souris doivent subir une déméthylation du génome paternel dans les stades précoces du développement embryonnaire qui est suivi d'une méthylation *de novo* de l'ADN après le stade blastula. Cette déméthylation est nécessaire à la différenciation des cellules germinales primaires et donc au bon développement embryonnaire. Ces 5mC sont retrouvées majoritairement au niveau d'îlots CpG qui sont des régions génomiques mesurant jusqu'à 1 000 pb et présentant une densité plus élevée en C et G que le reste du génome. Environ 70% des promoteurs comportent des îlots CpG qui régulent l'expression des gènes par le maintien de la structure chromatinienne ainsi que par l'accessibilité des facteurs de transcription. La méthylation de ces îlots a été associée à une répression génique due à la non-accessibilité des facteurs de transcription sur le promoteur entraînant une répression transcriptomique.

Aussi, il a été démontré une hypométhylation globale du génome dans les tumeurs avec cependant une hyperméthylation locale des promoteurs de gènes suppresseurs de tumeur (Révisé dans (Moore et al., 2013; Zhou, 2012).

b) Les modifications d'histones

Chacune des histones peut être modifiée de manière post-traductionnelle sur sa queue N-terminale par divers acteurs. L'ensemble de ces modifications est représenté dans la **Figure 2**. Ces modifications modulent l'état de compaction de la chromatine et créent différents territoires chromatinien avec des caractéristiques spécifiques : l'euchromatine pour la chromatine ouverte et l'hétérochromatine pour la chromatine fermée (détaillé au chapitre suivant). Ces modifications sont des ajouts de groupements acétyles (-COCH₃) et méthyles (-CH₃) sur les lysines (K) ou arginines (R) ou phosphoryles (PO₃²⁻) sur les sérines (S) ou thréonines (T).

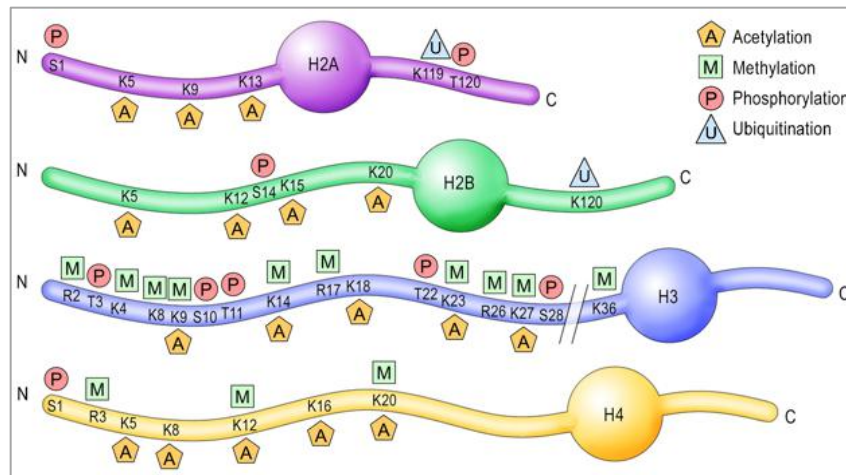


Figure 2: Ensemble des modifications des queues N-terminales des Histones. (Gräff and Mansuy, Behav Brain Res, 2008)

(1) Acétylation

(a) Effet et localisation des histones acétylées

L'ajout de groupements acétyles sur la queue N-terminale des histones diminue leur charge globale positive, diminuant ainsi leur affinité pour l'ADN chargé négativement. Cette répulsion de charge crée une structure chromatinienne ouverte et ainsi augmente l'accessibilité des facteurs de transcription à l'ADN. L'acétylation des histones est donc généralement associée à une activation de la transcription et à l'expression génique (Barnes et al., 2019; Chen et al., 2015; Lee and Grant, 2019). Des études d'immuno-précipitation de la chromatine ont permis de localiser les marques d'histones acétylées à proximité des sites de début de transcription de gènes transcriptionnellement actifs. Cependant, il existe différents sites d'acétylation qui varient selon leur localisation dans le génome. En effet, les marques H3K9ac et H3K27ac sont retrouvées au niveau des promoteurs de gènes tandis que les marques H4K12ac et H4K16ac sont retrouvées à la fois au niveau des promoteurs et dans le corps des gènes transcriptionnellement actifs. La marque H3K27ac est retrouvée au niveau de régions régulatrices distales, par exemple des enhancers (Heintzman et al., 2007; Schübeler et al., 2004; Wang et al., 2008).

(b) Régulation de l'acétylation des histones

Les groupements acétyles viennent du cofacteur acétyl-CoA et sont transférés sur les lysines par une famille de protéines appelées Histone Acetyl-Transferases (HAT) (Brownell et al., 1996; Kleff et al., 1995). Les HATs forment une superfamille d'enzymes composée d'au moins cinq

sous-familles basées sur la divergence de séquence de leurs domaines catalytiques. Malgré leurs divergences de séquence, les sous-familles de HAT partagent une structure conservée de leur région centrale permettant la liaison à l'acétyl-CoA. Cependant chacune possède une stratégie catalytique différente pour le transfert du groupement acétyle. La plupart de ces enzymes sont associées à la régulation de la transcription et sont regroupées au sein de 3 complexes multiprotéiques majeurs: le complexe SAGA, le complexe NURD et le complexe Sin3 (Révisé dans (Marmorstein and Zhou, 2014).

Les lysines acétylées sont reconnues ou « lues » par le module « Bromodomaine » présent dans divers types de protéines. Chez les mammifères, 61 bromodomaines ont été décrits dans 46 protéines comprenant des enzymes de modification des histones telles que des HATs ou des méthyltransférases, des hélicases, des facteurs de transcription et des protéines appartenant à des complexes de remodelage de la chromatine renforçant l'association entre l'acétylation des lysines et l'état chromatinien. Huit familles de bromodomaines ont été décrites avec peu de conservation de séquence mais avec la conservation d'un résidu asparagine critique pour la liaison à la lysine acétylée (Dhalluin et al., 1999; Sanchez and Zhou, 2009; Schultz et al., 2000).

Les groupements acétylés sont retirés des lysines par une famille de protéines appelées Histone Désacetylases (HDAC). Les HDACs appartiennent généralement à de gros complexes multiprotéiques. Chez les mammifères, 18 HDACs ont été décrites et ont été classées en 4 classes distinctes selon leur ressemblance avec leur homologue de levure (Bolden et al., 2006; Verdin and Ott, 2015; Verdin et al., 2003). Les classes I, II et IV regroupent les protéines HDAC 1 à 11 et présentent des caractéristiques communes telles que leur dépendance au zinc²⁺ pour leur activité catalytique. Elles sont inhibées par la TrichoStatin A (TSA) et le SAHA. Les HDACs de classe III regroupent les protéines SIRT 1 à 7 et sont dépendantes du NAD⁺ pour leur activité catalytique. Contrairement aux classes I, II et IV, les HDACs de classe III ne sont pas inhibées par la TSA (Scroggins et al., 2007). Les HDACs sont retrouvées surexprimées dans un grand nombre de cancers (tels que le cancer du sein, de l'ovaire et de la vessie) et vont promouvoir la carcinogénèse (Hayashi et al., 2010; Müller et al., 2013; Poyet et al., 2014). Des inhibiteurs de HDACs ont été développés et sont utilisés en chimiothérapie comme le Vorinostat dans le traitement du lymphome (Duvic and Vu, 2007; Tan et al., 2010).

Les HATs, HDACs et Bromodomaines sont aussi impliqués dans l'acétylation/désacétylation et la reconnaissance de protéines acétylées autres que les histones telles que la protéine suppresseur de tumeur P53 ou le facteur de transcription NF- κ B et peuvent parfois être appelées Lysines Acetyl-Transferase/Deacetylases (KAT/KDAC).

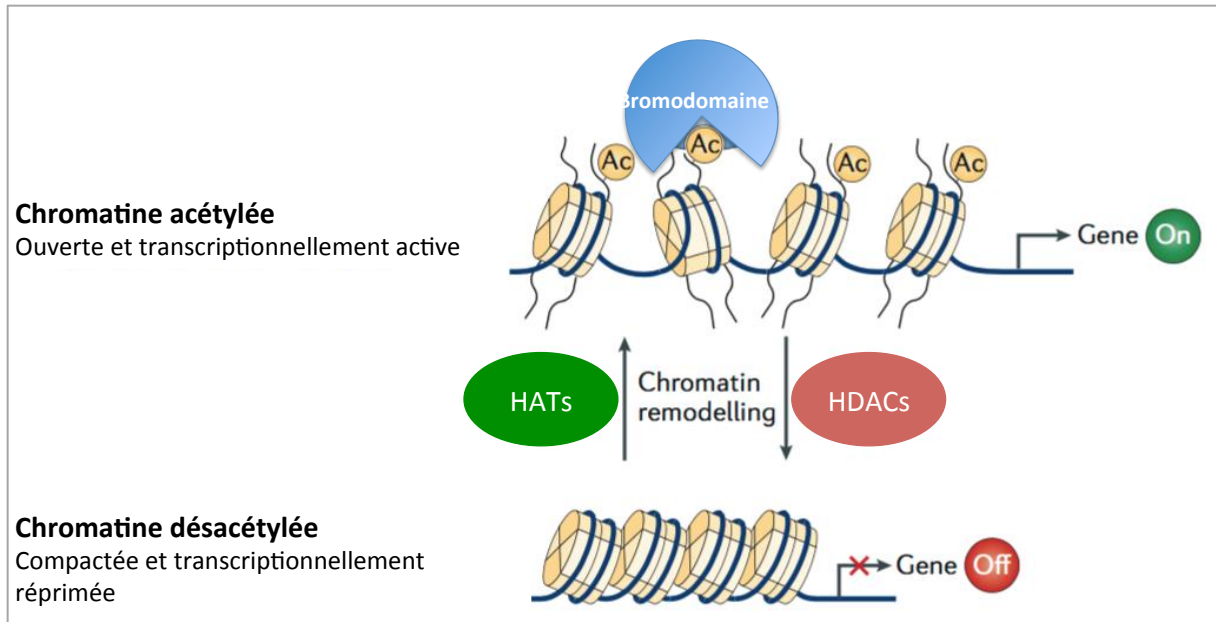


Figure 3: Acétylation des histones, condensation de la chromatine et expression génique. (Adapté de(Verdin and Ott, 2015)) L'acétylation des histones par les HATs entraîne une décondensation de la chromatine et active la transcription des gènes en aval tandis que la désacétylation par les HDACs entraîne une condensation de la chromatine et une répression génique. Les protéines portant un bromodomaine peuvent reconnaître les lysines acétylées et permettre le recrutement de facteurs de transcriptions et de complexes de remodelage de la chromatine.

(2) Méthylation

(a) Localisation des différentes marques de méthylation

Contrairement à l'acétylation, la méthylation ne modifie pas la charge globale des histones. La méthylation des histones peut donc être associée à de l'activation et à de la répression transcriptomique. Son effet va dépendre de l'acide aminé modifié ainsi que du nombre de groupements méthyles déjà présents sur ce même résidu. En effet, les lysines peuvent être mono-, di- ou tri-méthylées. Ainsi la di- ou tri-méthylation de la lysine 4 de l'histone H3 (H3K4me3/2) ainsi que la marque H3K36me3 sont associées à une activation de la transcription tandis que les marques H3K9me2/3, H3K27me2/3 et H4K20me3 sont associées à une répression transcriptomique. Cependant certaines marques sont spécifiques de régions génomiques précises et la méthylation est un déterminant clé de l'état chromatinien (Black et al., 2012; Mikkelsen et al., 2007; Schübeler et al., 2004).

En effet, H3K4me3 est très enrichie au niveau des promoteurs de gènes, actifs ou non (Shilatifard, 2012). H3K4me3 ne marque pas seulement les gènes actifs mais aussi les gènes susceptibles de s'activer. Dans ce cas, les promoteurs sont co-marqués par H3K27me3 et sont

bits bivalents (Bernstein et al., 2006; Vastenhouw and Schier, 2012). La présence de H3K4me3 au niveau des promoteurs permet le recrutement de facteurs de transcription facilitant le recrutement de l'ARN Polymérase II (Ng et al., 2003). Les séquences directement à proximité du site d'initiation de la transcription sont enrichies en H3K4me1/2 (Heintzman et al., 2007; Kim and Buratowski, 2009). Ces promoteurs actifs sont associés à la présence de marques H4K20me1 et H3K36me3 dans le corps des gènes en aval avec une augmentation d'H3K36me3 en 3' des gènes actifs tandis que la méthylation de H3K79 diminue (Barski et al., 2007; Black et al., 2012; Mikkelsen et al., 2007; Schübeler et al., 2004). La présence de la marque H3K27me3 dans le corps des gènes est associée à une répression de la transcription due à une inhibition de l'élongation de la transcription, ce qui suggère un rôle antagoniste des marques H3K36me3 et H3K27me3 (Chen et al., 2012b; Young et al., 2011). Cependant, il a été montré chez la drosophile et *C. elegans* que les gènes actifs localisés dans des régions hétérochromatiniennes présentent une augmentation des marques H3K9me2/3 dans le corps de ces gènes alors que cette marque est absente des gènes euchromatiniens (Ahringer and Gasser, 2018; modENCODE Consortium et al., 2010).

Aussi, les marques H3K9me3 et H4K20me3 sont enrichies au niveau des régions répétées telles que les séquences satellites de la chromatine péricentromérique (Barski et al., 2007; Ernst et al., 2011; Filion et al., 2010; Gerstein et al., 2010; Kharchenko et al., 2011; Liu et al., 2011; Mikkelsen et al., 2007; modENCODE Consortium et al., 2010). Même si le rôle de ces régions d'ADN n'est pas encore compris, ces marques y ont été conservées au cours de l'évolution et participent au maintien de la stabilité génomique.

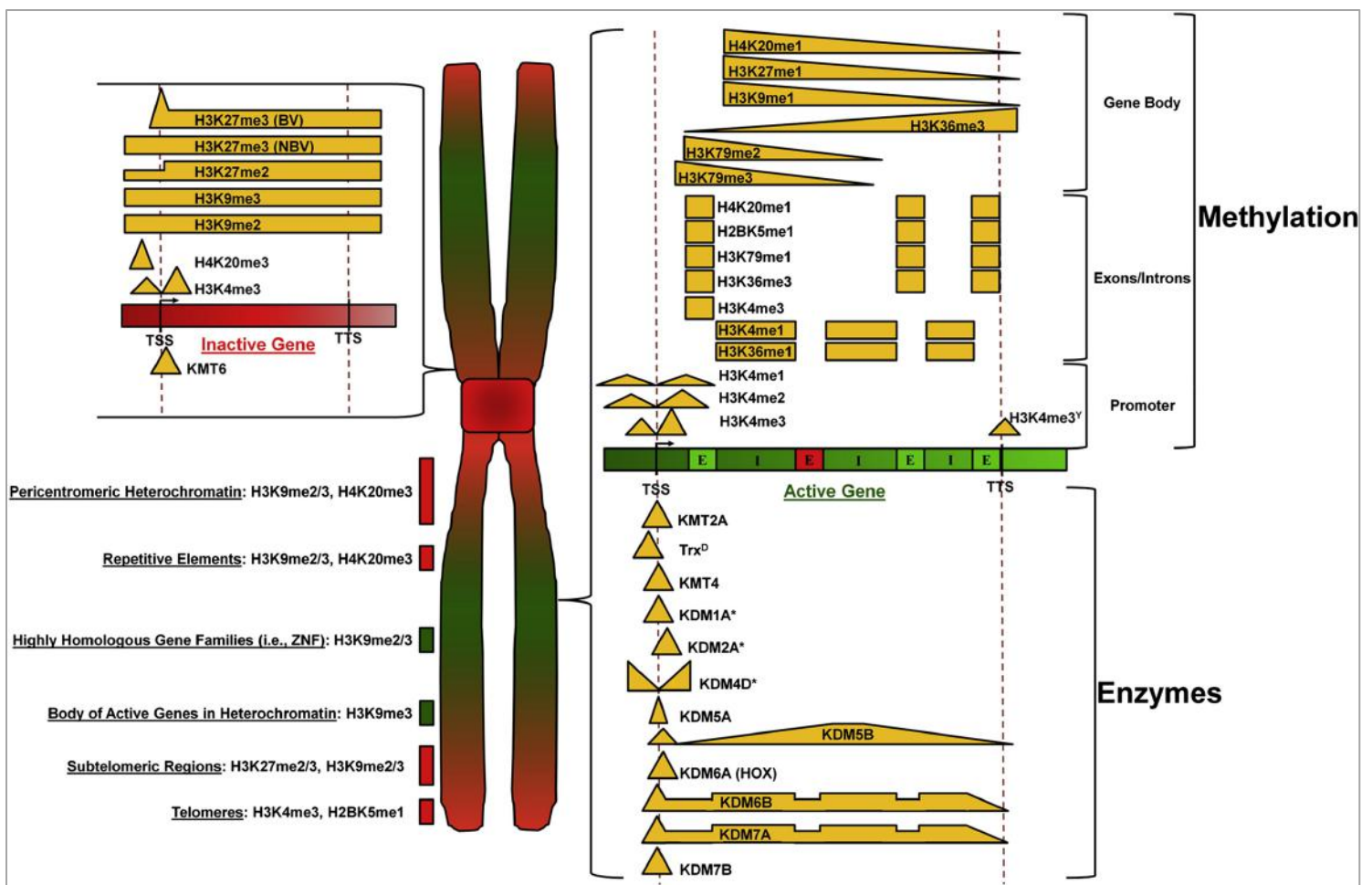


Figure 4: Distribution des marques de méthylation des histones, des HMTs et des HDMs tout au long du chromosome (Black et al., 2012). Les régions d'euchromatine sont représentées en vert et les régions d'hétérochromatine en rouge.

(b) Régulation de la méthylation des histones

La méthylation des histones consiste en l'ajout d'un ou plusieurs groupements méthyles à partir du cofacteur S-Adenosyl-Methionine (SAM) sur des résidus lysines ou arginines. Cette réaction est catalysée par une famille de protéines appelées Histones MethylTransferases (HMTs) divisée en deux classes. La première classe regroupe des enzymes possédant un domaine SET, domaine catalytique spécifique de 130 aa, conservé de la levure à l'Homme. La deuxième classe de HMTs ne possède pas de domaine SET (Dillon et al., 2005; Jenuwein, 2006; Nguyen and Zhang, 2011; Okada et al., 2005).

Chaque HAT possède un haut degré de spécificité pour la lysine à modifier ainsi que pour le degré de méthylation. Par exemple, MLL1 di- ou tri-méthyle la lysine 4 de l'histone H3 (H3K4me2/3) à partir d'une lysine non méthylée (Patel et al., 2008, 2009; Shilatifard, 2012; Shinsky et al., 2015). SUV39H1 et 2 triméthylent la lysine 9 de l'histone H3 (H3K9me3) à partir d'une mono-méthylation (H3K9me1) et ce mécanisme est nécessaire pour la mise en place de l'hétérochromatine ainsi que le maintien de la stabilité des télomères chez la souris (Benetti et

al., 2007; García-Cao et al., 2004; Rea et al., 2000). Il en est de même avec l'établissement de la marque H4K20me3 par Suv420H1/2. En effet, chez la souris, l'inhibition de Suv39H1/2 et Suv420H1/2 entraîne une augmentation de la taille des télomères ainsi qu'une plus grande fréquence de recombinaison télomérique (Benetti et al., 2007; García-Cao et al., 2004).

La méthylation des lysines est retirée par des Histones Déméthylases (HDMs). Les HDMs sont divisées en deux classes. La première classe regroupe des enzymes contenant un domaine amine oxydase dépendant d'une flavine adénine dinucléotide (FAD). La deuxième classe regroupe des enzymes possédant un domaine JmjC capable de catalyser la déméthylation par l'oxydation du groupement méthyle et cette réaction dépend de la présence d' α -cétoglutarate, d'oxygène et de ferII. Comme pour les HMTs, chaque HDM possède une spécificité pour une lysine particulière et un degré de méthylation (Révisé dans (Black et al., 2012; Hyun et al., 2017; Torres and Fujimori, 2015).

La balance entre les HMTs et HDMs permet la mise en place de territoires chromatinien pouvant être distingués par leur profil de méthylation. Aussi, il existe de plus en plus d'évidences en faveur de la collaboration des machineries de méthylation des histones et de l'ADN pour la mise en place et le maintien de ces territoires, ce qui serait nécessaire au bon développement embryonnaire (Li et al., 2015; Patel, 2016; Rose and Klose, 2014)(**Figure 5**).

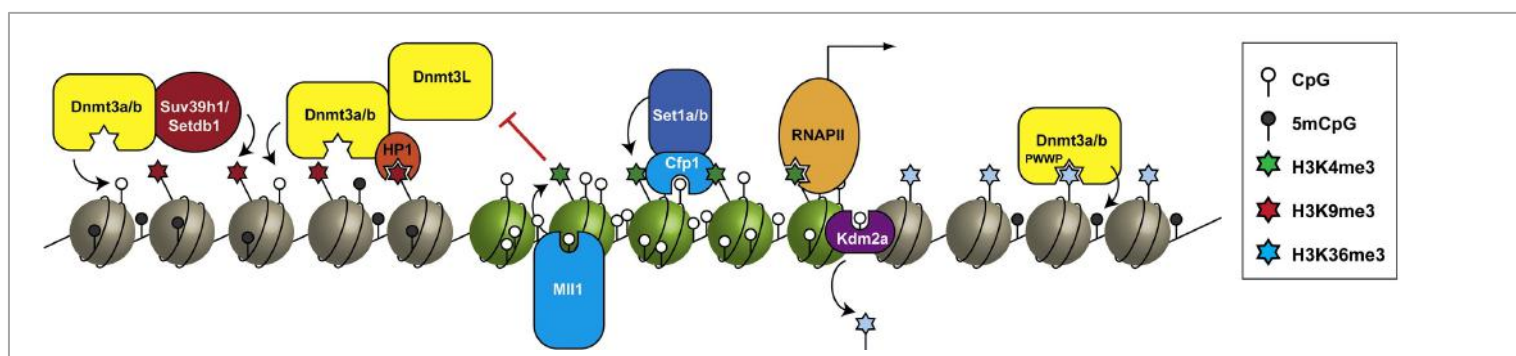


Figure 5: Lien entre la méthylation des histones et la méthylation de l'ADN (Rose and Klose, 2014).

Les régions réprimées sont enrichies en H3K9me3 et en méthylation de l'ADN tandis que les régions promotrices (en vert) sont enrichies en H3K4me3 et l'ADN y est déméthylé. Dans le corps des gènes, la présence de H3k36me3 va permettre le recrutement de la machinerie de méthylation de l'ADN.

(3) Phosphorylation

L'ajout d'un groupement phosphate sur la queue N-terminale des histones est catalysée par des kinases sur des résidus sérine et thréonine, et en est retiré par des phosphatases. Comme pour l'acétylation, le phosphate négativement chargé diminue l'affinité des histones pour l'ADN, conduisant à l'ouverture de la chromatine. La phosphorylation est donc

généralement associée à une activation de la transcription (Révisé dans (Ellenbroek and Youn, 2016; Rossetto et al., 2012).

(4) Ubiquitination

Les histones H2A et H2B sont les protéines ubiquitinées les plus abondantes du noyau. Il est estimé que 5 à 15% de H2A et 1-2% de H2B sont ubiquitinées chez les vertébrés tandis qu'environ 10% des histones H2B sont ubiquitinées chez la levure. Les histones sont majoritairement mono-ubiquitinées sur leurs chaînes C-terminales au niveau de la lysine 119 de l'histone H2A (H2Aub) et de la lysine 120 de l'histone H2B (H2Bub). Des études d'immuno-précipitation de la chromatine montrent que H2Aub est enrichi dans les régions satellites du génomes et est corrélé avec une extinction génique tandis que H2Bub est retrouvé dans le corps des gènes transcriptionnellement actifs et est généralement associé avec une activation de la transcription (Révisé dans (Cao and Yan, 2012; Shukla et al., 2009; Swatek and Komander, 2016).

(5) Le code histone

En plus de la séquence primaire d'ADN, l'ensemble des modifications des queues N-terminales des histones ainsi que leurs différentes combinaisons induisent donc une modification locale de l'état chromatinien capable de réguler l'expression des gènes. Ce niveau de régulation est appelé le « code histone ». Ce code nécessite différentes protéines pour le lire (readers), l'écrire (writers) et l'effacer (erasers) de manière spécifique pour chaque type de modification des histones. La plupart des acteurs sont connus pour chaque marque et ont été décrits pour chaque type de modification précédemment.

De manière intéressante, des analyses par spectrométrie de masse ont montré que certaines marques ont plus de chances d'être associées ensemble tandis que d'autres vont être mutuellement exclusives, suggérant une régulation croisée entre ces marques. De plus, ces régulations vont pouvoir se faire en *cis* sur la même queue d'histone ou en *trans* entre deux histones d'un même nucléosome ou de deux nucléosomes adjacents. La combinaison de ces différentes modifications post-traductionnelles des histones permet la création de territoires génomiques distincts selon leur état de compaction chromatinien et leur activité transcriptionnelle.

Par exemple, les régions génomiques ouvertes et actives transcriptionnellement sont globalement enrichies en acétylation des histones H3 et H4 ainsi qu'en marques de méthylation

H3K4me3, H3K79me3 et H3K36me3. Cependant, sur des régions génomiques plus précises, il existe aussi des associations spécifiques entre ces marques. En effet, les promoteurs des gènes sont fortement enrichis en H3K4me3 et hyperacétylés pour réguler l'initiation de la transcription tandis que la présence de la marque H3K36me3 dans le corps des gènes permet de recruter des HDACs et sera donc associée à une désacétylation de ces régions pour éviter une ré-initiation de la transcription. A l'inverse, les régions génomiques compactées et peu transcrites sont désacétylées et riches en marques répressives de méthylation H3K9me3 et H3K27me3 de manière mutuellement exclusive, générant respectivement l'hétérochromatine constitutive et facultative qui sera étudié plus en détail au chapitre suivant (Barski et al., 2007; Black et al., 2012; Mikkelsen et al., 2007; Schübeler et al., 2004; Shukla et al., 2009).

Finalement, il existe des boucles de régulation positive dans les régions transcriptionnellement actives entre les protéines qui mettent en place les marques activatrices et les marques elles-mêmes, quel que soit le type de marque (méthylation, acétylation et phosphorylation). Il en est de même pour les régions réprimées avec des boucles d'inhibition. De plus, il existe une inhibition mutuelle entre les acteurs et marques épigénétiques associées à des rôles opposés concernant l'activité de transcription de la région génomique, permettant le maintien d'états chromatiniens distincts entre les domaines.

(6) Les ARNs non-codants

Les ARNs non-codants (ncARN) sont de petites molécules d'ARN simple brin nucléaires capables de cibler spécifiquement des régions d'ADN complémentaires. Chez la levure, il a été rapporté que ces ncARN sont capables de se lier à la machinerie de méthylation des histones, pouvant ainsi jouer un rôle dans la formation de l'hétérochromatine. Ce mécanisme est appelé inhibition transcriptionnelle de gène (Transcriptional gene silencing ou TGS). Il a été suggéré que l'évolution de ce mécanisme permettrait la protection du génome face à des ADN étrangers ou des virus qui pourraient induire une instabilité génétique s'ils ne sont pas correctement inhibés (Martienssen and Moazed, 2015; Volpe et al., 2002).

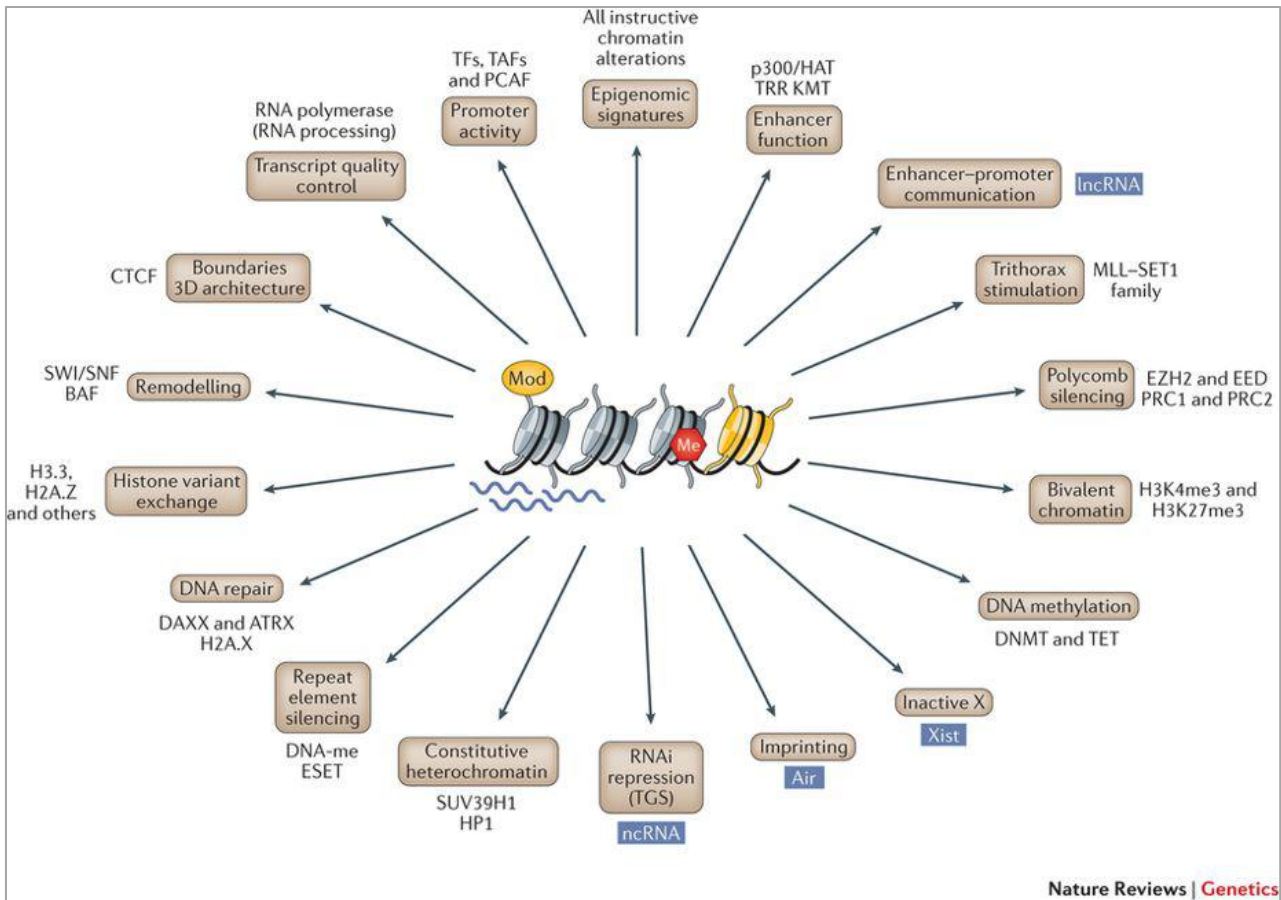


Figure 6: Les différentes modifications épigénétiques sont impliquées dans de nombreux mécanismes cellulaires et sont nécessaires au bon fonctionnement de la cellule (Allis and Jenuwein, 2016). Les facteurs majeurs de modification de la chromatine sont mentionnés.

3. Les territoires structuraux de la chromatine

a) L'euchromatine

L'euchromatine est le nom donné à la chromatine sous forme décondensée en interphase, c'est l'état de chromatine « actif ». Elle est riche en gènes et présente une forte activité transcriptionnelle.

Dans les régions d'euchromatine, l'ADN est hypométhylé et les marques d'histones principales sont des marques d'activation telles que des marques d'acétylation ainsi que de méthylation telle que H3K4me3 (Barski et al., 2007; Black et al., 2012; Mikkelsen et al., 2007; Schübeler et al., 2004; Shukla et al., 2009).

b) L'hétérochromatine

Le terme hétérochromatine a été introduit pour la première fois en 1928 par Emil Heitz en développant des colorants de la chromatine. Il a découvert que certaines régions ne sont colorées qu'en télophase (l'euchromatine) tandis que d'autres régions sont colorées tout au long du cycle cellulaire (l'hétérochromatine). De plus, il a observé que certaines régions n'étaient colorées que dans certains types cellulaires, ce qui a été appelé plus tard l'hétérochromatine facultative, contrairement à d'autres régions constitutivement colorées, l'hétérochromatine constitutive (Passarge, 1979).

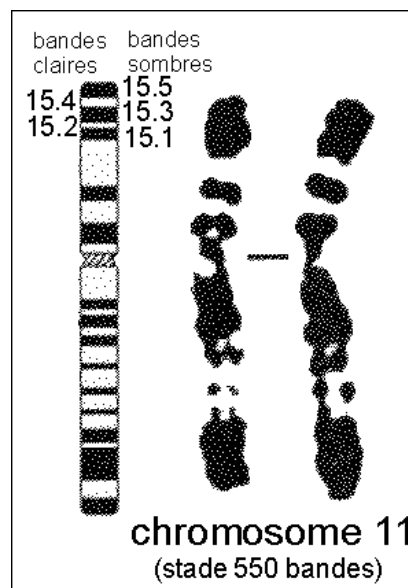


Figure 7: Observation de chromosome métaphasique après coloration au Giemsa (Biologie-Géologie, 1996). Nous pouvons observer distinctivement des bandes claires d'euchromatine et des bandes sombres d'hétérochromatine.

L'hétérochromatine est un élément très conservé et fondamental pour l'architecture nucléaire. Elle représente entre 25 et 90% des génomes eucaryotes multicellulaires. C'est un état chromatinien très condensé, fermé, pauvre en gènes qui présentent une faible activité de transcription. Ces régions sont très riches en séquences répétées (Microsatellites, Majorsatellites, transposons, ADN ribosomiaux) ainsi qu'en éléments transposables tels que les LINEs ou les SINEs (Allshire and Madhani, 2018; Lander et al., 2001; Vicient and Casacuberta, 2017).

L'hétérochromatine peut être séparée en 2 catégories: l'hétérochromatine constitutive et l'hétérochromatine facultative, avec chacune leurs marques d'histones et leurs caractéristiques propres.

(1) L'hétérochromatine constitutive

(a) Structure

L'hétérochromatine constitutive (c-Het) est essentielle pour le maintien structurel et la stabilité du génome. Elle se trouve principalement au niveau de séquences répétées telles que des transposons ou l'ADN satellite, des centromères et des télomères et permet le maintien de la forte compaction de ces régions. Il a longtemps été pensé que cette structure était stable mais il a été montré récemment que l'hétérochromatine est une structure très dynamique (Larson et al., 2017; Strom et al., 2017). Elle est très enrichie en marque d'histone H3K9me2/3 et en protéine « Heterochromatin Protein 1 » ou HP1 alors que les histones sont généralement hypo-acétylées (Alper et al., 2013; James and Elgin, 1986; Taddei et al., 2001). Chez l'homme, HP1 est une famille de 3 protéines essentielles à la mise en place de l'hétérochromatine (HP1 α , HP1 β et HP1 γ), représentant les composants principaux de l'hétérochromatine constitutive. Ces 3 isoformes possèdent l'organisation caractéristique de HP1 mais diffèrent dans leurs fonctions, leurs expressions et leurs localisations. Par exemple, HP1 α et HP1 β sont principalement associées à l'hétérochromatine, telles que les régions centromériques et télomériques, et participent à l'inhibition de la transcription dans ces régions. Bien qu'elle soit présente dans l'hétérochromatine, HP1 γ est aussi largement retrouvée localisée à l'euchromatine où elle joue un rôle dans l'élongation de la transcription et la maturation des ARN. Aussi, il a été démontré que la méthylation de l'histone H3K9 peut conduire à la méthylation de l'ADN de ces régions, augmentant ainsi la stabilité de l'hétérochromatine (Li et al., 2015; Patel, 2016; Rose and Klose, 2014). L'hétérochromatine constitutive s'organise sous forme de territoires dans le noyau des cellules. Par exemple, dans les cellules de souris, les régions de c-Het sont directement visualisables comme des zones plus intenses après simple marquage d'ADN ; ces régions sont appelées chromocentres car ils sont plus intenses en fluorescence. Dans les cellules humaines, l'hétérochromatine constitutive est principalement observable autour des nucléoles et en périphérie nucléaire (Mayer et al., 2005).

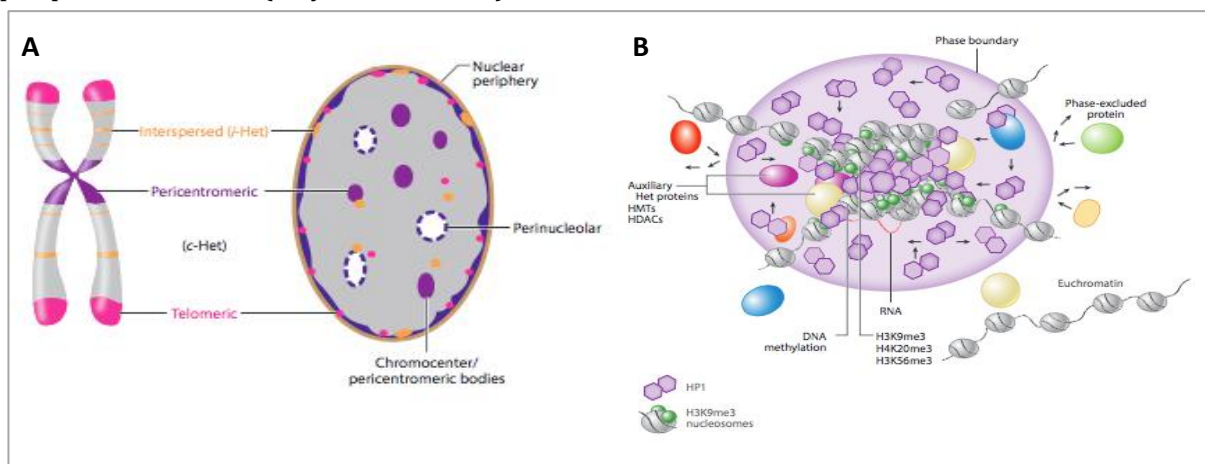


Figure 8 : L'hétérochromatine se localise dans des territoires distincts du noyau (Janssen et al., 2018). A) Dans les cellules murines, l'hétérochromatine se localise préférentiellement en périphérie du noyau, autour des nucléoles et dans les chromocentres. B) Organisation d'un chromocentre regroupant l'ADN condensé ainsi que toutes les protéines constitutives de l'hétérochromatine.

(b) Mise en place et maintien de l'hétérochromatine constitutive

La mise en place de l'hétérochromatine constitutive se fait en trois étapes : l'initiation, la propagation et le maintien (Allshire and Madhani, 2018; Hall et al., 2002; Janssen et al., 2018; Wang et al., 2016).

Au cours de l'initiation, des protéines de liaison à l'ADN spécifiques de séquence ou des ARNs non-codants sont recrutés au niveau de centres de nucléation. Cela induit le recrutement de HDACs et de HMTs-H3K9 appartenant à la famille Suv39 résultant en une hypo-acétylation et une hyperméthylation des histones présentes dans ces sites de nucléation.

L'hétérochromatine s'étend ensuite aux régions adjacentes au point de nucléation. En effet, les HMTs-H3K9 possèdent souvent des chromodomaines leur permettant de se lier à H3K9me3, ce qui va engendrer la propagation de cette marque. De plus, les protéines HP1 possèdent aussi un chromodomaine leur permettant de se lier à H3K9me3 et servent de plateforme pour le recrutement d'autres enzymes de remodelage de la chromatine incluant Suv420H pour l'apposition de marques H4K20me3, des HDACs ainsi que des DNMTs conduisant à la méthylation de l'ADN. La combinaison de ces deux actions entraîne la propagation indépendante de séquence de l'hétérochromatine sur de larges domaines chromatiniens.

Au cours de la réplication, les histones parentales porteuses de modifications post-traductionnelles sont incorporées de manière aléatoire entre les deux copies d'ADN. La présence d'H3K9me induit de manière spécifique le recrutement des HMTs-H3K9, permettant le maintien de cet état épigénétique au cours des générations.

Enfin, plusieurs mécanismes de « barrières », tels que la présence d'ARNs de transfert ou la présence de régions transcriptionnellement actives, existent pour empêcher la propagation de l'hétérochromatine qui serait délétère pour les régions d'euchromatine alentour (Ahmad and Henikoff, 2002; Allshire and Madhani, 2018).

(2) L'hétérochromatine facultative

L'hétérochromatine facultative est la chromatine dont l'état de compaction est modulable en fonction du contexte spatial, temporel et de l'empreinte parentale. De manière similaire à l'hétérochromatine constitutive, l'hétérochromatine facultative est très peu active en terme de transcription. Cependant, elle contient quelques gènes, principalement développementaux, qui peuvent être exprimés en fonction du type cellulaire. Son état de

compaction peut alors être modulé pour activer l'expression de ces gènes en fonction du contexte temporel (développement), spatial (localisation nucléaire) et génétique (expression génique mono-allélique). L'hétérochromatine facultative est fortement enrichie en marques d'histones H3K27me3 et H4K20me1 ainsi qu'en protéines du groupe Polycomb. Comme pour l'hétérochromatine constitutive, l'ADN y est hyper-méthylé et hypo-acétylé. Par exemple, le cluster de gènes *Hox* est très fortement exprimé au cours du développement embryonnaire puis est réprimé après différenciation cellulaire par hétérochromatinisation (Trojer and Reinberg, 2007).

Le mécanisme de mise en place de l'hétérochromatine facultative est peu compris. Récemment, il a été montré chez la levure que sa mise en place dépendrait de protéines appartenant au complexe des shelterines, formant la coiffe des télomères, qui seraient capables de recruter la machinerie de remodelage de la chromatine au niveau d'origines de réplication tardive. En effet, l'inhibition d'une de ces protéines empêche la formation d'hétérochromatine dans ces régions, conduisant à une dérégulation des gènes présents (Zofall et al., 2016). Ce résultat démontre un lien entre la mise en place de l'hétérochromatine et le contrôle de la réplication et souligne donc l'importance de ce mécanisme dans le maintien de la stabilité génomique.

B. La réplication de l'ADN chez les eucaryotes supérieurs

Les mécanismes de réplication de l'ADN sont fortement conservés au cours de l'évolution. La réplication s'initie au niveau d'origines de réplication et se fait de manière bidirectionnelle à partir de celles-ci, toujours dans le sens 5' vers 3'. C'est un mécanisme semi-conservatif se servant d'un brin d'ADN comme matrice pour la synthèse d'un nouveau brin.

1. Le réplisome

Le réplisome désigne l'ensemble des protéines localisées à la fourche de réplication qui vont contribuer à la réplication de l'ADN. Il est constitué de plusieurs enzymes portant différentes activités impliquées directement dans la réplication de l'ADN ou facilitant le processus (Burgers and Kunkel, 2017).

Chez les eucaryotes, le réplisome canonique est constitué d'une dizaine de protéines :

- Un chargeur de clamp, RCF, permettant le chargement des autres protéines sur l'ADN.
- Une hélicase permettant l'ouverture de la double hélice d'ADN.

- Une topoisomérase permettant de dérouler la double hélice d'ADN au fur et à mesure de l'avancée de la fourche.
- Deux facteurs de processivité PCNA, un sur chaque brin d'ADN. Ce sont des anneaux protéiques augmentant la processivité des ADN polymérases.
- Sur le brin direct, la Polymérase répliquative ϵ (Pol ϵ) réplique l'ADN de manière continue.
- Sur le brin indirect, la Polymérase répliquative δ (Pol δ) ainsi que la Primase α (Pol α) répliquent l'ADN de manière discontinue avec la synthèse d'une amorce ARN par Pol α suivie de l'élongation par Pol δ .
- La RNase H et une ligase, sur le brin indirect, permettent l'excision de l'amorce ARN et la ligature des multiples fragments d'Okazaki.

Chaque origine de réplication possède deux réplisomes, un de chaque côté de la fourche.

2. Les ADN polymérases

Les ADN polymérases sont des protéines capables de répliquer l'ADN. Chez les eucaryotes, il existe six familles de polymérases classées selon leurs caractéristiques structurelles et fonctionnelles récapitulées dans le tableau suivant (Burgers et al., 2001; Garcia-Diaz and Bebenek, 2007; Hubscher et al., 2002; Jain et al., 2018; Yang and Gao, 2018).

La famille A regroupe les polymérases γ , ν et θ . Pol γ est une polymérase de haute-fidélité. Elle est la seule polymérase responsable de la réplication et de la réparation de l'ADN mitochondrial. Pol ν et Pol θ sont, au contraire, peu fidèles et sont impliquées dans la synthèse translésionnelle ainsi que dans la réparation de l'ADN. En particulier, Pol θ est essentielle pour la voie d'« alternative end-joining » et c'est la seule polymérase eucaryote possédant un domaine hélicase.

La famille B regroupe les ADN polymérases répliquatives α , δ et ϵ responsables de la synthèse d'ADN au cours de la réplication physiologique ainsi que la polymérase ζ , une polymérase translésionnelle appartenant à cette famille de par sa ressemblance structurale avec les polymérases répliquatives.

La famille X regroupe les polymérases impliquées dans la réparation de l'ADN. Ces polymérases sont responsables de la resynthèse d'ADN à la suite de l'excision d'une base ou d'un brin d'ADN endommagé par les protéines des différents mécanismes de réparation.

La famille Y regroupe les polymérases translésionnelles impliquées dans le mécanisme de tolérance des dommages de l'ADN. Ces polymérases permettent la synthèse d'ADN à travers une lésion et assurent ainsi la continuité de la réplication même en présence de dommage.

La famille RT regroupe les polymérases dotées d'une activité de transcriptase inverse leur permettant de synthétiser de l'ADN à partir d'une matrice ARN, contrairement aux autres polymérases capables uniquement de synthétiser de l'ADN à partir d'une matrice ADN.

La famille AEP (Archeo-Eukaryotic primase) ne contient que la protéine Primpol, une polymérase découverte récemment, possédant une double activité de primase ADN et de polymérase translésionnelle.

Les polymérases répliquatives et translésionnelles seront étudiées plus en détail dans les paragraphes suivants.

Famille	Polymérase	Fonction
A	Pol γ	Réplication de l'ADN mitochondrial
	Pol θ	Réparation
	Pol ν	Réparation
B	Pol α	Primase et Réparation
	Pol δ	Réplication
	Pol ϵ	Réplication
	Pol ζ (Rev3)	Synthèse translésionnelle
X	Pol β	Réparation
	Pol λ	Réparation
	Pol μ	Réparation
	Pol σ	Réparation
Y	Pol η	Synthèse translésionnelle
	Pol ι	Synthèse translésionnelle
	Pol κ	Synthèse translésionnelle
	Rev1	Synthèse translésionnelle
RT	Reverse Transcriptase	Copie des rétrotransposons
	Telomerase	Elongation des télomères
AEP	Primpol	Primase et Synthèse translésionnelle

Figure 9: Tableau récapitulatif des différentes familles de polymérases eucaryotes et leurs fonctions. La dénomination des polymérases humaines est indiquée. Elles possèdent toutes un homologue procaryote à l'exception de Pol θ , Pol ν et la transcriptase inverse.

Chez les procaryotes, il existe deux familles de polymérases supplémentaires : la famille C regroupant PolIII et S. pyo DnaE chez les bactéries et la famille D composée uniquement de PolD chez les archéobactéries.

3. La réplication de l'ADN

a) La réplication de l'ADN en conditions physiologiques

(1) Les ADN polymérases répliquatives

Les ADN polymérases répliquatives sont responsables de la réplication de l'ADN en conditions physiologiques. Elles appartiennent toutes à la famille B. Chez l'Homme, Pol α , Pol δ et Pol ϵ sont des complexes enzymatiques composés de quatre sous-unités protéiques dont une seule porte l'activité de polymérisation de l'ADN (respectivement POLA1, POLD1 et POLE). Ces enzymes sont très fidèles avec un taux d'erreur variant de 10^{-4} à 10^{-6} selon la séquence répliquée et le pool de nucléotides disponible au moment de la réplication (Hubscher et al., 2002; Jain et al., 2018; Yang and Gao, 2018).

La structure par cristallographie de ces trois sous-unités catalytiques a révélé une structure commune aux polymérases répliquatives de la famille B comprenant 5 domaines structuraux comparés à une main humaine (Jain et al., 2018):

- une paume contenant les résidus catalytiques,
- les doigts entourant le nouveau brin synthétisé,
- un pouce qui se lie au petit sillon de l'ADN,
- un domaine exonucléase actif chez Pol δ et Pol ϵ mais inactif chez Pol α ,
- un domaine N-terminal.

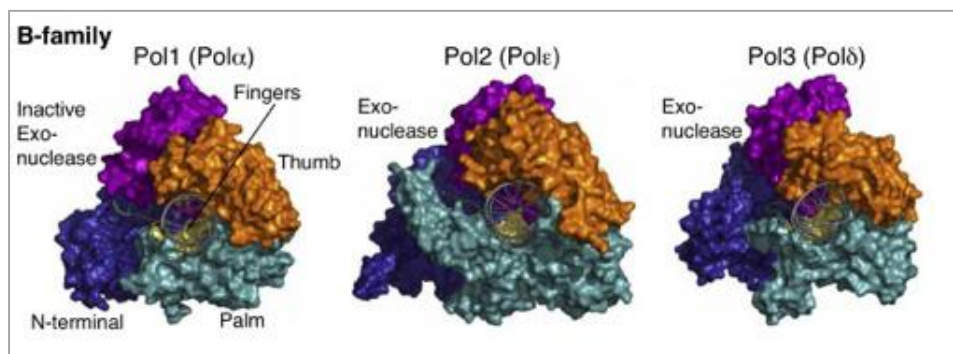


Figure 10: Structure par cristallographie des sous-unités catalytiques des polymérases répliquatives de la famille B (Jain et al., 2018).

(a) Pol α

L'ADN polymérase α est un hétérotétramère composé d'une sous-unité polymérase (POLA1 et POLA2) et d'une sous-unité primase (PRIM1 et PRIM2). Chacune de ces sous-unités est composée d'une sous-unité catalytique (POLA1 et PRIM1) et d'une sous-unité régulatrice (POLA2 et PRIM2). PRIM1 et POLA1 portent respectivement les activités ARN et ADN-

polymérase mais, contrairement à Pol δ et Pol ϵ , POLA1 ne possède pas d'activité exonucléase, la rendant ainsi plus mutagène. Sur les brins direct et indirect, Pol α synthétise un primer d'ARN d'une longueur de 9 nucléotides grâce à son activité primase puis switcher vers son activité polymérase et continuer la réplication d'ADN sur une longueur de 20 à 30 nucléotides. Sur le brin indirect, répliqué de manière discontinue, ces primers d'ARN sont appelés fragments d'Okazaki.

(b) Pol δ

L'ADN polymérase δ est composée de 4 sous-unités, POLD1, POLD2, POLD3 et POLD4. POLD1 est la sous-unité catalytique portant une activité polymérase ainsi qu'une activité 3',5'-exonucléase, les autres protéines ayant des fonctions régulatrices. Cette activité exonucléase confère à Pol δ la capacité de revenir en arrière si un mauvais nucléotide a été inséré jusqu'à 5 bases en amont. Pol δ est responsable de la réplication du brin indirect de manière discontinue à la suite des fragments d'Okazaki synthétisés par Pol α .

(c) Pol ϵ

L'ADN polymérase ϵ est composée de 4 sous-unités, POLE1, POLE2, POLE3 et POLE4. La sous-unité catalytique est la protéine POLE1 qui est environ deux fois plus grosse que POLD1. POLE1 possède une activité polymérase et exonucléase. Pol ϵ est intrinsèquement plus processive que Pol δ et est responsable de la synthèse en continu du brin direct. Des mutations germinales des domaines exonucléases des sous-unités catalytiques de Pol δ et Pol ϵ ont été associées avec l'apparition de cancers « hypermutés », porteurs d'un grand nombre de mutations.

(2) Déroulé de la réplication

(a) Initiation

L'initiation de la réplication se fait au niveau des origines de réplication, des séquences d'ADN spécifiques sur lesquelles se fixent les protéines du complexe ORC.

L'initiation peut se résumer en 4 étapes :

- Fixation des protéines ORC sur les origines de réplication.

- Association des complexes CMG en double hexamère constitués des hélicases MCM2-7 associées par leurs extrémités N-terminales, CDC45 et le complexe GINS sur les sites marqués par les protéines ORC.
- Les extrémités C-terminales des complexes CMG portent une activité ATPase responsable de la translocation des complexes et de l'ouverture de la double hélice d'ADN.
- Les deux complexes CMG couissent alors l'un par rapport à l'autre de telle sorte que leur extrémité N-terminale soit placée en tête par rapport au sens de la réplication.

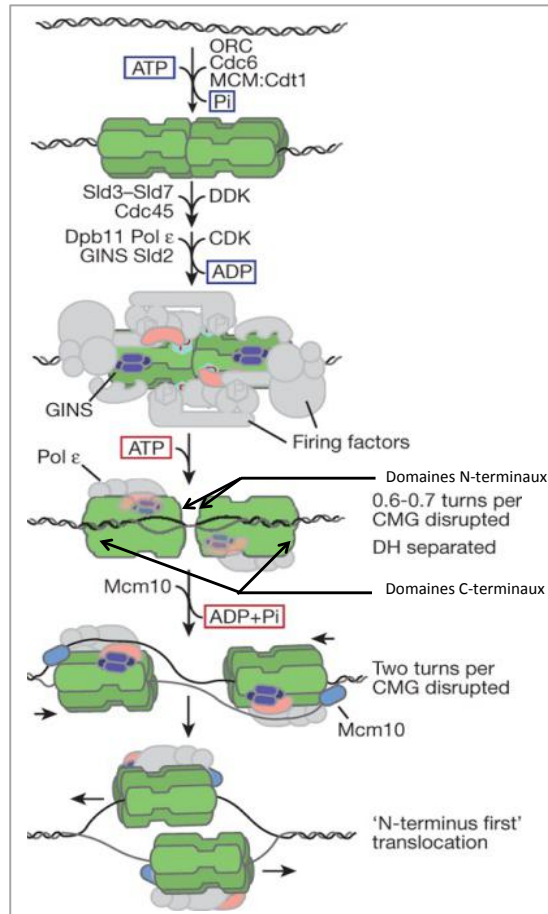


Figure 11 : Initiation de la réplication de l'ADN (Adapté de (Douglas et al., 2018)).

Contrairement à l'ancien modèle d'initiation où les hélicases suivaient le sens de réplication par leur extrémité C-terminale, ce nouveau modèle de coulissement des complexes CMG permet de réconcilier le positionnement des ADN polymérases répliquatives par rapport au sens de la réplication. En effet, Pol α s'associant avec l'extrémité C-terminale de l'hélicase MCM2-7 et Pol ϵ avec l'extrémité C-terminale de celle-ci, ce modèle permet la synthèse d'ADN pour chacun des brins (direct et indirect) directement dès l'ouverture de la double hélice et limite ainsi la formation d'ADN simple brin à l'avant de la fourche (Douglas et al., 2018; Goswami et al., 2018).

(b) Élongation

La réplication de l'ADN se fait toujours dans le sens 5'OH vers 3'P. Le brin direct 5'3' est répliqué de manière continue par Pol ϵ tandis que le brin indirect 3'5' est répliqué de manière discontinue par Pol α et Pol δ pour suivre le sens de la fourche unidirectionnelle grâce à la formation d'une boucle du brin matrice d'ADN (Burgers and Kunkel, 2017; Kelly and Callegari, 2019). Ce dogme est aujourd'hui remis en question. En effet, plusieurs études chez la levure ainsi qu'*in vitro* ont montré que Pol δ serait capable d'initier la réplication du brin direct avant que Pol ϵ ne reprenne la suite de la réplication (Burgers and Kunkel, 2017; Johnson et al., 2015; Stillman, 2015).

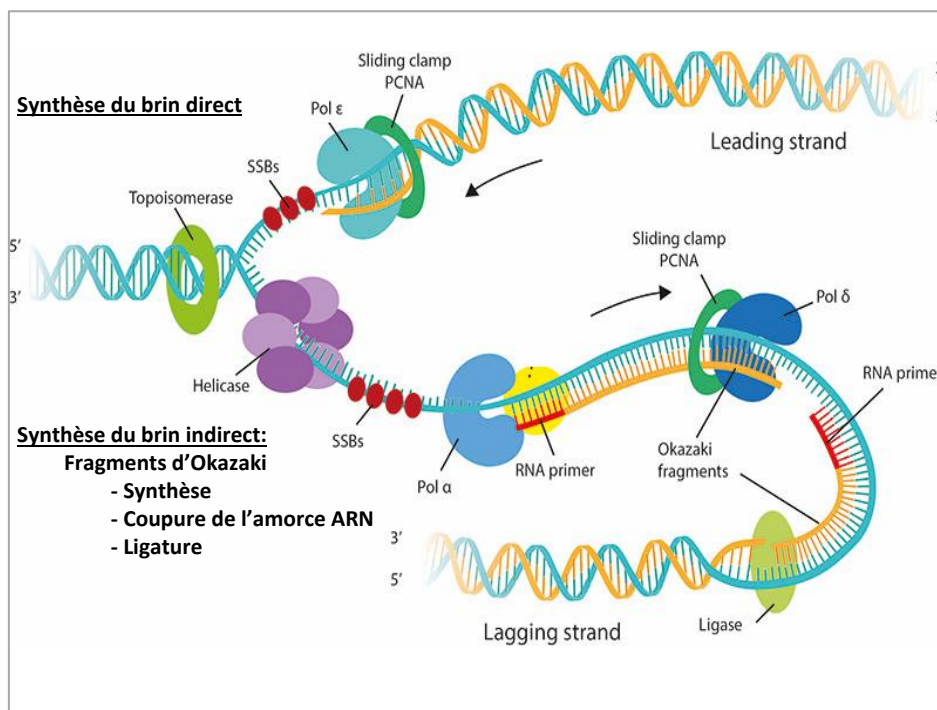


Figure 12: Représentation du réplisome en cours d'élongation au niveau d'une fourche de réplication (Adapté de <https://www.mechanobio.info>).

(c) Terminaison

La terminaison de la réplication se fait lorsque deux fourches de réplication se rejoignent, la réplication étant achevée de part et d'autre. Les deux fourches convergent jusqu'à ce que la totalité de l'ADN soit déroulé, les régions non répliquées restantes sont remplies, ligaturées et les protéines du réplisome relâchées de l'ADN. Chez certaines bactéries, cette terminaison se fait au niveau de séquences spécifiques appelées *Ter*. Chez les eucaryotes, à cause de l'activation aléatoire des origines de réplication, les sites de terminaison sont aussi aléatoires et déterminés par l'activation des origines alentour (Dewar and Walter, 2017).

b) Cas particulier des régions centromériques et péri-centromériques

(1) Structure

Les centromères sont les points de fusion entre deux chromatides nécessaires au recrutement des cohésines et à la ségrégation adéquate des chromosomes en mitose. Dans les cellules humaines, les centromères sont constitués de longues séquences d'ADN répétées presque identiques appelées α -satellites compliquant l'analyse des centromères par séquençage. Les centromères portent spécifiquement le variant d'histone H3 CENP-A qui joue un rôle critique pour la spécification des centromères. CENP-A entre dans la composition des nucléosomes centromériques en formant un octamère avec les histones canoniques H4, H2A et H2B. Les nucléosomes comportant CENP-A représentent 4% de l'ADN centromérique (Forsburg and Shen, 2017; Guenatri et al., 2004; Romeo et al., 2016).

Dans les cellules de mammifères, les régions d'ADN entourant les centromères sont appelées les régions péri-centromériques. Ces régions contribuent à la cohésion chromosomique, peuvent réguler les fonctions des centromères et créent une barrière entre les régions géniques chromosomiques et non-géniques centromériques. Les régions péri-centromériques sont des régions d'hétérochromatine constitutive enrichies en ADN hypométhylé et en marque H3K9me3 créant un environnement répressif pour l'expression des gènes présents. Comme expliqué précédemment, sa mise en place dépend des histones méthyltransférases Suv39 et du recrutement de la protéine HP1 (Baumann et al., 2008; Déjardin, 2015; Guenatri et al., 2004; Saksouk et al., 2015; Strålfors and Ekwall, 2011). Il n'existe pas de séquence consensus de ces régions mais elles sont toujours constituées de courtes répétitions satellites en tandem. Chez la souris, l'hétérochromatine péri-centromérique mesure environ 1 Mb et est constituée de répétitions en tandem de 234 pb riches en AT appelées majorsatellites regroupées au niveau des chromocentres.

Cependant, la composition des régions péri-centromériques est très dynamique et est capable de faire la transition de l'hétérochromatine constitutive vers de l'hétérochromatine facultative dans certains tissus et dans les cellules cancéreuses. Cette transition serait due au recrutement de protéines Polycomb capables de recruter des protéines régulatrices de la transcription (Déjardin, 2015). La structure des centromères et des régions péri-centromériques est fondamentale pour le recrutement des cohésines, la mise en place des kinétochores et le bon déroulement de la mitose. Le gain ou la perte d'un centromère entraîne des problèmes de ségrégation et de cassures chromosomiques pouvant conduire à une aneuploïdie ainsi que d'autres désordres génétiques entraînant potentiellement le développement d'un cancer.

(2) Réplication

La réplication des régions centromériques et péricentromériques nécessite un travail coordonné des machineries de réplication, de mise en place de l'hétérochromatine et de séparation des chromosomes pour maintenir la stabilité génomique. La perte d'hétérochromatine au cours de la réplication dans ces régions les rend très sensibles aux défauts de réplication pouvant induire des réarrangements chromosomiques. Le variant d'histone CENP-A étant essentiel pour le maintien des centromères, la chromatine portant CENP-A doit être préservée au cours des divisions cellulaires afin de maintenir l'intégrité centromérique. Dans les cellules humaines, CENP-A est déposé *de novo* au niveau des centromères en début de phase G1 par sa chaperonne spécifique HJURP. Au cours de la réplication des centromères, CENP-A est éjecté de la chromatine en phase S tardive par le réplisome. MCM-2 et HJURP seraient alors capables de retenir CENP-A sur la chromatine centromérique après son éjection par la fourche de réplication. CENP-A va alors être recyclé sur l'ADN parental de manière couplée à la réplication par CAF-1 (une protéine chaperon impliquée dans la déposition des histones au cours de la réplication) et MCM2-7 (une hélicase réplivative) (Zasadzińska et al., 2018). La protéine CENP-C de CCAN, un sous-complexe du kinétochore capable de se lier aux centromères, va aussi jouer un rôle critique pour le re-chargement de CENP-A au niveau des centromères après le passage de la fourche de réplication (Nechemia-Arbely et al., 2019). L'ensemble de ces mécanismes permet ainsi le maintien de l'hétérochromatine centromérique au cours des divisions cellulaires.

De manière intéressante, les régions péricentromériques nécessitent la présence de la protéine chaperon télomérique TRF2 pour assister le réplisome. TRF2 va se lier aux différentes structures non-B de l'ADN dans ces régions représentant des barrières pour la fourche de réplication. La présence de TRF2 sur ces structures empêche alors l'activation de la voie de réparation ATM et va recruter l'hélicase RTEL1 ainsi que d'autres enzymes pour supprimer ces barrières, continuer la réplication et assurer la stabilité des régions péricentromériques (Mendez-Bermudez et al., 2018).

c) Cas particulier des télomères

(1) Structure

Les télomères sont les extrémités physiques des chromosomes linéaires eucaryotes. Chez les mammifères, les télomères sont des régions d'hétérochromatine constitutive constituées de répétitions (TTAGGG)_n en tandem d'une longueur allant de 5-15kb chez l'Homme à plus de

40kb chez la souris. Elles sont constituées d'une extrémité 3' sortante susceptible d'être digérée par des exonucléases. Le raccourcissement des télomères est responsable du vieillissement cellulaire et la taille des télomères est considérée comme une « horloge » moléculaire. En effet, lorsque les télomères atteignent une taille minimale, une senescence répliquative est induite pouvant conduire à l'arrêt de la prolifération cellulaire. Pour ralentir ce processus, l'ADN télomérique forme une boucle à la fin du chromosome, la t-loop, où l'extrémité 3' sortante est invaginée dans le double brin d'ADN en amont par un mécanisme de déplacement de brin. De plus, cette structure est protégée par les protéines dites de « la coiffe », 6 protéines formant le complexe des shelterines : Trf1, Trf2, Rap1, Tin2, Tpp1, Pot1. Trf1 et Trf2 se lient directement à l'ADN télomérique double brin ainsi qu'à Tin2 et Tin2 se lie à Tpp1 qui se lie à Pot1. Trf2 recrute aussi Rap1 aux télomères. Ce mécanisme de protection des télomères empêche donc leur dégradation mais aussi une fusion télomérique entre deux extrémités 3' sortantes qui pourrait être responsable d'une forte instabilité génétique (Shay and Wright, 2019).

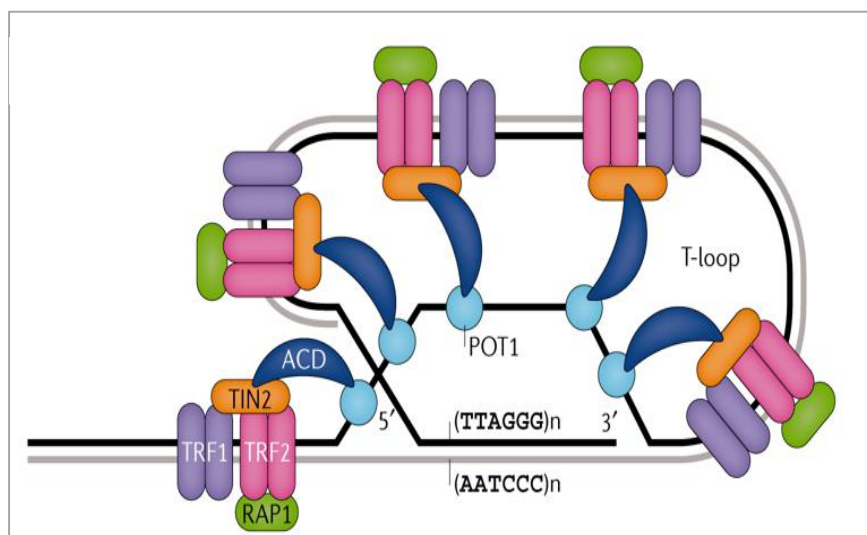


Figure 13: Structure des télomères de mammifères sous forme de t-loop associés aux protéines du complexe shelterin (Shay and Wright, 2019).

(2) Réplication

Dans les cellules de mammifères, les télomères sont répliqués tout au long de la phase S, contrairement à la levure où ils sont répliqués en phase S tardive. La réplication des télomères se fait de manière semi-conservative, unidirectionnelle et à partir d'une origine centromère-proximale. A chaque division cellulaire, les télomères raccourcissent à cause d'une réplication incomplète de l'ADN linéaire par les ADN polymérases répliquatives. Ce phénomène est appelé le problème de réplication des extrémités et est spécifiquement dû à un mécanisme de résection après la réplication du brin direct afin de recréer une extrémité 3' sortante et la T-loop (Maestroni et al., 2017). Lors du développement embryonnaire, pour pallier cette perte

irréversible lors des divisions cellulaires successives, une transcriptase inverse capable d'étendre l'extrémité 3' sortante par l'addition de répétitions télomériques, la télomérase, est exprimée et va être recrutée au niveau des télomères par les shelterines.

La télomérase est une nucléoprotéine composée au minimum d'un ARN guide TERC complémentaire de la séquence des télomères et d'une sous-unité catalytique, la transcriptase inverse TERT. Après l'addition d'une répétition, la télomérase peut alors soit se déplacer en 3' et recommencer l'opération, soit se décrocher des télomères. En absence de télomérase, les télomères se raccourcissent à chaque division jusqu'à entraîner une inhibition de la prolifération et la sénescence cellulaire ou l'apoptose en fonction du contexte cellulaire. Or, une des caractéristiques majeures des cellules cancéreuses est leur capacité de prolifération infinie. Ce phénomène est dû à la réactivation de la télomérase dans 85%-90% des tumeurs ou à un mécanisme alternatif de croissance des télomères, la voie ALT, par recombinaison homologue dans 10-15% des tumeurs (Pfeiffer and Lingner, 2013; Saretzki, 2018; Shay, 2016; Shay and Wright, 2019).

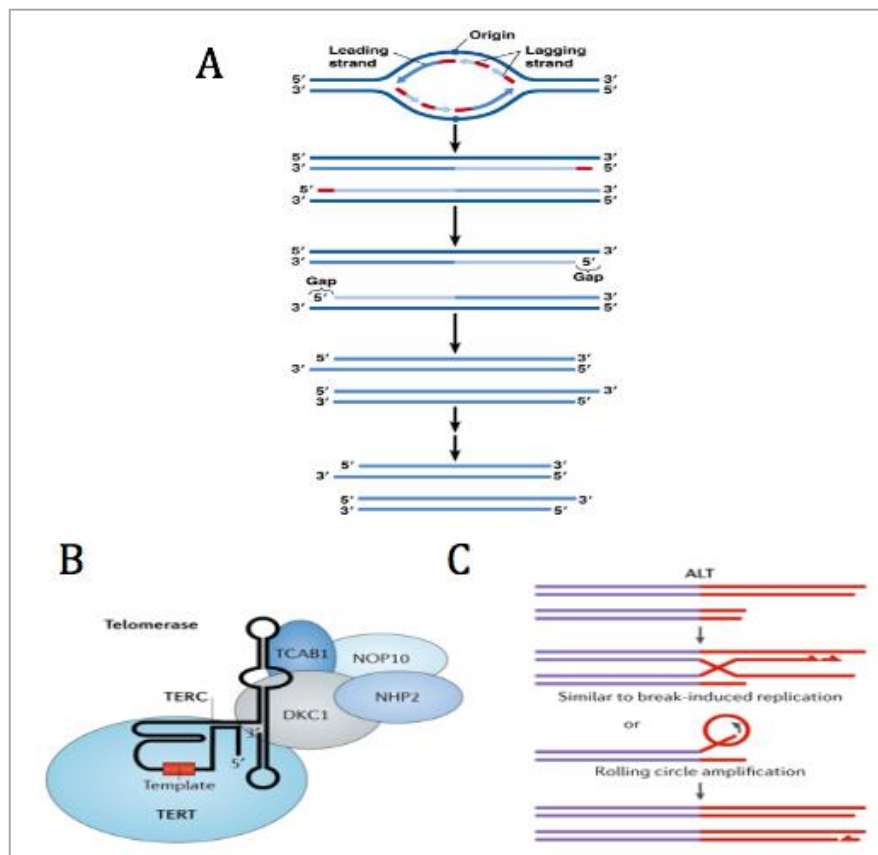


Figure 14: Réplication des télomères (Pearson Education, 2012 ; (Shay and Wright, 2019) A) Problème de réplication de la fin des chromosomes linéaires. Alors que le brin direct peut être répliqué en totalité, le brin indirect va subir un raccourcissement dû à l'excision de l'amorce ARN en 5'. Aucune polymérase ne peut venir répliquer cette extrémité car il n'y a pas d'extrémité 3'OH pour l'ajout d'un nucléotide. Chaque cycle de réplication va donc engendrer des molécules d'ADN de plus en plus courtes. B) Structure de la télomérase réactivée dans 85-90% des tumeurs. C) La voie ALT, mécanisme alternatif d'élongation des télomères par recombinaison homologue, activé dans 10-15% des tumeurs.

d) Drogues affectant la réplication de l'ADN

(1) Aphidicoline

L'aphidicoline (APH) est un antibiotique isolé à partir du champignon *Nigrospora sphaerica*. Cette drogue est un inhibiteur des polymérase répliquatives α , δ , et ϵ . Elle se lie à leur site actif et bloque spécifiquement l'incorporation de dCTPs. L'APH est donc spécifique des cellules en cours de réplication. Les cellules en dehors de la phase S peuvent continuer leur cycle cellulaire jusqu'au point de contrôle G1/S suivant où elles vont s'accumuler. A faible concentration (1 μ M), l'APH induit un ralentissement de la machinerie de réplication, les cellules mettant alors plus de temps pour terminer leur phase S. À forte dose (>5 μ M), l'APH bloque la progression des polymérase et entraîne un blocage complet de la fourche de réplication. Les polymérase sont bloquées tandis que l'hélicase continue à ouvrir les deux brins d'ADN, créant ainsi un découplage du réplisome. L'ADN simple brin créé constitue un domaine vulnérable aux cassures chromosomiques, en particulier au niveau des sites fragiles communs, et peut induire une instabilité génétique. Cependant, les effets de l'APH sont réversibles si les cellules ne sont cultivées en sa présence que pendant pas plus de 2 générations (Vesela et al., 2017).

(2) Hydroxyurée

L'hydroxyurée (HU) a été synthétisée pour la première fois au cours du 19^{ème} siècle. C'est un inhibiteur de la ribonucléotide réductase (RNR), une enzyme capable de convertir les NTPs en dNTPs qui est l'étape limitante de la synthèse nucléotidique. Chaque type de dNTPs va être affecté de manière différente. En inhibant la RNR, l'HU induit une diminution du pool de nucléotides qui va entraîner une diminution du nombre d'évènements d'initiation de la réplication ainsi que de la vitesse des fourches de réplication engendrant finalement la création d'ADN simple brin plus vulnérable. Cette propriété de l'HU est donc spécifique des cellules en réplication. Cependant, elle va aussi affecter la réparation de l'ADN ainsi que la synthèse d'ADN mitochondrial indépendamment du cycle cellulaire. L'HU induit aussi l'expression de sites fragiles précocement répliqués. L'HU est utilisée en médecine pour traiter des patients atteints de thrombocytose, de désordre myéloprolifératif, de psoriasis et peut être utilisée comme agent antirétroviral chez des patients atteints de SIDA (Vesela et al., 2017).

(3) Cisplatine

Le cisplatine (*cisPt*) est un composé inorganique synthétisé par Michel Peyrone qui induit des pontages intrabrins (96%) et interbrins (1%), des mono-adduits d'ADN (2%) ainsi que des pontages ADN-protéines (<1%). Ces modifications structurales de l'ADN empêchent l'ouverture des deux bris ainsi que le déroulement de la double hélice d'ADN, deux événements nécessaires à la réplication et à la transcription de l'ADN. Le *cisPt* induit donc un blocage des fourches de réplication, inhibe la transcription et conduit à la mort cellulaire par apoptose à concentration élevée. En clinique, le *cisPt* est largement utilisé en chimiothérapie dans un vaste spectre de cancers tels que le cancer des testicules, de l'ovaire, de la vessie, de l'œsophage et le cancer du poumon non à petites cellules (Vesela et al., 2017).

(4) Camptothécine

La camptothécine (CPT) a été isolée à partir d'un arbre chinois *Camptotheca acuminata* par Wall et al. La CPT cible spécifiquement la topoisomérase I (TopoI), l'enzyme responsable du déroulement de l'ADN superenroulé pour faciliter divers processus tels que la réplication, la transcription et la recombinaison. La TopoI permet de réduire la torsion de l'ADN en créant une cassure simple brin qui sera religaturée après le passage du second brin à travers la cassure. La CPT se lie de manière covalente au complexe ADN-TopoI et empêche la fermeture de la cassure simple brin, augmentant ainsi le nombre de cassures dans le génome. La CPT est une drogue spécifique des cellules en phase S. A faible concentration, la CPT induit principalement un ralentissement des fourches de réplication ainsi que des réversions de fourches. A forte concentration, la CPT induit un blocage de la réplication de l'ADN mais aussi la formation de cassures double brin dues à la collision des fourches de réplication avec le complexe ADN-TopoI (Vesela et al., 2017).

e) Rôle de la réplication dans le processus d'oncogénèse

Malgré la grande fidélité des ADN polymérase répliquatives, des mutations peuvent tout de même survenir au cours de la réplication de l'ADN. Si ces néo-mutations sont engendrées au niveau de gènes limitant la prolifération cellulaire, cela pourra conférer un avantage sélectif à la cellule mutée pouvant conduire à la mise en place d'un processus oncogénique. Une fois ce mécanisme établi, l'accumulation de nouvelles mutations va permettre à la cellule tumorale d'acquérir de nouvelles capacités, comme par exemple une prolifération incontrôlée et la résistance aux traitements anti-cancéreux. La réplication de l'ADN joue donc un rôle

fondamental pour le maintien de la stabilité génétique, c'est pourquoi ce mécanisme doit être finement régulé à la fois de manière spatiale et temporelle.

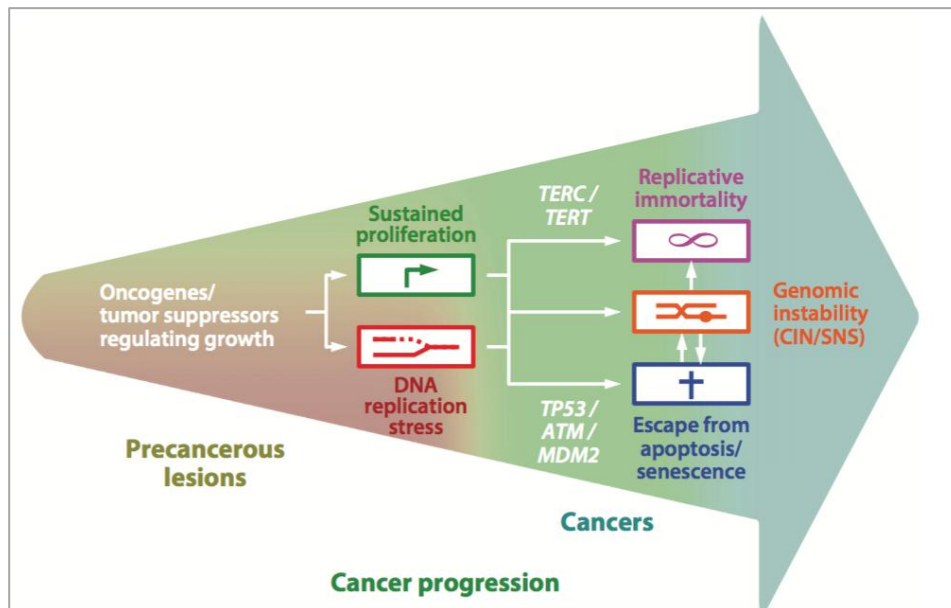


Figure 15 : Modèle de développement tumoral dû au stress réplcatif (Macheret and Halazonetis, 2015). Au cours de la réplication, l'accumulation de mutations dans des gènes régulant la croissance cellulaire conduit à l'activation de deux voies majeures de l'oncogénèse : la prolifération anarchique des cellules et l'induction d'un stress réplcatif.

4. Régulation spatio-temporelle de la réplication

a) Régulation spatiale

La réplication de l'ADN commence au niveau d'origines de réplication. Il en existe deux types, les origines constitutives, activées à chaque cycle cellulaire, et les origines dormantes qui vont pouvoir être activées en cas de diminution de la vitesse de la fourche de réplication, par exemple en cas de stress réplcatif, pour permettre d'assurer la fin de la réplication avant l'entrée en mitose. Dans un cycle cellulaire, toutes les origines ne vont pas être activées et chaque origine ne va être activée qu'une seule fois par cycle pour éviter de la re-réplication (Parker et al., 2017). Chez la levure, ces origines sont des séquences consensus riches en AT. Chez les mammifères, il n'existe pas de séquence consensus de ces origines mais elles partagent des caractéristiques structurales communes. En effet, les origines de réplication sont enrichies en guanine ainsi qu'en îlots CpG, favorisant la formation de structures G-quadruplexes et la structure secondaire de l'ADN au niveau de l'origine de réplication serait nécessaire à sa reconnaissance par le complexe ORC (Miotto et al., 2016; Parker et al., 2017; Warner et al., 2017).

De plus, la molécule d'ADN est organisée sous forme de boucles dans le noyau. Ainsi, les origines de réplication se regroupent dans des foyers de réplication où sont concentrés les réactifs nécessaires à l'initiation de la réplication de l'ADN (Fragkos et al., 2015).

b) Régulation temporelle

Le programme temporel de réplication est un niveau supérieur de régulation épigénétique orchestrant l'ordre dans lequel les origines de réplication vont s'activer tout au long de la phase S.

(1) Méthodes d'analyses

(a) *Incorporation d'analogues de thymidine*

Il est possible de suivre la réplication de l'ADN en ajoutant dans le milieu de culture des cellules un analogue de thymidine tel que l'EdU, l'IdU ou le BrdU capables de s'incorporer dans l'ADN à la place d'un T au cours de la réplication et qui vont pouvoir être marqués par immunofluorescence. Seules les cellules en phase S sont marquées. En fonction de l'état d'avancement de la cellule dans la phase S au moment où l'analogue a été ajouté, il est possible de distinguer si la cellule se trouve en début, en milieu ou en fin de phase S. En effet, chaque sous-phase de la phase S présente un pattern de réplication distinct. Au début de la phase S, on observe un marquage diffus dans tout le noyau mais exclu des nucléoles. Au milieu, les contours du noyau ainsi que les nucléoles sont marqués. En fin de phase S, on n'observe plus que très peu de foyers marqués de manière très intense et majoritairement en périphérie du noyau. Les mêmes résultats sont obtenus après un marquage par immunofluorescence du facteur de processivité PCNA. Cependant, cette technique ne permet d'observer que la réplication globale de la cellule sans permettre de distinguer précisément les régions répliquées.

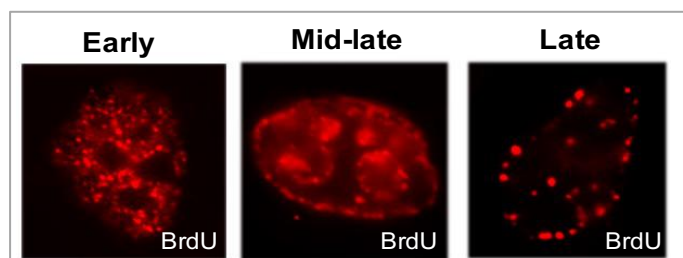


Figure 16: Pattern de réplication de l'ADN par incorporation d'analogue de thymidine dans des cellules humaines.

(b) Timing de réplication à l'échelle du génome

Cette technique permet de comparer entre elles les régions répliquées en début et en fin de phase S. Elle est aussi basée sur une incorporation d'analogue de la thymidine, le BrdU, suivie d'un tri cellulaire par cytométrie en flux basé sur la quantité d'ADN grâce à un marquage à l'iodure de propidium (IP) permettant de séparer les cellules en début (une copie d'ADN) et fin de phase S (deux copies d'ADN). Après immunoprécipitation des brins d'ADN néosynthétisés grâce à un anticorps anti-BrdU dans ces deux fractions, l'ADN est marqué par des fluorochromes de couleurs différentes et amplifié par « Whole Genome Amplification » (WGA). Enfin, les deux fractions sont hybridées sur une puce à ADN couvrant l'intégralité du génome (Repli-chip) ou peuvent directement être séquencées (Repli-seq).

Dans le cas de la puce à ADN, le ratio des intensités relatives des signaux des deux fractions permet de déterminer si la région est répliquée de manière précoce ou tardive. Par exemple, si la fraction de début de phase S est marquée en rouge et la fraction de fin de phase S en vert comme dans l'exemple ci-dessous, pour une région donnée, le ratio rouge/vert sera positif si la région est répliquée de manière précoce tandis qu'il sera négatif si la région est répliquée de manière tardive. Cette technique permet d'obtenir des courbes lissées du timing de réplication pour l'intégralité du génome.

Dans le cas du séquençage, le ratio de la couverture de la région entre les deux fractions permet de déterminer le programme de réplication de la région. De plus, le Repli-seq permet d'éviter les limitations d'espèces des puces à ADN disponibles et possède une meilleure sensibilité dans le cas d'études spécifiques de séquences telles que les études de polymorphismes simples (SNP). Cependant les résultats obtenus par Repli-seq ne donnent pas des courbes lissées comme celles obtenues par Repli-chip mais des courbes « en dents de scie » (Marchal et al., 2018).

Dans les deux cas, le nombre de fractions de cellules en phase S triées par FACS peut varier de 2 à 8 jusqu'à présent. L'augmentation du nombre de fractions n'apporte que peu pour l'augmentation de la résolution du profil de réplication mais elle peut donner des informations sur la synchronicité d'une population cellulaire au cours de la phase S qui serait indétectable car moyennée avec uniquement deux fractions. Cependant, les résultats obtenus pour 2 à 6 fractions par les deux techniques donnent sensiblement les mêmes profils après lissage et normalisation.

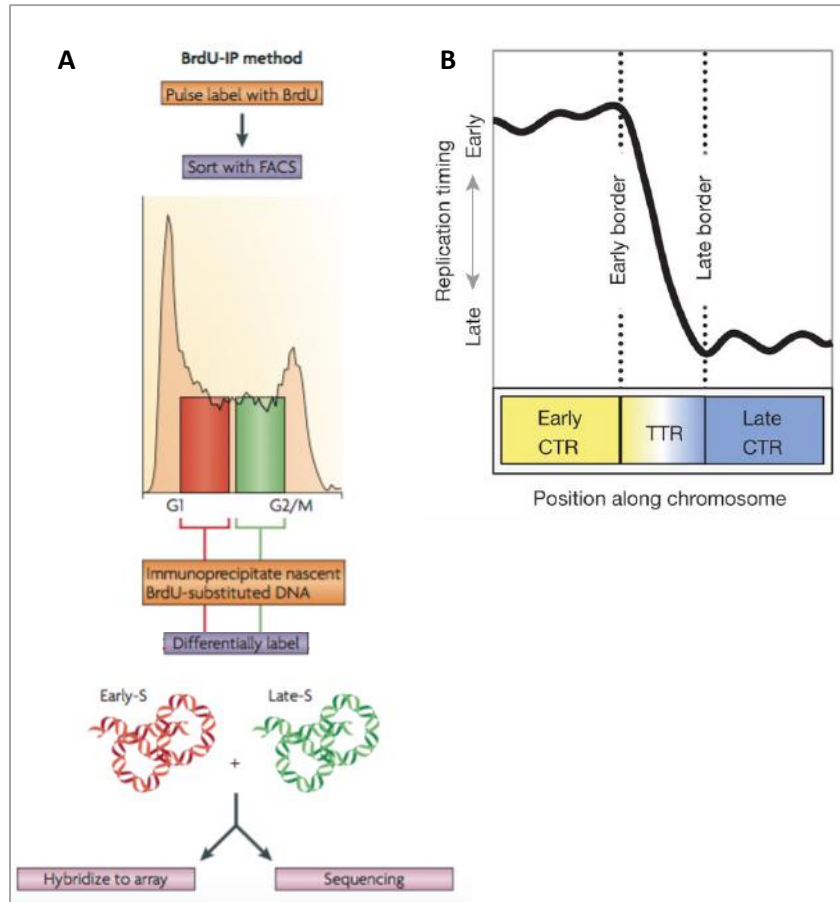


Figure 17: Analyse du timing de répllication (Gilbert, 2010). A) Résumé du protocole de Repli-chip. Les régions répliquées en début de phase S sont marquées en rouge et les régions répliquées en fin de phase S en vert. B) Exemple de profil de répllication. En ordonnées : \log_2 du ratio rouge/vert. En abscisses : Marche sur le chromosome. On observe des domaines de répllication constants (CTR) et des domaines de transitions (TTR).

(2) Corrélation entre la programme temporel de la répllication, la séquence d'ADN et la transcription

Les régions d'ADN répliquées en phase S précoce sont riches en GC ainsi qu'en gènes transcrits. A l'inverse, les régions répliquées en phase S tardive sont pauvres en gènes mais riches en éléments répétés tels que les LINEs et sont très faiblement transcrites (Hiratani et al., 2009; MacAlpine et al., 2004; Rhind and Gilbert, 2013; Schmegner et al., 2007). Cependant, certaines régions d'ADN pouvant changer de moment de répllication au cours de la différenciation cellulaire, la séquence primaire d'ADN n'est pas suffisante pour expliquer le moment de répllication (Hiratani et al., 2010; Ryba et al., 2011).

(3) Corrélation entre le paysage épigénétique et le programme temporel de la réplication

Les régions répliquées de manière précoce sont transcriptionnellement actives et donc associées avec des marques de chromatine ouverte et activatrices de la transcription telles que H3K4me3 ainsi que des marques d'acétylation. A l'inverse, les régions répliquées de manière tardive sont associées à des marques de chromatine fermée et d'inhibition de la transcription telles que H3K9me3 (Hiratani et al., 2010; Ryba et al., 2011).

(4) Influence du programme temporel de la réplication sur le paysage mutationnel

Au cours de l'évolution du génome humain, l'accumulation de mutations s'est faite majoritairement dans les régions répliquées tardivement au cours de la phase S. En effet, une étude sur la divergence humains-primates et sur le polymorphisme humain a révélé une augmentation du taux de mutation de respectivement 22% et 53% dans des régions répliquées de manière tardive (Stamatoyannopoulos et al., 2009). De manière intéressante, chez l'Homme, les gènes impliqués dans le développement, la différenciation et la prolifération cellulaire sont répliqués de manière tardive en phase S (Tomkova et al., 2018). Cela suggère que l'augmentation du taux de mutation dans ces régions aurait un coût évolutif important dans le but de réprimer ces gènes potentiellement dangereux.

Aussi, il a été montré une augmentation du taux de mutation dans des régions répliquées tardivement au cours de l'évolution tumorale (Donley and Thayer, 2013; Sima and Gilbert, 2014). En effet, cette accumulation a été montrée dans le cancer du poumon muté EGFR au cours du développement de la résistance aux traitements (Jia et al., 2013). Ces résultats ont aussi été observés en comparant une lignée germinale humaine et plusieurs lignées cellulaires cancéreuses (Liu et al., 2013; Woo and Li, 2012). A l'inverse, les régions d'euchromatine répliquées en début de phase S sont anti-corrélées avec la charge mutationnelle dans les cellules cancéreuses (Du et al., 2019). De plus, les régions répliquées tardivement au cours de la phase S sont plus souvent le site de réarrangements chromosomiques entraînant des variants du nombre de copies de gènes (CNV) (Miotto et al., 2016).

(5) Les régions de transition temporelle

Dans les cellules de mammifères, les régions de transition temporelle (TTRs) représentent des régions de plusieurs centaines de kilobases séparant les régions de timing

constant (CTRs) répliquées en début et en fin de phase S. Ces régions sont caractérisées par une très faible densité en origines de réplication et une réplication majoritairement unidirectionnelle à partir d'une origine activée en phase S précoce (Guan et al., 2009). Récemment, il a été mis en évidence la possibilité d'une activation séquentielle d'origines au sein des TTRs en cas de diminution de la vitesse de la fourche (Guilbaud et al., 2011). Les TTRs sont donc répliquées de plus en plus tardivement des régions précoces vers les régions tardives flanquantes. A l'inverse des CTRs dont le timing de réplication est relativement uniforme d'une cellule à l'autre, les TTRs sont les régions dont le timing de réplication peut changer au cours du développement ou de la différenciation cellulaire. Ces régions sont majoritairement constituées d'hétérochromatine facultative portant des marques épigénétiques bivalentes à la fois d'euchromatine et d'hétérochromatine telles que H3K4me3, H3K27me3 et la protéine HP1 (Hiratani et al., 2010; Rhind and Gilbert, 2013; Ryba et al., 2011).

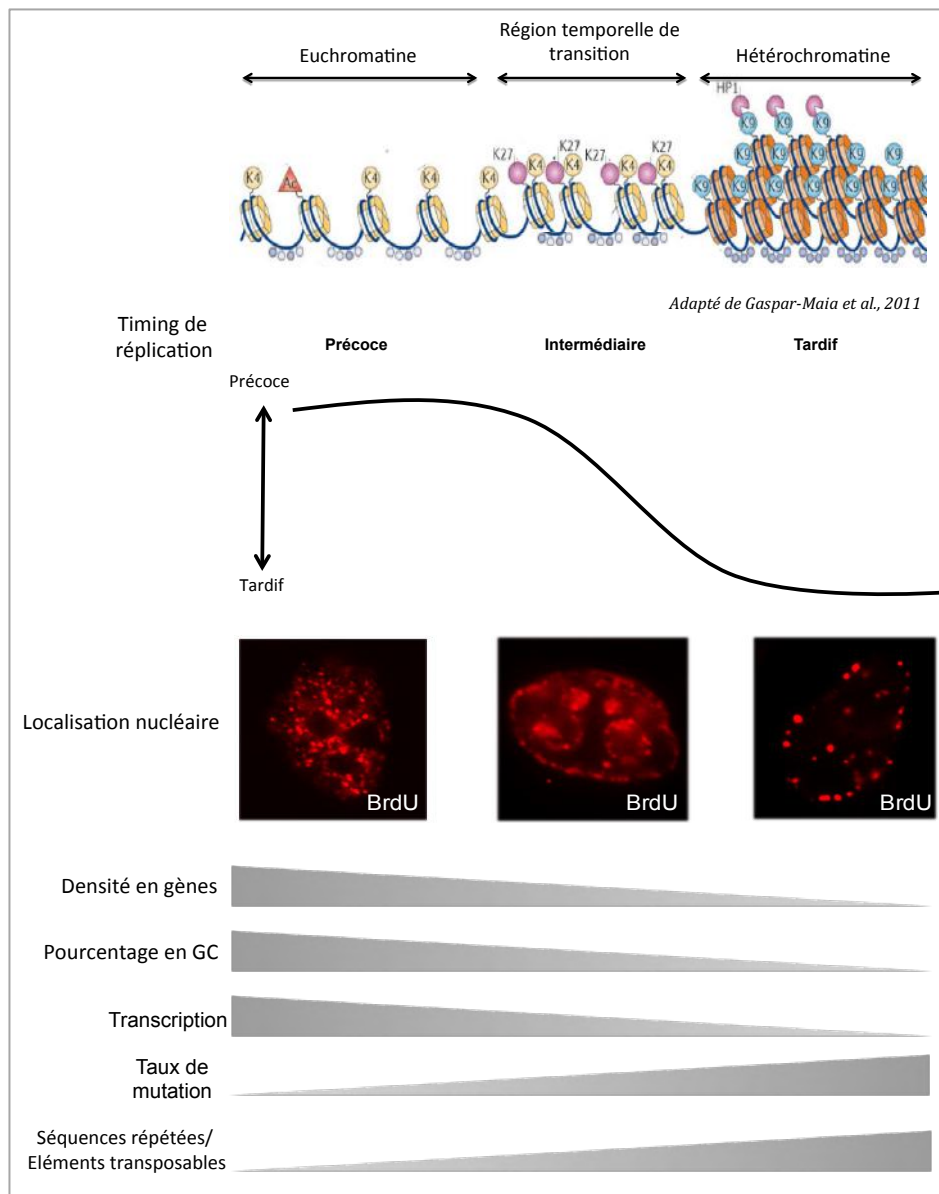


Figure 18: Corrélations entre timing de réplication, paysage génétique et épigénétique.

Étant donné que l'organisation chromatinienne ainsi que la nature de la séquence nucléotidique sont extrêmement différentes dans l'euchromatine vs l'hétérochromatine, il est possible que la composition du réplisome change au cours de la phase S dans le but d'assister les polymérase répliquatives conventionnelles pour répliquer les différents types de régions génomiques rencontrées. Par exemple, la présence naturelle dans le génome de structures secondaires de l'ADN et de séquences d'ADN répétées représente un obstacle à la progression des fourches de réplication, ces régions sont appelées des régions difficiles à répliquer. Notre hypothèse est que la distribution spatiale de l'apparition des mutations dans le génome reflète l'implication temporelle de polymérase spécialisées translésionnelles au cours de la réplication non perturbée des cellules mammifères dont la fidélité est réduite par rapport aux polymérase répliquatives. Cette hypothèse représente le point de départ de mon projet de thèse et sera discutée tout au long de ce manuscrit.

C. Problèmes au cours de la réplication : causes et conséquences sur l'intégrité du génome

1. Les différents obstacles de la fourche de réplication

L'ADN est constamment soumis à des agressions pouvant nuire à l'intégrité du génome en causant des lésions représentant des obstacles à la progression des fourches de réplication. Ces agressions peuvent venir de l'environnement extérieur, telles que les rayons ionisants, la fumée de cigarette... ou directement du métabolisme intracellulaire, telles que les espèces réactives de l'oxygène (ROS). Si ces lésions ne sont pas correctement prises en charge au cours de la réplication de l'ADN, elles peuvent conduire à l'introduction de mutations ponctuelles, voire à des remaniements chromosomiques. Ces agents dommageables de l'ADN déclenchent un stress génotoxique défini comme induisant une modification de l'information génétique par l'introduction de mutations pouvant conduire à l'apparition de maladies. L'établissement de ces lésions représente donc un grand pas vers le processus de carcinogenèse (Garaycochea et al., 2018).

De plus, l'ADN n'est pas sous forme linéaire dans le noyau mais il est capable d'adopter des structures secondaires, voire tertiaires, qui vont aussi former des barrières naturelles à la progression des fourches de réplication. Ces structures ne dépendent d'aucun facteur extérieur mais sont simplement dues à la séquence primaire de nucléotides.

a) Les lésions endogènes

Le métabolisme cellulaire et en particulier les NADPH oxydases et la respiration mitochondriale produisent des radicaux libres appelés espèces réactives de l'oxygène (ROS). Les ROS comprennent l'anion superoxide ($O_2^{\cdot-}$), le peroxyde d'hydrogène (H_2O_2), et le radical hydroxyle (OH^{\cdot}). Un niveau élevé intracellulaire de ROS est appelé stress oxydant. Les ROS possèdent des propriétés chimiques inhérentes à leur structure qui leur permettent de réagir avec différentes cibles protéiques ou l'ADN. En cas de stress oxydant, les ROS peuvent induire des dommages de l'ADN tels que la 8-oxo-7,8-dihydro-guanine, des pontages ADN-protéines et des cassures double brin, chacun ayant un impact biologique propre. D'autre part, les ROS ont un rôle dans l'activation de voies pro-oncogéniques impliquées dans la progression anti-tumorale, l'angiogenèse et la survie cellulaire (Reczek and Chandel, 2017; Zuo et al., 2015).

b) Les lésions exogènes

(1) Les agents physiques

Les agents physiques créent des lésions directement sur les bases de l'ADN.

(a) Les rayons ultraviolets

La source principale de rayons ultraviolets (UVs) sur Terre est le soleil et peuvent être classés en 3 catégories selon leur longueur d'onde : les UVA de 320 à 400nm, les UVB de 290 à 320 nm et les UVC inférieurs à 290 nm. Les rayons UVs sont responsables de la majorité des cancers de la peau. Les lésions majoritaires dues aux UVs sont les dimères de pyrimidine cyclobutane (CPDs) et les photoproduits pyrimidine(6-4)pyrimidone (6-4PPs). Les 6-4PPs peuvent être convertis en leur isomère Dewar après exposition aux UVA ou UVB. Toutes ces lésions sont formées par une réaction photochimique dont l'efficacité dépend de la longueur d'onde des UVs. La charge maximale de CPDs et 6-4PPs apparaît autour de 260nm (Ravanat and Douki, 2016). Ces types de lésions vont être pris en charge au cours de la réplication et induire une substitution de base C→T et plus rarement des substitutions de base en tandem CC→TT. Ces mutations représentent la signature mutationnelle des UVs (Alexandrov et al., 2013). Les CPDs peuvent être pris en charge de manière fidèle par Polη qui est capable d'introduire un A en face de T lésé. Les patients porteurs d'une mutation du gène *POLH* codant pour Polη développent le syndrome Xeroderma Pigmentosum Variant (XPV) caractérisé par une forte sensibilité au soleil et une prédisposition au cancer de la peau (Stary et al., 2003). Secondairement, les rayons UV peuvent induire un stress oxydant avec production de ROS (Cadet and Davies, 2017).

(b) Les rayonnements ionisants

Les rayonnements ionisants (RI) proviennent naturellement de rayons cosmiques, de la radioactivité naturelle ou créée par l'Homme (centrales nucléaires) ainsi que de certains traitements médicaux tels que les Rayons X des radiographies. Les RI induisent principalement des cassures simple ou double brin de l'ADN. Secondairement, les RI peuvent induire une radiolyse de l'eau conduisant à un stress oxydant (Ravanat and Douki, 2016).

(2) Les agents chimiques

A l'inverse des agents physiques, les agents chimiques ont un effet indirect sur l'ADN. En effet, ils ont généralement besoin d'être métabolisés par la cellule et les sous-produits de cette métabolisation induisent des lésions de l'ADN. Il existe une grande quantité d'agents chimiques mutateurs de l'ADN. Nous pouvons citer par exemple le benzo[*a*]pyrène contenu dans la fumée de cigarette. Aussi, les agents chimiques utilisés en chimiothérapie tels que le méthanesulfonate et le temozolomide induisent des lésions conduisant à l'apparition de mutations dans les cellules cancéreuses inhibant leur prolifération et les tuant (Kucab et al., 2019). D'autres agents chimiques tels que le formaldéhyde induisent des pontages ADN-ADN et ADN-protéines qui empêchent la progression des fourches de réplication (Garaycochea et al., 2018).

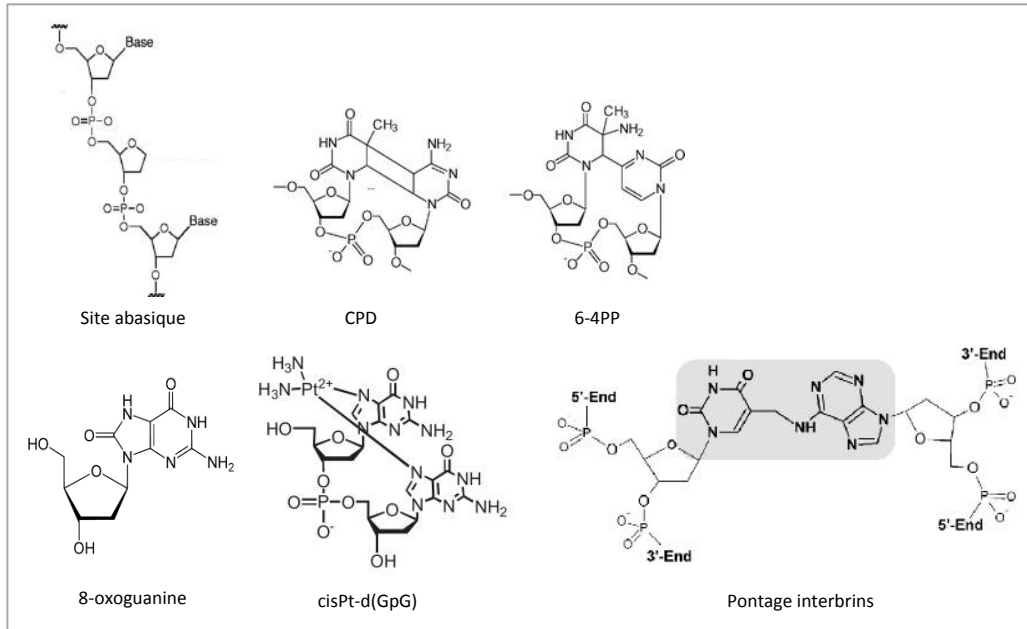


Figure 19 : Structure des principales lésions de l'ADN. Le site abasique représente l'absence de la base azotée sur un nucléotide. C'est une lésion commune pouvant être spontanée ou induite après alkylation. C'est aussi un intermédiaire de réparation de l'ADN. Les CPD et 6-4PP sont les lésions majoritaires après exposition aux rayons UV sont induits par une réaction photochimique. La 8-oxo-guanine résulte de l'oxydation d'une guanine pouvant être induite après un stress oxydant par les ROS. Le *cisPt-dGpG* est un pontage intrabrin entre deux guanines adjacentes sur le même brin d'ADN produit par le cisplatine. Les pontages interbrins résultent de la formation d'une liaison covalente entre deux bases sur des brins d'ADN complémentaires. Ils peuvent être produits par plusieurs agents tels que le cisplatine, le formaldéhyde, le psoralène...

c) Les barrières naturelles de la fourche de réplication et régions difficiles à répliquer

(1) Les G-quadruplexes

Les séquences d'ADN simple brin riches en guanine peuvent se replier sur elles-mêmes pour adopter une structure à 4 brins appelée G-quadruplex (G4) provenant de l'empilement d'au minimum deux quartets de G. Il existe une grande variabilité structurelle des G4s selon l'orientation des brins, le nombre de boucles et de quartets. Ces structures sont majoritairement formées au moment de la phase S et sont retrouvées au niveau des promoteurs de gènes et des télomères. Elles jouent un rôle important dans la régulation génique ainsi que dans la maintenance des télomères. Aussi, en l'absence d'une hélicase capable de résoudre cette structure, la présence de G4s empêche la progression des ADN polymérases réplcatives et entraîne un blocage des fourches de réplication pouvant conduire à une instabilité génétique. Chez la souris, il a été montré que la protéine RTEL1 est capable de résoudre les G4s télomériques et maintient ainsi leur intégrité. De plus, les polymérases Rev1 et Primpol ont été impliquées dans le maintien épigénétique des G4s et leur inhibition peut donc activer ou inhiber les gènes en fonction de la localisation du G4. Des ligands spécifiques des G4s sont capables de les stabiliser et ainsi induire des cassures chromosomiques par blocage des fourches. Cette instabilité chromosomique peut conduire à des insertions ou délétions de nucléotides à cette position. Finalement, les G4s représentent des régions d'instabilité génomique qui sont enrichies dans les cellules immortalisées et cancéreuses (Hänsel-Hertsch et al., 2017).

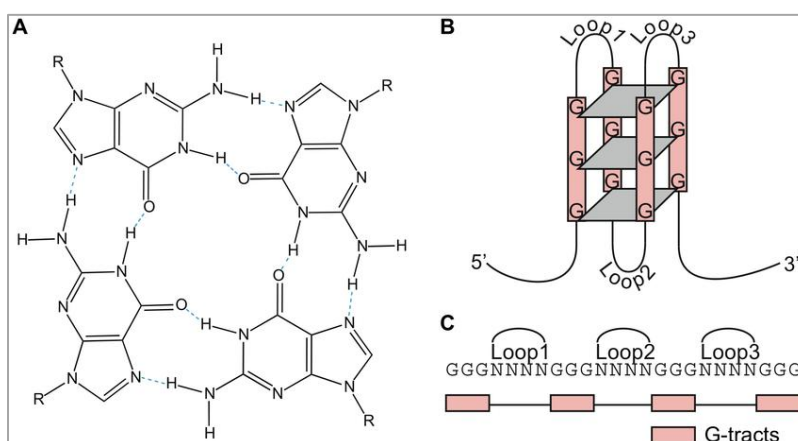


Figure 20: Structure d'un G-quadruplex (Capra et al., 2010) A) Structure d'un quartet de G où les 4 guanines sont reliées par des liaisons hydrogènes. B) Structure d'un G4 composé de 3 quartets de G. Chaque quartet est représenté comme un plateau. C) Séquence primaire d'un G4 composé de 4 stretches de G séparés par 3 boucles.

(2) Les R-loops

Au cours de la transcription, le brin d'ARN naissant est capable de s'apparier avec le brin d'ADN matrice, laissant ainsi le brin d'ADN non-matrice sous forme simple brin. Cette structure à trois brins est appelée une R-loop. Les R-loops ont été impliquées dans divers processus biologiques tels que la régulation des gènes, la réparation de ADN et la terminaison de la transcription. Cependant, la génération d'un simple brin d'ADN dans la R-loop représente un risque pour la stabilité génétique. En effet, le simple brin est plus sensible à des agents dénaturants de l'ADN ainsi qu'aux nucléases pouvant induire des cassures simple brin. Aussi, la désamination spontanée des Cytosines vers un Uracile sur ce simple brin serait responsable de cassures chromosomiques (Crossley et al., 2019; Skourti-Stathaki and Proudfoot, 2014).

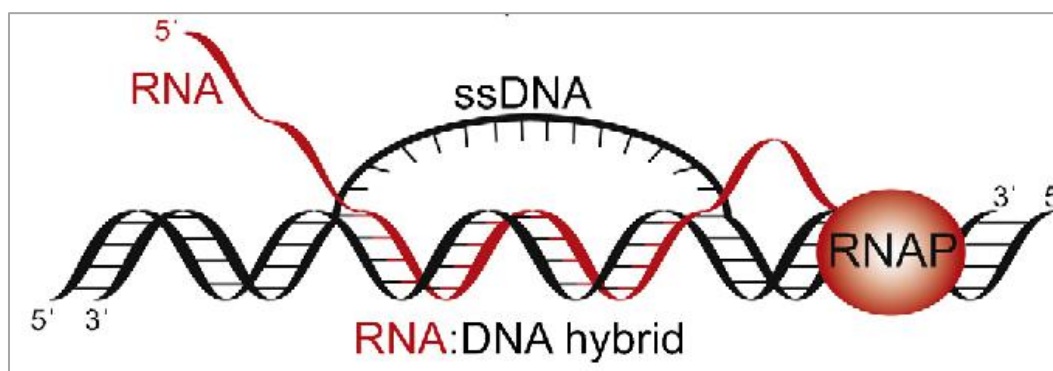


Figure 21 : Représentation schématique de la structure d'une R-loop (Hamperl and Cimprich, 2014). L'ARN en cours de polymérisation par l'ARN polymérase, en rouge, s'hybride avec le brin d'ADN matrice en créant un hybride ADN:ARN et un ADN simple brin.

(3) Les sites fragiles communs

Les sites fragiles communs (CFSs) ont été décrits pour la première fois en 1984 comme étant des sites privilégiés de cassures chromosomiques après un faible stress réplicatif sur des chromosomes en métaphase. Lorsqu'un de ces sites est le lieu d'une cassure chromosomique, on dit que le site fragile commun est exprimé. Certains CFSs sont communs à tous les types cellulaires mais exprimés de manière plus ou moins forte tels que FRA16B et WWOX. Cependant, il existe des CFSs exprimés spécifiquement dans certains types cellulaires, par exemple, chez l'Homme, le site FRA3B est fortement exprimé dans les lymphocytes mais ne l'est pas dans les fibroblastes. Les CFSs sont localisés dans des grands gènes, transcrits et riches en A-T et répétitions microsatellites capables de former des structures secondaires pouvant empêcher la progression des fourches de réplication. Ce sont des sites de forte instabilité génétique répliqués en phase S tardive, avec une faible concentration en origines de réplication. Si la réplication de ces sites n'est pas complète en fin de mitose, par exemple lorsque la réplication est ralentie en

présence d'un stress réplicatif, l'ADN de ces régions ne sera pas entièrement répliqué et conduira à une cassure double brin en mitose. Les CFSs sont souvent associés à des sites de remaniements chromosomiques dans les cancers (Glover et al., 2017).

(4) Les conflits transcription-réplication

La réplication et la transcription de l'ADN se font de manière concomitante à partir de la même matrice ADN en provoquant parfois la rencontre des deux machineries. La machinerie de transcription représente donc un obstacle à la progression des fourches de réplication. De plus, l'activité transcriptionnelle génère des obstacles à la progression des fourches tels que des hybrides ADN-ARN, une torsion de l'ADN et des structures non-B de l'ADN (G4s). Toutes ces barrières vont induire un blocage de la réplication qui peut conduire à une instabilité génétique avec des cassures chromosomiques si le conflit n'est pas résolu (García-Muse and Aguilera, 2016).

(5) Modification de la composition du réplisome au cours de la phase S

Comme expliqué précédemment, l'organisation de la chromatine ainsi que la séquence nucléotidique est extrêmement différente entre l'euchromatine et l'hétérochromatine. En effet, les régions hétérochromatiniennes telles que les régions péri-centromériques ou les télomères sont très compactées et peuvent adopter des structures non B de l'ADN représentant des barrières naturelles pour la progression des fourches de réplication. Nous pouvons donc penser que la composition du réplisome peut changer au cours de la phase S, en fonction de la présence de lésion ou des structures d'ADN rencontrées, dans pour aider les polymérase répliquatives peu efficaces dans ces régions à répliquer l'intégralité du génome. Dans ce cas, les polymérase translésionnelles prendraient la place des polymérase répliquatives au niveau de ces régions difficiles à répliquer et seraient donc impliquées dans la réplication non perturbée de l'ADN.

2. Mécanismes de signalisation et prise en charge des lésions de l'ADN

a) Signalisation des lésions au cours de la phase S

(1) Mécanismes généraux des points de contrôle du cycle cellulaire

Il existe tout au long du cycle cellulaire des points de contrôles permettant de vérifier l'intégrité du génome avant de passer à la phase suivante.

Des capteurs sont capables de détecter et de signaler s'il y a un problème au cours de la phase et dans ce cas retarder l'entrée de la cellule dans la phase suivante du cycle jusqu'à ce que le problème soit réparé. Quatre points de contrôles principaux régulent les transitions du cycle cellulaire G1/S, S/G2, G2/M et au cours de la mitose la transition métaphase/anaphase.

En résumé, le point de contrôle G1/S vérifie que tout est en place dans la cellule pour permettre la réplication de l'ADN ; le point de contrôle S/G2 vérifie que tout l'ADN a été répliqué et qu'il ne persiste pas de lésion ; et le point de contrôle métaphase/anaphase vérifie que tous les chromosomes sont bien alignés sur la plaque métaphasique. Ces points de contrôle impliquent trois catégories de protéines :

- Des capteurs reconnaissent l'anomalie, la signalent et activent une cascade de signalisation.
- Des transducteurs, en général des kinases, amplifient le message en phosphorylant plusieurs protéines jusqu'à la cible finale.
- Des effecteurs, les cibles finales, résolvent le problème.

Je ne décrirai ici que le point de contrôle S/G2 car il relève du cadre de ma thèse.

(2) Le point de contrôle S/G2

La protéine « Ataxia telangiectasia and Rad3-related » ou ATR est une kinase présente en phase S activée en cas de stress répliatif et/ou de dommages de l'ADN conduisant à l'arrêt du cycle cellulaire en fin de phase S ou en début de phase G2. Cet arrêt permet de donner plus de temps à la cellule pour finir la réplication de son génome ou réparer les lésions et permet donc le maintien de l'intégrité génomique.

ATR et son partenaire ATRIP (« ATR-interacting protein »), nécessaire à l'activité d'ATR, sont recrutés au niveau des fourches de réplication bloquées grâce à une interaction directe avec la protéine RPA (« Replication Protein A) qui recouvre l'ADN simple brin. La kinase ATR peut être activée par TOPBP1 ainsi qu'ETAA1 recrutés de manière indépendante respectivement aux jonctions double/simple brin ou sur de l'ADN simple brin recouvert de RPA. Une fois activée, ATR phosphoryle plusieurs protéines effectrices. Par exemple, la phosphorylation de Chk1

enclenche une cascade de signalisation de plusieurs effecteurs aboutissant à l'arrêt du cycle cellulaire avec l'inhibition des origines de réplication alentour, l'activation d'origines dormantes et la diminution de la vitesse de la fourche de réplication. ATR va aussi phosphoryler plusieurs protéines du replisome ainsi que des facteurs de réparation permettant la stabilisation et la réparation des fourches de réplication bloquées (Saldivar et al., 2017; Shiotani and Zou, 2016; Yazinski and Zou, 2016).

Récemment, il a été montré qu'ATR jouerait aussi un rôle dans le point de contrôle S/G2 non perturbé et permettrait de contrôler la fin de la réplication de l'ADN afin de coordonner la réplication et la mitose. ATR serait activée par ETAA1 au niveau de l'ADN simple brin généré normalement au cours de la réplication, inhibant les protéines CDK1 et FOXM1. La perte d'activité d'ATR lorsque la réplication de l'ADN est complète entraînerait alors une activation de CDK1 et FOXM1 donnant le signal d'entrée en phase G2 et mitose (Saldivar et al., 2018).

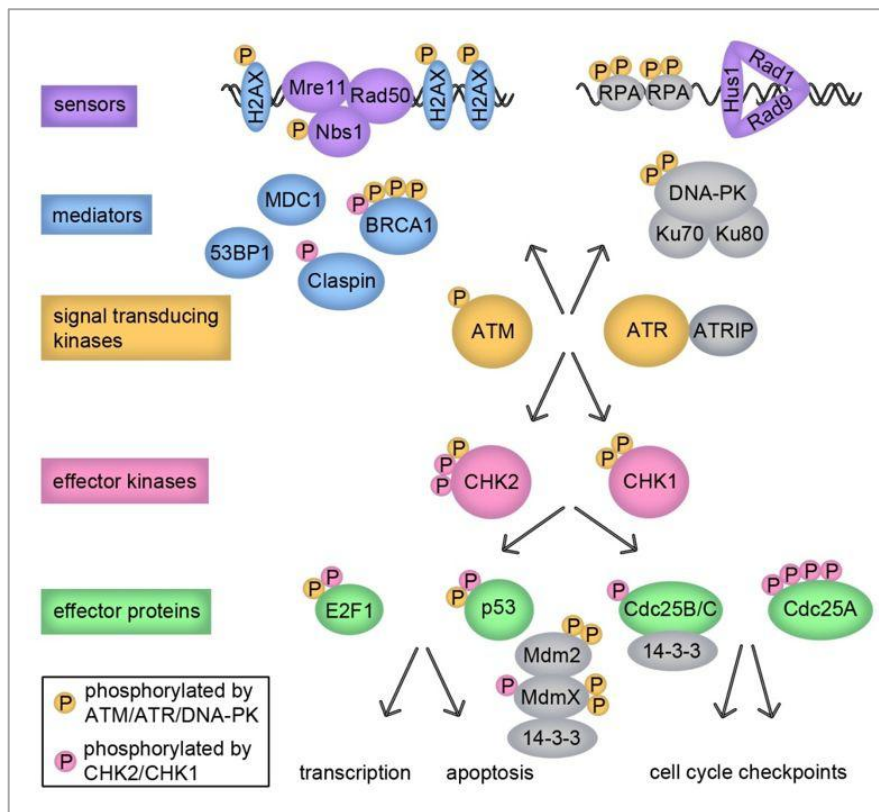


Figure 22 : Vue simplifiée de la réponse aux dommages de l'ADN (Freeman and Monteiro, 2010). Les dommages de l'ADN induisent une cascade de signalisation impliquant de multiples kinases pouvant conduire à la transcription, l'apoptose ou l'arrêt du cycle cellulaire.

b) La synthèse translésionnelle

Les dommages de l'ADN peuvent être induits par différents facteurs environnementaux ou métaboliques qui vont perturber la réplication. La présence de barrières de la fourche de

réplication telles que l'ADN structuré ou la présence de complexes ADN-protéines peuvent aussi entraîner un blocage de la réplication. Ces lésions/barrières représentent donc des dangers potentiels pour la cellule et doivent être réparés. Si elles sont rencontrées au cours de la réplication, il existe des mécanismes de tolérance aux dommages permettant de continuer la réplication sans retirer la lésion. Le premier mécanisme de tolérance des dommages au cours de la réplication est le changement de matrice d'ADN (« Template switching ») non mutagène. Le brin naissant se découple de la matrice parentale utilisée pour sa réplication puis va utiliser le brin naissant sur la chromatide sœur comme matrice pour continuer son élongation. Ce mécanisme résulte en la formation de structures de pseudo double jonctions de Holliday qu'il est nécessaire de résoudre pour le maintien de la stabilité génétique. Ces structures sont alors prises en charge par des endonucléases structure-spécifiques telles que SLX4. De manière intéressante, une étude chez la levure a montré que l'absence de SLX4 entraîne une augmentation de 5 fois du taux de mutagenèse et cette augmentation est dépendante du second mécanisme de tolérance aux dommages qui va compenser cette perte, la synthèse TLS mutagène que je vais présenter ici, et plus particulièrement de Rev3/Pol ζ dont je parlerai en détail plus loin (Flott et al., 2007).

(1) Les ADN polymérases translésionnelles

L'ADN est constamment soumis à des stress génotoxiques capables d'endommager les bases. Les ADN polymérases translésionnelles sont capables de prendre en charge ces bases endommagées grâce à leur site catalytique plus large que chez les polymérases réplcatives et mettre une base en face de cette liaison, continuant ainsi la réplication de l'ADN au détriment de l'apparition d'une mutation. Ces dommages n'auraient pu être pris en charge par les polymérases réplcatives et auraient induit un blocage de la fourche de réplication (Jain et al., 2018; Yang and Gao, 2018).

Les ADN polymérases translésionnelles η , ι , κ et Rev1 appartiennent à la famille Y alors que Pol ζ appartient à la famille B comme les polymérases réplcatives. Chaque polymérase translésionnelle est spécialisée pour le franchissement d'un type de lésion mais n'y est pas restreinte. Pol ζ ne sera pas discutée dans cette partie mais sera étudiée en détail par la suite.

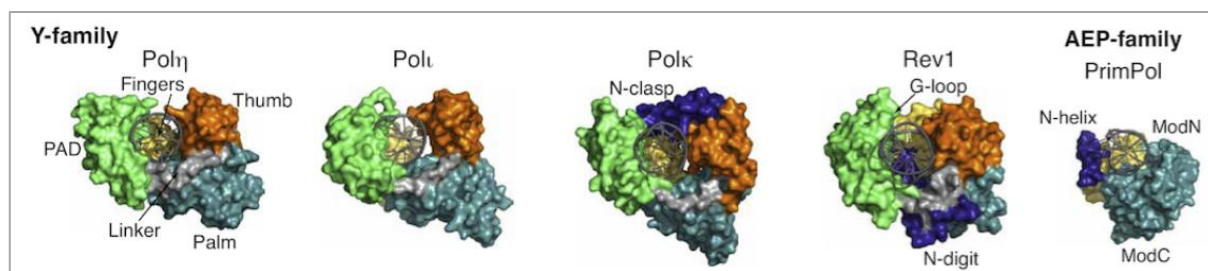


Figure 23: Structure par cristallographie des domaines catalytiques des polymérases translésionnelles (Jain et al., 2018). La structure de Polζ n'a pas encore été établie.

(a) *Polη*

Polη est principalement caractérisée pour son rôle dans le franchissement du principal type de lésion causé par les rayons UVs, les dimères de pyrimidine cyclobutane (CPDs), de manière fidèle (Stary et al., 2003). Ce rôle canonique de *Polη* est induit par la mono-ubiquitination de PCNA par son E3 ubiquitine-ligase RAD18 (Kannouche and Lehmann, 2006, 2004). *Polη* est aussi capable de bypasser des sites abasiques, des dommages oxydatifs ainsi que des adduits d'ADN créés par des composants dérivés du platine tel que le cisplatine et serait impliquée dans la résistance tumorale face à ces agents chimiothérapeutiques dérivés du platine (Choi et al., 2010; Srivastava et al., 2015; Vaisman et al., 2000; Zlatanou et al., 2011). Aussi, *Polη* est impliquée dans le mécanisme d'hypermutation somatique des gènes d'immunoglobulines où elle va induire des mutations A:T et augmenter ainsi le répertoire des anticorps (Delbos et al., 2007).

Malgré que cette polymérase soit entre 100 et 1000 fois plus mutagène que les polymérases répliquatives sur une matrice d'ADN non endommagée (McCulloch and Kunkel, 2008), des résultats récents du laboratoire semblent montrer qu'elle est capable de s'associer au brin naissant dans une partie des foyers de réplication et qu'elle voyage avec la fourche au cours de la réplication non perturbée (Despras et al., 2016).

Les mutations héréditaires du gène codant pour *Polη* chez l'homme sont responsables du syndrome Xeroderma Pigmentosum Variant (XPV), caractérisée par des défauts de réplication UV-induits et une hypermutabilité. Cette maladie génétique est donc caractérisée par une sensibilité accrue aux rayons UV et une forte prédisposition pour les cancers cutanés (Kannouche et al., 2001; Lin et al., 2006).

(b) *Polι*

Polι incorpore plus fidèlement des nucléotides en face d'une purine (A ou G) que d'une pyrimidine (C ou T) et permet donc la réplication à travers des adduits de purines tels que le

1,N⁶-ethenodeoxyadenosine et le N1-methyl-deoxyadenosine. De plus, notre équipe a montré que l'inhibition de Pol ι dans des cellules humaines entraîne une sensibilité accrue au stress oxydatif induit par le peroxyde d'hydrogène (H₂O₂) et au ménadione suggérant un rôle de Pol ι dans la protection des cellules face à ce type de dommages (Petta et al., 2008).

(c) *Polk*

Polk est spécialisé pour l'extension d'un primer après la réplication d'une lésion. En effet, le mécanisme de réplication translésionnelle nécessite parfois deux polymérases : une pour répliquer à travers la lésion puis une seconde pour étendre ce primer. Cependant, Polk est aussi capable de réaliser une synthèse translésionnelle en face d'une lésion sur l'azote 2 d'une guanine (N2-dG) induite, par exemple, par le benzopyrène (Vaisman and Woodgate, 2017).

(d) *Rev1*

Rev1 est une cytosine-transférerase qui incorpore uniquement un C en face d'un G et est donc capable de bypasser des adduits de guanine. Rev1 possède aussi des fonctions non-catalytiques et pourrait servir comme protéine d'assemblage dans des complexes multiprotéiques où elle servirait de plateforme pour le recrutement d'autres polymérases translésionnelles (Vaisman and Woodgate, 2017). Il a été montré que Rev1 est capable d'interagir de manière simultanée avec PCNA et Rev7 (composante de Pol ζ) (Acharya et al., 2005; Gallego-Sánchez et al., 2012; Hara et al., 2010; Kikuchi et al., 2012; Lemontt, 1971a; Murakumo et al., 2001). Aussi, un domaine d'interaction avec Rev1 a été récemment découvert dans POLD3 suggérant un rôle de Rev1 dans un mécanisme d'échange de polymérases au niveau du réplisome (Pustovalova et al., 2016).

(e) *Primpol*

Primpol est la dernière polymérase classifiée en tant que polymérase translésionnelle. Elle possède à la fois une activité primase et une activité de polymérase translésionnelle. Face à une lésion de l'ADN, cette double activité lui permet de soit répliquer l'ADN à travers la lésion soit contourner la lésion pour reprendre la synthèse d'ADN de novo en aval de la lésion. De plus, son activité primase utilise préférentiellement des dNTPs et conduit donc à la synthèse d'une amorce ADN contrairement aux autres primases, telles que Pol α , qui synthétisent des amorces ARN. Par son activité TLS, Primpol est capable de bypasser des lésions UV-induites ainsi que la 8-oxoguanine induite par un stress oxydatif (Rudd et al., 2014).

(2) Mécanisme de synthèse translésionnelle

La synthèse translésionnelle se fait en deux étapes impliquant généralement deux polymérase TLS différentes. La première étape d'insertion nécessite une polymérase TLS capable d'accommoder et d'utiliser une base endommagée pour l'incorporation directe d'un dNTP face à la lésion. Cette étape est généralement réalisée par une polymérase translésionnelle mutagène de la famille Y, qui va introduire une mutation en face de la lésion entraînant une distorsion de l'hélice d'ADN. Dans un deuxième temps, l'étape d'extension va recruter une seconde polymérase TLS capable de prendre en charge cette amorce distordue et va pouvoir l'étendre, fixant ainsi la mutation. Cette étape est généralement réalisée par la seule polymérase translésionnelle appartenant à la famille B, Pol ζ (Vaisman and Woodgate, 2017; Yang and Gao, 2018; Zhao and Washington, 2017).

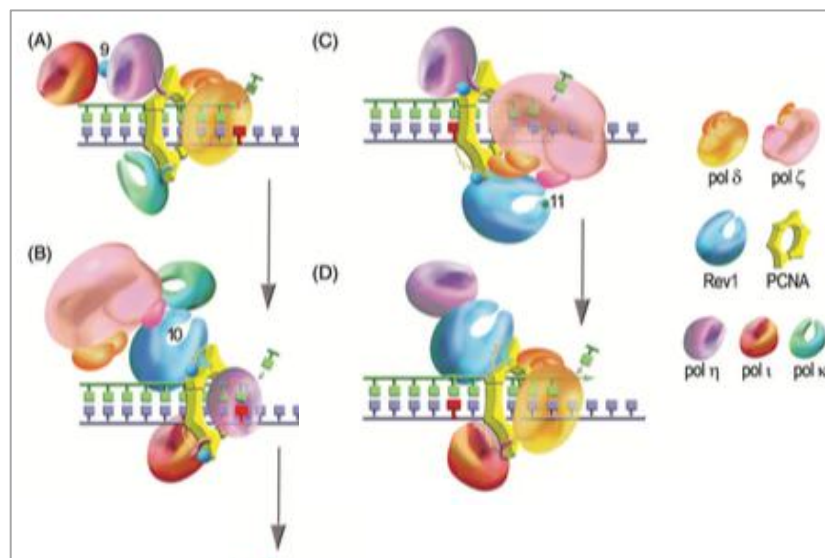


Figure 24: Modèle du changement de polymérase impliquant deux polymérase translésionnelles pour le franchissement d'une lésion (Vaisman and Woodgate, 2017) A) La présence d'une lésion sur l'ADN induit au cours de la répllication un blocage des polymérase répliquatives. B) L'ubiquitination de PCNA au site de blocage induit le recrutement d'une première polymérase TLS de la famille Y (ici Pol η) qui prend le relais et insère un nucléotide face à la lésion. C) Pol ζ étend l'amorce sur quelques nucléotides. D) La désubiquitination de PCNA induit le relargage de la polymérase TLS puis les polymérase répliquatives reprennent la répllication.

c) Prise en charge et réparation des lésions

Les lésions restantes après la répllication sont prises en charge par différents mécanismes de réparation en fonction de la lésion. Il existe cinq principaux mécanismes de réparation : la réparation par excision de base (BER), la réparation par excision de nucléotides (NER), la réparation des mésappariements de bases (MMR), la jonction d'extrémités non-homologues (NHEJ) ainsi que la recombinaison homologue (RH). Tous ces mécanismes, à

l'exception de la RH, entraînent l'excision de la lésion, la resynthèse d'ADN soit par les polymérase répliquatives (NER et MMR), soit par les polymérase TLS (BER), puis la ligature du fragment d'ADN néosynthétisé. La RH utilise la séquence du chromosome homologue de la région lésée ou cassée comme matrice pour sa répliquation. La RH et le NHEJ permettent la réparation des cassures double brin et des pontages interbrins tandis que les autres mécanismes prennent en charge les bases lésées et les pontages intrabrins. Ces mécanismes peuvent donc être une source de mutagenèse si la réparation n'est pas fidèle, par exemple par l'action des polymérase translésionnelles (Chatterjee and Walker, 2017).

3. Conséquences sur l'intégrité du génome

Toute perturbation de la dynamique de répliquation induit un stress répliquatif qui peut conduire à une instabilité génomique et au développement d'une maladie. Les sources de stress répliquatif sont variées telles qu'une mauvaise régulation des origines de répliquation, la présence de barrières de la fourche de répliquation ou une répliquation en mauvaises conditions métaboliques. Cependant ces perturbations peuvent induire les mêmes conséquences.

a) Instabilité génétique

Si les lésions de l'ADN ne sont pas correctement réparées, elles peuvent être tolérées par le mécanisme de synthèse translésionnelle où les polymérase TLS sont capables de continuer la synthèse de l'ADN à travers la lésion sans la réparer. Ces polymérase étant très mutagènes, ce mécanisme est responsable de l'apparition de mutations dans le génome, coût nécessaire pour la survie de la cellule. La synthèse TLS tient donc un rôle prépondérant dans la mutagenèse qui est à la fois un mécanisme nécessaire à l'évolution des génomes mais aussi dangereux pouvant conduire à l'apparition de mutations et de maladies si les régions génomiques touchées sont des régions codantes (Waters et al., 2009).

b) Instabilité chromosomique

L'un des marqueurs principaux du stress répliquatif est l'accumulation d'ADN simple brin aux fourches de répliquation et l'activation de la voie de signalisation ATR qui va s'assurer du redémarrage de la fourche de répliquation une fois que le stress a été supprimé. D'autres mécanismes alternatifs peuvent prendre le relais si la voie ATR n'est pas efficace tels que l'activation d'origines dormantes. Cependant, dans les régions pauvres en origines de répliquation

comme les zones de transitions ou les sites fragiles communs, le ralentissement de la fourche de réplication conduit à une réplication incomplète de l'ADN. Si les cellules arrivent toutefois en mitose tandis que la réplication est incomplète ou qu'il reste des intermédiaires de réparation de l'ADN non résolus, il se forme des ponts ultrafins d'ADN en anaphase créant un lien physique entre deux chromatides sœurs. La séparation des pools de chromosomes en anaphase va donc induire une tension mécanique sur ces chromatides sous-répliquées conduisant à des cassures chromosomiques. La cassure d'un pont ultrafin conduit à une ségrégation non équilibrée des chromosomes cassés dans les deux cellules filles et peut aussi conduire à des fusions ou des translocations chromosomiques. De plus, il a été montré une corrélation entre le stress réplicatif et la formation de centrosomes supplémentaires en mitose conduisant à une mitose multipolaire avec une séparation non équilibrée des chromosomes associée à des ponts d'ADN, une mauvaise condensation des chromosomes et un arrêt métaphasique prolongé (Gaillard et al., 2015; Jeggo et al., 2016; Magdalou et al., 2014; Zeman and Cimprich, 2014).

D. L'ADN polymérase zêta

1. Le complexe ADN polymérase ζ 4

Pol ζ est la seule polymérase TLS qui partage une structure similaire ainsi qu'une homologie de séquence avec les polymérases réplicatives et appartient donc à la famille B des polymérases (Burgers et al., 2001). Le complexe de l'ADN polymérase ζ est composé de 4 sous-unités, appelé Pol ζ 4. La protéine Rev3L est la sous-unité catalytique portant l'activité polymérase. Elle forme un hétérodimère avec la protéine Rev7, ou Mad2L2, composant la structure minimale Pol ζ 2. Rev7 représente la sous-unité régulatrice augmentant l'activité catalytique de Rev3L de 20 à 30 fois (Nelson et al., 1996) par un mécanisme encore inconnu. Cette interaction Rev3L/Rev7 se fait au niveau de séquences consensus $\phi\phi x PxxxxPSR$ (où ϕ représente un acide aminé aliphatique) dans deux domaines d'interaction adjacents en positions 1877 et 2009 de la protéine Rev3L (Hanafusa et al., 2010; Tomida et al., 2015). La structure cristallographique de Rev7 en liaison avec les deux fragments correspondants de Rev3L a été obtenue. La dimérisation de Rev7 semble contribuer à la fonction biologique de Pol ζ chez les cellules mammifères. Cependant, chez la levure, les sites de liaison Rev3/Rev7 restent encore à découvrir et on ne sait pas si Rev7 forme un dimère dans le complexe Pol ζ de levure (Gómez-Llorente et al., 2013; Hara et al., 2010; Nelson et al., 1996).

Rev3L est aussi capable de se lier à une des sous unités régulatrices de la polymérase δ , la protéine POLD2, se liant elle-même à POLD3 (Baranovskiy et al., 2012a; Johnson et al., 2012; Lee et al., 2014; Makarova et al., 2012). En effet, le domaine C-terminal de Rev3L possède une forte

homologie de séquence avec celui de la polymérase δ . De plus, des mutations de ce domaine CysB chez la levure entraîne la perte de l'interaction entre Rev3 et Pol31 conduisant à un défaut de la mutagenèse induite (Makarova et al., 2012).

Une structure à faible résolution du complexe Pol ζ de levure a été obtenue par microscopie électronique à coloration négative et suggère que les trois sous-unités régulatrices forment un complexe où Rev7 est capable d'interagir directement avec Pol32 et Pol32 interagit directement avec Pol31 (Gómez-Llorente et al., 2013). Cette interaction POLD3/REV7 n'a pas été décrite chez les mammifères. Cependant, le rôle de ces sous-unités régulatrices dans Pol ζ reste largement inexploré. Un rôle important serait de stimuler le recrutement de Rev1, un régulateur majeur de la synthèse translésionnelle. En effet, Rev1 est capable d'interagir à la fois avec Rev7 et POLD3 et cette interaction est nécessaire au bon fonctionnement avec Pol ζ .

In vitro, Pol ζ 4 montre une activité catalytique supérieure à Pol ζ 2 de 5-10 fois chez l'Homme et de plus de 30 fois chez la levure et ce complexe Pol ζ 4 est nécessaire pour la synthèse TLS dépendante de PCNA (Lee et al., 2014; Makarova et al., 2012). Finalement, ce partage de sous-unités entre Pol δ et Pol ζ permettrait un échange simple entre les deux polymérases pour passer de la réplication physiologique à une réplication translésionnelle.

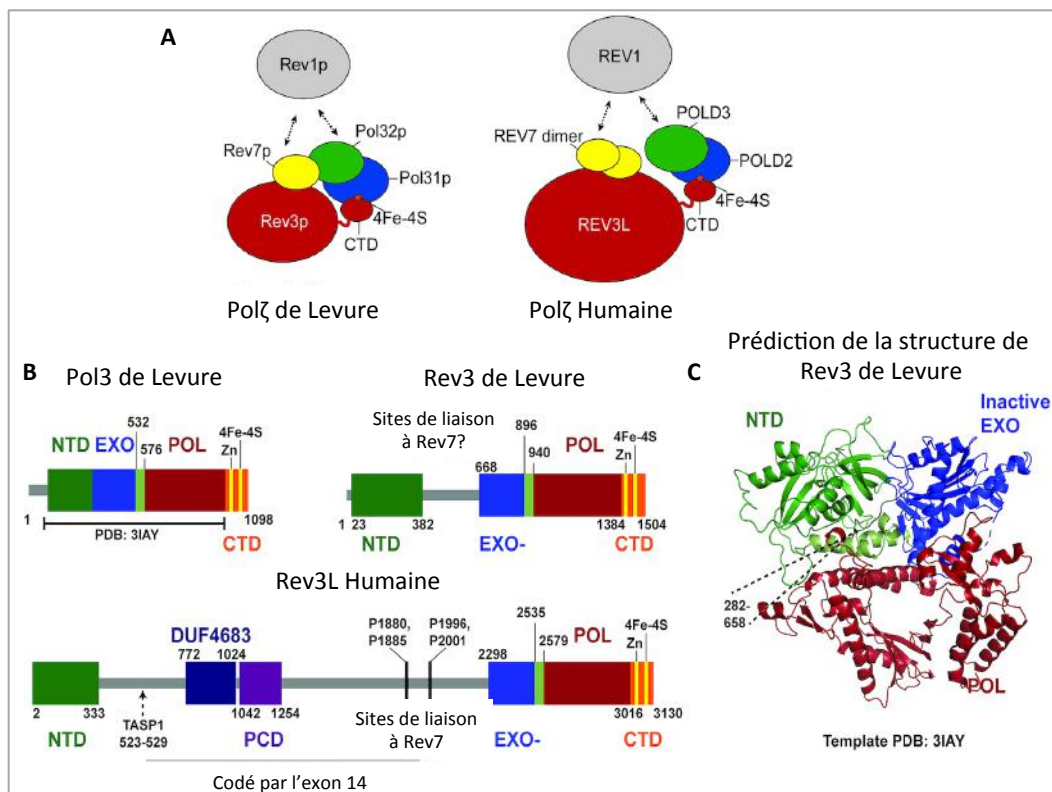


Figure 25: Structure de la Polymérase ζ 4 (Martin and Wood, 2019). A) Représentation schématique complexe Pol ζ 4 chez la levure et chez l'homme. B) Représentation schématique des sous-unités catalytique de Pol δ (Pol3) et Pol ζ (Rev3) chez la levure ainsi que Rev3L humaine. NTP : Domaine N-terminal, EXO : Domaine exonucléase, POL : Domaine polymérase, CTD : Domaine C-terminal, DUF : Domaine de fonction inconnue (Domain of unknown function), PCD : Domaine chargé positivement. C) Modèle de structure de Rev3 obtenu grâce au logiciel PHYRE2 à partir de la structure de Pol3 comme modèle compte tenu de la forte homologie entre les deux protéines. La séquence 382-668 entre le NTD et le domaine exonucléase n'a pas été intégré dans l'analyse.

2. Le gène et la protéine Rev3L

a) Saccharomyces cerevisiae

Chez la levure *S. cerevisiae*, le gène *REV3* est localisé sur le chromosome XVI sur le brin indirect, mesure 4 515 pb et n'est constitué que d'un exon. Ce gène code pour une protéine de 1 504 acides aminés (173 kDa).

b) Mus musculus

Chez la souris, le gène *Rev3L* est localisé sur le chromosome 10 en position chr10 :39,732,118-39,875,211 dans la bande qB1. Ce gène est situé sur le brin direct et comprend 143 143 pb. Il est composé de 32 exons et code pour au moins deux transcrits principaux dont le majoritaire conduit à la formation d'une protéine de 3 122 acides aminés (350 kDa).

c) Homo sapiens

Chez l'Homme, le gène *REV3L* est localisé sur le chromosome 6 en position chr6 :111,299,031-111,483,715 dans la bande q21. Ce gène est situé sur le brin indirect et comprend 184 685 pb. Il est composé de 41 exons et code pour au moins deux transcrits principaux donnant lieu à deux protéines de 3 130 et 3 052 acides aminés. L'isoforme alternative possède une insertion de 128pb entre les positions 139 et 140 de l'ADN complémentaire conduisant à l'apparition de trois codons stop en phase (Morelli et al., 1998). Un codon d'initiation alternatif au sein de cette insertion a été proposé pour produire une protéine Rev3L tronquée de 78 acides aminés du côté N-terminal. La forme la plus longue de la protéine Rev3L est prédite pour avoir une masse moléculaire de 353 kDa. Seule l'isoforme de référence semble capable de produire une protéine *in vitro* ainsi que *in cellulo* (Kawamura et al., 2001). Finalement, REV3L est exprimé de manière ubiquitaire dans un grand nombre de tissus conduisant à la production d'une faible quantité de protéines et il existe un élément de réponse pouvant lier la protéine p53 en amont du promoteur de *Rev3L* (Gibbs et al., 1998).

3. Structure de la protéine Rev3L de mammifère

La protéine catalytique Rev3L est homologue aux sous-unités catalytiques des polymérases répliquatives au niveau des régions N et C-terminales (Baranovskiy et al., 2012a;

Johnson et al., 2012; Lee et al., 2014; Makarova et al., 2012) et est composée de 6 domaines structuraux majeurs :

- Un domaine N-terminal en positions 1 à 333 contient un signal de localisation à la mitochondrie (MLS) en position 79 (Figure 25, Vert).
- Un domaine chargé positivement en positions 1 042 à 1 254 (Figure 25, Violet).
- Deux domaines de liaison à Rev7 en positions 1 877 et 2 009 (Figure 25, Gris).
- Un domaine exonucléase inactif en positions 2 294 à 2 532 (Figure 25, Bleu).
- Un domaine polymérase en positions 2 558 à 3 006 (Figure 25, Bordeaux).
- Un domaine C-terminal caractéristique de la famille B des polymérases composé respectivement de 2 clusters riches en cystéines, CysA et CysB, formant 2 doigts de zinc et de 4 clusters Fer-Soufre. CysB est nécessaire pour la liaison de POLD2 à Rev3L (Figure 25, Rouge/Jaune).

La structure par cristallographie du domaine catalytique de Rev3L n'a pas encore été découverte à l'exception d'un fragment complexé avec Rev7 (Hara et al., 2010). Cependant, en se basant sur la conservation de séquence entre Rev3L et les sous-unités catalytiques des polymérases répliquatives, il est possible de prédire que Pol ζ possède la même topologie que celles-ci avec un domaine N-terminal coiffant le domaine exonucléase et contenant des résidus capables d'interagir avec l'ADN (Coloma et al., 2016; Gómez-Llorente et al., 2013; Hogg et al., 2014; Swan et al., 2009). Chez les autres polymérases de la famille B, le domaine N-terminale est situé directement en amont du domaine exonucléase entraînant une proximité physique de ces domaines. Cependant dans Rev3L ces deux domaines sont séparés par une large insertion, il est donc intéressant que cette proximité physique ait été conservé malgré la distance physique. Aussi, chez la levure, Pol δ possède une séquence d'environ 45 acides aminés en position C-terminal du domaine exonucléase capable de se replier dans son domaine N-terminal. Cette séquence est conservée chez Rev3L et la modélisation structurelle prédit qu'elle serait capable de se replier de la même manière dans le domaine N-terminal de Rev3 (Martin and Wood, 2019).

L'insertion séparant le domaine N-terminal du domaine exonucléase inactif chez Rev3L est codée par un exon unique dont la taille varie selon les espèces, résultant en une forte variation de taille de la protéine Rev3L. Dans le gène REV3L mammifère, cet exon code pour une séquence de 1386 acides aminés conduisant à une protéine deux fois plus grande que chez la levure (Lange et al., 2016). Tandis que cet exon contient plusieurs régions prédites de désordre structural, il contient aussi un domaine chargé positivement (PCD) nécessaire à l'activité polymérase de la protéine (Lee et al., 2014). Aussi, une région d'environ 250 acides aminés en amont du PCD présente une forte homologie au gène *NEXMIF* (*KIAA2022*) cependant le rôle de ce domaine dans Rev3L ou NEXMIF n'est pas encore compris.

4. Conservation au cours de l'évolution du gène *Rev3L*

Le gène *Rev3L* est très conservé au cours de l'évolution et se retrouve de la plante à l'homme. Le fait que la protéine *Rev3L* de mammifère mesure environ deux fois la taille de celle de levure provient principalement de l'intégration de l'exon 14 qui code pour 388 aa (Gibbs et al., 1998). Cette intégration se serait faite au cours de l'évolution au niveau du poisson zèbre. En dehors de l'exon 14, les principaux domaines structuraux de *Rev3L* ont été conservés. Comme chez la levure, la protéine *Rev3L* de mammifère contient un domaine N-terminal, deux domaines de liaison à *Rev7* ainsi qu'un domaine polymérase C-terminal. Le domaine N-terminal possède 36% d'identité entre la levure et les mammifères, le domaine de liaison à *Rev7* 29% et le domaine polymérase 39% (Gan et al., 2008; Makarova and Burgers, 2015).

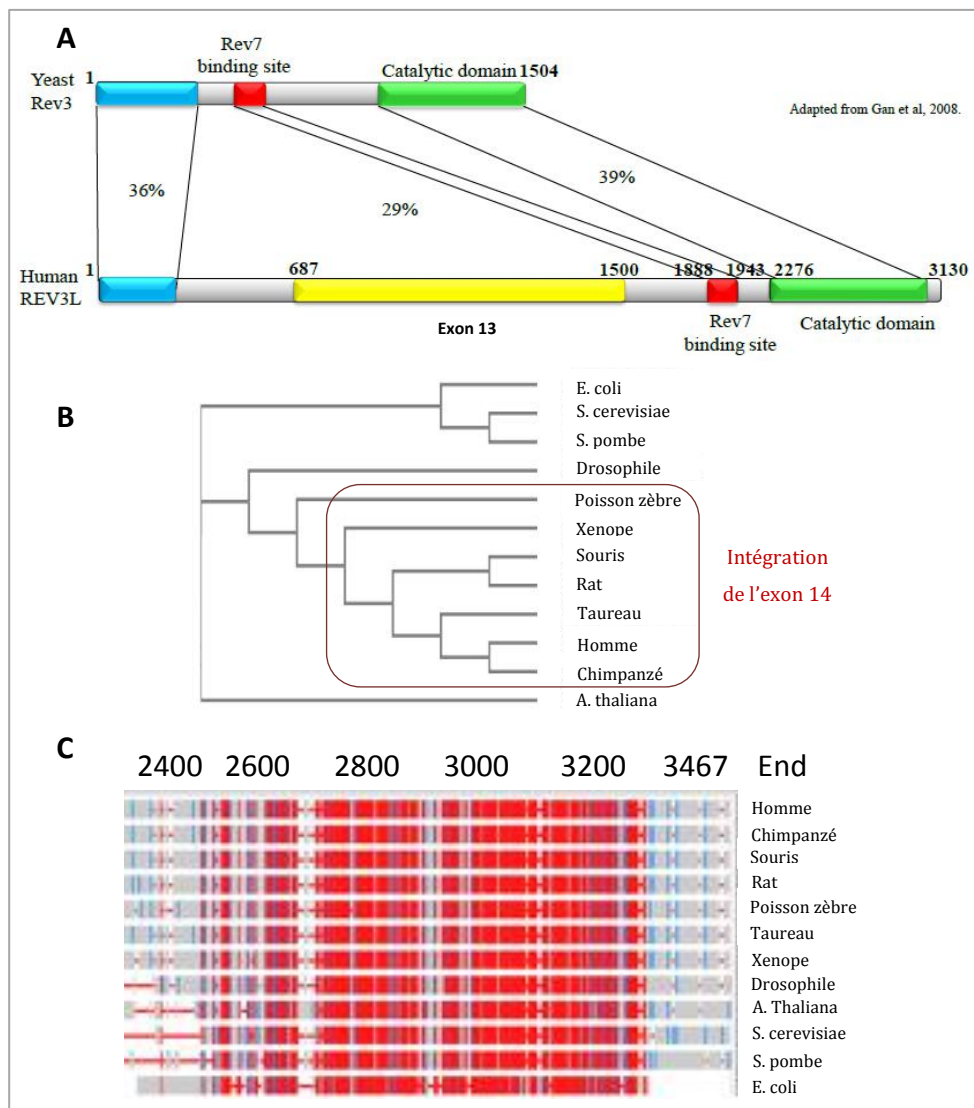


Figure 26: *Rev3L* au cours de l'évolution. A) Adapté de (Gan et al., 2008). Conservation des domaines de *Rev3L* au cours de l'évolution de la levure à l'homme à l'exception de l'intégration de l'exon 14. Bleu : Domaine N-terminal, Jaune : Exon 14, Rouge, Site de liaison à *Rev7*, Vert : Domaine catalytique. B) L'arbre phylogénétique de la protéine *Rev3L* montre l'intégration de l'exon 14 au niveau du poisson zèbre. C) L'alignement des séquences géniques de l'extrémité C-terminale de *Rev3L* montre une forte conservation de ce domaine au cours de l'évolution.

5. Rôles connus de Pol ζ

a) Synthèse Translésionnelle

Chez la levure, il a été montré que Pol ζ est capable de prendre directement en charge quelques lésions spécifiques telles que les dimères de thymine, les thymine-glycols, les sites abasiques ainsi que la présence de ribonucléotides dans l'ADN (Haracska et al., 2001; Johnson et al., 2003; Lazzaro et al., 2012; Nelson et al., 1996; Stone et al., 2011). Comme expliqué précédemment, la synthèse translésionnelle se déroule en deux étapes avec le franchissement de la lésion suivi de l'extension de l'amorce d'ADN. Or, Pol ζ n'est pas efficace pour l'incorporation de nucléotide directement en face d'une lésion et joue donc principalement un rôle d'extenseur. Ce mécanisme a été mis en évidence pour la réplication des 6-4 PP, des CPD, des sites abasiques avec différentes polymérases prenant en charge la première étape de synthèse et Pol ζ catalysant l'extension de l'amorce (Haracska et al., 2001, 2003; Johnson et al., 2000, 2003). Par exemple, *in vitro*, Pol η est capable d'insérer un nucléotide en face d'un pontage intrabrin 1,2(dGpG) induit par le cisplatine tandis que Pol ζ en est incapable (Lee et al., 2014). Cependant, Pol η n'est pas efficace pour étendre une amorce après une lésion tandis que Pol ζ est très efficace. Pol η et Pol ζ possèdent donc des activités biochimiques complémentaires suggérant une coopération de celles-ci dans le franchissement des lésions. De manière intéressante, chez la levure, il a été rapporté que Pol ζ est capable de synthétiser de l'ADN jusqu'à 1kb après une lésion UV-induite. Après irradiation aux UVC, des régions d'ADN simple brin de cette taille sont observées dans des lignées de levure déficientes pour le mécanisme de réparation NER ce qui suggère que Pol ζ est responsable de la synthèse d'une large proportion de ces régions (Kochenova et al., 2015; Lopes et al., 2006). La synthèse translésionnelle par Pol ζ peut donc se faire au cours de la réplication du génome ou en arrière de la fourche de réplication.

b) Rôle dans la mutagenèse et signature mutationnelle

Le domaine exonucléase de Rev3L étant inactif (Nelson et al., 1996), Pol ζ ne possède pas d'activité de relecture et de ce fait est très mutagène. Son taux de mutation varie entre 10^{-2} et 10^{-4} et elle est donc 100 à 10 000 fois plus mutagène que les polymérases répliquatives dont le taux de mutation est de 10^{-6} sur un substrat non endommagé (Zhong et al., 2006). De manière intéressante, Pol ζ est 5 à 30 fois moins mutagène que les autres polymérases TLS appartenant à la famille Y (Zhong et al., 2006). Cependant, la plupart de ces analyses ont été faites *in vitro* en utilisant le complexe Pol ζ 2 (Rev3/Rev7), avant la découverte du complexe Pol ζ 4, et devraient être remises à jour.

(1) Mutagenèse

Chez la levure, c'est au cours d'une étude génétique visant à découvrir de nouveaux gènes responsables de la mutagenèse que Jeffrey Lemontt a décrit pour la première fois les gènes « *Rev* » pour leur phénotype de non réversion aux UVs. Après irradiation des levures à différentes doses d'UV, il a mis en évidence 3 gènes appelés *Rev1*, *Rev2* et *Rev3* pour lesquels il y a une diminution de la mutagenèse UV-induite lorsqu'ils sont mutés. En effet, *Rev3* est responsable de 80% de la mutagenèse induite par les UVs chez la levure. De manière intéressante, *Rev3* est responsable de 60% de la mutagenèse spontanée, ce qui démontre son importance pour la réplication des lésions endogènes chez la levure ainsi qu'après l'exposition à un stress réplicatif. Cependant, *Rev3* ne tient qu'un rôle modéré dans la sensibilité aux UVs et son absence n'induit qu'une faible mortalité après UVs (Lemontt, 1971a, 1971b, 1971c, 1972; Morrison et al., 1989; Quah et al., 1980). En plus des UVs, Pol ζ a aussi été impliquée dans la mutagenèse induite par les rayons gamma, les pontages interbrins causés par le psoralène ainsi que plusieurs agents alkylants anticancéreux (McKee and Lawrence, 1979; Räschle et al., 2008). Les cellules de mammifères déficientes en *Rev3L* présentent aussi une forte diminution du taux de mutations mais sont cependant jusqu'à 10 fois plus sensibles que les cellules non traitées à l'irradiation aux UV ainsi qu'au traitement par le cisplatine (Gueranger et al., 2008; Lee et al., 2014; Segovia et al., 2017).

(2) Signature mutationnelle

Puisque la majorité de la mutagenèse spontanée implique l'action de polymérase TLS, il est fondamental d'identifier les profils de mutations introduits par ces polymérase. Chez la levure, sur un substrat non endommagé, Pol ζ est moins mutagène que d'autres polymérase TLS telles que Pol η mais entraîne une augmentation des transitions C \rightarrow G (0,21%), T \rightarrow A (0,091%), T \rightarrow G (0,08%) (Pavlov et al., 2001; Sakamoto et al., 2007).

Les mutations multinucléotidiques sont des mutations changeant plusieurs nucléotides à la fois. Chez la levure, et dans des cellules humaines, Pol ζ génère des clusters de mutations, définis comme au moins 2 mutations séparées de moins de 10 nucléotides, ainsi que des mutations de bases en tandem après exposition aux UV (Abdulovic et al., 2008; Harfe and Jinks-Robertson, 2000; Minesinger and Jinks-Robertson, 2005; Northam et al., 2010, 2014; Zhong et al., 2006). La mutation dinucléotidique (DNM) GC \rightarrow AA/TT est le plus fréquent polymorphisme humain et est associée à plusieurs maladies. De manière intéressante, ce DNM est fortement corrélé à l'activité de Pol ζ . En effet, cette mutation est retrouvée dans des levures porteuses d'un allèle *Rev3* mutateur (*rev3-L979F*) et déficientes pour la voie de réparation NER ainsi que dans des cellules

murines porteuses d'allèles *Rev3* WT (Sakamoto et al., 2007; Stone et al., 2009, 2012). Cette mutation GC→AA/TT représente donc la signature mutationnelle de Polζ. Cependant, il a été montré que l'environnement nucléotidique direct de la mutation, c'est-à-dire les nucléotides en position -1 et +1 de la mutation, sont importants pour l'action des polymérases. Aussi, le motif WGCW, où W représente soit un A soit un T, semble être un point chaud de mutation pour Polζ dans les cellules de mammifères.

c) Rôle dans la synthèse d'ADN après exposition à un stress répliatif

Comme expliqué précédemment, chez la levure, Polζ est responsable de la majorité de la mutagenèse spontanée et UV-induite suggérant que les cellules utilisent Polζ pour répliquer l'ADN lésé de manière endogène ainsi que suite à un stress répliatif. Une source de stress répliatif peut provenir d'une diminution de la quantité de nucléotides disponibles, par exemple après une incubation avec de l'hydroxyurée (HU). Chez la levure, l'exposition à l'HU entraîne une augmentation de la mutagenèse dépendante de *Rev3* (Northam et al., 2010). De plus, la synthèse d'ADN par Polζ n'est que faiblement altérée par les changements de niveau de dNTPs *in vitro* et *in vivo* contrairement aux autres polymérases (Kochenova et al., 2017). Ces observations suggèrent que Polζ est capable de synthétiser de l'ADN en réponse à un stress répliatif, incluant des circonstances où la quantité de dNTPs est assez faible pour bloquer la répliation de l'ADN. Une autre source de stress répliatif peut provenir de mutations d'un des composant du réplisome entraînant sa déstabilisation et conduisant à une augmentation de la mutagenèse. Ce mécanisme est appelé Mutagenèse induite par un réplisome défectueux (DRIM). De manière intéressante, la vaste majorité de cette mutagenèse est dépendante de *Rev3* (Northam et al., 2006, 2010). Aussi, l'interaction *Rev3/Pol31/Pol32* est essentielle à ce mécanisme suggérant que le complexe Polζ4 est impliqué (Szwajczak et al., 2017).

Finalement, une dernière source de stress répliatif provient de la présence de séquences génomiques difficiles à répliquer, comme par exemple des séquences capables de former des structures secondaires (Cooper et al., 2011; Lu et al., 2015). *In vivo*, les mutations dépendantes de *Rev3* sont retrouvées à des sites prédits pour former des structures en épingles à cheveux (Northam et al., 2014). Ces mêmes structures bloquent la progression de Polδ *in vitro*. Ces observations suggèrent que l'introduction de mutations par Polζ peut être spécifique de la séquence d'ADN rencontrée et pourrait expliquer la variation du taux de mutations induites par *Rev3* observée en fonction de la région génomique étudiée. Un exemple de séquence difficile à répliquer est la présence de séquences d'ADN répétées qui peuvent ralentir la progression des fourches de répliation. Chez la levure, en utilisant un système de gène rapporteur *URA3*, il a été montré que *Rev3* serait impliquée dans la répliation de répétitions télomériques intégrées au

sein des chromosomes appelées « séquences télomériques interstitielles » (ITS) (Moore et al., 2018), et des mutations dépendantes de Rev3 ont été retrouvées jusqu'à 8kb en aval d'une séquence Alu inversée quasi-palindromique ainsi qu'à proximité de répétitions de tri-nucléotides et ce phénomène est accentué en présence de mutants de Pol2 ou Pol3 (Shah and Mirkin, 2015; Shah et al., 2012; Tang et al., 2013). De plus, il a été montré que des cassures doubles brins sont engendrées au sein de ces régions mais il n'a pas encore été établi de lien entre la formation de ces cassures et la mutagenèse dépendante de Rev3.

d) Rôle dans le maintien de la stabilité des sites fragiles communs

Dans les cellules humaines, la déplétion de REV3L conduit à la formation de quatre fois plus de ponts anaphasiques par rapport aux cellules WT, ce qui suggère un problème de sous-réplication de l'ADN. De plus, ces cellules présentent aussi une augmentation de cinq fois du nombre de cassures chromosomiques ainsi qu'une augmentation de l'expression des sites fragiles communs après de faibles doses d'aphidicoline par rapport aux cellules WT. Ces observations sont associées avec une augmentation du nombre de cassures double brin. Les mêmes résultats ont été observés dans des fibroblastes embryonnaires de souris (MEFs) (Bhat et al., 2013a). L'ensemble de ces résultats suggère que l'absence de REV3L induit une instabilité génomique similaire à celle observée après un stress répliatif. REV3L joue donc un rôle dans la réplication des sites fragiles communs et est donc nécessaire à la stabilité du génome.

e) Rôle dans la réparation de l'ADN

Il est connu que des mutations dépendantes de Rev3 sont retrouvées localisées à proximité de cassures doubles brins induites par l'endonucléase HO (Holbeck and Strathern, 1997; Rattray et al., 2002). Cependant, Rev3 n'est pas essentiel à la réparation de ces cassures posant la question de la fonction de Pol ζ dans la synthèse d'ADN à proximité de cassures doubles brins (McVey et al., 2016). Pol ζ est nécessaire à la prise en charge d'un type spécifique de cassures doubles brins à extrémité simple brin, appelée réplication induite de cassures médiées par micro-homologie (MMBIR). Comme pour la réplication induite par cassure (Break-induced replication, BIR), le MMBIR utilise le brin d'ADN homologue comme matrice qui va être envahit par le brin cassé grâce à la présence d'un primer comportant une séquence de micro-homologie. Ce type de cassures peut provenir d'un blocage de fourches de réplication ou de la dégradation des télomères et conduit à l'apparition de mutations complexes à proximité du site d'invasion de brin. En l'absence de Rev3, ces mutations complexes sont absentes (Sakofsky et al., 2015).

Cependant, il serait intéressant de mesurer le niveau de mutations ponctuelles dépendantes de Rev3 en aval d'une cassure réparée par le MMBIR afin de mesurer l'étendue de la synthèse d'ADN par Pol ζ et ainsi comprendre sa fonction potentielle dans ce mécanisme.

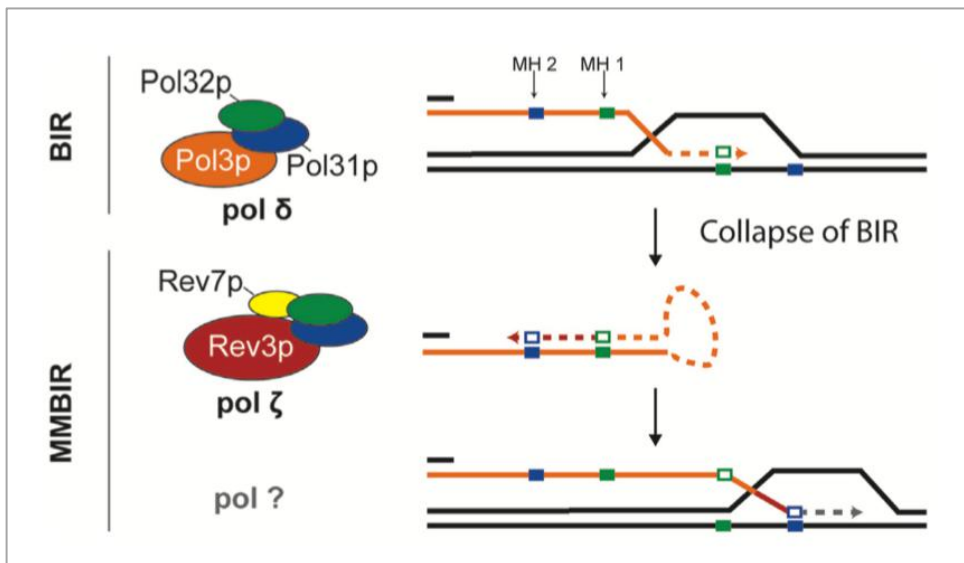


Figure 27 : Modèle du rôle de Pol ζ dans la réplication induite de cassures médiées par microhomologie (MMBIR) (Martin and Wood, 2019).

f) Rôle dans la mitochondrie

Chez la levure, une séquence de localisation à la mitochondrie (MLS) a été décrite dans Rev3 tandis que, chez les mammifères, l'isoforme de référence de Rev3L ne contient pas de MLS fonctionnel. Un MLS a été identifié dans l'isoforme alternative tronquée de Rev3L. L'expression ectopique d'un peptide contenant cette séquence MLS conduit le peptide à la mitochondrie (Singh et al., 2015; Zhang et al., 2006). Cependant, comme expliqué précédemment, il n'a toujours pas été démontré que cette isoforme conduise à la formation de protéine. Aussi, la présence de Rev3L ainsi que d'autres protéines appartenant au complexe Pol ζ n'a jamais été rapporté au sein de la mitochondrie.

Finalement, (Singh et al., 2015) ont montré que Pol ζ peut s'associer avec Poly, la polymérase responsable de la synthèse d'ADN mitochondriale, et l'ADN mitochondrial et protège le génome mitochondrial face aux dommages. Aussi, l'inactivation de Rev3L conduit à une réduction du potentiel de membrane mitochondriale, une réduction de l'activité de phosphorylation oxydative et une augmentation de la consommation de glucose par les cellules. L'ensemble de ces observations semble suggérer un rôle de Pol ζ dans la stabilité mitochondriale.

6. *Rev3L in vivo*

En 2000, trois équipes indépendantes ont montré que l'inactivation du gène *Rev3L* chez la souris entraîne une létalité embryonnaire précoce en utilisant trois stratégies différentes d'inactivation du gène. En effet, les embryons meurent *in utero* entre 8,5 et 12,5 jours post-fécondation. Le corps des embryons *Rev3L*^{-/-} est de taille réduite par rapport aux embryons *Rev3L*^{+/+} et *Rev3L*^{+/-} de la même portée et présentent un retard de développement qui ne dépasse jamais le stade 4-6 somites. Les phénotypes observés vont de l'absence complète de corps embryonnaire jusqu'à des embryons au stade 6 somites. L'absence de *Rev3L* induit aussi des défauts de la morphogenèse avec des défauts des structures cranio-faciales et abdominales. Cependant, les structures extra-embryonnaires sont normalement développées. De plus, les MEFs dérivés d'embryons *Rev3L* KO présentent un défaut de prolifération cellulaire induisant une sénescence ou de l'apoptose après 3-4 générations. La croissance de ces cellules n'est donc possible qu'après leur immortalisation par l'antigène T du virus SV40 (Bemark et al., 2000; Esposito et al., 2000; Lange et al., 2012; Wittschieben et al., 2000). Cependant, même si l'immortalisation permet aux cellules de continuer à proliférer en inactivant les voies apoptotiques, l'absence de *Rev3L* entraîne une augmentation de 10 fois du nombre de cassures doubles brins et de réarrangements chromosomiques dans les cellules mammifères.

Pol ζ est donc la seule polymérase translésionnelle dont l'absence entraîne une létalité embryonnaire chez les mammifères, suggérant son rôle fondamentale et non-redondant au moins dans le développement embryonnaire. Aussi le caractère essentiel de *Rev3L* chez les mammifères contraste-t-il avec le fait que les levures mutées dans le gène *Rev3* ne montrent pas de défaut de croissance, ont moins de mutations et sont uniquement plus sensibles aux agents génotoxiques. *Rev3L* a donc acquis de nouvelles fonctions nécessaires à la vie au cours de l'évolution entre la levure et la souris.

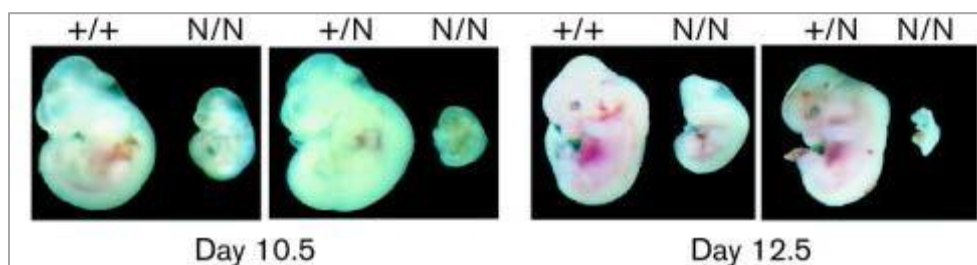


Figure 28: L'absence de *Rev3L* entraîne une létalité embryonnaire précoce chez la souris (Bemark et al., 2000). Les embryons *Rev3L* KO sont significativement plus petits et sous-développés que les embryons WT et hétérozygotes du même âge.

Les souris adultes sont capables de survivre après une inactivation ciblée de *Rev3L* dans les kératinocytes épithéliaux par un système de recombinaison Cre sous le contrôle d'un

promoteur kératin-5 spécifiquement activé dans ce tissu (Lange et al., 2013). Cependant, ces animaux présentent des défauts de peau, incluant un déficit du nombre de cellules au niveau de l'épiderme, un cycle pileux irrégulier, ainsi qu'une perte graduelle de pelage. Aussi, les kératinocytes de ces souris présente un niveau d'instabilité génétique élevé de manière endogène et ne survivent pas lorsqu'elles sont placées en culture. De manière intéressante, ces animaux sont le modèle de souris le plus sensible aux UVs connu aujourd'hui, plus sensible que les souris présentant un défaut du mécanisme de réparation par excision de nucléotide. Cet excès de sensibilité en l'absence de Rev3L provient de deux sources : un défaut du mécanisme de TLS en réponse aux dommages de l'ADN induits par l'irradiation aux UV ainsi qu'un défaut de prolifération des kératinocytes pour réparer la peau après irradiation. Cette dernière observation a été validée après un test de cicatrisation au cours duquel l'épiderme présentant un défaut en Rev3L présente un défaut de cicatrisation (Lange et al., 2018). De plus, après irradiation aux UVs et test de cicatrisation, l'épiderme de ces animaux présente une pigmentation de l'épiderme en absence de Rev3L. Cette observation est intéressante car, de manière différente de chez les humains, l'épiderme basal des souris ne contient pas de mélanocytes. La pigmentation observée provient du stress cellulaire causé par les dommages à l'ADN développé dans les kératinocytes Rev3L^{-/-}. Ces signaux de stress induit une migration des mélanocytes de manière p53 dépendante au niveau de l'épiderme. Dans les kératinocytes, p53 stimule la production d'eumélanine et augmente ainsi la pigmentation (McGowan et al., 2008).

7. REV3L et cancer

REV3L est exprimé de manière ubiquitaire dans des tissus sains ou tumoraux avec des niveaux d'expression variables (Kawamura et al., 2001; Xiao et al., 1998). Des mutations de Rev3L ont été retrouvées dans 10% de l'ensemble des lignées cancéreuses NCI-60. Une analyse de lignées cancéreuses via cBioportal prenant en compte les mutations et les variants de nombres de copies rapporte que Rev3L est altéré dans environ 2% des cancers dont 18% sont associés à une perte d'hétérozygotie. Le gène *REV3L* est sous-exprimé dans des cancers du côlon, du poumon, de l'estomac, de la prostate et du rein comparativement aux tissus sains adjacents et cette sous-expression pourrait être un des premiers évènements du développement tumoral entraînant l'accumulation de dommages de l'ADN (Brondello et al., 2008). Cependant, REV3L est surexprimée dans le gliome, le mélanome et le cancer de l'œsophage. Cette surexpression serait un mécanisme de réponse aux dommages de l'ADN mis en place par la cellule tumorale, entraînant une résistance aux traitements anti-cancéreux par le

franchissement des lésions induites ainsi que l'accumulation de mutations (Song et al., 2017; Wang et al., 2009; Zhu et al., 2016a).

Des polymorphismes de REV3L ont été retrouvés dans des familles humaines et ont été associés à des prédispositions au cancer du poumon, du sein, colorectal et au mésothéliome (Esteban-Jurado et al., 2016; Goričar et al., 2014; Huang et al., 2016; Varadi et al., 2011; Zhang et al., 2013). Aussi, la présence de Pol ζ est associée à une progression tumorale ainsi qu'à une résistance aux traitements chimiothérapeutiques et donc à un mauvais pronostic vital dans plusieurs types de cancers (Shi et al., 2013; Song et al., 2017; Sui et al., 2010; Wang et al., 2009, 2015; Zhu et al., 2016a). A l'inverse, de par son rôle fondamental dans le maintien de la stabilité génétique et chromosomique, l'inhibition de REV3L est associée à un arrêt de prolifération des cellules cancéreuses conduisant à l'apoptose de ces cellules et permet la sensibilisation des tumeurs à différents traitements anticancéreux tels que le cisplatine et le 5-FluoroUracil (Brondello et al., 2008; Knobel et al., 2011; Kong et al., 2018; Wang et al., 2015; Yang et al., 2015). De par ses fonctions protectrices du génome, Rev3L serait donc un gène suppresseur de tumeurs. En effet, la délétion conditionnelle ou mosaïque de Rev3L dans divers tissus chez la souris est associée à une augmentation de la tumorigenèse dans le tissu correspondant. Par exemple, chez des souris mutantes pour p53, la délétion mosaïque de Rev3L dans la lignée lymphocytaire conduit à une augmentation de l'incidence ainsi qu'à une réduction du temps de latence de lymphomes. 90% des souris portant la délétion conditionnelle de Rev3L dans les kératinocytes développent un carcinome après 1 an. Aussi, la délétion partielle de Rev3L dans la glande mammaire est associée à une augmentation de la tumorigenèse (Lange et al., 2013; Wittschieben et al., 2010). Il est probable que l'augmentation de tumeurs observées en absence de Rev3L soit due à une augmentation des réarrangements chromosomiques oncogéniques au sein de ces tissus.

Pol ζ joue donc un rôle important dans les processus de progression tumorale et de chimiorésistance. REV3L représente à la fois un biomarqueur intéressant de la sensibilité tumorale aux traitements chimiques (chimiothérapie) ou physiques (radiations) ainsi qu'une cible thérapeutique potentielle, en collaboration de la chimiothérapie, permettant de supprimer la chimiorésistance de certaines tumeurs. Pol ζ représente donc une cible thérapeutique potentielle pour le traitement des cancers et une meilleure compréhension des différentes fonctions de Pol ζ et de sa régulation dans les cellules mammifères est nécessaire dans le but de développer des inhibiteurs spécifiques de cette protéine.

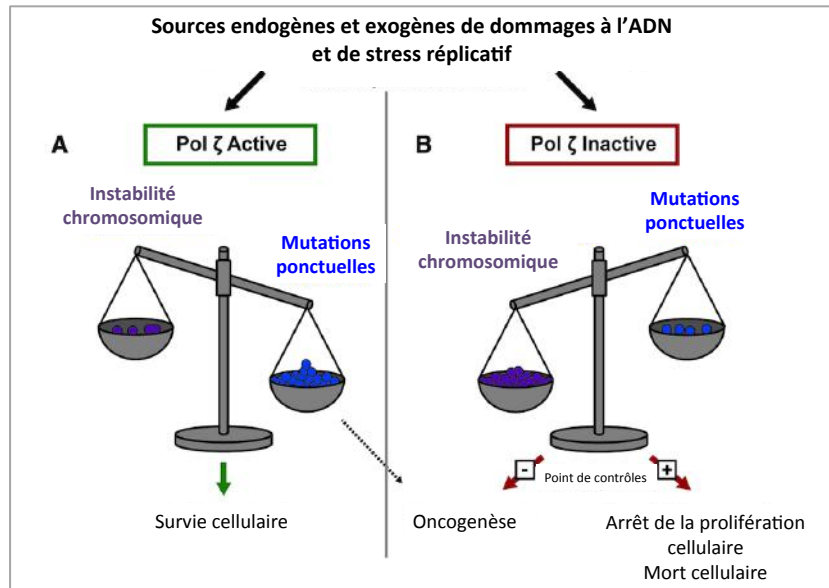


Figure 29 : Polζ agit comme suppresseur de tumeur en préservant la stabilité chromosomique au prix d'une augmentation de la mutagenèse chez les mammifères (Martin and Wood, 2019). A) Les cellules possédant une Polζ active ont besoin de cette polymérase pour prévenir l'apparition de cassures doubles brins et de réarrangement chromosomiques en présence de stress réplicatif permettant ainsi la survie cellulaire mais au prix d'une mutagenèse élevée. B) En absence de Polζ, les cellules vont accumuler des cassures doubles brins et de potentiels réarrangement chromosomique après exposition à un stress réplicatif. Cette instabilité chromosomique conduit à un arrêt de la prolifération cellulaire dans les cellules possédant des points de contrôles fonctionnels. Les cellules ayant perdu leurs points de contrôles du cycle cellulaire, telles que les cellules cancéreuses, vont survivre et les réarrangements chromosomiques pourront conduire à l'établissement d'un mécanisme oncogénique.

Au cours de cette dernière partie j'ai présenté l'état des connaissances avant le début de ma thèse concernant la caractérisation et les différents rôles connus de Polζ. Cependant des questions essentielles restaient toutefois en suspens telles que : Polζ est-elle impliquée dans la réplication non-perturbée des cellules mammifères ? Par quel mécanisme ? A quel niveau ? Comment Polζ est-elle recrutée à la chromatine ? Si cette polymérase est si mutagène mais si fondamentale pour le maintien de la stabilité génétique, quels sont les mécanismes de régulation mis en place par les cellules de mammifères pour limiter son activité ? C'est l'ensemble de ces questions auxquelles j'ai essayé de répondre et dont je vais présenter les quelques réponses obtenues au cours de ma thèse.

E. Objectifs de la thèse

Bien qu'il soit clairement établi que Pol ζ est impliquée dans de multiples processus biologiques afin de maintenir la stabilité du génome chez les eucaryotes supérieurs, beaucoup de questions restent en suspens. En effet, sa fonction biologique n'est toujours pas clairement définie et nous ne connaissons que très peu de choses sur les mécanismes qui contrôlent la régulation de cette ADN polymérase. Ceci est principalement dû aux difficultés techniques d'étudier la sous-unité catalytique de Pol ζ , la protéine REV3L. Quatre problèmes majeurs sont à prendre en considération pour l'étude de Pol ζ . 1) REV3L étant essentielle pour le développement embryonnaire, il est impossible d'obtenir des organismes modèles KO de mammifères viables pour cette protéine, 2) l'inhibition de REV3L par manipulation génétique ou par ARN interférant dans des cultures primaires de cellules de mammifères induit un arrêt de la prolifération cellulaire au bout de quelques générations, empêchant son étude à long terme. 3) l'absence d'anticorps commercial spécifique dirigé contre REV3L empêche sa détection cellulaire. 4) La grande taille de cette protéine (>350 kDa) limite les approches biochimiques et réduit fortement sa capacité à être surexprimée de manière ectopique dans les cellules de mammifères. L'ensemble de ces difficultés techniques ont freiné l'étude et la caractérisation de cette TLS polymérase pourtant essentielle.

Toutefois des outils biologiques importants ont été développés. En effet, avant le début de ma thèse, le laboratoire du Dr Richard D. Wood (MDACC, USA) avait développé deux lignées cellulaires permettant de lever certains verrous techniques : 1) Une lignée cellulaire de fibroblastes embryonnaires de souris (MEFs) KO pour REV3L pour laquelle la prolifération cellulaire a été restaurée par leur immortalisation par l'antigène T du virus SV40, ce qui nous permettait d'étudier l'impact de son absence sur plusieurs générations, 2) Une lignée de MEFs pour laquelle une étiquette FLAG (3xFLAG) a été introduite au niveau du premier exon du gène *Rev3l* permettant ainsi de détecter la protéine grâce à un anticorps anti-Flag, ceci après plusieurs mois de mise au point.

Le développement de ces lignées cellulaires murines nous a permis d'élaborer un projet de thèse afin de répondre à deux grandes questions auxquelles nous ne pouvions répondre auparavant :

1) Quelle est la fonction de Pol ζ dans les cellules mammifères?

La majorité des connaissances sur Pol ζ ont été obtenues soit par des études d'interactions génétiques chez la levure soit par des approches *in vitro* en utilisant la protéine Rev3 de levure. Ainsi, étant donné que Pol ζ /REV3L a une fonction essentielle dans le développement embryonnaire et la prolifération cellulaire chez les mammifères, il semblait donc primordial d'étudier son rôle dans la réplication non perturbée de l'ADN ainsi qu'après exposition à un stress génotoxique étant donné sa fonction de synthèse translésionnelle.

2) Quels sont les mécanismes de régulation de Polζ dans les cellules mammifères?

Polζ a été impliquée dans des mécanismes de protection cellulaire face aux dommages de l'ADN induits par différents stress génotoxiques. Plus particulièrement, il a été suggéré que Polζ serait impliquée dans la résistance aux traitements chimio-thérapeutiques. Il est donc nécessaire de mieux comprendre la régulation de cette protéine qui pourrait, dans certains types de cancer, potentiellement être ciblée en parallèle de la chimiothérapie dans le but de re-sensibiliser la tumeur.

Ces deux questions ont représenté le point de départ ainsi que le fil conducteur de mes travaux de thèse. Dans la première partie de ma thèse, j'ai donc voulu comprendre les fonctions cellulaires de Polζ pouvant expliquer son rôle essentiel à la vie. Je me suis alors intéressée à l'impact de l'absence de la polymérase ζ sur la stabilité génétique, épigénétique et chromosomique afin de comprendre quels sont les défauts induits par cette absence qui entraîne une mortalité embryonnaire précoce chez la souris. Les résultats de ces travaux sont présentés dans l'Article 1 (*soumis*). Dans une deuxième partie, j'ai étudié les différents mécanismes de régulation de Polζ, en conditions physiologiques et en présence de différents stress réplicatifs, dans les cellules mammifères. Les résultats obtenus pour cette partie sont présentés dans l'Article 2 (*en préparation*).

Je dois cependant préciser que le travail présenté dans la partie I avait été initié par une doctorante (Sana Ahmed-Seghir) avant son départ en stage post-doctoral.

II. Résultats

- A. Partie I - Conséquences de l'absence de Rev3L sur la stabilité génétique et épigénétique (Article 1, soumis)..... p.77
- B. Partie II - Caractérisation fonctionnelle et régulatrice de l'ADN Polymérase zêta (Article 2, en cours de préparation)..... p.140

Partie I - Conséquences de l'absence de Rev3L sur la stabilité génétique et épigénétique.

Résumé Article 1: Au cours de la première partie de ma thèse, je me suis intéressée à la caractérisation des fonctions biologiques de l'ADN polymérase ζ dans les cellules mammifères. Pour ce faire, j'ai utilisé une lignée cellulaire de fibroblastes embryonnaires de souris dans laquelle le gène *Rev3l* a été inactivé dans le but de déterminer les défauts cellulaires associés à son absence et d'en déduire les fonctions de la protéine REV3L, et par extension celles de Pol ζ . Nous avons montré que dans les MEFs *Rev3l* KO, la progression de la phase S est perturbée avec une accumulation des cellules en milieu-fin de phase S. La technique de « timing de réplication » à l'échelle du génome nous a permis de mettre en évidence une réplication retardée dans 5% du génome. Ces retards sont localisés dans des domaines réplicatifs particuliers, appelés « régions de transition temporelle », pour lesquels une réplication unidirectionnelle a été mise en évidence, fragilisant ces régions génomiques. Ceci suggère que, dans des conditions physiologiques, Pol ζ participe à la réplication de certaines régions du génome difficiles à répliquer. De façon intéressante, nous avons observé que dans les MEFs *Rev3l* KO, ce défaut de réplication est associé à une augmentation de la mutagenèse, probablement due à une réparation infidèle de cassures de l'ADN induites par des blocages de la fourche de réplication en absence de Pol ζ . Nous avons également montré que Pol ζ joue un rôle critique dans le maintien épigénétique en interagissant avec des composants de l'hétérochromatine tels que les marques répressives d'histones. Grâce à son interaction directe avec la protéine HP1 (heterochromatin protein 1), Pol ζ localise au niveau des régions péri-centromériques riches en répétitions microsatellites. En lien avec ces résultats, nous avons mis en évidence que dans les MEFs *Rev3l* KO, les nombreuses cassures d'ADN double brin spontanées sont préférentiellement localisées dans l'hétérochromatine. Finalement, des résultats préliminaires suggèrent un rôle de Pol ζ dans la réplication des régions télomériques, riches en hétérochromatine et en ADN répété. Ainsi, nous proposons que Pol ζ a été sélectionnée au cours de l'évolution pour assister les polymérases réplicatives à répliquer des régions hétérochromatiennes compactes, afin de limiter l'instabilité génomique.

1. DNA polymerase zeta contributes to heterochromatin replication to prevent genome instability (submitted)

Barbara Ben Yamin^{1,8}, Sana Ahmed-Seghir^{1,8,10}, Junya Tomida^{2,8,12}, Caroline Pouvelle¹, Guillaume Meurice³, Nathalie Droin⁴, Philippe Dessein³, Didier Goidin⁵, Sabine S. Lange^{2,11}, Maité T Mitjavila-Garcia⁶, Giuseppe Baldacci⁷, Jean Charles Cadoret^{7,9} Richard D Wood^{2,9} and Patricia L Kannouche^{1,9,13,*}

¹Paris-Sud University, CNRS-UMR8200, Equipe labellisée Ligue Contre le Cancer, Gustave Roussy, Villejuif, France.

²Department of Epigenetics and Molecular Carcinogenesis, The University of Texas MD Anderson Cancer Center and The University of Texas Graduate School of Biomedical Sciences at Houston, Houston, TX 77030, USA

³Bioinformatics Platform, Gustave Roussy, Villejuif, France.

⁴Genomics Facility, Gustave Roussy, Villejuif, France.

⁵Agilent technologies, France

⁶INSERM UMR935, SFR André Lwoff, University Paris Sud, Paul Brousse Hospital, Villejuif, France.

⁷Institut Jacques Monod, UMR7592, CNRS and University Paris-Diderot, 15 Rue Hélène Brion, Paris, Cedex 13 75205, France.

⁸Co-first authors

⁹Co-last authors

¹⁰Present address: Molecular Biology Program and Department of Radiation Oncology, Memorial Sloan-Kettering Cancer Center, New York, NY 10065, USA

¹¹Present address: Texas Commission on Environmental Quality, Austin, TX 78753 USA

¹²Present address: Department of Biological Sciences, University of North Carolina at Charlotte, Charlotte NC 28223, USA

¹³Lead contact

* Correspondence: patricia.kannouche@gustaveroussy.fr (P.L.K)

Key words: DNA replication, Replication timing, TLS polymerase, mutational events, heterochromatin, Rev3, Rev7, polymerase zeta

Running title: DNA polymerase zeta contributes to heterochromatin replication

ABSTRACT

The DNA polymerase zeta (Pol ζ) plays a role in bypassing DNA damage. REV3L, the catalytic subunit of Pol ζ is also essential in mouse embryonic development and cell proliferation for reasons that remain unclear. Here we show that S-phase progression is impaired in *Rev3l*^{-/-} cells. *Rev3l* loss disrupts replication timing at specific genomic domains that replicate in mid-to-late S phase and correlates with an increase of mutational events likely due to replication-dependent fork breakage at these loci. We also revealed that REV3L protein interacts with heterochromatin components including repressive histone marks, and localizes in DNA repeat-rich pericentromeric regions through direct interaction with HP1 dimer. Consistent with this, spontaneous double-strand breaks (DSBs) caused by REV3L loss are preferentially located in heterochromatin regions. We propose that in mammalian cells, Pol ζ assists high-fidelity DNA polymerases in replicating heterochromatin regions thus preventing genome instability.

INTRODUCTION

It is well established that DNA polymerase delta (Pol δ) and epsilon (Pol ϵ) catalyze the high-fidelity duplication of the genome (reviewed in (Burgers and Kunkel, 2017)). However, an emerging concept is that translesion synthesis (TLS) DNA polymerases, known for their error-prone lesion-bypass properties, can also facilitate DNA synthesis of non-damaged DNA. TLS polymerases may be employed when replicative DNA polymerases are not able to pass through stalled sites of replication at structured DNA sequences and/or hard-to-replicate genomic regions (reviewed in (Tsao and Eckert, 2018)). For example, studies have shown that DNA polymerase eta (Pol η) and zeta (Pol ζ) are required to maintain common fragile site stability (CFS) in human cells (Bergoglio et al., 2013; Bhat et al., 2013; Despras et al., 2016). It is not known whether specialized DNA polymerases are recurrently recruited to DNA more broadly during S phase to assist replicative DNA polymerases for replicating unconventional DNA structures preventing thus genome instability at the cost of increased point mutations. Intriguingly, growing evidence suggests that replication timing influences genomic mutation rates with an increasing gradient of single-nucleotide substitutions that correlates with late replicating regions (Koren et al., 2012; Polak et al., 2015; Stamatoyannopoulos et al., 2009), but the link between mutation rate and spatiotemporal organization of the genome has been poorly understood.

Pol ζ is a TLS polymerase complex in eukaryotes, consisting of four subunits: Rev3, the catalytic subunit, Rev7 that enhances Rev3 activity, and two subunits shared with the replicative polymerase Pol δ (Baranovskiy et al., 2012; Johnson et al., 2012; Lee et al., 2014; Makarova et al., 2012). In contrast to other TLS enzymes that belong to the Y-family of polymerases, Pol ζ belongs to the B-family that includes the highly accurate replicative DNA Pol δ and Pol ϵ (Gan et al., 2008). Pol ζ lacks an intrinsic 3'-5' exonuclease activity, making this TLS polymerase error-prone with a spontaneous mutation rate that is 10 to 100-fold greater than that of replicative DNA polymerases in yeast (McCulloch and Kunkel, 2008; Stone et al., 2009). Pol ζ is a key player in translesion DNA synthesis by elongating primer termini that are positioned opposite base damage and non-instructional lesions (Johnson et al., 2000). This DNA polymerase has been extensively characterized in the budding yeast *S. cerevisiae*, showing that spontaneous as well as damage-induced mutagenesis is largely dependent on Pol ζ (Makarova and Burgers, 2015).

Counterparts of the yeast *REV* genes have been identified in other eukaryotes. Mouse and human Rev3-like (*Rev3l*) orthologs have a large extra segment which is not conserved in yeast *REV3*, and are thus about twice the size of the 173-kDa yeast Rev3 (350 and 353 kDa, respectively). Despite the established participation of *Rev3l* in important cellular processes (Martin and Wood, 2019), the role of REV3L protein is incompletely understood and studies

have been hampered by the inability to detect this large protein in cells. Pol ζ is unique among TLS polymerases in mammalian cells, because inactivation of *Rev3l* gene leads to embryonic lethality in mice (Esposito et al., 2000; Van Sloun et al., 2002; Wittschieben et al., 2000). *Rev3l*^{-/-} mouse embryonic stem cells are not viable and primary mouse embryonic fibroblasts (MEFs) obtained from conditional *Rev3L* knockouts show genome instability and growth defects without external damage to DNA (Lange et al., 2012; Wittschieben et al., 2006). All these data underscore an essential role of Pol ζ in mammalian cells.

In this study we investigated the biological function of Pol ζ during S phase under normal growth conditions. We show that S phase progression is impaired in *Rev3l*^{-/-} MEFs. Genome-wide profiling of replication timing revealed that a set of genomic loci located in temporal transition regions are delayed in response to *Rev3l* inactivation, suggesting an involvement of Pol ζ in these specific regions that replicate in mid-to-late S phase where replication forks progress unidirectionally (Desprat et al., 2009) and known to be enriched in repressive histone marks (Picard et al., 2014). Moreover, changes in replication timing are associated with an increase of mutational events probably due to replication-dependent fork breakage in these genomic regions. We also revealed that REV3L interacts with heterochromatin components and localizes in DNA repeat-rich pericentromeric regions through direct interaction with HP1 dimer. Consistent with this, loss of *Rev3l* causes spontaneous DSBs that are preferentially located in heterochromatin regions. We therefore propose that in vertebrates, Pol ζ has been selected to assist high-fidelity DNA polymerases in replicating heterochromatin regions during mid-to-late S phase preventing thus genome instability.

RESULTS

S phase progression is altered in *Rev3l*^{-/-} MEFs

Polζ is critical for proliferation of normal primary cells suggesting that Polζ can be required during unchallenged DNA replication (Lange et al., 2012). To gain molecular insight into its function during DNA replication, we investigated the dynamics of S phase progression in the absence of REV3L, the catalytic subunit of Polζ. We used *Rev3l*^{-/-} and *Rev3l*^{+/+} mouse embryonic fibroblasts (MEFs) (Lange et al., 2012). We first monitored S phase progression in *Rev3l*^{-/-} and control *Rev3l*^{+/+} cells. Unchallenged *Rev3l*^{-/-} MEFs were pulse-labeled with the thymidine analog 5'-bromo-2'-deoxyuridine (BrdU) for 15 min, released into fresh medium, and harvested at the indicated time points. FACS analysis showed that cell cycle progression was altered in *Rev3l*^{-/-} MEFs (**Figure 1A**). A higher percentage of BrdU-positive *Rev3l*^{-/-} cells was detected in late S/G2 phase at 4, 6 and 8 h after the BrdU pulse as compared to control cells. This is mirrored by a depletion of BrdU-positive cells in G1/early S at the same time periods (**Figure 1A**, see red arrows in upper panel and histograms on bottom panel). In *Rev3l*^{+/+} MEFs, BrdU-negative cells in G2 phase needed 2 h to re-enter G1 phase while the re-entry of *Rev3l*^{-/-} cells into G1 phase took longer (**Figure 1A**, see green arrows in upper panel). Quantification (**Figure 1A**, lower panel) showed a 2 h delay in the progression of the S phase in *Rev3l*^{-/-} MEFs (7-8 h for *Rev3l*^{+/+} MEFs vs 9-10 h for *Rev3l*^{-/-} MEFs). Similar results were obtained using REV3L-depleted human cells (**Figure S1**). Our results on the altered progression of *Rev3l*^{-/-} cells through S phase prompted us to investigate patterns of spatio-temporal replication in *Rev3l* null MEFs.

S phase can be divided into different stages by immunofluorescence observation of characteristic thymidine analog incorporation patterns corresponding to early, mid or late S phase (Dimitrova and Berezney, 2002; Guenatri et al., 2004). To analyze the spatial organization of DNA replication in *Rev3l*^{-/-} MEFs, cells were pulsed-labelled with IdU and the replication patterns typically observed in mouse cells were identified and quantified (Guenatri et al., 2004). Representative replication patterns for each category are represented: early S-phase patterns with a high density of small foci distributed throughout the nucleus (stage I), early-mid S phase patterns with small foci throughout the nucleus excepted nucleolus (stage II), then mid-late S phase patterns with a typical ring-shaped labeling around pericentromeric heterochromatin domains (stage III), and finally late S phase patterns with a few dots located mainly at the nuclear periphery (stage IV) (**Figure 1B**). We found that in the absence of REV3L, the proportion of cells with an early replication spatial pattern (SI) decreased whereas the proportion of cells in mid and late S phase increased (**Figure 1B**). A similar effect was found after depletion of REV3L in human cells (**Figure 1C**). These results indicate a delay of DNA replication for mid- and late-replicating regions.

Loss of *Rev3l* disrupts replication timing in specific genomic loci that replicate in mid-to-late S phase

Since loss of *Rev3l* can affect the DNA replication program, we examined replication timing in *Rev3l*^{+/+} and *Rev3l*^{-/-} cells after a few (p5, p7) or many passages (p60) in cell culture. To perform genome-wide profiling of replication timing, cells were pulse-labeled with BrdU and sorted into early and late S phase fractions by flow cytometry (see methods, (Fernandez-Vidal et al., 2014)). Newly-synthesized DNA of each fraction was immunoprecipitated and labeled before hybridization on microarrays. The Agilent CGH algorithm determines the replication timing of genomic domains by analyzing the log ratio value of early fraction vs late fraction. When the log ratio of a domain is positive, the replication timing is early and when the log ratio is negative, the replication timing is late (Ryba et al., 2011). Genome-wide replication timing analysis identified a set of genomic domains that changed replication timing (RT) in response to *Rev3l* inactivation. A striking effect of *Rev3l* deficiency was boundary shifts as exemplified in **Figure 2A**, corresponding to a delay in regions that lie between early and late replicating domains called temporal transition regions (TTR) by Gilbert and colleagues (Hiratani et al., 2008). 157 genomic regions showed this delay (Mid-to-Late) representing ~5% of the whole genome. Notably we never detected advanced replication timing for these transition regions (Mid-to-Early). Observed changes in replication timing were reproducible between replicates. In addition, there were no significant variations in replication timing in *Rev3l*^{-/-} MEFs between p6 and p60 (**Figure 2A, grey lines**) suggesting that the delay in TTR is an early event after *Rev3l* inactivation.

Analysis of the correlation between genomic regions that changed replication timing and the coverage for GC and LINE-1 contents (**Figure 2B**) as well as H3K27ac, H3K4me3, H3K27me3 marks (**Figure 2C**) showed a molecular signature intermediate between that of early and late domains, strengthening our hypothesis that these regions correspond to TTR. These results reveal that *Rev3l* inactivation induces changes in the temporal replication program, especially in particular TTRs that are replicated later than the control cells and suggested that REV3L/Polζ might contribute to replicate these specific genomic domains.

Delay in replication timing correlates with an increase of mutational events in *Rev3l*-deficient cells

In yeast, Polζ contributes to >50% of spontaneous mutagenesis (Quah et al., 1980). We wondered whether the number of mutational events is modified in specific genomic loci showing a delay in RT in *Rev3l*-deficient cells compared to wt cells. To answer this question, specific genomic capture and sequencing was carried out for three chromosomal regions (25 Mb), including two harboring genomic domains with RT changed between *Rev3l*^{+/+} and *Rev3l*^{-/-} MEFs (**Figure 3B**). Mutational events in these chromosomal regions were analyzed by counting

mutations (including nucleotide substitutions, insertion, deletions) differentially found in *Rev3l*^{+/+} and *Rev3l*^{-/-} genomes. This analysis was performed on 4 different groups: *Rev3l*^{+/+} cells after a few (p7) or many (p60) culture passages and *Rev3l*^{-/-} cells after a few (p5) or many (p60) culture passages (**Figure 3A**). Interestingly, we observed a significant difference in mutational events in *Rev3l*^{-/-} compared to *Rev3l*^{+/+} MEFs only in genomic regions displaying a delay in RT, with an increase of mutational events in these regions (**Figure 3C, Figure S2**). This is consistent with the high level of spontaneous DSBs and genome instability found in *Rev3l*^{-/-} cells (Lange et al., 2012; Wittschieben et al., 2006). We also observed that the number of mutational events in temporal transition regions and late replicating regions is significantly lower between p5 and p60 only in *Rev3l*^{-/-} MEFs. This suggests that mutations arising from *Rev3l* loss may impact cell fitness, resulting in a clonal population of *Rev3l*^{-/-} MEFs over time.

***Rev3l* inactivation results in epigenetic changes with down-regulation of numerous developmentally regulated genes**

Links between transcription and replication timing have been well documented (Hiratani et al., 2008; Sima et al., 2019). Genome-wide analyses have identified a strong correlation between early replication and high transcriptional potential (Farkash-Amar et al., 2008; Hiratani et al., 2008; MacAlpine et al., 2004). We reasoned that changes in replication timing in the absence of *Rev3l* may impact the transcriptional program. We therefore performed microarray-based transcriptome profiling from *Rev3l*^{+/+} and *Rev3l*^{-/-} MEFs (**Figure 4A and Figure S3**). Gene altered by at least three-fold between conditions were considered (FDR<0.05). Analysis of microarray data indicated a total of 317 genes with altered mRNA expression in *Rev3l*-deficient cells as compared to control cells, with 112 genes significantly up-regulated and 205 genes that were significantly down-regulated. Genes displaying a down-regulation in *Rev3l*^{-/-} MEFs had significant enrichment for Gene Ontology terms related to organ and system development (**Figure S3A**). Prominent repressed genes are in the *HoxB* gene cluster, including *Hoxb2* to *Hoxb9* as well as two *HoxA* genes (**Figure 4A-B**). *Rev3l* inactivation also resulted in the decrease of expression of approximately 14 genes from imprinted loci (corresponding to 6.8% of down-regulated genes, **Figure S3B-C**). These included paternally expressed genes *Dlk1*, *Snrpn*, *Peg3*, *Mest* and *Dcn* (**Figure S3B**, blue) and maternally expressed genes *H19*, *Meg3* and *Cdkn1c* (**Figure S3B**, pink). *WT1* is a well-characterized developmental gene that is mutated in Wilms' tumor and was identified as being imprinted (Schwienbacher et al., 2000). This gene was significantly down-regulated by at least 50-fold in *Rev3l*^{-/-} MEFs as compared to control cells. qRT-PCR from independent cultures of *Rev3l*^{-/-} and *Rev3l*^{+/+} MEFs confirmed the down-regulation of multiple *HoxB* and imprinted genes (**Figure 4B and S3D**).

We next asked if genes showing altered mRNA expression in the absence of REV3L are located in genomic regions that displayed a RT delay. For this, we integrated the data on gene expression and genome-wide profiling of replication timing. We found that 24 genes (corresponding to 7.6% of deregulated genes) were located in these specific genomic domains (**Table 1**) and all showed a down-regulation. These observations are consistent with the fact that late replicating genes are often silenced. Altogether these data suggest that the loss of *Rev3l* affects expression of genes involved in growth and development.

In addition to gene expression, various genomic features are linked to replication timing. These include histone modification, DNA methylation, DNA repeat sequences, ordered chromatin structure and nuclear compartmentalization (Aladjem, 2007; Hiratani et al., 2008). Changes in replication timing as well as deregulation of gene expression in response to *Rev3l* inactivation prompted us to assess whether these modifications were associated with variations in global levels of histone modifications. Total histones from *Rev3l*^{+/+} and *Rev3l*^{-/-} MEFs were screened by immunoblotting with a panel of histone modification-specific antibodies. REV3L depletion resulted in a four-fold up-regulation of histone H3K27 tri-methylation (**Figure 4C**), and a more than two-fold increase in H3K9me3 and H3K4me3. While H3K27me3 and H3K9me3 are repressive marks, H3K4me3 is associated to active genes.

We then examined repressive histone modifications (H3K27me3 and H3K9me3) at gene loci down-regulated in *Rev3l*^{-/-} MEFs (including *HoxB2*, *HoxB8*, *WT1*). By ChIP-qPCR we found that H3K27me3 and H3K9me3 levels were higher in *Rev3l*^{-/-} than in *Rev3l*^{+/+} MEFs (**Figure 4D**). We also observed enrichment of repressive histone marks in gene loci that showed a delay in replication timing in *Rev3l*^{-/-} cells such as *FoxG1* and *GATA6* genes supporting the idea that the downregulation of genes observed in *Rev3l*^{-/-} cells can be caused by epigenetic silencing.

Constitutive heterochromatin exhibits a complex chromatin landscape, marked by high levels of H3K9me3, DNA methylation and histone hypoacetylation. The pericentromere is a heterochromatic domain that provides a structural scaffold for centromere formation and plays a crucial role in genome stability (Allshire and Karpen, 2008). In mouse cells, pericentromeric heterochromatin consists of ~10⁵ major satellite DNA repeats that are methylated, decorated by H3K9me3 and enriched in the heterochromatin protein 1 alpha (HP1 α) (Maison et al., 2002). Since we observed an up-regulation of H3K9me3 levels in *Rev3l*-deficient cells, we next investigated the DNA methylation pattern at major satellite repeats. Genomic DNA was extracted from *Rev3l*^{+/+} and *Rev3l*^{-/-} MEFs, digested with methylation-sensitive restriction enzymes MaeII, and DNA blots hybridized with major satellite probes. In *Rev3l*-deficient cells, a fraction of DNA satellite repeats was resistant to enzymatic digestion, indicating an increase of methylation in the absence of REV3L (**Figure 4E**). Therefore, our data suggest that REV3L can participate in the duplication of DNA repeat-rich pericentromeric regions known to be replicated in mid-to-late S

phase (Guenatri et al., 2004). Absence of REV3L altered epigenetic marks in these heterochromatic regions.

REV3L interacts with heterochromatin components and repressive histone marks

We reasoned that if REV3L is implicated in heterochromatin replication, it may interact with heterochromatin components. No commercially available antibodies can detect endogenous levels of REV3L, and so we generated a MEF cell line expressing REV3L protein with a 3X Flag tag at the endogenous locus (Flag-*Rev3l*) (**Figure S4A**). We first verified that we were able to detect endogenous Flag-REV3L in this cell line (**Figure S4B-C**). We then used co-immunoprecipitation to examine interaction of REV3L with heterochromatin proteins. We found that both HP1 α and ATRX, a chromatin remodelling protein known to localize in heterochromatin (Eustermann et al., 2011), co-immunoprecipitated with REV3L. As control, we confirmed that REV7, the regulatory subunit of Pol ζ was also co-immunoprecipitated (**Figure 5A**). To strengthen these observations, we performed a proximity ligation assay (PLA) that detects protein associations within 40 nm. As expected, we observed specific PLA signals between REV7 and Flag-REV3L in the nucleus. We also detected PLA signals between HP1 α and Flag-REV3L (**Figure 5B-C**). We then tested whether REV3L was spatially close to particular modified histone marks. We found REV3L in close proximity to H3K9me3, H3K27me3, H3K4me3, H4K20me3, but not to H3K9ac (**Figure 5D-E**). Together, these results show that REV3L localizes in heterochromatin regions and interacts with heterochromatin proteins and repressive histone marks.

Direct interaction of REV3L with HP1 dimer allows REV3L recruitment at pericentromeric regions

To better characterize the molecular mechanisms linking Pol ζ /REV3L to pericentromeric heterochromatin we examined the subcellular localization of REV3L. In mouse cells, pericentromeric regions are easily identified by their intense DNA staining with the dye DAPI. They are clearly discernible during interphase when major satellites from different chromosomes associate to form clusters (so called chromocenters), that co-localize with HP1 α (Maison et al., 2002). To explore REV3L subcellular localization, we generated a panel of eGFP-tagged REV3L truncated proteins (**Figure 6A, top panel**). All constructs were detected mainly in the nucleus and localized into chromocenters, visualized by dense DAPI staining (**Figure 6A, bottom panel**). Interestingly, the smaller REV3L construct (REV3L⁷⁶¹⁻¹⁰²⁹) consisting of 268 residues was sufficient to target most, if not all the fusion protein into the nucleus, suggesting that this polypeptide fragment contains a putative nuclear localization signal (NLS); the amino-acid sequence KSRKRRKMSKLLPP at position 960-973 is a good candidate (Nguyen and

Lavenier, 2009). More importantly this small construct is also sufficient to target fusion protein into chromocenters. We then confirmed that location in pericentromeric heterochromatin by visualizing the co-localization of Rev3⁷⁶¹⁻¹⁰²⁹ with HP1 α and H3K9me3 by immunofluorescence (**Figure 6B-C**). In the mouse, the association of HP1 α at pericentromeric heterochromatin is severely compromised in cells lacking the H3 Lys9 methyltransferases, *Suv39h1* and *Suv39h2* (Peters et al., 2001). We therefore investigated REV3L localization in *Suv39h* double null (dn) cells. HP1 α was no longer enriched at pericentromeric heterochromatin in the *Suv39h* dn cells as previously reported (Maison et al., 2002). Likewise, Rev3⁷⁶¹⁻¹⁰²⁹ displayed a diffuse staining throughout the nucleus (**Figure 6C**). We also observed that Rev3⁷⁶¹⁻¹⁰²⁹ staining was disrupted when cells were treated with trichostatin A (TSA), a histone deacetylase inhibitor which reversibly increases the acetylation level of histone H3K9 at pericentromeric regions (**Figure S5A**). In TSA-treated cells, Rev3⁷⁶¹⁻¹⁰²⁹ followed the HP1 α distribution which dissociates from heterochromatic domains as previously reported (Taddei et al., 2001). These changes reversed rapidly after drug removal. Collectively, these results strongly suggest that HP1 α is required for targeting REV3L to pericentromeric heterochromatin.

We therefore assessed whether REV3L directly interacts with HP1 α . The chromoshadow domain of HP1 α interacts with proteins containing a pentapeptide motif, PxVxL (Murzina et al., 1999; Thiru et al., 2004). Such a motif is present at position 800-804 in the 268-amino-acid polypeptide (REV3L⁷⁶¹⁻¹⁰²⁹) that can target pericentromeric heterochromatin (**Figure 6D**). The PxVxL sequence, and the putative NLS in this region are both completely conserved in vertebrate REV3L [Fig S1 in (Lange et al., 2016), (**Figure S5B**)]. We tested if this motif was critical for REV3L localization into chromocenters by introducing point mutations in the PSVVL motif (**Figure 6E**). Localization of a transfected mutant GFP fusion protein (e.g. ASAVA) was strongly impaired. Two alternative candidate PxVxL motifs are present near the C-terminus of REV3L. Expression of a REV3L construct containing these two motifs (GFP-REV3L²⁷³²⁻³¹³⁰) in mouse cells showed a diffuse staining in both nucleus and cytoplasm indicating, that these motifs alone cannot target Rev3L to pericentromeric heterochromatin. (**Figure S5C**). Taken together, these data indicate that the PxVxL motif at position 800-804 is important for REV3L localization at pericentromeric heterochromatin.

To confirm a direct interaction between REV3L and HP1, we performed a GST pull-down assay by using fusion polypeptides purified from *E. coli*. GST-XPA was used as negative control. While GST-REV3L^{700 - 900} constructs bound to His-HP1 α and His-HP1 γ , the V802D mutant REV3L^{700 - 900} did not (**Figure 6F**) demonstrating that REV3L directly interacts with HP1 via its PxVxL motif at position 800-804. We confirmed these results by transfecting 293T cells with plasmids expressing FH-REV3L^{700 - 900}, V802D FH-REV3L^{700 - 900} or empty vector. WT REV3L^{700 - 900} but not the mutant, co-immunoprecipitated using HP1 γ antibody (**Figure 6G**).

Homodimerization of HP1 through its chromoshadow domain is critical for HP1 binding of H3K9me3 chromatin and chromatin condensation (Hiragami-Hamada et al., 2016). To gain insight into the molecular interaction between REV3L and HP1, we asked whether HP1 dimerization is mandatory for REV3L interaction with HP1. 293T cells were co-transfected with FH-REV3L^{700 - 900} or empty vector and V5-HP1 γ , IY165-168EE V5-HP1 γ mutant (unable to dimerize), or empty vector. We found that dimerization of HP1 is required for its interaction with REV3L (**Figure 6H**). We confirmed this result using F-H-tagged full length Rev3L (**Figure S5D**). These data indicate that REV3L binds HP1 when HP1 forms condensed chromatin with H3K9me3.

Inactivation of *Rev3l* provokes spontaneous DSBs in heterochromatin and a slow repair of heterochromatin-associated DSBs induced by ionizing radiation

Given that *Rev3l* loss leads to elevated levels of DNA damage and replication-DSBs in primary *Rev3l*^{-/-} cells (Lange et al., 2012), we asked whether DSBs occur in heterochromatin regions. To do this, we examined if DSB markers γ H2AX and 53BP1 were spatially close to heterochromatin-associated factor, HP1 using a PLA approach (**Figure 7**). We observed that specific PLA signals between γ H2AX and HP1 β (**Figure 7A**), and 53BP1 and HP1 α (**Figure 7B**) were significantly higher in *Rev3l*^{-/-} cells than in *Rev3l*^{+/+} cells suggesting that replication forks in heterochromatic regions are prone to collapse in the absence of REV3L/Pol ζ . We next tested the repair kinetics of heterochromatin-associated DSBs induced by X-rays. Levels of PLA foci γ H2AX-HP1 β (**Figure 7C**), and 53BP1-HP1 α (**Figure 7D**) were higher in *Rev3l*^{-/-} cells than in *Rev3l*^{+/+} cells soon after 2 Gy irradiation and remain elevated 6 h post-irradiation indicating a slow repair of heterochromatin-associated DSBs in absence of REV3L. These results suggest that REV3L not only prevents genome instability in heterochromatic regions but may also facilitate repair of heterochromatin-associated DSBs.

To gain insight into the genome stability function of REV3L, we examined the integrity of mitotic chromosomes from cultured *Rev3l*^{-/-} and *Rev3l*^{+/+} MEFs by analyzing metaphase chromosome spreads stained with DAPI. Spontaneous chromosomal aberrations were significantly increased in *Rev3l*^{-/-} MEFs with 6.5 events per metaphase in *Rev3l*^{-/-} cells vs ~1.5 in *Rev3l*^{+/+} cells (**Figure 7E**). Intriguingly, 60% of *Rev3l* null MEFs metaphases showed a missing long chromosome arm (**Figure 7F**). This is consistent with a previous study which reported that the distribution of chromosome aberrations was not entirely random in *Rev3l*^{-/-}; *p53*^{-/-} cells with rearrangements occurring preferentially within chromosome long arms (Wittschieben et al., 2006). Moreover there was a 3.5-fold increase of spontaneous chromosomal break/gaps in the absence of REV3L (**Figure 7G**). Interestingly, a large proportion (30%) of these chromatid

breaks and gaps occurred in close proximity of pericentromeric regions reinforcing the hypothesis that REV3L contributes to replication of heterochromatic regions.

DISCUSSION

Unique among translesion synthesis DNA polymerases, Pol ζ is essential during embryogenesis and cell proliferation (Lange et al., 2018; Lange et al., 2012; Wittschieben et al., 2000). However, the mechanisms of essentiality under normal growth conditions are not clear. In this study we discovered that REV3L associates specifically with heterochromatin via a direct interaction with HP1 dimer. This prevents spontaneous DNA and chromosome break formation and facilitates DNA repair in heterochromatin. The absence of REV3L leads to further consequences that may be ascribed to heterochromatin replication-associated functions of Pol ζ . Replication timing is disrupted in specific genomic domains located in TTRs where replication forks can progress unidirectionally. This is correlated with changes in epigenetic landscape and transcriptional control of developmental and imprinted genes. Altogether these results reveal a new function of Pol ζ in assisting high-fidelity DNA polymerases to replicate heterochromatin during mid-to-late S phase, thus preventing chromosome instability.

Specific disruption of replication timing in absence of Pol ζ .

In *Rev3l*-deficient cells, replication timing defects do not affect the whole genome, but only specific loci in temporal transition regions corresponding to 5% of the whole genome. This is a contrast to the change in replication patterns following inactivation of Rif1 (Repressor/activator site binding protein-Interacting Factor 1), with a shift in replication timing in over 40% of the replication segments (Cornacchia et al., 2012; Hayano et al., 2012). Thus, our data suggest an involvement of REV3L/Pol ζ in the replication of particular TTRs that lie between early-replicating DNA and late-replicating DNA and, therefore, often replicate in mid-to-late S phase. Replication in these regions is achieved by either sequential activation of a series of origins (Guilbaud et al., 2011) or of single long unidirectional fork that initiates at an adjacent early origin (Hiratani et al., 2008; Norio et al., 2005; Schultz et al., 2010). It has been reported that the replication machinery propagates at a slower velocity in some TTRs (Desprat et al., 2009; Farkash-Amar et al., 2008), suggesting that one major consequence is an increase in the probability of replication fork stalling and breaks/gaps as have been proposed for fragile sites regions (Letessier et al., 2011; Ozeri-Galai et al., 2014; Watanabe and Maekawa, 2010). Consistent with this idea, syntenic breakpoints in the mouse and human genomes appear to occur predominantly in transition regions (Watanabe et al., 2004), implying that these regions may be sources of breakage during the generation of new chromosomes during evolution (Donley and Thayer, 2013). Thus it appears that certain TTRs are difficult to replicate for

replicative polymerases and that REV3L/Pol ζ may facilitate progression of replication forks. In line with this, studies showed that yeast Pol ζ contributes to the replication of DNA secondary structures that impede replicative polymerases (Northam et al., 2014). From our results we hypothesize that delay in replication timing may reflect an exacerbated slow progression of replication fork and/or fork stalling in specific TTRs. This may explain the high level of spontaneous DSBs in *Rev3l*-deficient fibroblasts, as proposed for common fragile sites (Bhat et al., 2013). Interestingly, changes in replication timing are associated with an increase of mutational events likely reflecting an error-prone repair of accumulated DNA damage.

REV3L loss impacts on epigenetic landscape and transcriptional program

Mid-replicating regions are enriched in facultative heterochromatin, H3K27me3 (Hiratani et al., 2009; Julienne et al., 2013; Picard et al., 2014). Deposition and maintenance of H3K27me3 are key to the spatiotemporal control of developmental gene expression. This is achieved by the Polycomb Repressive Complex 2 (PRC2) that catalyzes both di- and tri-methylation of Histone H3 at lysine 27 (reviewed in (van Mierlo et al., 2019)). Our results showed that *Rev3l* deficiency impacts histone modification profiles, with a substantial increase of H3K27me3 levels and, to a lesser extent, H3K4me3 and H3K9me3. Whether REV3L/Pol ζ contributes to the regulation of PRC2 activities and/or H3K27me3 maintenance during cell cycle is largely unexplored. As a direct or indirect consequence, our transcriptome analysis revealed that in *Rev3l*-deficient cells numerous repressed genes had significant enrichment for Gene Ontology terms related to organ and system development such as multiple *Hox* genes known to be regulated by H3K27me3 (Vieux-Rochas et al., 2015). Intriguingly, we also found that several imprinted genes are down-regulated in the absence of *REV3L*. This may help explain the observations of Rajewsky and colleagues who suggested that overall phenotype of *Rev3l*^{-/-} embryos appears comparable to androgenetic embryos (Fundele and Surani, 1994), raising the possibility that Pol ζ participates in imprinting (Esposito et al., 2000). Thus, our results on the transcriptional program may help explain the embryonic lethality observed in *Rev3l* KO mice.

Pol ζ contributes to heterochromatin replication

Our study also provides valuable insights on the subcellular localization of REV3L. We revealed that REV3L is in close proximity to heterochromatin components including HP1, ATRX and repressive histone marks. Moreover, we found that REV3L can target pericentromeric heterochromatin through a direct interaction with HP1 mediated by its PxVxL motif. We demonstrated that dimerization of HP1 is necessary for interacting with REV3L indicating that REV3L binds HP1 only when HP1 is engaged on chromatin through its interaction with H3K9me3 (Hiragami-Hamada et al., 2016). Interestingly, the PxVxL motif located at position

800-804 in the human REV3L sequence is embedded in a larger region that is well-conserved in vertebrates but absent in invertebrate and fungal genomes. This indicates that the function of Pol ζ described in this study evolved in concert with larger genomes having more complex controls. Consistent with temporal regions that are disrupted in the absence of REV3L, pericentromeric heterochromatin is replicated in mid-to-late S phase (Guenatri et al., 2004; Natale et al., 2018), and replication dynamics through these regions can be weakened by their intrinsic features consisting of highly repetitive DNA sequences that are tightly packed into heterochromatin. Our data suggest that REV3L/Pol ζ contributes to the replication of these heterochromatic regions. In the absence of REV3L, fork movements in pericentromeric heterochromatin can slow down or stall, and a delay in restarting forks may cause heterochromatin propagation (Jasencakova et al., 2010; Saveliev et al., 2003). This is consistent with the global increase of H3K9me3 and the aberrant DNA methylation pattern at major satellite repeats observed in *Rev3l null* cells. This could also explain frequent pericentromeric breaks/gaps detected in *Rev3l*^{-/-} metaphases and spontaneous DSBs in heterochromatin regions. Note that the REV7 subunit also is associated with HP1 α (Tomida et al., 2018), suggesting another mode of interaction of Pol ζ with heterochromatin. A recent study showed that the shelterin subunit TRF2, in association with the helicase RTEL1, is required for progression of replication forks through pericentromeric heterochromatin (Mendez-Bermudez et al., 2018). Further investigation is warranted to define whether REV3L/Pol ζ works together with TRF2 and RTEL1 to facilitate fork progression in such heterochromatic regions.

Interplay between replication timing and mutagenesis

A growing body of evidence suggests that the replication timing program strongly influences the spatial distribution of mutagenic events during both species and cancer evolution with an increasing gradient of single-nucleotide substitutions that correlates with late replicating heterochromatic regions (Koren et al., 2012; Schuster-Bockler and Lehner, 2012; Stamatoyannopoulos et al., 2009). Given the error-prone activity of Pol ζ , our results suggest that REV3L/Pol ζ may contribute to this increased mutagenesis in heterochromatin regions. This is consistent with a recent bioinformatic analysis showing that mutations consistent with a Pol ζ signature increase in late replicating regions of the human genome (Seplyarskiy et al., 2015). Moreover, experiments in yeast established that late-replicating regions of the genome also have higher rates of spontaneous mutagenesis than early-replicating regions. Deletion of the Y-family DNA polymerase Rev1, a partner of Pol ζ , significantly lowers mutation rate specifically in these late-replicating regions suggesting that Rev1/Pol ζ are involved in the gradient of mutagenesis in lower eukaryotes (Lang and Murray, 2011).

Altogether our results unveil important insights on the biological function of Polζ in mammalian cells during unchallenged S phase paving the way for future research on the link between heterochromatin replication, mutagenesis and cancer.

SUPPLEMENTAL INFORMATION

Supplemental information includes five figures.

ACKNOWLEDGMENTS

We thank T. Jenuwein (Max Planck Institute, Germany) for providing *Suv39h dn* cells. We are also grateful to Niels de Wind (Leiden University, Netherland) for sharing unpublished results. We thank Nikola Djordjevic, Charlène Garandeanu and Loélia Babin for experimental help and members of the Kannouche lab for helpful discussions.

The P.L.K lab is supported by La Ligue Nationale contre le Cancer (Equipe labellisée). This work was supported by grants from Institut National du Cancer (INCa) PLBIO16-011 (P.L.K. and J-C.C), INCa-DGOS-Inserm 12551 and ANR-14-CE10-0008-01 (P.L.K. and J-C.C). S.A.S received support from the Ministère de l'Enseignement Supérieur et de la Recherche and the Fondation ARC and B.B received support from the Ministère de l'Enseignement Supérieur et de la Recherche and the Fondation pour la Recherche Médicale. The J-C.C lab is supported by the IdEX Université de Paris ANR-18-IDEX-0001. R.D.W. and J.T were funded by National Institutes of Health grant CA193124, Department of Defense grant W81XWH-17-10239, and the Grady F. Saunders Ph.D. Distinguished Research Professorship.

AUTHOR CONTRIBUTIONS

B.B performed experiments for protein interaction (co-IP and PLA) and ChIP-qPCR, SAS performed GFP-fused protein localization experiments and FACS analyses. J-C.C. performed genome-wide analysis of RT. J-C.C, P.L.K, G.M., N.D. and D.G. designed experiment and performed mutational analysis. J.T. carried out GST pull-down and Co-IP experiments involving HP1 protein. S.S.L. designed the endogenous tagged Rev3l mouse and derived all MEF cell lines used in this research. M.T.M-G. helped in the design of cellular studies and C.P brought technical support in many experiments. P.L.K and R.D.W. wrote the manuscript and all authors reviewed it. P.L.K., R.D.W. J.C.C. and J.T. conceived and planned the study.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

- Aladjem, M.I. (2007). Replication in context: dynamic regulation of DNA replication patterns in metazoans. *Nat Rev Genet* 8, 588-600.
- Allshire, R.C., and Karpen, G.H. (2008). Epigenetic regulation of centromeric chromatin: old dogs, new tricks? *Nat Rev Genet* 9, 923-937.
- Baranovskiy, A.G., Lada, A.G., Siebler, H.M., Zhang, Y., Pavlov, Y.I., and Tahirov, T.H. (2012). DNA polymerase delta and zeta switch by sharing accessory subunits of DNA polymerase delta. *J Biol Chem* 287, 17281-17287.
- Bergoglio, V., Boyer, A.S., Walsh, E., Naim, V., Legube, G., Lee, M.Y., Rey, L., Rosselli, F., Cazaux, C., Eckert, K.A., *et al.* (2013). DNA synthesis by Pol eta promotes fragile site stability by preventing under-replicated DNA in mitosis. *J Cell Biol* 201, 395-408.
- Bhat, A., Andersen, P.L., Qin, Z., and Xiao, W. (2013). Rev3, the catalytic subunit of Polzeta, is required for maintaining fragile site stability in human cells. *Nucleic Acids Res* 41, 2328-2339.
- Burgers, P.M.J., and Kunkel, T.A. (2017). Eukaryotic DNA Replication Fork. *Annu Rev Biochem* 86, 417-438.
- Cornacchia, D., Dileep, V., Quivy, J.P., Foti, R., Tili, F., Santarella-Mellwig, R., Antony, C., Almouzni, G., Gilbert, D.M., and Bonanno, S.B. (2012). Mouse Rif1 is a key regulator of the replication-timing programme in mammalian cells. *EMBO J* 31, 3678-3690.
- Despras, E., Sittewelle, M., Pouvelle, C., Delrieu, N., Cordonnier, A.M., and Kannouche, P.L. (2016). Rad18-dependent SUMOylation of human specialized DNA polymerase eta is required to prevent under-replicated DNA. *Nat Commun* 7, 13326.
- Desprat, R., Thierry-Mieg, D., Lailier, N., Lajugie, J., Schildkraut, C., Thierry-Mieg, J., and Bouhassira, E.E. (2009). Predictable dynamic program of timing of DNA replication in human cells. *Genome Res* 19, 2288-2299.
- Dimitrova, D.S., and Berezney, R. (2002). The spatio-temporal organization of DNA replication sites is identical in primary, immortalized and transformed mammalian cells. *J Cell Sci* 115, 4037-4051.
- Donley, N., and Thayer, M.J. (2013). DNA replication timing, genome stability and cancer: late and/or delayed DNA replication timing is associated with increased genomic instability. *Semin Cancer Biol* 23, 80-89.
- Esposito, G., Godindagger, I., Klein, U., Yaspo, M.L., Cumano, A., and Rajewsky, K. (2000). Disruption of the Rev3l-encoded catalytic subunit of polymerase zeta in mice results in early embryonic lethality. *Curr Biol* 10, 1221-1224.
- Eustermann, S., Yang, J.C., Law, M.J., Amos, R., Chapman, L.M., Jelinska, C., Garrick, D., Clynes, D., Gibbons, R.J., Rhodes, D., *et al.* (2011). Combinatorial readout of histone H3 modifications specifies localization of ATRX to heterochromatin. *Nat Struct Mol Biol* 18, 777-782.
- Farkash-Amar, S., Lipson, D., Polten, A., Goren, A., Helmstetter, C., Yakhini, Z., and Simon, I. (2008). Global organization of replication time zones of the mouse genome. *Genome Res* 18, 1562-1570.
- Fernandez-Vidal, A., Guitton-Sert, L., Cadoret, J.C., Drac, M., Schwob, E., Baldacci, G., Cazaux, C., and Hoffmann, J.S. (2014). A role for DNA polymerase theta in the timing of DNA replication. *Nat Commun* 5, 4285.
- Fundele, R.H., and Surani, M.A. (1994). Experimental embryological analysis of genetic imprinting in mouse development. *Dev Genet* 15, 515-522.
- Gan, G.N., Wittschieben, J.P., Wittschieben, B.O., and Wood, R.D. (2008). DNA polymerase zeta (pol zeta) in higher eukaryotes. *Cell Res* 18, 174-183.
- Guenatri, M., Bailly, D., Maison, C., and Almouzni, G. (2004). Mouse centric and pericentric satellite repeats form distinct functional heterochromatin. *J Cell Biol* 166, 493-505.
- Guilbaud, G., Rappailles, A., Baker, A., Chen, C.L., Arneodo, A., Goldar, A., d'Aubenton-Carafa, Y., Thermes, C., Audit, B., and Hyrien, O. (2011). Evidence for sequential and increasing activation of replication origins along replication timing gradients in the human genome. *PLoS Comput Biol* 7, e1002322.

- Hayano, M., Kanoh, Y., Matsumoto, S., Renard-Guillet, C., Shirahige, K., and Masai, H. (2012). Rif1 is a global regulator of timing of replication origin firing in fission yeast. *Genes Dev* 26, 137-150.
- Hiragami-Hamada, K., Soeroes, S., Nikolov, M., Wilkins, B., Kreuz, S., Chen, C., De La Rosa-Velazquez, I.A., Zenn, H.M., Kost, N., Pohl, W., *et al.* (2016). Dynamic and flexible H3K9me3 bridging via HP1beta dimerization establishes a plastic state of condensed chromatin. *Nat Commun* 7, 11310.
- Hiratani, I., Ryba, T., Itoh, M., Yokochi, T., Schwaiger, M., Chang, C.W., Lyou, Y., Townes, T.M., Schubeler, D., and Gilbert, D.M. (2008). Global reorganization of replication domains during embryonic stem cell differentiation. *PLoS Biol* 6, e245.
- Hiratani, I., Takebayashi, S., Lu, J., and Gilbert, D.M. (2009). Replication timing and transcriptional control: beyond cause and effect--part II. *Curr Opin Genet Dev* 19, 142-149.
- Jasencakova, Z., Scharf, A.N., Ask, K., Corpet, A., Imhof, A., Almouzni, G., and Groth, A. (2010). Replication stress interferes with histone recycling and predeposition marking of new histones. *Mol Cell* 37, 736-743.
- Johnson, R.E., Prakash, L., and Prakash, S. (2012). Pol31 and Pol32 subunits of yeast DNA polymerase delta are also essential subunits of DNA polymerase zeta. *Proc Natl Acad Sci U S A* 109, 12455-12460.
- Johnson, R.E., Washington, M.T., Haracska, L., Prakash, S., and Prakash, L. (2000). Eukaryotic polymerases iota and zeta act sequentially to bypass DNA lesions. *Nature* 406, 1015-1019.
- Julienne, H., Zoufir, A., Audit, B., and Arneodo, A. (2013). Human genome replication proceeds through four chromatin states. *PLoS Comput Biol* 9, e1003233.
- Koren, A., Polak, P., Nemesh, J., Michaelson, J.J., Sebat, J., Sunyaev, S.R., and McCarroll, S.A. (2012). Differential relationship of DNA replication timing to different forms of human mutation and variation. *Am J Hum Genet* 91, 1033-1040.
- Lang, G.I., and Murray, A.W. (2011). Mutation rates across budding yeast chromosome VI are correlated with replication timing. *Genome Biol Evol* 3, 799-811.
- Lange, S.S., Bhetawal, S., Reh, S., Powell, K.L., Kusewitt, D.F., and Wood, R.D. (2018). DNA polymerase zeta deficiency causes impaired wound healing and stress-induced skin pigmentation. *Life Sci Alliance* 1. doi: 10.26508/lsa.201800048
- Lange, S.S., Tomida, J., Boulware, K.S., Bhetawal, S., and Wood, R.D. (2016). The Polymerase Activity of Mammalian DNA Pol zeta Is Specifically Required for Cell and Embryonic Viability. *PLoS Genet* 12, e1005759.
- Lange, S.S., Wittschieben, J.P., and Wood, R.D. (2012). DNA polymerase zeta is required for proliferation of normal mammalian cells. *Nucleic Acids Res* 40, 4473-4482.
- Lee, Y.S., Gregory, M.T., and Yang, W. (2014). Human Pol zeta purified with accessory subunits is active in translesion DNA synthesis and complements Pol eta in cisplatin bypass. *Proc Natl Acad Sci U S A* 111, 2954-2959.
- Letessier, A., Millot, G.A., Koundrioukoff, S., Lachages, A.M., Vogt, N., Hansen, R.S., Malfoy, B., Brison, O., and Debatisse, M. (2011). Cell-type-specific replication initiation programs set fragility of the FRA3B fragile site. *Nature* 470, 120-123.
- MacAlpine, D.M., Rodriguez, H.K., and Bell, S.P. (2004). Coordination of replication and transcription along a Drosophila chromosome. *Genes Dev* 18, 3094-3105.
- Maison, C., Bailly, D., Peters, A.H., Quivy, J.P., Roche, D., Taddei, A., Lachner, M., Jenuwein, T., and Almouzni, G. (2002). Higher-order structure in pericentric heterochromatin involves a distinct pattern of histone modification and an RNA component. *Nat Genet* 30, 329-334.
- Makarova, A.V., and Burgers, P.M. (2015). Eukaryotic DNA polymerase zeta. *DNA Repair (Amst)* 29, 47-55.
- Makarova, A.V., Stodola, J.L., and Burgers, P.M. (2012). A four-subunit DNA polymerase zeta complex containing Pol delta accessory subunits is essential for PCNA-mediated mutagenesis. *Nucleic Acids Res.*
- Martin, S.K., and Wood, R.D. (2019). DNA polymerase zeta in DNA replication and repair. *Nucleic Acids Res.*

- McCulloch, S.D., and Kunkel, T.A. (2008). The fidelity of DNA synthesis by eukaryotic replicative and translesion synthesis polymerases. *Cell Res* 18, 148-161.
- Mendez-Bermudez, A., Lototska, L., Bauwens, S., Giraud-Panis, M.J., Croce, O., Jamet, K., Irizar, A., Mowinckel, M., Koundrioukoff, S., Nottet, N., *et al.* (2018). Genome-wide Control of Heterochromatin Replication by the Telomere Capping Protein TRF2. *Mol Cell* 70, 449-461 e445.
- Murzina, N., Verreault, A., Laue, E., and Stillman, B. (1999). Heterochromatin dynamics in mouse cells: interaction between chromatin assembly factor 1 and HP1 proteins. *Mol Cell* 4, 529-540.
- Natale, F., Scholl, A., Rapp, A., Yu, W., Rausch, C., and Cardoso, M.C. (2018). DNA replication and repair kinetics of Alu, LINE-1 and satellite III genomic repetitive elements. *Epigenetics Chromatin* 11, 61.
- Nguyen, V.H., and Lavenier, D. (2009). PLAST: parallel local alignment search tool for database comparison. *BMC Bioinformatics* 10, 329.
- Norio, P., Kosiyatrakul, S., Yang, Q., Guan, Z., Brown, N.M., Thomas, S., Riblet, R., and Schildkraut, C.L. (2005). Progressive activation of DNA replication initiation in large domains of the immunoglobulin heavy chain locus during B cell development. *Mol Cell* 20, 575-587.
- Northam, M.R., Moore, E.A., Mertz, T.M., Binz, S.K., Stith, C.M., Stepchenkova, E.I., Wendt, K.L., Burgers, P.M., and Shcherbakova, P.V. (2014). DNA polymerases zeta and Rev1 mediate error-prone bypass of non-B DNA structures. *Nucleic Acids Res* 42, 290-306.
- Ozeri-Galai, E., Tur-Sinai, M., Bester, A.C., and Kerem, B. (2014). Interplay between genetic and epigenetic factors governs common fragile site instability in cancer. *Cell Mol Life Sci* 71, 4495-4506.
- Peters, A.H., O'Carroll, D., Scherthan, H., Mechtler, K., Sauer, S., Schofer, C., Weipoltshammer, K., Pagani, M., Lachner, M., Kohlmaier, A., *et al.* (2001). Loss of the Suv39h histone methyltransferases impairs mammalian heterochromatin and genome stability. *Cell* 107, 323-337.
- Picard, F., Cadoret, J.C., Audit, B., Arneodo, A., Alberti, A., Battail, C., Duret, L., and Prioleau, M.N. (2014). The spatiotemporal program of DNA replication is associated with specific combinations of chromatin marks in human cells. *PLoS Genet* 10, e1004282.
- Polak, P., Karlic, R., Koren, A., Thurman, R., Sandstrom, R., Lawrence, M., Reynolds, A., Rynes, E., Vlahovicek, K., Stamatoyannopoulos, J.A., *et al.* (2015). Cell-of-origin chromatin organization shapes the mutational landscape of cancer. *Nature* 518, 360-364.
- Quah, S.K., von Borstel, R.C., and Hastings, P.J. (1980). The origin of spontaneous mutation in *Saccharomyces cerevisiae*. *Genetics* 96, 819-839.
- Ryba, T., Battaglia, D., Pope, B.D., Hiratani, I., and Gilbert, D.M. (2011). Genome-scale analysis of replication timing: from bench to bioinformatics. *Nat Protoc* 6, 870-895.
- Saveliev, A., Everett, C., Sharpe, T., Webster, Z., and Festenstein, R. (2003). DNA triplet repeats mediate heterochromatin-protein-1-sensitive variegated gene silencing. *Nature* 422, 909-913.
- Schultz, S.S., Desbordes, S.C., Du, Z., Kosiyatrakul, S., Lipchina, I., Studer, L., and Schildkraut, C.L. (2010). Single-molecule analysis reveals changes in the DNA replication program for the POU5F1 locus upon human embryonic stem cell differentiation. *Mol Cell Biol* 30, 4521-4534.
- Schuster-Bockler, B., and Lehner, B. (2012). Chromatin organization is a major influence on regional mutation rates in human cancer cells. *Nature* 488, 504-507.
- Schwienbacher, C., Angioni, A., Scelfo, R., Veronese, A., Calin, G.A., Massazza, G., Hatada, I., Barbanti-Brodano, G., and Negrini, M. (2000). Abnormal RNA expression of 11p15 imprinted genes and kidney developmental genes in Wilms' tumor. *Cancer Res* 60, 1521-1525.
- Seplyarskiy, V.B., Bazykin, G.A., and Soldatov, R.A. (2015). Polymerase zeta Activity Is Linked to Replication Timing in Humans: Evidence from Mutational Signatures. *Mol Biol Evol* 32, 3158-3172.
- Sima, J., Chakraborty, A., Dileep, V., Michalski, M., Klein, K.N., Holcomb, N.P., Turner, J.L., Paulsen, M.T., Rivera-Mulia, J.C., Trevilla-Garcia, C., *et al.* (2019). Identifying cis Elements for Spatiotemporal Control of Mammalian DNA Replication. *Cell* 176, 816-830 e818.

- Stamatoyannopoulos, J.A., Adzhubei, I., Thurman, R.E., Kryukov, G.V., Mirkin, S.M., and Sunyaev, S.R. (2009). Human mutation rate associated with DNA replication timing. *Nat Genet* 41, 393-395.
- Stone, J.E., Kissling, G.E., Lujan, S.A., Rogozin, I.B., Stith, C.M., Burgers, P.M., and Kunkel, T.A. (2009). Low-fidelity DNA synthesis by the L979F mutator derivative of *Saccharomyces cerevisiae* DNA polymerase zeta. *Nucleic Acids Res* 37, 3774-3787.
- Taddei, A., Maison, C., Roche, D., and Almouzni, G. (2001). Reversible disruption of pericentric heterochromatin and centromere function by inhibiting deacetylases. *Nat Cell Biol* 3, 114-120.
- Thiru, A., Nietlispach, D., Mott, H.R., Okuwaki, M., Lyon, D., Nielsen, P.R., Hirshberg, M., Verreault, A., Murzina, N.V., and Laue, E.D. (2004). Structural basis of HP1/PXVXL motif peptide interactions and HP1 localisation to heterochromatin. *EMBO J* 23, 489-499.
- Tomida, J., Takata, K.I., Bhetawal, S., Person, M.D., Chao, H.P., Tang, D.G., and Wood, R.D. (2018). FAM35A associates with REV7 and modulates DNA damage responses of normal and BRCA1-defective cells. *EMBO J* 37.
- Tsao, W.C., and Eckert, K.A. (2018). Detours to Replication: Functions of Specialized DNA Polymerases during Oncogene-induced Replication Stress. *Int J Mol Sci* 19.
- van Mierlo, G., Veenstra, G.J.C., Vermeulen, M., and Marks, H. (2019). The Complexity of PRC2 Subcomplexes. *Trends Cell Biol*.
- Van Sloun, P.P., Varlet, I., Sonneveld, E., Boei, J.J., Romeijn, R.J., Eeken, J.C., and De Wind, N. (2002). Involvement of mouse Rev3 in tolerance of endogenous and exogenous DNA damage. *Mol Cell Biol* 22, 2159-2169.
- Vieux-Rochas, M., Fabre, P.J., Leleu, M., Duboule, D., and Noordermeer, D. (2015). Clustering of mammalian Hox genes with other H3K27me3 targets within an active nuclear domain. *Proc Natl Acad Sci U S A* 112, 4672-4677.
- Watanabe, Y., Ikemura, T., and Sugimura, H. (2004). Amplicons on human chromosome 11q are located in the early/late-switch regions of replication timing. *Genomics* 84, 796-805.
- Watanabe, Y., and Maekawa, M. (2010). Spatiotemporal regulation of DNA replication in the human genome and its association with genomic instability and disease. *Curr Med Chem* 17, 222-233.
- Wittschieben, J., Shivji, M.K., Lalani, E., Jacobs, M.A., Marini, F., Gearhart, P.J., Rosewell, I., Stamp, G., and Wood, R.D. (2000). Disruption of the developmentally regulated Rev3l gene causes embryonic lethality. *Curr Biol* 10, 1217-1220.
- Wittschieben, J.P., Reshmi, S.C., Gollin, S.M., and Wood, R.D. (2006). Loss of DNA polymerase zeta causes chromosomal instability in mammalian cells. *Cancer Res* 66, 134-142.

EXPERIMENTAL PROCEDURES**KEY RESSOURCES TABLE****Antibodies**

Antibody	Source	Identifier
Mouse Monoclonal anti-Flag M2	Sigma-Aldrich	F1804
Mouse Monoclonal anti-flag M2 affinity gel	Sigma-Aldrich	A2220
Rabbit Polyclonal anti-Flag	Sigma-Aldrich	F7425
Rabbit Polyclonal anti-Histone H3 (tri methyl k4)	Abcam	ab8580
Goat anti-Histone H3	Abcam	ab12079
Rabbit polyclonal anti-Histone H4	Active Motif	39269
Rabbit Polyclonal anti-ATRX	Santa Cruz	sc-15408
Mouse Monoclonal anti-BrdU (for FACS) BU20a	Dako	M0744
Purified Mouse Anti-BrdU (for IF) clone B44	BD Biosciences	347580
Rabbit Polyclonal anti-Histone H3 (tri methyl k9)	Abcam	ab8898
Rabbit Polyclonal anti-acetyl Histone H3 (Lys 9/14)	Cell signaling technology	9677
Mouse Monoclonal anti-Histone H3 (tri methyl K27)	Abcam	ab6002
Rabbit Polyclonal anti-Histone H3 (tri methyl K36)	Abcam	ab9050
Mouse Monoclonal anti-Histone H4 (tri methy K20)	Active Motif	39672
Rabbit polyclonal anti-Histone H4 (acetyl K12)	Abcam	ab46983
Rabbit monoclonal anti-Histone H4 (acetyl K16)	Abcam	ab109463
Rabbit monoclonal anti-Histone H4 (acetyl K5)	Abcam	ab51997
Mouse monoconal anti-HP1 α	Millipore	05-689
Mouse monoconal anti-HP1 β	Abcam	ab10478
Mouse Monoclonal anti-phospho Histone H2A.X (ser139)	Millipore	05-636
Rabbit Polyclonal anti-53BP1	Abcam	ab21083
Rabbit monoclonal anti-Mad2L2/Rev7	Abcam	ab180579

Primers for ChIP-qPCR:

<u>Gene</u>	<u>Forward</u>	<u>Reverse</u>
Primers for ChIPqPCR		
<i>Hoxb2</i>	CCGAGTGAGTCCGTTTGGTT	GACTCCAGCCATGAGTAGCC
<i>Hoxb8</i>	GCCGGATGCAAATACCGAC	GAGAAGATGTGGGGTGGGTG
<i>Hoxb9</i>	CCCTTTCACAGGTGAGTCCC	GGTGTCCACAGGAAGAGCAA
<i>WT1</i>	GAGGGAGGGAGATGAGAGGC	CCTGCATCTCAGGGCACTTT
<i>FoxG1</i>	TGAGGACAGGCCAGGAAAAC	GTCAAGGCTTCCATGTGTGC
<i>Gata6</i>	CGACTTGGGAGGACCTGTTG	TGGTACATTTCTCCGGCTG
PRIMERS for screening KI mouse line		
<i>Rev3l</i>	ACGAGTTCGCGGGTCCTTAGAGGTC	ACTTTCTACAGCCACAGCATCTC

List of TaqMan Gene Expression Assays used in this study

(TaqMan Gene Expression Assays consist of forward and reverse primers with TaqMan minor groove binder probe for each gene)

Gene	Catalog #
<i>Hoxb2</i>	Mm04209931-m1
<i>Hoxb8</i>	Mm00439368-m1
<i>Hoxb9</i>	Mm01700220-m1
<i>WT1</i>	Mm01337048-m1
<i>Dlk1</i>	Mm00494477-m1
<i>Slc18a4</i>	Mm0045056-m1
<i>Dcn</i>	Mm0051435-m1
<i>H19</i>	Mm011567221-g1
<i>GAPDH</i>	Mm999999915-g1

Depository of microarray data from gene expression.

The microarray datasets analyzed in this study have been deposited at EBI Array express database with the Accession Number E-MTAB-8338.

Cell culture

Immortalized mouse embryonic fibroblast (MEF) *Rev3l*^{-/lox} and *Rev3l*^{+/-lox} (referred to in the text respectively as *Rev3l*^{-/-} and *Rev3l*^{+/-}) were previously described (Lange et al., 2012), and were cultivated in D-MEM (Dulbecco's Modified Eagle medium) containing 100 U/ml penicillin, 100

µg/ml streptomycin and 10% FCS in an atmosphere containing 5% CO₂. MEFs expressing functional REV3L with a 3X Flag epitope tag at the endogenous locus were derived from knock-in mice as described below. HeLa cells were cultivated in D-MEM containing sodium pyruvate, penicillin/streptomycin and 10% FCS in a 5% CO₂ atmosphere. Cells were regularly tested for mycoplasma.

Generation of knock-in mice expressing 3x FLAG tag REV3L

Ethics Statement: All animal work in this study was done according to The University of Texas, MD Anderson Cancer Center Institutional Animal Care and Use Committee guidelines, and approved by the MD Anderson Animal Care and Use Committee (IACUC).

Construction of the Targeting Vector. The targeting vector construction was designed and performed by genOway (Lyon, France). The *Rev3l* targeting vector was constructed from 129Sv/Pas mouse strain genomic DNA with a long (5.8 kb) homology arm upstream of exon 1, and a short (1.5 kb) homology arm downstream of exon 1. A 3x FLAG peptide sequence was inserted in frame with the ATG codon. A positive selection neomycin gene was flanked by loxP sites. The targeting vector also incorporated a diphtheria toxin negative selection cassette.

Screening of Rev3l targeted knock-in ES Cell clones. Linearized targeting vector was transfected into 129Sv ES cells (genOway, Lyon, France) according to genOway's electroporation procedures (i.e. 5 x 10⁶ ES cells with 40 µg of linearized plasmid, 260 V, 500 µF). Positive selection started 48 h after electroporation in medium containing 200 µg/ml of G418 (150 µg/ml of active component, Life Technologies, Inc.). *Rev3l* resistant clones were isolated and amplified in 96-well plates. Duplicates of 96-well plates were made. The set of plates containing ES cell clones amplified on gelatin were genotyped by both PCR and Southern blot analysis.

For PCR analysis, one primer pair was designed to validate the presence of the 3xFLAG tag within the 5' homology arm. This primer pair was designed to specifically amplify the targeted locus: Forward (Neo cassette): 5'- CCTGCTCTTTACTGAAGGCTCTTTACTATTGC-3'

Reverse: 5'- GGAACCCACAGTGGTTGTCCTAGTGC-3'

PCR products were sequenced in order to validate the presence of all elements. The targeted locus was confirmed by Southern blotting using a 3' probe (**Figure S4A**). Six clones were identified as correctly targeted and containing the 3x FLAG tag at the *Rev3l* locus.

Generation of mosaic mice and breeding scheme. Clones were microinjected into C57BL/6J blastocysts, and gave rise to male mosaics with a significant ES cell contribution (as determined by agouti coat color). Mice were bred to C57BL/6 mice expressing the Cre recombinase to remove the Neo cassette.

Genotyping of the knock-in mouse line. The following primers were used to monitor the Cre-mediated excision event and served for genotyping.

Forward: 5'- ACGAGTTCGCGGGTCCTTAGAGGTC-3'

Reverse: 5'- ACTTTCTACAGCCACAGCATCTCCGG-3'

The wild-type allele generates a product of 195 bp, the recombined allele 1958 bp, and the Cre-excised allele 289 bp.

Heterozygous mutant mice (which were maintained as pure C57BL/6 mice) were used to produce MEFs from embryos isolated at days 13.5 dpc. Seven heterozygous clones were obtained and T-antigen immortalized as described (Lange et al., 2012).

Plasmid construction

The human *Rev3l* full-length cDNA was kindly provided by C. Lawrence (University of Rochester Medical Center, Rochester, NY, USA). To generate the series of GFP-Rev3L plasmids, constructs were obtained by polymerase chain reaction (PCR) amplification from *Rev3l* cDNA as a XhoI-EcoRI fragment and cloned into peGFP-C3 expression vector (Clontech). Point mutants were generated by PCR-based methods using a QuikChange® site-directed mutagenesis kit (Stratagene) according to the manufacturer's instructions.

The cDNA encoding REV3L amino acids 700-900 was PCR amplified from REV3L/pETDuet-1 (Tomida et al., 2015) as a XhoI-NotI fragment with 5' REV3L (XhoI) primer and 3' REV3L (NotI) primers (5'- CCGCTCGAGATGAATACATTGGGCAAAAATTCTTTC and 5'- TAAAAGCGGCCGCTTATCCAAAGTGACAGTCTATAAAAC) to clone into pGEX6P-1 (GE Healthcare). The V802D mutation was introduced into pGEX6P-1 using a site-directed mutagenesis kit. The GST fusion protein was expressed and purified from *E. coli* as described (Tomida et al., 2015). The HP1 α (CBX5) and HP1 γ (CBX3) cDNAs were obtained from ThermoFisher Scientific. The cDNAs were cloned into vectors pETDuet-1 and pCDH-EF1 α -V5-MCS-IRES-Puro for His and V5-tagging, respectively. pCDH-EF1 α -V5-MCS-IRES-Puro was cloned into pCDH-EF1 α -V5-REV7-IRES-Puro from pRSFDuet-1 (Novagen) XhoI-NotI fragment. pCDH-EF1 α -V5-REV7-IRES-Puro was initially PCR amplified from pOZN REV7 (Tomida et al., 2018) with 5' V5-1st primer and 3' pOZ 3' primers (5'- ATTCTACGCTCGAGATGACCACGCTCACACGACAAG and 5'- CGGAATTGATCCGCTAGAG). Three subsequent rounds of PCR used the following primer sets: 5' V5 2nd primer and 3' pOZ 3' primers (5'- CTCTCCTCGGTCTCGATTCTACGCTCGAGATGACCACGC and 5'- CGGAATTGATCCGCTAGAG); 5' V5 3rd primer and 3' pOZ 3' primers (5'- AGCCTATCCCTAACCTCTCCTCGGTCTCGATTCTAC and 5'- CGGAATTGATCCGCTAGAG); 5' V5 4th primer (EcoRI) and 3' pOZ 3' primers (5'- GGAATTCATGGGTAAGCCTATCCCTAACCTCTCCTCG and 5'- CGGAATTGATCCGCTAGAG). The resulting V5-REV7 fragment (XhoI-NotI) was cloned into pCDH-EF1 α -MCS-IRES-Puro (System Biosciences). Mutations were introduced using a site-directed mutagenesis kit (Stratagene). His-XPA was as described (Manandhar et al., 2017). All

constructs were verified by DNA sequencing. Details on all constructs are available upon request.

Transfection

For transient transfection, plasmids were transfected using jetPEI (Polyplus), according to the manufacturer's instructions, and cells were processed 24-48 hrs later.

siRNAs were purchased from Dharmacon: siRNA *mRev3l* (ON-TARGETplus Smart pool), non-targeting siRNA (ON-TARGETplus non target pool). Cells were transfected with 30nM of siRNAs using Interferin (Polyplus) according to the manufacturer's instructions, and were processed 72h later.

Protein extraction, Immunoprecipitation and western blotting

Histone extraction was carried out in 10 volumes of CSK buffer (200 mM NaCl, 300mM sucrose, 3 mM MgCl₂, 10 mM TrisHCl pH 7, 1 mM EGTA, 0.5% Triton X-100, 1 mM PMSF, 10 mM Na-Butyrate) during 5min at 37°C. After centrifugation, the pellet was resuspended in 0.4 mM H₂SO₄ supplemented with Na-Butyrate and PMSF for 1 h, then centrifuged before adding acetone to supernatant and incubated overnight in cold room. After centrifugation, the pellet was resuspended in distilled water and histone extracts were kept at -80 C.

For immunoprecipitation, cells were lysed in NETN buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 1 mM EDTA, 0.5% NP40, anti-proteases) for 30 min on ice and sonicated twice at 29% for 10s. Samples were cleared by centrifugation for 5 min at 4°C. Immunoprecipitations were performed with Anti-FLAG M2 Agarose Beads (Sigma-Aldrich) overnight at 4°C on a wheel. Beads were extensively washed in NETN, with 300 mM NaCl for the final wash, and denatured in 2X Laemmli buffer.

Proteins were separated on 5% or 15% acrylamide SDS-PAGE, transferred on Nylon membrane (GE Healthcare) and detected with the indicated antibodies and ECL reagents (ThermoFisher Scientific).

GST- pull-down assay

GST-tagged REV3L fragments and His-tagged test proteins were incubated together with Glutathione-Sepharose 4B (GS4B) beads (GE Healthcare) at 4 °C for 4 h in 500 µl of buffer 0.1B (100 mM KCl, 20 mM Tris-HCl [pH 8.0], 5 mM MgCl₂, 10% glycerol, 1 mM PMSF, 0.1% Tween-20, 10 mM 2-mercaptoethanol). The beads were washed three times with 0.1B and eluted with 30 µl of 2 x SDS loading buffer (100 mM Tris-HCl [pH 6.8], 4% SDS, 0.2% bromophenol blue, 20% glycerol, 200 mM DTT).

Flow cytometry

In order to study the S phase progression in *Rev3L*^{-/-} and *Rev3L*^{+/+} fibroblasts, cells were incubated with 50 µM bromodeoxyuridine (BrdU) prior to harvest at different time points. Cells were collected, washed in PBS and fixed in 80% ice-cold ethanol overnight at -20°C. Total DNA was stained in 2.5 µg propidium iodide (IP) and 50 µg/µl RNase (Sigma) in PBS for 30 min at room temperature. Samples were analyzed on a C6 flow cytometer using the C6 Flow software (BD Accuri).

Immunofluorescence

For the visualization of DNA replication sites, cells seeded on glass coverslips were incubated with 50 µM of thymidine analog (IdU or BrdU) for 15 min, rinsed twice with cold PBS, then fixed with 4% paraformaldehyde at RT. Cells were then treated with 2 M HCl at 37°C for 20 min, then neutralized in 0.1 M borate buffer (pH 8.5). After washing in PBS, cells were incubated for 1 h at room temperature with anti-BrdU mAb (BD Biosciences). After washes, secondary antibodies coupled to Alexa Fluor was incubated for 30 min at RT. After washing three times with PBS, cells were mounted with mounting medium (Dako) supplemented with DAPI (Sigma).

For detection of heterochromatin-associated proteins and repressive histone marks, cells were pre-extracted or not with Triton-X 100 0.2% for 5 min on ice with gentle agitation then fixed with 4% paraformaldehyde. GFP-REV3L mutants were detected by autofluorescence and interested proteins were visualized by immunodetection with indicated primary antibodies for 1h at RT then processes as described above. Images were acquired using Axio Imager Z1 microscope (Zeiss).

Metaphase spreads

Cells were resuspended in hypotonic solution (0.05 M KCl), incubated for 15 min at 37°C, then fixed in an ethanol/acetic acid solution (3:1). Cell suspension was dropped on slides to obtain spread chromosomes. Metaphase slides were incubated with RNase A (100 µg/ml IN 2X SCC) for 1 h at 37°C and dehydrated in successive ethanol bath (70, 85 and 100%). After 2 washes in PBS at 55°C for 10 min and one wash at room temperature, coverslips were mounted in mounting medium (Dako) supplemented with DAPI (Sigma). Images were acquired using Axio Imager Z1 microscope (Zeiss).

Proximity Ligation Assay (PLA)

Mouse Embryonic Fibroblasts (MEFs) Flag-Rev3L were seed on coverslips for 24 hours. Cells were washed with PBS, protein pre-extraction was carried out with CSK 50mM NaCl (PLA Flag-Histones) or CSK 100mM NaCl (PLA DSB) for 5min on ice, fixed with 4% paraformaldehyde for

20 min, permeabilized with 0.5% Triton X-100 for 15 min and blocked with PBS-BSA 3% for at least 30 min. Coverslips were incubated with primary antibodies (see Table 1 for dilution) for 1h at RT. Proximity ligation was performed using the Duolink® In Situ Red Starter Kit Mouse/Rabbit (Sigma-Aldrich) according to the manufacturer's protocol. The oligonucleotides and antibody-nucleic acid conjugates used were those provided in the Sigma-Aldrich PLA kit (DUO92101). Images were quantified by counting the number of foci per nucleus using ImageJ software.

For the PLA post-irradiation, *Rev3L*^{-/-} and *Rev3L*^{+/+} fibroblasts were seeded on coverslips then 24 h later irradiated by X-rays (Xrad 320) at 2 Gy in PBS then put back in fresh medium for indicated time.

Southern blot

DNA methylation status at major satellites was performed on 5 µg of genomic DNA (gDNA) extracted from *Rev3L*^{-/-} and *Rev3L*^{+/+} fibroblasts and digested with the methylation sensitive enzyme MaeII (NEB) overnight at 65°C. After precipitation with EtOH, the digested gDNA was run on an agarose gel overnight and successively depurinated with a 250 mM HCL solution, denaturated with a solution containing 0.5 M NaOH and 1.5 M NaCl. After neutralization, the gDNA was transferred onto a Hybond-N+ membrane (GE Healthcare Amersham) overnight. The membrane was prehybridized with the hybridization buffer (6x SSC, 5X Denhardt's, 0.5% SDS) containing 100 µg/ml salmon sperm DNA for 1 h at 65°C and incubated overnight at 65°C with the hybridization buffer containing radiolabeled probes against major satellites (5'-GTGAAATATGGCGAGGAAAAC-3'). The membrane was washed once with 5x SSC, 0.1% SDS at 65°C for 10 min, twice with 3X SSC, 0.1% SDS at 65°C for 10 min and twice with 2X SSC, 0.1% SDS at room temperature for 10 min before visualization using X-ray film.

RNA extraction and RT-qPCR

Total RNAs were extracted from cell cultures, purified using RNeasy Mini Kit (QIAGEN), and concentration was measured with a Qubit Fluorometer (Invitrogen, Carlsbad, CA, USA). Total RNAs were then reverse transcribed using Superscript III (Invitrogen). cDNAs were used for quantitative PCR using Taqman Gene Expression Assays (Applied Biosystems) on an ABI Prism 7500 apparatus (Applied Biosystems). *GAPDH* was used as an internal control. The comparative threshold (Δ Ct) method was used.

ChIP qPCR

Rev3L^{-/-} and *Rev3L*^{+/+} fibroblasts were seeded at 1.5×10^6 and 10^6 cells per dish, respectively. After 24 hours, cells were cross-linked with 1% formaldehyde for 10 min at RT and

formaldehyde was quenched with Glycine 0.1M for 5 min at RT. Cells were washed twice with PBS + protease inhibitor, scraped in PBS + protease inhibitor then centrifuged at 800 x g – 4°C for 5 min. Cell pellets were lysed in 1 ml of Cell Lysis Buffer (20mM Tris-HCl [pH 8], 85 mM KCl, 0.5% NP40) for 15 min on ice and centrifuged at 800 x g at 4°C for 5 min. Pellets were then resuspended in 1 ml of Nuclei Lysis Buffer (Tris-HCl pH 8 50 mM, EDTA pH 8 10 mM, SDS 1%), sonicated using QSonica, Amplitude 40% for 10 cycles of 15 s ON/45 s OFF, and samples were centrifuged at 10000 x g – 4°C for 10 min. For each IP, supernatants were diluted 10 times in Dilution Buffer (0.01% SDS, 1.1% Triton X100, 1.2 mM EDTA, 16.7 mM Tris-HCl [pH 8], 150 mM NaCl) and incubated with 5 µg antibody, 20 µl Dynabeads Magnetic beads and 1% protease inhibitor cocktail overnight at 4°C with rotation. After magnetic separation, supernatant was removed and beads were washed successively with Low Salt Immune Complex Wash Buffer, High Salt Immune Complex Wash Buffer and LiCl Immune Complex Wash Buffer followed by magnetic separation. The last wash was performed in 200 µl TE Buffer for 3min at 4°C with rotation then beads were resuspended in 100 µl TE Buffer. Samples were then deproteinated by adding 2 µl Proteinase K and incubated at 55°C overnight with shaking. After proteinase K inactivation, beads were removed by magnetic separation to keep the supernatant. DNA purification was performed by phenol/chloroform extraction followed by ethanol precipitation and DNA resuspension in EB buffer. DNA samples were analyzed by quantitative real-time PCR on the Applied Biosystems Real Time PCR system (7500 system) using the Maxime SYBR Green/Rox qPCR Mastermix (ThermoFisher Scientific). The primers used are listed below.

Replication timing experiments and microarrays

Cells were incubated with 50 µM of BrdU for 1.5 h and collected, washed three times with DPBS (Dulbecco's Phosphate Buffered Saline, Gibco Life Technologies), and then fixed in ethanol 75%. Fixed cells were re-suspended in DPBS with RNase (0.5 mg/ml) and then with propidium iodide (50 µg/ml) followed by incubation in the dark at room temperature for 30 min with low agitation. Two fractions of 100,000 cells, S1 and S2 corresponding to Early and Late S-phase fractions respectively, were sorted by flow cytometry using INFLUX 500 (Cytocopia BD Biosciences). Whole DNA was extracted with a lysis buffer (50 mM Tris [pH 8], 10 mM EDTA, 300 mM NaCl, 0.5% SDS, and 0.2 mg/ml of Proteinase K) during 2h at 65°C. Neo-synthesized DNA were immunoprecipitated with BrdU antibodies (Anti-BrdU Pure, BD Biosciences, # 347580) as previously described (Fernandez-Vidal et al., 2014). To control the quality of enrichment of early and late fractions in S1 and S2, qPCR was performed with BMP1 oligonucleotides (early control) and with Dppa2 oligonucleotides (late control; data not shown, (Hiratani et al., 2008)). Microarray hybridization requires a minimum of 1000 ng of DNA. To obtain sufficient specific immunoprecipitated DNA for this hybridization step, whole genome

amplification was conducted (WGA, Sigma) on immunoprecipitated DNA. A post WGA qPCR was performed to confirm the maintaining of specific enrichment in both S1 and S2 fractions. Then early and late amplified neo-synthesized DNA were labeled with Cy3 and Cy5 ULS molecules, respectively (Genomic DNA labeling Kit, Agilent) as recommended by the manufacturer. The hybridization was performed according to the manufacturer instructions on 4×180K mouse microarrays (SurePrint G3 Mouse CGH Microarray Kit, 4x180K, AGILENT Technologies, reference genome: mm9). Microarrays were scanned with an Agilent High-Resolution C Scanner using a resolution of 3 µm and the autofocus option. Feature extraction was performed with the Feature Extraction 9.1 software (Agilent Technologies). For each experiment, the raw data sets were automatically normalized by the Feature extraction software. Analysis was performed with the Agilent Genomic Workbench 5.0 software. The log₂-ratio timing profiles were smoothed using the Triangular Moving Average option of the Agilent Genomic Workbench 5.0 software with the linear algorithm and 500 kb windows.

The raw data are available in geo database with this accession number GEXXXXXX (in progress).

Identification of replication domains and segments where replication timing profiles change.

To determine the replication domains in control cells and in *Rev3L*^{-/-} MEFs, the algorithms from CGH applications in the Agilent Genomic Workbench 5.0 software were exploited as previously described (Fernandez-Vidal et al., 2014). A comparison was conducted between early and late domains from both cell lines in order to determine segments where replication timing changes.

GC content

For the GC content, two steps were required: i) loading of the DNA sequence from interval files was performed with the Extract Genomic DNA tool with the mm9 genome annotation and ii) calculation of the GC content for each interval by the GeeCee EMBOSS tool installed on the GALAXY website.

Intersection and coverage with H3K27ac, H3K4me3 and H3K27me3 marks

Data are loaded from the UCSC website. Peak coordinates of H3K27me3 come from the Broad Institute in the mm8 genome annotation (http://genome-euro.ucsc.edu/cgi-bin/hgTrackUi?hgsid=199709818_uvGUfB28yATavzb9eAlxuHR5KGbQ&c=chr12&g=broadChromatinChIPSeq). The conversion of coordinates to mm9 genome annotation was performed with the Lift-over tool available in the GALAXY website (<https://usegalaxy.org/>). The peak coordinates of H3K4me3 and H3K27ac were obtained from the Ludwig Institute for Cancer Research via the UCSC website (<http://genome-euro.ucsc.edu/cgi->

bin/hgTrackUi?hgsid=208559602_SARngKJBy54VXgefriIgd8rJaWFD&c=chr12&g=wgEncodeLicrHistone) in MEF mouse cell line and then used knowingly for intersection and coverage with the Intersection and Coverage tools installed on the GALAXY website. The statistical analysis was performed with the R package.

LINE-1 content

We obtained the line1 coordinates from the RepeatMasker track via the UCSC web site. Intersection and coverage were performed with the Intersection and Coverage tools installed on the GALAXY website.

Next generation sequencing

A custom design Agilent © sure select capture kit (0730681) have been designed to catch the selected genomic region. 1 µg of genomic DNA was sheared with the Covaris S220 system to obtain 500 base Gaussian distribution using Peak power 140, Duty factor 10, Cycle/Burst 200 during 110 seconds (LGC Genomics/Kbioscience). DNA fragments were end-repaired, extended with an 'A' base on the 3' end, ligated with paired-end adaptors with the Bravo Platform (Agilent) and amplified (six cycles). Adaptor-ligated libraries were hybridized for 40 h with biotinylated oligo RNA baits using Custom Capture (Agilent), and enriched with streptavidin-conjugated magnetic beads. The final libraries were indexed (PCR 12 cycles), quantified by qPCR, pooled and sequenced using the onboard cluster method (V3), as paired-end sequencing (2x300 bp reads) on Illumina MiSeq sequencer at Gustave Roussy.

Analysis of mutational events

To count mutational events, FASTQ files were first cleaned using trimmomatic v0.32, with the MINLEN set to 36, SLIDINGWINDOW set to 4:15, TRAILING set to 3, LEADING set to 3 and ILLUMINACLIP set to TruSeq_adaptator_NextFlex_96_plus_revcomp.fa:2:30:10. Cleaned FASTQ files were then aligned to the selected genomic regions (chr2-73981219-89071252, chr9-47242391-52849181, and chr19-13221902-16369973) extracted from the July 2007 mouse (*Mus musculus*) genome (mm9/NCBI37) using bwa-mem (v.0.7.10), by marking shorter split hits as secondary (M), with a matching score (-A) of 2, with a gap extension penalty (-E) of 1. To get a summary of the coverage of mapped reads on a reference sequence at a single base pair resolution, we used samtools mpileup (v1.1) without skipping anomalous read pairs in variant calling (-A), disabling probabilistic realignment for the computation of base alignment quality (-B) and with a minimum mapping quality for an alignment (-q) of 20. The mpileup files were transformed into vcf files with a home-made script written in python (v2.7.3), the vcf files were then used to count for mutational events (insertions, deletions, transversions and transitions)

occurring in each reads (without any filtering on depth). Since we are counting each variants occurring in each single reads, we have to normalize the mutation events counts by the library size. Furthermore, to limit the number of sequencing artifact, all libraries have been sequenced twice. Only common variations between the two sequencing run were selected, by comparing the vcf file (previously indexed with tabix) using bcftools. The distribution of the number of mutational events have been performed using the non-parametric Wilcoxon rank sum test.

Gene expression profiling data analysis

RNA was extracted using Qiagen column (RNeasy kit) according to the manufacturer's protocol's. The quantity and purity of the total RNA were evaluated using a NanoDrop spectrophotometer. Quality of RNA samples was assessed using Lab-on-a-chip Bioanalyser 2000 technology (Agilent Technologies), based on the 28S/18S ribosomal RNAs ratio. All samples included in this study displayed a ratio of ribosomal RNAs between 1.5 and 2. Labeling of RNA samples was done according to Agilent oligo Cy5 or Cy3 probes labeling protocol using the Agilent Low Input QuickAmp labelling kit (ref 5190-2306) adapted for small amount of total RNA (100 ng total RNA per reaction). Hybridization was performed using the Agilent Hybridization Protocol (Gene expression Hyb kit Large - ref 5188-5280). Scanning was performed with an Agilent G2505C DNA Microarray scanner using default parameters: 20 bits mode, 3 um resolution, at 20 C in low ozone concentration environment. Microarray images were analyzed by using Feature extraction software version 10.7.3.1 from Agilent technologies and Agilent normalization protocol GE2_107_Sep09 with 028005_D_F_20130207 as design. Default settings are used. A quantile array normalization was performed on raw MedianSignal values with in house scripts using Bioconductor LIMMA package. Controls and flags spots were filtered. Missing values were completed with KNN imputation method. Then the median of all probes for a given transcript was computed to summarize the data. Normalized data (intensities in log2) were imported in BrB Arrays Tools (<https://brb.nci.nih.gov/BRB-ArrayTools/>). Class Comparison was performed as weighted *t*-test and an FDR adjustment (0.05) for multiple testing between *Rev3l* wt and KO samples (Class Comparison in BrB Arrays Tools 4.6.0) [https://brb.nci.nih.gov/BRB-ArrayTools]. Gene expressions altered by at least three-fold between conditions were considered.

REFERENCES

Fernandez-Vidal, A., Guitton-Sert, L., Cadoret, J.C., Drac, M., Schwob, E., Baldacci, G., Cazaux, C., and Hoffmann, J.S. (2014). A role for DNA polymerase theta in the timing of DNA replication. *Nat Commun* 5, 4285.

Manandhar, M., Lowery, M.G., Boulware, K.S., Lin, K.H., Lu, Y., and Wood, R.D. (2017). Transcriptional consequences of XPA disruption in human cell lines. *DNA Repair (Amst)* 57, 76-90.

Tomida, J., Takata, K., Lange, S.S., Schibler, A.C., Yousefzadeh, M.J., Bhetawal, S., Dent, S.Y., and Wood, R.D. (2015). REV7 is essential for DNA damage tolerance via two REV3L binding sites in mammalian DNA polymerase zeta. *Nucleic Acids Res* 43, 1000-1011.

Tomida, J., Takata, K.I., Bhetawal, S., Person, M.D., Chao, H.P., Tang, D.G., and Wood, R.D. (2018). FAM35A associates with REV7 and modulates DNA damage responses of normal and BRCA1-defective cells. *EMBO J* 37.

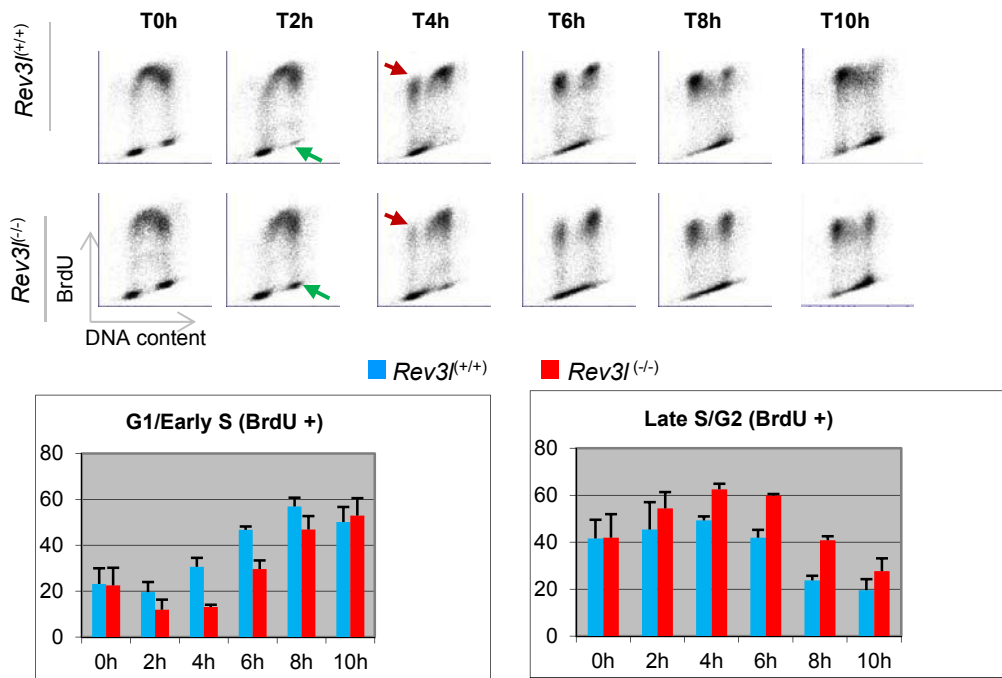
Table 1. Loci displaying both delay in RT and misregulation in *Rev3I*^{-/-} MEFs

Domain	Chr	Size (bp)	Replication timing domain	Gene	Misregulation (fold change*)
1	1	1 042,476	TTR	Pax3	-3.57
2	2	2 573 606	TTR	Plcb4	-4.17
3	6	6 628 128	TTR	Frmd4b	-3.13
4	8	671 566	TTR	Hand2	-7.69
5	8	641 165	TTR	Lpl	-27.03
6	10	567 222	TTR	Ptprb	-5.56
6	10	567 222	TTR	Tspan8	-7.14
7	11	920 184	TTR	Meis1	-5.56
8	12	487 988	TTR	Foxg1	-20.41
9	12	548 987	TTR	Dlk1	-6.25
9	12	548 987	TTR	Meg3	-5.26
9	12	548 987	TTR	Mirg	-6.67
10	12	1 372 498	TTR	Mycn	-5.88
11	13	1 179 925	TTR	Sfrp4	-6.25
12	14	671 779	TTR	Wnt5a	-4.55
13	15	1 054 278	TTR	Fbxl7	-8.33
14	15	1 145 389	TTR	Has2	-4.55
15	17	808 766	TTR	Rftn1	-12.35
16	18	402 349	TTR	Gata6	-12.82
17	18	539 392	TTR	Pcdhb20	-3.45
17	18	539 392	TTR	Pcdhb21	-5.56
18	18	1 232 897	TTR	Ppp2r2b	-8.33
19	19	509 162	TTR	Aldh1a1	-11.33
19	19	509 162	TTR	Aldh1a7	-3.57

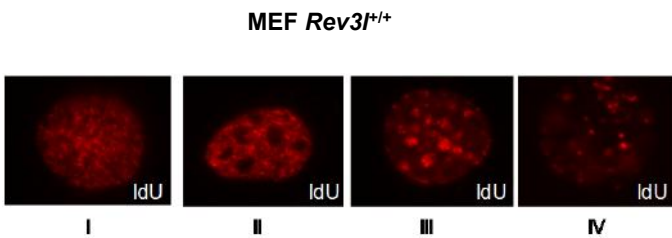
* FDR<0.05

Figure 1

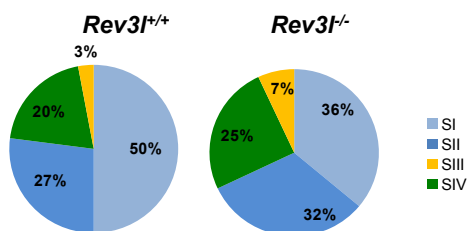
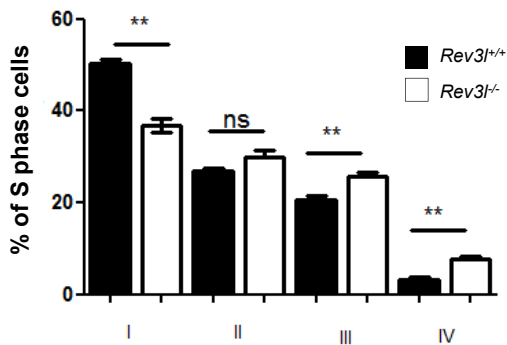
A



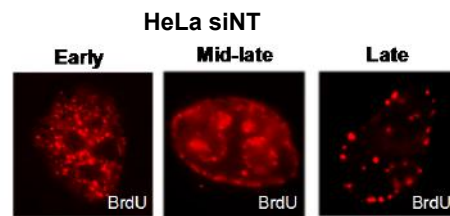
B



Distribution of cells in S phase



C



Distribution of cells in S phase

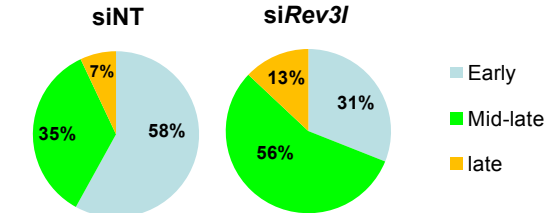
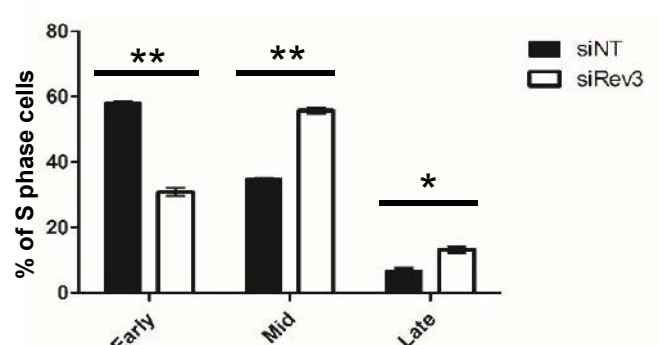


Figure 1. S phase progression is impaired in *Rev3l*^{-/-} MEFs

(A) *Rev3l*^{+/+} and *Rev3l*^{-/-} MEFs were seeded and pulse-labelled with BrdU prior to harvesting and analyzed by flow cytometry at different time points. The analysis was focused on S phase divided in 3 parts: G1/early S, middle S and late S/G2 BrdU⁺ cells. Histograms represent the percentage of cells in G1/early and late S/ G2 phase after the BrdU pulse. Error bars represent standard error for three independent experiments. **(B)** *Rev3l*^{+/+} and *Rev3l*^{-/-} MEFs were pulse-labelled with IdU during 10min, fixed and stained for IdU (red). S-phase sub-stages from I to IV were evaluated by visual inspection of the cycling population. Histograms and pie charts show the relative proportion (percentage of total S) of each sub-stage from I to IV. Error bars indicate standard error of the mean for two independent experiments (n=2). *, P<0.05 and ** P<0.01 by Student's *t*-test. **(C)** After 72h of transfection with non-target (NT) or *Rev3l* siRNA, HeLa cells were pulse-labelled with BrdU for 10 min, fixed and stained for BrdU (red). As previously, S-phase sub-stages were evaluated by visual inspection of the cycling population. Histograms and pie charts show the relative proportion (percentage of total S) of early, middle-late and late S-phase. Error bars indicate standard error of the mean for three independent experiments (n=3). *, P<0.05, ** P<0.01 and *** P< 0.001 by Student's *t*-test.

Figure 2

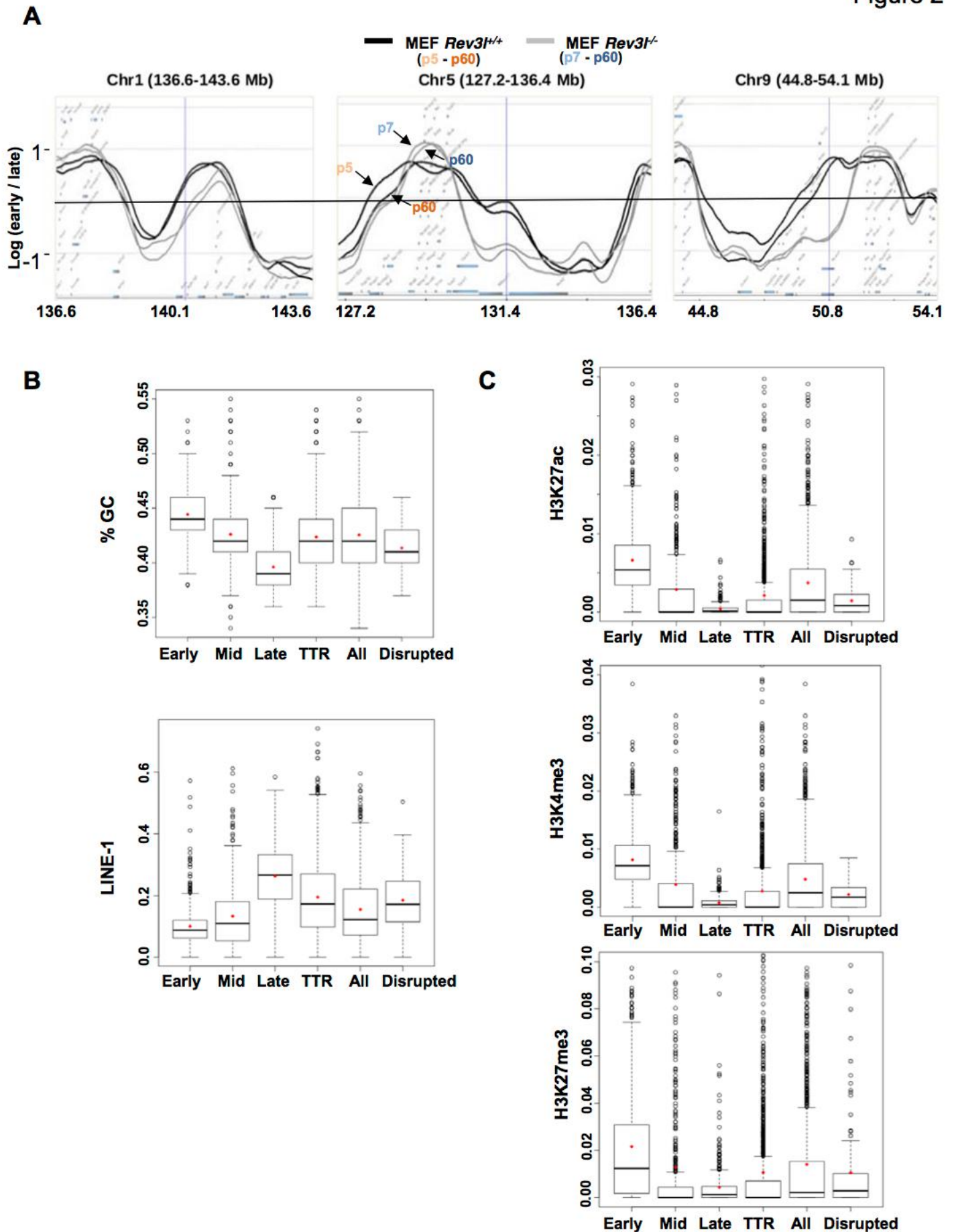
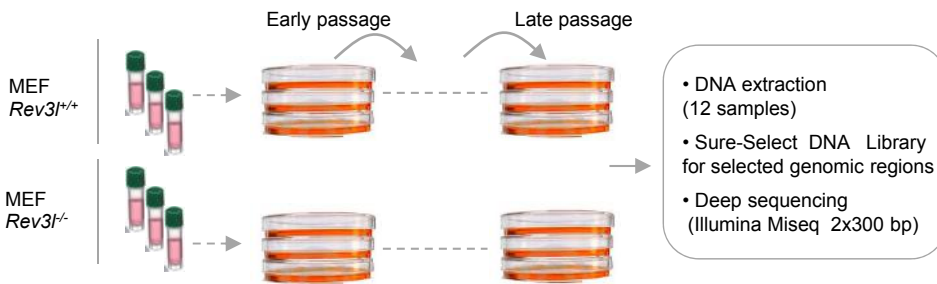


Figure 2. Loss of *Rev3l* disrupts replication timing in specific genomic loci.

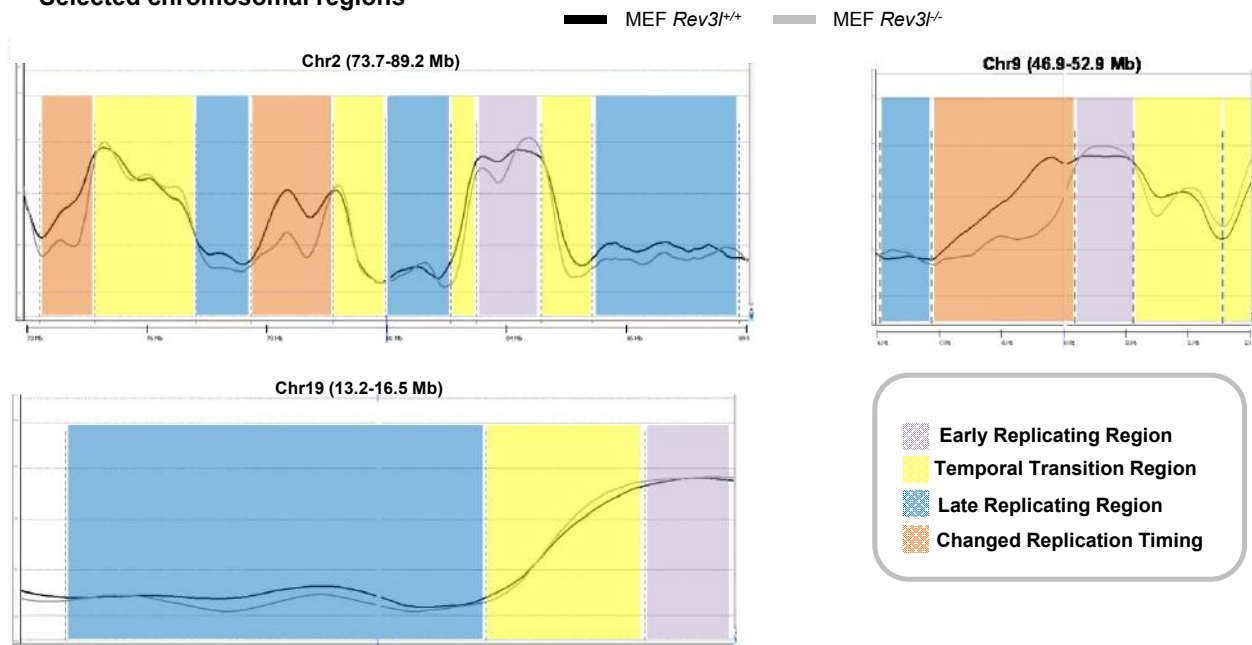
(A) Cells were pulse-labelled with BrdU for 1h30 and sorted by flow cytometry in two fractions, S1 and S2, corresponding to early and late S-phase fractions, respectively. Neo-synthesized DNA was immunoprecipitated with BrdU antibodies. Early and late neo-synthesized DNAs were labeled with Cy3 and Cy5 and hybridized on microarrays. After processing analyzing as described in methods, replication-timing profiles can be obtained. Shown are the zoomed microarray profiles of the timing of replication on chromosome 1 (136.6-143.6 Mb), chromosome 5 (127.2-136.4 Mb) and chromosome 9 (44.8-54.1 Mb) from *Rev3l*^{+/+} and *Rev3l*^{-/-} MEFs overlaid. Black lines represent replication timing from *Rev3l*^{+/+} MEFs and grey lines represent *Rev3l*^{-/-} MEFs. For each group, replication timing at early passage (p5 or p7) and late passage (p60) were analyzed (corresponding to double line on zoomed microarray profiles for each cell line). **(B)** Analysis of GC and Line-1 content in early, mid, late, TTR and all domains in mouse genome compared to changed replicating regions (disrupted) found in *Rev3l*^{-/-} MEFs. **(C)** Analysis of the active marks H3K27ac and H3K4me3 and facultative heterochromatin mark H3K27me3 content in early, mid, late, TTR and all domains in mouse genome compared to changed replicating regions (disrupted) found in *Rev3l*^{-/-}.

Figure 3

A Experimental procedure



B Selected chromosomal regions



C

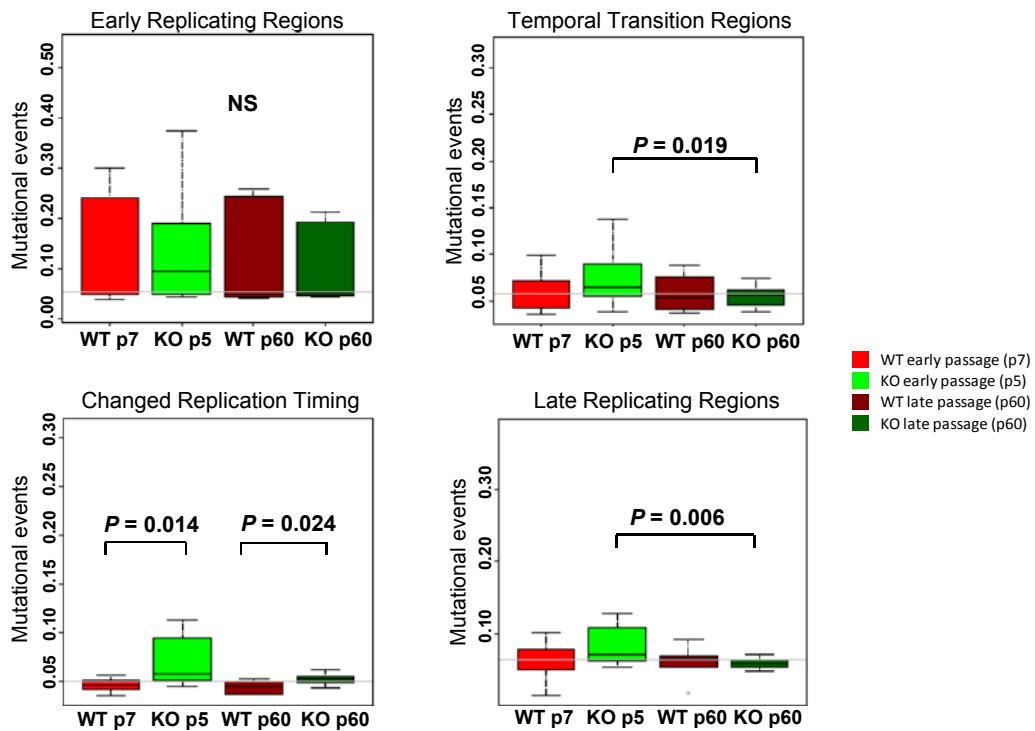


Figure 3. Delay in replication timing correlate with an increase of mutational events in *Rev3l*^{-/-} MEFs

(A) Schematic representation of the experimental procedure. Three independent cultures of *Rev3l*^{+/+} and *Rev3l*^{-/-} MEFs were propagated for few passages (p7 for *Rev3l*^{+/+} and p5 for *Rev3l*^{-/-}) or many passages (p60 for both cell lines). DNA was extracted from these 12 samples then selected genomic regions were deep sequenced. **(B)** Zoomed microarray profiles of the replication timing from selected genomic regions that have been sequenced. These genomic regions were sub-divided in Early (purple), Late (blue), Temporal Transition (yellow) regions according to their replication timing. Changed replication Timing (Orange) regions represent regions displaying delays in *Rev3l*^{-/-} cells. **(C)** Total mutational events (insertions, deletions, transversions and transitions) normalized by reads are shown as boxplots for each category (early, late, temporal transition regions and changed replication timing regions according to their replication timing). The horizontal black line in the box represents the median; p-values are for comparisons of median values (Wilcoxon rank sum test), ns: not significant.

Figure 4

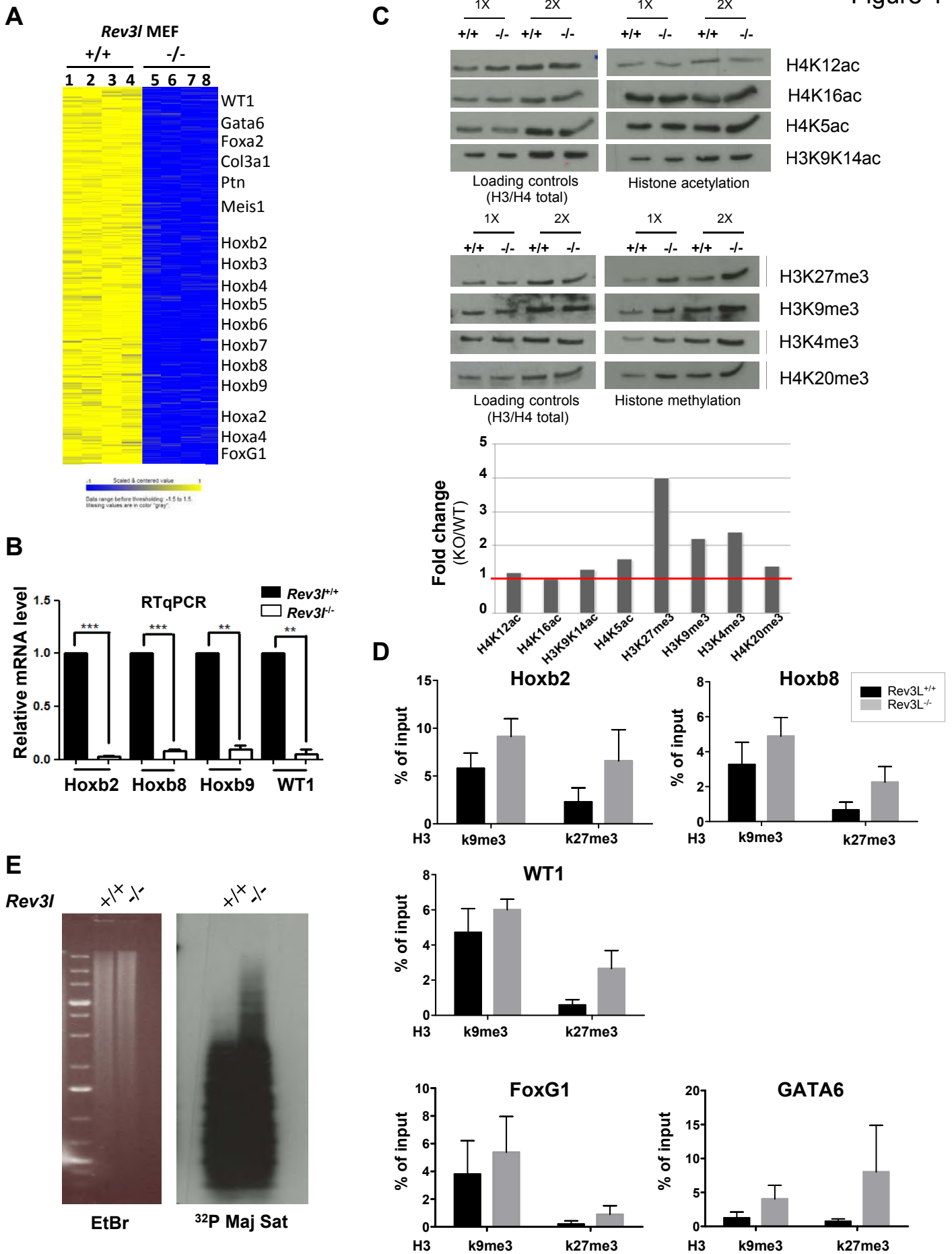


Figure 4. Loss of *Rev3l* results in epigenetic changes and down-regulation of numerous developmentally regulated genes

(A) Heat map showing log₂ fold change in differentially expressed genes in *Rev3l*^{+/+} and *Rev3l*^{-/-} MEFs (from 2 independent biological experiments with 2 technical replicates). Several repressed developmentally regulated genes in *Rev3l*^{-/-} MEFs are indicated on the left. Yellow and blue indicate high and low mRNA expression levels, respectively. **(B)** Relative mRNA levels of four genes involved in development (*Hoxb2*, *8*, *9* and *WT1*) were validated through qRT-PCR from *Rev3l*^{+/+} and *Rev3l*^{-/-} MEFs samples. The data were normalized to the amount of GAPDH mRNA. Error bars indicate standard error of the mean for three independent experiments. ***, P<0.001 and **, P<0.005 by Student's *t*-test. **(C)** Total histones were acid extracted from *Rev3l*^{+/+} and *Rev3l*^{-/-} MEFs. Samples were analyzed by western-blot using indicated antibodies specific to histone marks. An anti-pan H4 or H3 was used as loading control. Histograms show WT/KO *Rev3l* fold change. **(D)** Histone mark levels were examined at selected genes repressed in *Rev3l*^{-/-} MEFs (*Hoxb2*, *Hoxb8*, *WT1*). Histograms represent enrichment of H3k9me3 and H3k27me3 at indicated loci assessed by CHIP-qPCR in *Rev3l*^{+/+} and *Rev3l*^{-/-} MEFs. Each experiment was repeated at least three times. Histone mark levels were also evaluated at selected genes that localized in genomic loci displaying replication timing delay and downregulated in *Rev3l*^{-/-} MEFs (*FoxG1*, *GATA6*). Histograms represent enrichment of H3K9me3 and H3K27me3 assessed by CHIP-qPCR at indicated loci in *Rev3l*^{+/+} and *Rev3l*^{-/-} MEFs. Each experiment was repeated at least three times. **(E)** Southern blot analysis of genomic DNA extracted from *Rev3l*^{+/+} and *Rev3l*^{-/-} MEFs and restricted with the methylation-sensitive enzyme MaeII (5' -ACGT- 3'). Blot was hybridized with radiolabeled probes specific to major satellites.

Figure 5

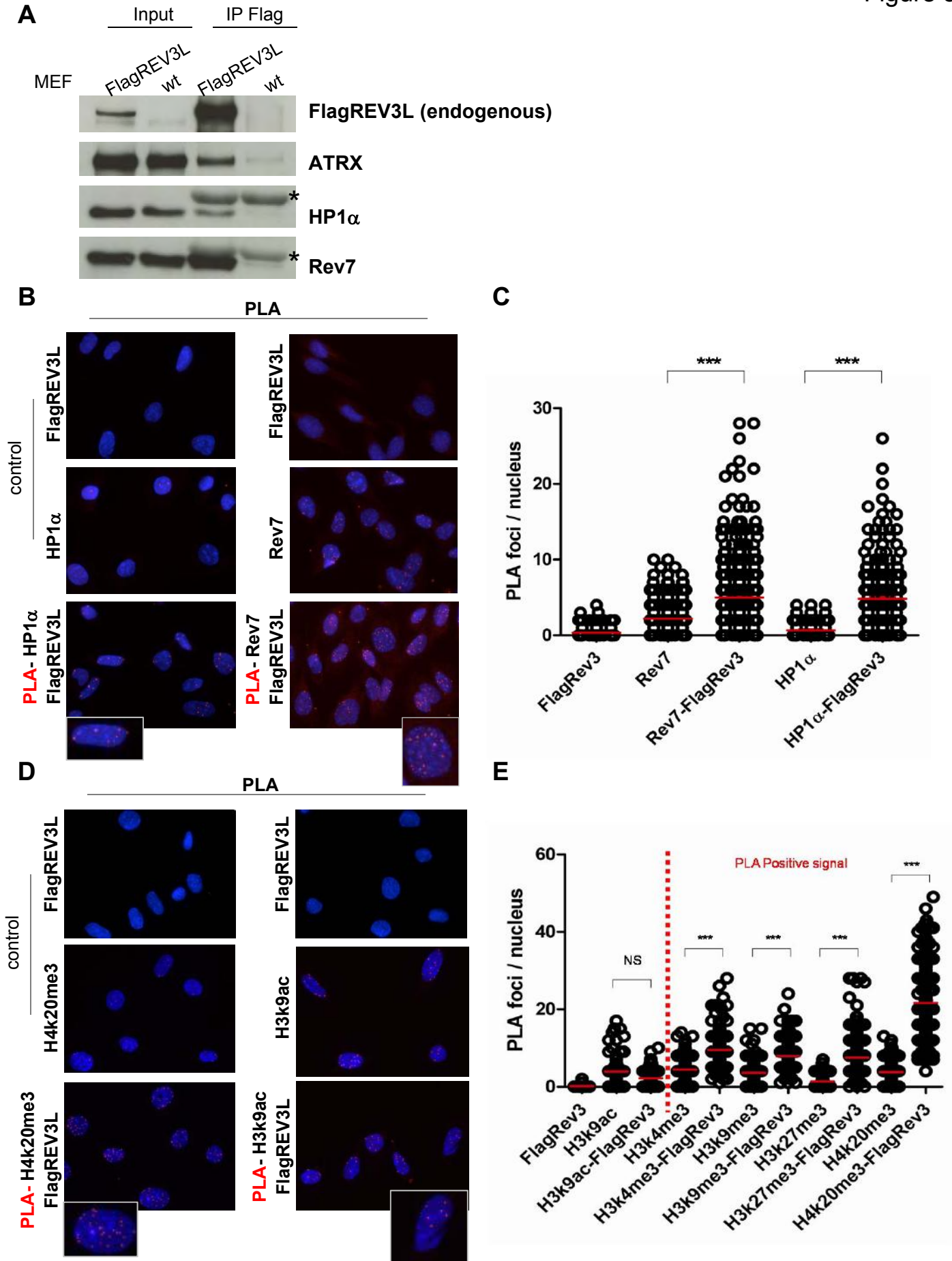


Figure 5: REV3L interacts with heterochromatin components and repressive histone marks

(A) Asynchronous MEFs expressing endogenous Flag-REV3L were lysed and REV3L were immunoprecipitated using anti-Flag (M2) antibodies. Co-immunoprecipitated proteins were analyzed by immunoblotting using the indicated antibodies. *: IgG light chain **(B)** Representative images of PFA fixed MEFs subjected to *in situ* proximity ligation assay (PLA). Cells showed PLA interactions (red fluorescent spots) between REV3L and HP1 α (lower panel, on the left), and between REV3L and REV7 (lower panel, on the right). Upper and middle panels are shown as PLA control (only one antibody used for the procedure). **(C)** Images from each PLA condition were quantified, counting the number of interactions per nucleus. Experiment was repeated three times (n>150, ns: not significant, ***: p<0.001, Mann-Whitney test). **(D)** Representative images as (A) showing PLA interactions between REV3L and H4K20me3 (lower panel, on the left), but not between REV3L and H3K9ac (lower panel, on the right). Upper and middle panels are shown as PLA control (only one antibody used for the procedure). Note that H3K9ac antibody alone made a high background **(E)** Images from each PLA condition including H4K20me3, H3K9me3, H3K27me3, H3K4me3 and H3K9ac were quantified, counting the number of interactions per nucleus. Experiment was repeated at least three times (n>150, ns: not significant, ***: p<0.001, Mann-Whitney test).

Figure 6

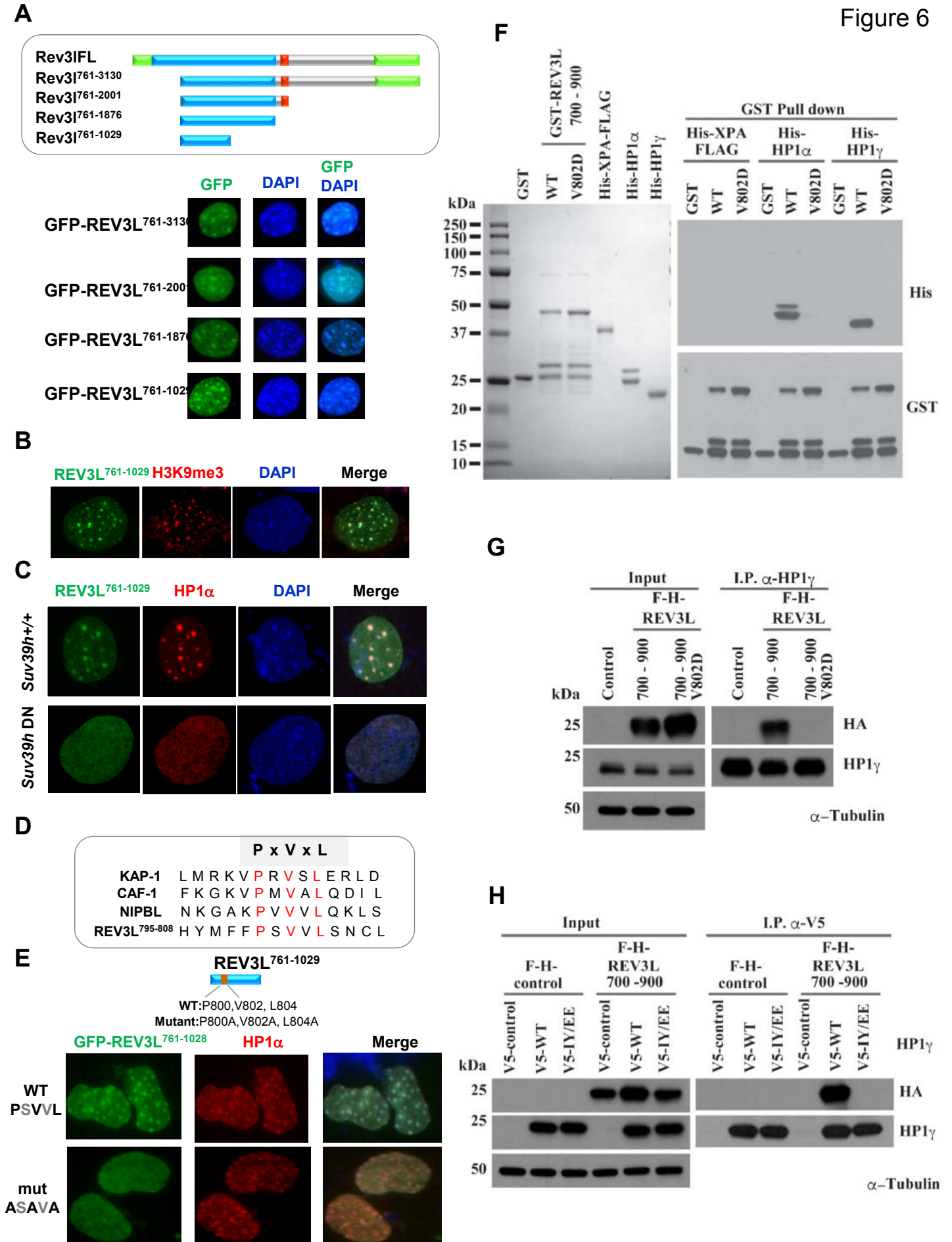


Figure 6. REV3L localizes to pericentromeric heterochromatin through a direct interaction with HP1 dimer

(A) Schematic representation of the human wild-type catalytic subunit of Pol ζ , REV3L and truncated constructs. Conserved domains between yeast and human REV3L in green are the N-terminal domain (NTD, 1-333 aa) and the catalytic domain (2276-3130 aa), the REV7 interacting domains are in red (1880-2001 aa). A large region in blue not found in yeast protein is almost exclusively encoded by exon 14. All the constructs lack the NTD domain and were fused to GFP. MEFs cells were transiently transfected with various GFP-REV3L constructs and fixed with 4% formaldehyde. The distribution of the GFP-REV3L mutants was detected by autofluorescence and nuclei were visualized using DAPI staining. **(B)** MEFs cells were transiently transfected with GFP-REV3L⁷⁶¹⁻¹⁰²⁹ and fixed with 4% formaldehyde. The distribution of GFP-REV3L⁷⁶¹⁻¹⁰²⁹ was detected by autofluorescence, chromocenters were visualized by H3K9me3 immunostaining (red) and nuclei were visualized by DAPI staining. **(C)** WT cells and MEFs lacking Suv39h (*Suv39h* dn) were transiently transfected with GFP-REV3L⁷⁶¹⁻¹⁰²⁹ construct then fixed with 4% formaldehyde after 0.5% Triton X-100 pre-extraction. The distribution of GFP-REV3L⁷⁶¹⁻¹⁰²⁹ was detected by autofluorescence, HP1 α immunodetection was processed using specific antibodies (red) and nuclei were visualized by DAPI staining. **(D)** Sequence alignment of proteins containing the PxVxL motif important for an interaction with HP1, with canonical residues shown in red. **(E)** Schematic representation of REV3L⁷⁶¹⁻¹⁰²⁹ and punctual mutations made for this study (PxVxL mutant: P800A, V802A, L804A) is shown (top panel). MEFs cells were transfected with GFP-REV3L⁷⁶¹⁻¹⁰²⁹ WT PSVVL or ASAVA mutant. Cells were fixed with 4% formaldehyde, the distribution of GFP-REV3L constructs was detected by autofluorescence and nuclei were visualized by DAPI staining (bottom panel). **(F)** GST, GST-REV3L⁷⁰⁰⁻⁹⁰⁰, mutant V802D FH-REV3L⁷⁰⁰⁻⁹⁰⁰, His-XPA-FLAG, His-HP1 α and His-HP1 γ were purified from *E. coli* and stained with Coomassie brilliant blue. These were used with glutathione beads for GST pulldown experiments. After electrophoresis, samples were immunoblotted with anti-His or anti-GST as indicated. **(G)** Human 293T cells were transfected with either FH-REV3L⁷⁰⁰⁻⁹⁰⁰, mutant V802D FH-REV3L⁷⁰⁰⁻⁹⁰⁰ or empty vector. HP1 γ -IP of 293T cells transfected. 48 h after transfection cell lysates were made and used for immunoprecipitation with HP1 antibody. Western blot was processed and membranes were immunoblotted with the indicated antibodies. **(H)** FH-REV3L⁷⁰⁰⁻⁹⁰⁰ or empty vector was co-transfected with V5-HP1 γ , mutant IY165-168EE V5-HP1 or empty vector into 293T cells. 48 h after transfection cell lysates were made and used for immunoprecipitation with anti-V5 antibody.

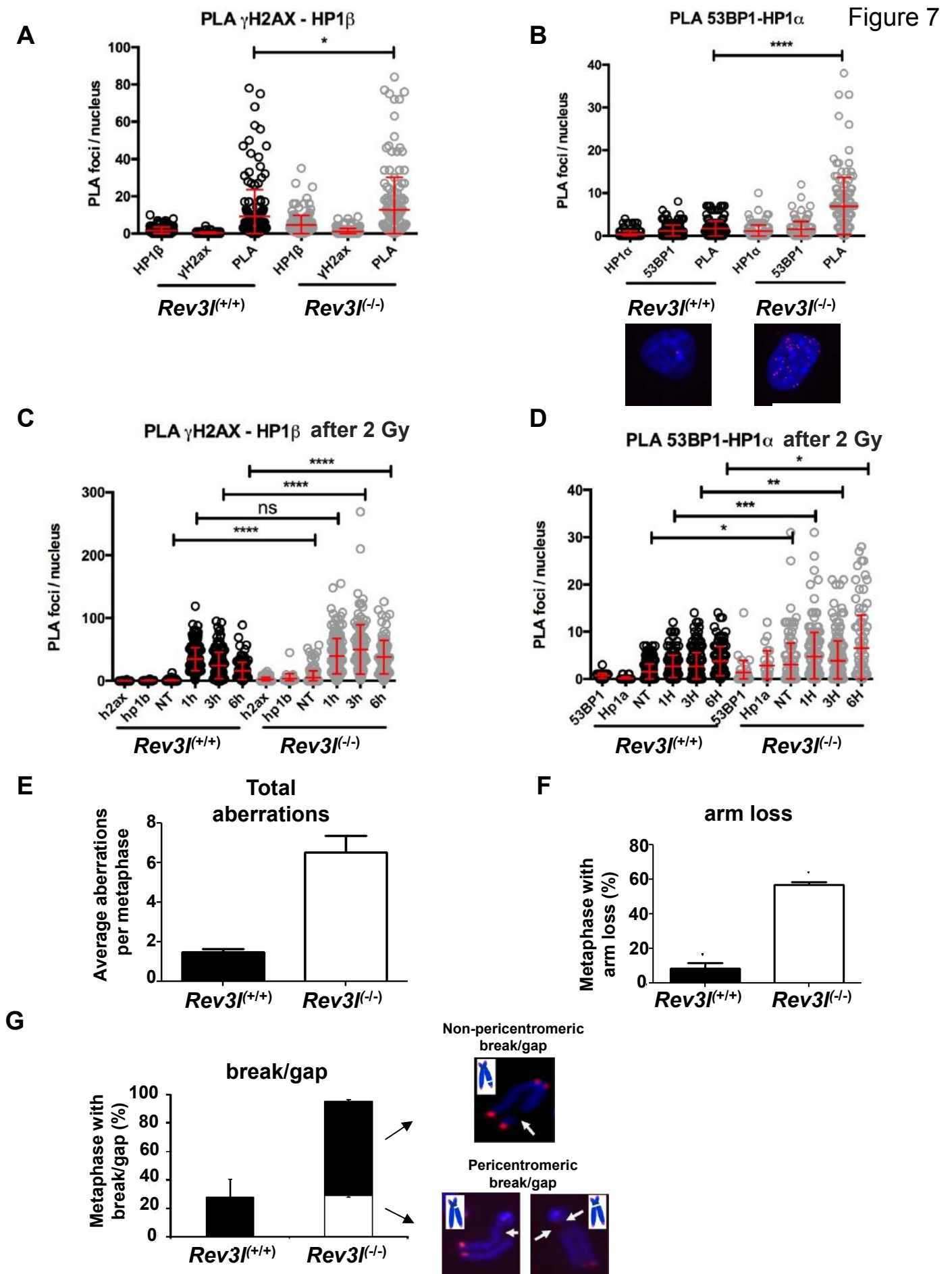
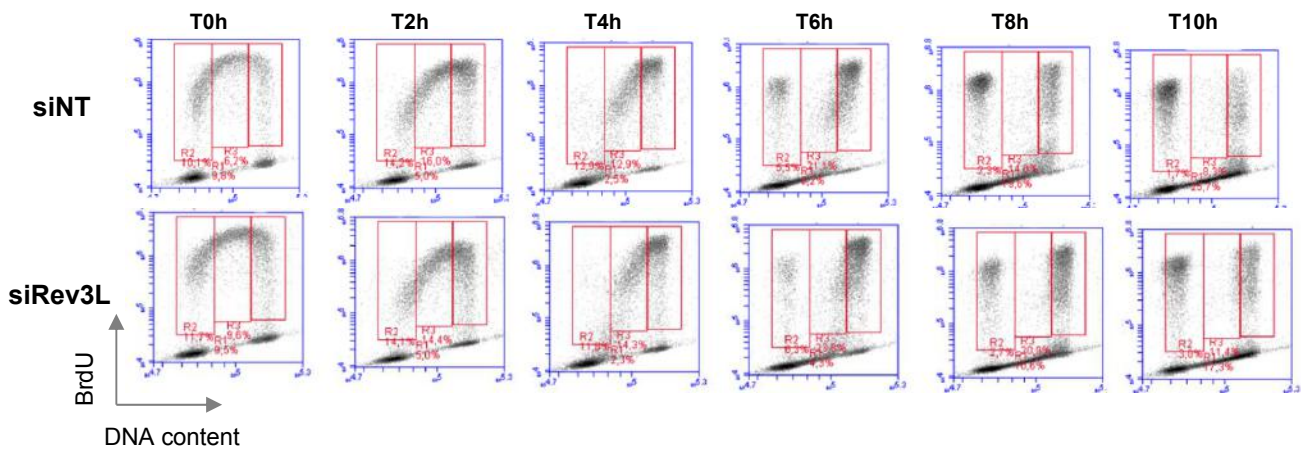


Figure 7. Inactivation of *Rev3l* causes spontaneous and X-rays-induced heterochromatin-associated DSBs

(A) Asynchronous *Rev3l*^{+/+} and *Rev3l*^{-/-} MEFs were fixed with PFA then subjected to *in situ* proximity ligation assay (PLA) using γ H2AX and HP1 β antibodies. PLA foci indicating interaction between proteins were counted in both *Rev3l*^{+/+} and *Rev3l*^{-/-} cell lines. **(B)** PLA was processed as in (A) by using 53BP1 and HP1 α antibodies then PLA foci were counted in both *Rev3l*^{+/+} and *Rev3l*^{-/-} cell lines. Representative images are shown. Experiment was repeated three times (n>150, *: p<0.05, ****: p<0.0001, Mann-Whitney test). **(C-D)** Asynchronous *Rev3l*^{+/+} and *Rev3l*^{-/-} MEFs were irradiated at 2 Gy, fixed with PFA at different time point then subjected to *in situ* proximity ligation assay (PLA) using γ H2AX and HP1 β antibodies (C) or 53BP1 and HP1 α antibodies (D). PLA foci were counted in both *Rev3l*^{+/+} and *Rev3l*^{-/-} cell lines. Experiment was repeated three times (n>150, *: p<0.05, ****: p<0.0001, Mann-Whitney test). **(E-H)** Metaphase were spreads from *Rev3l*^{+/+} and *Rev3l*^{-/-} MEFs. Histograms showed the quantitative analysis of **(E)** average aberrations per metaphases in *Rev3l*^{-/-} versus *Rev3l*^{+/+} MEFs. **(F)** Percentage of metaphases associated with long arms loss in *Rev3l*^{-/-} compared to *Rev3l*^{+/+} MEFs. **(G)** Percentage of metaphases with breaks/gaps in *Rev3l*^{+/+} versus *Rev3l*^{-/-} MEFs (left). Representative images of *Rev3l*^{-/-} MEFs with pericentromeric and non-pericentromeric breaks/gaps. Error bars represent standard error for 3 independent experiments.

SUPPLEMENTAL INFORMATION

Supplemental Figure 1



Cells in different fractions of the S-phase

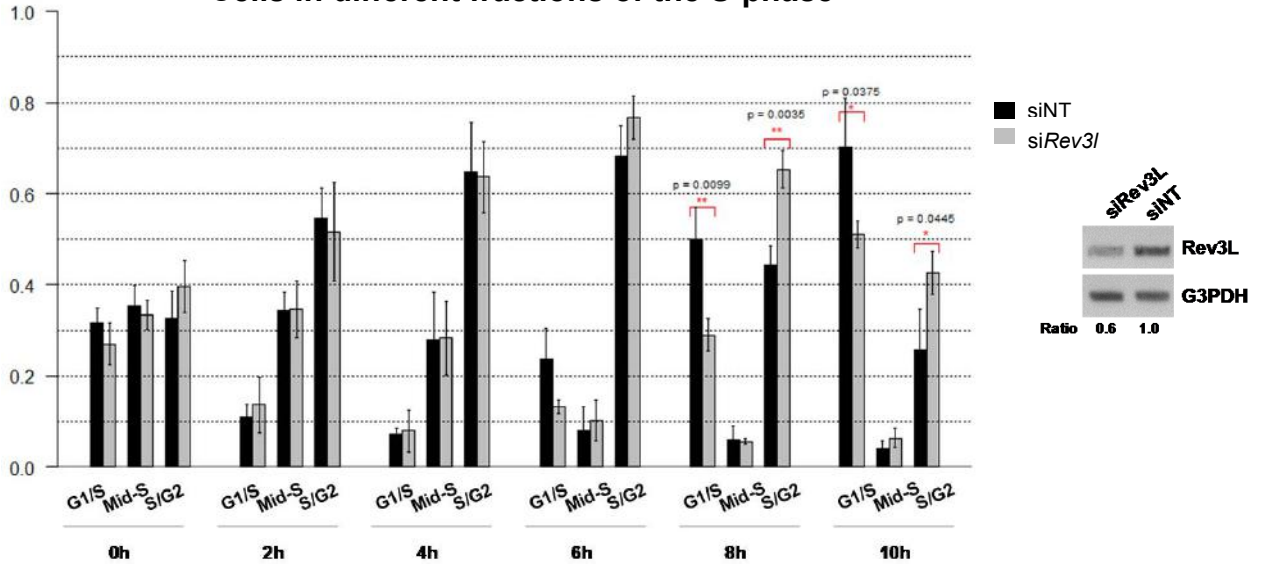
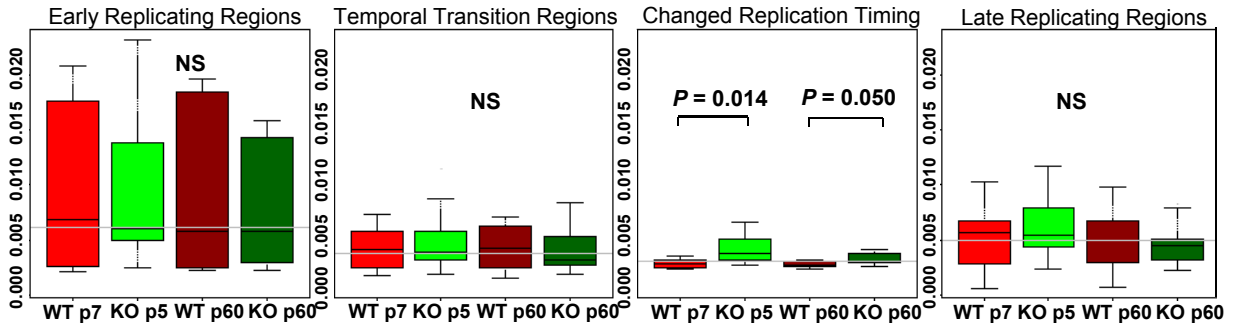


Figure S1, related to Figure 1. S phase progression is impaired in human cells after down regulation of *Rev3l*

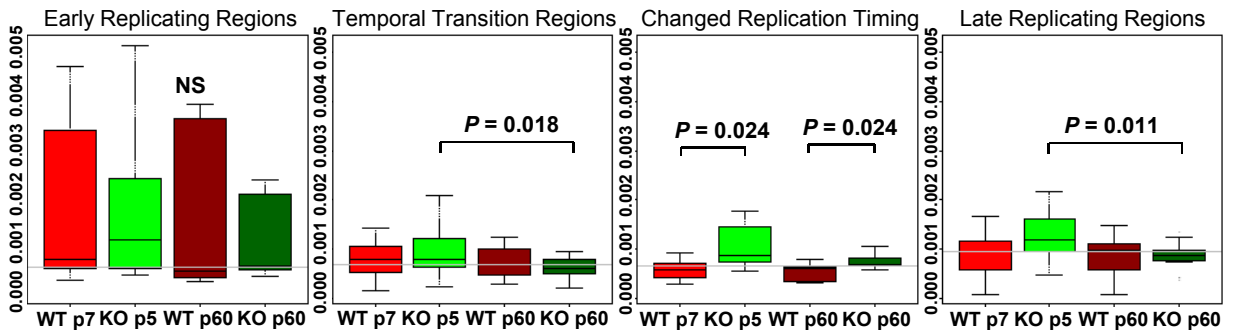
After 48h of transfection with untargeted (NT) siRNA or *Rev3l* siRNA, HeLa cells were pulse-labeled with BrdU prior to harvest and analyzed by flow-cytometry at different time points. Analysis was focused on S phase divided in 3 parts: G1/early S, middle S and late S/G2 BrdU+ cells. Histograms represent the percentage of cells in G1/early and late S/ G2 phase after the BrdU pulse. Error bars indicate standard error of the mean from three independent experiments (Student's *t*-test). qRT-PCR was performed to verify the downregulation of *Rev3l* in HeLa cells.

Supplemental Figure 2

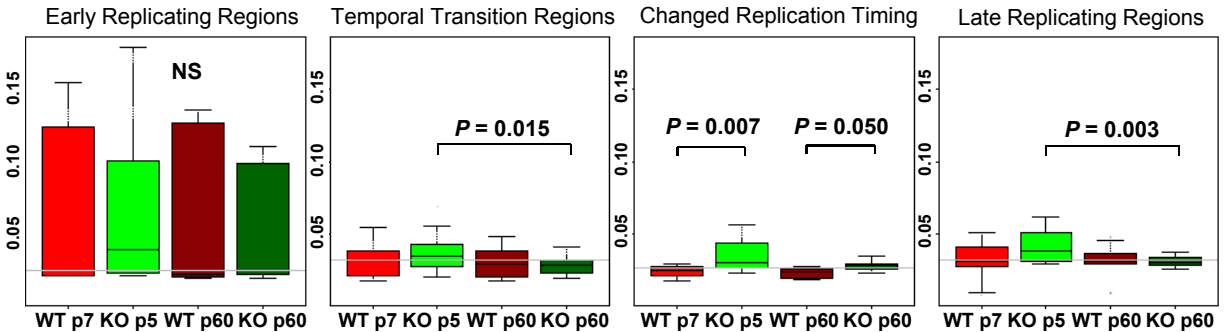
Deletion



Insertion



Transition



Tranversion

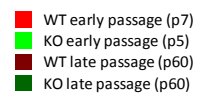
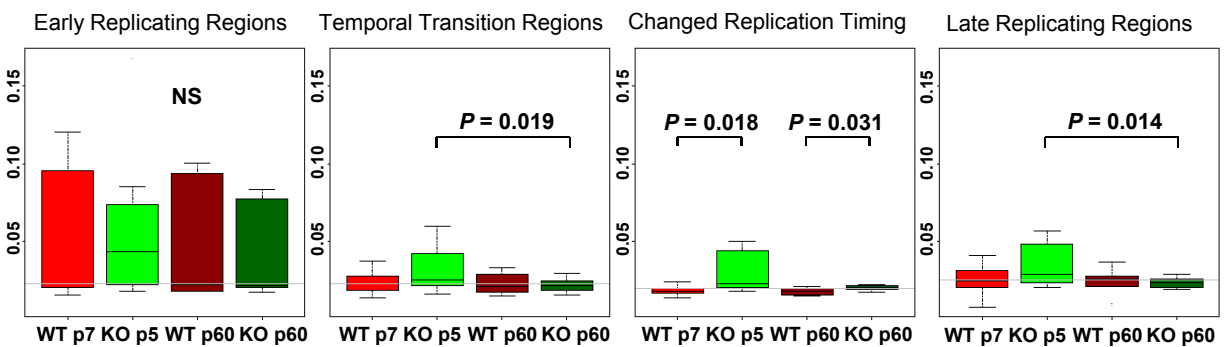
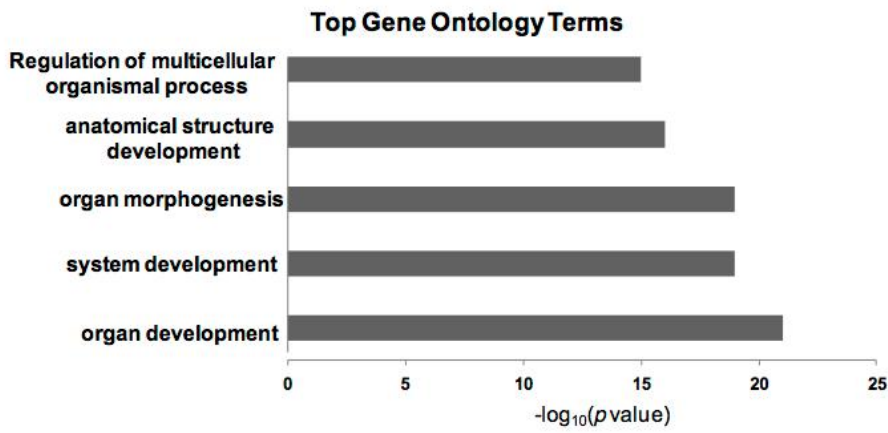


Figure S2, related to Figure 3. Loss of REV3L results in increased mutational events in genomic regions displaying delays in replication timing.

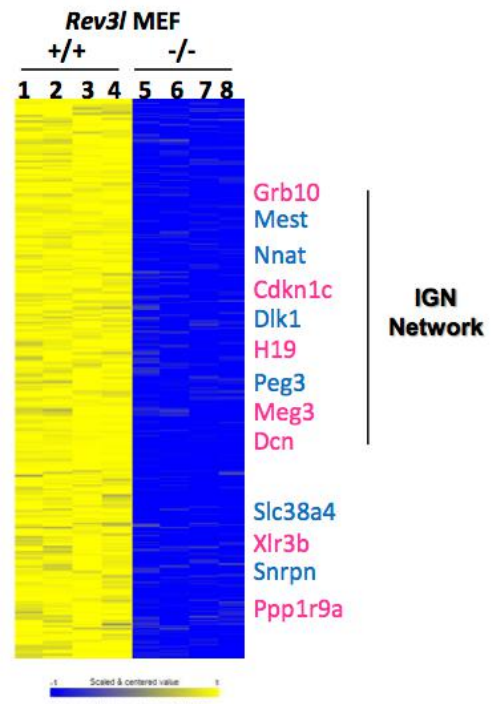
Assessment of mutational events in *Rev3l*^{+/+} (WT) and *Rev3l*^{-/-} (KO) MEFs was performed as described in Methods in specific genomic regions depicted in Figure 3. Deletions, insertions, transitions and transversions were quantified and normalized by the number of reads as shown by boxplots for each category (early, late, temporal transition regions and changed replication timing regions according to their replication timing). The horizontal black line in the box represents the median; p-values are for comparisons of median values (Wilcoxon rank sum test), ns: not significant.

Supplemental Figure 3

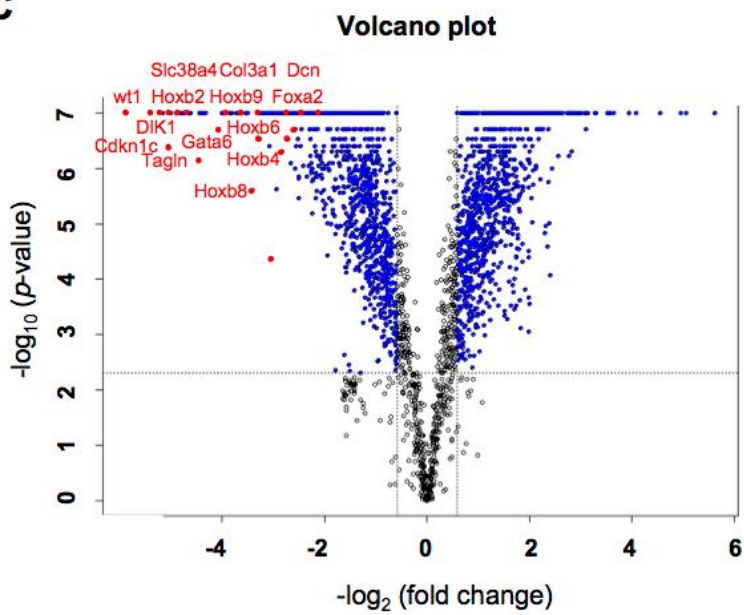
A



B



C



D

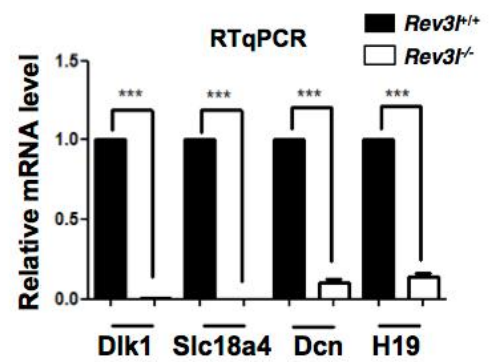
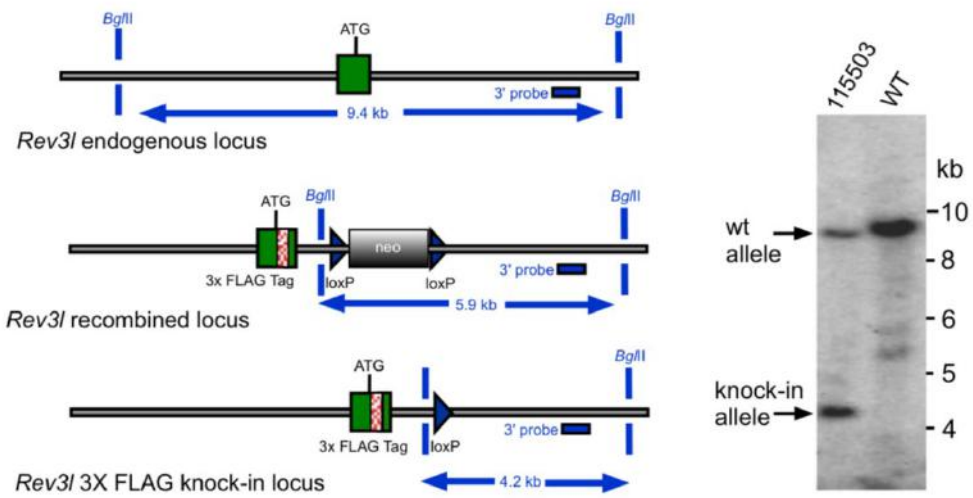


Figure S3, related to Figure 4: Inactivation of *Rev3l* impairs expression of numerous developmentally regulated genes

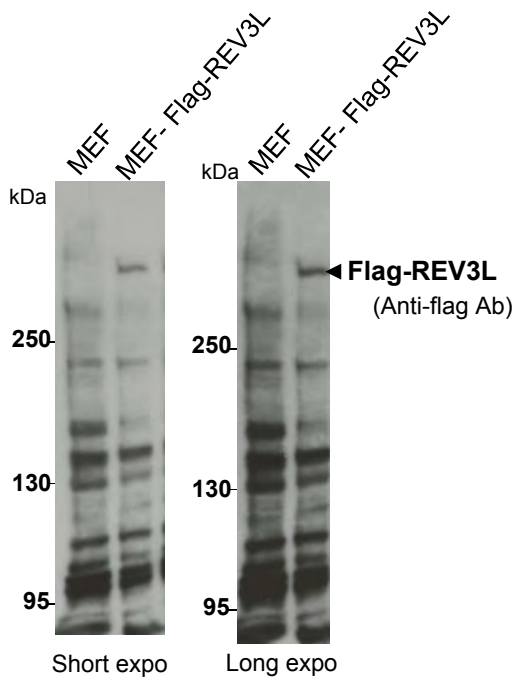
(A) Top five biological process gene ontology (GO) terms of genes found deregulated in *Rev3l*^{-/-} MEFs transcriptome analysis. **(B)** Heat map showing log₂ fold change of differentially expressed genes in *Rev3l*^{+/+} and *Rev3l*^{-/-} MEFs. Several imprinted genes repressed in *Rev3l*^{-/-} MEFs are indicated on the left. Yellow and blue indicate high and low mRNA expression levels, respectively. IGN: imprinted gene network. **(C)** Volcano plot shows in red genes involved in development and imprinting with high fold changes ≥ 3 (FDR <0.05). **(D)** Relative mRNA levels of genes involved in imprinting (paternal expressed genes: *Dlk1* and *Slc18a4* and maternal expressed genes: *Dcn1* and *H19*) were validated through qRT-PCR from *Rev3l*^{+/+} and *Rev3l*^{-/-} MEFs samples. The data were normalized to the amount of *GAPDH* mRNA. Error bars indicate standard error of the mean for three independent experiments. ***, P <0.001 and **, P <0.005 by Student's *t*-test.

A

MEFs expressing endogenous Flag-tagged REV3L



B



C

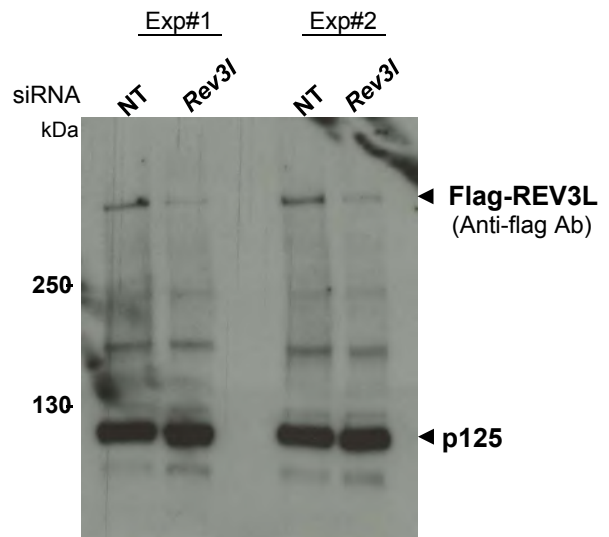
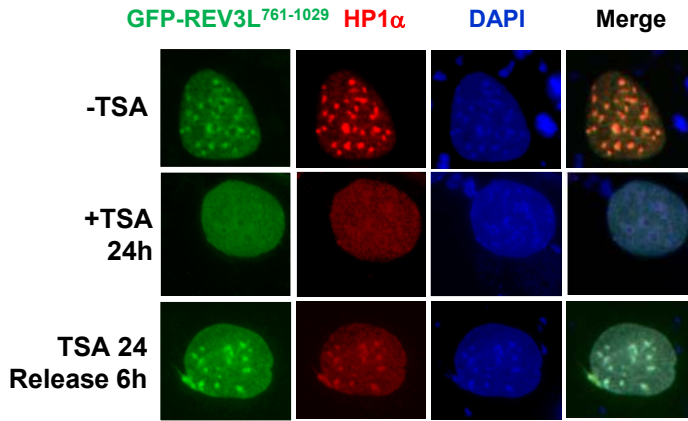


Figure S4, related to Figure 5. Generation of mouse embryonic cell line expressing endogenous Flag-REV3L protein.

(A) (Left) Diagram of the mouse *Rev3l* 3x FLAG knock-in allele. The wild-type (WT) locus is shown at the top. A green rectangle shows the first exon with the initiator ATG sequence. The gray line represents chromosomal sequence. The middle diagram shows the targeted locus after homologous recombination. The neomycin positive selection cassette (neo) is shown in gray. The cassette was excised by breeding mosaic animals with C57BL/6 Cre deleter mice, to produce the 3x FLAG tagged knock-in allele shown at bottom. (Right) Example of Southern blot analysis of the knock-in locus. Genomic DNA of the tested animals (number 115503 shown here) was compared with C57BL/6 wild-type genomic DNA (WT). *Bgl*II digested DNA was blotted on a nylon membrane and hybridized with the external 3' probe with the position shown at the left. Restriction fragments of 9.4 kb and 4.2 kb were observed for the wild-type, and knock-in locus, respectively. Genomic DNA was further analyzed extensively and confirmed by specific PCR assays and complete DNA sequencing as described in the Materials and Methods. **(B)** Total proteins were extracted from MEFs or MEFs expressing endogenous Flag-REV3L then REV3L protein was detected by immunoblotting using anti-Flag (M2) antibody. Specific signal above 250 kDa is detected only in MEF Flag-REV3L. **(C)** Specificity of Flag-REV3L detection was confirmed by down-regulating the REV3L protein. MEF Flag-REV3L was transfected with non-targeted siRNA (NT) or siRNA targeting *mRev3l*. 72h later, cells were lysed then proteins were analyzed by immunoblotting using anti-Flag (M2) antibody or anti-p125 antibody as loading control. Experiment was repeated twice (Exp#1 and #2).

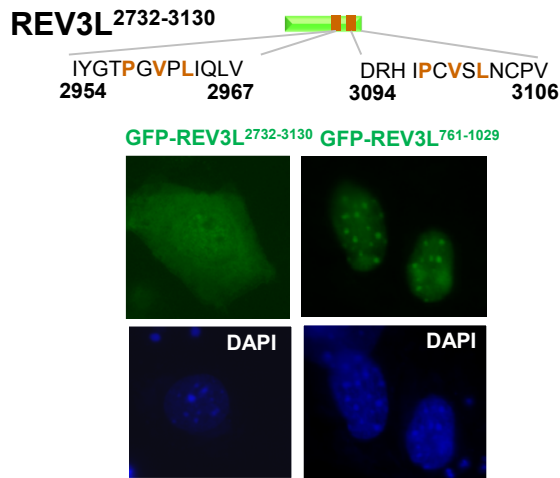
A



B

	HP1-Binding motif PxVxL		
REV3L, <i>H. sapiens</i>	775	RYEEFQEHKTEKPSLSQQA AHY MFF PSVVL SNCLTRPQKLSPVTYKLQPG	824
REV3L, <i>C. lupus</i>	696	RYEEFQEHKTEKPSLSQQA AHY MFF PSVVL SNCLSRPQKLSPVTYKLQPS	745
REV3L, <i>B. taurus</i>	775	RYEEFQEHKTEKPSLSQQA AHY MFF PSVVL SNCLSRPQKLSPVTYKLQPG	824
REV3L, <i>M. musculus</i>	774	RYEEFQEHKMEKPSLSQQA AHY MFF PSVVL SNCLTRPQKLSPVTYKLQSG	823
REV3L, <i>R. norvegicus</i>	774	RYEEFQEHKMERPSLSQQA AHY MFF PSVVL SNCLTRPQKLSPVTYKLQPG	823
REV3L, <i>G. gallus</i>	764	RYEEFQEHKA EKPSLSQQA AHY MFF PSVVL SNCLSRPQKLAPVTYKLQQG	813
REV3L, <i>X. laevis</i>	767	RSEELQQHATDMTSGNQQA THYKFF PSVVL SNCLNRPQKLAPVTYKVQQC	816

C



D

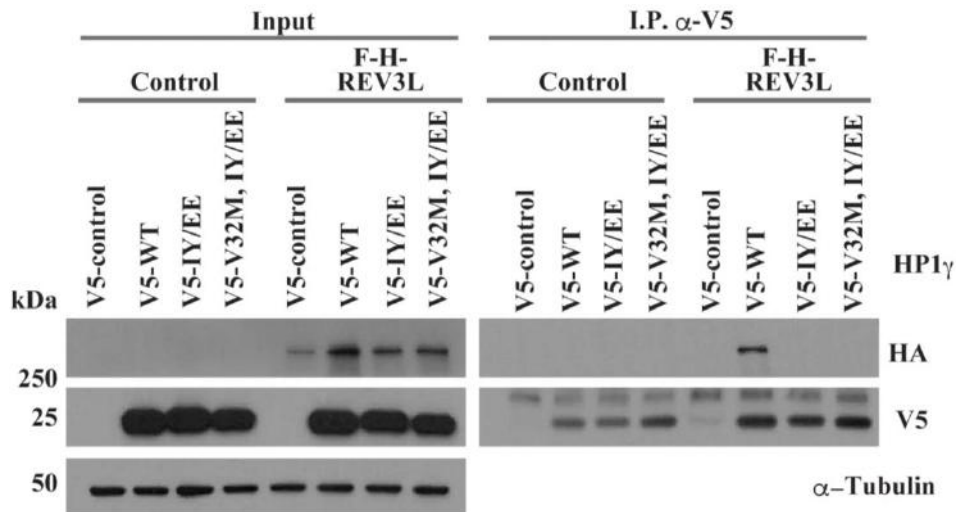


Figure S5, related to Figure 6. PxVxL motif targets Rev3L to pericentromeric heterochromatin through HP1 interaction

(A) WT MEFs were transiently transfected with GFP-REV3L⁷⁶¹⁻¹⁰²⁹ construct and treated or not for 24h with trichostatin A (TSA). Cells were released or not with fresh medium during 6 hours and fixed with 4% formaldehyde after 0.5% Triton X-100 pre-extraction. The distribution of GFP-REV3L⁷⁶¹⁻¹⁰²⁹ was detected by autofluorescence, HP1 α was immunodetected using antibodies (red) and nuclei were visualized using DAPI staining. **(B)** Sequence alignment of REV3L among different species containing the PxVxL motif with canonical residues shown in blue. **(C)** Schematic representation of REV3L²⁷³²⁻³¹³⁰ construct. MEFs cells were transiently transfected with GFP-REV3L²⁷³²⁻³¹³⁰ or GFP-REV3L⁷⁶¹⁻¹⁰²⁹ constructs and fixed with 4% formaldehyde. The distribution of GFP-REV3L constructs was detected by autofluorescence and nuclei with DAPI staining. **(D)** 293T cells were co-transfected with Flag-HA-REV3L full length or empty vector and V5-HP1 γ , mutant IY165 168EE V5-HP1 γ , mutant V32M IY165 168EE V5-HP1 γ or empty vector. 48 h after transfection cell lysates were made and used for immunoprecipitation with V5 antibody. After electrophoresis, samples were immunoblotted with anti-HA, anti-HP1 γ or anti- α Tubulin as indicated.

2. Addendum: Rev3L might contribute to telomere stability.

INTRODUCTION

In eukaryotes, telomeres are ribonucleoprotein structures capping the end of every linear chromosome that shorten with each mitotic cell division leading to senescence and cell death. Human telomeres span about 10-15kb while mouse telomeres are 25-50 kb long (Blasco, 2005). In all vertebrates they are composed of thousands TTAGGG repeats coated with specific protecting proteins called shelterins. Indeed, chromosome ends pose a problem for DNA replication, leading to a 3' single stranded overhang that can fold back and invade the double stranded telomeric region which will result in the formation of the T-loop, preventing the telomere ends from being recognized by the DNA repair machinery. The T-loop is held together by the shelterin complex composed by six core proteins (TRF1, TRF2, POT1, TPP1, RAP1 and TIN2) that contributes to the protection of chromosome ends (Shay and Wright, 2019).

It has been reported that telomeric DNA is a problematic template for the replication machinery, most likely due to formation of G-G Hoogsteen base-paired G4 DNA in the G-rich telomeric repeat strand. Moreover epigenetic regulation of the telomeric chromatin template critically impacts on telomere function and telomere-length homeostasis (Maestroni et al., 2017; Shay and Wright, 2019).

As previously mentioned we have found that REV3L interacts with ATRX, HP1 α , H4K20me3 and H3K9me3 and localizes partially into chromocenters rich in long stretches of repetitive tandem DNA. Given that these heterochromatin components participate to the telomere maintenance (Wong, 2010; Wong et al., 2010), we decided to explore further these particular heterochromatin regions. Our hypothesis is that Pol ζ could be recruited to help replicative polymerases at maintaining telomeres stability, composed of highly condensed and "difficult to replicate" chromatin, in order to maintain cell proliferation.

RESULTS AND DISCUSSION

We first tested if REV3L localizes at telomeric sequences by performing PLA between Flag-REV3L and one of the shelterin complex. We observed specific PLA signals between Flag-REV3L and TRF1 in physiological conditions (**Figure 1**), suggesting that Pol ζ can localize at telomere sequences.

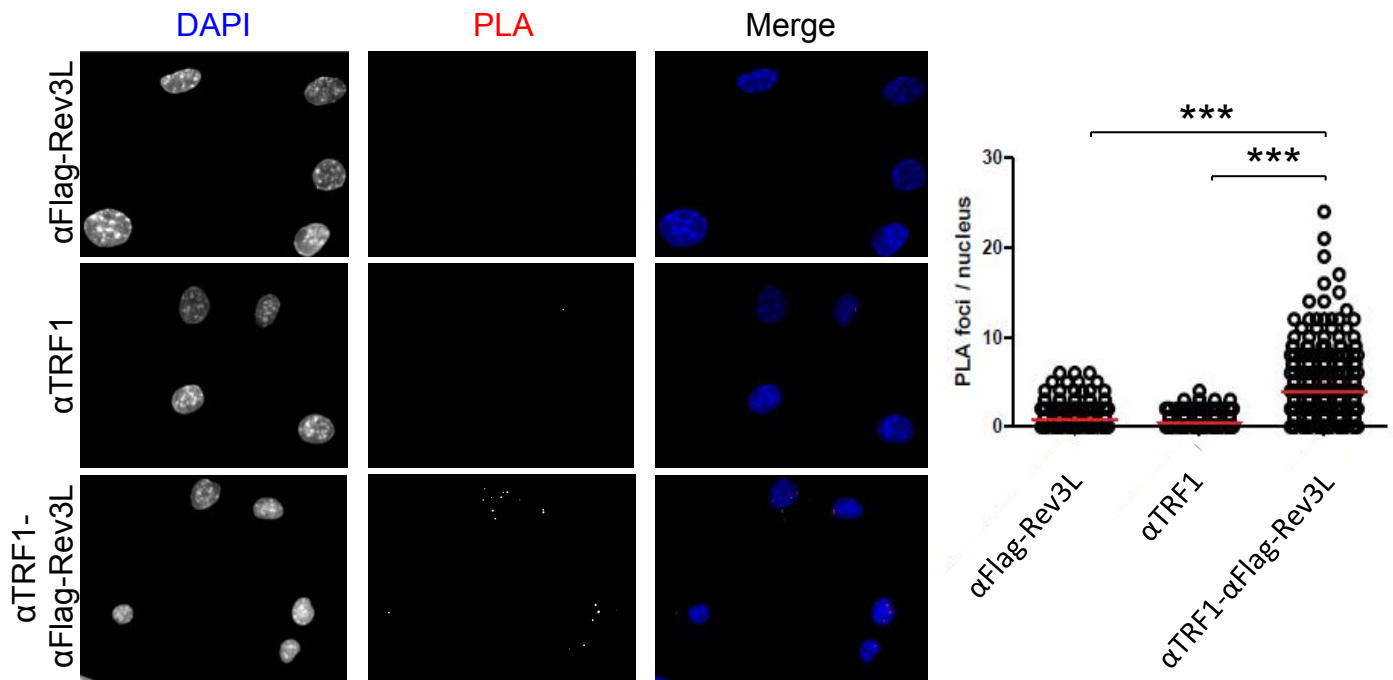


Figure 1: REV3L interacts with TRF1 in physiological conditions.

Proximity ligation assay between Flag-REV3L and TRF1. Bar chart show number of PLA foci per nucleus from two independent experiments (n=2). ** P<0,05 by Student's *t*-test.

Dysfunctional telomeres are associated with DNA-damage response factors such as 53BP1, γ -H2AX, ATM, and MRE11 (Takai et al., 2003). When the shelterin complex is compromised, telomeres become dysfunctional and display DNA-damage signals referred to as telomere dysfunction-induced foci (TIFs) (de Lange, 2005; Takai et al., 2003). Therefore we checked if loss of REV3L induced TIF by examining the co-localization of 53BP1 and TRF1 in unstressed cells. We found that the number of TIFs was significantly higher in MEF *Rev3l*^{-/-} than in MEF *Rev3l*^{+/+} (**Figure 2A**). These results were then validated by performing PLA between TRF1 and 53BP1, showing a higher proximity of the two proteins in absence of REV3L (**Figure 2B**). These results suggest that telomere function is impaired in absence of Pol ζ , leading to DSBs and DNA damage Response, corroborating the increase of telomere loss in MEF *Rev3l*^{-/-}.

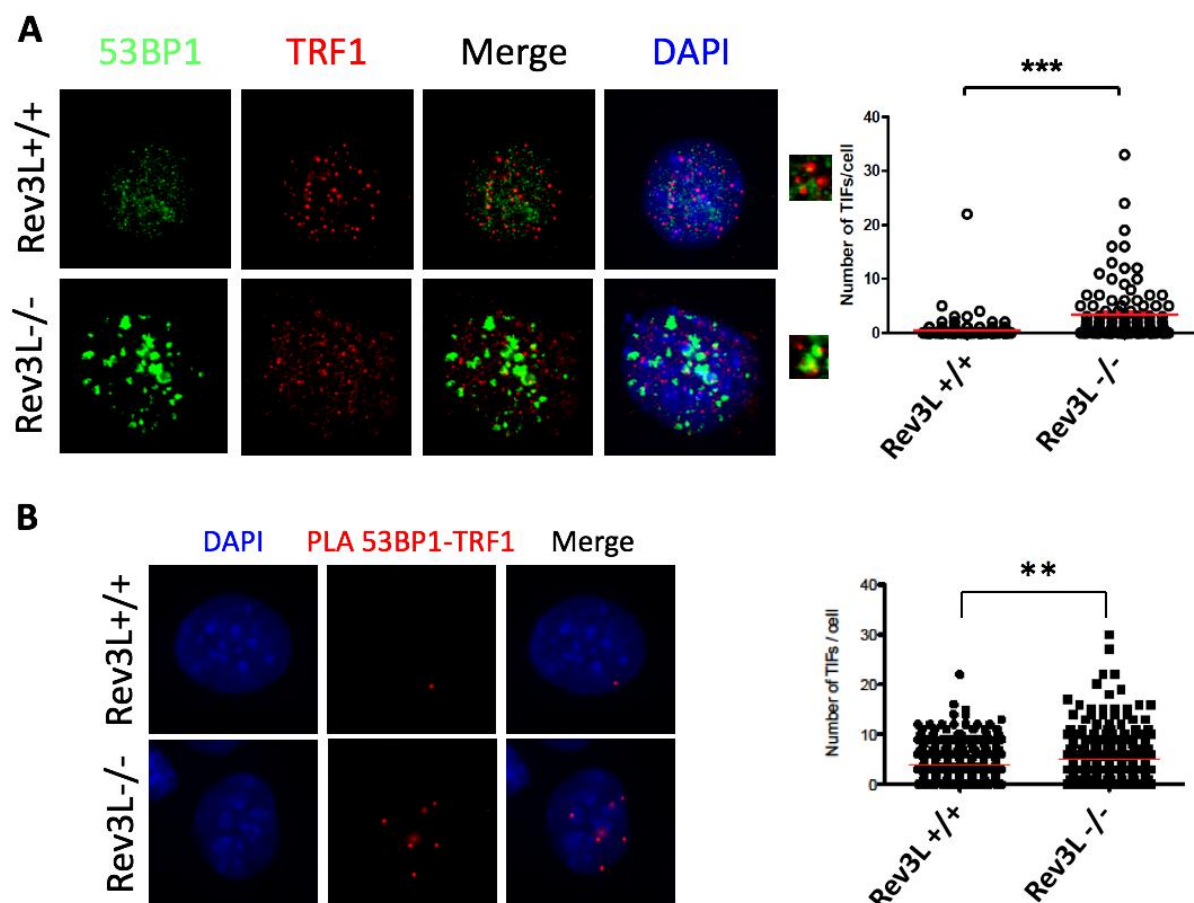


Figure 2: Telomere dysfunction is associated to REV3L loss

A) Co-immunofluorescence between 53BP1 (Green) labelling DNA double strand breaks, and TRF1 (Red), a component of telomeres capping, in MEFs Rev3L WT and Rev3L KO. Co-localization between the two signals indicated DNA breaks in telomeres regions, so-called telomere dysfunction-induced foci (TIFs). Zooms: presence or absence of co-localization are shown. Bar charts show number of TIFs per nucleus for Rev3L WT and Rev3L KO MEFs for two independent experiments (n=2). *** P<0,005 by Student's *t*-test. B) Proximity ligation assay between 53BP1 and TRF1. Bar charts show number of PLA foci per nucleus for Rev3L WT and Rev3L KO MEFs for two independent experiments (n=2). ** P<0,05 by Student's *t*-test.

To assess telomere length in absence of REV3L, we used the Q-fish method (Ourliac-Garnier and Londoño-Vallejo, 2017; Poon and Lansdorp, 2001). To do so, we performed FISH staining on MEF *Rev3l*^{-/-} than in MEF *Rev3l*^{+/+} after metaphase spreading using specific telomeric PNA probe TTAGGG (Figure 3A). Q-Fish allows to quantify target sequence. Since telomeres are composed of thousands TTAGGG repeats, signal intensity is proportional to telomeres length. For each telomere, signal intensity was quantified using ImageJ software. Interestingly, we found a significant reduction of the average telomere signal intensity in MEF *Rev3l*^{-/-} compared to MEF *Rev3l*^{+/+} (Figure 3B), indicating that telomere dysfunction is correlated with the reduction of telomere length.

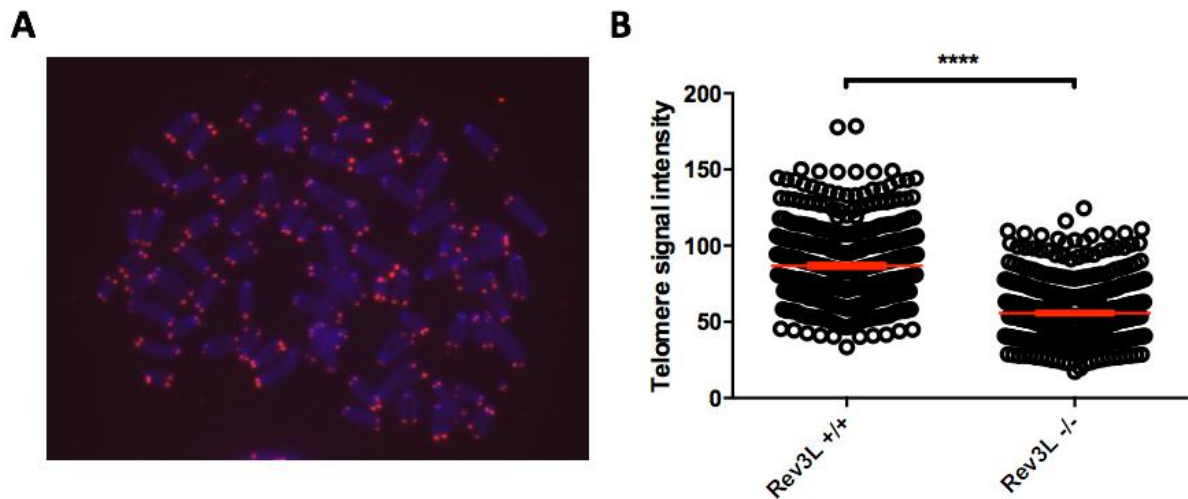


Figure 3: Loss of REV3L leads telomere shortening in MEFs.

A) Representative image of metaphase spreading performed on MEF *Rev3l*^{-/-} than in MEF *Rev3l*^{+/+} (only MEF wt is shown). Telomeres were labelled by fluorescent in situ hybridization (FISH) using a specific TTAGGG telomeric PNA probe in red. B) Q-fish method using metaphase chromosomes to estimate telomere length. The average intensity of the telomeric signal was calculated using ImageJ (arbitrary unit). Each dot of the chart represents one telomere intensity value. Two metaphases per genotype were analyzed. **** P<0,0001 by Mann-Whitney test.

In conclusion, we showed that REV3L is associated to shelterin proteins and its absence induces telomere dysfunction shown by the increased number of TIFs and telomere breaks resulting in telomere shortening, reinforcing our hypothesis that Polζ contributes to telomere replication, even if much work remains to be done to fully understand the role of Polζ in this process.

EXPERIMENTAL PROCEDURES

Proximity Ligation Assay (PLA)

PLAs were performed as previously described (Ben Yamin et al., submitted) using the mouse monoclonal anti-TRF1 (Abcam, ab10579) at 1:100.

Telomere Induced Foci (TIFs)

MEF *Rev3l*^{-/-} than in MEF *Rev3l*^{+/+} seeded on coverslips were washed with PBS, protein pre-extraction was carried out using CSK 100mM NaCl for 5min on ice, fixed with 4% paraformaldehyde for 20min and permeabilized with 0,5% Triton X-100 for 15. Coverslips were incubated with primary antibodies anti-53BP1 (1:500) and anti-TRF1 (1:100) for 1h at RT, then washed with PBS and incubated with fluorescent secondary antibodies for 45min at RT. Coverslips were then mounted in DAKO-DAPI mounting medium. Images were quantified by counting the number of co-localization signals per nucleus using the colocalization plugin of ImageJ.

REFERENCES

- Blasco, M.A. (2005). Telomeres and human disease: ageing, cancer and beyond. *Nat. Rev. Genet.* 6, 611–622.
- de Lange, T. (2005). Shelterin: the protein complex that shapes and safeguards human telomeres. *Genes Dev.* 19, 2100–2110.
- Maestroni, L., Matmati, S., and Coulon, S. (2017). Solving the Telomere Replication Problem. *Genes (Basel)* 8.
- Ourliac-Garnier, I., and Londoño-Vallejo, A. (2017). Telomere Length Analysis by Quantitative Fluorescent in Situ Hybridization (Q-FISH). *Methods Mol. Biol.* 1587, 29–39.
- Poon, S.S.S., and Lansdorp, P.M. (2001). Quantitative fluorescence in situ hybridization (Q-FISH). *Curr Protoc Cell Biol Chapter 18*, 18.4.1-18.4.21.
- Shay, J.W., and Wright, W.E. (2019). Telomeres and telomerase: three decades of progress. *Nat. Rev. Genet.* 20, 299–309.
- Takai, H., Smogorzewska, A., and de Lange, T. (2003). DNA damage foci at dysfunctional telomeres. *Curr. Biol.* 13, 1549–1556.
- Wong, L.H. (2010). Epigenetic regulation of telomere chromatin integrity in pluripotent embryonic stem cells. *Epigenomics* 2, 639–655.
- Wong, L.H., McGhie, J.D., Sim, M., Anderson, M.A., Ahn, S., Hannan, R.D., George, A.J., Morgan, K.A., Mann, J.R., and Choo, K.H.A. (2010). ATRX interacts with H3.3 in maintaining telomere structural integrity in pluripotent embryonic stem cells. *Genome Res.* 20, 351–360.

Partie II – Caractérisation
fonctionnelle et régulatrice de
l'ADN Polymérase zêta.

Dans cette deuxième partie, j'ai voulu étudier les mécanismes de régulation de cette polymérase à la fois dans des conditions physiologiques et après l'exposition à un stress génotoxique dans les cellules de mammifères. Les travaux entrepris au cours de cette seconde partie m'ont conduit à m'intéresser à l'implication de l'endopeptidase TASP1 dans ce mécanisme. Il m'est donc nécessaire d'introduire ici les caractéristiques de cette protéine, ses cibles déjà connues ainsi que les voies cellulaires dans lesquelles elle est impliquée.

1. Introduction : L'endopeptidase Threonine ASpartase 1 (TASP1)

a) Le gène et la protéine TASP1

Chez l'Homme, le gène *TASP1* est localisé sur le chromosome 20 sur le brin indirect en localisation 20p12.1 et compte 24 exons. Le gène mesure 534 177pb et code pour une protéine de 420aa. TASP1 est la seule protéase appartenant à la famille d'enzymes possédant un domaine d'homologie Asparaginases₂. En effet, TASP1 aurait apparemment évolué à partir d'enzymes métabolisant l'asparagine ou les protéines N-glycosylées pour reconnaître et cliver une protéine après un motif conservé avec un aspartate en position P1. TASP1 est exprimée de manière ubiquitaire dans la majorité des tissus et est très conservée au cours de l'évolution de la levure à l'homme.

b) Structure de la protéine TASP1

Le gène *TASP1* code pour un précurseur de 50 kDa. Après traduction, TASP1 subit un clivage post-traductionnel auto-protéolytique de manière à donner deux sous-unités α (28 kDa) et β (22 kDa) capables de s'assembler pour donner la forme active de TASP1 sous forme d'un hétérodimère α 2- β 2. Le site actif de TASP1 est la thréonine en position 234 N-terminale de la sous-unité β utilisée comme nucléophile pour la protéolyse.

c) Fonctions et cibles principales de TASP1

TASP1 est une endopeptidase spécifique de site capable de cliver ses substrats après un résidu aspartate (D) dans un motif IXQL(V)D/G (Bier et al., 2011). Comme pour la protéine TASP1 elle-même, tous ses substrats sont traduits sous forme de précurseurs qui doivent subir un clivage protéolytique post-traductionnel pour former des hétérodimères α/β matures (Hsieh et al., 2003a, 2003b). De manière intéressante, les cibles de TASP1 jouent un rôle dans la régulation transcriptionnelle de multiples voies cellulaires, suggérant un rôle important de

TASP1 pour orchestrer ces voies à travers son activité de clivage. Contrairement à la phosphorylation, l'acétylation, etc., le clivage protéolytique catalysé par TASP1 est une modification post-traductionnelle irréversible. Chez l'Homme, les cibles de TASP1 identifiées à l'heure actuelle sont MLL1, MLL2, TFIIA, ALF et chez la drosophile HCF1 (Capotosti et al., 2007; Hsieh et al., 2003b; Takeda et al., 2006; Zhou et al., 2006).

(1) MLL1 et MLL2

Les protéines « Mixed-lineage leukemia » MLL1 et MLL2 sont des méthyl-transférases spécifiques de la lysine 4 sur l'histone 3 (H3K4). MLL1 est transcrit sous forme d'un précurseur nécessitant un clivage protéolytique pour générer un fragment N-terminal de 320 kDa et C-terminal de 180 kDa. Ces deux fragments sont capables de s'hétérodimériser par la présence de domaines conservés pour former un complexe actif (Hsieh et al., 2003a, 2003b). MLL1 est impliquée dans l'expression des gènes *Hox* nécessaires à la différenciation cellulaire. En effet, l'expression orchestrée des différents gènes du cluster *Hox* détermine la mise en place des segments du corps chez les vertébrés. De plus, par son activité méthyltransférase, MLL1 contrôle l'expression de plusieurs gènes et en particulier des cyclines E, A et B impliquées dans la progression du cycle cellulaire. Chez l'Homme, une translocation chromosomique de la bande 11q23 entraînant une disruption du gène *MLL1* altère l'expression des gènes *HOX* et conduit à une leucémie (Oyama et al., 2013).

(2) TFIIA

La protéine « Transcription Factor II A » ou TFIIA est un facteur de transcription composé de trois sous-unités α , β et γ . Le précurseur TFIIA α - β est transcrit à partir d'un seul gène avant de subir une protéolyse post-traductionnelle pour générer un hétérodimère mature composé des deux sous-unités α et β . Le clivage du précurseur α - β est impliqué dans sa stabilité mais pas dans sa transcription, ni dans le développement embryonnaire chez *Xenopus laevis* (Zhou et al., 2006). En effet, le précurseur non clivé est plus stable que la forme clivée qui est ciblée pour une dégradation protéasomique. TFIIA actif peut former un complexe avec les protéines se liant aux TATA-box des promoteurs des gènes, initiant ainsi la transcription de ces gènes de manière dépendante de RNAPolIII. TFIIA joue un rôle dans la régulation de l'expression de différents gènes impliqués dans plusieurs processus biologiques tels que la spermiogenèse (Oyama et al., 2013).

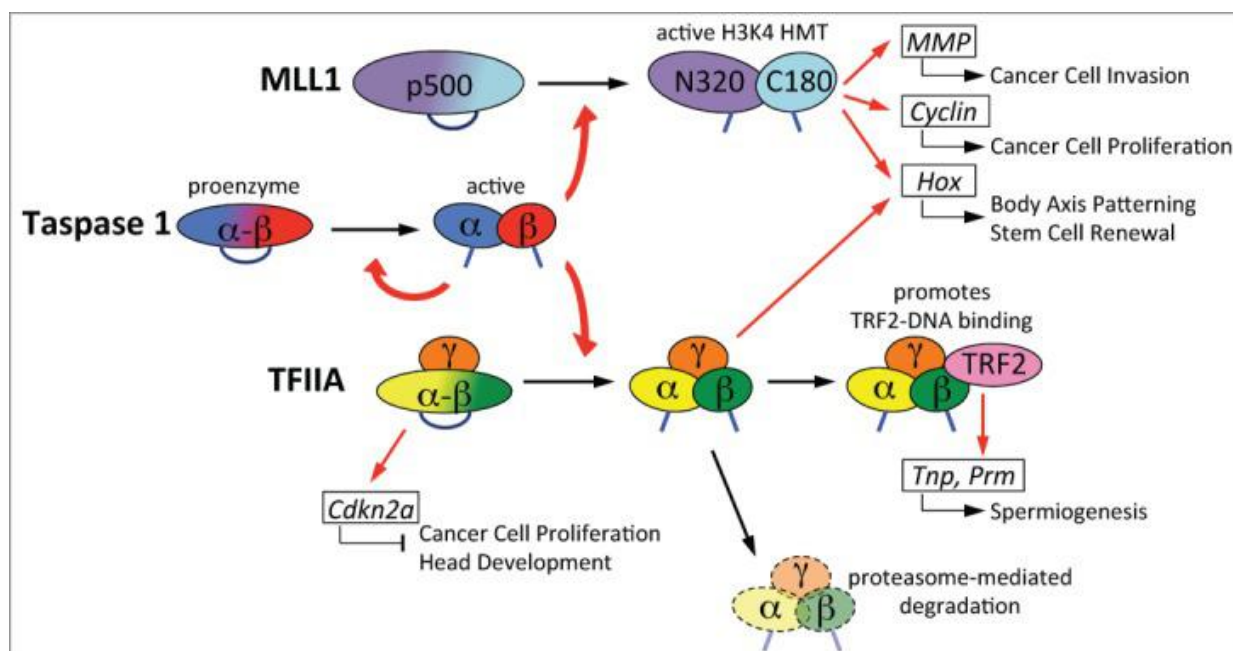


Figure 30: TASP1 régule divers mécanismes cellulaires à travers son clivage des facteurs de transcription MLL1 et TFIIA (Niizuma et al., 2015). La pro-enzyme TASP1 α - β est capable de s'auto-protéolyser et TASP1 mature clive MLL1 et TFIIA pour les activer.

d) Tasp1 in vivo

(1) Chez la souris

Takeda et al. ont généré des souris *Tasp1*^{+/-} en éliminant l'exon 9 porteur du site actif Thr 234 de Tasp1 par un mécanisme de recombinaison Cre-Lox dans un contexte génétique mixte 129Svj/C57BL/6. La délétion de cet exon 9 entraîne la perte complète de la protéine Tasp1 chez ces souris. Les individus hétérozygotes ne présentent pas de malformations évidentes. Le croisement de ces souris hétérozygotes a donné naissance à la première génération de souris *Tasp1* KO dans des proportions légèrement plus faibles que le ratio mendélien, cependant la majorité des nouveau-nés meurent un ou deux jours après la naissance. Après trois semaines, il ne reste que 13% des individus *Tasp1* KO.

A la naissance, les souris *Tasp1* KO sont plus petites que les souris WT et hétérozygotes de la portée, avec des malformations du crâne et de la langue, ainsi que des malformations du squelette typiques d'une transformation homéotique telles que l'apparition d'une paire de côtes supplémentaires au niveau du sternum et des transformations vertébrales. La diminution de la taille est due à une diminution de la prolifération cellulaire en absence de Tasp1 observable par la diminution de 54% du nombre de thymocytes *in vivo* et une diminution de trois fois du nombre de MEFS *Tasp1* KO après sept jours en culture *in vitro*. Cette diminution de la prolifération n'est pas due à une augmentation de la sénescence ou de l'apoptose mais à un

problème de réentrée dans le cycle à la transition G0-G1 et un blocage à la transition G1/S. Ces problèmes de cycle cellulaire sont associés à la diminution de l'expression des cyclines A, B et E due à une diminution de la marque H3K4me3 au niveau de leurs promoteurs en absence de MLL1 actif du fait de son non-clivage par Tasp1. De plus, l'absence de Tasp1 entraîne une augmentation de l'expression du gène *P16^{lnk4a}* sans modification de la méthylation au niveau de son promoteur, suggérant une régulation par une autre voie que MLL (Takeda et al., 2006).

A l'âge adulte, les mâles *Tasp1*^{-/-} survivants sont tous infertiles tandis que les femelles *Tasp1*^{-/-} sont fertiles. Cette infertilité des mâles est associée à une réduction de la taille des testicules, une absence de spermatozoïdes matures au niveau de l'épididyme et une réduction du nombre de couches de cellules germinales dans les tubes séminifères avec une absence de spermatides en élongation mais aucun problème hormonal. La diminution des spermatozoïdes matures est liée à l'incapacité des spermatides post-méiotiques de dépasser le stade 11 de la spermiogenèse résultant en une augmentation de l'apoptose des spermatides dans les tubes séminifères qui se retrouvent réduits de 33% dans les souris *Tasp1*^{-/-}. Les spermatides immatures retrouvées dans les souris *Tasp1*^{-/-} présentent un défaut de condensation de l'ADN avec une diminution de l'expression des protéines de transition (Tnp1) et des protamines (Prm1) nécessaires à cette condensation et donc à la spermiogenèse. En effet, en absence de Tasp1, le complexe TFIIA/TRF2 (TBP-related factor 2, facteur de transcription nécessaire à la spermiogenèse) n'est pas actif et se fixe de manière réduite sur les promoteurs des gènes *Tnp1* et *Prm1*, entraînant une diminution de leur expression (Oyama et al., 2013).

Dans un contexte génétique C57BL6 seul, les souris *Tasp1*^{-/-} présentent un vaste spectre de malformations de la tête allant de petits yeux (microphthalmie) à l'absence complète des structures antérieures aux oreilles (acéphalie). Ces malformations sont associées à un défaut de prolifération cellulaire du prosencéphale. Après une étude génétique, les auteurs ont relié les malformations observées à l'absence de clivage de TFIIA par TASP1 qui va stabiliser TFII α - β non clivé. TFIIA non clivé se fixe sur les promoteurs des gènes *p16lnk4a* and *p19Arf* au cours de la morphogenèse du crâne, augmentant fortement leur expression alors qu'ils sont normalement réprimés au cours de l'embryogenèse (Takeda et al., 2015).

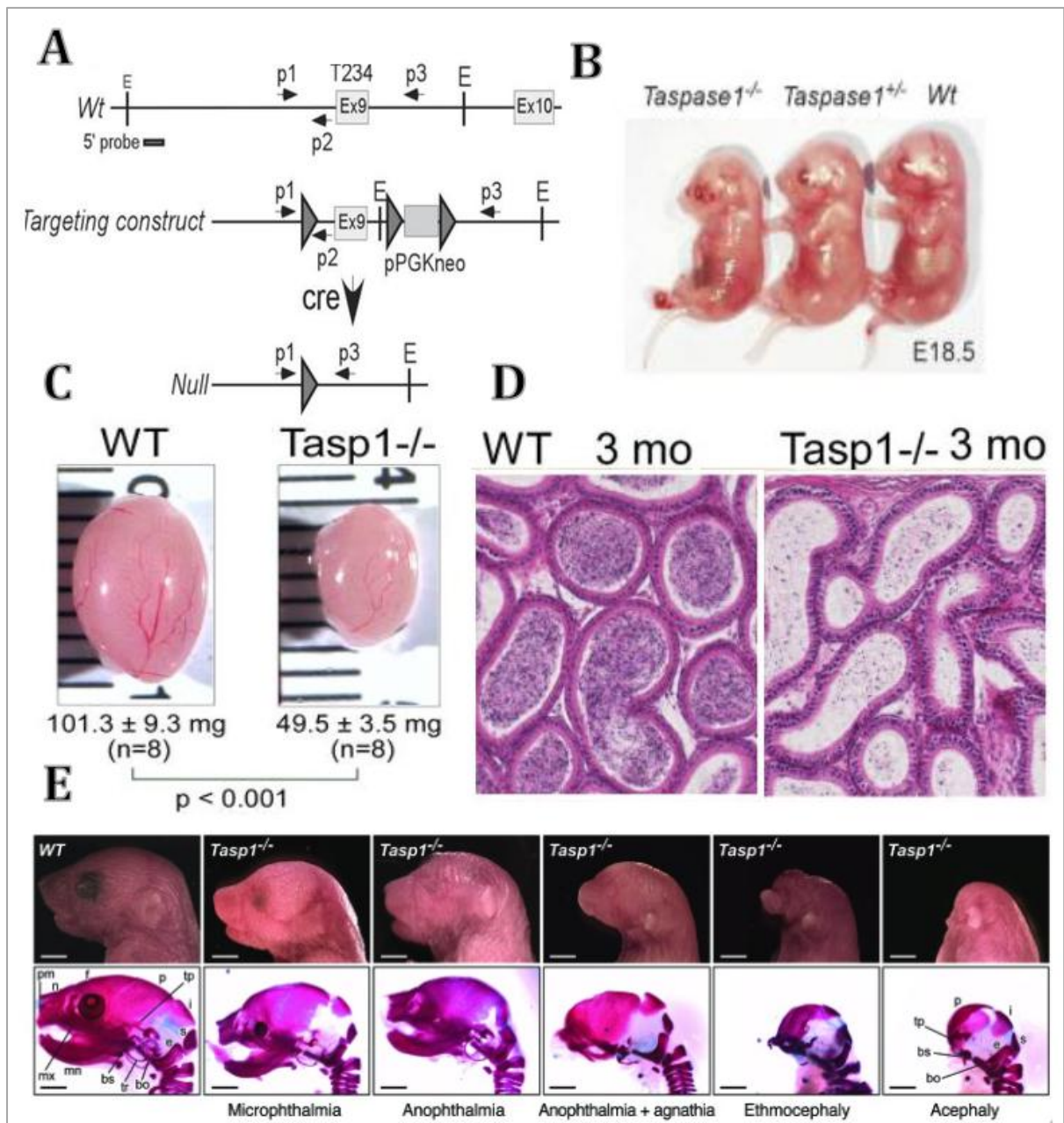


Figure 31: Phénotypes associés aux souris *Tasp1*^{-/-} (Oyama et al., 2013; Takeda et al., 2006, 2015). A) Représentation schématique du mécanisme d'inactivation de *Tasp1* par un système Cre-Lox pour générer des souris *Tasp1*^{-/-}. B) Diminution de la taille des souris *Tasp1*^{-/-} par rapport aux souris WT de la même portée. C) Réduction de la taille des testicules dans les souris *Tasp1*^{-/-}. D) L'épididyme des mâles matures *Tasp1*^{-/-} ne présente aucun spermatozoïde mature. E) Malformations cranio-faciales observées dans les souris C57BL6 *Tasp1*^{-/-}.

(2) Chez l'homme

En 2018, (Suleiman et al., 2018) ont décrit pour la première fois un patient porteur d'une délétion homozygote de la région 20p12.1 comportant les exons 5 à 10 du gène *TASP1*. Cet enfant présente un retard du développement, une microcéphalie ainsi que plusieurs anomalies congénitales du squelette, du cœur et du système urinaire. De plus, certains phénotypes retrouvés chez ce patient sont aussi observés chez les patients atteints du syndrome de Wiedemann-Steiner causé par un défaut de *MLL1*, une cible de *TASP1*.

e) TASP1 et cancer

TASP1 est retrouvée surexprimée dans un grand nombre de cancers mais pas mutée et semble nécessaire à la prolifération tumorale (Chen et al., 2010). En effet, TASP1 initie le développement tumoral par son clivage des protéines MLL1/2 et TFIIA qui jouent ensuite un rôle dans la prolifération cellulaire. L'absence de TASP1 dans des cellules cancéreuses de mélanome ou de glioblastome entraîne une augmentation de l'apoptose et inhibe la prolifération cellulaire. Aussi, une inhibition de TASP1 en association avec un traitement chimiothérapeutique entraîne une meilleure efficacité du traitement à tuer les cellules cancéreuses (Chen et al., 2010). De plus, dans des lignées d'hépatocarcinomes, l'activité complète d'histone méthyl-transférase de MLL1 est nécessaire à l'activation de métalloprotéases matricielles (MMP1 et MMP3) qui sont des régulateurs clés de l'invasion tumorale (Takeda et al., 2013).

Aussi, un inhibiteur de TASP1 dérivé d'arsenic (appelé NC48300) entraîne une diminution de la croissance tumorale *in vivo*, dans des modèles HER2/neu de tumeurs du sein ainsi que sur des xénogreffes de glioblastome chez la souris (Chen et al., 2012a). Aussi, la délétion conditionnelle de Tasp1 dans les glandes mammaires de souris prévient la tumorigénèse dépendante de HER2/neu de manière similaire aux souris portant des allèles non clivables de la protéine MLL1 (Dong et al., 2014). A travers son clivage de MLL1, TASP1 est donc est donc requis dans le développement du cancer du sein dépendant de HER2/neu.

Finalement, TASP1 n'est pas un oncogène classique puisqu'elle n'entraîne pas la transformation cellulaire, même en coopération avec d'autres oncogènes tels que E1a, Ras, Myc ou p53 dans des fibroblastes embryonnaires de souris. Cependant, son absence induit une résistance à la transformation cellulaire par ces mêmes oncogènes (Takeda et al., 2006). TASP1 est un facteur non oncogénique d'addiction pour les cellules cancéreuses qui en ont besoin pour la survie tumorale et représente donc une cible thérapeutique attractive pour de nouveaux traitements anticancéreux.

Résumé Article 2: Au cours de la seconde partie de ma thèse, je me suis intéressée à la régulation cellulaire la sous-unité catalytique de Pol ζ , REV3L, au cours de la réplication non perturbée de l'ADN ou après exposition des cellules à un stress génotoxique. Pour ce faire, j'ai utilisé une lignée cellulaire de MEFs (MEF Flag-REV3L) pour laquelle une étiquette FLAG (3xFLAG) a été introduite au niveau du premier exon du gène *Rev3l* permettant ainsi de détecter la protéine grâce à un anticorps anti-Flag. Nous avons observé que REV3L est finement régulée avec des temps de demi-vie courte de la protéine comme de l'ARN messager. Ceci a pour conséquence une diminution rapide de la quantité de protéine REV3L suite à une irradiation UV, due à une inhibition globale de sa transcription. De façon inattendue, nous avons mis en évidence un mécanisme de régulation unique de la protéine REV3L, de poids moléculaire de 360 kDa. En effet, nous avons démontré que REV3L est naturellement clivée par l'endopeptidase TASP1 au niveau de sa séquence consensus QLDGTAD en position 525. Ce clivage conduit à la formation d'un hétérodimère stable (REV3L p70/p300), composé des deux fragments résultants du clivage, suggérant que cet hétérodimère correspond à la forme active de REV3L. De plus, ce complexe est capable d'interagir avec REV7, la sous-unité régulatrice de Pol ζ , et de localiser aux fourches de réplication, sur les brins naissants dans des conditions physiologiques, suggérant que le potentiel de Pol ζ d'introduire des mutations dans le génome est beaucoup plus élevé que si elle était limitée à sa fonction de polymérase translésionnelle sur de l'ADN endommagé. De manière intéressante, le fragment N-terminal est hyperphosphorylé en réponse à un stress génotoxique. Des résultats préliminaires laissent penser que le clivage de REV3L par TASP1 est nécessaire pour la prolifération cellulaire, la résistance aux UV ainsi que le maintien de la stabilité génomique dans les cellules mammifères. En conclusion, nous avons découvert un mécanisme unique pour contrôler la fonction d'une TLS polymérase dans les cellules de mammifères.

2. *REV3L is proteolytically processed by Taspase1 (TASP1) to form an active heterodimer required for pol ζ activity. (in preparation)*

Barbara Ben Yamin¹, Caroline Pouvelle¹, Dorian Bochaton¹, Pierre David², Richard D Wood³, Emmanuelle Despras¹,....., Patricia L Kannouche^{1*}

¹Paris-Sud University, CNRS-UMR8200, Equipe labellisée Ligue Contre le Cancer, Gustave Roussy, Villejuif, France.

²LEAT - Institut Imagine, Boulevard du Montparnasse, Paris, France

³Department of Epigenetics and Molecular Carcinogenesis, The University of Texas MD Anderson Cancer Center and The University of Texas Graduate School of Biomedical Sciences at Houston, Houston, TX 77030, USA

* Correspondence: patricia.kannouche@gustaveroussy.fr (P.L.K)

HIGHLIGHTS

- REV3L protein is proteolytically processed by the endopeptidase TASP1
- TASP1-mediated REV3L cleavage generates two polypeptides that heterodimerize to form a stable complex
- Cleaved N-Terminal REV3L fragment is hyperphosphorylated following genotoxic stress
- Cleaved REV3L fragment is enriched on nascent DNA in physiologic conditions and after replication stress
- Abrogation of TASP1 cleavage site in *Rev3l* gene leads to a novel isoform of REV3L protein
- Alteration of the native TASP1-mediated cleavage of REV3L influences cell proliferation, UV sensitivity and genome stability in mammalian cells

INTRODUCTION

Genomic instability is a double-edged sword. It is the main source of genetic diversity with mutations that are beneficial for evolution, but it is also defined as a major driving force for tumorigenesis. Information about the mechanisms that generate mutations in eukaryotes is likely to be useful for understanding human health concerns, such as genotoxic stress and cancer. Mutations are largely the outcome of attacks by endogenous and environmental agents, and cells have evolved sophisticated repair systems to remove DNA damage. However, if such DNA damage persists in the genome during replication, they can create replication blocks leading to fork collapse, double-strand breaks, and genomic rearrangements (Zeman and Cimprich, 2014). To prevent this scenario, specialized DNA polymerases (Y-family pol η , pol ι , pol κ and REV1, and B-family pol ζ) help cells tolerate DNA damage by replicating opposite the lesions or filling single-stranded gaps left after replication in a process called “translesion synthesis” (TLS), at the expense of increased mutations (Lehmann et al., 2007; Zhu et al., 2003).

DNA polymerase ζ (pol ζ) is a key player in mutagenic TLS in eukaryotic cells. It was initially identified in yeast as a heterodimer comprising two subunits: Rev3, the catalytic subunit and Rev7, the regulatory subunit (Nelson et al., 1996). However, several labs recently reported that pol ζ is consisting of four subunits. In addition to Rev3 and Rev7, pol ζ shares two subunits with the replicative polymerase Pol δ , Pol31p/POLD2 and Pol32p/POLD3 (Baranovskiy et al., 2012b; Johnson et al., 2012; Lee et al., 2014; Makarova et al., 2012). Despite being a member of the B-family DNA polymerases, pol ζ lacks an intrinsic 3'-5' exonuclease activity, making this TLS polymerase highly error prone with a spontaneous mutation rate 10 to 100-fold higher than that of replicative DNA polymerases (Fortune et al., 2005; McCulloch and Kunkel, 2008; Shcherbakova et al., 2003). Pol ζ is essential for the bypass of most lesions acting predominately as an extender of aberrant primer termini formed at the lesion site (Guo et al., 2001; Haracska et al., 2001; Johnson et al., 2000). In yeast, DNA damage-induced mutagenesis and spontaneous mutagenesis are drastically reduced in pol ζ -deficient cells (Harfe and Jinks-Robertson, 2000; Lawrence and Christensen, 1979; Lemontt, 1971c; Quah et al., 1980; Roche et al., 1994). Beside its function in bypassing DNA damage, yeast pol ζ has been implicated in the replication of non-damaged DNA after replication fork stalling preventing thus the formation of double-stranded (Makarova and Burgers, 2015; Northam et al., 2014).

Unlike yeast *REV3*, which is not essential for viability (Morrison et al., 1989), disruption of *Rev3l* gene in mice causes embryonic lethality revealing an important role of pol ζ in cell proliferation and genomic stability even in the absence of environmentally induced damage (Bemark et al., 2000; Esposito et al., 2000; Wittschieben et al., 2000). Consistent with this, we recently showed that pol ζ assists high-fidelity DNA polymerases in replicating heterochromatin regions

preventing thus genome instability (Ben Yamin et al, submitted). However, despite the established involvement of REV3L in many cellular processes including lesion bypass, DNA replication and repair processes, little is known about how pol ζ is controlled in mammalian cells. Studies on mammalian REV3L protein have been hampered by the inability to express and/or detect this very large polypeptide in cells. This is particularly crucial since pol ζ is an important determinant for tumor resistance to chemotherapeutic agents (Doles et al., 2010; Xie et al., 2010; Zhu et al., 2016b), and thus could represent an attractive target for cancer therapy. Therefore, characterizing the mechanisms of regulation of pol ζ function under normal growth condition and after DNA damage insults are an important area for cancer biology.

Site-specific proteolysis regulates a variety of physiological and cellular processes. It can remove regulatory proteins when they are not needed, while transforming others from the dormant into the biologically active state (Ehrmann and Clausen, 2004). The endopeptidase Threonine Aspartase 1 peptidase (TASP1) is a 50-kDa endopeptidase of a family of hydrolases possessing an asparaginase 2 homology domain (Hsieh et al., 2003c; Khan et al., 2005), and was originally identified as a protease responsible for cleaving mixed-lineage leukemia 1 (MLL) protein for proper regulation of *HOX* gene expression (Hsieh et al., 2003a, 2003c). Genuine TASP1 substrates include MLL1, MLL2 (KMT2B), TFIIA α - β (GTF2A1), ALF α - β (TFIIA-like factor, GTF2A1L), and all TASP1 substrates are translated as α - β precursors that undergo proteolysis to form mature α / β heterodimers, like TASP1 itself (Hsieh et al., 2003a; Zhou et al., 2006). Interestingly, all of these substrates so far are broad-acting nuclear factors that play important roles in transcription regulation, suggesting a critical role for TASP1 in the orchestration of various genetic programs through the cleavage of its substrates (Niizuma et al., 2015).

In this study we reported that native REV3L protein is proteolytically processed by TASP1 in mammalian cells to generate two polypeptides that heterodimerize to form a stable complex that associates with REV7, likely representing the active complex of pol ζ . Cleaved N-terminal REV3L fragment responds to DNA replication stress by being hyperphosphorylated and enriched on nascent DNA. Importantly, cleaved REV3L localizes at replication forks even under physiological conditions, indicating that the potential for pol ζ to cause mutations is far greater than would be the case if pol ζ was restricted to TLS function on damaged DNA templates. Moreover, alteration of the native TASP1-mediated cleavage of REV3L influences cell proliferation, UV sensitivity and genome stability in mammalian cells. Altogether, these findings highlight a unique mechanism to control the function of an error-prone polymerase in mammalian cells.

RESULTS

In physiologic conditions, both REV3L protein and mRNA have a short half-life.

Despite the central role of REV3L/pol ζ in maintaining genome integrity, little is known about *Rev3l* gene and protein regulation. Therefore, we decided to examine the stability of REV3L protein in cells. For that we used the MEF Flag-REV3L which is a unique cell line expressing 3xFLAG-tagged REV3L protein at the endogenous level. Given that none REV3L antibody are commercially available, MEF Flag-REV3L is an unprecedented tool allowing the detection of endogenous REV3L protein in mammalian cells (Figure S4 in Ben Yamin *et al.*, submitted article). We measured the half-life of REV3L protein by treating cells with cycloheximide (CHX), an inhibitor of protein synthesis. Protein level was then monitored at different time points up to 6 h. We observed that REV3L protein levels drastically decreased within the first hours of treatment (**Figure 1A**). The linear regression analysis showed that the half-life of REV3L protein was of ~ 90 min (**Figure 1B**) demonstrating that albeit REV3L is a large protein (>360 kDa), it has a relatively short half-life in mouse embryonic fibroblasts.

We then examined the half-life of *Rev3l* mRNA using the Click-iT Nascent RNA Capture approach. Thus, we incubated MEFs-FlagREV3L cells with 5-ethynyluridine (5-EU), a Uridine analog, during one cell cycle in order to incorporate EU in all mRNAs. Cells were then incubated in fresh media and EU-containing mRNAs were recovered by click chemistry at different time points. After reverse transcription, levels of EU-containing *Rev3L* mRNA were analyzed by quantitative PCR (**Figure 1C**). The linear regression analysis showed that the half-life of *Rev3l* mRNA was of ~ 3 h30 (**Figure 1D**). Given that the median estimated half-life for all genes is about 7 h in ES cells (Sharova *et al.*, 2009), we concluded that *Rev3l* mRNA has a short half-life. These results suggest that *Rev3l* gene is closely regulated both at the transcriptional and translational levels allowing a rapid reduction of REV3L protein levels in mammalian cells. Consequently, any internal or external cellular stimuli that alter transcriptional or/and translational machineries could impact on REV3L protein levels.

Drastic decrease of REV3L p360 protein levels upon UV-irradiation

Cells with defective pol ζ are sensitive to a variety of DNA damaging agents including UV radiation, cisplatin and aflatoxin, which result in distorting or bulky adducts on DNA (Lin *et al.*, 2016; Tomida *et al.*, 2015; Zander and Bemark, 2004). To better understand the functional importance of pol ζ upon UV irradiation, we examined the regulation of its catalytic subunit, REV3L after exposure cells to UVC radiation. Unexpectedly, we observed a strong decrease of REV3L full length (p360) protein levels 4 h after UV-irradiation using immunofluorescence and western blot analysis (**Figure 2 A-B**). In contrast, aphidicolin treatment, a DNA polymerase

inhibitor, did not affect REV3L protein levels, indicating that the decline of REV3L protein levels observed after UVC irradiation is not due to DNA replication inhibition, but rather UV-induced cellular insults. Interestingly, the decrease of REV3L p360 protein levels followed a precise and reproducible kinetics, with a minimal level observed 6-8 h after UV-irradiation then an increase reaching basal levels 16 h later (**Figure 2 C**). This was dose-dependent, and mimics the kinetics of REV3L p360 protein after CHX treatment (**Figure 2 D-E**). Given the intrinsic short half-life of REV3L p360 (**Figure 1A-B**), it is likely that the decrease of REV3L p360 upon UV radiation is caused by changes in *Rev3l* gene expression.

To test this hypothesis, we monitored *Rev3l* mRNA levels in MEF Flag-REV3L irradiated at 20J/m² at different time points by qRT-PCR. We found that following UV-irradiation, changes of *Rev3l* mRNA levels were similar to those of REV3L protein, with a minimal amount observed 6-8 h post-irradiation (**Figure 2F**), and this was dependent of UV-dose (**Figure 2G**). We reasoned that the reduction of *Rev3l* mRNA levels upon UV-irradiation might be due to either transcription inhibition or *Rev3l* mRNA destabilization. To test the latter hypothesis, we examined the *Rev3l* mRNA half-life by click chemistry after incubation of cells with 5-EU during 24 h then UV- irradiation. Levels of EU-containing *Rev3L* mRNA were analyzed at different time points by qRT-PCR (**Figure S1A, top panel**). The linear regression analysis showed that *Rev3l* mRNA levels decrease at the same speed in mock and irradiated cells (**Figure S1A, bottom panel**), indicating that UV treatment does not impair *Rev3l* mRNA stability. We then tested *Rev3l* transcription inhibition using luciferase reporter assay. For that, we searched the proximal region of the transcription start site of *Rev3l* gene using appropriated software, then human promoter was cloned in a pGL4 luciferase reporter vector. After transfection in HEK293T cells, luciferase activity was measured from unirradiated cells or 6 h after UV-irradiation. We found a relative reduction of luciferase activity (~50%) in irradiated cell extracts compared to unirradiated cell extracts (**Figure S1B**), suggesting that exposure cells to UV irradiation inhibits *Rev3l* gene transcription, leading to a strong decrease of REV3L protein levels. We also confirmed that this UV-response on *Rev3l* gene expression was not cell line specific since we obtained comparable results with mouse and human primary, transformed and tumor cells (**Figure S1C-F**). It is noteworthy that these data are counterintuitive if we assume that polζ plays an important role in response to UV-irradiation.

REV3L protein is proteolytically processed by the endopeptidase TASP1

To better understand how REV3L protein functions in mammalian cells, we constructed a series of GFP-tagged REV3L fragments (**Figure 3A**). After transfection in HEK293T cells and western blot analysis, we observed that all constructs encompassing the N-terminal domain between aa 400-600 were systematically cleaved generating a stable N-terminal product of ~70 kDa

(corresponding to 90 kDa in GFP-fused constructs). This is in agreement with results obtained by others (Lee et al., 2014; Tomida et al., 2015). The simplest explanation is that REV3L may undergo a proteolytic cleavage for limiting protein levels in mammalian cells.

The protease Threonine Aspartase 1 (TASP1), has been identified as a threonine endopeptidase that cleaves various regulatory proteins, including the mixed-lineage leukemia-MLL1 and 2- proteins and the general transcription factor TFIIA, to form active complexes that orchestrate different biological processes (Takeda et al., 2006, 2015). TASP1 recognizes a consensus sequence (Q₃(F/I/L/V)₂D₁G₁X₂X₃D₄) and cleaves its targets between the Aspartic acid and the Glycine residues of this cleavage site. Interestingly Bier and colleagues used an *in silico* approach to identify targets of TASP1 and predicted REV3L as a substrate with a putative cleavage site located at position 523-529 (QLDGTAD) in human REV3L sequence (Bier et al., 2011) (**Figure 3B**, top panel). By comparing the putative cleavage site (CS) in REV3L to that of MLL1 and TFIIA proteins (**Figure 3B**, bottom panel), we noticed that the CS in REV3L is similar to the CS in the MLL1 (QVDGAAD) and TFIIA (QVDGTDG) proteins, known to be naturally cleaved at D/G by TASP1. This similarity strongly suggests that REV3L protein might be cleaved by the same protease.

Therefore, we assessed whether TASP1 is responsible for proteolytic cleavage observed on overexpressed REV3L constructs. Specific siRNAs were used to downregulate mTASP1 protein then 48h later, cells were transfected with GFP-REV3L⁴⁰⁰⁻⁶⁰⁰ construct. We observed a decrease in the endogenous TASP1 levels of about 40-50%. This led to an accumulation of the noncleaved (nc) REV3L⁴⁰⁰⁻⁶⁰⁰ fragment associated to a small decrease in the cleaved products (REV3L⁴⁰⁰⁻⁶⁰⁰N) in MRC5 and HEK293 cells (**Figure 3C**). Note that as GFP has been fused to the N-terminus of REV3L constructs, only cleaved N-products can be detected with GFP antibody. To strengthen these results, we used primary fibroblasts from patients mutated in *TASP1* gene (homozygous nonsense variant, c.199 C>T, p.Arg67*, (Suleiman et al., 2018)). We first verified the absence of TASP1 protein in these cells (**Figure S2A**), before transfecting primary fibroblasts with vector expressing GFP-REV3L⁴⁰⁰⁻⁶⁰⁰ then analyzed cell extracts by western blotting. While in normal fibroblasts we observed a large amount of N-terminal cleaved product (REV3L⁴⁰⁰⁻⁶⁰⁰N), in *TASP1* mutated cells mainly noncleaved REV3L⁴⁰⁰⁻⁶⁰⁰ product was detected (**Figure 3D**). These findings indicate that overexpressed REV3L protein is cleaved specifically by TASP1 in mammalian cells.

Next, we examined whether the putative cleavage site identified in REV3L (QLDGTAD) is essential for TASP1 cleavage. For that, we generated noncleavable (nc) REV3L fragments by mutating QLDGTAD sequence to QLAATAD, to abolish proteolytic cleavage by TASP1. HEK293T cells were then transfected with plasmids expressing GFP-REV3L⁴⁰⁰⁻⁶⁰⁰wt or GFP-REV3L⁴⁰⁰⁻⁶⁰⁰nc. As shown in **Figure 3E**, mutations of REV3L cleavage site completely abolish the formation of N-

terminal cleaved products resulting in a significant increase of noncleavable REV3L⁴⁰⁰⁻⁶⁰⁰ fragments. We confirmed these data by transfecting cells with plasmids expressing GFP-REV3L¹⁻¹²⁰⁰wt or GFP-REV3L¹⁻¹²⁰⁰nc. We observed that REV3L¹⁻¹²⁰⁰ construct possessing mutated cleavage site was no longer processed as evidenced by the disappearance of N-terminal cleaved products as well as an abundance of REV3L¹⁻¹²⁰⁰nc (**Figure 2F**). Moreover, we were able to detect the C-terminal cleaved product from GFP-REV3L¹⁻¹²⁰⁰wt cell extract. Indeed, we have generated a mouse monoclonal antibody against residues 636-655 of human REV3L sequence, which failed to detect endogenous REV3L protein, but could detect overexpressed REV3L fragments. Using this antibody, we detected noncleaved REV3L¹⁻¹²⁰⁰ as well as the C-terminal portion of REV3L (**Figure 2F**). These results demonstrate that both N-terminal and C-terminal products of REV3L are stable after proteolytic cleavage. Altogether these findings indicate that overexpressed human REV3L protein is proteolytically processed by TASP1 that cleaves REV3L at residues D525/G526. Interestingly, we found that the sequence of cleavage site is highly conserved in REV3L homologs from zebrafish to mammals (**Figure S1B**). Moreover, it was previously reported that in MLL and TFIIA proteins, TASP1 CS is located in close proximity to an acidic amino-acids stretch which may play a role in TASP1 recognition site, by facilitating its docking or favoring the positioning of TASP1 active site on the cleavage site (Zhou et al., 2006). In REV3L sequence, the acidic amino-acids stretch is located upstream of the TASP1 CS while in TFIIA and MLL protein sequences, it is located downstream.

TASP1-mediated REV3L cleavage generates two polypeptides that heterodimerize to form a stable complex

To further explore the regulation of native REV3L protein in mammalian cells, we examined if TASP1-mediated cleavage of REV3L takes place at the endogenous level. To do that, cell extracts from asynchronous MEF Flag-REV3L were analyzed by western blotting using anti-Flag antibody. In addition to detect REV3L full length (p360), we also observed a specific band migrating at ~70 kDa which strongly decreased after *Rev3l* siRNA, and was completely missing in mouse fibroblasts used as a negative control (**Figure 4A**). We assume that this band probably corresponds to the cleaved N-terminal products processed by TASP1. However, we noticed that it run at a higher molecular weight (70 kDa) than expected since REV3L is supposed to be cleaved at residues 525-526. Using cell extracts expressing F-hREV3L¹⁻⁵²⁵ we showed that the ectopic REV3L polypeptide run at a similar size than endogenous mREV3L p70 (**Figure S3A**), indicating that mREV3L p70 may correspond to the N-terminal products processed by TASP1.

To determine whether REV3L p70 form is specifically produced by TASP1 cleavage, we downregulated *TASP1* gene using specific siRNA in MEF Flag-REV3L and examined REV3L protein levels 72h later. Interestingly, downregulation of TASP1 led to an increase of REV3L

p360 levels suggesting an important role of TASP1 cleavage in the regulation of REV3L p360 levels. This can explain its short half-life monitored in **Figure 1**. Intriguingly, we did not observe a decrease of REV3L p70 levels upon siTASP1 (**Figure 4B**), suggesting that in contrast to REV3L full length (p360), the cleaved product REV3L p70 has a long half-life allowing its maintenance in cells for many hours. To test this hypothesis, we treated MEF Flag-REV3L cells with CHX and monitored REV3L p70 protein levels. We observed that REV3L p70 is very stable with a half-life longer than 8 hours (**Figure 4C and Figure S3B**), which contrasts with the short half-life of REV3L p360. However, these data did not formally demonstrate that REV3L p70 form is a cleaved product processed by TASP1. Indeed, we cannot rule out that REV3L p70 is a REV3L isoform produced by alternative splicing.

To definitely prove the contribution of TASP1 endopeptidase in the formation of REV3L p70 form, we inactivated *TASP1* gene in MEF Flag-REV3L using CRISPR/Cas9 approach, and isolated independent clones. We also generated KO *Rev3l* clones as positive control. Western blot analysis revealed the disappearance of REV3L p70 form and a great abundance of REV3L p360 in KO *TASP1* clones. As control, we confirmed that the noncleaved TFII2A protein, a TASP1 substrate, accumulated in KO *TASP1* cells (**Figure S3C**). Moreover, we observed a complete loss of both REV3L p70 and REV3L p360 forms in KO *Rev3l* clone (**Figure 4D**). Altogether, these findings demonstrate that REV3L is naturally cleaved by TASP1 and that REV3L p70 is the N-terminal cleaved products. Given that REV3L p70 is extremely stable, we assume that the REV3L cleavage is not a proteolytic process for protein degradation, but rather a dynamic process to form an active complex of processed N-terminal and C-terminal fragments of REV3L as reported for MLL1 or TFIIA complexes (Niizuma et al., 2015).

Therefore, we explored whether the two N- and C-terminal fragments of REV3L could interact with each other to form an α/β mature heterodimer. For that, we fused the N-terminal and C-terminal portion of REV3L with a Flag- and GFP-tag, respectively (**Figure 4E**, top panel). Using co-immunoprecipitation approach (GFP-Trap), we showed that N-terminal domain of REV3L¹⁻⁵²⁵ interacts with the C-terminal fragment REV3L⁷⁶¹⁻³¹³⁰, but not with the constructs REV3L⁷⁶¹⁻¹⁰²⁹, REV3L⁷⁶¹⁻¹⁸⁷⁶ or REV3L⁷⁶¹⁻²⁰⁰¹, indicating that the interacting region is located in the C-terminal fragment between residues 2001 and 3130 (**Figure 4E**, bottom panel).

Ongoing experiments

Define more precisely the interacting region in C-terminal fragment of REV3L that interacts with N-terminal portion of REV3L.

We reasoned that if the catalytic subunit of pol ζ is naturally cleaved by TASP1 to form an active heterodimer, the accessory protein REV7 should interact with this complex. Given that the interacting REV7 domains are located in a region between aa 1870 and 2010 in REV3L protein (**Figure 3B**), REV7 binds the C-terminal fragment of REV3L (p300). We then asked if we can

recover the N-terminal fragment of REV3L after co-immunoprecipitation using anti-REV7 antibody from MEF Flag-REV3L cell extracts. As shown in **Figure 3F**, in addition to REV3L p360, REV3L p70 was detected among proteins that co-immunoprecipitate with REV7, strongly suggesting that REV3L p70/p300 and REV7 form a stable complex in cells.

Ongoing experiments

In an attempt to demonstrate that REV3L p70/p300 and REV7 is an active complex that possesses DNA polymerase activity, we have undertaken two different biochemical approaches. First, in collaboration with Agnès Cordonnier (IBSB, Illkirch) DNA polymerase activity will be assessed on cell extracts from MEF F-REV3Lwt, MEF F-REV3L^{KO TASP1}, and MEF F-REV3L^{KO Rev3l} cells (as negative control). For that, the minimal pol ζ_2 (REV3L/REV7) complex will be enriched by co-immunoprecipitation using REV7 antibody from 10⁸ cells then specific polymerase activity of this complex will be analyzed by primer extension assay. If the cleavage of REV3L is required to promote DNA synthesis activity, we should not see any activity in MEF F-REV3L^{KO TASP1} cell extracts. With the second biochemical approach, we want to purify the minimal hpol ζ_2 complex and test its DNA polymerase activity. This work will be achieved in collaboration with the Charbonnier Lab (I2BC, CEA Saclay) and the platform of Genetic engineering and protein biochemistry – CIGEx (CEA FAR). To do that, plasmids expressing REV3L and its accessory protein MBP-tagged REV7 will be expressed in mammalian cells Epi293. To prevent REV3L cleavage in cells, we will use Epi293 cells in which TASP1 gene will have been inactivated by CRISPR/Cas9 approach. Therefore, we will be able to purify noncleaved pol ζ_2 using tagged-REV7. We will then test the specific polymerase activity of this complex using primer extension assay, after addition or not of purified TASP1 endopeptidase in the samples. This is a way to elucidate whether TASP1-mediated cleavage of REV3L is mandatory for inducing pol ζ polymerase activity.

Cleaved N-Terminal REV3L fragment is hyperphosphorylated following UV-irradiation

As shown above, levels of REV3L p360 levels strongly decrease after UV irradiation due to the short half-life of the protein. Since we found that REV3L p70 has a longer half-life, we asked if its stability was affected in response to UV-irradiation. Remarkably, we observed that REV3L p70 was highly post-translationally modified upon UV irradiation, which was illustrated by specific slower migrating forms detected on western blotting 4 h after 20J/m² (**Figure 5A**). These modifications occur within the first hour and persist for at least 24 h after UV-irradiation (**Figure 5B-C**). Using cell fractionation, we showed that both unmodified and modified REV3L p70 were highly enriched in the insoluble fraction (**Figure 5D**), suggesting that REV3L p70 actively participates in DNA transaction in normal conditions and in response to UV-irradiation. Altogether, these data indicate that REV3L p360 and REV3L p70 respond in a different way to

UV-induced DNA damage, suggesting that these two REV3L polypeptides might play distinct roles in DNA damage tolerance mechanisms.

We further investigated if REV3L p70 was modified in response to other replication stress including aphidicolin (APH) that inhibits activities of replicative polymerases, hydroxyurea (HU) that affects dNTP pools, camptothecin (CPT) that targets topoisomerase I, and formaldehyde (FA), a chemical drug that induces DNA adducts as well as protein-DNA and DNA-DNA crosslinks. Cells were treated and harvested 6 h later, then protein extracts were analyzed by western blot. As shown in **Figure 5E**, all five genotoxic agents elicited REV3L p70 post-translational modifications, even if each drug treatment induced a specific pattern. This suggests that modifications of REV3L p70 is associated with the type of DNA damage. Moreover, we observed that the decrease of REV3L p360 levels was restricted to UV and FA treatment, suggesting that these genotoxic agents induce global transcription inhibition, limiting thus *Rev3l* mRNA amount.

Given that several phosphoresidues have been detected in the N-terminal region of human REV3L (<http://www.phosphosite.org>), we examined whether post-translational modifications of REV3L p70 correspond to an hyperphosphorylation. Therefore, protein extracts from cells treated with genotoxic agents were incubated with or without alkaline phosphatase then analyzed by western blot. Interestingly, phosphatase treatment abolished electrophoretic mobility shifts of REV3L p70 in total extracts as well as in chromatin-bound fraction (**Figure 5F-G**), indicating that REV3L p70 is hyperphosphorylated in response to genotoxic stress. It is interesting to note that slight mobility shifts of REV3L p70 were detected in unstressed cells (**Figure 5G**), suggesting that cleaved N-terminal fragment of REV3L is phosphorylated in physiologic conditions.

A family of phosphoinositol-3-phosphate kinase-like kinases including ataxia telangiectasia-mutated (ATM), ATM and Rad3-related (ATR) and DNA-dependent protein kinase (DNA-PK) are activated in response to DNA damage, and subsequently phosphorylate targets involved in DNA damage response (Blackford and Jackson, 2017).

We thus tested if REV3L p70 is a substrate for one of these kinases by using specific kinase inhibitors. As demonstrated in **Figure 5H**, phosphorylation of REV3L p70 remains unchanged after treatment with kinase inhibitors while an absence of phosphorylation was observed for targets known to be modified by these kinases (including RPA32, KAP1 and Chk1) when genotoxic treatment was associated with kinase inhibitors.

Ongoing experiments

To identify residues that are phosphorylated in physiologic conditions and after genotoxic stress, we performed an immunoprecipitation using Flag antibody from untreated MEF Flag-REV3L or UV-irradiated cells. Immunoprecipitated complexes were loaded on SDS-PAGE gel and after

migration and Coomassie blue coloration, the gel pieces around 70-90 kDa (i.e. modified REV3L p70) were cut and sent to MS.

Cleaved REV3L fragment is enriched on nascent DNA in physiologic conditions and after replication stress

We previously revealed that pol ζ contributes to heterochromatin replication during unperturbed S phase (Ben Yamin *et al.*, submitted article), suggesting that this polymerase can be recruited to replication forks in absence of DNA damage. To assess the biological relevance of the TASP1-cleaved fragment, REV3L p70, we investigated whether it is associated with replication forks in unstressed conditions using iPOND approach. Nascent strands were pulse-labelled with the thymine analogue EdU followed by conjugation of biotin on EdU and purification by streptavidin pull-down. Proteins associated to labelled DNA were then analyzed by western blot. Interestingly REV3L p70 was retrieved in the sample harvested immediately after the pulse but lost in the thymidine-chased sample (**Figure 6A**). This behaviour is similar to the pol ζ subunit, REV7, as well as the one of known replisome components, PCNA, demonstrating that REV3L p70, likely in association with REV3L p300 and REV7, travels with replication forks during unperturbed S phase. We then explored if REV3L p70 was retained at replication forks after replication stress by treating cells with UVC or HU. We found that phosphorylated REV3L p70 was enriched on nascent DNA after 1 h HU treatment. In contrast REV3L p70 was weakly detected immediately after UV irradiation while it was enriched after chase treatment suggesting that phosphorylated REV3L p70 contributes to fill in gap left opposite UV lesions. As a control, Rad51 was also found enriched behind replication forks after UV irradiation as part of a post-replication repair mechanism.

To confirm that REV7 protein still form a complex with REV3L p70 after replication stress, MEF Flag-REV3L cells were exposed to UV or HU treatment then protein complexes enriched after REV7 co-immunoprecipitation were analyzed (**Figure 6B**). We observed that REV7 remains associated with phosphorylated forms of REV3L p70 after replication stress, emphasizing that REV3L p70/p300 and REV7 might be the active form of pol ζ .

Abrogation of TASP1 cleavage site in *Rev3l* gene leads to a novel isoform of REV3L protein

To go further, we decided to explore the biological importance of the proteolytic cleavage of REV3L by TASP1. We thus generated a noncleavable REV3L cell line by mutating at the genomic level the QLDGTAD CS into a QLAATAD noncleavable sequence (REV3L KI^{nc}) using CRISPR-Cas9 technology along with a single stranded DNA oligonucleotide template containing the mutations. We first isolated 2 independent clones. By using Sanger sequencing to verify the mutations in TASP1 CS, we realized that the clones contained different *Rev3l* mutations (**Figure 7A**). While

the clone 21 carried the expected mutations coding for QLAATAD sequence, the clone 14 harbored a 3bp deletion in the genomic sequence, producing thus the amino-acid sequence QDGTAD. Nevertheless, both clones contained a putative noncleavable form of REV3L. We then checked by western blot the loss of REV3L p70 (**Figure 7B**). As expected, we observed a total disappearance of REV3L p70 product in both clones. However, we did not find an increase of REV3L p360 levels. Instead, we detected a clear band running at ~240kDa which was completely missing in MEF Flag-REV3L wt, and was strongly reduced after *Rev3l* siRNA (**Figure 7C**), confirming that this band corresponds to a stable REV3L polypeptide that appears specifically when TASP1 consensus cleavage site is mutated. As REV3L p240 can be detected by the Flag antibody, we concluded that this REV3L fragment contains the N-terminal portion of REV3L.

One possible hypothesis to explain the formation of REV3L p240 fragment is that when REV3L cannot be cleaved by TASP1 at its native site, another endopeptidase might process REV3L at an unknown cleavage site in C-terminal region of REV3L. This hypothesis is rather unlikely, because REV3L p240 should be detected in MEF Flag-REV3L^{KO} *TASP1*. Alternatively, REV3L may contain a cryptic TASP1 cleavage site which is processed only when the canonical CS is mutated. We tested this hypothesis and searched in the amino-acid sequence of REV3L if such a cryptic site may exist. We found a degenerate TASP1 cleavage site (QIDGPSL) at the position 2266-2272 (**Figure S4A**). Note that expected molecular weight of the cleaved REV3L fragment is consistent with the band size detected in REV3L KI^{nc} cell extracts. To test if REV3L p240 was produced by TASP1 cleavage, we inactivated *TASP1* gene in both REV3L KI^{nc} clones, 14 and 21 using CRISPR/Cas9 technique. Positive populations sorted by FACS were then analyzed by western blot. Unfortunately, inactivation of TASP1 enzyme was not associated with disappearance of REV3L p240 (**Figure S4B**), indicating that this polypeptide is not formed by TASP1 cryptic site. This is consistent with the fact that we never detected such a cleaved product when we overexpressed the C-terminal fragment of REV3L (**Figure 4E**). A third hypothesis that could explain the formation of REV3L p240 fragment is that the mutations introduced in the genome to abolish TASP1 CS were located in close proximity to an intron-exon junction, and REV3L p240 could be formed by an aberrant exon skipping event.

Ongoing experiments

We are currently cloning cDNAs isolated from REV3L KI^{nc} cell lines to sequence them and verify if they correspond to exon skipping-induced Rev3l isoforms.

Short half-life of REV3L p360 is likely due to its TASP1-mediated cleavage. Therefore, we should observe an increase of the REV3L p360 half-life in MEF Flag-REV3L^{KO} *TASP1* cell lines. We thus examined the stability of REV3L forms in different clones by treating cells with CHX for indicated time (**Figure 7D**). We found that REV3L full length p360 is extremely stable in MEF

Flag-REV3L^{KO TASP1} cells as compared to wt cells, indicating that the proteolytic cleavage of REV3L p360 is indeed responsible for its short half-life. In addition, REV3L p360 is more stable in KI^{nc} cells than in wt cells, suggesting that molecular mechanism inducing REV3L p240 is less efficient than TASP1-mediated cleavage of REV3L that operates in wt cells. Interestingly, we noticed that after UV-irradiation, REV3L p360 and to a lesser extent p240 were highly phosphorylated in MEF Flag-REV3L^{KO TASP1} and KI^{nc} cell lines, respectively (**Figure 7E**), suggesting that phosphorylation of N-terminal region of REV3L in response to replication stress is independent of REV3L cleavage. In wt cells, phosphorylation of REV3L p360 cannot be detected due to its short half-life. We also confirmed that REV7 is found in a complex with the three forms of REV3L p360, p240 and p70 (**Figure 7F**). Altogether, these data reinforce our hypothesis that REV3L p360 is a precursor which forms a stable and active heterodimer when it is processed by TASP1.

Alteration of the native TASP1-mediated cleavage of REV3L influences cell proliferation, UV sensitivity and genome stability in mammalian cells

We next sought to understand the functional relevance of the TASP1-mediated cleavage of REV3L in mammalian cells. Several studies reported that loss of REV3L protein impairs cell proliferation (Lange et al., 2012, 2013). We thus wondered if inhibition of REV3L cleavage could impact on cell growth. By monitoring the cell proliferation of the different MEF Flag-REV3L cell lines (**Figure 8A**), we observed that the growth rate of MEF Flag-REV3L^{KO TASP1} cells was similar to that of MEF Flag-REV3L^{KO Rev3l} cells, both cell lines grew slower than MEF Flag-REV3L wt. This is consistent with study reported by Hsieh and colleagues showing that TASP1 regulates the expression of key cell cycle regulators (Takeda et al., 2006). However, we do not know if the growth defect observed in REV3L^{KO TASP1} cells is caused by REV3L p360 accumulation or due to the inability to cleave other TASP1 targets. Intriguingly, the clones KI^{nc} displayed different growth rates. Indeed, while the clone KI^{nc} #14 grew faster than the wt cell line, cell growth of clone KI^{nc} #21 was comparable to that of MEF Flag-REV3L^{KO TASP1} cells.

Given that inactivation of *Rev3l* causes sensitivity to DNA damaging agents (Lange et al., 2012; Zander and Bemark, 2004), we asked whether cells impaired in the native TASP1-mediated cleavage of REV3L are sensitive to genotoxic agents. Cells were thus exposed to UVC radiation then relative cell survival was measured 72 h later. Unexpectedly, MEF Flag-REV3L^{KO TASP1} cells tend to be more resistant to UV treatment than the wt cells (**Figure 8B**), suggesting that accumulation of REV3L p360 due to TASP1 inactivation may protect cells from genotoxic insults. Unfortunately, the two REV3L KI^{nc} clones again showed different behaviors. While the clone KI^{nc} #14 displayed a strong UV sensitivity similar to that of REV3L^{KO Rev3l} cells, clone KI^{nc} #21 showed the same sensitivity to UV treatment than wt cells.

Absence of REV3L is known to increase spontaneous chromosomal instability in mammalian cells (Wittschieben et al., 2006). Therefore, we asked if these features are found in cells showing impairment of canonical REV3L cleavage. We first examined spontaneous double-strand breaks (DSB) using the DSB marker 53BP1 which was detected by immunofluorescence (**Figure 8B**). We observed a significant increase of spontaneous 53BP1 foci in both KI^{nc} clones while MEFs lacking TASP1 displayed a comparable number of 53BP1 foci than wt cells (**Figure 8C**), suggesting that mutations in TASP1 cleavage site of REV3L may disrupt REV3L protein folding leading to detrimental effects on genome stability. Finally, we detected a small increase in micronuclei formation in tested cells (**Figure 8D**), except for REV3L^{KO Rev3l} cells as expected (Wittschieben et al., 2006).

Ongoing experiments

Because we found a discrepancy in cell proliferation and UV sensitivity between the two KI^{nc} clones, we assume that one of them has undergone off-target events by CRISPR/Cas9. We are currently analyzing three other independent clones.

DISCUSSION

Although human Pol ζ was discovered >20 years ago (Gibbs et al., 1998), there is still little understanding of how this TLS polymerase functions in mammalian cells. Here we reveal that REV3L protein is naturally proteolytically processed by the endopeptidase TASP1 for generating two polypeptides that heterodimerize to form a stable complex. This finding explains why the level of full length REV3L (p360) protein is low in mammalian cells. Interestingly, cleaved N-terminal REV3L p70 responds to DNA replication stress by being hyperphosphorylated and enriched on nascent DNA. Importantly, our data indicate that the function of REV3L in DNA replication is not limited to exogenous stress since cleaved REV3L localizes at replication forks, even under physiological conditions. This is consistent with our previous results showing that REV3L contributes to heterochromatin replication (Ben Yamin et al., submitted article). Furthermore, alteration of TASP1-mediated cleavage of REV3L influences cell proliferation, UV sensitivity and genome stability in mammalian cells. Altogether, these findings uncover a unique biological process to regulate an error-prone DNA polymerase in mammalian cells.

REV3L is regulated at transcriptional and translational levels

We have shown that expression of *Rev3l* gene can be regulated at transcriptional level with a relatively short mRNA half-life. REV3L p360 protein has also a short half-life, likely due to its TASP1-mediated cleavage. Thus any internal or external cellular stimuli that alter transcriptional and/or translational machineries could impact on REV3L protein levels, exemplified in this study by UV and formaldehyde treatment. This observation is important taking in consideration that potent anti-cancer agents target the basal transcription machinery for cancer therapy (e.g. CDK9 inhibitor, (Morales and Giordano, 2016), which may reduce REV3L protein levels, resulting in a decrease of chemotherapy-induced mutagenesis. However, persistent decrease of REV3L may have deleterious side effects with increased chromosome breaks and chromosome rearrangements (Wittschieben et al., 2006). We have also uncovered that REV3L is cleaved and that the N-terminal processed fragment of REV3L is hyperphosphorylated in response to genotoxic stress, indicating that the activity of REV3L can be modulated by post-translational modifications, such as phosphorylation, and proteolysis. Identifying the kinases involved in this pathway would be of great interest.

TASP1 is the primary protease for REV3L.

In this study, we have provided several lines of evidence that TASP1 is the protease for REV3L. First, TASP1 cleaves REV3L efficiently *in cellulo*, whereas the REV3L cleavage site mutants D525A/G526A cannot be cleaved by TASP1. Second, inactivation of endogenous TASP1

abolishes cleavage of REV3L. Third, REV3L cleavage cannot be detected in *TASP1* KO MEFs, which unequivocally demonstrates that TASP1 is the primary protease for REV3L. Inactivation of *TASP1* greatly increases REV3L p360 protein levels and impairs cell proliferation. Whether reduced cell proliferation is due to an accumulation of uncleaved REV3L is unclear, given that cleavage of other TASP1 substrates, such as MLL or TFIIA, play a role in proliferation (Takeda et al., 2006, 2015). Further investigation is needed to clarify this aspect. *TASP1*^{-/-} knockout mice in which only the uncleaved REV3L form is present die soon after birth. The few surviving *TASP1*^{-/-} mice exhibit small body sizes, craniofacial malformations and impaired telencephalic cell proliferation (Takeda et al., 2015). This indicates that the uncleaved REV3L p360 is not a nonfunctional precursor, but instead plays a critical role in early embryonic development preventing thus embryonic lethality seen in *Rev3l*^{-/-} mice (Esposito et al., 2000; Van Sloun et al., 2002; Wittschieben et al., 2000). Interestingly, recent reports showed that deletion or loss-of-function mutations in human *TASP1* gene lead to developmental delay, microcephaly, distinctive facial features, and multiple congenital anomalies in infant (Suleiman et al., 2018, 2019). Whether aberrant accumulation of uncleaved REV3L contributes to these pathophysiologic features remains to be determined.

REV3L is proteolytically processed to form an active heterodimer

Several lines of evidence support the hypothesis that REV3L forms a stable heterodimer, p70/p300, which corresponds to the catalytically active form of polζ. Indeed, we have shown that C-terminal portion of REV3L is detectable by western blot after TASP1 cleavage (**Figure 3F**), indicating that this region is stable in cells. We also observed that N-terminal fragment of REV3L is able to interact with the C-terminal part of REV3L (**Figure 4E**), supporting the notion that N- and C-terminal cleaved products form a heterodimer. In line with this, the Yang lab has reported the purification of human polζ with accessory subunits using a series of human REV3L deletion constructs expressed in HEK293 cells. Even if they assumed that REV3L cleavage was due to a degradation process, they observed that the C-terminal part of REV3L was stable and could be co-purified with REV7 and the two Polδ subunits, p50 and p66 (Lee et al., 2014). More importantly, when they assessed the polymerase activity of different REV3L variants, they found that uncleaved REV3L (due to deletion in region encompassing TASP1 recognition sequence) has a specific polymerase activity ~2-10% of that of REV3L wt on a DNA template-primer pair, strongly suggesting that TASP1-mediated cleavage of REV3L is mandatory to form an active polymerase complex.

Our results are also consistent with the structural modeling of yeast Rev3 (Rev3p) predicting that C-terminal region immediately downstream of the inactive exonuclease domain folds with the N-terminal domain (Martin and Wood, 2019). In mammalian REV3L protein, a

large insertion (1386 amino acids) corresponding to a single exon separates the N-terminal domain from the C-terminal portion of REV3L which contains most of the DNA polymerase domain. We speculate that this insert may have disrupted the folding of the protein, hampering the catalytic activity of REV3L. One might envision a scenario in which the 360-kDa precursor REV3L (REV3L p360) undergoes TASP1-mediated proteolytic cleavage, which gives rise to the mature REV3L p70/p300 heterodimer with an appropriate folding, restoring thus REV3L enzymatic activity. This is consistent with the fact that conserved recognition sequence for TASP1 is absent in yeast and co-evolved in vertebrates with the emergence of the large exon in *Rev3l* gene.

Polζ travels with replication forks

Our finding that REV3L p70, REV7 (and presumably polζ complex) are found at DNA replication forks in unstressed cells is a strong indication that this TLS polymerase can replicate undamaged templates *in vivo*. This could explain the involvement of polζ in the generation of spontaneous mutations (Lawrence, 2002). We also found that REV3L p70 responds to HU-induced replication stress by being phosphorylated and enriched on nascent DNA. This is consistent with recent work from Kochenova *et al* that demonstrated that yeast polζ synthesis represents a unique mechanism of rescuing stalled replication when dNTP supply is low (Kochenova *et al.*, 2017), contributing thus to HU-induced mutagenesis (Northam *et al.*, 2010). Thus, the potential for polζ to cause mutations is far greater than would be the case if polζ was restricted to TLS function on damaged DNA templates. Further investigation is warranted to identify the functions of REV3L at replication forks, and to define the dependencies for REV3L recruitment to normal and stressed replication forks.

Inhibition of TASP1-mediated cleavage of REV3L as a potential therapeutic target (?)

An important question emerging from this study is whether the cleavage of REV3L by TASP1 constitutes an essential axis for spontaneous and/or DNA damage-induced mutagenesis. One can expect that the precursor REV3L p360 has rather a structural role limiting chromosomal instability, while the mature REV3L p70/p300 heterodimer has an error-prone polymerase activity. Therefore, determining the mutagenesis in MEF REV3L wt as well as in *TASP1* KO, or *KI^{nc}* cell lines after genotoxic treatment, such as cisplatin, should reveal the impact of uncleaved REV3L on drug-induced mutation rate. A potentially exciting prospect for cancer therapy is to inhibit polζ polymerase activity in tumor cells, which should sensitize tumors to therapy as well as prevent the emergence of chemoresistance. Thus, administering TASP1 inhibitors in conjunction with chemotherapeutic agents might produce therapeutic benefits for cancer patients. Small-molecule inhibitors of TASP1 are currently developed for cancer therapy (Boom *et al.*, 2014).

SUPPLEMENTAL INFORMATION

Supplemental information includes four figures.

REFERENCES

- Baranovskiy, A.G., Lada, A.G., Siebler, H.M., Zhang, Y., Pavlov, Y.I., and Tahirov, T.H. (2012). DNA polymerase δ and ζ switch by sharing accessory subunits of DNA polymerase δ . *J. Biol. Chem.* *287*, 17281–17287.
- Bier, C., Knauer, S.K., Klapthor, A., Schweitzer, A., Rekik, A., Krämer, O.H., Marschalek, R., and Stauber, R.H. (2011). Cell-based analysis of structure-function activity of threonine aspartase 1. *J. Biol. Chem.* *286*, 3007–3017.
- Blackford, A.N., and Jackson, S.P. (2017). ATM, ATR, and DNA-PK: The Trinity at the Heart of the DNA Damage Response. *Mol. Cell* *66*, 801–817.
- Boom, J. van den, Mamić, M., Baccelliere, D., Zweerink, S., Kaschani, F., Knauer, S., Bayer, P., and Kaiser, M. (2014). Peptidyl Succinimidyl Peptides as Taspase 1 Inhibitors. *ChemBioChem* *15*, 2233–2237.
- Doles, J., Oliver, T.G., Cameron, E.R., Hsu, G., Jacks, T., Walker, G.C., and Hemann, M.T. (2010). Suppression of Rev3, the catalytic subunit of Pol{zeta}, sensitizes drug-resistant lung tumors to chemotherapy. *Proc. Natl. Acad. Sci. U.S.A.* *107*, 20786–20791.
- Ehrmann, M., and Clausen, T. (2004). Proteolysis as a regulatory mechanism. *Annu. Rev. Genet.* *38*, 709–724.
- Esposito, G., Godindagger, I., Klein, U., Yaspo, M.L., Cumano, A., and Rajewsky, K. (2000). Disruption of the Rev3l-encoded catalytic subunit of polymerase zeta in mice results in early embryonic lethality. *Curr. Biol.* *10*, 1221–1224.
- Fortune, J.M., Pavlov, Y.I., Welch, C.M., Johansson, E., Burgers, P.M.J., and Kunkel, T.A. (2005). *Saccharomyces cerevisiae* DNA polymerase delta: high fidelity for base substitutions but lower fidelity for single- and multi-base deletions. *J. Biol. Chem.* *280*, 29980–29987.
- Gibbs, P.E., McGregor, W.G., Maher, V.M., Nisson, P., and Lawrence, C.W. (1998). A human homolog of the *Saccharomyces cerevisiae* REV3 gene, which encodes the catalytic subunit of DNA polymerase zeta. *Proc. Natl. Acad. Sci. U.S.A.* *95*, 6876–6880.
- Guo, D., Wu, X., Rajpal, D.K., Taylor, J.S., and Wang, Z. (2001). Translesion synthesis by yeast DNA polymerase zeta from templates containing lesions of ultraviolet radiation and acetylaminofluorene. *Nucleic Acids Res.* *29*, 2875–2883.
- Haracska, L., Unk, I., Johnson, R.E., Johansson, E., Burgers, P.M., Prakash, S., and Prakash, L. (2001). Roles of yeast DNA polymerases delta and zeta and of Rev1 in the bypass of abasic sites. *Genes Dev.* *15*, 945–954.
- Harfe, B.D., and Jinks-Robertson, S. (2000). DNA polymerase zeta introduces multiple mutations when bypassing spontaneous DNA damage in *Saccharomyces cerevisiae*. *Mol. Cell* *6*, 1491–1499.
- Hsieh, J.J.-D., Cheng, E.H.-Y., and Korsmeyer, S.J. (2003a). Taspase1: a threonine aspartase required for cleavage of MLL and proper HOX gene expression. *Cell* *115*, 293–303.
- Hsieh, J.J.-D., Ernst, P., Erdjument-Bromage, H., Tempst, P., and Korsmeyer, S.J. (2003b). Proteolytic cleavage of MLL generates a complex of N- and C-terminal fragments that confers protein stability and subnuclear localization. *Mol. Cell Biol.* *23*, 186–194.
- Johnson, R.E., Washington, M.T., Haracska, L., Prakash, S., and Prakash, L. (2000). Eukaryotic polymerases iota and zeta act sequentially to bypass DNA lesions. *Nature* *406*, 1015–1019.
- Johnson, R.E., Prakash, L., and Prakash, S. (2012). Pol31 and Pol32 subunits of yeast DNA polymerase δ are also essential subunits of DNA polymerase ζ . *Proc. Natl. Acad. Sci. U.S.A.* *109*, 12455–12460.
- Khan, J.A., Dunn, B.M., and Tong, L. (2005). Crystal structure of human Taspase1, a crucial protease regulating the function of MLL. *Structure* *13*, 1443–1452.
- Kochenova, O.V., Bezalel-Buch, R., Tran, P., Makarova, A.V., Chabes, A., Burgers, P.M.J., and Shcherbakova, P.V. (2017). Yeast DNA polymerase ζ maintains consistent activity and mutagenicity across a wide range of physiological dNTP concentrations. *Nucleic Acids Res.* *45*,

1200–1218.

Lange, S.S., Wittschieben, J.P., and Wood, R.D. (2012). DNA polymerase zeta is required for proliferation of normal mammalian cells. *Nucleic Acids Res.* *40*, 4473–4482.

Lange, S.S., Bedford, E., Reh, S., Wittschieben, J.P., Carbajal, S., Kusewitt, D.F., DiGiovanni, J., and Wood, R.D. (2013). Dual role for mammalian DNA polymerase ζ in maintaining genome stability and proliferative responses. *Proc. Natl. Acad. Sci. U.S.A.* *110*, E687–696.

Lawrence, C.W. (2002). Cellular roles of DNA polymerase zeta and Rev1 protein. *DNA Repair (Amst.)* *1*, 425–435.

Lawrence, C.W., and Christensen, R.B. (1979). Ultraviolet-induced reversion of *cyc1* alleles in radiation-sensitive strains of yeast. III. *rev3* mutant strains. *Genetics* *92*, 397–408.

Lee, Y.-S., Gregory, M.T., and Yang, W. (2014). Human Pol ζ purified with accessory subunits is active in translesion DNA synthesis and complements Pol η in cisplatin bypass. *Proc. Natl. Acad. Sci. U.S.A.* *111*, 2954–2959.

Lehmann, A.R., Niimi, A., Ogi, T., Brown, S., Sabbioneda, S., Wing, J.F., Kannouche, P.L., and Green, C.M. (2007). Translesion synthesis: Y-family polymerases and the polymerase switch. *DNA Repair (Amst.)* *6*, 891–899.

Lemontt, J.F. (1971). Pathways of ultraviolet mutability in *Saccharomyces cerevisiae*. II. The effect of *rev* genes on recombination. *Mutat. Res.* *13*, 319–326.

Lin, Y.-C., Owen, N., Minko, I.G., Lange, S.S., Tomida, J., Li, L., Stone, M.P., Wood, R.D., McCullough, A.K., and Lloyd, R.S. (2016). DNA polymerase ζ limits chromosomal damage and promotes cell survival following aflatoxin exposure. *Proc. Natl. Acad. Sci. U.S.A.* *113*, 13774–13779.

Makarova, A.V., and Burgers, P.M. (2015). Eukaryotic DNA polymerase ζ . *DNA Repair (Amst.)* *29*, 47–55.

Makarova, A.V., Stodola, J.L., and Burgers, P.M. (2012). A four-subunit DNA polymerase ζ complex containing Pol δ accessory subunits is essential for PCNA-mediated mutagenesis. *Nucleic Acids Res.* *40*, 11618–11626.

Martin, S.K., and Wood, R.D. (2019). DNA polymerase ζ in DNA replication and repair. *Nucleic Acids Res.* *47*, 8348–8361.

McCulloch, S.D., and Kunkel, T.A. (2008). The fidelity of DNA synthesis by eukaryotic replicative and translesion synthesis polymerases. *Cell Res.* *18*, 148–161.

Morales, F., and Giordano, A. (2016). Overview of CDK9 as a target in cancer research. *Cell Cycle* *15*, 519–527.

Morrison, A., Christensen, R.B., Alley, J., Beck, A.K., Bernstine, E.G., Lemontt, J.F., and Lawrence, C.W. (1989). *REV3*, a *Saccharomyces cerevisiae* gene whose function is required for induced mutagenesis, is predicted to encode a nonessential DNA polymerase. *J. Bacteriol.* *171*, 5659–5667.

Nelson, J.R., Lawrence, C.W., and Hinkle, D.C. (1996). Thymine-thymine dimer bypass by yeast DNA polymerase zeta. *Science* *272*, 1646–1649.

Niizuma, H., Cheng, E.H., and Hsieh, J.J. (2015). Taspase 1: A protease with many biological surprises. *Mol Cell Oncol* *2*, e999513.

Northam, M.R., Robinson, H.A., Kochenova, O.V., and Shcherbakova, P.V. (2010). Participation of DNA polymerase zeta in replication of undamaged DNA in *Saccharomyces cerevisiae*. *Genetics* *184*, 27–42.

Northam, M.R., Moore, E.A., Mertz, T.M., Binz, S.K., Stith, C.M., Stepchenkova, E.I., Wendt, K.L., Burgers, P.M.J., and Shcherbakova, P.V. (2014). DNA polymerases ζ and Rev1 mediate error-prone bypass of non-B DNA structures. *Nucleic Acids Res.* *42*, 290–306.

Quah, S.K., von Borstel, R.C., and Hastings, P.J. (1980). The origin of spontaneous mutation in *Saccharomyces cerevisiae*. *Genetics* *96*, 819–839.

Ran, F.A., Hsu, P.D., Wright, J., Agarwala, V., Scott, D.A., and Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. *Nat Protoc* *8*, 2281–2308.

Roche, H., Gietz, R.D., and Kunz, B.A. (1994). Specificity of the yeast *rev3* delta antimutator and *REV3* dependency of the mutator resulting from a defect (*rad1* delta) in nucleotide excision repair. *Genetics* *137*, 637–646.

Sharova, L.V., Sharov, A.A., Nedorezov, T., Piao, Y., Shaik, N., and Ko, M.S.H. (2009). Database for

- mRNA half-life of 19 977 genes obtained by DNA microarray analysis of pluripotent and differentiating mouse embryonic stem cells. *DNA Res.* 16, 45–58.
- Shcherbakova, P.V., Pavlov, Y.I., Chilkova, O., Rogozin, I.B., Johansson, E., and Kunkel, T.A. (2003). Unique error signature of the four-subunit yeast DNA polymerase epsilon. *J. Biol. Chem.* 278, 43770–43780.
- Suleiman, J., Mundt, M., Sampath, S., and El-Hattab, A.W. (2018). TASP1 is deleted in an infant with developmental delay, microcephaly, distinctive facial features, and multiple congenital anomalies. *Clin. Genet.* 94, 170–173.
- Suleiman, J., Riedhammer, K.M., Jicinsky, T., Mundt, M., Werner, L., Gusic, M., Burgemeister, A.L., Alsaif, H.S., Abdulrahim, M., Moghrabi, N.N., et al. (2019). Homozygous loss-of-function variants of TASP1, a gene encoding an activator of the histone methyltransferases KMT2A and KMT2D, cause a syndrome of developmental delay, happy demeanor, distinctive facial features, and congenital anomalies. *Hum. Mutat.* 40, 1985–1992.
- Takeda, S., Chen, D.Y., Westergard, T.D., Fisher, J.K., Rubens, J.A., Sasagawa, S., Kan, J.T., Korsmeyer, S.J., Cheng, E.H.-Y., and Hsieh, J.J.-D. (2006). Proteolysis of MLL family proteins is essential for taspase1-orchestrated cell cycle progression. *Genes Dev.* 20, 2397–2409.
- Takeda, S., Sasagawa, S., Oyama, T., Searleman, A.C., Westergard, T.D., Cheng, E.H., and Hsieh, J.J. (2015). Taspase1-dependent TFIIA cleavage coordinates head morphogenesis by limiting *Cdkn2a* locus transcription. *J. Clin. Invest.* 125, 1203–1214.
- Tomida, J., Takata, K., Lange, S.S., Schibler, A.C., Yousefzadeh, M.J., Bhetawal, S., Dent, S.Y.R., and Wood, R.D. (2015). REV7 is essential for DNA damage tolerance via two REV3L binding sites in mammalian DNA polymerase ζ . *Nucleic Acids Res.* 43, 1000–1011.
- Van Sloun, P.P.H., Varlet, I., Sonneveld, E., Boei, J.J.W.A., Romeijn, R.J., Eeken, J.C.J., and De Wind, N. (2002). Involvement of mouse Rev3 in tolerance of endogenous and exogenous DNA damage. *Mol. Cell. Biol.* 22, 2159–2169.
- Wittschieben, J., Shivji, M.K., Lalani, E., Jacobs, M.A., Marini, F., Gearhart, P.J., Rosewell, I., Stamp, G., and Wood, R.D. (2000). Disruption of the developmentally regulated Rev3l gene causes embryonic lethality. *Curr. Biol.* 10, 1217–1220.
- Wittschieben, J.P., Reshmi, S.C., Gollin, S.M., and Wood, R.D. (2006). Loss of DNA polymerase zeta causes chromosomal instability in mammalian cells. *Cancer Res.* 66, 134–142.
- Xie, K., Doles, J., Hemann, M.T., and Walker, G.C. (2010). Error-prone translesion synthesis mediates acquired chemoresistance. *Proc. Natl. Acad. Sci. U.S.A.* 107, 20792–20797.
- Zander, L., and Bemark, M. (2004). Immortalized mouse cell lines that lack a functional Rev3 gene are hypersensitive to UV irradiation and cisplatin treatment. *DNA Repair (Amst.)* 3, 743–752.
- Zeman, M.K., and Cimprich, K.A. (2014). Causes and consequences of replication stress. *Nat. Cell Biol.* 16, 2–9.
- Zhou, H., Spicuglia, S., Hsieh, J.J.-D., Mitsiou, D.J., Høiby, T., Veenstra, G.J.C., Korsmeyer, S.J., and Stunnenberg, H.G. (2006). Uncleaved TFIIA is a substrate for taspase 1 and active in transcription. *Mol. Cell. Biol.* 26, 2728–2735.
- Zhu, F., Jin, C.-X., Song, T., Yang, J., Guo, L., and Yu, Y.-N. (2003). Response of human REV3 gene to gastric cancer inducing carcinogen N-methyl-N'-nitro-N-nitrosoguanidine and its role in mutagenesis. *World J. Gastroenterol.* 9, 888–893.
- Zhu, X., Zou, S., Zhou, J., Zhu, H., Zhang, S., Shang, Z., Ding, W.-Q., Wu, J., and Chen, Y. (2016). REV3L, the catalytic subunit of DNA polymerase ζ , is involved in the progression and chemoresistance of esophageal squamous cell carcinoma. *Oncol. Rep.* 35, 1664–1670.

EXPERIMENTAL PROCEDURES

Cell culture

SV40-transformed MEF Flag-REV3L (described in Ben Yamin et al, submitted), tumor cell U2OS and primary human TASP1^{-/-} mutated fibroblasts were cultivated in D-MEM (Dulbecco's Modified Eagle medium - Gibco) containing 10% SVF, 1% penicillin-streptomycin. HEK293T human cell line was cultivated in D-MEM containing 10% SVF and 1% penicillin-streptomycin, 1% sodium pyruvate and 1% non-essential amino acids. Primary human normal fibroblasts 405-VI were cultivated in MEM containing 10% SVF, 1% penicillin-streptomycin. All cell lines were cultivated under 5% CO₂.

Drug treatments

For ultraviolet-C irradiation (254 nm), cells were plated 24 hours prior treatment. Before treatment, the culture medium was removed and cells were washed once with PBS and then exposed to 20J/m² (except if indicated otherwise) without any liquid. After irradiation, cells were grown in the old culture medium at 37°C for indicated time and then scraped for protein or RNA extraction. For all genotoxic treatments, cells were plated 24 h prior treatment. Cells were then treated either with 3μM Aphidicolin (APH), 1mM Hydroxyurea (HU), 1μM Camptothecin (CPT) or 200μM Formaldehyde (FA). After treatment, cells were grown at 37°C for indicated time, then scraped for protein extraction. For determining protein half-life, MEFs Flag-REV3L were treated with cycloheximide (CHX) at the final concentration of 10μg/ml and cells were harvested at different time points. All chemical products were purchased from Sigma-Aldrich.

Alkaline phosphatase treatment

After cell lysis in NET-N buffer + protease/phosphatase inhibitors, 20μg of proteins were resuspended in the 1X Phosphatase Buffer with or without 1 unit of FastAP Thermosensitive Alkaline Phosphatase (Thermo Fisher) and incubated at 37°C for 1 hour. After incubation, dephosphorylated proteins were analyzed by western blot.

PI3 Kinases inhibition

MEFs Flag-REV3L were incubated with 2μM NU7441 (DNAPK inhibitor), 10μM KU-55933 (ATM inhibitor) or 10μM VE-821 (ATR inhibitor) independently or in combination when indicated for 1h at 37°C before UV-irradiation at 20J/m². Protein extracts were then analyzed by western blot.

Cell extracts

Cells were lysed with NET-N buffer (150mM NaCl, 1mM EDTA, 50mM Tris pH 7.5, 0.5% NP40) + protease/phosphatase inhibitors for 30 min on ice and sonicated 3 times 10s amplitude 30% using a Vibracell sonicator. Protein extracts were kept in -20°C or loaded on SDS-PAGE gel for western-blot analysis.

Cell Fractionation

MEFs Flag-Rev3L were washed once in cold PBS, scraped in PBS and centrifuged 3 min, 7000rpm at 4°C. For the insoluble fraction, cell pellet was resuspended in 200µl CSK100 (100mM NaCl, 300mM Sucrose, 3mM MgCl₂, 10mM Pipes pH 6,8, 1mM EGTA pH8, 0,2% Triton) + protease inhibitor cocktail, incubated for 5min on ice, and centrifuged for 5min, 5000rpm at 4°C. Pellets were incubated in NET-N for 30min on ice, then DNA was sonicated twice 10s with 30% amplitude. Insoluble protein fraction was then boiling in 1 × Laemmli buffer at 95 °C for 10 min before loading.

Co-immunoprecipitation, GFP-trap and western blot

MEFs Flag-REV3L were washed twice in cold PBS and lysed in NET-N buffer + protease/phosphatase inhibitors for 30 min on ice with gentle shaking then scraped. Sonication was performed using Vibracell, amplitude 30% for 3 cycles of 10s. Cell debris were removed by centrifugation for 10 min at 12000g at 4°C. 10% of the supernatant was used as Input and the remaining supernatant was incubated overnight at 4°C with 2µg anti-Mad2L2/REV7 antibody (ab180579, Abcam) for REV7 co-immunoprecipitation (anti-IgG control antibody was used as negative control). Samples were additionally incubated 2h with 60µl of Protein G Sepharose Beads at 4°C on rollers then beads were washed twice with NET-N 150mM NaCl at 4°C followed by one wash with NET-N 300mM NaCl. Immunoprecipitated proteins were eluted in Laemmli buffer, boiled and loaded on SDS-PAGE (5, 8 or 10% gels). Proteins were then transferred on PVDF membrane, blocked in 5% milk and then incubated overnight at 4°C with primary antibodies (α -GFP ab290 Abcam 1:2000; α -Flag-M2 F1804 Sigma-Aldrich 1:2000; α -Mad2L2 ab180579 Abcam 1:1000; α -Tasp1 α 1:500 ; α -p125 sc8797 Santa Cruz 1:2000; α - β -catenin 610153 BD Transduct 1:10000). Secondary antibodies (Goat polyclonal anti-mouse HRP, Goat polyclonal anti-rabbit HRP, Donkey polyclonal anti-goat HRP; DAKO) were used at 1:5000 and incubated 45 min at RT and revealed using ECL Pierce substrate.

For GFP trap, 25µl of GFP Nanobody/V_HH coupled to agarose beads (Chromotek) were used to pull-down protein complexes. After washing, immunoprecipitated proteins were boiled in laemmli buffer before being loaded.

Immunofluorescence

For Flag-REV3L detection, MEFs Flag-REV3L seeded on coverslips were washed with PBS then pre-extracted using CSK 50 (50mM NaCl, 300mM Sucrose, 3mM MgCl₂, 10mM Pipes pH 6,8, 1mM EGTA pH8, 0,2% Triton) + protease inhibitor cocktail, fixed with 4% paraformaldehyde for 20min, and blocked with PBS-BSA 3% for at least 30 min. Coverslips were incubated with primary antibody anti-Flag M2 (Sigma-Aldrich, F1804) diluted at 1:1000 for 1h at RT. After cells were incubated with the secondary antibody anti-mouse conjugated with the fluorescent protein Alexa-fluor 492 (Life Technologies) for 45 minutes at RT. Coverslips were washed three times with PBS before mounting using a DAKO-DAPI fluorescent mounting medium.

For 53BP1 foci detection, cells seeded on coverslips were washed with PBS, fixed with 4% paraformaldehyde for 20min, permeabilized with 0.5% Triton X-100 for 15 min then blocked with PBS-BSA 3% for at least 30 min. Coverslips were then incubated with primary rabbit antibody anti-53BP1 (Abcam, ab21083) diluted at 1:500 for 1h at RT, then coverslips were processed as described above. Images were quantified by counting the number of foci per nucleus using ImageJ software. (>100 cells)

Cell growth assay

Cells from MEF Flag-REV3L wt, Flag-REV3L^{KO Rev3l}, Flag-REV3L KI^{nc} clones #14 and #21 and MEF Flag-REV3L^{KO TASP1} clones #23, #32 and #33 were plated in triplicate at 2.5.10⁴ cells per well in 6-well plates, then 24h later cells were UV-irradiated or not at 10 J/m². Cells were counted 72 h after irradiation and normalized to the number of nonirradiated cells. For cell proliferation, cells were counted everyday up to 4 days with trypan blue and hemocytometer.

iPOND

MEFs Flag-REV3L were pulse-labelled with 10 µM EdU (Invitrogen) for 10 min. Immediately after the pulse or after a 1 h chase in fresh medium supplemented with 10 µM thymidine (Sigma), cells were crosslinked with 1% formaldehyde (Sigma) in PBS for 15 min at room temperature under gentle agitation. Crosslink was stopped by addition of 125 mM glycine for 5 min. Cells were harvested by scrapping and washed in ice-cold PBS. Pellets were permeabilized in PBS with 0.5% Triton × 100 for 30 min at room temperature. Biotin-azide (Molecular Probes) was conjugated to EdU by click chemistry for 2 h in click reaction buffer (10 mM sodium-l-ascorbate, 10 µM biotin-azide, 2 mM CuSO₄ in PBS). Cells were lysed in iPOND lysis buffer (10 mM Hepes-NaOH pH 7.9, 100 mM NaCl, 2 mM EDTA, 1 mM EGTA, 1 mM PMSF, 0.2% SDS, 0.1% sarkozyl, antiproteases) and sonicated on a Bioruptor device (30 cycles of 30 s on/30 s off at the highest setting). Solubilized chromatin was retrieved by centrifugation at 16,000g for 10 min and supernatant was further incubated overnight with magnetic streptavidin

beads (Dynabeads MyOne Streptavidine C1, Invitrogen). Beads were washed once in lysis buffer, once in 500 mM NaCl and twice in lysis buffer. Proteins were eluted by boiling in 1 × Laemmli buffer at 95 °C for 30 min.

RNA extraction and qRT-PCR

Total RNA extraction was performed using Maxwell RSC simply RNA cell kit (Promega) according to the manufacturer's instructions. 200µg RNA was reverse transcribed to cDNA using RevertAid H minus Reverse Transcriptase (Thermo Scientific) using the following amplification program: 5min at 25°C, 60min at 42°C, 5 min at 70°C. qRT-PCR was performed using the CFX96 Real-time system (Biorad) with 1µl cDNA in Maxima SYBR Green/Rox qPCR Mastermix. Quantities of *Rev3l* mRNA was calculated by dCt method with CFX Maestro software and normalized to GAPDH levels. Following primers were used for qPCR: mGAPDH: F-AGTGCCAGCCTCGTCCCGTAGACAAAATG and R-AAGTGGGCCCCGGCCTTCTCCAT; mRev3L: F-ACAGCGGAGAAGAAACAGGA and R-ACTCTTGGGATCCATTGCTG; hRev3L: F-ACCCCGGAGTACCACTTATCCAGC and R-CCGGAGATATGGTGCCTTGGACA. hGAPDH: F-ACCACAGTCCATGCCATCAC and R-TCCACCACCCTGTTGCTGTA.

Click-iT Nascent RNA Capture

MEF Flag-REV3L were incubated for 24 h with 0.1mM of 5'-Ethylnyluridine (EU) added in culture medium. Cells were washed three times with PBS to stop the incorporation then incubated in fresh culture medium for indicated time. Cells were harvested and total RNA extraction as described above. EU-RNAs were biotinylated by Click reaction and biotinylated RNAs were immunoprecipitated using Dynabeads MyOne Streptavidin T1 magnetic beads according to the manufacturer's instructions (Click-It Nascent RNA Capture kit, Molecular probes). qRT-PCR was performed as described above.

Plasmids constructions and siRNA transfections

The human *Rev3l* full-length cDNA was kindly provided by C. Lawrence (University of Rochester Medical Center, Rochester, NY, USA). To generate the series of GFP-REV3L plasmids, constructs were obtained by polymerase chain reaction (PCR) amplification from *Rev3l* cDNA as a XhoI-EcoRI fragment and cloned into peGFP-C3 expression vector (Clontech). Alternatively, *Rev3l* cDNA fragments were cloned into pFLAG-CMV-4 (Addgene) using EcoRI/KpnI restriction sites. Point mutants were generated by PCR-based methods using a QuikChange® site-directed mutagenesis kit (Stratagene) according to the manufacturer's instructions. All constructs were verified by DNA sequencing. Plasmids were transfected using jetPEI (Polyplus), according to the manufacturer's instructions, and cells were processed 24-48 hrs later. For siRNA transfection,

cells were transfected with 10nM of siRNAs using Interferin (Polyplus) according to the manufacturer's instructions, and were processed 72h later. All siRNAs were purchased from Dharmacon: siRNA mouse *Rev3l* (ON-TARGETplus SMART pool): GAAAUGGAGUGGACCGAUA; AGUCA AUGAUGUUGCGUAU; CGACAU AUCCCGUGUGUUU; ACCCAGAAGUUGUAUAAU, siRNA mouse *TASP1* (ON-TARGETplus Smart pool): CUUAGAUGGGUGAGCGGAA; CUGAAGAGAAGACGACAAU; GAGAAUGUUCACACGCUUU; CGAGCUUGUCAGAAGGCAA, non-target siRNA (ON-TARGETplus non targeting siRNA •1): UGGUUUACAUGUCGACUAA.

Generation of Flag-REV3L^{KI^{nc}} and Flag-REV3L^{KO TASP1} MEFs using CRISPR/Cas9 technique

For MEF Flag-REV3L^{KO TASP1} clones

Single guide RNA (sgRNA) was designed using appropriate software and two single stranded complementary oligos: F- CACCGGCCCTGTAAGGAGAAACGAG and R – AAACCTCGTTTCTCCTTACAGGGCC were annealed, cloned into pX458 plasmid (addgene, Zhang Lab) and processed as previously described (Ran et al., 2013). *Plasmid was then electroporated in MEF Flag-REV3L using* MEF Nucleofection kit II (Amaxa) then 48 h later, GFP positive cells were sorted by FACS and plated at low density to generate independent clones which were expanded and assessed for TASP1 loss by western blot and Sanger sequencing.

For MEF Flag-REV3L^{KI^{nc}} clones

RNP complexes containing purified CAS9 protein (HiFi Cas9 Nuclease V3, IDT) and sgRNA: mU*mC*mU* rGrCrA rGrUrU rCrCrA rUrCrU rArArC rUrGrG rUrUrU rUrArG rArGrC rUrArG rArArA rUrArG rCrArA rGrUrU rArArA rArUrA rArGrG rCrUrA rGrUrC rCrGrU rUrArU rCrArA rCrUrU rGrArA rArArA rGrUrG rGrCrA rCrCrG rArGrU rCrGrG rUrGrC mU*mU*mU* rU (CRISPR-Cas9 sgRNA, IDT) were mixed with a single stranded DNA oligonucleotide template containing the mutations (IDT): G*C*T TTT TGC CAA TCT CTC TAT ACC TCA GTT AGC TGC AAC TGC AGA TGA AAA TAG TGG TAA GAA AAC AAA CCT A*T*T and HDR Enhancer solution (IDT). pEGFP plasmid (0.2 ug) was added to RNP complexes then mix was *electroporated in MEF Flag-REV3L using* MEF Nucleofection kit II (Amaxa). GFP positive cells were sorted 48 h later by FACS and plated at low density to generate independent clones which were expanded and assessed for noncleavable REV3L mutations by western blot and Sanger sequencing.

Genotyping of REV3L-noncleavable KI MEFs

Genomic DNA extraction from each clone was performed using the Nucleospin Tissue kit (Macherey-Nagel). To amplify the genomic region containing mutations, PCR was carried out using specific primers: F1-TTAACCGCTAAGCCACCTC and R2-ACCAAGCAACTCACAAATATCC.

The resulting PCR product was purified with GeneJET PCR Purification kit (Thermo Scientific) and sent for Sanger sequencing (Eurofins Genomics).

***Rev3l* promoter reporter assay**

To determine the proximal region of the transcription start site of human *Rev3l* gene, different softwares have been used: Neural Network promoter prediction, WWW promoter scan, Promoter 2.0, CpG islands prediction and Gpminer, and a 2 kb minimal *Rev3l* promoter region has been identified. Genomic DNA was extracted from human lymphoblastoid cell line, JEFF and PCR was performed using specific primers: F-ATCGAAGCTTGAAGCCCTCGAGCTTTCGTC and R-ATCGGCTAGCAGGGGCTATTGTCAGGTTT containing HindIII and NheI restriction site, respectively. The amplified fragment was inserted in the pGL4 luciferase reporter gene vector (promega), then pRev3l-Luciferase plasmid was co-transfected in HEK293 cells with plasmid containing the Renilla reporter gene as expression control. Twenty-four hours later, cells were rinsed and lysed in Passive Lysis Buffer from the Dual-Luciferase Reporter Assay System according to the manufacturer's instructions (promega). Luciferase and Renilla activity were measured using Wallac Victor 1240® luminometer (Perkin Elmer). The ratio Luciferase / Renilla was calculated to determine Luciferase activity.

Figure 1

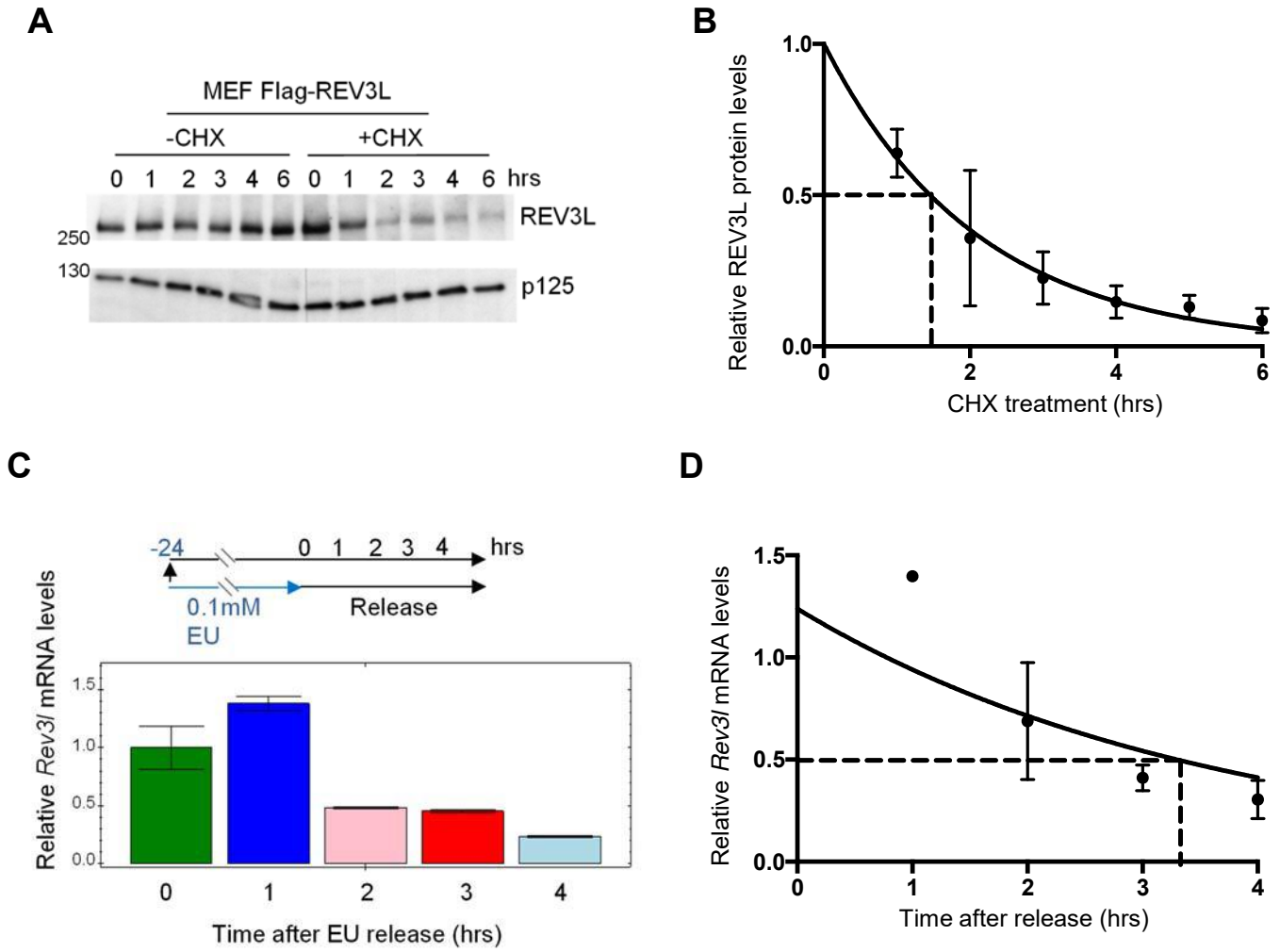


Figure 1: Both REV3L protein and mRNA have a short half-life.

(A) MEF Flag-REV3L were treated or not with 10 μ g/ml Cycloheximide (CHX) and harvested at the indicated time points. Protein extracts were analyzed by western blot using anti-Flag antibody. **(B)** Relative quantification of REV3L protein levels by calculating REV3L signal intensity using ImageJ software and normalized on p125 signal intensity. Error bars indicate +/- sem from 4 independent experiments (n=4) **(C)** MEF Flag-REV3L were incubated with 0.1mM 5'-Ethynyluridine (EU) for 24 h. EU-containing RNAs were processed at the indicated time points after EU release, then reverse-transcribed. cDNAs were then used to quantify *Rev3l* mRNA levels by qRT-PCR using specific primers. Bar chart shows *Rev3l* mRNA levels calculated using CFX Maestro software and normalized on GAPDH levels. **(D)** Linear regression was determined from two independent experiments. Error bars indicate +/- sem (n=2).

Figure 2

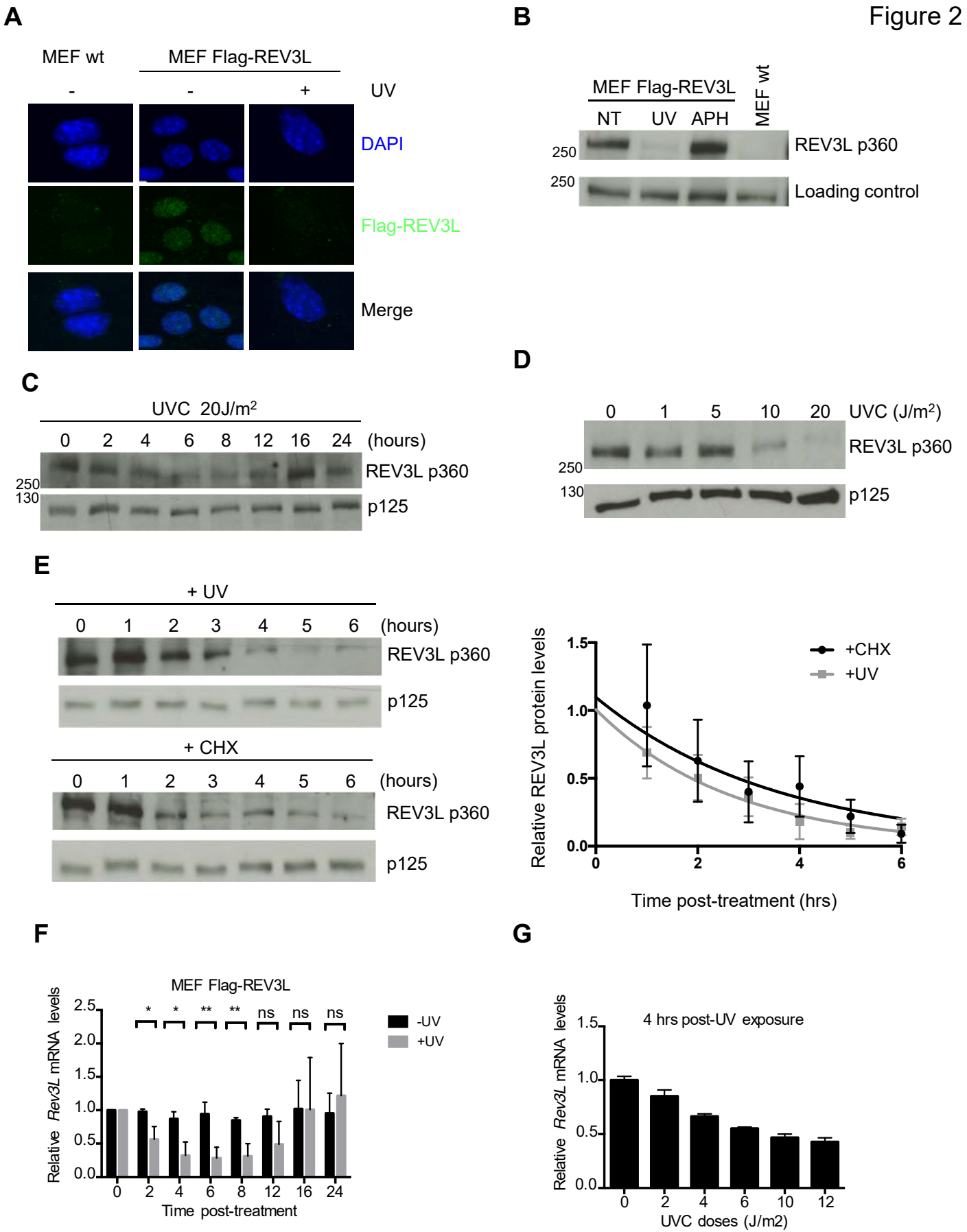


Figure 2: REV3L protein levels strongly decrease after UV exposure in a dose-dependent manner.

(A) Detection of REV3L protein in mock or UV-irradiated cells 4h after 20J/m² with anti-Flag antibody (green staining) in MEF Flag-REV3L using immunofluorescence. MEF wt were used as a negative control. **(B)** Cell extracts from MEF Flag-REV3L non-treated (NT), exposed to 20J/m² UVC (UV) or treated with 3μM aphidicolin (APH) for 4 h were analyzed by western blotting using anti-flag and anti-p125 (loading control) antibodies, **(C)** MEF Flag-REV3L were UV-irradiated at 20J/m², then harvested at different time points up to 24 h. REV3L proteins levels were detected as in (B), **(D)** MEF Flag-REV3L were UV-irradiated at different UV doses then harvested 4 h post-treatment. REV3L proteins levels were detected as in (B), **(E)** MEF Flag-REV3L were UV-irradiated at 20J/m², then harvested at different time points up to 24 h. REV3L proteins levels were detected as in (B). Linear regression of three independent experiments is shown on the right panel, error bars indicate standard deviation (n=3), **(F)** MEF Flag-REV3L were irradiated or not at 20J/m² UVC, then cells were harvested at indicated time points. Total RNAs were extracted, reverse-transcribed and RT-qPCR was performed using specific *Rev3l* primers. Relative *Rev3l* mRNA levels were calculated using CFX Maestro software and normalized to GAPDH mRNA levels. Error bars indicate standard deviation for three independent experiments (n=3) *P<0,05 and ** P<0,005 by Student's t-test. **(G)** MEF Flag-REV3L were exposed to various UV doses, total RNAs were extracted then relative *Rev3l* mRNA levels were determined as in (F) from two independent experiments (n=2).

Figure 3

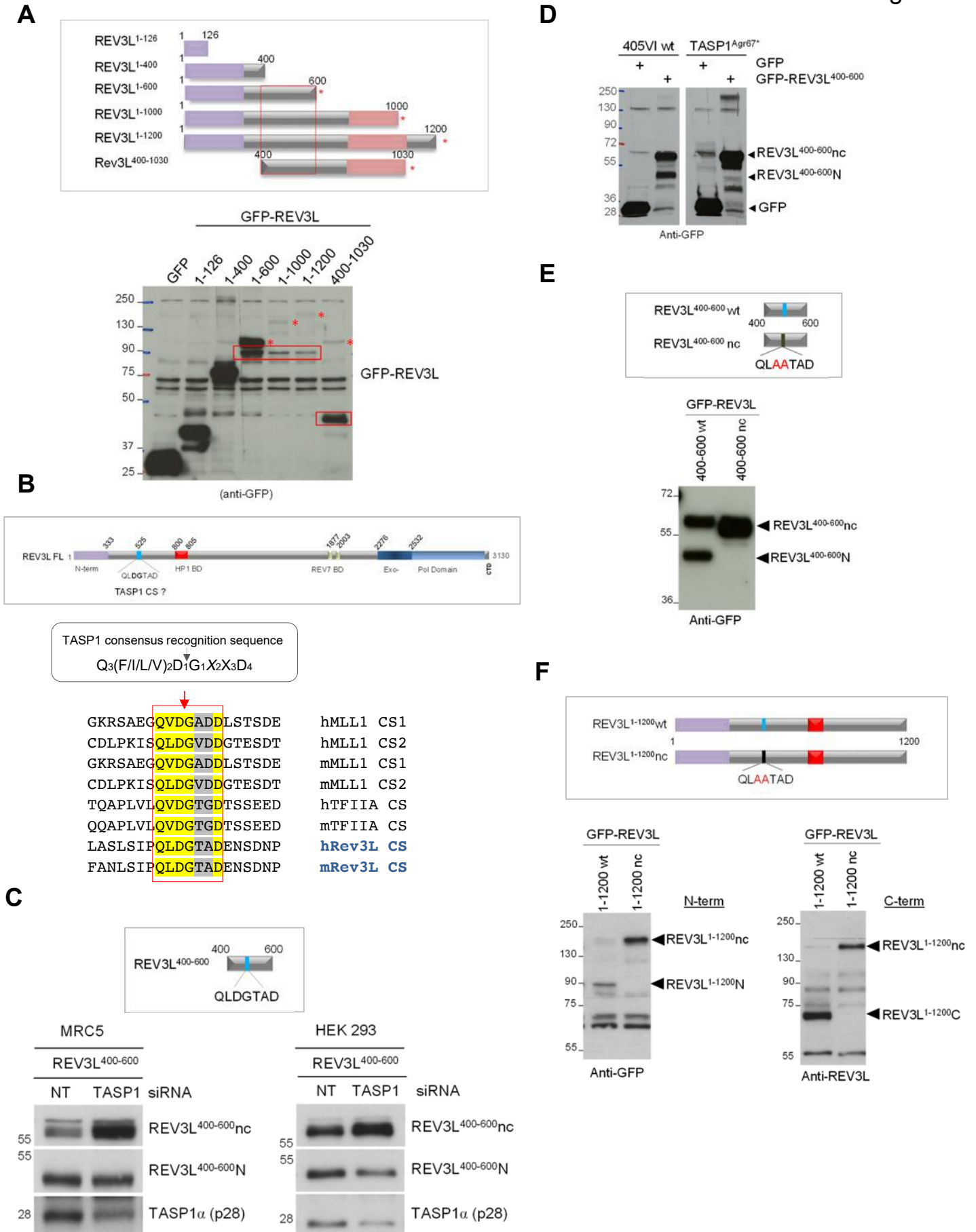


Figure 3: Overexpressed REV3L is cleaved by the endopeptidase TASP1 in mammalian cells.

(A) Scheme of various GFP fusion proteins containing different REV3L constructs (top panel). GFP-REV3L constructs were transfected in HEK293T human cells, then cells were harvested 24 h later. Protein extracts were analyzed by western-blotting using anti-GFP antibody (bottom panel), **(B)** Schematic representation of REV3L full-length protein showing various domains of interest (top panel). Sequence alignment of proteins containing TASP1 recognition sequence, with canonical residues shown in yellow (bottom panel), **(C)** HEK293T and MRC5 human cells were transfected with siRNA against *TASP1* before being transfected with plasmid expressing GFP-REV3L⁴⁰⁰⁻⁶⁰⁰. Protein extracts were then analyzed by western blotting anti-GFP and anti-TASP1 antibodies. Noncleaved (nc) and cleaved N-terminal (N) REV3L fragments were detected **(D)** Primary human normal fibroblasts or primary human fibroblasts mutated in *TASP1* gene were transfected with plasmid expressing GFP-REV3L⁴⁰⁰⁻⁶⁰⁰. Noncleaved (nc) and cleaved N-terminal (N) REV3L fragments were detected by western blotting using anti-GFP antibody, **(E)** HEK293T cells were transfected with plasmids expressing GFP-REV3L^{400-600wt} or GFP-REV3L^{400-600nc} (mutated in TASP1 cleavage site), then Noncleaved (nc) and cleaved N-terminal (N) REV3L fragments were detected by western blotting using anti-GFP antibody, **(F)** HEK293T cells were transfected with plasmids expressing GFP-REV3L^{1-1200wt} or GFP-REV3L^{1-1200nc} (mutated in TASP1 cleavage site), then Noncleaved (nc) and cleaved N-terminal (N) REV3L fragments were detected by western blotting using anti-GFP or anti-REV3L antibodies.

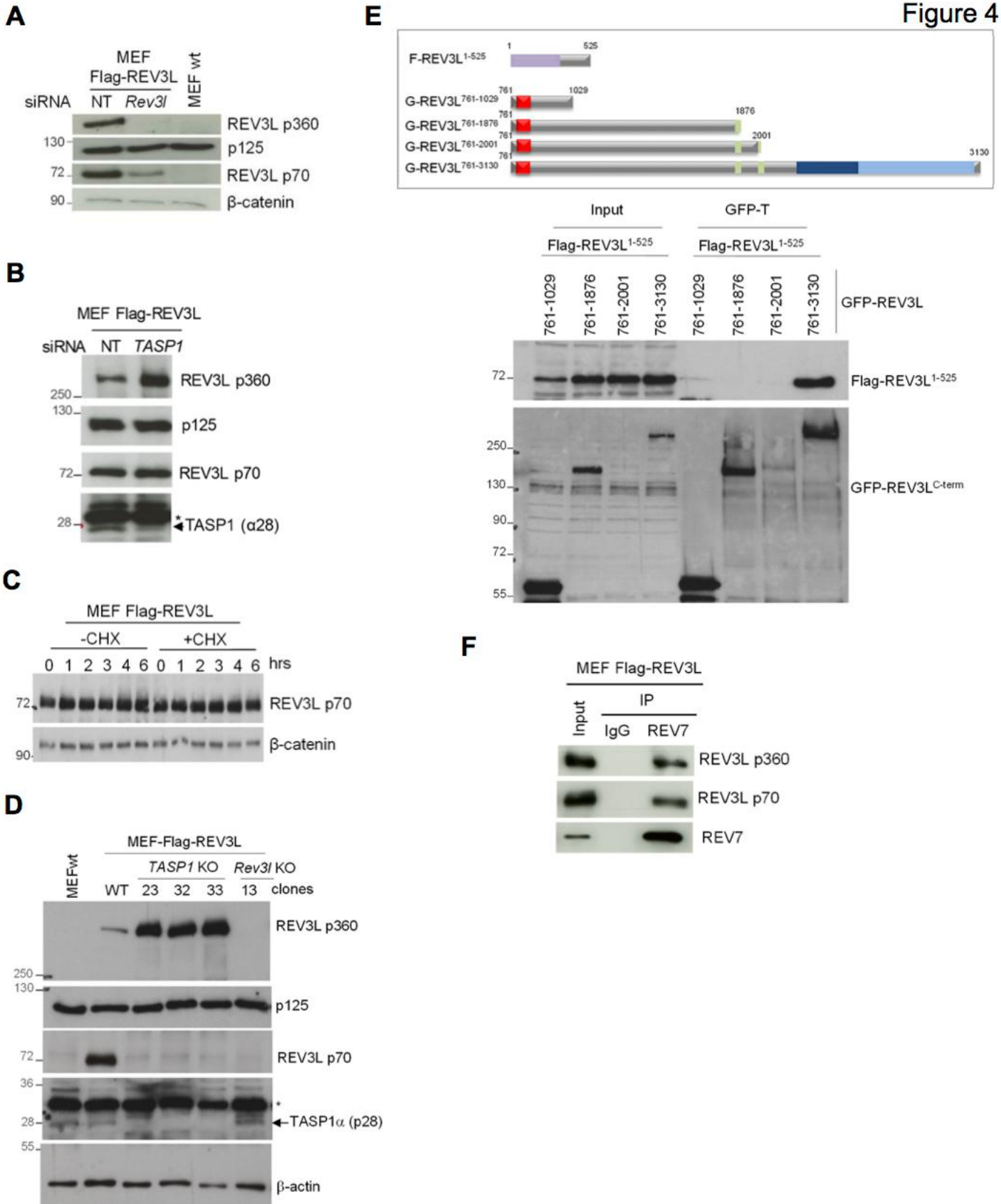


Figure 4: TASP1-mediated cleavage of native REV3L generates a stable cleaved N-terminal fragment that interacts with C-terminal portion of REV3L.

(A) MEF Flag-REV3L were transfected with siRNA against *Rev3l* before being harvested 2 h later. Protein extracts were analyzed by western blotting and REV3L p360 and p70 were detected using anti-Flag antibody. β -catenin and p125 were used as loading control, and MEF wt were used as a negative control. **(B)** MEF Flag-REV3L were transfected with siRNA against *TASP1* before being harvested 2 h later. Protein extracts were analyzed by western blotting. TASP1 downregulation was verified using anti-TASP1 antibody and REV3L p360 and p70 were detected using anti-Flag antibody, **(C)** MEF Flag-REV3L were treated or not with 10 μ g/ml CHX and harvested at indicated time points. Protein extracts were analyzed by western blotting using an anti-Flag antibody. β -catenin was used as loading control. **(D)** MEF Flag-REV3L^{KO} *TASP1* and MEF Flag-REV3L^{KO} *Rev3l* were generated using CRISPR/Cas9 approach then REV3L proteins were detected by western blot using anti-Flag antibody. TASP1 inactivation was verified using anti-TASP1 antibody, **(E)** Scheme of various REV3L constructs (top panel). HEK293T were co-transfected with vectors expressing Flag-REV3L¹⁻⁵²⁵ and one of the four plasmids expressing GFP-tagged C-terminal REV3L constructs as indicated. Co-immunoprecipitation of Flag-REV3L¹⁻⁵²⁵ with GFP-Trap was performed, then analyzed by western blot using anti-Flag and anti-GFP antibodies, **(F)** Immunoprecipitation using anti-Mad2L2/REV7 antibody (or anti-IgG for negative control) was carried out on protein extracts from MEF Flag-REV3L, then immunoblots were performed using anti-REV7 and anti-Flag antibodies.

Figure 5

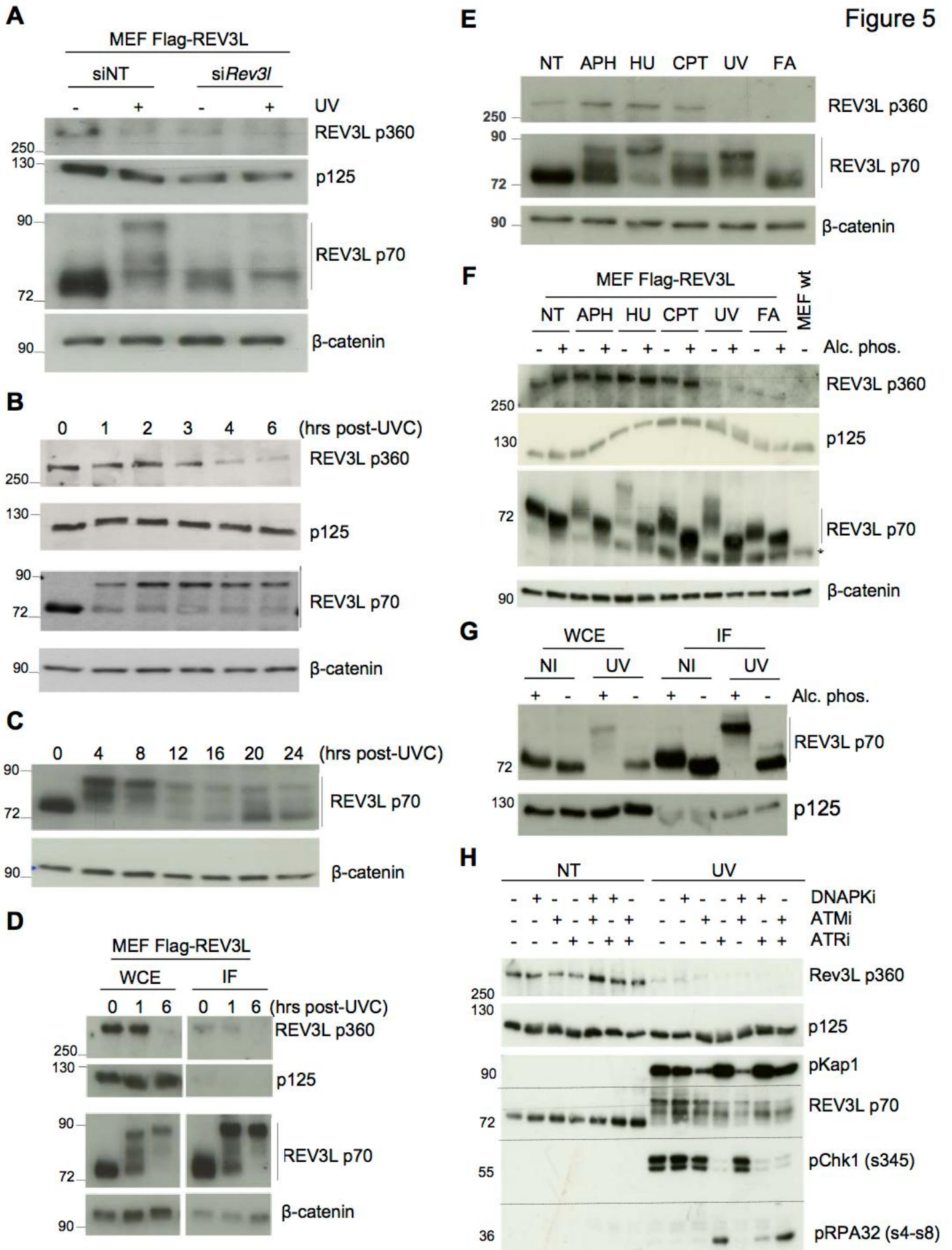


Figure 5: REV3L p70 is hyperphosphorylated following genotoxic stress.

(A) MEF Flag-REV3L were transfected with non-targeting siRNA (NT) or specific *Rev3l* siRNA then UV-irradiated or not at 20J/m² 72 h later. Cells were harvested 6h post-irradiation then protein extracts were analyzed by western blot using anti-Flag antibody to detect REV3L p360 and REV3L p70. β -catenin and p125 were used as loading control. **(B)** MEF Flag-REV3L were UV- irradiated at 20J/m² and harvested at indicated time points. Protein extracts were analyzed as (a), **(C)** MEF Flag-REV3L were UV-irradiated at 20J/m² and harvested at indicated time points. Protein extracts were analyzed as (a), **(D)** MEF Flag-REV3L were UV-irradiated at 20J/m² and insoluble protein fraction (IF) was recovered after cell fractionation and compared with whole cell extracts (WCE) by western blotting using anti-Flag antibody, **(E)** MEF Flag-REV3L were nontreated (NT) or treated with either 3 μ M Aphidicolin (APH), 1mM Hydroxyurea (HU), 1 μ M Camptothecin (CPT), 200 μ M Formaldehyde (FA) or with UV at 20 J/m² (UV), then cells were harvested 6h later. Protein extracts were analyzed by western blot using an anti-Flag antibody. β -catenin was used as loading control. **(F)** MEF Flag-REV3L were treated as indicated in (E) and protein extracts were treated or mock-treated with alkaline phosphatase for 1 h before western blot analysis using anti-Flag antibody. β -catenin and p125 were used as loading control, **(G)** MEF Flag-REV3L were UV- irradiated at 20J/m² then harvested h post-irradiation. whole cell extracts (WCE) or insoluble protein fraction (IF) were treated or mock-treated with alkaline phosphatase before being analyzed by western blotting using anti-Flag antibody, **(H)** MEF Flag-REV3L were treated with either one or two combined kinase inhibitors DNAPKi, ATMi or ATRi for 1 h prior UV-irradiation at 20J/m². Cells were harvested 6 h post-irradiation then protein extracts were analyzed by western blotting using anti-Flag antibody. pKap1, pChk1 and pRPA32 were used as positive controls for kinase inhibitors.

Figure 6

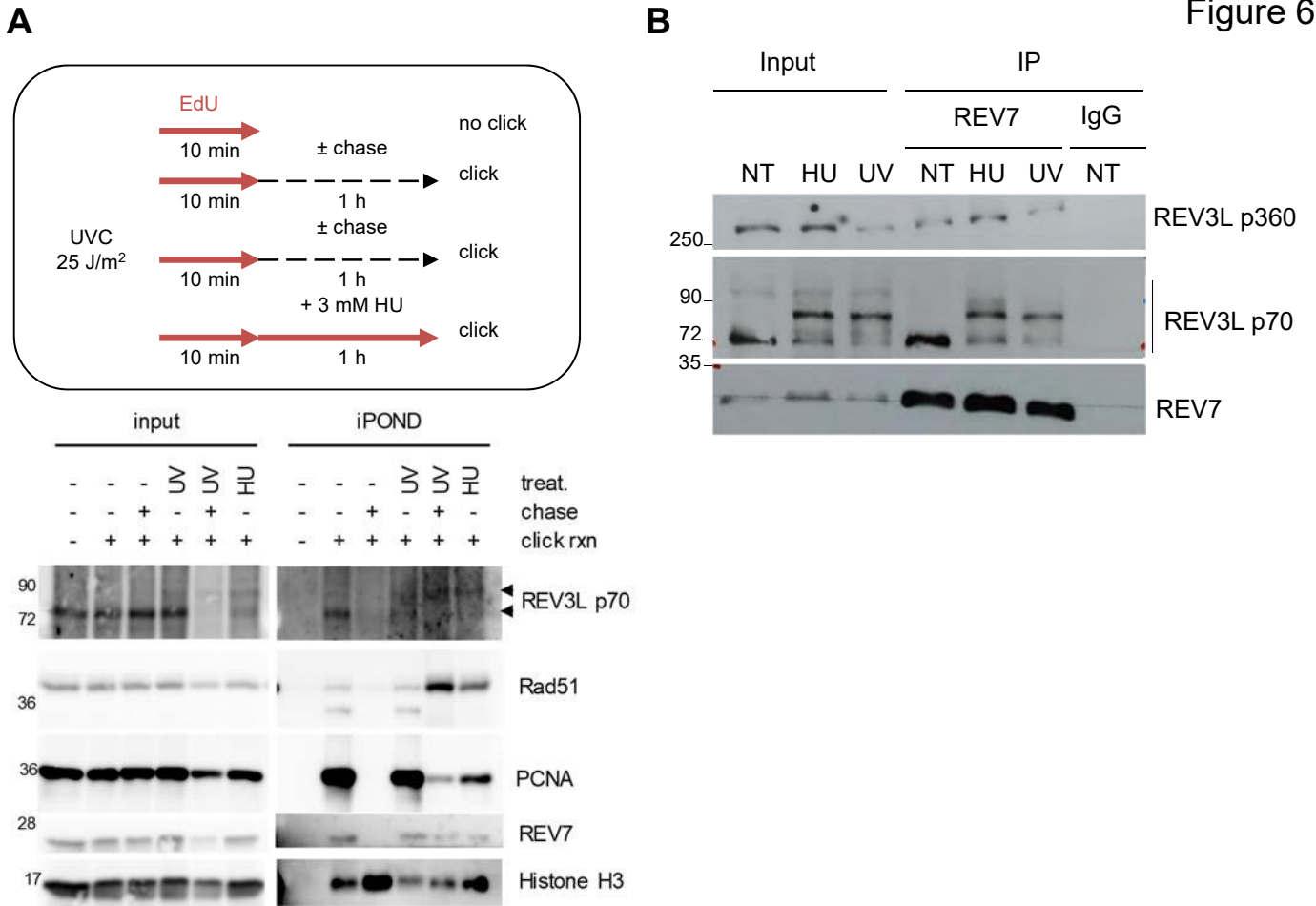


Figure 6: REV3L p70 is enriched on nascent DNA in physiologic conditions and after replication stress.

(A) Scheme for iPOND procedure is shown (Top panel). MEF Flag-REV3L were irradiated or not at $25\text{J}/\text{m}^2$ UVC and pulse-labelled with EdU for 10 min. Cells were then harvested immediately or after 1h thymidine chase to allow replication forks moving away from the labeled DNA. Alternatively, cells were pre-incubated for 10 min with EdU then treated for 1h with 3mM HU and EdU. A whole cell extract (input) and EdU-associated proteins (iPOND) were analyzed by western blot using the indicated antibodies. PCNA was used as positive control for the immunoprecipitation of the replisome, Histone H3 as marker of chromatin, and Rad51 as part of a post-replication repair mechanism. **(B)** MEF Flag-REV3L were nontreated (NT), UV-irradiated at $20\text{J}/\text{m}^2$ (UV) or treated with 1mM Hydroxyurea (HU) for 4 h then cells were harvested. Immunoprecipitation using anti-Mad2L2/REV7 antibody (or anti-IgG for negative control) was carried out on cell extracts, then immunoblots were performed using anti-REV and anti-Flag antibodies.

Figure 7

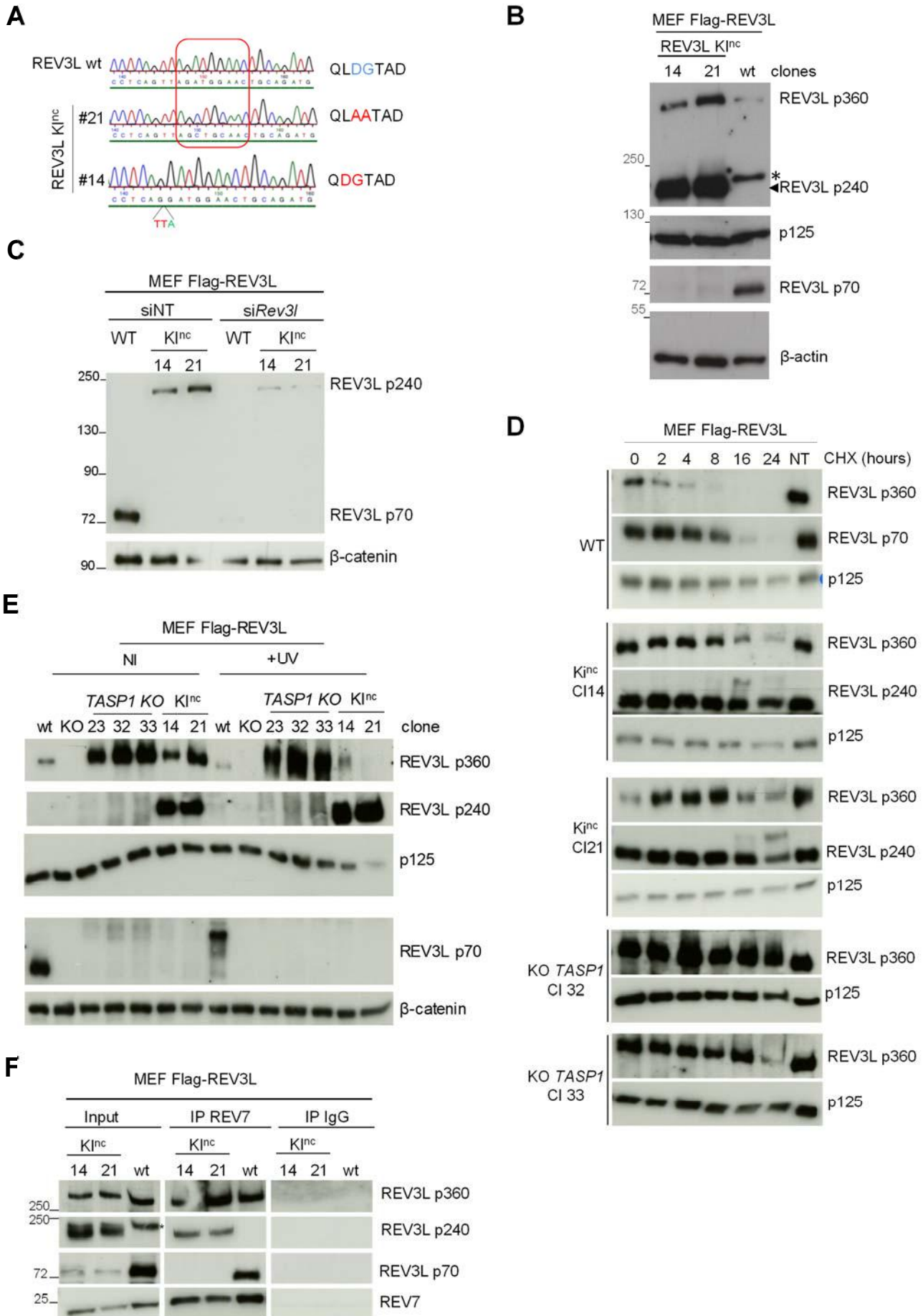


Figure 7: Biochemical characterization of REV3L KI^{nc} and TASP1 KO clones.

(A) Sanger sequencing revealed mutations in the two MEF Flag-REV3L KI^{nc} clones #14 and #21 obtained after CRISPR/Cas9 approach. Corresponding amino acid sequences are noted on the right. **(B)** Cell extracts from MEF Flag-REV3L wt and Flag-REV3L KI^{nc} clones #14 and #21 were analyzed by western blotting using anti-Flag antibody. β -actin and p125 were used as loading control. **(C)** MEF Flag-REV3L wt and Flag-REV3L KI^{nc} clones #14 and #21 were transfected with non-targeting siRNA (NT) or specific *Rev3l* siRNA, then cell extracts were analyzed 2h later by western blot using anti-Flag antibody. β -catenin was used as loading control, **(D)** MEF Flag-REV3L wt, Flag-REV3L KI^{nc} clones #14 and #21 and Flag-REV3L^{KO TASP1} clones #32 and #33 were treated with 10 μ g/ml CHX and harvested at indicated time points. Protein extracts were then analyzed by western blot using anti-Flag antibody. p125 was used as a loading control, **(E)** MEF Flag-REV3L wt, Flag-REV3L^{KO Rev3l}, Flag-REV3L KI^{nc} clones #14 and #21 and MEF Flag-REV3L^{KO TASP1} clones #23, #32 and #33 were UV-irradiated (+UV) or not (NI) at 20J/m² UV and harvested 4 h later. Cell extracts were analyzed by western blotting using anti-Flag antibody. p125 and β -catenin were used as loading controls, **(F)** Immunoprecipitation using anti-Mad2L2/REV7 antibody (or anti- IgG as negative control) was carried out on cell extracts from MEF Flag-REV3L wt and Flag-REV3L KI^{nc} clones #14 and #21 then immunoblots were performed using anti-REV7 and anti-Flag antibodies.

Figure 8

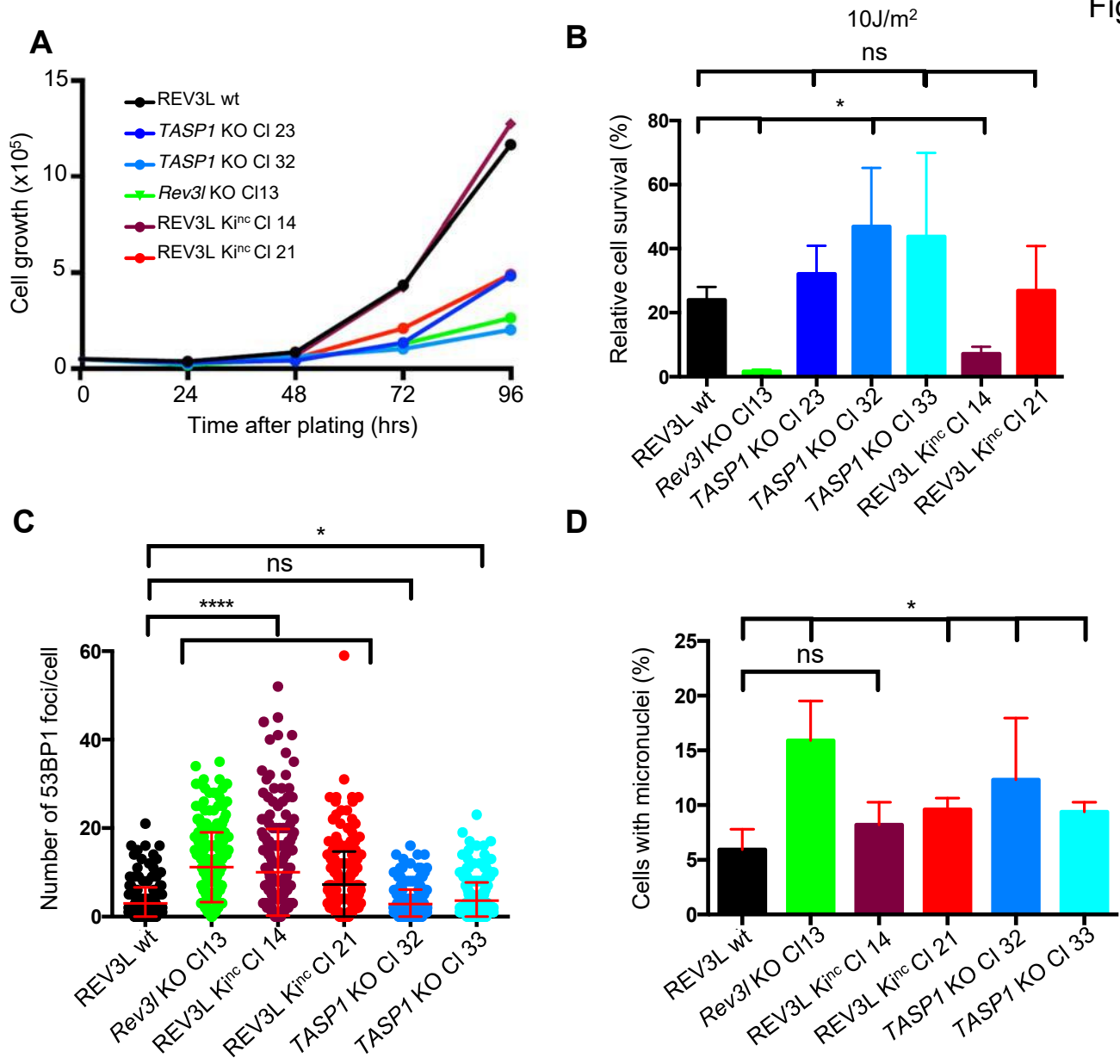


Figure 8: Functional characterization of *TASP1* KO and Rev3L K^{INC} cell lines.

(A) 10⁵ cells of MEF Flag-REV3L wt, Flag-REV3L^{KO Rev3l}, Flag-REV3L K^{INC} clones #14 and #21 and MEF Flag-REV3L^{KO TASP1} clones #23, #32 were plated in duplicate, then cells were counted every 24 h for 96h. **(B)** 10⁵ cells of MEF Flag-REV3L wt, Flag-REV3L^{KO Rev3l}, Flag-REV3L K^{INC} clones #14 and #21 and MEF Flag-REV3L^{KO TASP1} clones #23, #32 and #33 were plated in duplicate then 24h later, cells were UV-irradiated or not at 10 J/m². Cells were counted 72 h after irradiation and normalized to the number of nonirradiated cells. **(C)** Spontaneous 53BP1 foci were detected in MEF Flag-REV3L wt, Flag-REV3L^{KO Rev3l}, Flag-REV3L K^{INC} clones #14 and #21 and MEF Flag-REV3L^{KO TASP1} clones #32 and #33 by immunofluorescence. The graphical representation shows the number of 53BP1 foci per nucleus automatically counted using ImageJ software. Error bars indicate standard error of the mean for two independent experiments (n=2). **(D)** MEF Flag-REV3L wt, Flag-REV3L^{KO Rev3l}, Flag-REV3L K^{INC} clones #14 and #21 and MEF Flag-REV3L^{KO TASP1} clones #32 and #33 were fixed on coverslip and processed. The number of cells displaying at least one micronucleus was reported to the total number of counted cells (n>100 cells). The graphical representation shows the percentage of cells containing at least one micronucleus. Error bars indicate standard deviation for 4 independent experiments.

SUPPLEMENTAL INFORMATIONS

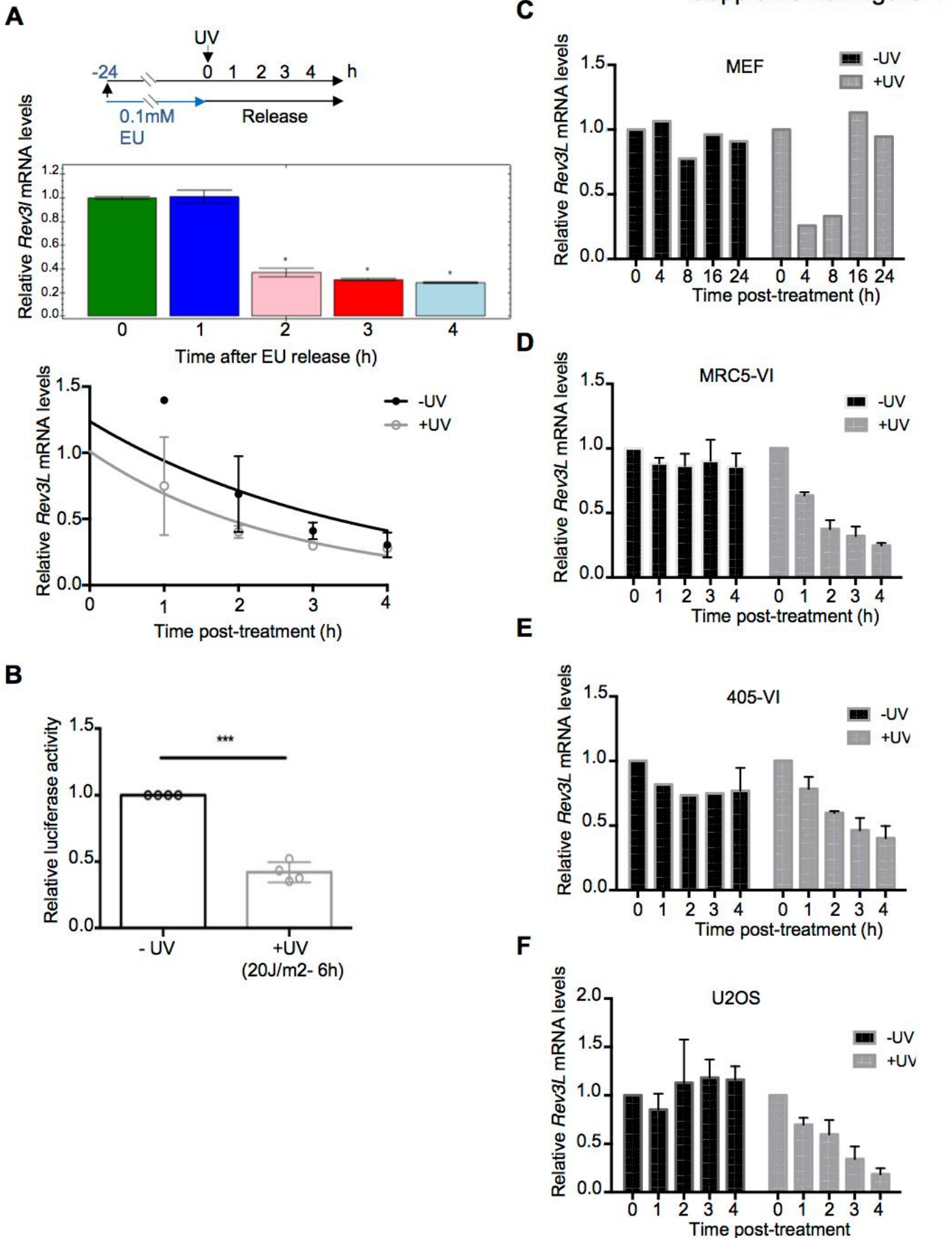
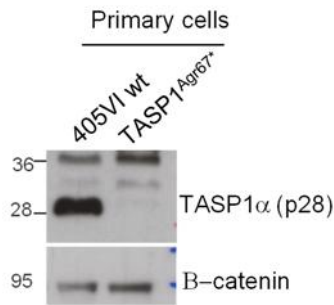


Figure S1, related to Figure 2: UV irradiation inhibits *Rev3l* gene transcription

(A) MEF Flag-REV3L were incubated with EU for 24 h then UV-irradiated or not at 20 J/m². EU-containing RNAs were recovered at indicated time using click chemistry. After reverse-transcription, qPCR was performed using specific *Rev3l* primers. Representative bar chart showing relative quantities of *Rev3l* mRNA calculated by dCt method with CFX Maestro software and normalized to GAPDH levels (top panel). Linear regression with error bars from two independent experiments is shown (bottom panel), **(B)** HEK293T cells were transfected with plasmid expressing luciferase reporter gene under the control of *Rev3l* promoter. Cells were then UV-irradiated or not at 20 J/m² and harvested 4 h later, lysed and luciferase activity was assessed. Error bars indicate standard deviation for 4 independent experiments. *** P<0,0005 by Student's t-test. **(C-E)** MEF wt, SV40-transformed fibroblasts (MRC5), human primary fibroblasts (405VI) and tumor cells (U2OS) were UV-irradiated or not at 20 J/m² then cells were harvested at indicated time points. Total RNAs were extracted and RTqPCR was performed using specific *Rev3l* primers. Relative quantities of *Rev3l* mRNA calculated by dCt method were normalized to GAPDH levels.

A



B

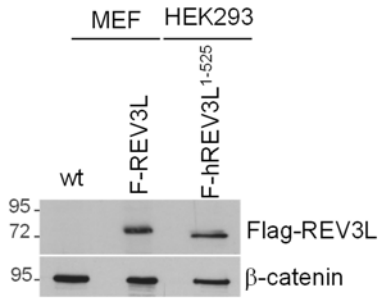
		Acidic stretch			
H. sapiens	486	SLCRNTHRSSTEDDDSSSGEFMEWSDNSLLLASLSIP	Q	LDGTAD	ENSNDP 535
P. Troglodytes	486	SLCRNTHRSSTEDDDSSSGEEMWSDNSLLLASLSIP	Q	LDGTAD	ENSNDP 535
M. mulatta	486	SLCRNTHRSSTEDDDSSSEEMWSDNSLLANLSIP	Q	LDGTAD	ENSNDP 535
C. Lupus	486	SLCRNTHRSSTEDDDSSSEEMWSDNSLLANLSIP	Q	LDGTAD	ENSNDP 535
B. taurus	487	SLCRNTHRSSTEDDDSSSEEMWSDNSLLANLSIP	Q	LDGTAD	ENSNDP 536
M. musculus	486	SLCRNAHRSSTEDDDSSSEEMWTDNSLLFANLSIP	Q	LDGTAD	ENSNDP 535
R. Norvegicus	486	SLCRNAHRSSTEDDDSSSEEMWTDNSLLFANLSIP	Q	LDGTAD	ENSNDP 535
G. gallus	467	SAGKNMHRSTEE-DSSSEEMWSGSNMLLTNLSIP	Q	LDGTAD	ENGDNP 515
D. Rerio	488	LFMK-ANESSEEQQASSDEEGNWKGINSFQTDLSIP	Q	LDGAAD	ESSDSS 536
X. Tropicalis	487	F--KSIIESSPEDED-STDEMEWSRNRLLLADLSIP	Q	LDGTAD	ENSNDP 533

Figure S2, related to Figure 3: TASP1 cleavage site and an acidic stretch in REV3L are well conserved in vertebrates

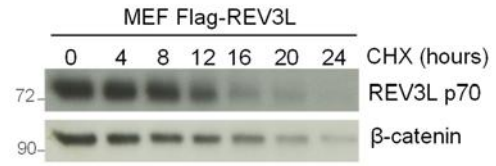
(A) Loss of TASP1 protein in human fibroblasts from patient harboring germinal homozygous mutation in TASP1 gene (c.199 C>T, p.Arg67*). Protein extracts were analyzed by western blot using anti-TASP1 antibody. β -catenin was used as a loading control. **(B)** Sequence alignment of REV3L protein across species shows that, in addition to TASP1 cleavage site, an upstream acidic stretch is also conserved in REV3L in vertebrates.

Supplemental Figure 3

A



B



C

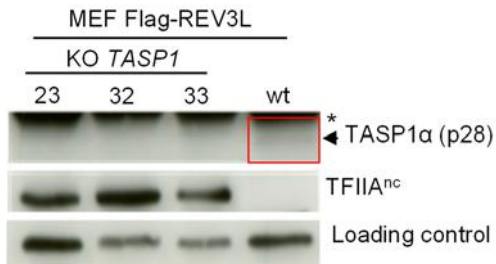


Figure S3, related to Figure 4: Characterization of REV3L p70 in mammalian cells
(A) Validating the molecular weight of REV3L p70. HEK293T cells were transfected with plasmid expressing Flag-hREV3L¹⁻⁵²⁵ then cells were lysed 24 h later. Protein extracts were analyzed by western blot using anti-Flag antibody to compare the molecular weight of endogenous Flag-mREV3L p70 and ectopic Flag-hREV3L¹⁻⁵²⁵ (MEF wt was used as negative control), **(B)** MEF Flag-REV3L were treated or not with 10µg/ml Cycloheximide (CHX) and harvested at the indicated time points. Protein extracts were analyzed by western blot using anti-Flag antibody. β-catenin was used as a loading control. **(C)** Validating the inactivation of *TASP1* gene in MEF Flag-REV3L^{TASP1 KO} clones. Cell extracts from MEF Flag-REV3L^{TASP1 KO} clones #23, #32, #33 and MEF Flag-REV3L wt were analyzed by western blot using anti-TASP1 and anti-TFIIA antibodies. Noncleaved TFIIA (nc) were detected in cells inactivated for *TASP1* gene.

A**Supplemental Figure 4**

Putative non canonical TASP1 cleavage site

		2266	↓	2272	
H.sapiens	2246	LL-TQAKNQFAAVNTPQKETS	Q	IDGPSLNNTYGFKVSIQNLQEA	KALHEI 2294
P.troglodytes	2246	LL-TQAKNQFAAVNTPQKETS	Q	IDGPSLNNTYGFKVSIQNLQEA	KALHEI 2294
M.mulatta	2243	LL-TQAKNQFAAVNTPQKETS	Q	IDGPSLNNTYGFKVSIQNLQEA	KALHEI 2291
C.lupus	2252	LL-TQAKNQFAAVNTPKKETS	Q	IDGPSLNNTYGFKVSVQNLQEA	KALHEI 2300
B.taurus	2249	LL-TQAKNQFAAVNTPKKETS	Q	IDGPSLNNTYGFKVSIQNLLEEA	KALHEI 2297
M.musculus	2237	LLTTQVENQFAAVNTPKKETS	Q	IDGPSLNNTYGFKVSIQNLQEA	KALHEI 2286
R.norvegicus	2247	LLTTQVENQFAAVNTPKKETS	Q	IDGPSLNNTYGFKVSIQNLQEA	KALHEI 2296
G.gallus	2261	LLTTQMKNQFAALGAPKRETS	Q	IEGPSLNNTYGFKVSVQNLQEA	KS LHEV 2310
D.rerio	2066	LLSTQVKQNQLINMNVPKKENS	Q	IEGPFSCNSYGFKVSQNLQEA	KALHEV 2115
X.tropicalis	2246	LLTTQMKNQFATMAQLKRETS	Q	IEGPSLNNSYGFKVSVQNLQDA	KALHEV 2295

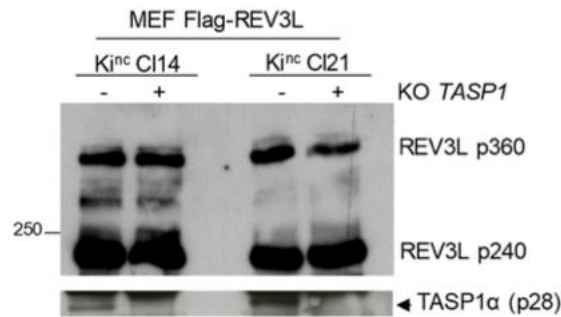


Figure S4, related to Figure 7: Formation of REV3L p240 is not due to a cryptic TASP1 cleavage site in C-terminus of REV3L protein.

(A) Location of putative cryptic TASP1 cleavage site in REV3L sequence at position 2266-2272. Sequence alignment of REV3L across species shows that degenerate recognition sequence of TASP1 “QIDGPSL” is conserved across mammals. **(B)** Inactivation of *TASP1* gene using CRISPR/Cas9 approach was performed in MEF Flag-REV3L KI^{nc} clones #14 and #21, and GFP+ populations were analyzed by western blot using anti-Flag and anti-TASP1 antibodies. This indicates that the formation of REV3L p240 is unaffected by TASP1 loss.

III. Discussion / Perspectives

Dans un premier temps, mes travaux de thèse ont permis de mettre en évidence, dans les cellules de mammifères, un nouveau rôle de l'ADN polymérase ζ dans la réplication non perturbée du génome au niveau de régions hétérochromatiniennes spécifiques, les régions de transitions temporelles (RTTs). Dans un deuxième temps, mes recherches ont permis de mieux comprendre les mécanismes de régulation de la sous-unité catalytique de Pol ζ , REV3L, dans des conditions physiologiques et suite à stress génotoxique dans les cellules de mammifères, ce qui était jusqu'alors inconnu. Ces résultats ont permis de mettre en lumière un nouveau mécanisme de régulation unique au sein de la famille des ADN polymérases dans lequel REV3L p360 serait activée par clivage post-traductionnel par l'endopeptidase TASP1. Je propose que la forme active de REV3L corresponde à l'hétérodimère REV3L p70/p300 formé par les deux fragments résultant de ce clivage.

Je soulève dans cette partie III les nouvelles questions engendrées suite aux résultats obtenus au cours de ma thèse et comment, dans le futur, nous pourrions y répondre.

Implication de Pol ζ dans la réplication non perturbée de l'hétérochromatine.

Il était déjà envisagé que Pol ζ soit impliquée dans la réplication non perturbée des cellules de mammifères. En effet, l'absence de Pol ζ au cours du développement embryonnaire conduit à une létalité embryonnaire précoce démontrant son rôle primordial au cours du développement. De plus, l'absence de Pol ζ dans les cellules de mammifères conduit à une augmentation de l'expression des sites fragiles communs en absence de stress répliatif, renforçant son rôle au cours de la réplication non perturbée de ces régions génomiques particulières. Cependant, il n'y avait jusqu'alors aucune démonstration du rôle effectif de Pol ζ au cours de la réplication physiologique des cellules de mammifères. Nous avons mis en évidence, en absence de REV3L, un retard du timing de réplication associé à une augmentation de la mutagenèse dans 5% du génome localisés spécifiquement au niveau des RTTs. Ces régions étant répliquées plus lentement, majoritairement par une fourche de réplication unidirectionnelle unique, notre hypothèse est qu'en absence de REV3L, les fourches de réplication sont davantage ralenties dans ces régions, conduisant à un blocage du réplisome et potentiellement à une augmentation du nombre de cassures double brin. Cette hypothèse est renforcée par notre observation d'une augmentation des cassures spontanées de l'ADN spécifiquement au niveau de l'hétérochromatine en absence de REV3L. Afin de vérifier cette hypothèse, il serait nécessaire de mesurer la vitesse des fourches de réplication spécifiquement au niveau des régions dont le timing de réplication est retardé en absence de REV3L par peignage moléculaire de l'ADN de MEFs *Rev3l* WT et KO. Des sondes ADN spécifiques des régions choisies et couplées à des fluorophores seraient ensuite hybridées par la technique de FISH sur les fibres d'ADN pour reconnaître spécifiquement ces régions génomiques et y calculer la vitesse des fourches de

réplication (Mendez-Bermudez et al., 2018). Si notre hypothèse est valide, en absence de REV3L la vitesse des fourches de réplication dans ces régions devrait être plus lente que dans les cellules WT. Il serait aussi intéressant de calculer la vitesse de réplication dans d'autres régions hétérochromatiniennes connues comme difficiles à répliquer telles les régions péri-centromériques et les télomères dont on ne peut pas déterminer le timing de réplication par la technique de Répli-chip que nous avons utilisée. Des sondes PNA commerciales spécifiques de ces régions sont disponibles.

Pol ζ est spécifiquement localisée dans l'hétérochromatine grâce à l'interaction de REV3L avec plusieurs protéines décrites comme spécifiques de ce domaine telles que HP1 α et ATRX. Cependant, son mécanisme de recrutement reste encore inconnu. Deux hypothèses sont envisageables : (1) REV3L est pré-chargée au niveau de l'hétérochromatine grâce à son interaction avec la protéine HP1 α et pourra prendre le relais des polymérase répliquatives lorsque le réplisome s'engage dans la réplication de ces régions génomiques. La présence de régions difficiles à répliquer engendrerait un ralentissement des fourches pour les polymérase répliquatives, ce qui entraînerait une modification post-traductionnelle de PCNA et le recrutement de Pol ζ à proximité. (2) REV3L se déplace avec le réplisome au cours de la réplication et le stress répliquatif engendré par ces régions difficiles à répliquer permet l'échange entre polymérase répliquatives et Pol ζ . Ce scénario est facilement envisageable pour la réplication du brin « retardé » étant donné que Pol ζ partage deux sous-unités (p50/p66) avec Pol δ , responsable de la synthèse de ce brin d'ADN. Nos études préliminaires par iPOND semblent montrer que REV3L serait capable de voyager avec la fourche de réplication. Il est cependant nécessaire de répéter l'expérience. Cette étude n'a trouvé que REV3L p70 (et non p360) complexée avec REV7, et associée à des protéines du réplisome comme PCNA.

Notre hypothèse est donc que REV3L p360 aurait plutôt un rôle structural et de plateforme pour d'autres protéines permettant le maintien de la stabilité chromosomique/hétérochromatinienne. En effet, HP1 α et ATRX sont connues pour contribuer à l'organisation, la compaction et le maintien des territoires hétérochromatiniens. Au cours de nos différentes expériences, nous avons observé la coexistence des deux formes REV3L p70 et p360 au sein des cellules mammifères. Il est donc possible que REV3L p360, de par son interaction avec ces deux protéines, soit aussi essentielle au maintien de l'organisation de l'hétérochromatine. Cette hypothèse est soutenue par l'augmentation de l'hétérochromatinisation observée en absence de REV3L, qui servirait alors de barrière pour réguler ces régions spécifiques. D'un autre côté, le complexe REV3L p70/p300/REV7 représenterait la forme active de Pol ζ possédant l'activité polymérase et participant à la résolution des structures secondaires dans l'hétérochromatine au cours de la réplication non perturbée de l'ADN et à la mutagenèse spécifique de Pol ζ . De plus, l'observation de la forme

hyperphosphorylée de REV3L p70 associée aux fourches de réplication suite à une déplétion du pool de nucléotides suggère que l'activité du complexe REV3L p70/p300/REV7 est modulée par modifications post-traductionnelles suite à un stress réplicatif. Même si nous avons des évidences expérimentales que ce complexe REV3L p70/p300/Rev7 pourrait être la forme active Pol ζ , une démonstration par des approches biochimiques est indispensable. Nous avons entamé des collaborations en ce sens (voir Article 2). Comme nous l'avons fait pour les MEF *Rev3l^{+/+}* et *Rev3l^{-/-}*, il serait aussi très intéressant de comparer la mutagenèse induite par des agents génotoxiques, type cisplatine, dans les lignées cellulaires KO *TASP1* et *REV3L* K^{inc} afin de déterminer si le clivage de REV3L par TASP1 est indispensable pour induire la mutagenèse dépendante de Pol ζ . Ceci démontrerait *in cellulo*, l'activité catalytique du complexe REV3L p70/p300/REV7.

De manière intéressante, nous avons trouvé après analyse par spectrométrie de masse, une interaction entre REV3L et plusieurs protéines de la réparation notamment la protéine ScaI. Inversement, une étude par spectrométrie de masse de ScaI a révélé REV3L comme partenaire majeur (Isobe et al., 2017). Nous avons confirmé cette interaction par co-immunoprécipitation. De façon intéressante, il a été rapporté que ScaI promeut la réparation des cassures double brin localisées au niveau de l'hétérochromatine de manière dépendante de 53BP1 (Hansen et al., 2016). Or, nos résultats suggèrent également que REV3L pourrait faciliter la réparation des cassures double brin localisées dans l'hétérochromatine. Des investigations sont en cours au laboratoire pour comprendre la fonction de l'interaction REV3L-ScaI en condition physiologique et suite à l'induction de cassures double brin. Une hypothèse serait que REV3L faciliterait la réparation des cassures double brin dans l'hétérochromatine grâce à son interaction avec différents partenaires de la réparation tels que REV7, ScaI et d'autres protéines identifiées en spectrométrie de masse. REV3L p360 pourrait servir de plateforme pour recruter ces protéines afin de promouvoir la réparation des cassures de l'ADN dans les régions hétérochromatiniennes. Pour vérifier cette hypothèse, il serait intéressant d'étudier la localisation de chacune de ces protéines en absence de REV3L afin de déterminer si REV3L sert bien de plateforme après irradiation des cellules aux rayons X.

Nous avons également montré que REV3L localise à proximité de TRF1, une protéine chaperon télomérique, suggérant une localisation télomérique de Pol ζ . De plus, l'absence de REV3L induit une augmentation des cassures télomériques ainsi qu'une diminution de la taille des télomères. Des candidats potentiels identifiés dans notre spectrométrie de masse sont en cours de caractérisation pour mieux comprendre la fonction de Pol ζ dans la réplication des régions télomériques. L'expérience de peignage moléculaire associée à un FISH télomérique en présence ou en absence de REV3L pourrait apporter un début de réponse à cette question en permettant de calculer la vitesse des fourches de réplication spécifiquement au niveau des télomères. Une

deuxième expérience envisageable serait un test de répllication télomérique *in vitro* en présence ou en absence de REV3L.

Mise en évidence d'un nouveau mécanisme de régulation de REV3L

Au cours de la seconde partie de ma thèse, je me suis intéressée aux différents mécanismes régulant Pol ζ et plus particulièrement sa sous-unité catalytique REV3L. L'observation d'une demi-vie courte à la fois de l'ARN et de la protéine REV3L a été surprenante compte tenu de la grande taille du gène et de la protéine REV3L nécessitant un certain temps à le transcrire et à la traduire. En fait, nous avons montré que c'est uniquement la protéine REV3L p360 qui a une demi-vie courte. Ceci est dû à sa protéolyse par TASP1 alors que la forme clivée REV3L p70 a une demi-vie relativement longue. Par conséquent, une altération du processus de transcription entraîne une diminution rapide de REV3L p360, mais pas de ses formes clivées. Nous l'avons montré suite à une irradiation des cellules aux UVs. En effet, les UVs entraînent une inhibition globale de la transcription à l'exception de certains gènes de réponse aux dommages de l'ADN. Sachant que Pol ζ joue un rôle important dans le franchissement des lésions induites par les UVs, il nous est apparu paradoxal d'observer une disparition de sa sous-unité catalytique, REV3L p360, après irradiation. Cette observation renforce notre hypothèse que le complexe REV3L p70/p300/REV7 représente la forme active de Pol ζ .

Nous avons également décrit un nouveau mécanisme de régulation post-traductionnelle où REV3L est clivée par TASP1 en générant un fragment de 70kDa correspondant à la partie N-terminale de REV3L. Nous n'avons cependant pas pu détecter la partie C-terminale de REV3L résultant du clivage. Il serait donc nécessaire soit de produire un nouvel anticorps spécifique de cette région, soit, comme cela a été fait pour les MEFs Flag-REV3L, introduire une étiquette endogène à l'extrémité C-terminale du gène *Rev3l*. Toutefois, les sous-unités p50/p60 interagissent avec REV3L via son extrémité C-terminale. Ainsi, l'introduction d'une étiquette pourrait altérer la formation du complexe Pol ζ . La modélisation de la structure de REV3L de levure prédit qu'une région de 45 acides aminés immédiatement en aval du domaine exonucléase situé en C-terminal de REV3L s'insère dans le domaine N-terminal. Notre hypothèse est que cette séquence en position 2535-2579 dans la séquence de REV3L humaine serait la région minimale d'interaction entre les fragments clivés N- et C-terminal. Nous allons donc ôter cette séquence de notre construction GFP-REV3L C-terminale (761-3130) puis vérifier par co-immunoprécipitation si le fragment N-terminal (1-525) interagit ou pas avec le fragment C-terminal de REV3L.

Après fractionnement cellulaire, nous avons observé que les formes clivées et la forme non-clivée de REV3L coexistent sur la chromatine. Si notre hypothèse précédente suggérant que la forme clivée de REV3L est la forme active de la polymérase, une nouvelle question est de savoir comment REV3L est activée par son clivage par TASP1 dans le noyau ? Est-elle clivée lorsqu'elle

est localisée sur la chromatine ? dans le nucléoplasme ? quel signal active TASP1 pour cliver REV3L ? Peut-on imaginer qu'une modification post-traductionnelle de REV3L (type phosphorylation) facilite le recrutement de TASP1, et le clivage de REV3L ?

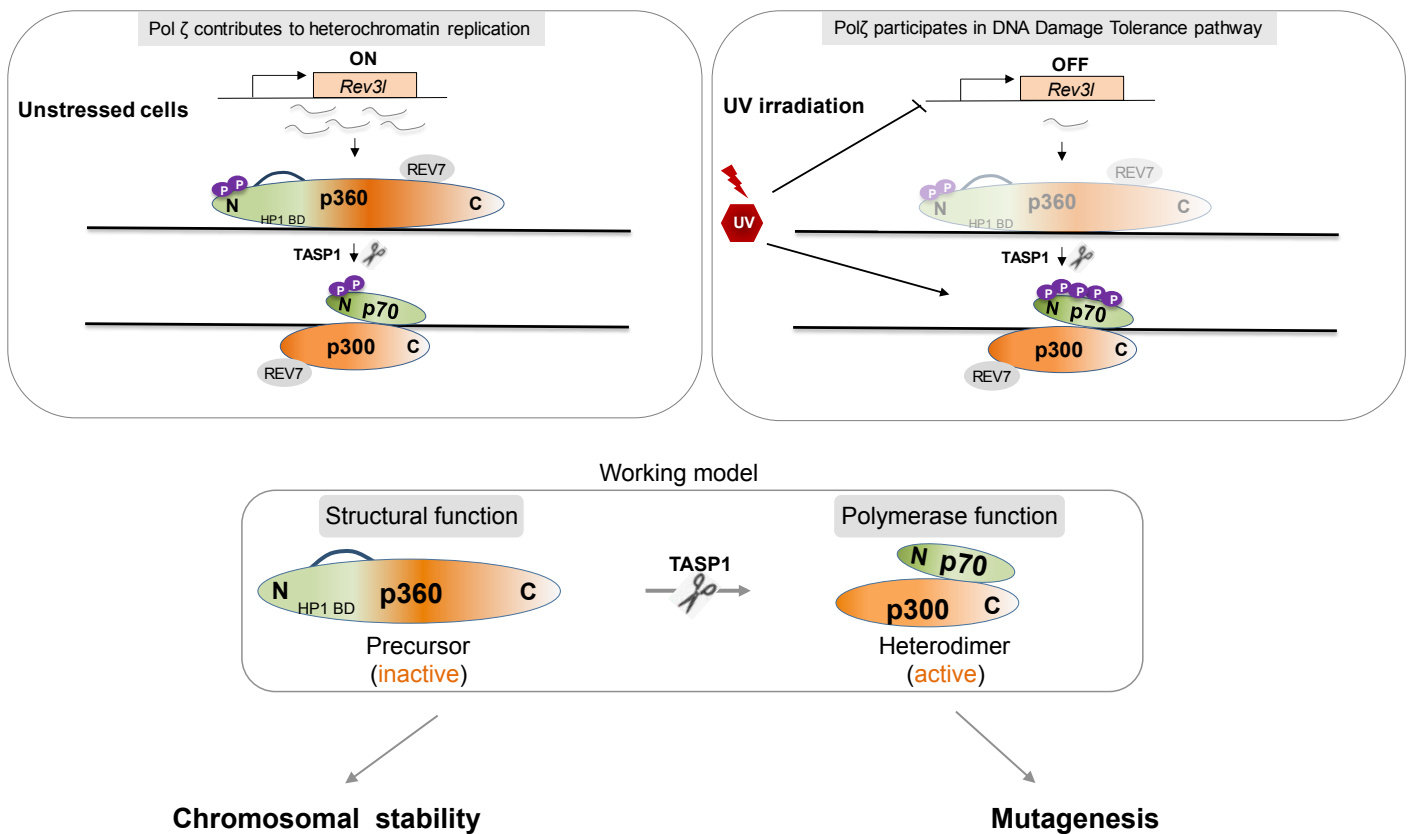
Si, outre son activité catalytique, Pol ζ possède aussi une fonction structurale comme évoquée précédemment, la présence de la protéine dépourvue d'activité catalytique à la chromatine pourrait avoir un impact sur la stabilité génétique. Lorsque le clivage de REV3L est inhibé dans les MEF *REV3L* K^{INC} ou KO *TASP1*, la présence de la protéine non activable pourrait représenter un obstacle à la progression de la fourche de réplication et pourrait alors être encore plus délétère que l'absence complète de la protéine ou bien d'autres mécanismes pourraient être mis en place pour contrebalancer son absence. Cependant, nous savons que les souris *TASP1* KO meurent précocement après la naissance, suggérant que la forme non clivée de REV3L p360 joue un rôle important dans le développement embryonnaire, sinon nous aurions dû observer une létalité embryonnaire comme cela est le cas pour les souris *Rev3l* KO. Nous établissons actuellement des lignées de souris *REV3L* K^{INC}, ce qui nous donnera une première réponse de l'importance du clivage de REV3L *in vivo*. Toutefois, comme nous l'avons observé dans les MEFs *REV3L* K^{INC}, la présence de l'isoforme REV3L p240 induite par la mutation QLDGTAD>QLAATAD pourrait générer une activité résiduelle et ne nous permettra pas de conclure sur l'absence totale de clivage de REV3L. Il sera donc nécessaire de comprendre la provenance de cette nouvelle isoforme. Après avoir rejeté l'hypothèse d'un second site cryptique de REV3L pour TASP1, nous étudions actuellement l'hypothèse d'un problème au niveau de l'ARN. La position de la mutation induite étant proche d'un site donneur d'épissage, une modification de la structure de cette région pourrait induire un phénomène d'échappement d'exon entraînant la formation d'une protéine plus courte. La présence d'un messenger tronqué va donc être investiguée par extraction d'ARN suivie d'une RT-PCR et d'un séquençage dans les MEFs WT et *REV3L* K^{INC}.

Finalement, Pol ζ a été impliquée dans le franchissement de divers types de lésions de l'ADN ainsi que dans les mécanismes de résistance à la chimiothérapie, c'est pourquoi la compréhension de la régulation de REV3L après exposition à un agent génotoxique est importante. Pour ce faire, il serait intéressant d'étudier des différents partenaires de REV3L par des approches de spectrométrie de masse en incluant une analyse des modifications post-traductionnelles retrouvées après différents stress génotoxiques. Cette étude nous permettrait d'élucider les différents motifs d'hyperphosphorylation observés et d'identifier potentiellement les kinases responsables de ces modifications. Elle nous permettrait également d'identifier les partenaires spécifiques en réponse à chaque stress génotoxique et ainsi de déterminer les voies de signalisations activées selon la drogue utilisée pour potentiellement éviter les phénomènes de chimiorésistance.

IV. Conclusion

Pol ζ est une ADN polymérase spécialisée dont la fonction canonique est la synthèse translésionnelle. Cependant, nous avons montré au cours de ce travail qu'elle possède aussi un rôle dans la réplication physiologique de l'hétérochromatine dans les cellules de mammifères dans lesquelles cette polymérase tient un rôle prépondérant pour le maintien de la stabilité génétique et épigénétique. Ces observations sont cohérentes avec le caractère crucial que tient Pol ζ dans le développement embryonnaire. Cependant, une dérégulation de cette polymérase entraîne une instabilité génétique capable d'induire une carcinogenèse ou, plus tard, de la chimiorésistance. C'est pourquoi l'étude de la régulation de Pol ζ dans des conditions physiologiques ou suite à un stress génotoxique permettra dans le futur, nous l'espérons, de mieux comprendre ce mécanisme spécifique de carcinogenèse et, peut-être, de mettre au point de nouvelles thérapies pour certains des cancers. Finalement, si nous arrivons à montrer que le clivage de REV3L par TASP1 est nécessaire à son activité, l'utilisation d'inhibiteurs de TASP1 couplée à la chimiothérapie représenterait une solution de traitement contre certains cancers.

A



B

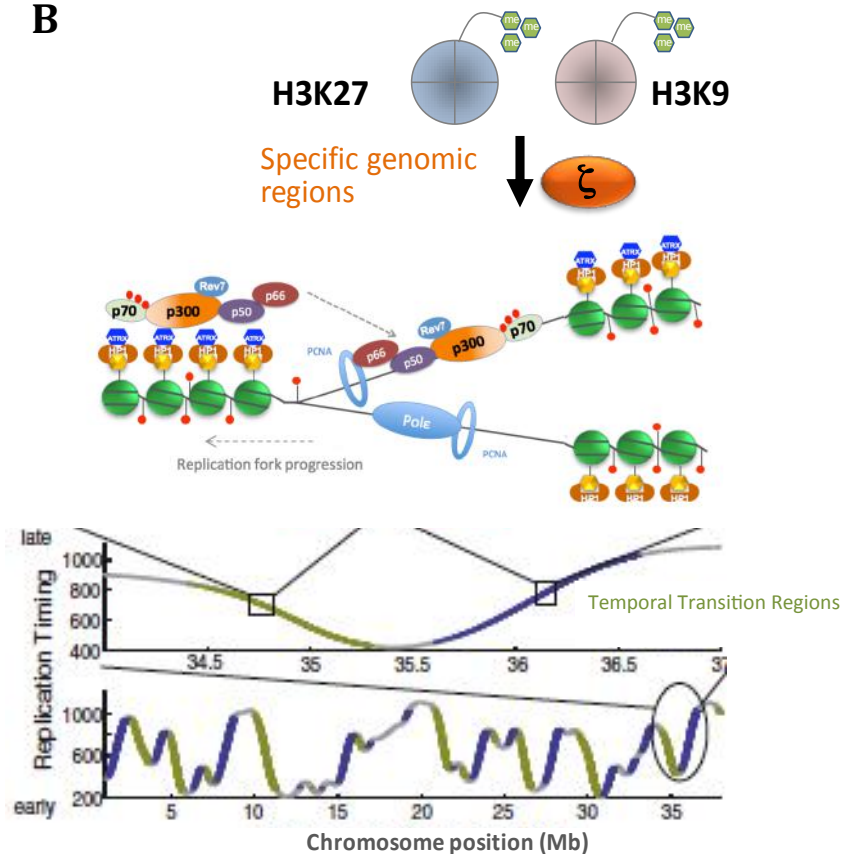


Figure 32 : Conclusion A) Synthèse graphique des résultats obtenus au cours de ma thèse. Modèle des différents mécanismes de régulation de REV3L au niveau transcriptionnel et post-traductionnel et modèle d'activation de Pol ζ . REV3L p360 est un précurseur inactif possédant un rôle structural nécessaire à la stabilité chromosomique. Il va être clivé par TASP1 de manière post-traductionnelle pour obtenir la polymérase active possédant l'activité polymérase nécessaire à la réplication de l'hétérochromatine et à la mutagenèse. B) Pol ζ est nécessaire à la réplication de l'hétérochromatine, spécifiquement au niveau de régions de transitions temporelles, grâce à son interaction avec HP1 α et ATRX.

V. Références

- Abdulovic, A.L., Minesinger, B.K., and Jinks-Robertson, S. (2008). The effect of sequence context on spontaneous Polzeta-dependent mutagenesis in *Saccharomyces cerevisiae*. *Nucleic Acids Res.* *36*, 2082–2093.
- Acharya, N., Haracska, L., Johnson, R.E., Unk, I., Prakash, S., and Prakash, L. (2005). Complex formation of yeast Rev1 and Rev7 proteins: a novel role for the polymerase-associated domain. *Mol. Cell. Biol.* *25*, 9734–9740.
- Ahmad, K., and Henikoff, S. (2002). Histone H3 variants specify modes of chromatin assembly. *Proc. Natl. Acad. Sci. U. S. A.* *99 Suppl 4*, 16477–16484.
- Ahringer, J., and Gasser, S.M. (2018). Repressive Chromatin in *Caenorhabditis elegans*: Establishment, Composition, and Function. *Genetics* *208*, 491–511.
- Alexandrov, L.B., Nik-Zainal, S., Wedge, D.C., Aparicio, S.A.J.R., Behjati, S., Biankin, A.V., Bignell, G.R., Bolli, N., Borg, A., Børresen-Dale, A.-L., et al. (2013). Signatures of mutational processes in human cancer. *Nature* *500*, 415–421.
- Allis, C.D., and Jenuwein, T. (2016). The molecular hallmarks of epigenetic control. *Nat. Rev. Genet.* *17*, 487–500.
- Allshire, R.C., and Madhani, H.D. (2018). Ten principles of heterochromatin formation and function. *Nat. Rev. Mol. Cell Biol.* *19*, 229–244.
- Alper, B.J., Job, G., Yadav, R.K., Shanker, S., Lowe, B.R., and Partridge, J.F. (2013). Sir2 is required for Ctr4 to initiate centromeric heterochromatin assembly in fission yeast. *EMBO J.* *32*, 2321–2335.
- Baranovskiy, A.G., Lada, A.G., Siebler, H.M., Zhang, Y., Pavlov, Y.I., and Tahirov, T.H. (2012a). DNA polymerase δ and ζ switch by sharing accessory subunits of DNA polymerase δ . *J. Biol. Chem.* *287*, 17281–17287.
- Baranovskiy, A.G., Lada, A.G., Siebler, H.M., Zhang, Y., Pavlov, Y.I., and Tahirov, T.H. (2012b). DNA polymerase δ and ζ switch by sharing accessory subunits of DNA polymerase δ . *J. Biol. Chem.* *287*, 17281–17287.
- Barnes, C.E., English, D.M., and Cowley, S.M. (2019). Acetylation & Co: an expanding repertoire of histone acylations regulates chromatin and transcription. *Essays Biochem.* *63*, 97–107.
- Barski, A., Cuddapah, S., Cui, K., Roh, T.-Y., Schones, D.E., Wang, Z., Wei, G., Chepelev, I., and Zhao, K. (2007). High-resolution profiling of histone methylations in the human genome. *Cell* *129*, 823–837.
- Baumann, C., Schmidtman, A., Muegge, K., and De La Fuente, R. (2008). Association of ATRX with pericentric heterochromatin and the Y chromosome of neonatal mouse

spermatogonia. *BMC Mol. Biol.* *9*, 29.

Bemark, M., Khamlichi, A.A., Davies, S.L., and Neuberger, M.S. (2000). Disruption of mouse polymerase zeta (Rev3) leads to embryonic lethality and impairs blastocyst development in vitro. *Curr. Biol. CB* *10*, 1213–1216.

Benetti, R., Gonzalo, S., Jaco, I., Schotta, G., Klatt, P., Jenuwein, T., and Blasco, M.A. (2007). Suv4-20h deficiency results in telomere elongation and derepression of telomere recombination. *J. Cell Biol.* *178*, 925–936.

Bernstein, B.E., Mikkelsen, T.S., Xie, X., Kamal, M., Huebert, D.J., Cuff, J., Fry, B., Meissner, A., Wernig, M., Plath, K., et al. (2006). A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* *125*, 315–326.

Bhat, A., Andersen, P.L., Qin, Z., and Xiao, W. (2013a). Rev3, the catalytic subunit of Pol ζ , is required for maintaining fragile site stability in human cells. *Nucleic Acids Res.* *41*, 2328–2339.

Bhat, A., Andersen, P.L., Qin, Z., and Xiao, W. (2013b). Rev3, the catalytic subunit of Pol ζ , is required for maintaining fragile site stability in human cells. *Nucleic Acids Res.* *41*, 2328–2339.

Bier, C., Knauer, S.K., Klapthor, A., Schweitzer, A., Reikik, A., Krämer, O.H., Marschalek, R., and Stauber, R.H. (2011). Cell-based analysis of structure-function activity of threonine aspartase 1. *J. Biol. Chem.* *286*, 3007–3017.

Black, J.C., Van Rechem, C., and Whetstone, J.R. (2012). Histone lysine methylation dynamics: establishment, regulation, and biological impact. *Mol. Cell* *48*, 491–507.

Blackford, A.N., and Jackson, S.P. (2017). ATM, ATR, and DNA-PK: The Trinity at the Heart of the DNA Damage Response. *Mol. Cell* *66*, 801–817.

Blasco, M.A. (2005). Telomeres and human disease: ageing, cancer and beyond. *Nat. Rev. Genet.* *6*, 611–622.

Bolden, J.E., Peart, M.J., and Johnstone, R.W. (2006). Anticancer activities of histone deacetylase inhibitors. *Nat. Rev. Drug Discov.* *5*, 769–784.

Boom, J. van den, Mamić, M., Baccelliere, D., Zweerink, S., Kaschani, F., Knauer, S., Bayer, P., and Kaiser, M. (2014). Peptidyl Succinimidyl Peptides as Taspase 1 Inhibitors. *ChemBioChem* *15*, 2233–2237.

Brondello, J.-M., Pillaire, M.J., Rodriguez, C., Gourraud, P.-A., Selves, J., Cazaux, C., and Piette, J. (2008). Novel evidences for a tumor suppressor role of Rev3, the catalytic subunit of Pol zeta. *Oncogene* *27*, 6093–6101.

Brownell, J.E., Zhou, J., Ranalli, T., Kobayashi, R., Edmondson, D.G., Roth, S.Y., and Allis,

- C.D. (1996). Tetrahymena histone acetyltransferase A: a homolog to yeast Gcn5p linking histone acetylation to gene activation. *Cell* 84, 843–851.
- Burgers, P.M.J., and Kunkel, T.A. (2017). Eukaryotic DNA Replication Fork. *Annu. Rev. Biochem.* 86, 417–438.
- Burgers, P.M., Koonin, E.V., Bruford, E., Blanco, L., Burtis, K.C., Christman, M.F., Copeland, W.C., Friedberg, E.C., Hanaoka, F., Hinkle, D.C., et al. (2001). Eukaryotic DNA polymerases: proposal for a revised nomenclature. *J. Biol. Chem.* 276, 43487–43490.
- Cadet, J., and Davies, K.J.A. (2017). Oxidative DNA Damage & Repair: An Introduction. *Free Radic. Biol. Med.* 107, 2–12.
- Cao, J., and Yan, Q. (2012). Histone ubiquitination and deubiquitination in transcription, DNA damage response, and cancer. *Front. Oncol.* 2, 26.
- Capotosti, F., Hsieh, J.J.-D., and Herr, W. (2007). Species selectivity of mixed-lineage leukemia/trithorax and HCF proteolytic maturation pathways. *Mol. Cell. Biol.* 27, 7063–7072.
- Capra, J.A., Paeschke, K., Singh, M., and Zakian, V.A. (2010). G-quadruplex DNA sequences are evolutionarily conserved and associated with distinct genomic features in *Saccharomyces cerevisiae*. *PLoS Comput. Biol.* 6, e1000861.
- Chatterjee, N., and Walker, G.C. (2017). Mechanisms of DNA damage, repair, and mutagenesis. *Environ. Mol. Mutagen.* 58, 235–263.
- Chen, D.Y., Liu, H., Takeda, S., Tu, H.-C., Sasagawa, S., Van Tine, B.A., Lu, D., Cheng, E.H.-Y., and Hsieh, J.J.-D. (2010). Taspase1 functions as a non-oncogene addiction protease that coordinates cancer cell proliferation and apoptosis. *Cancer Res.* 70, 5358–5367.
- Chen, D.Y., Lee, Y., Van Tine, B.A., Searleman, A.C., Westergard, T.D., Liu, H., Tu, H.-C., Takeda, S., Dong, Y., Piwnica-Worms, D.R., et al. (2012a). A pharmacologic inhibitor of the protease Taspase1 effectively inhibits breast and brain tumor growth. *Cancer Res.* 72, 736–746.
- Chen, H.P., Zhao, Y.T., and Zhao, T.C. (2015). Histone deacetylases and mechanisms of regulation of gene expression. *Crit. Rev. Oncog.* 20, 35–47.
- Chen, S., Ma, J., Wu, F., Xiong, L.-J., Ma, H., Xu, W., Lv, R., Li, X., Villen, J., Gygi, S.P., et al. (2012b). The histone H3 Lys 27 demethylase JMJD3 regulates gene expression by impacting transcriptional elongation. *Genes Dev.* 26, 1364–1375.
- Choi, J.-Y., Lim, S., Kim, E.-J., Jo, A., and Guengerich, F.P. (2010). Translesion synthesis across abasic lesions by human B-family and Y-family DNA polymerases α , δ , η , ι , κ , and REV1. *J. Mol. Biol.* 404, 34–44.
- Coloma, J., Johnson, R.E., Prakash, L., Prakash, S., and Aggarwal, A.K. (2016). Human DNA

- polymerase α in binary complex with a DNA:DNA template-primer. *Sci. Rep.* *6*, 23784.
- Cooper, D.N., Bacolla, A., Férec, C., Vasquez, K.M., Kehrer-Sawatzki, H., and Chen, J.-M. (2011). On the sequence-directed nature of human gene mutation: the role of genomic architecture and the local DNA sequence environment in mediating gene mutations underlying human inherited disease. *Hum. Mutat.* *32*, 1075–1099.
- Crossley, M.P., Bocek, M., and Cimprich, K.A. (2019). R-Loops as Cellular Regulators and Genomic Threats. *Mol. Cell* *73*, 398–411.
- Deans, C., and Maggert, K.A. (2015). What Do You Mean, “Epigenetic”? *Genetics* *199*, 887–896.
- Déjardin, J. (2015). Switching between Epigenetic States at Pericentromeric Heterochromatin. *Trends Genet. TIG* *31*, 661–672.
- Delbos, F., Aoufouchi, S., Faili, A., Weill, J.-C., and Reynaud, C.-A. (2007). DNA polymerase eta is the sole contributor of A/T modifications during immunoglobulin gene hypermutation in the mouse. *J. Exp. Med.* *204*, 17–23.
- Dewar, J.M., and Walter, J.C. (2017). Mechanisms of DNA replication termination. *Nat. Rev. Mol. Cell Biol.* *18*, 507–516.
- Dhalluin, C., Carlson, J.E., Zeng, L., He, C., Aggarwal, A.K., and Zhou, M.M. (1999). Structure and ligand of a histone acetyltransferase bromodomain. *Nature* *399*, 491–496.
- Dillon, S.C., Zhang, X., Trievel, R.C., and Cheng, X. (2005). The SET-domain protein superfamily: protein lysine methyltransferases. *Genome Biol.* *6*, 227.
- Doles, J., Oliver, T.G., Cameron, E.R., Hsu, G., Jacks, T., Walker, G.C., and Hemann, M.T. (2010). Suppression of Rev3, the catalytic subunit of Pol{zeta}, sensitizes drug-resistant lung tumors to chemotherapy. *Proc. Natl. Acad. Sci. U. S. A.* *107*, 20786–20791.
- Dong, Y., Van Tine, B.A., Oyama, T., Wang, P.I., Cheng, E.H., and Hsieh, J.J. (2014). Taspase1 cleaves MLL1 to activate cyclin E for HER2/neu breast tumorigenesis. *Cell Res.* *24*, 1354–1366.
- Donley, N., and Thayer, M.J. (2013). DNA replication timing, genome stability and cancer: late and/or delayed DNA replication timing is associated with increased genomic instability. *Semin. Cancer Biol.* *23*, 80–89.
- Douglas, M.E., Ali, F.A., Costa, A., and Diffley, J.F.X. (2018). The mechanism of eukaryotic CMG helicase activation. *Nature* *555*, 265–268.
- Du, Q., Bert, S.A., Armstrong, N.J., Caldon, C.E., Song, J.Z., Nair, S.S., Gould, C.M., Luu, P.-L., Peters, T., Khoury, A., et al. (2019). Replication timing and epigenome remodelling are associated with the nature of chromosomal rearrangements in cancer. *Nat. Commun.* *10*, 416.

- Duvic, M., and Vu, J. (2007). Vorinostat: a new oral histone deacetylase inhibitor approved for cutaneous T-cell lymphoma. *Expert Opin. Investig. Drugs* *16*, 1111–1120.
- Ehrmann, M., and Clausen, T. (2004). Proteolysis as a regulatory mechanism. *Annu. Rev. Genet.* *38*, 709–724.
- Ellenbroek, B., and Youn, J. (2016). Chapter 5 - Environment Challenges and the Brain. In *Gene-Environment Interactions in Psychiatry*, B. Ellenbroek, and J. Youn, eds. (San Diego: Academic Press), pp. 107–139.
- Ernst, J., Kheradpour, P., Mikkelson, T.S., Shores, N., Ward, L.D., Epstein, C.B., Zhang, X., Wang, L., Issner, R., Coyne, M., et al. (2011). Mapping and analysis of chromatin state dynamics in nine human cell types. *Nature* *473*, 43–49.
- Esposito, G., Godindagger, I., Klein, U., Yaspo, M.L., Cumano, A., and Rajewsky, K. (2000). Disruption of the Rev31-encoded catalytic subunit of polymerase zeta in mice results in early embryonic lethality. *Curr. Biol. CB* *10*, 1221–1224.
- Esteban-Jurado, C., Franch-Expósito, S., Muñoz, J., Ocaña, T., Carballal, S., López-Cerón, M., Cuatrecasas, M., Vila-Casadesús, M., Lozano, J.J., Serra, E., et al. (2016). The Fanconi anemia DNA damage repair pathway in the spotlight for germline predisposition to colorectal cancer. *Eur. J. Hum. Genet. EJHG* *24*, 1501–1505.
- Filion, G.J., van Bommel, J.G., Braunschweig, U., Talhout, W., Kind, J., Ward, L.D., Brugman, W., de Castro, I.J., Kerkhoven, R.M., Bussemaker, H.J., et al. (2010). Systematic protein location mapping reveals five principal chromatin types in *Drosophila* cells. *Cell* *143*, 212–224.
- Flott, S., Alabert, C., Toh, G.W., Toth, R., Sugawara, N., Campbell, D.G., Haber, J.E., Pasero, P., and Rouse, J. (2007). Phosphorylation of Slx4 by Mec1 and Tel1 Regulates the Single-Strand Annealing Mode of DNA Repair in Budding Yeast. *Mol. Cell. Biol.* *27*, 6433–6445.
- Forsburg, S.L., and Shen, K.-F. (2017). Centromere Stability: The Replication Connection. *Genes* *8*.
- Fortune, J.M., Pavlov, Y.I., Welch, C.M., Johansson, E., Burgers, P.M.J., and Kunkel, T.A. (2005). *Saccharomyces cerevisiae* DNA polymerase delta: high fidelity for base substitutions but lower fidelity for single- and multi-base deletions. *J. Biol. Chem.* *280*, 29980–29987.
- Fragkos, M., Ganier, O., Coulombe, P., and Méchali, M. (2015). DNA replication origin activation in space and time. *Nat. Rev. Mol. Cell Biol.* *16*, 360–374.
- Freeman, A.K., and Monteiro, A.N. (2010). Phosphatases in the cellular response to DNA damage. *Cell Commun. Signal. CCS* *8*, 27.

- Gaillard, H., García-Muse, T., and Aguilera, A. (2015). Replication stress and cancer. *Nat. Rev. Cancer* *15*, 276–289.
- Gallego-Sánchez, A., Andrés, S., Conde, F., San-Segundo, P.A., and Bueno, A. (2012). Reversal of PCNA ubiquitylation by Ubp10 in *Saccharomyces cerevisiae*. *PLoS Genet.* *8*, e1002826.
- Gan, G.N., Wittschieben, J.P., Wittschieben, B.Ø., and Wood, R.D. (2008). DNA polymerase zeta (pol zeta) in higher eukaryotes. *Cell Res.* *18*, 174–183.
- Garaycochea, J.I., Crossan, G.P., Langevin, F., Mulderrig, L., Louzada, S., Yang, F., Guilbaud, G., Park, N., Roerink, S., Nik-Zainal, S., et al. (2018). Alcohol and endogenous aldehydes damage chromosomes and mutate stem cells. *Nature* *553*, 171–177.
- García-Cao, M., O’Sullivan, R., Peters, A.H.F.M., Jenuwein, T., and Blasco, M.A. (2004). Epigenetic regulation of telomere length in mammalian cells by the Suv39h1 and Suv39h2 histone methyltransferases. *Nat. Genet.* *36*, 94–99.
- García-Díaz, M., and Bebenek, K. (2007). Multiple functions of DNA polymerases. *Crit. Rev. Plant Sci.* *26*, 105–122.
- García-Muse, T., and Aguilera, A. (2016). Transcription-replication conflicts: how they occur and how they are resolved. *Nat. Rev. Mol. Cell Biol.* *17*, 553–563.
- Gerstein, M.B., Lu, Z.J., Van Nostrand, E.L., Cheng, C., Arshinoff, B.I., Liu, T., Yip, K.Y., Robilotto, R., Rechtsteiner, A., Ikegami, K., et al. (2010). Integrative analysis of the *Caenorhabditis elegans* genome by the modENCODE project. *Science* *330*, 1775–1787.
- Gibbs, P.E., McGregor, W.G., Maher, V.M., Nisson, P., and Lawrence, C.W. (1998). A human homolog of the *Saccharomyces cerevisiae* REV3 gene, which encodes the catalytic subunit of DNA polymerase zeta. *Proc. Natl. Acad. Sci. U. S. A.* *95*, 6876–6880.
- Gilbert, D.M. (2010). Evaluating genome-scale approaches to eukaryotic DNA replication. *Nat. Rev. Genet.* *11*, 673–684.
- Glover, T.W., Wilson, T.E., and Arlt, M.F. (2017). Fragile sites in cancer: more than meets the eye. *Nat. Rev. Cancer* *17*, 489–501.
- Gómez-Llorente, Y., Malik, R., Jain, R., Choudhury, J.R., Johnson, R.E., Prakash, L., Prakash, S., Ubarretxena-Belandia, I., and Aggarwal, A.K. (2013). The architecture of yeast DNA polymerase ζ . *Cell Rep.* *5*, 79–86.
- Goričar, K., Kovač, V., and Dolžan, V. (2014). Polymorphisms in translesion polymerase genes influence treatment outcome in malignant mesothelioma. *Pharmacogenomics* *15*, 941–950.
- Goswami, P., Abid Ali, F., Douglas, M.E., Locke, J., Purkiss, A., Janska, A., Eickhoff, P.,

- Early, A., Nans, A., Cheung, A.M.C., et al. (2018). Structure of DNA-CMG-Pol epsilon elucidates the roles of the non-catalytic polymerase modules in the eukaryotic replisome. *Nat. Commun.* *9*, 5061.
- Guan, Z., Hughes, C.M., Kosiyatrakul, S., Norio, P., Sen, R., Fiering, S., Allis, C.D., Bouhassira, E.E., and Schildkraut, C.L. (2009). Decreased replication origin activity in temporal transition regions. *J. Cell Biol.* *187*, 623–635.
- Guenatri, M., Bailly, D., Maison, C., and Almouzni, G. (2004). Mouse centric and pericentric satellite repeats form distinct functional heterochromatin. *J. Cell Biol.* *166*, 493–505.
- Gueranger, Q., Stary, A., Aoufouchi, S., Faili, A., Sarasin, A., Reynaud, C.-A., and Weill, J.-C. (2008). Role of DNA polymerases eta, iota and zeta in UV resistance and UV-induced mutagenesis in a human cell line. *DNA Repair* *7*, 1551–1562.
- Guilbaud, G., Rappailles, A., Baker, A., Chen, C.-L., Arneodo, A., Goldar, A., d'Aubenton-Carafa, Y., Thermes, C., Audit, B., and Hyrien, O. (2011). Evidence for sequential and increasing activation of replication origins along replication timing gradients in the human genome. *PLoS Comput. Biol.* *7*, e1002322.
- Guo, D., Wu, X., Rajpal, D.K., Taylor, J.S., and Wang, Z. (2001). Translesion synthesis by yeast DNA polymerase zeta from templates containing lesions of ultraviolet radiation and acetylaminofluorene. *Nucleic Acids Res.* *29*, 2875–2883.
- Hall, I.M., Shankaranarayana, G.D., Noma, K.-I., Ayoub, N., Cohen, A., and Grewal, S.I.S. (2002). Establishment and maintenance of a heterochromatin domain. *Science* *297*, 2232–2237.
- Hamperl, S., and Cimprich, K.A. (2014). The contribution of co-transcriptional RNA:DNA hybrid structures to DNA damage and genome instability. *DNA Repair* *19*, 84–94.
- Hanafusa, T., Habu, T., Tomida, J., Ohashi, E., Murakumo, Y., and Ohmori, H. (2010). Overlapping in short motif sequences for binding to human REV7 and MAD2 proteins. *Genes Cells Devoted Mol. Cell. Mech.* *15*, 281–296.
- Hänsel-Hertsch, R., Di Antonio, M., and Balasubramanian, S. (2017). DNA G-quadruplexes in the human genome: detection, functions and therapeutic potential. *Nat. Rev. Mol. Cell Biol.* *18*, 279–284.
- Hara, K., Hashimoto, H., Murakumo, Y., Kobayashi, S., Kogame, T., Unzai, S., Akashi, S., Takeda, S., Shimizu, T., and Sato, M. (2010). Crystal structure of human REV7 in complex with a human REV3 fragment and structural implication of the interaction between DNA polymerase zeta and REV1. *J. Biol. Chem.* *285*, 12299–12307.
- Haracska, L., Unk, I., Johnson, R.E., Johansson, E., Burgers, P.M., Prakash, S., and Prakash,

- L. (2001). Roles of yeast DNA polymerases delta and zeta and of Rev1 in the bypass of abasic sites. *Genes Dev.* *15*, 945–954.
- Haracska, L., Prakash, S., and Prakash, L. (2003). Yeast DNA polymerase zeta is an efficient extender of primer ends opposite from 7,8-dihydro-8-Oxoguanine and O6-methylguanine. *Mol. Cell. Biol.* *23*, 1453–1459.
- Harfe, B.D., and Jinks-Robertson, S. (2000). DNA polymerase zeta introduces multiple mutations when bypassing spontaneous DNA damage in *Saccharomyces cerevisiae*. *Mol. Cell* *6*, 1491–1499.
- Hayashi, A., Horiuchi, A., Kikuchi, N., Hayashi, T., Fuseya, C., Suzuki, A., Konishi, I., and Shiozawa, T. (2010). Type-specific roles of histone deacetylase (HDAC) overexpression in ovarian carcinoma: HDAC1 enhances cell proliferation and HDAC3 stimulates cell migration with downregulation of E-cadherin. *Int. J. Cancer* *127*, 1332–1346.
- Heintzman, N.D., Stuart, R.K., Hon, G., Fu, Y., Ching, C.W., Hawkins, R.D., Barrera, L.O., Van Calcar, S., Qu, C., Ching, K.A., et al. (2007). Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nat. Genet.* *39*, 311–318.
- Hiratani, I., Takebayashi, S., Lu, J., and Gilbert, D.M. (2009). Replication timing and transcriptional control: beyond cause and effect--part II. *Curr. Opin. Genet. Dev.* *19*, 142–149.
- Hiratani, I., Ryba, T., Itoh, M., Rathjen, J., Kulik, M., Papp, B., Fussner, E., Bazett-Jones, D.P., Plath, K., Dalton, S., et al. (2010). Genome-wide dynamics of replication timing revealed by in vitro models of mouse embryogenesis. *Genome Res.* *20*, 155–169.
- Hogg, M., Osterman, P., Bylund, G.O., Ganai, R.A., Lundström, E.-B., Sauer-Eriksson, A.E., and Johansson, E. (2014). Structural basis for processive DNA synthesis by yeast DNA polymerase ϵ . *Nat. Struct. Mol. Biol.* *21*, 49–55.
- Holbeck, S.L., and Strathern, J.N. (1997). A role for REV3 in mutagenesis during double-strand break repair in *Saccharomyces cerevisiae*. *Genetics* *147*, 1017–1024.
- Holliday, R., and Pugh, J.E. (1975). DNA modification mechanisms and gene activity during development. *Science* *187*, 226–232.
- Hotchkiss, R.D. (1948). The quantitative separation of purines, pyrimidines, and nucleosides by paper chromatography. *J. Biol. Chem.* *175*, 315–332.
- Hsieh, J.J.-D., Ernst, P., Erdjument-Bromage, H., Tempst, P., and Korsmeyer, S.J. (2003a). Proteolytic cleavage of MLL generates a complex of N- and C-terminal fragments that confers protein stability and subnuclear localization. *Mol. Cell. Biol.* *23*, 186–194.

- Hsieh, J.J.-D., Cheng, E.H.-Y., and Korsmeyer, S.J. (2003b). Taspase1: a threonine aspartase required for cleavage of MLL and proper HOX gene expression. *Cell* *115*, 293–303.
- Hsieh, J.J.-D., Cheng, E.H.-Y., and Korsmeyer, S.J. (2003c). Taspase1: a threonine aspartase required for cleavage of MLL and proper HOX gene expression. *Cell* *115*, 293–303.
- Huang, K.K., Jang, K.W., Kim, S., Kim, H.S., Kim, S.-M., Kwon, H.J., Kim, H.R., Yun, H.J., Ahn, M.J., Park, K.U., et al. (2016). Exome sequencing reveals recurrent REV3L mutations in cisplatin-resistant squamous cell carcinoma of head and neck. *Sci. Rep.* *6*, 19552.
- Hubscher, U., Maga, G., and Spadari, S. (2002). Eukaryotic DNA polymerases. *Annu. Rev. Biochem.* *71*, 133–163.
- Hyun, K., Jeon, J., Park, K., and Kim, J. (2017). Writing, erasing and reading histone lysine methylations. *Exp. Mol. Med.* *49*, e324.
- Jain, R., Aggarwal, A.K., and Rechkoblit, O. (2018). Eukaryotic DNA polymerases. *Curr. Opin. Struct. Biol.* *53*, 77–87.
- James, T.C., and Elgin, S.C. (1986). Identification of a nonhistone chromosomal protein associated with heterochromatin in *Drosophila melanogaster* and its gene. *Mol. Cell. Biol.* *6*, 3862–3872.
- Janssen, A., Colmenares, S.U., and Karpen, G.H. (2018). Heterochromatin: Guardian of the Genome. *Annu. Rev. Cell Dev. Biol.* *34*, 265–288.
- Jeggo, P.A., Pearl, L.H., and Carr, A.M. (2016). DNA repair, genome stability and cancer: a historical perspective. *Nat. Rev. Cancer* *16*, 35–42.
- Jenuwein, T. (2006). The epigenetic magic of histone lysine methylation. *FEBS J.* *273*, 3121–3135.
- Jia, P., Jin, H., Meador, C.B., Xia, J., Ohashi, K., Liu, L., Pirazzoli, V., Dahlman, K.B., Politi, K., Michor, F., et al. (2013). Next-generation sequencing of paired tyrosine kinase inhibitor-sensitive and -resistant EGFR mutant lung cancer cell lines identifies spectrum of DNA changes associated with drug resistance. *Genome Res.* *23*, 1434–1445.
- Johnson, R.E., Washington, M.T., Haracska, L., Prakash, S., and Prakash, L. (2000). Eukaryotic polymerases iota and zeta act sequentially to bypass DNA lesions. *Nature* *406*, 1015–1019.
- Johnson, R.E., Yu, S.-L., Prakash, S., and Prakash, L. (2003). Yeast DNA polymerase zeta (zeta) is essential for error-free replication past thymine glycol. *Genes Dev.* *17*, 77–87.
- Johnson, R.E., Prakash, L., and Prakash, S. (2012). Pol31 and Pol32 subunits of yeast DNA polymerase δ are also essential subunits of DNA polymerase ζ . *Proc. Natl. Acad. Sci. U. S. A.* *109*, 12455–12460.

- Johnson, R.E., Klassen, R., Prakash, L., and Prakash, S. (2015). A Major Role of DNA Polymerase δ in Replication of Both the Leading and Lagging DNA Strands. *Mol. Cell* 59, 163–175.
- Kannouche, P., and Lehmann, A. (2006). Localization of Y-family polymerases and the DNA polymerase switch in mammalian cells. *Methods Enzymol.* 408, 407–415.
- Kannouche, P.L., and Lehmann, A.R. (2004). Ubiquitination of PCNA and the polymerase switch in human cells. *Cell Cycle Georget. Tex* 3, 1011–1013.
- Kannouche, P., Broughton, B.C., Volker, M., Hanaoka, F., Mullenders, L.H., and Lehmann, A.R. (2001). Domain structure, localization, and function of DNA polymerase eta, defective in xeroderma pigmentosum variant cells. *Genes Dev.* 15, 158–172.
- Kawamura, K., O-Wang, J., Bahar, R., Koshikawa, N., Shishikura, T., Nakagawara, A., Sakiyama, S., Kajiwara, K., Kimura, M., and Tagawa, M. (2001). The error-prone DNA polymerase zeta catalytic subunit (Rev3) gene is ubiquitously expressed in normal and malignant human tissues. *Int. J. Oncol.* 18, 97–103.
- Kelly, T., and Callegari, A.J. (2019). Dynamics of DNA replication in a eukaryotic cell. *Proc. Natl. Acad. Sci. U. S. A.* 116, 4973–4982.
- Khan, J.A., Dunn, B.M., and Tong, L. (2005). Crystal structure of human Taspase1, a crucial protease regulating the function of MLL. *Struct. Lond. Engl.* 1993 13, 1443–1452.
- Kharchenko, P.V., Alekseyenko, A.A., Schwartz, Y.B., Minoda, A., Riddle, N.C., Ernst, J., Sabo, P.J., Larschan, E., Gorchakov, A.A., Gu, T., et al. (2011). Comprehensive analysis of the chromatin landscape in *Drosophila melanogaster*. *Nature* 471, 480–485.
- Kikuchi, S., Hara, K., Shimizu, T., Sato, M., and Hashimoto, H. (2012). Structural basis of recruitment of DNA polymerase ζ by interaction between REV1 and REV7 proteins. *J. Biol. Chem.* 287, 33847–33852.
- Kim, T., and Buratowski, S. (2009). Dimethylation of H3K4 by Set1 recruits the Set3 histone deacetylase complex to 5' transcribed regions. *Cell* 137, 259–272.
- Kleff, S., Andrulis, E.D., Anderson, C.W., and Sternglanz, R. (1995). Identification of a gene encoding a yeast histone H4 acetyltransferase. *J. Biol. Chem.* 270, 24674–24677.
- Knobel, P.A., Kotov, I.N., Felley-Bosco, E., Stahel, R.A., and Marti, T.M. (2011). Inhibition of REV3 expression induces persistent DNA damage and growth arrest in cancer cells. *Neoplasia N. Y. N* 13, 961–970.
- Kochenova, O.V., Dae, D.L., Mertz, T.M., and Shcherbakova, P.V. (2015). DNA polymerase ζ -dependent lesion bypass in *Saccharomyces cerevisiae* is accompanied by error-prone copying of long stretches of adjacent DNA. *PLoS Genet.* 11, e1005110.

- Kochenova, O.V., Bezalel-Buch, R., Tran, P., Makarova, A.V., Chabes, A., Burgers, P.M.J., and Shcherbakova, P.V. (2017). Yeast DNA polymerase ζ maintains consistent activity and mutagenicity across a wide range of physiological dNTP concentrations. *Nucleic Acids Res.* *45*, 1200–1218.
- Kong, L., Murata, M.M., and Digman, M.A. (2018). Absence of REV3L promotes p53-regulated cancer cell metabolism in cisplatin-treated lung carcinoma cells. *Biochem. Biophys. Res. Commun.* *496*, 199–204.
- Kucab, J.E., Zou, X., Morganella, S., Joel, M., Nanda, A.S., Nagy, E., Gomez, C., Degasperi, A., Harris, R., Jackson, S.P., et al. (2019). A Compendium of Mutational Signatures of Environmental Agents. *Cell* *177*, 821-836.e16.
- Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., et al. (2001). Initial sequencing and analysis of the human genome. *Nature* *409*, 860–921.
- Lange, S.S., Wittschieben, J.P., and Wood, R.D. (2012). DNA polymerase zeta is required for proliferation of normal mammalian cells. *Nucleic Acids Res.* *40*, 4473–4482.
- Lange, S.S., Bedford, E., Reh, S., Wittschieben, J.P., Carbajal, S., Kusewitt, D.F., DiGiovanni, J., and Wood, R.D. (2013). Dual role for mammalian DNA polymerase ζ in maintaining genome stability and proliferative responses. *Proc. Natl. Acad. Sci. U. S. A.* *110*, E687-696.
- Lange, S.S., Tomida, J., Boulware, K.S., Bhetawal, S., and Wood, R.D. (2016). The Polymerase Activity of Mammalian DNA Pol ζ Is Specifically Required for Cell and Embryonic Viability. *PLoS Genet.* *12*.
- Lange, S.S., Bhetawal, S., Reh, S., Powell, K.L., Kusewitt, D.F., and Wood, R.D. (2018). DNA polymerase ζ deficiency causes impaired wound healing and stress-induced skin pigmentation. *Life Sci. Alliance* *1*.
- de Lange, T. (2005). Shelterin: the protein complex that shapes and safeguards human telomeres. *Genes Dev.* *19*, 2100–2110.
- Larson, A.G., Elnatan, D., Keenen, M.M., Trnka, M.J., Johnston, J.B., Burlingame, A.L., Agard, D.A., Redding, S., and Narlikar, G.J. (2017). Liquid droplet formation by HP1 α suggests a role for phase separation in heterochromatin. *Nature* *547*, 236–240.
- Lawrence, C.W. (2002). Cellular roles of DNA polymerase zeta and Rev1 protein. *DNA Repair* *1*, 425–435.
- Lawrence, C.W., and Christensen, R.B. (1979). Ultraviolet-induced reversion of *cyc1* alleles in radiation-sensitive strains of yeast. III. *rev3* mutant strains. *Genetics* *92*, 397–408.

- Lazzaro, F., Novarina, D., Amara, F., Watt, D.L., Stone, J.E., Costanzo, V., Burgers, P.M., Kunkel, T.A., Plevani, P., and Muzi-Falconi, M. (2012). RNase H and postreplication repair protect cells from ribonucleotides incorporated in DNA. *Mol. Cell* *45*, 99–110.
- Lee, C.Y., and Grant, P.A. (2019). Chapter 1-1 - Role of Histone Acetylation and Acetyltransferases in Gene Regulation. In *Toxicogenetics*, S.D. McCullough, and D.C. Dolinoy, eds. (Academic Press), pp. 3–30.
- Lee, Y.-S., Gregory, M.T., and Yang, W. (2014). Human Pol ζ purified with accessory subunits is active in translesion DNA synthesis and complements Pol η in cisplatin bypass. *Proc. Natl. Acad. Sci. U. S. A.* *111*, 2954–2959.
- Lehmann, A.R., Niimi, A., Ogi, T., Brown, S., Sabbioneda, S., Wing, J.F., Kannouche, P.L., and Green, C.M. (2007). Translesion synthesis: Y-family polymerases and the polymerase switch. *DNA Repair* *6*, 891–899.
- Lemontt, J.F. (1971a). Mutants of yeast defective in mutation induced by ultraviolet light. *Genetics* *68*, 21–33.
- Lemontt, J.F. (1971b). Pathways of ultraviolet mutability in *Saccharomyces cerevisiae*. I. Some properties of double mutants involving *uvs9* and *rev*. *Mutat. Res.* *13*, 311–317.
- Lemontt, J.F. (1971c). Pathways of ultraviolet mutability in *Saccharomyces cerevisiae*. II. The effect of *rev* genes on recombination. *Mutat. Res.* *13*, 319–326.
- Lemontt, J.F. (1972). Induction of forward mutations in mutationally defective yeast. *Mol. Gen. Genet. MGG* *119*, 27–42.
- Li, C., Guo, S., Zhang, M., Gao, J., and Guo, Y. (2015). DNA methylation and histone modification patterns during the late embryonic and early postnatal development of chickens. *Poult. Sci.* *94*, 706–721.
- Lin, Q., Clark, A.B., McCulloch, S.D., Yuan, T., Bronson, R.T., Kunkel, T.A., and Kucherlapati, R. (2006). Increased susceptibility to UV-induced skin carcinogenesis in polymerase η -deficient mice. *Cancer Res.* *66*, 87–94.
- Lin, Y.-C., Owen, N., Minko, I.G., Lange, S.S., Tomida, J., Li, L., Stone, M.P., Wood, R.D., McCullough, A.K., and Lloyd, R.S. (2016). DNA polymerase ζ limits chromosomal damage and promotes cell survival following aflatoxin exposure. *Proc. Natl. Acad. Sci. U. S. A.* *113*, 13774–13779.
- Liu, L., De, S., and Michor, F. (2013). DNA replication timing and higher-order nuclear organization determine single-nucleotide substitution patterns in cancer genomes. *Nat. Commun.* *4*, 1502.
- Liu, T., Rechtsteiner, A., Egelhofer, T.A., Vielle, A., Latorre, I., Cheung, M.-S., Ercan, S.,

- Ikegami, K., Jensen, M., Kolasinska-Zwierz, P., et al. (2011). Broad chromosomal domains of histone modification patterns in *C. elegans*. *Genome Res.* *21*, 227–236.
- Lopes, M., Foiani, M., and Sogo, J.M. (2006). Multiple mechanisms control chromosome integrity after replication fork uncoupling and restart at irreparable UV lesions. *Mol. Cell* *21*, 15–27.
- Lu, S., Wang, G., Bacolla, A., Zhao, J., Spitzer, S., and Vasquez, K.M. (2015). Short Inverted Repeats Are Hotspots for Genetic Instability: Relevance to Cancer Genomes. *Cell Rep.* *10*, 1674–1680.
- MacAlpine, D.M., Rodríguez, H.K., and Bell, S.P. (2004). Coordination of replication and transcription along a *Drosophila* chromosome. *Genes Dev.* *18*, 3094–3105.
- Macheret, M., and Halazonetis, T.D. (2015). DNA replication stress as a hallmark of cancer. *Annu. Rev. Pathol.* *10*, 425–448.
- Maestroni, L., Matmati, S., and Coulon, S. (2017). Solving the Telomere Replication Problem. *Genes* *8*.
- Magdalou, I., Lopez, B.S., Pasero, P., and Lambert, S.A.E. (2014). The causes of replication stress and their consequences on genome stability and cell fate. *Semin. Cell Dev. Biol.* *30*, 154–164.
- Makarova, A.V., and Burgers, P.M. (2015). Eukaryotic DNA polymerase ζ . *DNA Repair* *29*, 47–55.
- Makarova, A.V., Stodola, J.L., and Burgers, P.M. (2012). A four-subunit DNA polymerase ζ complex containing Pol δ accessory subunits is essential for PCNA-mediated mutagenesis. *Nucleic Acids Res.* *40*, 11618–11626.
- Marchal, C., Sasaki, T., Vera, D., Wilson, K., Sima, J., Rivera-Mulia, J.C., Trevilla-García, C., Nogues, C., Nafie, E., and Gilbert, D.M. (2018). Genome-wide analysis of replication timing by next-generation sequencing with E/L Repli-seq. *Nat. Protoc.* *13*, 819–839.
- Marmorstein, R., and Zhou, M.-M. (2014). Writers and readers of histone acetylation: structure, mechanism, and inhibition. *Cold Spring Harb. Perspect. Biol.* *6*, a018762.
- Martienssen, R., and Moazed, D. (2015). RNAi and heterochromatin assembly. *Cold Spring Harb. Perspect. Biol.* *7*, a019323.
- Martin, S.K., and Wood, R.D. (2019). DNA polymerase ζ in DNA replication and repair. *Nucleic Acids Res.* *47*, 8348–8361.
- Mayer, R., Brero, A., von Hase, J., Schroeder, T., Cremer, T., and Dietzel, S. (2005). Common themes and cell type specific variations of higher order chromatin arrangements in the mouse. *BMC Cell Biol.* *6*, 44.

- McCulloch, S.D., and Kunkel, T.A. (2008). The fidelity of DNA synthesis by eukaryotic replicative and translesion synthesis polymerases. *Cell Res.* *18*, 148–161.
- McGowan, K.A., Li, J.Z., Park, C.Y., Beaudry, V., Tabor, H.K., Sabnis, A.J., Zhang, W., Fuchs, H., de Angelis, M.H., Myers, R.M., et al. (2008). Ribosomal mutations cause p53-mediated dark skin and pleiotropic effects. *Nat. Genet.* *40*, 963–970.
- McKee, R.H., and Lawrence, C.W. (1979). Genetic analysis of gamma-ray mutagenesis in yeast. I. Reversion in radiation-sensitive strains. *Genetics* *93*, 361–373.
- McVey, M., Khodaverdian, V.Y., Meyer, D., Cerqueira, P.G., and Heyer, W.-D. (2016). Eukaryotic DNA Polymerases in Homologous Recombination. *Annu. Rev. Genet.* *50*, 393–421.
- Mendez-Bermudez, A., Lototska, L., Bauwens, S., Giraud-Panis, M.-J., Croce, O., Jamet, K., Irizar, A., Mowinckel, M., Koundrioukoff, S., Nottet, N., et al. (2018). Genome-wide Control of Heterochromatin Replication by the Telomere Capping Protein TRF2. *Mol. Cell* *70*, 449-461.e5.
- Mikkelsen, T.S., Ku, M., Jaffe, D.B., Issac, B., Lieberman, E., Giannoukos, G., Alvarez, P., Brockman, W., Kim, T.-K., Koche, R.P., et al. (2007). Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* *448*, 553–560.
- Minesinger, B.K., and Jinks-Robertson, S. (2005). Roles of RAD6 epistasis group members in spontaneous polzeta-dependent translesion synthesis in *Saccharomyces cerevisiae*. *Genetics* *169*, 1939–1955.
- Miotto, B., Ji, Z., and Struhl, K. (2016). Selectivity of ORC binding sites and the relation to replication timing, fragile sites, and deletions in cancers. *Proc. Natl. Acad. Sci. U. S. A.* *113*, E4810-4819.
- modENCODE Consortium, Roy, S., Ernst, J., Kharchenko, P.V., Kheradpour, P., Negre, N., Eaton, M.L., Landolin, J.M., Bristow, C.A., Ma, L., et al. (2010). Identification of functional elements and regulatory circuits by *Drosophila* modENCODE. *Science* *330*, 1787–1797.
- Moore, A., Dominska, M., Greenwell, P., Aksenova, A.Y., Mirkin, S., and Petes, T. (2018). Genetic Control of Genomic Alterations Induced in Yeast by Interstitial Telomeric Sequences. *Genetics* *209*, 425–438.
- Moore, L.D., Le, T., and Fan, G. (2013). DNA methylation and its basic function. *Neuropsychopharmacol. Off. Publ. Am. Coll. Neuropsychopharmacol.* *38*, 23–38.
- Morales, F., and Giordano, A. (2016). Overview of CDK9 as a target in cancer research. *Cell Cycle Georget. Tex* *15*, 519–527.
- Morelli, C., Mungall, A.J., Negrini, M., Barbanti-Brodano, G., and Croce, C.M. (1998).

Alternative splicing, genomic structure, and fine chromosome localization of REV3L. *Cytogenet. Cell Genet.* 83, 18–20.

Morrison, A., Christensen, R.B., Alley, J., Beck, A.K., Bernstine, E.G., Lemontt, J.F., and Lawrence, C.W. (1989). REV3, a *Saccharomyces cerevisiae* gene whose function is required for induced mutagenesis, is predicted to encode a nonessential DNA polymerase. *J. Bacteriol.* 171, 5659–5667.

Müller, B.M., Jana, L., Kasajima, A., Lehmann, A., Prinzler, J., Budczies, J., Winzer, K.-J., Dietel, M., Weichert, W., and Denkert, C. (2013). Differential expression of histone deacetylases HDAC1, 2 and 3 in human breast cancer--overexpression of HDAC2 and HDAC3 is associated with clinicopathological indicators of disease progression. *BMC Cancer* 13, 215.

Murakumo, Y., Ogura, Y., Ishii, H., Numata, S., Ichihara, M., Croce, C.M., Fishel, R., and Takahashi, M. (2001). Interactions in the error-prone postreplication repair proteins hREV1, hREV3, and hREV7. *J. Biol. Chem.* 276, 35644–35651.

Nechemia-Arbely, Y., Miga, K.H., Shoshani, O., Aslanian, A., McMahon, M.A., Lee, A.Y., Fachinetti, D., Yates, J.R., Ren, B., and Cleveland, D.W. (2019). DNA replication acts as an error correction mechanism to maintain centromere identity by restricting CENP-A to centromeres. *Nat. Cell Biol.* 21, 743–754.

Nelson, J.R., Lawrence, C.W., and Hinkle, D.C. (1996). Thymine-thymine dimer bypass by yeast DNA polymerase zeta. *Science* 272, 1646–1649.

Ng, H.H., Robert, F., Young, R.A., and Struhl, K. (2003). Targeted recruitment of Set1 histone methylase by elongating Pol II provides a localized mark and memory of recent transcriptional activity. *Mol. Cell* 11, 709–719.

Nguyen, A.T., and Zhang, Y. (2011). The diverse functions of Dot1 and H3K79 methylation. *Genes Dev.* 25, 1345–1358.

Niizuma, H., Cheng, E.H., and Hsieh, J.J. (2015). Taspase 1: A protease with many biological surprises. *Mol. Cell. Oncol.* 2, e999513.

Northam, M.R., Garg, P., Baitin, D.M., Burgers, P.M.J., and Shcherbakova, P.V. (2006). A novel function of DNA polymerase zeta regulated by PCNA. *EMBO J.* 25, 4316–4325.

Northam, M.R., Robinson, H.A., Kochenova, O.V., and Shcherbakova, P.V. (2010). Participation of DNA polymerase zeta in replication of undamaged DNA in *Saccharomyces cerevisiae*. *Genetics* 184, 27–42.

Northam, M.R., Moore, E.A., Mertz, T.M., Binz, S.K., Stith, C.M., Stepchenkova, E.I., Wendt, K.L., Burgers, P.M.J., and Shcherbakova, P.V. (2014). DNA polymerases ζ and Rev1

- mediate error-prone bypass of non-B DNA structures. *Nucleic Acids Res.* 42, 290–306.
- Okada, Y., Feng, Q., Lin, Y., Jiang, Q., Li, Y., Coffield, V.M., Su, L., Xu, G., and Zhang, Y. (2005). hDOT1L links histone methylation to leukemogenesis. *Cell* 121, 167–178.
- Ourliac-Garnier, I., and Londoño-Vallejo, A. (2017). Telomere Length Analysis by Quantitative Fluorescent in Situ Hybridization (Q-FISH). *Methods Mol. Biol. Clifton NJ* 1587, 29–39.
- Oyama, T., Sasagawa, S., Takeda, S., Hess, R.A., Lieberman, P.M., Cheng, E.H., and Hsieh, J.J. (2013). Cleavage of TFIIA by Taspase1 activates TRF2-specified mammalian male germ cell programs. *Dev. Cell* 27, 188–200.
- Parker, M.W., Botchan, M.R., and Berger, J.M. (2017). Mechanisms and regulation of DNA replication initiation in eukaryotes. *Crit. Rev. Biochem. Mol. Biol.* 52, 107–144.
- Passarge, E. (1979). Emil Heitz and the concept of heterochromatin: longitudinal chromosome differentiation was recognized fifty years ago. *Am. J. Hum. Genet.* 31, 106–115.
- Patel, D.J. (2016). A Structural Perspective on Readout of Epigenetic Histone and DNA Methylation Marks. *Cold Spring Harb. Perspect. Biol.* 8, a018754.
- Patel, A., Vought, V.E., Dharmarajan, V., and Cosgrove, M.S. (2008). A conserved arginine-containing motif crucial for the assembly and enzymatic activity of the mixed lineage leukemia protein-1 core complex. *J. Biol. Chem.* 283, 32162–32175.
- Patel, A., Dharmarajan, V., Vought, V.E., and Cosgrove, M.S. (2009). On the mechanism of multiple lysine methylation by the human mixed lineage leukemia protein-1 (MLL1) core complex. *J. Biol. Chem.* 284, 24242–24256.
- Patrushev, L.I., and Minkevich, I.G. (2008). The problem of the eukaryotic genome size. *Biochem. Biokhimiia* 73, 1519–1552.
- Pavlov, Y.I., Shcherbakova, P.V., and Kunkel, T.A. (2001). In vivo consequences of putative active site mutations in yeast DNA polymerases alpha, epsilon, delta, and zeta. *Genetics* 159, 47–64.
- Petta, T.B., Nakajima, S., Zlatanou, A., Despras, E., Couve-Privat, S., Ishchenko, A., Sarasin, A., Yasui, A., and Kannouche, P. (2008). Human DNA polymerase iota protects cells against oxidative stress. *EMBO J.* 27, 2883–2895.
- Pfeiffer, V., and Lingner, J. (2013). Replication of telomeres and the regulation of telomerase. *Cold Spring Harb. Perspect. Biol.* 5, a010405.
- Poon, S.S.S., and Lansdorp, P.M. (2001). Quantitative fluorescence in situ hybridization (Q-FISH). *Curr. Protoc. Cell Biol. Chapter 18*, 18.4.1-18.4.21.
- Poyet, C., Jentsch, B., Hermanns, T., Schweckendiek, D., Seifert, H.-H., Schmidpeter, M.,

- Sulser, T., Moch, H., Wild, P.J., and Kristiansen, G. (2014). Expression of histone deacetylases 1, 2 and 3 in urothelial bladder cancer. *BMC Clin. Pathol.* *14*, 10.
- Pustovalova, Y., Magalhães, M.T.Q., D'Souza, S., Rizzo, A.A., Korza, G., Walker, G.C., and Korzhnev, D.M. (2016). Interaction between the Rev1 C-Terminal Domain and the PolD3 Subunit of Pol ζ Suggests a Mechanism of Polymerase Exchange upon Rev1/Pol ζ -Dependent Translesion Synthesis. *Biochemistry* *55*, 2043–2053.
- Quah, S.K., von Borstel, R.C., and Hastings, P.J. (1980). The origin of spontaneous mutation in *Saccharomyces cerevisiae*. *Genetics* *96*, 819–839.
- Ran, F.A., Hsu, P.D., Wright, J., Agarwala, V., Scott, D.A., and Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. *Nat. Protoc.* *8*, 2281–2308.
- Räschle, M., Knipscheer, P., Knipscheer, P., Enoiu, M., Angelov, T., Sun, J., Griffith, J.D., Ellenberger, T.E., Schäfer, O.D., and Walter, J.C. (2008). Mechanism of replication-coupled DNA interstrand crosslink repair. *Cell* *134*, 969–980.
- Rattray, A.J., Shafer, B.K., McGill, C.B., and Strathern, J.N. (2002). The roles of REV3 and RAD57 in double-strand-break-repair-induced mutagenesis of *Saccharomyces cerevisiae*. *Genetics* *162*, 1063–1077.
- Ravanat, J.-L., and Douki, T. (2016). UV and ionizing radiations induced DNA damage, differences and similarities. *Radiat. Phys. Chem.* *128*, 92–102.
- Rea, S., Eisenhaber, F., O'Carroll, D., Strahl, B.D., Sun, Z.W., Schmid, M., Opravil, S., Mechtler, K., Ponting, C.P., Allis, C.D., et al. (2000). Regulation of chromatin structure by site-specific histone H3 methyltransferases. *Nature* *406*, 593–599.
- Reczek, C.R., and Chandel, N.S. (2017). The Two Faces of Reactive Oxygen Species in Cancer. *Annu. Rev. Cancer Biol.* *1*, 79–98.
- Rhind, N., and Gilbert, D.M. (2013). DNA replication timing. *Cold Spring Harb. Perspect. Biol.* *5*, a010132.
- Roche, H., Gietz, R.D., and Kunz, B.A. (1994). Specificity of the yeast rev3 delta antimutator and REV3 dependency of the mutator resulting from a defect (rad1 delta) in nucleotide excision repair. *Genetics* *137*, 637–646.
- Romeo, F., Falbo, L., and Costanzo, V. (2016). Replication, checkpoint suppression and structure of centromeric DNA. *Nucl. Austin Tex* *7*, 540–546.
- Rose, N.R., and Klose, R.J. (2014). Understanding the relationship between DNA methylation and histone lysine methylation. *Biochim. Biophys. Acta* *1839*, 1362–1372.
- Rossetto, D., Avvakumov, N., and Côté, J. (2012). Histone phosphorylation: a chromatin modification involved in diverse nuclear events. *Epigenetics* *7*, 1098–1108.

- Rudd, S.G., Bianchi, J., and Doherty, A.J. (2014). PrimPol-A new polymerase on the block. *Mol. Cell. Oncol.* *1*, e960754.
- Ryba, T., Hiratani, I., Sasaki, T., Battaglia, D., Kulik, M., Zhang, J., Dalton, S., and Gilbert, D.M. (2011). Replication timing: a fingerprint for cell identity and pluripotency. *PLoS Comput. Biol.* *7*, e1002225.
- Sakamoto, A.N., Stone, J.E., Kissling, G.E., McCulloch, S.D., Pavlov, Y.I., and Kunkel, T.A. (2007). Mutator alleles of yeast DNA polymerase zeta. *DNA Repair* *6*, 1829–1838.
- Sakofsky, C.J., Ayyar, S., Deem, A.K., Chung, W.-H., Ira, G., and Malkova, A. (2015). Translesion Polymerases Drive Microhomology-Mediated Break-Induced Replication Leading to Complex Chromosomal Rearrangements. *Mol. Cell* *60*, 860–872.
- Saksouk, N., Simboeck, E., and Déjardin, J. (2015). Constitutive heterochromatin formation and transcription in mammals. *Epigenetics Chromatin* *8*, 3.
- Saldivar, J.C., Cortez, D., and Cimprich, K.A. (2017). The essential kinase ATR: ensuring faithful duplication of a challenging genome. *Nat. Rev. Mol. Cell Biol.* *18*, 622–636.
- Saldivar, J.C., Hamperl, S., Bocek, M.J., Chung, M., Bass, T.E., Cisneros-Soberanis, F., Samejima, K., Xie, L., Paulson, J.R., Earnshaw, W.C., et al. (2018). An intrinsic S/G2 checkpoint enforced by ATR. *Science* *361*, 806–810.
- Sanchez, R., and Zhou, M.-M. (2009). The role of human bromodomains in chromatin biology and gene transcription. *Curr. Opin. Drug Discov. Devel.* *12*, 659–665.
- Saretzki, G. (2018). Telomeres, Telomerase and Ageing. *Subcell. Biochem.* *90*, 221–308.
- Schmegner, C., Hameister, H., Vogel, W., and Assum, G. (2007). Isochores and replication time zones: a perfect match. *Cytogenet. Genome Res.* *116*, 167–172.
- Schübeler, D., MacAlpine, D.M., Scalzo, D., Wirbelauer, C., Kooperberg, C., van Leeuwen, F., Gottschling, D.E., O'Neill, L.P., Turner, B.M., Delrow, J., et al. (2004). The histone modification pattern of active genes revealed through genome-wide chromatin analysis of a higher eukaryote. *Genes Dev.* *18*, 1263–1271.
- Schultz, J., Copley, R.R., Doerks, T., Ponting, C.P., and Bork, P. (2000). SMART: a web-based tool for the study of genetically mobile domains. *Nucleic Acids Res.* *28*, 231–234.
- Scroggins, B.T., Robzyk, K., Wang, D., Marcu, M.G., Tsutsumi, S., Beebe, K., Cotter, R.J., Felts, S., Toft, D., Karnitz, L., et al. (2007). An acetylation site in the middle domain of Hsp90 regulates chaperone function. *Mol. Cell* *25*, 151–159.
- Segovia, R., Shen, Y., Lujan, S.A., Jones, S.J.M., and Stirling, P.C. (2017). Hypermutation signature reveals a slippage and realignment model of translesion synthesis by Rev3 polymerase in cisplatin-treated yeast. *Proc. Natl. Acad. Sci. U. S. A.* *114*, 2663–2668.

- Shah, K.A., and Mirkin, S.M. (2015). The hidden side of unstable DNA repeats: Mutagenesis at a distance. *DNA Repair* 32, 106–112.
- Shah, K.A., Shishkin, A.A., Voineagu, I., Pavlov, Y.I., Shcherbakova, P.V., and Mirkin, S.M. (2012). Role of DNA polymerases in repeat-mediated genome instability. *Cell Rep.* 2, 1088–1095.
- Sharova, L.V., Sharov, A.A., Nedorezov, T., Piao, Y., Shaik, N., and Ko, M.S.H. (2009). Database for mRNA half-life of 19 977 genes obtained by DNA microarray analysis of pluripotent and differentiating mouse embryonic stem cells. *DNA Res. Int. J. Rapid Publ. Rep. Genes Genomes* 16, 45–58.
- Shay, J.W. (2016). Role of Telomeres and Telomerase in Aging and Cancer. *Cancer Discov.* 6, 584–593.
- Shay, J.W., and Wright, W.E. (2019). Telomeres and telomerase: three decades of progress. *Nat. Rev. Genet.* 20, 299–309.
- Shcherbakova, P.V., Pavlov, Y.I., Chilkova, O., Rogozin, I.B., Johansson, E., and Kunkel, T.A. (2003). Unique error signature of the four-subunit yeast DNA polymerase epsilon. *J. Biol. Chem.* 278, 43770–43780.
- Shi, T.-Y., Yang, L., Yang, G., Tu, X.-Y., Wu, X., Cheng, X., and Wei, Q. (2013). DNA polymerase ζ as a potential biomarker of chemoradiation resistance and poor prognosis for cervical cancer. *Med. Oncol. Northwood Lond. Engl.* 30, 500.
- Shilatifard, A. (2012). The COMPASS family of histone H3K4 methylases: mechanisms of regulation in development and disease pathogenesis. *Annu. Rev. Biochem.* 81, 65–95.
- Shinsky, S.A., Monteith, K.E., Viggiano, S., and Cosgrove, M.S. (2015). Biochemical reconstitution and phylogenetic comparison of human SET1 family core complexes involved in histone methylation. *J. Biol. Chem.* 290, 6361–6375.
- Shiotani, B., and Zou, L. (2016). Signaling of DNA Replication Stress Through the ATR Checkpoint. In *DNA Replication, Recombination, and Repair: Molecular Mechanisms and Pathology*, F. Hanaoka, and K. Sugawara, eds. (Tokyo: Springer Japan), pp. 405–428.
- Shukla, A., Chaurasia, P., and Bhaumik, S.R. (2009). Histone methylation and ubiquitination with their cross-talk and roles in gene expression and stability. *Cell. Mol. Life Sci. CMLS* 66, 1419–1433.
- Sima, J., and Gilbert, D.M. (2014). Complex correlations: replication timing and mutational landscapes during cancer and genome evolution. *Curr. Opin. Genet. Dev.* 25, 93–100.
- Singh, B., Li, X., Owens, K.M., Vanniarajan, A., Liang, P., and Singh, K.K. (2015). Human REV3 DNA Polymerase Zeta Localizes to Mitochondria and Protects the Mitochondrial

Genome. *PloS One* *10*, e0140409.

Skourti-Stathaki, K., and Proudfoot, N.J. (2014). A double-edged sword: R loops as threats to genome integrity and powerful regulators of gene expression. *Genes Dev.* *28*, 1384–1396.

Song, L., McNeil, E.M., Ritchie, A.-M., Astell, K.R., Gourley, C., and Melton, D.W. (2017). Melanoma cells replicate through chemotherapy by reducing levels of key homologous recombination protein RAD51 and increasing expression of translesion synthesis DNA polymerase ζ . *BMC Cancer* *17*, 864.

Srivastava, A.K., Han, C., Zhao, R., Cui, T., Dai, Y., Mao, C., Zhao, W., Zhang, X., Yu, J., and Wang, Q.-E. (2015). Enhanced expression of DNA polymerase eta contributes to cisplatin resistance of ovarian cancer stem cells. *Proc. Natl. Acad. Sci. U. S. A.* *112*, 4411–4416.

Stamatoyannopoulos, J.A., Adzhubei, I., Thurman, R.E., Kryukov, G.V., Mirkin, S.M., and Sunyaev, S.R. (2009). Human mutation rate associated with DNA replication timing. *Nat. Genet.* *41*, 393–395.

Stary, A., Kannouche, P., Lehmann, A.R., and Sarasin, A. (2003). Role of DNA polymerase eta in the UV mutation spectrum in human cells. *J. Biol. Chem.* *278*, 18767–18775.

Stillman, B. (2015). Reconsidering DNA Polymerases at the Replication Fork in Eukaryotes. *Mol. Cell* *59*, 139–141.

Stone, J.E., Kissling, G.E., Lujan, S.A., Rogozin, I.B., Stith, C.M., Burgers, P.M.J., and Kunkel, T.A. (2009). Low-fidelity DNA synthesis by the L979F mutator derivative of *Saccharomyces cerevisiae* DNA polymerase zeta. *Nucleic Acids Res.* *37*, 3774–3787.

Stone, J.E., Kumar, D., Binz, S.K., Inase, A., Iwai, S., Chabes, A., Burgers, P.M., and Kunkel, T.A. (2011). Lesion bypass by *S. cerevisiae* Pol ζ alone. *DNA Repair* *10*, 826–834.

Stone, J.E., Lujan, S.A., Kunkel, T.A., and Kunkel, T.A. (2012). DNA polymerase zeta generates clustered mutations during bypass of endogenous DNA lesions in *Saccharomyces cerevisiae*. *Environ. Mol. Mutagen.* *53*, 777–786.

Strålfors, A., and Ekwall, K. (2011). Heterochromatin and Euchromatin—Organization, Boundaries, and Gene Regulation. In *Reviews in Cell Biology and Molecular Medicine*, (American Cancer Society), p.

Strom, A.R., Emelyanov, A.V., Mir, M., Fyodorov, D.V., Darzacq, X., and Karpen, G.H. (2017). Phase separation drives heterochromatin domain formation. *Nature* *547*, 241–245.

Sui, Y., Li, Y.-J., Jin, C.-X., and Xu, F. (2010). [Roles of REV3 in proliferation and genomic stability of colon carcinoma cells]. *Yi Chuan Hered.* *32*, 467–472.

Suleiman, J., Mundt, M., Sampath, S., and El-Hattab, A.W. (2018). TASP1 is deleted in an infant with developmental delay, microcephaly, distinctive facial features, and multiple

congenital anomalies. *Clin. Genet.* *94*, 170–173.

Suleiman, J., Riedhammer, K.M., Jicinsky, T., Mundt, M., Werner, L., Gusic, M., Burgemeister, A.L., Alsaif, H.S., Abdulrahim, M., Moghrabi, N.N., et al. (2019). Homozygous loss-of-function variants of TASP1, a gene encoding an activator of the histone methyltransferases KMT2A and KMT2D, cause a syndrome of developmental delay, happy demeanor, distinctive facial features, and congenital anomalies. *Hum. Mutat.* *40*, 1985–1992.

Swan, M.K., Johnson, R.E., Prakash, L., Prakash, S., and Aggarwal, A.K. (2009). Structural basis of high-fidelity DNA synthesis by yeast DNA polymerase delta. *Nat. Struct. Mol. Biol.* *16*, 979–986.

Swatek, K.N., and Komander, D. (2016). Ubiquitin modifications. *Cell Res.* *26*, 399–422.

Szwajczak, E., Fijalkowska, I.J., and Suski, C. (2017). The CysB motif of Rev3p involved in the formation of the four-subunit DNA polymerase ζ is required for defective-replisome-induced mutagenesis. *Mol. Microbiol.* *106*, 659–672.

Taddei, A., Maison, C., Roche, D., and Almouzni, G. (2001). Reversible disruption of pericentric heterochromatin and centromere function by inhibiting deacetylases. *Nat. Cell Biol.* *3*, 114–120.

Takai, H., Smogorzewska, A., and de Lange, T. (2003). DNA damage foci at dysfunctional telomeres. *Curr. Biol. CB* *13*, 1549–1556.

Takeda, S., Chen, D.Y., Westergard, T.D., Fisher, J.K., Rubens, J.A., Sasagawa, S., Kan, J.T., Korsmeyer, S.J., Cheng, E.H.-Y., and Hsieh, J.J.-D. (2006). Proteolysis of MLL family proteins is essential for taspase1-orchestrated cell cycle progression. *Genes Dev.* *20*, 2397–2409.

Takeda, S., Liu, H., Sasagawa, S., Dong, Y., Trainor, P.A., Cheng, E.H., and Hsieh, J.J. (2013). HGF-MET signals via the MLL-ETS2 complex in hepatocellular carcinoma. *J. Clin. Invest.* *123*, 3154–3165.

Takeda, S., Sasagawa, S., Oyama, T., Searleman, A.C., Westergard, T.D., Cheng, E.H., and Hsieh, J.J. (2015). Taspase1-dependent TFIIA cleavage coordinates head morphogenesis by limiting Cdkn2a locus transcription. *J. Clin. Invest.* *125*, 1203–1214.

Tan, J., Cang, S., Ma, Y., Petrillo, R.L., and Liu, D. (2010). Novel histone deacetylase inhibitors in clinical trials as anti-cancer agents. *J. Hematol. Oncol.* *J Hematol Oncol* *3*, 5.

Tang, W., Dominska, M., Gawel, M., Greenwell, P.W., and Petes, T.D. (2013). Genomic deletions and point mutations induced in *Saccharomyces cerevisiae* by the trinucleotide repeats (GAA·TTC) associated with Friedreich's ataxia. *DNA Repair* *12*, 10–17.

Tomida, J., Takata, K., Lange, S.S., Schibler, A.C., Yousefzadeh, M.J., Bhetawal, S., Dent,

- S.Y.R., and Wood, R.D. (2015). REV7 is essential for DNA damage tolerance via two REV3L binding sites in mammalian DNA polymerase ζ . *Nucleic Acids Res.* *43*, 1000–1011.
- Tomkova, M., Tomek, J., Kriaucionis, S., and Schuster-Böckler, B. (2018). Mutational signature distribution varies with DNA replication timing and strand asymmetry. *Genome Biol.* *19*, 129.
- Torres, I.O., and Fujimori, D.G. (2015). Functional coupling between writers, erasers and readers of histone and DNA methylation. *Curr. Opin. Struct. Biol.* *35*, 68–75.
- Trojer, P., and Reinberg, D. (2007). Facultative heterochromatin: is there a distinctive molecular signature? *Mol. Cell* *28*, 1–13.
- Vaisman, A., and Woodgate, R. (2017). Translesion DNA polymerases in eukaryotes: what makes them tick? *Crit. Rev. Biochem. Mol. Biol.* *52*, 274–303.
- Vaisman, A., Masutani, C., Hanaoka, F., and Chaney, S.G. (2000). Efficient translesion replication past oxaliplatin and cisplatin GpG adducts by human DNA polymerase η . *Biochemistry* *39*, 4575–4580.
- Van Sloun, P.P.H., Varlet, I., Sonneveld, E., Boei, J.J.W.A., Romeijn, R.J., Eeken, J.C.J., and De Wind, N. (2002). Involvement of mouse Rev3 in tolerance of endogenous and exogenous DNA damage. *Mol. Cell. Biol.* *22*, 2159–2169.
- Varadi, V., Bevier, M., Grzybowska, E., Johansson, R., Enquist, K., Henriksson, R., Butkiewicz, D., Pamula-Pilat, J., Tecza, K., Hemminki, K., et al. (2011). Genetic variation in genes encoding for polymerase ζ subunits associates with breast cancer risk, tumour characteristics and survival. *Breast Cancer Res. Treat.* *129*, 235–245.
- Vastenhouw, N.L., and Schier, A.F. (2012). Bivalent histone modifications in early embryogenesis. *Curr. Opin. Cell Biol.* *24*, 374–386.
- Verdin, E., and Ott, M. (2015). 50 years of protein acetylation: from gene regulation to epigenetics, metabolism and beyond. *Nat. Rev. Mol. Cell Biol.* *16*, 258–264.
- Verdin, E., Dequiedt, F., and Kasler, H.G. (2003). Class II histone deacetylases: versatile regulators. *Trends Genet. TIG* *19*, 286–293.
- Vesela, E., Chroma, K., Turi, Z., and Mistrik, M. (2017). Common Chemical Inductors of Replication Stress: Focus on Cell-Based Studies. *Biomolecules* *7*.
- Vicient, C.M., and Casacuberta, J.M. (2017). Impact of transposable elements on polyploid plant genomes. *Ann. Bot.* *120*, 195–207.
- Volpe, T.A., Kidner, C., Hall, I.M., Teng, G., Grewal, S.I.S., and Martienssen, R.A. (2002). Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi.

Science 297, 1833–1837.

Wang, H., Zhang, S.-Y., Wang, S., Lu, J., Wu, W., Weng, L., Chen, D., Zhang, Y., Lu, Z., Yang, J., et al. (2009). REV3L confers chemoresistance to cisplatin in human gliomas: the potential of its RNAi for synergistic therapy. *Neuro-Oncol.* 11, 790–802.

Wang, J., Jia, S.T., and Jia, S. (2016). New Insights into the Regulation of Heterochromatin. *Trends Genet.* TIG 32, 284–294.

Wang, W., Sheng, W., Yu, C., Cao, J., Zhou, J., Wu, J., Zhang, H., and Zhang, S. (2015). REV3L modulates cisplatin sensitivity of non-small cell lung cancer H1299 cells. *Oncol. Rep.* 34, 1460–1468.

Wang, Z., Zang, C., Rosenfeld, J.A., Schones, D.E., Barski, A., Cuddapah, S., Cui, K., Roh, T.-Y., Peng, W., Zhang, M.Q., et al. (2008). Combinatorial patterns of histone acetylations and methylations in the human genome. *Nat. Genet.* 40, 897–903.

Warner, M.D., Azmi, I.F., Kang, S., Zhao, Y., and Bell, S.P. (2017). Replication origin-flanking roadblocks reveal origin-licensing dynamics and altered sequence dependence. *J. Biol. Chem.* 292, 21417–21430.

Waters, L.S., Minesinger, B.K., Wiltrout, M.E., D’Souza, S., Woodruff, R.V., and Walker, G.C. (2009). Eukaryotic Translesion Polymerases and Their Roles and Regulation in DNA Damage Tolerance. *Microbiol. Mol. Biol. Rev.* MMBR 73, 134–154.

Wittschieben, J., Shivji, M.K., Lalani, E., Jacobs, M.A., Marini, F., Gearhart, P.J., Rosewell, I., Stamp, G., and Wood, R.D. (2000). Disruption of the developmentally regulated Rev3l gene causes embryonic lethality. *Curr. Biol.* CB 10, 1217–1220.

Wittschieben, J.P., Reshmi, S.C., Gollin, S.M., and Wood, R.D. (2006). Loss of DNA polymerase zeta causes chromosomal instability in mammalian cells. *Cancer Res.* 66, 134–142.

Wittschieben, J.P., Patil, V., Glushets, V., Robinson, L.J., Kusewitt, D.F., and Wood, R.D. (2010). Loss of DNA polymerase zeta enhances spontaneous tumorigenesis. *Cancer Res.* 70, 2770–2778.

Wong, L.H. (2010). Epigenetic regulation of telomere chromatin integrity in pluripotent embryonic stem cells. *Epigenomics* 2, 639–655.

Wong, L.H., McGhie, J.D., Sim, M., Anderson, M.A., Ahn, S., Hannan, R.D., George, A.J., Morgan, K.A., Mann, J.R., and Choo, K.H.A. (2010). ATRX interacts with H3.3 in maintaining telomere structural integrity in pluripotent embryonic stem cells. *Genome Res.* 20, 351–360.

Woo, Y.H., and Li, W.-H. (2012). DNA replication timing and selection shape the landscape

- of nucleotide variation in cancer genomes. *Nat. Commun.* **3**, 1004.
- Xiao, W., Lechler, T., Chow, B.L., Fontanie, T., Agustus, M., Carter, K.C., and Wei, Y.F. (1998). Identification, chromosomal mapping and tissue-specific expression of hREV3 encoding a putative human DNA polymerase zeta. *Carcinogenesis* **19**, 945–949.
- Xie, K., Doles, J., Hemann, M.T., and Walker, G.C. (2010). Error-prone translesion synthesis mediates acquired chemoresistance. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 20792–20797.
- Yang, W., and Gao, Y. (2018). Translesion and Repair DNA Polymerases: Diverse Structure and Mechanism. *Annu. Rev. Biochem.* **87**, 239–261.
- Yang, L., Shi, T., Liu, F., Ren, C., Wang, Z., Li, Y., Tu, X., Yang, G., and Cheng, X. (2015). REV3L, a promising target in regulating the chemosensitivity of cervical cancer cells. *PLoS One* **10**, e0120334.
- Yazinski, S.A., and Zou, L. (2016). Functions, Regulation, and Therapeutic Implications of the ATR Checkpoint Pathway. *Annu. Rev. Genet.* **50**, 155–173.
- Young, M.D., Willson, T.A., Wakefield, M.J., Trounson, E., Hilton, D.J., Blewitt, M.E., Oshlack, A., and Majewski, I.J. (2011). ChIP-seq analysis reveals distinct H3K27me3 profiles that correlate with transcriptional activity. *Nucleic Acids Res.* **39**, 7415–7427.
- Zander, L., and Bemark, M. (2004). Immortalized mouse cell lines that lack a functional Rev3 gene are hypersensitive to UV irradiation and cisplatin treatment. *DNA Repair* **3**, 743–752.
- Zasadzińska, E., Huang, J., Bailey, A.O., Guo, L.Y., Lee, N.S., Srivastava, S., Wong, K.A., French, B.T., Black, B.E., and Foltz, D.R. (2018). Inheritance of CENP-A Nucleosomes during DNA Replication Requires HJURP. *Dev. Cell* **47**, 348–362.e7.
- Zeman, M.K., and Cimprich, K.A. (2014). Causes and consequences of replication stress. *Nat. Cell Biol.* **16**, 2–9.
- Zhang, H., Chatterjee, A., and Singh, K.K. (2006). *Saccharomyces cerevisiae* polymerase zeta functions in mitochondria. *Genetics* **172**, 2683–2688.
- Zhang, S., Chen, H., Zhao, X., Cao, J., Tong, J., Lu, J., Wu, W., Shen, H., Wei, Q., and Lu, D. (2013). REV3L 3'UTR 460 T>C polymorphism in microRNA target sites contributes to lung cancer susceptibility. *Oncogene* **32**, 242–250.
- Zhao, L., and Washington, M.T. (2017). Translesion Synthesis: Insights into the Selection and Switching of DNA Polymerases. *Genes* **8**.
- Zhong, X., Garg, P., Stith, C.M., Nick McElhinny, S.A., Kissling, G.E., Burgers, P.M.J., and Kunkel, T.A. (2006). The fidelity of DNA synthesis by yeast DNA polymerase zeta alone and with accessory proteins. *Nucleic Acids Res.* **34**, 4731–4742.
- Zhou, F.C. (2012). DNA methylation program during development. *Front. Biol.* **7**, 485–494.

- Zhou, H., Spicuglia, S., Hsieh, J.J.-D., Mitsiou, D.J., Høiby, T., Veenstra, G.J.C., Korsmeyer, S.J., and Stunnenberg, H.G. (2006). Uncleaved TFIIA is a substrate for taspase 1 and active in transcription. *Mol. Cell. Biol.* 26, 2728–2735.
- Zhu, F., Jin, C.-X., Song, T., Yang, J., Guo, L., and Yu, Y.-N. (2003). Response of human REV3 gene to gastric cancer inducing carcinogen N-methyl-N'-nitro-N-nitrosoguanidine and its role in mutagenesis. *World J. Gastroenterol.* 9, 888–893.
- Zhu, X., Zou, S., Zhou, J., Zhu, H., Zhang, S., Shang, Z., Ding, W.-Q., Wu, J., and Chen, Y. (2016a). REV3L, the catalytic subunit of DNA polymerase ζ , is involved in the progression and chemoresistance of esophageal squamous cell carcinoma. *Oncol. Rep.* 35, 1664–1670.
- Zhu, X., Zou, S., Zhou, J., Zhu, H., Zhang, S., Shang, Z., Ding, W.-Q., Wu, J., and Chen, Y. (2016b). REV3L, the catalytic subunit of DNA polymerase ζ , is involved in the progression and chemoresistance of esophageal squamous cell carcinoma. *Oncol. Rep.* 35, 1664–1670.
- Zlatanou, A., Despras, E., Braz-Petta, T., Boubakour-Azzouz, I., Pouvelle, C., Stewart, G.S., Nakajima, S., Yasui, A., Ishchenko, A.A., and Kannouche, P.L. (2011). The hMsh2-hMsh6 complex acts in concert with monoubiquitinated PCNA and Pol η in response to oxidative DNA damage in human cells. *Mol. Cell* 43, 649–662.
- Zofall, M., Smith, D.R., Mizuguchi, T., Dhakshnamoorthy, J., and Grewal, S.I.S. (2016). Taz1-Shelterin Promotes Facultative Heterochromatin Assembly at Chromosome-Internal Sites Containing Late Replication Origins. *Mol. Cell* 62, 862–874.
- Zuo, L., Zhou, T., Pannell, B.K., Ziegler, A.C., and Best, T.M. (2015). Biological and physiological role of reactive oxygen species--the good, the bad and the ugly. *Acta Physiol. Oxf. Engl.* 214, 329–348.

Titre : Fonction et régulation de l'ADN polymérase zêta au cours de la réplication de l'ADN: conséquences sur la stabilité du génome.

Mots clés : Réplication de l'ADN ; polymérase translésionnelle ; Instabilité génétique ; Hétérochromatine ; Mutagenèse ; Stress génotoxique, TASP1

Résumé : La réplication de l'ADN est un processus cellulaire fondamental qui assure la duplication fidèle de l'information génétique. Différentes perturbations peuvent interférer avec la progression de la fourche de réplication menaçant ainsi l'intégrité du génome. Pour éviter le blocage du réplisome, les polymérases répliquatives peuvent être remplacées par des polymérases translésionnelles (TLS) mutagènes mais capables de franchir différents types de lésions. Parmi les polymérases TLS, Pol ζ est unique car son absence entraîne une létalité embryonnaire chez la souris suggérant qu'elle a acquis une fonction essentielle au cours de l'évolution. Cependant, sa fonction et sa régulation dans les cellules mammifères restent méconnues. Dans ce travail, nous avons montré que la phase S est perturbée en absence de REV3L, avec une modification du programme temporel de la réplication au niveau de régions génomiques répliquées en milieu-fin de phase S. Ce défaut de réplication est associé à une augmentation de la mutagenèse et des modifications du paysage épigénétique. Nous avons de plus mis en évidence que REV3L interagit avec les composants de l'hétérochromatine et est localisé au niveau des régions péri-centromériques, ce qui suggère que Pol ζ participe à la réplication de l'hétérochromatine et limite ainsi l'instabilité génomique. Dans une seconde partie, nous avons découvert que la protéine REV3L est clivée de manière post-traductionnelle par l'endopeptidase TASP1 générant deux polypeptides capables de s'hétérodimériser pour former un complexe stable qui, en association avec Rev7, représente probablement le complexe actif de Pol ζ . Aussi, nous avons observé que REV3L est finement régulée de manière endogène ou après exposition à un stress génotoxique à de multiples niveaux : (1) au niveau transcriptionnel, (2) par le clivage par TASP1, (3) par des phosphorylations post-traductionnelles. Finalement, l'ensemble de ces découvertes met en lumière un mécanisme de régulation unique contrôlant la fonction d'une polymérase mutagène dans les cellules mammifères. Ces résultats sont particulièrement importants étant donné que Pol ζ est un facteur impliqué dans les mécanismes de résistance des tumeurs face aux chimiothérapies.

Title : Function and regulation of the DNA polymerase zeta during the DNA replication: consequences on genome stability.

Keywords : DNA replication ; Translesion DNA polymerase ; Genetic instability ; Heterochromatin; Mutagenesis ; Genotoxic stress, TASP1

Abstract : DNA replication is a fundamental process that ensures accurate duplication of the genetic information. Various perturbations can impede replication fork progression, and thus threatening genome integrity. To prevent fork collapse, replicative DNA polymerases can be replaced by error-prone DNA polymerases called translesion (TLS) polymerases, able to bypass DNA damage at the cost of increased mutations. Among TLS polymerases, Pol ζ is unique because inactivation of its catalytic subunit, REV3L, leads to embryonic lethality in mice underscoring its biological importance. However, little is known about its function and regulation in mammalian cells. We showed that loss of REV3L impairs S phase progression with a disruption of replication timing at specific genomic loci that replicate in mid-late S-phase, and this is associated with increased mutagenic events and aberrant epigenetic landscape. We also revealed that REV3L interacts with heterochromatin components and localizes in pericentromeric regions, suggesting that Pol ζ contributes to replicate heterochromatin regions to limit genome instability. In a second part, we discovered that REV3L protein is proteolytically processed by the endopeptidase TASP1 to generate two polypeptides that heterodimerize to form a stable complex that associates with REV7, likely representing the active complex of Pol ζ . We also found that REV3L is finely regulated in physiological conditions and after genotoxic stress at multiple levels: (1) transcriptionally, (2) proteolytically by TASP1 and (3) post-translationally by phosphorylation. Altogether these findings highlight a unique mechanism to control the function of an error-prone polymerase in mammalian cells. These data are particularly important given that Pol ζ is an important factor for tumor resistance to chemotherapeutic agents.