



Influence des séquences subtélomériques sur la régulation des télomères : exemple du locus de la Dystrophie Facio-Scapulo-Humérale en 4q35 et implication en pathologie

Caroline Schluth-Bolard

► To cite this version:

Caroline Schluth-Bolard. Influence des séquences subtélomériques sur la régulation des télomères : exemple du locus de la Dystrophie Facio-Scapulo-Humérale en 4q35 et implication en pathologie. Biologie cellulaire. Ecole normale supérieure de lyon - ENS LYON, 2011. Français. <NNT : 2011ENSL0623>. <tel-00689555>

HAL Id: tel-00689555

<https://tel.archives-ouvertes.fr/tel-00689555>

Submitted on 19 Apr 2012

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.

N° d'ordre : 623

N° attribué par la bibliothèque : _ENSL623

THÈSE

en vue d'obtenir le grade de

Docteur de l'Université de Lyon – Ecole Normale Supérieure de Lyon

spécialité : Sciences de la vie

LABORATOIRE DE BIOLOGIE MOLECULAIRE DE LA CELLULE -UMR 5239-LBMC

Ecole Doctorale de Biologie Moléculaire Intégrative et Cellulaire

présentée et soutenue publiquement le 30/05/2011

par Madame Caroline SCHLUTH épouse SCHLUTH-BOLARD

Titre :

***Influence des séquences subtélomériques sur la régulation des télomères :
exemple du locus de la Dystrophie Facio-Scapulo-Humérale en 4q35 et implication en pathologie.***

Directeur de thèse : Monsieur Eric GILSON

*Après avis de : MME Kerstin BYSTRICKY
MME Mary CALLANAN*

Devant la Commission d'examen formée de :

*Madame Kerstin BYSTRICKY, Membre/Rapporteur
Madame Mary CALLANAN, Membre/Rapporteur
Monsieur Eric GILSON (directeur de thèse), Membre
Madame Frédérique MAGDINIER, Membre
Monsieur Damien SANLAVILLE, Membre
Monsieur Laurent SCHAEFFER, Président*

REMERCIEMENTS

A Monsieur le Professeur Eric Gilson, pour m'avoir accueillie dans son équipe, pour sa disponibilité et les conseils qu'il m'a apportés tout au long de cette thèse.

A Monsieur le Professeur Laurent Schaeffer, qui m'a permis de réaliser ce travail dans son laboratoire et qui a accepté de faire partie du jury de cette thèse.

A Frédérique Magdinier pour la qualité de son encadrement et son enthousiasme communicatif pour la génétique et l'épigénétique.

A toute l'équipe Régulations Epigénétiques et Télomériques, et plus particulièrement Amina Boussouar, Irimina Diala, Julia Morere, Alexandre Ottaviani, Jing Ye, Serge Bauwens, Delphine Rondier, Marie Monniaux et Nathalie Pion, pour leur aide et leur bonne humeur.

A l'équipe du PLATIM et à Sandrine Mouradian, pour leur soutien technique.

A Monsieur le Professeur Bruno Lina, qui m'a prêté un bureau calme et propice à la réflexion pour rédiger cette thèse.

A Mme Kerstin Bystricky et Mme Mary Callanan qui ont accepté d'être rapporteur de ce travail.

A Damien Sanlaville, pour l'intérêt qu'il a porté pour ce travail et pour avoir accepté de participer au jury de cette thèse.

A mes deux plus fidèles supporters, Jean-François et Lucas.

Ce travail a été soutenu par un financement couplé d'Assistant Hospitalier de Recherche HCL-CNRS.

RESUME

Les subtélomères forment la transition entre les séquences spécifiques des chromosomes et les répétitions télomériques terminales. Ils semblent capables d'influencer les fonctions télomériques mais les connaissances sur les mécanismes mis en jeu sont encore limitées. Les subtélomères sont pourtant associés à de nombreuses pathologies comme la myopathie facio-scapulo-humérale (FSHD), une dystrophie musculaire secondaire à la contraction de répétitions macrosatellites *D4Z4* dans la région subtélomérique 4q35.

Afin d'étudier les propriétés de la séquence subtélomérique *D4Z4*, nous avons créé des constructions reproduisant l'organisation génomique au locus 4q35. Nous avons montré que *D4Z4* est capable d'adresser un télomère à la périphérie du noyau. Cette activité est couplée à une activité insulatrice au niveau d'une séquence proximale de 80 pb et est dépendante de CTCF et des Lamines A. De plus, la relocalisation périphérique d'un télomère par *D4Z4* s'accompagne d'une réplication plus tardive de celui-ci. Par ailleurs, la recherche de séquences capables de s'opposer à l'effet de position télomérique (TPE) a identifié un élément de 30 pb contenant un site CTCF dans la séquence insulatrice proximale de *D4Z4*. De même, l'introduction d'un signal de poly-adénylation entre un gène rapporteur et les répétitions télomériques interfère avec le TPE et est accompagnée d'une diminution d'un transcrit hybride contenant le gène rapporteur et des répétitions télomériques, suggérant un rôle des transcrits télomériques TERRAs dans la régulation du TPE.

En conclusion, ce travail a permis de caractériser l'implication de séquences subtélomériques, et notamment *D4Z4*, dans la régulation des télomères, leur compartimentalisation nucléaire, la réplication ou l'effet de position télomérique. De plus, il apporte un éclairage nouveau sur la physiopathologie de la FSHD et ouvre des perspectives dans la compréhension d'autres pathologies liées aux subtélomères.

Mots-clés : Télomère ; subtélomère ; *D4Z4* ; dystrophie musculaire facio-scapulo-humérale; transcrits télomériques.

ABSTRACT

Subtelomeres form the transition between chromosome specific sequences and terminal telomeric repeats. They might influence telomeric functions but underlying mechanisms are still unclear. Nevertheless, subtelomeres are associated with a number of human pathologies such as facioscapulohumeral muscular dystrophy (FSHD), an autosomal dominant disease secondary to the contraction of an array of *D4Z4* macrosatellite repeats in the subtelomeric region 4q35.

In order to study the biological function of the *D4Z4* sequence, we created constructs that mimic the genomic organization of the 4q35 locus. We showed that *D4Z4* is able to localize a telomere at the nuclear periphery. This perinuclear activity was dependant on interactions with CTCF and A type lamins and lied within a 80 bp proximal sequence that harbors an insulator activity. Moreover, the peripheral positionning of a telomere by *D4Z4* is accompanied by a late replication timing of the telomere. We also searched for sequences able to counteract telomeric position effect (TPE) and identified a 30 bp element containing a CTCF binding site in the proximal region of *D4Z4*. In another construct, the introduction of a poly-adenylation signal between a reporter gene and telomeric repeats counteracted TPE. This effect is accompanied by the production of a hybrid transcript encompassing the reporter gene and telomeric repeats, suggesting a role for the TERRAs telomeric transcripts in TPE regulation.

This work contibuted to characterize the role of subtelomeric sequences, especially the *D4Z4* macrosatellite, in telomere regulation, their nuclear compartmentalization, their replication or the telomeric position effect. We will discuss the implications in the understanding of the pathophysiology of FSHD and other subtelomeric diseases.

Key-words : telomere; subtelomere; D4Z4; facioscapulohumeral muscular dystrophy; telomeric containing repeat RNA.

TABLE DES MATIERES

1. INTRODUCTION.....	11
1. 1. LES TELOMERES.....	12
1.1.1. Définition et historique.....	12
1.1.2. Structure des télomères.....	13
1.1.2.1. Séquence nucléotidique.....	13
1.1.2.2. Le complexe télosome ou shelterin.....	13
1.1.3. Fonctions des télomères.....	15
1.1.3.1. Protection des extrémités chromosomiques.....	15
1.1.3.1.1. Réplication des extrémités et télomérase.....	15
1.1.3.1.2. Protection contre les signaux de dommages à l'ADN et les systèmes de réparation de l'ADN.....	19
1.1.3.1.3. Conséquences : Sénescence, cancer et vieillissement.....	22
1.1.3.2. Télomères et reproduction.....	24
1.1.3.3. Effet de position télomérique.....	25
1.1.3.3.1. Rappels sur l'effet de position chromosomique.....	25
1.1.3.3.2. L'effet de position télomérique.....	27
1.1.4. Régulation épigénétique des télomères.....	29
1.1.4.1. La chromatine télomérique.....	30
1.1.4.2. Les transcrits télomériques ou TERRAs.....	32
1.1.4.3. Compartimentalisation nucléaire.....	36
1.1.4.3.1. Organisation tridimensionnelle des télomères dans le noyau.....	36
1.1.4.3.2. Rôle de la périphérie nucléaire.....	39
1. 2. LES SUBTELOMERES.....	44
1.2.1. Définition.....	44
1.2.2. Structure des subtélomères.....	44
1.2.3. Rôle des subtélomères.....	49
1.2.3.1. Transcrits et gènes subtélomériques.....	49
1.2.3.2. Subtélomères et effet de position télomérique.....	50
1.2.3.3. Le mécanisme ALT : "Alternative Lengthening of telomere".....	52
1. 3. EXEMPLE D'UN ELEMENT SUBTELOMERIQUE IMPLIQUE EN PATHOLOGIE HUMAINE : <i>D4Z4</i> ET LA MYOPATHIE FACIO-SCAPULO-HUMERALE.....	55
1.3.1. <i>D4Z4</i> , un macrosatellite subtélomérique.....	55
1.3.2. La myopathie facio-scapulo-humérale, aspects cliniques et biologiques.....	58
1.3.2.1. Clinique.....	58
1.3.2.2. Evolution.....	59
1.3.2.3. Biologie moléculaire.....	59
1.3.2.4. Diagnostic moléculaire.....	62
1.3.3. Les hypothèses physiopathologiques de la FSHD.....	62
1.3.3.1. La méthylation de l'ADN.....	63
1.3.3.2. Chromatine et histones.....	63
1.3.3.3. Complexe répresseur.....	64
1.3.3.4. Le modèle de boucle chromatinienne.....	65
1.3.3.5. A la recherche de gènes candidats.....	68
1.3.3.6. <i>DUX4</i> : Enfin la solution ?.....	69
1.3.3.7. <i>D4Z4</i> et périphérie nucléaire.....	70
1.3.3.8. FSHD : Gain de fonction de CTCF ?.....	71
2. OBJECTIFS DE L'ETUDE, MATERIELS ET METHODES.....	74
2.1. OBJECTIFS DE L'ETUDE.....	74
2.2. MATERIELS ET METHODES.....	76
2.2.1. Modèle cellulaire.....	76
2.2.2. Etude de l'organisation nucléaire en 3D-immuno-FISH.....	78
3. RESULTATS.....	79
3.1. INTERACTION DE <i>D4Z4</i> AVEC LA PERIPHERIE NUCLEAIRE.....	79
3.2. ROLES DES SEQUENCES SUBTELOMERIQUES DANS LA REGULATION DE LA REPLICATION DES TELOMERES.....	88
3.3. INFLUENCE DES SEQUENCES SUBTELOMERIQUES SUR L'EFFET DE POSITION TELOMERIQUE.....	94
3.3.1. Rôle du site insulateur proximal de <i>D4Z4</i>	95

3.3.2. Caractérisation d'un site CTCF dans la partie distale de D4Z4.....	98
3.3.3. Identification de séquences pouvant interférer avec l'effet de position télomérique.....	101
3.3.4. Exploration du rôle des transcrits télomériques TERRAs dans la régulation du TPE.....	104
4. DISCUSSION ET PERSPECTIVES.....	113
4.1. INFLUENCE DE LA LAMINA NUCLEAIRE SUR LES FONCTIONS TELOMERIQUES.....	113
4.2. REGULATION DE L'EFFET DE POSITION TELOMERIQUE CHEZ L'HOMME.....	118
4.2.1. Quel rôle pour les TERRAs dans le TPE ?	118
4.2.2. Rôle des insulateurs et de CTCF dans la régulation du TPE.....	121
4.3. LES PATHOLOGIES SUBTELOMERIQUES CONSTITUTIONNELLES.....	123
4.3.1. La FSHD : un modèle de pathologie subtélomérique, une nouvelle laminopathie, ou... ?	123
4.3.2. Les syndromes subtélomériques.....	127
5. CONCLUSION.....	131
ANNEXES	132
BIBLIOGRAPHIE.....	139

TABLE DES FIGURES ET DES TABLEAUX

Figure 1 : Structure des télomères	14
Figure 2 : Réplication des télomères	16
Figure 3 : Structure de la télomérase	18
Figure 4 : Mécanismes de réparation mis en jeu aux extrémités télomériques	21
Figure 5 : L'attrition des télomères contrôle la capacité proliférative des cellules somatiques	23
Figure 6 : Les différents types d'effet de position	26
Figure 7 : Les transcrits télomériques	33
Figure 8 : Organisation tridimensionnelle du noyau	37
Figure 9 : Structure des régions subtélomériques	45
Figure 10 : Dynamique des régions subtélomériques	48
Figure 11 : Maintenance des télomères par le mécanisme ALT	53
Figure 12 : Le macrosatellite <i>D4Z4</i>	57
Figure 13 : Mécanisme moléculaire de la FSHD	61
Figure 14 : Hypothèses physiopathologiques de la FSHD	66
Figure 15 : Cartographie des éléments régulateurs de <i>D4Z4</i>	72
Figure 16 : Matériels et méthodes	77
Figure 17 : Rôle de la séquence consensus 30pb dans la localisation nucléaire périphérique des chromosomes	84
Figure 18 : Etude du rôle de la séquence <i>D4S139</i> dans la localisation nucléaire périphérique de la région 4qter	86
Figure 19 : Rôle du site CTCF de l'insulateur proximal de <i>D4Z4</i> sur le TPE	96
Figure 20 : Caractérisation du site CTCF distal de <i>D4Z4</i>	100
Figure 21 : Influence de la séquence <i>D4S139</i> sur le TPE	102
Figure 22 : Mise en évidence d'un transcrit GFP-TERRA par northern-blot	106
Figure 23 : Mise en évidence d'un transcrit GFP-TERRA par transcription inverse	107
Figure 24 : Le signal de poly-adénylation n'affecte pas la stabilité des transcrits <i>GFP</i>	109
Figure 25 : Etude de la chromatine au niveau des télomères et du locus <i>eGFP</i>	112
Figure 26 : Modèle cellulaire conditionnel	117
Figure 27 : Proposition d'une hypothèse physiopathologique de la FSHD	126
Tableau I : Principales maladies dues à une mutation du gène <i>LMNA</i>	42
Tableau II : Gènes morbides présents dans le dernier mégabase des extrémités chromosomiques	129

ACRONYMES ET ABBREVIATIONS

A

AD	Autosomique Dominant
ADN	Acide désoxyribonucléique
ALT	Alternative Lengthening of Telomeres
APBs	ALT associated PML Bodies
AR	Autosomique Récessif
ARN	Acide ribonucléique

B

BAC	Chromosome Artificiel de Bactérie
BrdC	Bromodeoxycytidine
BrdU	Bromodeoxyuridine

C

CGH-array	Hybridation Génomique Comparative sur Puce ADN
ChIP	Immunoprécipitation de chromatine
CNV	Variation du nombre de copies
CPE	Effet de position chromosomique

D

DAPI	4',6'-diamidino-2-phénylindole
DH	Sites hypersensibles à la DNase I
DS	Déviation Standard

E

ES	Embryonnaire Souche
----	---------------------

F

FACs	Cytométrie de Flux
FISH	Hybridation <i>in situ</i> en fluorescence
FSHD	Dystrophie Facio-Scapulo-Humérale

H

HDAC	Histone désacétylase
------	----------------------

I

IPS	Induced Pluripotent Stem cells
-----	--------------------------------

K

kb	kilobase
----	----------

L

LAD	domaine associé à la lamina
lncRNA	long ARN non codant

M	
Mb	mégabase
miRNA	micro-ARN
N	
NAHR	recombinaison homologue non allélique
NHEJ	réparation par fusion d'extrémités non homologues
NMD	dégradation des ARNs non-sens
P	
PAC	Chromosome artificiel de phage
PCR	Polymerase Chain Reaction
PEV	Position Effect Variegation
pb	paire de base
Q	
qPCR	PCR quantitative
R	
ReDFISH	Replicative Detargetting FISH
RNAi	Interférence par l'ARN
RT-qPCR	Transcription inverse – PCR quantitative
S	
siRNA	petit ARN interférant
STR	Short Tandem Repeat
STS	Sequence Tagged Site
Srpt	duplicon subtélomérique subterminal
SNP	Single Nucleotide Polymorphism
S/MAR	Scaffold / Matrix Attachment Region
T	
TERRA	transcrit télomérique
TIF	Telomere dysfunction Induced Foci
TPE	effet de position télomérique
TSA	trichostatine A
T-SCE	échange de chromatides sœurs télomériques
V	
VNTR	Variable Number of Tandem Repeat
Y	
YAC	Chromosome artificiel de levure
3D	Tridimensionnel

1. INTRODUCTION

Chez les eucaryotes, l'information génétique est portée par des chromosomes linéaires. Les extrémités de ces chromosomes sont composées de structures nucléoprotéiques spécialisées, appelées télomères qui assurent la stabilité des chromosomes. Ces télomères sont séparés des séquences spécifiques du chromosome par des régions complexes de taille variable, les subtélomères. Si les télomères sont extensivement étudiés, les connaissances sur les séquences subtélomériques sont encore parcellaires, notamment du fait de leur grande hétérogénéité rendant leur étude difficile. Il apparaît cependant que les subtélomères peuvent subir l'influence des télomères mais pourraient également interférer avec les fonctions télomériques.

Nous rappellerons tout d'abord la structure et la fonction des télomères avant de présenter les séquences subtélomériques. Etant donné l'ampleur du sujet, nous nous limiterons aux mammifères, et en particulier à l'Homme. La référence à d'autres eucaryotes ne sera faite que ponctuellement. Ensuite nous détaillerons l'exemple de *D4ZA*, une séquence macrosatellite subtélomérique impliquée dans la myopathie facio-scapulo-humérale. Nous avons étudié ses fonctions dans la compartimentalisation nucléaire des télomères, leur réplication et le contrôle de l'effet de position télomérique. Nous discuterons enfin des relations télomères-subtélomères et de leurs conséquences potentielles en pathologie humaine constitutionnelle.

1. 1. Les télomères

1.1.1. Définition et historique

Le terme de télomère, du grec *telo* (fin) et *meros* (partie), a été introduit par H. Muller (Muller, 1938). Il désigne actuellement des structures nucléoprotéiques spécialisées localisées aux extrémités des chromosomes linéaires eucaryotes, nécessaires à leur stabilité et à leur réplication (Blackburn, 1984).

La notion de télomère est apparue à la fin des années 1930s à partir de travaux menés sur la drosophile (Muller, 1938) et le maïs (McClintock, 1941) qui ont remarqué que les extrémités des chromosomes ayant subi des cassures, induites par les UV ou des ponts anaphasiques, fusionnaient entre elles alors que les extrémités chromosomiques naturelles ne fusionnaient pas.

Ensuite, il a été observé que, du fait du mode semi-conservatif de la réplication de l'ADN, une molécule d'ADN linéaire est incomplètement répliquée et se raccourcit à chaque cycle de réplication (Watson, 1972). Ce problème de réplication des extrémités ("*end-replication problem*") a été mis en avant pour expliquer le nombre limité de divisions des cellules en culture (Hayflick, 1965; Olovnikov, 1973). Les liens entre raccourcissement télomérique et sénescence ont été étudiés plus tard, notamment chez la levure (Lundblad & Szostak, 1989).

La séquence des télomères a d'abord été identifiée chez un organisme cilié, *Tetrahymena thermophila* (Blackburn & Gall, 1978), puis dix ans plus tard chez l'Homme (Moyzis et al., 1988). Les différents composants fonctionnels du télomère ont ensuite été caractérisés comme la télomérase ("*Telomere Terminal Transferase*") (Greider & Blackburn, 1985) et les protéines associées aux télomères TRF1 (Chong et al., 1995) et TRF2 (Bilaud et al., 1997). Les publications sur le sujet n'ont cessé d'augmenter (Gilson & Segal-Bendirdjian, 2010) et les recherches sur "comment les chromosomes sont protégés par les télomères et la télomérase" ont été couronnées en 2009 par l'attribution du Prix Nobel de Médecine et Physiologie à E. Blackburn, C. Greider et J. Szostak.

1.1.2. Structure des télomères

1.1.2.1. Séquence nucléotidique

La structure des télomères est relativement bien conservée chez les eucaryotes et est composée, hormis quelques exceptions, de courtes répétitions double brin en tandem de 6 à 8 nucléotides, riches en G, se terminant par une extrémité 3' sortante simple brin (**Figure 1a**).

Chez les mammifères, et notamment chez l'Homme, la séquence télomérique est composée de répétitions (TTAGGG)_n (Moyzis et al., 1988). Le nombre de répétitions est variable en fonction des espèces. La taille des télomères est de 10 à 15 kilobases (kb) chez l'Homme (de Lange et al., 1990) et de 20 à 50 kb chez la souris (Kipling & Cooke, 1990). Ce nombre peut également varier selon le type cellulaire et d'un chromosome à l'autre dans le même organisme (de Lange et al., 1990).

Le brin riche en G (et dépourvu de C), encore appelé brin G, forme une extrémité simple brin 3' sortante dont la longueur est estimée de 130 à 210 bases chez l'Homme (Makarov et al., 1997; McElligott & Wellinger, 1997). Cette extrémité peut envahir la partie double brin pour former une structure en lasso appelée boucle-t (*t-loop*) (Griffith et al., 1999; Nikitina & Woodcock, 2004) (**Figure 1b**). L'extrémité 3' sortante est alors appariée avec le brin C et déplace le brin G en une boucle de déplacement (*D-loop*). Cette configuration particulière pourrait participer à la protection des extrémités chromosomiques en masquant la terminaison du télomère.

1.1.2.2. Le complexe télosome ou shelterin

Le complexe protéique spécialisé associé au télomère chez les mammifères s'appelle télosome ou shelterin (**Figure 1c**). Il est formé de six protéines (Liu et al., 2004) : TRF1 (Telomeric Repeat Binding Factor 1) (Chong et al., 1995), TRF2 (Telomeric Repeat Binding Factor 2) (Bilaud et al., 1997), POT1 (Protection Of Telomere 1) (Baumann & Cech, 2001), TIN2 (TRF1 Interacting factor 2) (Ye & de Lange, 2004), TPP1 (TriPeptidyl Peptidase 1) (Ye et al., 2004) et RAP1 (Repressor Activator Protein 1) (Li et al., 2000). Ces protéines peuvent se lier à l'ADN télomérique double brin ou à l'extrémité 3' sortante simple brin.

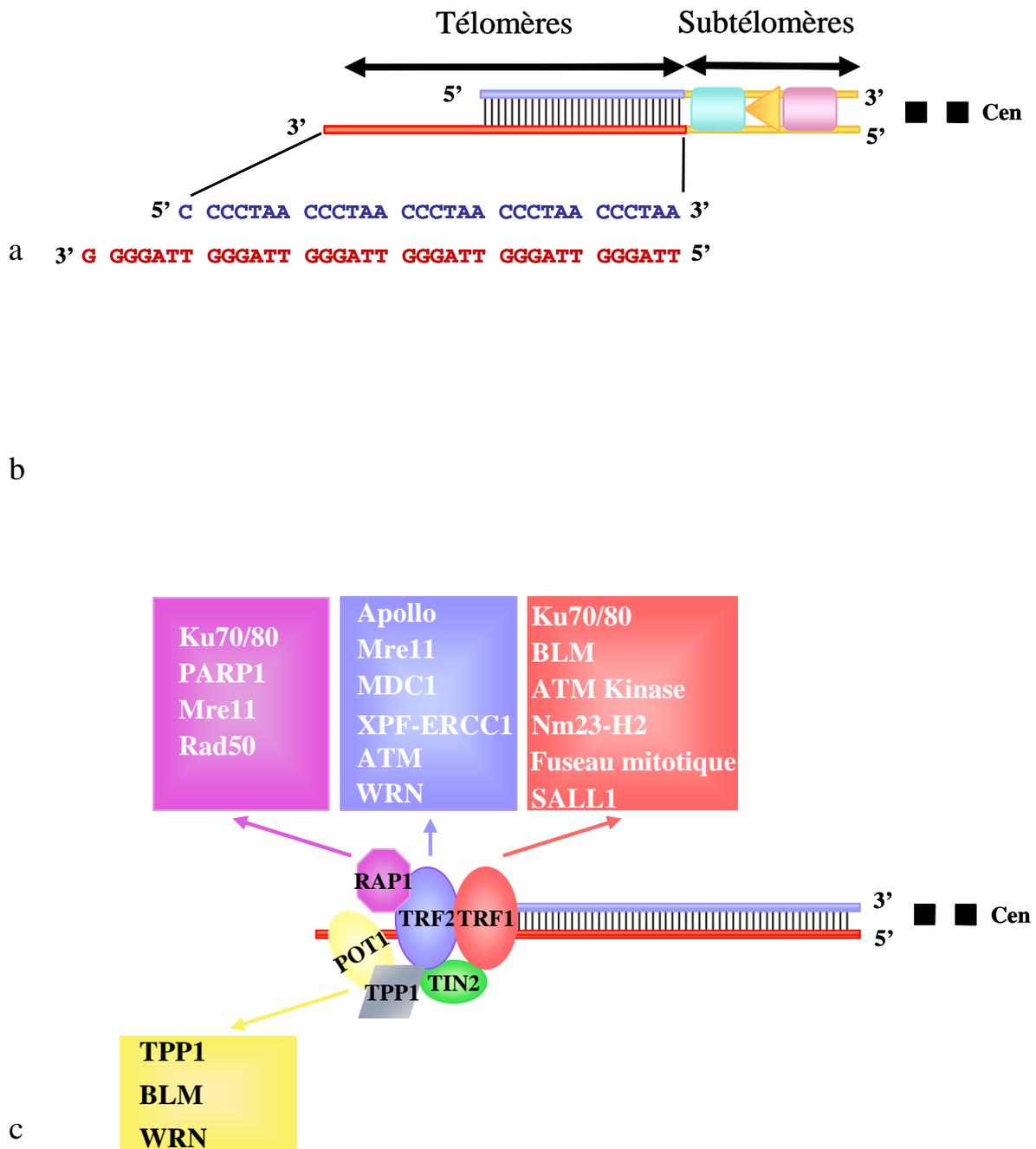


Figure 1 : Structure des télomères

a : Représentation schématique de la structure des extrémités chromosomiques. Les télomères sont séparés des séquences spécifiques des chromosomes par les régions subtélomériques. Le brin C est représenté en bleu et le brin G en rouge, avec l'extrémité 3' sortante. La séquence des répétitions télomériques est indiquée pour les deux brins. Les subtélomères sont composés d'un patchwork de séquences et forment la transition entre les télomères et les régions chromosomiques spécifiques.

(Cen: centromère)

b : Schéma de la boucle-t (*t-loop*) (d'après Palm & de Lange, 2008). L'extrémité 3' sortante envahit la partie double brin pour former la boucle-t (*t-loop*). Elle est alors appariée au brin C et déplace le brin G en une boucle de déplacement (*D-loop*).

c : Représentation schématique du télosome et exemples de facteurs accessoires associés. Le télosome est composé de 6 protéines : TRF1, TRF2, TIN2 et RAP1 associées à l'ADN double brin; TPP1 et POT1 associées à l'ADN simple brin. Des exemples de protéines interagissant avec les composants du télosome sont indiqués dans chaque cadre. La couleur du cadre correspond à la couleur de la protéine associée. (Cen: centromère)

La région double brin est reconnue spécifiquement par TRF1 et TRF2, deux protéines de la même famille qui se distinguent par leur domaine amino-terminal (acide pour TRF1, basique pour TRF2). Elles recrutent respectivement TIN2 et RAP1, TIN2 formant un pont entre TRF1 et TRF2. L'ADN simple brin est reconnu par TPP1 et POT1, TPP1 formant un pont entre TRF1-TRF2-TIN2 et POT1. Les éléments du télosome peuvent former des complexes en absence d'ADN télomérique et des sous-complexes ont également été identifiés dans des extraits cellulaires.

Le télosome intervient dans la formation de la boucle-t, dans la régulation de la longueur des télomères par la télomérase et dans la protection des extrémités télomériques contre les systèmes de signalisation et de réparation des dommages à l'ADN (Palm & de Lange, 2008; Jain & Cooper, 2010). Il interagit avec des facteurs accessoires, qui sont des protéines impliquées dans la signalisation des dommages à l'ADN (protéines ORC, RecQ hélicases ATM et BLM) (Kishi et al., 2001; Opresko et al., 2002), la réparation de l'ADN (Ku70/80, XPF/ERCC1, Apollo, Mre11, RAD50, PARP-1) (Hsu et al., 1999; Hsu et al., 2000; Zhu et al., 2000; Zhu et al., 2003; O'Connor et al., 2004; Lenain et al., 2006), ou la structure de la chromatine (HP1). Par exemple, TRF2 joue un rôle dans la formation de la boucle-t et l'inhibition de la voie ATM (Stansel et al., 2001). TPP1 et POT1 interagissent avec la voie ATR.

Récemment un nouveau complexe de trois protéines associé à la protection des télomères a été mis en évidence, le complexe CST (CTC1, STN1 et TEN1). Ses fonctions doivent encore être précisées (Miyake et al., 2009).

1.1.3. Fonctions des télomères

1.1.3.1. Protection des extrémités chromosomiques

1.1.3.1.1. Réplication des extrémités et télomérase

Du fait de l'incapacité de l'ADN polymérase classique à répliquer complètement l'ADN linéaire, l'extrémité des chromosomes eucaryotes est confrontée à une perte de matériel génétique à chaque cycle de réplication (Watson, 1972) (**Figure 2**).

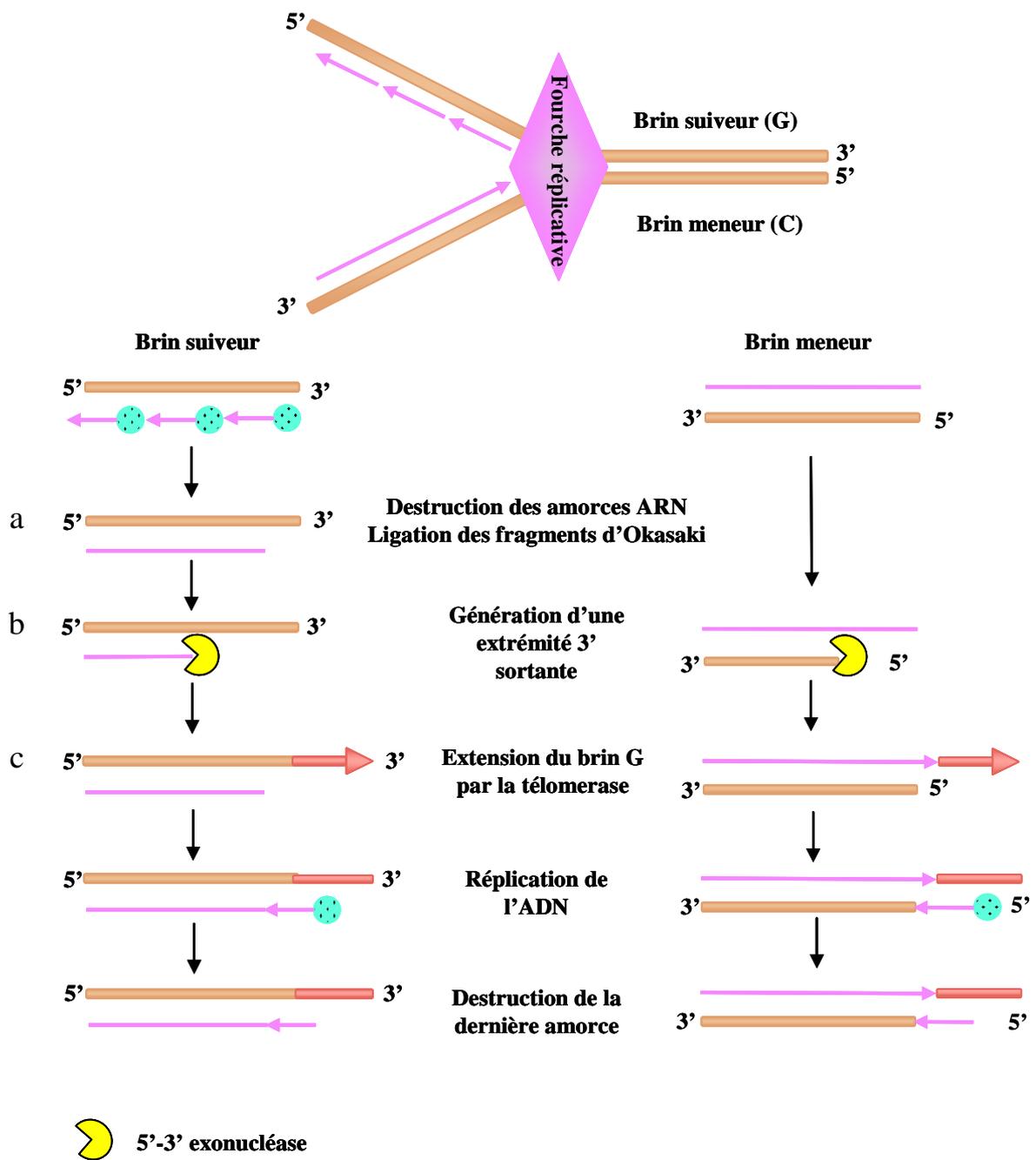


Figure 2 : Réplication des télomères

a : La réplication incomplète du brin suiveur est associée à la perte de nucléotides à chaque cycle.

b : L'extrémité 3' sortante est générée par l'action d'une exonucléase 5'-3'.

c : Dans certaines cellules, la perte nucléotidique secondaire est comblée par l'action de la télomérase.

En effet, lors de la réplication semi-conservative de l'ADN, le brin suiveur ("*lagging strand*") est synthétisé de façon discontinue par l'intermédiaire de fragments d'Okazaki. Chaque fragment d'Okazaki est initié par une amorce ARN qui est ensuite détruite, remplacée par de l'ADN et liguée. En revanche, l'amorce la plus distale ne peut pas être remplacée sur un ADN linéaire, ce qui a pour conséquence une délétion de quelques bases de matériel génétique terminal à chaque cycle réplicatif (Lundblad & Szostak, 1989).

Toutefois chez l'Homme, la perte télomérique à chaque cycle réplicatif est d'environ 100 à 200 pb, c'est-à-dire beaucoup plus importante que la taille d'une amorce ARN (Huffman et al., 2000). Par ailleurs, la réplication du brin meneur ("*leading strand*") est complète mais génère un double brin à bout franc. Or, il a été observé la présence d'un brin 3' sortant à toutes les extrémités télomériques (Makarov et al., 1997). L'ensemble de ces observations a fait émettre l'hypothèse d'un mécanisme post-réplicatif, faisant intervenir une activité nucléase 5' qui génère une nouvelle extrémité 3' sortante par résection du brin C (Sfeir 2005). Ce mécanisme est régulé par les protéines du télosome, notamment POT1 qui contrôle la détermination stricte du nucléotide terminal (Hockemeyer et al., 2005), et expliquerait l'attrition plus rapide des télomères.

Face au problème de réplication des extrémités, la plupart des eucaryotes ont développé un mécanisme, conservé au cours de l'évolution, pour lutter contre l'attrition télomérique : l'enzyme télomérase (Blackburn & Gall, 1978; Harrington et al., 1997; Nakamura et al., 1997). La télomérase est un complexe ribonucléoprotéique composé d'une sous-unité catalytique transcriptase inverse "télomérase" (hTERT), d'un ARN (hTR) (Weinrich et al., 1997) associé à des protéines accessoires qui facilitent l'assemblage et la stabilité du complexe catalytique comme la Dyskérine, NHP2, pontin-reptin, TCAB1 et NOP10 (Cohen et al., 2007; Fu & Collins, 2007; Venteicher et al., 2008; Venteicher et al., 2009) (**Figure 3**). La télomérase utilise la molécule d'ARN hTR comme matrice (AAUCCCAAUC), la complémentarité entre l'ARN hTR et l'ADN télomérique permettant un alignement correct. La télomérase peut ainsi catalyser plusieurs cycles d'alignement-extension permettant d'ajouter des centaines de nucléotides à la même extrémité par transcription inverse. Après l'extension de l'extrémité 3', la machinerie de réplication de l'ADN permet de compléter le brin opposé (Blackburn & Gall, 1978; Osterhage & Friedman, 2009).

Malgré la variabilité individuelle de la longueur des télomères, la longueur télomérique moyenne est maintenue dans un intervalle stable, reflétant un équilibre entre élongation et attrition des télomères. La régulation de l'élongation télomérique reposerait sur

une boucle de rétrocontrôle négatif où l'allongement du télomère s'accompagnerait d'une accumulation de facteurs inhibant la télomérase. Les composants du télosome, dont le nombre de copies associées au télomère augmente avec le nombre de répétitions (TTAGGG)_n, peuvent ainsi agir comme régulateurs en cis de la télomérase (van Steensel & de Lange, 1997; Smogorzewska & de Lange, 2004). Enfin, la régulation de la télomérase s'effectue au niveau transcriptionnel. Son expression est faible ou absente dans la plupart des cellules somatiques mais présente dans les cellules embryonnaires, certaines cellules souches et la plupart des tumeurs (Wright et al., 1996; Meyerson et al., 1997).

Figure 3: Structure de la télomérase (d'après Armanios, 2009)

hTERT: sous-unité catalytique transcriptase inverse

hTR: sous-unité ARN comportant

- une boîte H/ACA assurant sa stabilité et permettant son assemblage à la sous-unité hTERT
- une séquence matrice "*Template*" (AAUCCCAAUC) permettant l'extension des télomères

Protéines accessoires : Dyskerin, NHP2, NOP10, GAR1

1.1.3.1.2. Protection contre les signaux de dommages à l'ADN et les systèmes de réparation de l'ADN

D'un point de vue mécanique, les extrémités chromosomiques correspondent à des cassures d'ADN double brin. Les dommages de l'ADN induisent habituellement une série de réponses cellulaires qui activent les points de contrôle d'arrêt du cycle cellulaire pour permettre la réparation de l'ADN. En cas de dysfonction télomérique, l'activation de ces mécanismes de réparations à l'extrémité des chromosomes peut avoir des conséquences délétères : fusions télomériques, réarrangements chromosomiques et instabilité génomique globale. Les télomères possèdent donc des propriétés de protection spécifiques qui permettent d'échapper non seulement à la réponse aux dommages à l'ADN mais aussi aux mécanismes de réparation.

Chez les mammifères, deux principaux systèmes de réponse aux dommages à l'ADN peuvent être activés aux télomères : la voie ATM qui reconnaît les cassures double brin (Shiloh, 2003) et la voie ATR qui reconnaît l'accumulation d'ADN simple brin et pourrait être activée par le brin 3' sortant. ATM et ATR sont des protéines kinases qui phosphorylent l'histone H2AX et recrutent différents facteurs (MDC1, 53BP1, complexe Mre11) (d'Adda di Fagagna et al., 2003; Takai et al., 2003), entraînant l'apparition de foyers appelés TIFs ("*Telomere dysfunction Induced Foci*"). Par ailleurs, elles activent deux kinases nucléoplasmiques, Chk1 et Chk2, qui bloquent la progression du cycle cellulaire. Enfin, elles permettent l'activation de p53, qui inhibe le cycle cellulaire et induit la mort cellulaire. La voie ATM est inhibée par TRF2 (Karlseder et al., 1999; Celli & de Lange, 2005; Denchi & de Lange, 2007) alors que POT1 inhibe la voie ATR en collaboration avec TPP1 (Denchi & de Lange, 2007; Hockemeyer et al., 2007).

Les deux principaux modes de réparation de l'ADN qui peuvent menacer les télomères sont la réparation par fusion d'extrémités non homologues ("*Non Homologous End-Joining*", NHEJ) et la recombinaison homologue. Le NHEJ est un mécanisme de réparation non exact qui joint les deux extrémités franches d'une cassure double brin (Hartlerode & Scully, 2009). Ce mécanisme peut aboutir à la fusion télomérique de deux chromosomes, créant un chromosome dicentrique qui va initier à la prochaine division cellulaire un cycle de cassure-fusion-pont ("*breakage-fusion-bridge cycle*"), à l'origine d'une instabilité génomique (**Figure 4a**). En effet, à chaque mitose, la tension du fuseau mitotique sur les deux centromères du

chromosome dicentrique ainsi formé crée un pont anaphasique entraînant des cassures aléatoires suivies de nouvelles fusions. Ces cycles de cassures et réparation induisent donc à chaque cycle perte et gain de matériel chromosomique et d'importants remaniements à l'échelle du génome. La protéine TRF2 et à un moindre degré POT1 sont les inhibiteurs du NHEJ (van Steensel et al., 1998; Celli & de Lange, 2005).

La recombinaison homologue est une recombinaison entre deux séquences d'ADN identiques situées sur deux molécules différentes ou distantes l'une de l'autre sur la même molécule. Au niveau télomérique, elle peut avoir lieu à l'intérieur de la boucle-t ("*t-loop homologous recombination*") et entraîne la perte des télomères associée à l'accumulation de molécules télomériques circulaires de la taille d'une boucle-t (Wang et al., 2004) (**Figure 4b**). La recombinaison entre chromatides sœurs télomériques ("*telomere sister chromatid exchange*" T-SCE) survient après la réplication des télomères et peut aboutir à des chromatides de longueurs inégales en cas d'échange inégal (Celli et al., 2006) (**Figure 4c**). Enfin, la recombinaison ectopique entre un télomère et une séquence télomérique interne entraîne une délétion chromosomique de taille importante et la présence de matériel extra-chromosomique double-minute contenant de l'ADN télomérique (Zhu et al., 2003) (**Figure 4d**). Les protéines TRF2, Ku70, ERCC1, POT1, RAP1 et l'hélicase WRN participent à l'inhibition de ces différents mécanismes de recombinaison homologue (Zhu et al., 2003; Wang et al., 2004; Laud et al., 2005; Celli et al., 2006; Sfeir et al., 2010). Toutefois, le mécanisme de recombinaison homologue peut également permettre de réguler la taille des télomères, soit trop longs (processus de délétion rapide des télomères), soit trop courts (élongation alternative des télomères ou ALT, § 1.2.3.3.) (Jain & Cooper, 2010).

Figure 4 : Mécanismes de réparation mis en jeu aux extrémités télomériques

Le brin G est représenté en rouge et le brin C en vert (d'après Palm & de Lange, 2008)

a : Réparation par fusion d'extrémités non-homologues (NHEJ). La fusion de deux télomères aboutit à la formation d'un chromosome dicentrique et initie un cycle de cassure-fusion-pont. (télomère : carré jaune ; centromère : rond noir)

b : Recombinaison homologue dans la boucle t. La migration de l'extrémité télomérique à l'intérieur de la boucle-t induit la formation d'une jonction de Holliday (HJ). La résolution de cette dernière entraîne une perte télomérique et la présence de cercle télomérique

c : Echange de chromatides sœurs télomériques (T-SCE). Un échange inégal entre deux chromatides sœurs télomériques entraîne la ségrégation d'un chromosome au télomère allongé dans une cellule-fille et d'un chromosome au télomère raccourci dans l'autre cellule-fille.

d : Recombinaison ectopique avec un site interstitiel. L'échange inégal induit une perte de matériel chromosomique et la présence de chromosome double-minute.

1.1.3.1.3. Conséquences : Sénescence, cancer et vieillissement

Des dysfonctions des télomères et de la télomérase ont été décrites dans de nombreuses pathologies tant constitutionnelles qu'acquises.

D'une part, l'activation anormale de la télomérase peut être à l'origine de pathologies tumorales. Du fait de l'absence de télomérase dans les cellules somatiques, les télomères raccourcissent à chaque cycle cellulaire, jusqu'à atteindre une taille critique où le complexe télosome ne serait plus en quantité suffisante pour inhiber la réponse aux dommages à l'ADN (Allsopp et al., 1992; Palm & de Lange, 2008). Les cellules entrent alors en sénescence, une forme d'arrêt permanent du cycle cellulaire (**Figure 5**). Ce mécanisme se comporte comme une barrière anti-cancer en limitant la prolifération des cellules somatiques ayant potentiellement acquis des anomalies génomiques (d'Adda di Fagagna et al., 2003). Ce point d'entrée en sénescence correspond à la limite de Hayflick, qui définit le nombre maximal de divisions d'une cellule somatique primaire en culture. Toutefois, certaines cellules présentant une altération de la voie de contrôle du cycle cellulaire dépendant de p53 peuvent franchir cette limite et entreprendre 20 à 30 divisions supplémentaires, au cours desquelles les télomères continuent à raccourcir, entraînant l'apparition de réarrangements chromosomiques secondaires à des fusions télomériques et des cycles cassure-fusion-pont, par exemple, et générant une instabilité génomique (Hara et al., 1991). Ces cellules parviennent alors au stade de crise, caractérisé par une importante instabilité génomique et l'entrée en apoptose. Mais là encore, quelques rares clones peuvent émerger de la crise et activer un mécanisme d'élongation des télomères, qui permet de stabiliser les chromosomes. Dans 85 à 90% des cas, la télomérase est réactivée (Kim et al., 1994). Mais dans 10% des cas, un mécanisme alternatif appelé ALT ("*Alternative Lengthening of Telomeres*") (§ 1.2.3.3.) est mis en œuvre (Bryan et al., 1997). Ces clones sont à l'origine de la prolifération tumorale.

D'autre part, des mutations constitutives entraînant une perte de fonction de la télomérase sont à l'origine de pathologies constitutionnelles, regroupées sous le nom de syndromes de raccourcissement télomérique. Les conséquences se manifestent tout d'abord dans les tissus au renouvellement rapide, où les capacités répliquatives des cellules souches sont affectées. Les principaux signes caractéristiques sont un vieillissement prématuré (cheveux gris, dysplasie unguéale), une fibrose pulmonaire et/ou hépatique, des dysfonctions immunitaires et hématologiques, une augmentation du risque de cancer, une sensibilité accrue à la chimiothérapie et à la radiothérapie. La pathologie la plus anciennement connue est la

dyskératose congénitale (MIM #127550) secondaire à des mutations des sous-unités de la télomérase ou des protéines associées (*hTERT*, *hTR*, *DKC1*, *TINF2*, *NHP1*, *NOP10*). Plus récemment, le syndrome de Hoyeraal-Hreidarsson (MIM#300240) et des pathologies débutant chez l'adulte comme la fibrose pulmonaire idiopathique ou l'anémie aplasique ont été rattachées à des anomalies de la télomérase (Armanios, 2009).

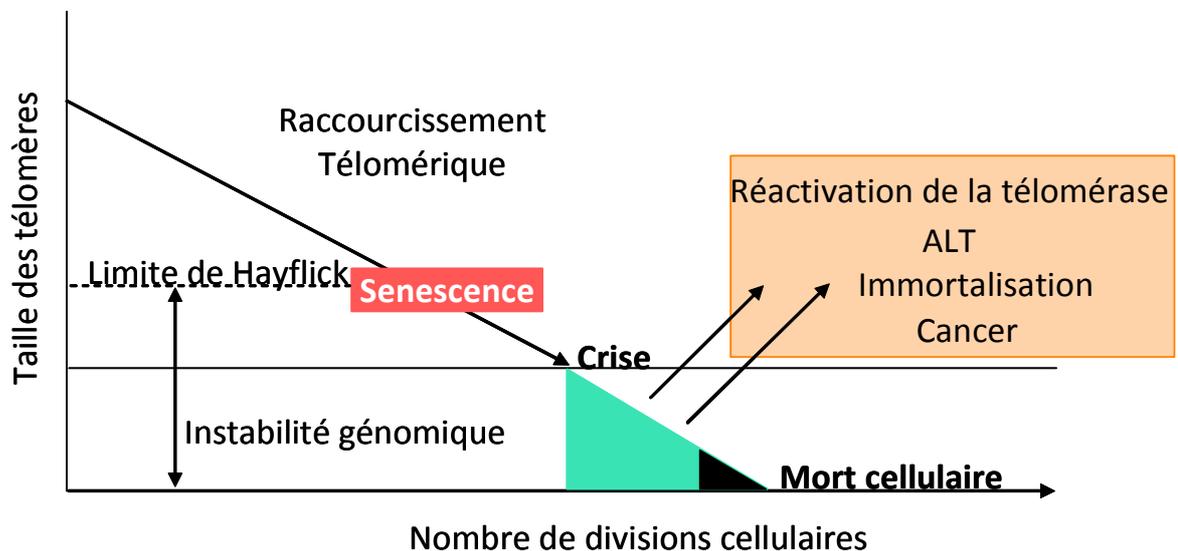


Figure 5 : L'attrition des télomères contrôle la capacité proliférative des cellules somatiques

1.1.3.2. Télomères et reproduction

Les télomères sont des structures indispensables à la protection des chromosomes linéaires des eucaryotes. Néanmoins, leur érosion à chaque division cellulaire peut conduire à la sénescence, au vieillissement ou au cancer en l'absence d'un rallongement par la télomérase. Ainsi, d'un point de vue évolutif, posséder et maintenir intacts des chromosomes linéaires nécessitent une énergie supplémentaire. Quel avantage peuvent-ils alors conférer? Les eucaryotes se distinguent des procaryotes par la reproduction sexuée qui permet un brassage de l'information génétique. Ce dernier est assuré au cours de la méiose, un processus qui produit quatre cellules filles haploïdes à partir d'une cellule mère diploïde. Alors qu'un chromosome circulaire, et a fortiori plusieurs, ont peu de chance de ségréger correctement au cours de la méiose, les chromosomes linéaires permettent un déroulement harmonieux de ce processus (Ishikawa & Naito, 1999).

La première division méiotique aboutit à la séparation des deux chromosomes homologues dans deux cellules filles (division réductionnelle). Cette étape nécessite l'attachement des télomères à la membrane nucléaire interne pendant la prophase I, qui conduit au stade de bouquet où les télomères confluent vers le centrosome. Cette conformation favorise l'alignement et l'appariement des chromosomes homologues et l'échange de matériel par recombinaison homologue, indispensable à la ségrégation correcte des chromosomes. Par ailleurs, la formation en bouquet permettrait de contrôler le fuseau méiotique (Tomita & Cooper, 2007).

Des télomères trop courts sont incapables d'accrocher les chromosomes à l'enveloppe nucléaire. Il a d'ailleurs été montré que la télomérase est active dans les cellules germinales et que les télomères sont allongés dans les cellules germinales mâles et femelles (Siderakis & Tarsounas, 2007). Par ailleurs, l'étude de générations successives de souris invalidées pour la télomérase (TR^{-/-}) a révélé un fort taux d'apoptose des spermatocytes chez les mâles et l'apparition d'ovocytes aberrants chez les femelles qui, une fois fécondés, aboutissent à des portées plus petites et des embryons porteurs d'anomalies du développement (Blasco et al., 1997; Lee et al., 1998; Herrera et al., 1999; Liu et al., 2002). En conclusion, le contrôle de l'intégrité des télomères dans les cellules germinales pourrait contribuer à assurer la fertilité.

1.1.3.3. Effet de position télomérique

1.1.3.3.1. Rappels sur l'effet de position chromosomique

L'effet de position décrit une modification de l'expression d'un gène par une altération de son environnement chromosomique. Il peut concerner plus largement la recombinaison ou la cinétique de réplication. Il existe deux types d'effet de position chromosomique (CPE), l'effet de position stable où la multimérisation d'une séquence induit le silence transcriptionnel (Sturtevant, 1925), et l'effet de position par variégation (PEV, "*Position Effect Variegation*") (Muller, 1930, 1932). Dans ce dernier cas, des réarrangements chromosomiques placent des gènes euchromatiniens dans des régions hétérochromatiniennes, comme les régions péri-centromériques par exemple, et induisent l'inactivation de ces gènes par diffusion de l'hétérochromatine à partir du point de cassure (**Figures 6a et 6b**). Ce silence transcriptionnel est alors héréditaire, inversement proportionnel à la distance par rapport au point de cassure et stochastique car la diffusion de l'hétérochromatine varie d'une cellule à l'autre, ce qui peut se manifester par un phénotype en mosaïque à l'échelle d'un tissu.

L'effet de position est un mécanisme complexe, faisant appel à des composants de l'hétérochromatine, des enzymes de modification des histones, des protéines non histones et des protéines de l'architecture nucléaire. La cellule a également développé des processus pour s'en protéger.

En effet, des régions de chromatine active et inactive sont souvent juxtaposées et leur identité respective est parfois maintenue par des éléments spécialisés tels que les insulateurs. Ces éléments sont définis comme une classe de séquences d'ADN qui possèdent la propriété commune de protéger les gènes de signaux inappropriés provenant de leur environnement proche (West et al., 2002; Gaszner & Felsenfeld, 2006). On en distingue classiquement deux types, ceux de type "*enhancer-blocking*" et ceux de type barrière de chromatine. Ces deux activités peuvent co-exister dans certains insulateurs, mais sont souvent découplées. Les insulateurs "*enhancer-blocking*" sont capables d'interférer avec les interactions *enhancer*-promoteur quand ils sont placés entre les deux. Les insulateurs barrière de chromatine protègent contre l'effet de position en empêchant la propagation de l'hétérochromatine.

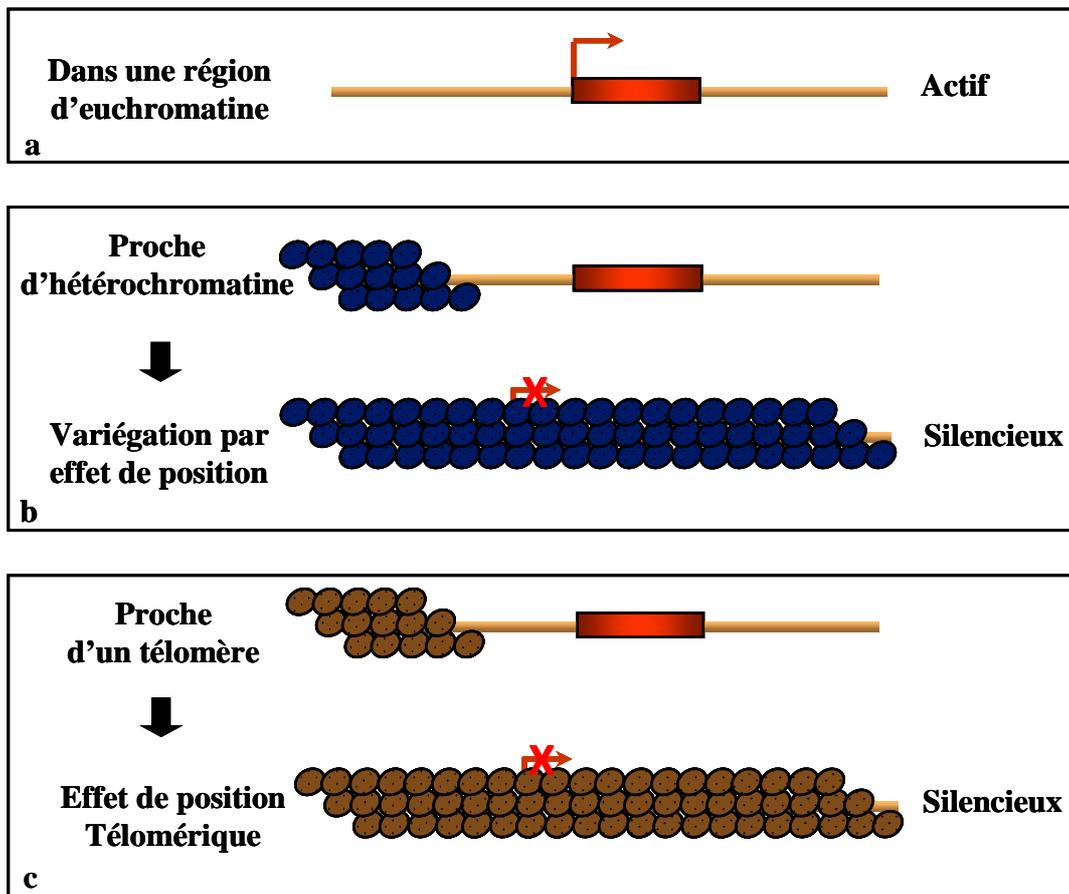


Figure 6: Les différents types d'effet de position

- a : Expression d'un gène dans une région euchromatinienne
- b : Répression par effet de position chromosomique
- c : Répression par effet de position télomérique

Chez les vertébrés, tous les insulateurs lient la protéine CTCF (CCCTC binding factor). Cette protéine, hautement conservée chez les eucaryotes, est caractérisée par la présence d'un domaine central de liaison à l'ADN composé de 11 doigts de zinc (Phillips & Corces, 2009). De 14 000 à 29 000 sites CTCF ont été identifiés dans le génome humain dont près de la moitié sont en position intergénique. Les régions contenant des promoteurs alternatifs sont enrichies en sites CTCF. Quatre-vingt pourcent des sites CTCF possèdent un motif consensus de 11 à 15 pb mais une séquence de 50 à 60 pb est nécessaire pour une liaison plus stable. CTCF est capable de reconnaître une vaste variété de séquences, peut établir différentes interactions grâce à différentes combinaisons de ses doigts de zinc et est susceptible de subir différentes modifications post-traductionnelles. Ainsi CTCF se comporte comme un facteur aux fonctions multiples. Il a été décrit comme participant non seulement aux fonctions d'insulation "*enhancer blocking*"/ barrière de chromatine mais aussi d'activation/ répression transcriptionnelle et d'interactions chromatiniennes à longue distance. CTCF a notamment été impliqué dans l'inactivation du chromosome X, dans l'empreinte de la région *IGF2/ H19* en 11p15 (Kurukuti et al., 2006) et dans le fonctionnement de l'insulateur 5'HS4 du locus β -globine du poulet (Bell et al., 1999). CTCF agirait notamment par la formation de boucles chromatiniennes et jouerait un rôle dans l'organisation de l'architecture chromatinienne.

1.1.3.3.2. L'effet de position télomérique

L'effet de position télomérique (TPE, "*Telomeric Position Effect*") est défini comme la répression transcriptionnelle réversible de gènes situés à proximité des télomères (**Figure 6c**). Bien que décrit initialement chez la drosophile (Gehring et al., 1984), le TPE a été extensivement étudié chez la levure *Saccharomyces cerevisiae*. Le TPE a été mis en évidence dans un modèle de levure où le gène *URA3* a été introduit à 1,3 kb du télomère du bras gauche du chromosome VII (VIII) (Gottschling et al., 1990). Cette souche de levure était non seulement capable de croître sur des boîtes déficientes en uracile mais montrait aussi une résistance à l'acide 5-fluoro-orotique (5-FOA) (33% des colonies), une drogue convertie en substance toxique par *URA3*, suggérant que le taux d'expression du gène *URA3* était réprimé sous le seuil de toxicité du 5-FOA. Le TPE est réversible, stochastique, dépend de la distance par rapport au télomère et implique des modifications des queues amino-terminales des

histones. Les protéines impliquées dans le TPE sont différentes de celles du PEV (Doheny et al., 2008). Bien que plusieurs protéines puissent moduler le TPE chez la levure, seules quelques-unes sont indispensables comme le complexe Sir (Silent Information Regulators Sir 2p, 3p, 4p), l'hétérodimère Ku (yKu70p, yKu80p) et le domaine C terminal de Rap1 (Aparicio et al., 1991; Kyrion et al., 1992; Boulton & Jackson, 1998; Ottaviani et al., 2008) qui forment un complexe répressif se propageant vers le subtélocère. Enfin, l'élongation des télomères renforce le TPE, probablement de façon secondaire au recrutement concomitant de facteurs répressifs. Par ailleurs, l'étude du TPE dans un contexte télomérique naturel a montré que sa force était extrêmement variable, et était notamment influencée par le type de séquences subtélomériques (Pryde & Louis, 1999).

Chez l'Homme, l'existence du TPE a longtemps été controversée (Bayne et al., 1994; Sprung et al., 1996). Un premier indice a été apporté par l'étude d'une délétion terminale du bras long du chromosome 22 (22qter) réparée par l'ajout de répétitions télomériques *de novo* ("telomere healing") (Ofir et al., 1999). Elle a mis en évidence une asynchronie de réplication entre les extrémités des deux chromosomes 22, avec une réplication plus tardive du chromosome délété, mais sans effet sur l'expression des gènes situés à 50 kb du télomère. Des modèles cellulaires ont ensuite permis de démontrer l'existence du TPE (Baur et al., 2001; Koering et al., 2002). Dans ces modèles, un gène rapporteur (luciférase, *eGFP*) est inséré à proximité d'un télomère néoformé par le processus de "fragmentation télomérique" (Farr et al., 1992) dans des lignées de cellules cancéreuses humaines (Hela, C33-A). Le plasmide linéaire contient un gène rapporteur adjacent à des répétitions télomériques (TTAGGG) de 1,6 kb. L'intégration du plasmide induit une cassure chromosomique suivie de l'extension de la séquence télomérique par la télomérase et la perte du fragment chromosomique distal natif. Les deux études ont mis en évidence une diminution stable de l'expression du gène rapporteur en localisation télomérique par rapport à une insertion chromosomique aléatoire.

Les mécanismes mis en œuvre dans le TPE chez l'Homme sont encore mal connus. Dans les modèles cellulaires, le traitement par un inhibiteur des histones désacétylases (HDAC), la trichostatine A (TSA) est responsable d'une augmentation de l'expression du gène rapporteur, associée à un déplacement des protéines d'hétérochromatine HP1 (Baur et al., 2001; Koering et al., 2002) suggérant un rôle des HDAC et de HP1. L'élongation des télomères par la surexpression de hTERT s'accompagne d'un renforcement du TPE (Baur et al., 2001). Cependant, ce résultat est discuté (Sprung et al., 1996; Ning et al., 2003). D'autres arguments seraient plus en faveur de l'implication d'une modification de l'environnement chromatinien télomérique. Une surexpression de TRF1 entraîne une augmentation transitoire

de l'expression du gène rapporteur, sans modifier la taille des télomères, probablement par dérégulation de la chromatine. Enfin, le TPE est dépendant de la distance par rapport au télomère. Il est significativement diminué au-delà de 100 kb du télomère dans un modèle de cellules de carcinome vésical (Kulkarni et al., 2010).

Plus récemment, un modèle murin a été développé dans des cellules embryonnaires souches (ES) puis des souris transgéniques issues de ces cellules ES (Pedram et al., 2006; Gao et al., 2007). L'expression du gène rapporteur est partiellement réprimée dans les cellules ES, mais complètement éteinte dans les tissus somatiques et les fibroblastes (MEF, "mouse embryonic fibroblasts") isolés à partir des souris transgéniques. Ce patron d'expression est corrélé à une forte méthylation du gène rapporteur dans les tissus somatiques, alors que la méthylation est faible dans les cellules ES. Contrairement à ce qui est observé chez l'Homme, la TSA n'a aucun effet sur le TPE, mais la 5-azacytidine, un agent déméthylant, peut lever le TPE chez la souris. Ainsi, il a été suggéré que la mise en place du TPE pourrait se dérouler en deux temps avec une méthylation secondaire, venant verrouiller la chromatine comme ce qui est observé lors du processus d'inactivation du chromosome X.

Les rôles du TPE ne sont pas clairement définis. Il pourrait n'être qu'un effet secondaire lié à la conformation particulière de la chromatine des télomères. Il pourrait jouer un rôle dans la sénescence et le vieillissement, en modifiant l'expression des gènes subtélomériques au fur et à mesure du raccourcissement télomérique (Hernandez-Caballero et al., 2009). Enfin, il pourrait intervenir dans la régulation de l'expression des gènes au cours du développement embryonnaire précoce (Gao et al., 2007).

1.1.4. Régulation épigénétique des télomères

Les modifications épigénétiques désignent l'ensemble des modifications héritables du phénotype secondaires à un mécanisme indépendant de la séquence primaire d'ADN. Les plus connues sont la méthylation de l'ADN et les modifications de la chromatine mais d'autres mécanismes comme, par exemple, les ARNs non codants ou encore la compartimentalisation nucléaire en font partie. Ces modifications jouent un rôle majeur dans la régulation des télomères.

1.1.4.1. La chromatine télomérique

La chromatine est composée de l'association d'ADN et de protéines structurales, les histones. Le nucléosome en est l'unité de base. Il est composé de 146 pb d'ADN enroulées autour d'un octamère d'histones composé de deux sous-unités de H3, H4, H2A et H2B. Les nucléosomes sont séparés par 20 à 50 pb d'ADN internucléosomal ("*linker*") qui peut être lié par l'histone H1 qui induit un degré de compaction supplémentaire. L'extrémité aminoterminal des histones dépasse du nucléosome et peut être l'objet de modifications covalentes post-traductionnelles (phosphorylation, méthylation, acétylation, ubiquitination, sumoylation...). Ces modifications permettent de recruter des effecteurs qui interviennent dans différentes voies biologiques. Ces modifications déterminent le code histone (Strahl & Allis, 2000). La chromatine est organisée en deux états distincts, l'euchromatine et l'hétérochromatine. Chaque état est caractérisé par son niveau de compaction, le code histone, la présence de variants d'histones, les modifications de l'ADN, l'association de protéines non histones, la réplication et sa localisation subnucléaire. L'euchromatine correspond à des régions d'ADN peu condensées, transcriptionnellement actives et de réplication précoce en phase S. L'hétérochromatine correspond à des régions compactées, silencieuses et de réplication tardive. L'hétérochromatine constitutive définit des régions toujours condensées et inactives, composées de séquences d'ADN répétées comme les régions péri-centromériques alors que l'hétérochromatine facultative désigne des régions qui peuvent être tantôt actives, tantôt inactives comme le chromosome X.

Bien que proche de l'hétérochromatine constitutive, la structure hétérochromatinienne des extrémités chromosomiques est particulière et certaines des caractéristiques de l'hétérochromatine constitutive, comme l'aspect condensé, ne sont pas retrouvées au niveau des télomères.

La taille de l'ADN internucléosomal des télomères est de 10 à 20 pb, soit de 20 à 40 pb inférieure à la moyenne (Tommerup et al., 1994; Pisano et al., 2008). Par ailleurs, l'histone H1 est moins présente au niveau des régions télomériques (Parseghian et al., 2001) et l'ADN télomérique possède une moindre affinité pour l'octamère d'histones (Filesi et al., 2000). Cette structure serait à l'origine d'une moindre compaction de la chromatine et permettrait l'accessibilité à d'autres complexes protéiques, notamment le télosome et les protéines associées. D'ailleurs, les protéines du télosome permettraient de réguler cette organisation. Une surexpression de TRF2 entraîne une diminution globale du nombre de nucléosome et une augmentation de l'espace internucléosomal (Benetti et al., 2008). De même, TRF1 pourrait

modifier la structure nucléosomale et induire une compaction de l'ADN télomérique (Pisano et al., 2010).

Bien que l'ADN télomérique ne puisse pas être méthylé du fait de l'absence de dinucléotide CpG, la chromatine télomérique porte des marques d'hétérochromatine caractérisées par un enrichissement en résidus lysine 9 de l'histone H3 (H3K9) di et triméthylés (Garcia-Cao et al., 2004), en résidus lysine 20 de l'histone H4 (H4K20) triméthylés (Gonzalo et al., 2005), en protéine HP1 (Koering et al., 2002; Garcia-Cao et al., 2004) et un appauvrissement en résidus d'histones H3 et H4 acétylés (Benetti et al., 2007). Ces modifications sont mises en place par des enzymes spécifiques. Les histones méthyltransférases SUV39-H1 et SUV39-H2 triméthylent les résidus H3K9. Ceux-ci permettent de recruter les protéines HP1. Les protéines HP1 permettent à leur tour le recrutement des histones méthyltransférases SUV4-20H1 et SUV4-20H2 qui sont responsables de la triméthylation de H4K20. Les protéines de la famille rétinoblastome (RB), connues pour leur rôle suppresseur de tumeur, permettent également de stabiliser les histones méthyltransférases SUV4-20H1 et SUV4-20H2 au télomère (Schotta et al., 2004). Il faut noter toutefois que ces enzymes agissent également sur la chromatine extra-télomérique. Plus récemment, une histone désacétylase de la famille des sirtuines, SIRT6, a été identifiée. Elle s'associe de façon spécifique avec les télomères et permet la désacétylation des résidus H3K9 et H3K56 (lysine 56 de l'histone H3) (Mischishita 2008). SIRT6 contribue à l'association de l'hélicase WRN au télomère, une protéine nécessaire à la réplication télomérique et dont les mutations sont responsables du syndrome de Werner (MIM #277700), caractérisé par un vieillissement prématuré et une prédisposition aux cancers (Tennen & Chua, 2011). D'ailleurs, les souris déficientes en *Sirt6* présentent une pathologie de vieillissement prématuré caractérisée par un décès précoce associé à une instabilité génomique et une sensibilité aux radiations (Mostoslavsky et al., 2006).

Des travaux chez la souris suggèrent que la chromatine télomérique interviendrait dans le contrôle de la taille des télomères. Dans des modèles murins, l'hypométhylation des histones est associée à des télomères longs. En effet, l'inactivation de SUV39-H1 et SUV39-H2 s'accompagne d'une diminution de la triméthylation des résidus H3K9, d'une diminution de HP1 et d'une élévation anormale des télomères. De même, l'inactivation des protéines de la famille RB s'accompagne d'une perte de triméthylation des résidus H4K20 et de télomères anormalement longs (Garcia-Cao et al., 2004; Gonzalo et al., 2005). Réciproquement, dans cette espèce, la longueur des télomères peut influencer la chromatine télomérique. Dans un modèle de souris déficientes en télomérase (*Terc*^{-/-}), Benetti et collaborateurs ont observé des

télomères raccourcis, l'apparition de marques d'euchromatine associée à une diminution de la triméthylation des résidus H3K9 et H4K20, une diminution de HP1 et une augmentation de l'acétylation de H3 et H4 (Benetti et al., 2007). Le relâchement de la chromatine concomitant au raccourcissement des télomères pourrait expliquer la réactivation de l'expression de gènes subtélomériques au cours du vieillissement et serait un effecteur de la régulation du TPE par la longueur des télomères.

En conclusion, les modifications de la chromatine télomérique et l'allongement des télomères sont étroitement corégulés.

1.1.4.2. Les transcrits télomériques ou TERRAs

Du fait de leur structure hétérochromatinienne, les télomères ont longtemps été considérés comme transcriptionnellement silencieux, jusqu'à la découverte récente des transcrits télomériques ou TERRAs ("*Telomeric repeat containing RNAs*"). Ces ARNs ont été décrits chez les mammifères (Homme, rat, souris) (Azzalin et al., 2007; Ho et al., 2008; Schoeftner & Blasco, 2008), ainsi que chez la levure (*Saccharomyces cerevisiae*) (Luke et al., 2008), le poisson-zèbre (*Danio rerio*) (Schoeftner & Blasco, 2008), les oiseaux (Solovei et al., 1994) et le trypanosome (Rudenko & Van der Ploeg, 1989). Leurs caractéristiques commencent à peine à être étudiées et de nombreux aspects de leur biologie, synthèse, régulation et fonctions, ne sont pas encore élucidés.

Chez l'Homme, les TERRAs sont de grands ARNs non codants de localisation nucléaire. Ils sont transcrits à partir du brin C, des régions subtélomériques vers les télomères (Azzalin et al., 2007; Schoeftner & Blasco, 2008) (**Figure 7**). Ils forment un ensemble de molécules de taille hétérogène, allant de 100 bases à 9 kb. La taille moyenne des répétitions télomériques (UUAGGG)_n contenues dans les TERRAs est de 200 bases (100 à 400 bases). La variabilité de taille proviendrait donc de la transcription à travers les régions subtélomériques (Porro et al., 2010).

La transcription des TERRAs est assurée par l'ARN polymérase II (Azzalin et al., 2007; Schoeftner & Blasco, 2008), bien que la participation d'autres polymérases ne soit pas exclue. Elle est initiée dans les régions subtélomériques, au niveau d'îlots de CpGs conservés sur différents subtélomères (Nergadze et al., 2009). L'état de méthylation de ces îlots est corrélé au niveau d'expression des TERRAs. En effet, le double *knock-out* (KO) des gènes codant pour les ADN-méthyltransférases DNMT1 et DNMT3b ainsi que le traitement des

cellules par la 5-azacytidine s'accompagnent d'une perte de méthylation et d'une augmentation des TERRAs (Nergadze et al., 2009). De même, les patients atteints du syndrome ICF (Immunodéficience, instabilité Centromérique, anomalies Faciales, MIM #242860), secondaire à une mutation du gène *DNMT3b* entraînant une déméthylation globale du génome, présentent un niveau élevé de TERRAs (Yehezkel et al., 2008; Deng et al., 2010). Ainsi, la méthylation de l'ADN des régions subtéломériques serait un élément de régulation des TERRAs.

Figure 7: Les transcrits télomériques (d'après Feuerhahn et al., 2010)

Les ARNs TERRAs sont de grands ARNs non-codants, dont la transcription est initiée au niveau d'îlots de CpGs dans les régions subtéломériques et s'étend à travers les répétitions télomériques. Les TERRAs possèdent une coiffe 7-méthylguanosine (m^7G) et 7 à 10 % d'entre eux sont polyadénylés.

Les TERRAs sont exclusivement nucléaires. Ils sont capables de former des foyers (Azzalin et al., 2007; Schoeftner & Blasco, 2008) qui colocalisent avec les télomères dans les noyaux interphasiques et sont présents à l'extrémité des chromosomes métaphasiques (Azzalin et al., 2007). Tous les TERRAs possèdent une coiffe 7-méthylguanosine (Porro et al., 2010). Sept à dix pourcent d'entre eux sont polyadénylés (poly-A+) (Azzalin et al., 2007; Porro et al., 2010), ce qui définit deux sous-populations aux propriétés biologiques différentes. Les TERRAs poly-A+ sont plus stables, avec une demi-vie de 8h, se terminent soit en UUA-3' soit en GGU-3' et sont essentiellement nucléoplasmiques. Les TERRAs poly-A- ont une demi-vie plus courte de 3h et se terminent en UUAGG-3'. Soixante pourcent des TERRAs poly-A- sont nucléoplasmiques et 40% sont associés à la chromatine, suggérant un rôle spécifique des TERRAs poly-A- dans l'assemblage ou la régulation de la chromatine télomérique (Azzalin et al., 2007; Schoeftner & Blasco, 2008; Porro et al., 2010).

Peu d'éléments sont connus concernant la régulation des transcrits télomériques. Chez la levure, les TERRAs sont dégradés par l'enzyme Rat1p, une exonucléase 5'-3' (Luke et al., 2008). A noter que la présence de transcrits anti-sens nommés "ARRET" a été mise en évidence dans cette espèce. Chez l'Homme, l'expression des TERRAs est régulée au cours du cycle cellulaire avec le niveau le plus élevé en phase G1 précoce et le plus bas en phase S tardive (Porro et al., 2010). Les protéines SMG ("*Suppressors with Morphogenetic defects in Genitalia*") UPF1, EST1A et SMG1, impliquées dans le mécanisme de dégradation des ARNs non-sens ("*non-sense mediated RNA decay*", NMD), sont capables de déplacer les TERRAs des télomères. Une invalidation des gènes codant pour ces protéines s'accompagne d'une augmentation du nombre de foyers TERRAs sans augmentation apparente ni de leur demi-vie ni de leur taux total (Azzalin et al., 2007).

Les TERRAs pourraient intervenir dans de nombreux processus comme le maintien de la chromatine télomérique, la régulation de la télomérase et la différenciation cellulaire. Tout d'abord, les TERRAs sont importants pour stabiliser les télomères puisque leur déplétion par interférence à l'ARN (RNAi) entraîne une augmentation des marques de dysfonctions télomériques (TIFs et anomalies télomériques métaphasiques), une modification de la chromatine télomérique (diminution de la diméthylation des résidus H3K9) et une diminution de la prolifération cellulaire (Deng et al., 2009). Les TERRAs sont d'ailleurs associés à diverses protéines du télomère comme TRF1, TRF2, la dyskerine, mais aussi des protéines comme ORC1, ORC2, ORC4, EBNA1, MECP2 et des protéines de réponse aux dommages à l'ADN (Deng et al., 2009). L'interaction entre TRF1 et TERRAs permettrait de réguler leur

transcription par l'ARN polymérase II (Schoeftner & Blasco, 2008; Deng et al., 2009). La surexpression de TRF2 s'accompagne d'une diminution des TERRAs (Benetti et al., 2008). Les TERRAs permettraient de maintenir l'hétérochromatine télomérique, notamment la triméthylation des résidus H3K9 et l'association de HP1, par l'intermédiaire d'une interaction avec les protéines TRF2 et ORC1. Par ailleurs, une augmentation des TERRAs a été mise en évidence dans des cellules déficitaires en histone méthyltransférases Suv39h et Suv4-20h (Schoeftner & Blasco, 2008). La protéine MLL, une histone méthyltransférase (résidu H3K4), associée entre autre à l'hétérochromatine télomérique, permettrait de réguler l'expression des TERRAs (Caslini et al., 2009). En effet, la diminution de MLL, soit par invalidation (siRNA) soit secondaire au raccourcissement télomérique au cours du vieillissement cellulaire, s'accompagne d'une diminution de la méthylation des résidus H3K4 télomériques et d'une diminution de la transcription des TERRAs. Cette voie pourrait être impliquée dans la réponse cellulaire aux dommages de l'ADN mise en jeu lors du raccourcissement télomérique (Caslini et al., 2009).

Ensuite, il a été constaté que les TERRAs sont plus abondants sur les télomères longs (Schoeftner & Blasco, 2008) et qu'ils pourraient participer à la régulation de la taille des télomères. L'association des TERRAs avec la télomérase a d'ailleurs été démontrée *in vitro* (Schoeftner & Blasco, 2008) et *in vivo* (Redon et al., 2010). Les TERRAs sont non seulement capables de se lier à la composante ARN hTR, en agissant comme compétiteur, mais aussi sur la composante transcriptase inverse hTERT, et ce de façon indépendante de la queue poly-A. L'activité télomérase est inhibée par les TERRAs *in vitro*. Les TERRAs pourraient donc participer à une boucle de rétrocontrôle négatif où ils inhiberaient la télomérase sur les télomères les plus longs. Cette activité serait liée à la capacité des TERRAs à former des structures G-quadruplex et des hybrides ARN-ADN (Xu et al., 2010). De plus, il a été suggéré que les TERRAs participeraient à la régulation de la réplication des télomères et à la formation du complexe télosome (Flynn et al., 2011).

Enfin, les TERRAs sont capables de former des foyers au niveau des bras longs des gonosomes Xq/Yq. Dans les cellules ES, murines et humaines, deux foyers de TERRAs sont présents. Au cours de la différenciation, un seul foyer est visible, soit sur le bras long du chromosome Y (Yq) (cellules XY) soit sur le bras long du chromosome X (Xq) inactif (cellules XX) (Zhang et al., 2009). La même étude a mis en évidence une augmentation du taux de TERRAs dans différentes situations, comme les anomalies du comptage des chromosomes X, l'irradiation des cellules ou l'absence de Dicer (machinerie de RNAi)

(Zhang et al., 2009), ouvrant des pistes de recherche sur le rôle des TERRAs au cours de la différenciation cellulaire, de la réponse au stress et de la voie d'interférence à l'ARN.

1.1.4.3. Compartimentalisation nucléaire

1.1.4.3.1. Organisation tridimensionnelle des télomères dans le noyau

L'arrangement spatial tridimensionnel des fibres de chromatine dans l'espace nucléaire est une notion qui est apparue au début du XXe siècle (Boveri, 1909) mais qui n'a pu être pleinement développée que plus tard avec l'avènement de l'hybridation *in situ* en fluorescence (FISH) (notamment des peintures chromosomiques) et les techniques de microscopie en trois dimensions (3D) (Cremer et al., 2008).

Dans le noyau interphasique, l'organisation de base est le territoire chromosomique (**Figure 8a**). Celui-ci est composé de repliement de fibres chromatiniennes qui seraient rattachées à la matrice nucléaire, un réseau ribonucléoprotéique parcourant l'ensemble du noyau, par des séquences d'attachement dénommées S/MARs (Cremer & Cremer, 2001; Nickerson, 2001). Cependant, l'organisation exacte des territoires chromosomiques de même que la notion de matrice nucléaire sont encore débattues. L'unité de base serait un domaine chromatinien d'environ 1 Mb, correspondant à des foyers réplicatifs pendant la phase S (Jackson & Pombo, 1998; Ma et al., 1998). Pour d'autres, il s'agirait plutôt de petites boucles de chromatine de 50 à 200 kb appelées rosettes (Munkel et al., 1999), ou encore de boucles géantes de plusieurs mégabases séparées par des régions de 200 kb ("*giant loop model*") (Sachs et al., 1995). Différents modèles architecturaux ont également été proposés (Chubb & Bickmore, 2003; Dehghani et al., 2005; Branco & Pombo, 2006), comme le modèle du territoire chromosomique-compartiment interchromatinien (Cremer et al., 2006) (**Figure 8b**). Le territoire chromosomique ressemble à une éponge traversée par un réseau continu de canaux et de lacunes, le compartiment interchromatinien. Ce dernier correspond à un réseau ribonucléoprotéique qui débute au niveau des pores nucléaires et s'étend entre les différents territoires chromosomiques et l'intérieur de ceux-ci (Rouquette et al., 2009). Les gènes activement transcrits sont exposés directement au contact de l'espace interchromatinien soit en bordure du territoire chromosomique, soit à l'intérieur de celui-ci (Verschure et al., 1999; Volpi et al., 2000; Mahy et al., 2002a; Mahy et al., 2002b; Kupper et al., 2007).

Figure 8 : Organisation tridimensionnelle du noyau

a : Détection simultanée des chromosomes dans les noyaux de fibroblastes humains (gauche) et en prométaphase (droite) (d'après Cremer & Cremer, 2010). Chaque territoire chromosomique et chromosome prométaphasique est représenté en fausse couleur et est annoté par son numéro.

b : Modèle d'architecture nucléaire territoire chromosomique (CT)-compartiment interchromatinien (IC) (d'après Cremer & Cremer, 2010). Les gènes (ronds gris) sont transcrits en périphérie des CT (machinerie transcriptionnelle en blanc). L'épissage à lieu au niveau de foyers spécifiques (*speckle*).

c : Lamina nucléaire (NL) et ses interactions avec le génome (d'après Kind & van Steensel, 2010). NPC : complexe de pore nucléaire; ONM : membrane nucléaire externe, INM : membrane nucléaire interne. Lap2 β , Emerin, LBR, MAN1 : protéines associées à la lamina. BAF : protéine associée à la chromatine. La chromatine associée à NL est transcriptionnellement inactive (enrichissement en H3K9me2 et H3K27me3). La bordure de certains LADs est limitée par des sites CTCF. Le domaine LEM est un domaine présent sur certaines protéines de la INM.

Chaque chromosome, et même chaque bras chromosomique, occupe un territoire distinct (Dietzel et al., 1998). La dynamique de la chromatine en interphase étant limitée (Chubb et al., 2002), la stabilité d'un territoire chromosomique une fois établie en interphase dure jusqu'à la mitose suivante (Walter et al., 2003). Cette propriété est conservée et non aléatoire (Cremer & Cremer, 2001). Cependant, elle peut varier en fonction des espèces et des types cellulaires étudiés. Ainsi, établir une relation entre l'organisation nucléaire et sa fonction peut s'avérer difficile. De nombreux paramètres ont été mis en rapport avec la position nucléaire comme l'activité transcriptionnelle, la cinétique de réplication ou le contenu en GC (Goetze et al., 2007a; Grasser et al., 2008; Heppenger et al., 2008).

L'organisation tridimensionnelle des territoires chromosomiques semble corrélée à la densité en gènes (Croft et al., 1999; Boyle et al., 2001; Murmann et al., 2005). Ainsi, le chromosome 19 riche en gène est localisé vers le centre du noyau alors que le chromosome 18, pauvre en gènes, est localisé vers l'enveloppe nucléaire. De même à l'échelle d'un territoire, les "ridges" (région génomique contiguë d'environ 80 gènes possédant une activité transcriptionnelle globale élevée) sont localisés de manière plus interne que les "anti-ridges", qui seraient plus périphériques (Goetze et al., 2007b). En revanche, la corrélation entre activité transcriptionnelle d'un gène et position nucléaire est plus débattue (Gilbert et al., 2004). Le repositionnement du gène *CFTR* a été observé en fonction de son état transcriptionnel. Il serait périphérique dans les cellules où il n'est pas exprimé et plus central dans les cellules où il est actif (Zink et al., 2004). Par ailleurs, l'observation d'hétérochromatine et de gènes silencieux à la périphérie nucléaire a renforcé le concept selon lequel la périphérie nucléaire est un environnement répressif pour la transcription et que l'intérieur du noyau est plus favorable à l'activité transcriptionnelle (Schneider & Grosschedl, 2007). Toutefois, il n'est pas établi si la position périphérique est la cause ou la conséquence de l'activité d'un gène. De plus, il est apparu que la périphérie nucléaire est un compartiment complexe où gènes actifs et inactifs peuvent être associés (Akhtar & Gasser, 2007; Kumaran & Spector, 2008; Deniaud & Bickmore, 2009; Luo et al., 2009).

La localisation des télomères dans le noyau est variable en fonction des espèces. Chez la levure, les 32 télomères sont regroupés en 4 à 6 foyers associés à l'enveloppe nucléaire (Gotta et al., 1996). Chez la souris, les télomères sont localisés plutôt en périphérie. Chez l'Homme, la localisation varie en fonction du type cellulaire. Les télomères sont regroupés à la périphérie nucléaire dans les spermatozoïdes (Zalensky et al., 1995) alors qu'ils sont plutôt internes dans les lymphocytes et les cellules Hela (Luderus et al., 1996; Weierich et al., 2003).

La formation de foyers est plus fréquente dans les cellules en G0 que dans les cellules répliquatives (Nagele et al., 2001; Weierich et al., 2003). Certains télomères présentent un comportement différent comme le télomère du bras long du chromosome 4 (4qter), impliqué dans la myopathie facio-scapulo-humérale, qui est localisé à la périphérie du noyau (Masny et al., 2004; Tam et al., 2004), ce qui pourrait lui conférer des propriétés particulières.

1.1.4.3.2. Rôle de la périphérie nucléaire

La périphérie nucléaire est un environnement particulier. Le noyau est délimité par une membrane nucléaire composée d'une bicouche phospholipidique entrecoupée de complexes de pores nucléaires qui permettent les échanges noyau-cytoplasme. La surface nucléoplasmique de la membrane nucléaire interne est recouverte d'un réseau filamenteux appelé lamina nucléaire (**Figure 8c**). Celle-ci est étroitement associée à la membrane nucléaire interne et rattachée à la périphérie par les complexes de pore nucléaire. La lamina nucléaire est composée de deux types de filaments intermédiaires de type V spécialisés, les lamines de type A et de type B. Les lamines possèdent un domaine central alpha-hélice entouré de deux domaines globulaires. Le domaine central permet de former des dimères de lamines qui s'assemblent ensuite en chaîne puis forment un feuillet d'orientation anti-parallèle. Les lamines de type B sont codées par deux gènes (*LMNB1* et *LMNB2*) et sont exprimées dans toutes les cellules. Les lamines de type A sont codées par un seul gène, *LMNA*, dont l'expression génère une pré-lamine qui subit des modifications post-traductionnelles comprenant farnésylation, méthylation et clivage et est à l'origine des lamines A, C, AΔ10 et C2. Les lamines A sont exprimées uniquement dans les cellules différenciées. Les lamines peuvent interagir avec de nombreuses protéines dont certaines font partie de la membrane nucléaire interne. Les interactions entre les lamines et la chromatine peuvent être soit directes avec les histones (Ruault et al., 2008), soit indirectes par l'intermédiaire de facteurs de transcription (Rb, SREBP1, MOK2, cFos) ou par l'interaction entre des protéines associées à la lamina et des protéines associées à la chromatine (LAP2-emerin-MAN1 et BAF, LBR et HP1). Les lamines forment également un réseau dans le nucléoplasme, qui pourrait intervenir dans la prolifération cellulaire, l'initiation de la réplication et l'épissage des ARNs (Dechat et al., 2008).

Les lamines jouent un rôle structural pour l'enveloppe nucléaire et permettent la transduction de signaux entre le nucléoplasme et le cytoplasme grâce à l'interaction avec les

protéines de la membrane nucléaire et du cytosquelette (Starr, 2009). De plus, la lamina nucléaire participe à l'organisation de la chromatine interphasique, la réplication de l'ADN, la transcription, la prolifération cellulaire, la différenciation et la réparation de l'ADN (Dechat et al., 2008). Cette multifonctionnalité est liée au fait que la lamina nucléaire est une structure non uniforme, dont les composants peuvent avoir différentes fonctions. Les pores nucléaires sont capables d'interagir avec l'euchromatine et des gènes actifs (Gilbert & Ramsahoye, 2005; Menon et al., 2005; Kalverda et al., 2010). Par ailleurs, les lamines de type B seraient plutôt associées à l'hétérochromatine et aux régions silencieuses, alors que les lamines de type A seraient associées à des régions d'euchromatine riches en gène (Shimi et al., 2008). Ceci expliquerait les résultats parfois discordants dans les corrélations entre localisation nucléaire et expression génique (Luo et al., 2009). Récemment, une cartographie des régions d'interaction entre la lamine B1 et l'ADN a été établie dans des cellules humaines par la technique damID (technique basée sur la méthylation des adénines des séquences d'ADN qui interagissent avec une protéine d'intérêt, ici la Lamine B1, fusionnée avec l'enzyme dam DNA adényl transférase) (Guelen et al., 2008). Elle a permis d'identifier 1 300 domaines associés à la lamina (LADs), soit environ 40% du génome. Ces domaines sont de relativement grandes tailles de 0,1 à 10 Mb (553 kb en moyenne) et sont pauvres en gènes. Les gènes des LADs sont 5 à 10 fois moins actifs que ceux qui ne sont pas associés à la lamina et sont dépourvus des marques de transcription active (polymérase II et diméthylation des résidus H3K4). Les LADs sont par ailleurs enrichis en résidus H3K27 triméthylés et H3K9 diméthylés et pauvres en histones acétylés, compatibles avec un environnement chromatinien répressif. L'interaction LAD-lamina pourrait être médiée par la liaison de la lamine B1 avec le facteur Oct1 dont les sites sont enrichis dans les LADs. La transition avec les régions non-LADs est brutale sur environ 10 kb. Trois marques, dans les 5 à 10 kb à l'extérieur des LADs, signalent préférentiellement ces frontières : la présence du site de liaison de la protéine insulatrice CTCF, l'augmentation de la densité de promoteurs orientés vers l'extérieur du LAD (associée à une augmentation de la polymérase II et des résidus H3K4 diméthylés) et l'enrichissement en îlots de CpGs. Ainsi les LADs permettraient de structurer l'organisation de domaines chromosomiques dans le noyau.

Enfin, les mutations constitutionnelles des lamines de type A (plus de 250) sont responsables d'un ensemble de pathologies appelées laminopathies. Elles regroupent des pathologies variées comme des dystrophies musculaires, des lipodystrophies, des neuropathies périphériques et des syndromes de vieillissement prématuré dont le plus connu est la progéria de Hutchinson-Gilford (Worman et al., 2009). Leurs principales caractéristiques cliniques

sont résumées dans le **tableau I**. Au niveau cellulaire, l'architecture nucléaire est perturbée avec des noyaux irréguliers, des hernies de l'enveloppe nucléaire, une perte des lamines de type B, des anomalies des pores nucléaires et une délocalisation de l'émerine. Une perte d'hétérochromatine périphérique, des anomalies de la méthylation des histones et des anomalies de localisation des territoires chromosomiques (Goldman et al., 2004; Galiova et al., 2008) sont également observées. La localisation subnucléaire, la taille et la structure chromatiniennne des télomères sont également perturbées, indiquant un rôle important des lamines de type A dans la biologie des télomères (Allsopp et al., 1992; Huang et al., 2008; Raz et al., 2008).

Maladie	Transmission	OMIM	Phénotype
Dystrophie musculaire			
Dystrophie musculaire d'Emery-Dreyfus ¹	AD/AR	#181350	Contractures musculaires, faiblesse musculaire, troubles du rythme et de la conduction cardiaque, cardiomyopathie
Cardiomyopathie dilatée 1A	AD	#115200	Cardiomyopathie
Dystrophie musculaire des ceintures LGMD1B	AD	#159001	Faiblesse musculaire des membres et des ceintures, cardiomyopathie dilatée
Dystrophie musculaire congénitale	AD	#613205	Atteinte musculaire diffuse dans la 1ère année de vie, cardiomyopathie
Syndrome main-cœur	AD	#610140	Brachydactylie mains et pieds, cardiomyopathie
Lipodystrophie			
Lipodystrophie partielle familiale de Dunnigan (FPLD2)	AD	#151660	Perte du tissu adipeux périphérique, insulino-résistance, diabète, hypertriglycéridémie, stéatose hépatique
Lipoatrophie avec diabète, stéatose hépatique, cardiomyopathie hypertrophique et papules leucomélanodermiques	AD	*150330	Perte générale du tissu adipeux, diabète insulino-résistant, stéatose hépatique, cardiomyopathie hypertrophique, papules disséminées
Dysplasie mandibulo-acrale ²	AR	#248370	Hypoplasie de la mandibule, acroostéolyse, raideur articulaire, atrophie cutanée mains et pieds, hypoplasie des clavicules. Persistance des sutures crâniennes et os wormien, alopecie, petite taille, lipodystrophie partielle
Neuropathie périphérique			
Maladie de Charcot-Marie-Tooth 2B1	AR	#605588	Faiblesse et aréflexie des membres inférieurs
Syndrome de vieillissement prématuré			
Progéria de Hutchinson-Gilford	AD	#176670	Vieillesse prématurée, retard de croissance, micrognathie, diminution du tissu adipeux sous-cutané, ostéoporose, pathologie vasculaire occlusive précoce
Syndrome de Werner atypique	AD	*150330	Vieillesse prématurée, petite taille, cataracte, sclérodermie, ostéoporose, pathologie vasculaire

AD: autosomique dominant ; AR: autosomique récessif

1: la dystrophie musculaire d'Emery-Dreyfus peut également être secondaire à des mutations du gène de l'Emeryine (EMD), récessives liées à l'X

2: la dysplasie mandibulo-acrale présente à la fois des signes de lipodystrophie et de vieillissement prématuré

Tableau I : Principales maladies dues à une mutation du gène *LMNA* (laminopathies)
(d'après Worman et al., 2009)

Cette première partie n'a pas pour but d'être exhaustive mais elle permet de rappeler la place fondamentale des télomères dans l'homéostasie cellulaire et la stabilité du génome ainsi que de présenter certains éléments clés de leur régulation. Dans de nombreux aspects, la physiologie des télomères est étroitement liée aux séquences qui leurs sont immédiatement adjacentes, les séquences subtélomériques, soit en les influençant, soit en subissant leur influence en retour. Nous détaillerons, dans la seconde partie de cette introduction, ces régions dont les connaissances sur la structure et les fonctions sont en plein développement.

1.2. Les subtélomères

1.2.1. Définition

Les subtélomères sont définis comme les régions de transition entre les séquences chromosomiques spécifiques et les répétitions télomériques terminales (TTAGGG)_n (Mefford & Trask, 2002) (**Figure 1**). Donner une définition plus précise est difficile car ces régions sont extrêmement dynamiques et variables. Elles sont composées d'un patchwork de séquences s'étendant sur une distance de 8 kb à 300 kb environ (Riethman et al., 2001) et sont particulièrement riches en duplications segmentaires (ou duplicons), qui sont des segments d'ADN génomique répétés d'au moins 1 kb de long et de plus de 90% de similarité de séquence. Les duplications segmentaires subtélomériques, encore appelées répétitions subtélomériques, représentent 25% des 500 kb les plus distaux et 80% des 100 kb les plus distaux des chromosomes (Riethman et al., 2004). Du fait de la présence de ces séquences répétées et de la grande variabilité allélique, l'identification et le clonage des séquences subtélomériques a été tardif, particulièrement difficile et n'est pas encore complet. De plus la sous-représentation de ces séquences dans les banques de BACs ("*Bacterial Artificial Chromosome*", chromosome artificiel de bactérie) qui ont permis le séquençage du génome a nécessité la mise au point de techniques spécifiques comme la production de demi-YACs ("*Yeast Artificial Chromosome*", chromosome artificiel de levure) (Riethman et al., 2001).

1.2.2. Structure des subtélomères

Les subtélomères sont composés d'un patchwork de séquences hautement évolutives et dynamiques. Les extrémités chromosomiques sont composées (dans le sens télomère-centromère) des répétitions (TTAGGG)_n terminales suivies par la région subterminale (<25kb) puis par la région subtélomérique étendue (Riethman, 2008b, a).

Figure 9 : Structure des régions subtélomériques

- a : Représentation schématique d'une région chromosomique subterminale (d'après Riethman, 2008a). Extrémité 3' sortante (rouge) et répétitions télomériques (noir); (TTAGGG) $_n$ dégénérées (noir et gris), séquence TAR1 (rose), séquences Sipt ou spécifique des chromosomes (bleu), transcrits subtélomériques (flèche verte).
- b : Exemple d'organisation des séquences dans différents subtélomères (d'après Riethman, 2008b) : (TTAGGG) $_n$ dégénérées (flèche noire), Sipt (bleu), duplicons subtélomériques (vert), séquence spécifique (jaune), région non séquencée (gap) (rose).
- c : Exemple de variations structurales d'un même subtélomère (d'après Riethman, 2008b). Séquence spécifique (jaune), variant de séquences spécifiques (jaune hachuré), répétitions subtélomériques (en couleur).

La région subterminale comprend des répétitions télomériques (TTAGGG)_n dégénérées, des séquences de type TAR1, qui sont des séquences satellites de 0 à 2 kb présentes dans les 2 kb adjacents aux répétitions télomériques (Brown et al., 1990) et enfin soit des séquences d'ADN spécifiques des chromosomes en copie unique (télomères 7q, 8q, 11q, 12q, 18q, Xp/Yp), soit une des six familles de duplicons subterminaux (Srpt A à F) (télomères 2p, 4p, 9p, 10q, 17p, 18p) (Ambrosini et al., 2007) (**Figure 9a**). Les duplicons subterminaux peuvent avoir des copies subterminales, subtélomériques ou à l'intérieur du génome. Certains de ces blocs contiennent des gènes, comme la famille des gènes WASH, dont l'expression est susceptible d'être directement affectée par la structure des télomères (§ 1.2.3.1.).

Les régions subtélomériques étendues sont très variables dans leur taille, leur séquence et leur organisation (**Figure 9b**). Elles peuvent contenir des répétitions télomériques dégénérées, des duplicons Srpt, d'autres duplications segmentaires et des séquences spécifiques de chromosome (Ambrosini et al., 2007). De plus, les régions subtélomériques sont enrichies en îlots de CpGs (séquence copie unique : 47,9%; Srpt : 44% ; duplicons 43 % ; génome 41,6%) (Riethman et al., 2004). Les séquences immédiatement centromériques aux îlots de CpGs subterminaux sont enrichies en sites CTCF, suggérant la présence de frontières chromatinienne (Riethman, 2008b). Les répétitions (TTAGGG)_n internes sont 25 fois plus fréquentes dans les régions subtélomériques que dans le reste du génome. Elles sont plus longues, de 150 à 200 pb en moyenne (maximum 823 pb), et plus conservées que dans le reste du génome, reflétant soit un événement de propagation récent soit une forte pression de sélection. Elles sont orientées vers les télomères et coïncident avec la frontière des duplicons. Les duplications segmentaires subtélomériques sont généralement plus grandes et plus nombreuses. Certaines sont spécifiques des subtélomères, d'autres sont plus répandues. En effet, 1151 blocs subtélomériques, dont 461 exclusivement subtélomériques, ont été recensés (Ambrosini et al., 2007). La densité en duplicons est très variable d'un subtélomère à un autre. Des délétions et insertions d'ADN en copie unique ont également été décrites.

Du fait des grandes disparités alléliques, les régions subtélomériques ont été parmi les premiers CNV (variation du nombre de copies) mis en évidence (Redon et al., 2006) (**Figure 9c**). Un exemple extrême est celui du bras court du chromosome 16, où la différence de taille entre le plus petit et le plus grand allèle est de 260 kb (Wilkie et al., 1991).

Cette grande variabilité allélique traduit les remaniements constants et les échanges de duplicons dont les subtélomères sont les objets. Par exemple, le bloc 7501 est présent en une

seule copie chez les primates non humains mais en 7 à 11 copies chez l'Homme. Il est porté par les extrémités 3qter, 15qter et 19pter mais beaucoup d'autres combinaisons ont été décrites en fonction de l'origine géographique. Ceci illustre une dynamique constante et une évolution récente (Mefford & Trask, 2002). Il a été estimé, en effet, que 49% des séquences subtélomériques ont été générées après la divergence Homme-chimpanzé (Linardopoulou et al., 2005). La formation des subtélomères proviendrait de translocations impliquant l'extrémité des chromosomes suivies de la transmission d'un complément déséquilibré à la descendance. La présence d'un polymorphisme segmentaire prédisposerait à de nouveaux réarrangements qui seraient à l'origine de nouveaux polymorphismes (Linardopoulou et al., 2005) (**Figure 10**). Le mécanisme majeur impliqué dans ces réarrangements est le NHEJ (92%) alors que la recombinaison homologue non-allélique ("*Non Allelic Homologous Recombination*" NAHR) n'interviendrait que rarement (8% des cas) (Linardopoulou et al., 2005). L'évolution des blocs chromosomiques n'est donc pas indépendante mais est marquée par de constantes interactions entre les chromosomes.

Figure 10 : Dynamique des régions subtélomériques : Génération de polymorphisme segmentaire par translocations (d'après Linardopoulou et al., 2005)

a : Une duplication/délétion terminale peut survenir lors de la transmission d'un chromosome dérivé de translocation et de son homologue intact à la descendance (c).

b : Une duplication/délétion segmentaire peut survenir lors d'un échange entre les chromosomes transloqués.

d et e : Les polymorphismes segmentaires peuvent favoriser d'autres réarrangements par l'intermédiaire de translocations entre régions homologues (d) ou non-homologues (e).

f et g : Des transferts de séquences homologues, réciproques ou non, sont alors possibles entre les différents chromosomes générés (l'astérisque indique les variants de séquences).

1.2.3. Rôle des subtélomères

Les fonctions des régions subtélomériques sont encore peu connues et assez difficiles à appréhender chez l'Homme. Les subtélomères ont été impliqués dans l'adaptation environnementale, la régulation du TPE et la recombinaison télomérique.

1.2.3.1. Transcrits et gènes subtélomériques

Les régions subtélomériques sont plus riches en transcrits (54 transcrits/Mb) que la moyenne du génome (43 transcrits /Mb). Plus de 1 000 transcrits ont été identifiés, certains issus de l'ADN en copie unique, d'autres des duplicons (Riethman et al., 2004). Parmi ces transcrits se trouvent les TERRAs dont la transcription est initiée au niveau des subtélomères et qui incluent une partie plus ou moins importante de séquence subtélomérique (§1.1.4.2.). Par ailleurs, de nombreux pseudogènes sont présents dans les subtélomères, mais certaines familles possèdent au moins une copie fonctionnelle.

Les régions subterminales contiennent les gènes *RPL23A7P*, *DEAD/H*, *DDX3L* (Costa et al., 2009) et la famille des gènes *WASH* ("Wiscott-Aldrich syndrome protein and Scar homolog") (Riethman et al., 2004). L'extrémité 3' des gènes *WASH* se termine dans les 5 kb avant le télomère. Il existe une forme courte, uniquement subterminale et une forme longue qui commence près d'îlots de CpGs situés à 20 kb du télomère. Cette ancienne famille est conservée au cours de l'évolution et est même indispensable à la survie chez *D. melanogaster* (Linardopoulou et al., 2007). Les produits de ces gènes participent à l'organisation du cytosquelette, la transduction du signal et la régulation de la polymérisation de l'actine (Millard et al., 2004; Takenawa & Suetsugu, 2007; Yamaguchi & Condeelis, 2007).

Les transcrits subtélomériques comprennent des familles de gènes codant par exemple pour les chaînes lourdes d'immunoglobulines, les récepteurs olfactifs (OR), les α et β -défensines et des protéines en doigts de zinc, mais aussi des gènes uniques comme le récepteur à l'interleukine 9 (*IL9R*) ou *DUX4* (Riethman et al., 2004). Les gènes des α et β -défensines codent pour des peptides cationiques possédant une activité anti-microbienne, anti-fongique et anti-virale (Ganz, 2003) dont les variations du nombre de copies (CNVs) ont été associées à des pathologies dysimmunitaires comme le lupus érythémateux disséminé, la

maladie de Crohn et le psoriasis (Bennett et al., 2003; Ishii et al., 2005; Fellermann et al., 2006; Hollox et al., 2008). Les récepteurs olfactifs sont une des grandes familles du génome incluant 400 gènes et pseudogènes (Glusman et al., 2001; Nîimura & Nei, 2007). Chaque neurone de l'épithélium olfactif exprime un seul allèle d'un seul gène OR et les axones des neurones exprimant le même gène convergent dans le bulbe olfactif. La combinaison des différents OR permet la détection de millions de chimiques. Des locis OR sont impliqués dans des remaniements génomiques récurrents. Par exemple, à l'extrémité du bras court du chromosome 8 (8pter), la recombinaison inégale entre deux groupes d'OR est à l'origine d'inversion-duplication, de marqueur surnuméraire, de délétion interstitielle et d'inversion submicroscopique (Giglio et al., 2001; Ciccone et al., 2006).

Dans d'autres espèces, la localisation subtélomérique des gènes permet une adaptation à l'environnement. Chez la levure, les subtélomères contiennent une famille de gènes dont les produits permettent d'utiliser différentes sources de carbone (SUC codant pour la β -fructofuranosidase, MAL pour l' α -glucosidase et MEL pour l' α -galactosidase) (Charron & Michels, 1988; Naumov et al., 1990). Les subtélomères de *P. falciparum* regroupent la plupart des membres des gènes var, codant pour l'antigène de surface APfEMP-1. La dynamique subtélomérique permet de générer une diversité des gènes et d'adapter la réponse à l'hôte (Freitas-Junior et al., 2000).

Chez l'Homme, le remaniement constant des télomères pourrait également contribuer à l'apparition de nouveaux gènes, les gènes dupliqués évoluant par mutations successives soit vers une sur-spécialisation, soit vers une nouvelle fonction, et permettrait ainsi l'adaptation à de nouvelles conditions environnementales. Par ailleurs, la proximité des télomères place les gènes subtélomériques sous l'influence directe du TPE (§1.2.3.2), qui pourrait également réguler la réponse aux stimuli environnementaux, les régions subtélomériques étant riches en gènes de l'immunité et de la réponse inflammatoire.

1.2.3.2. Subtélomères et effet de position télomérique

La structure hétérochromatinienne des télomères s'étend plus ou moins loin sur la région subtélomérique, notamment la triméthylation des résidus H3K9 et H4K20, l'enrichissement en protéine HP1 et le faible taux d'acétylation des histones (Tommerup et al., 1994; Gonzalo et al., 2006). De plus, contrairement à l'ADN télomérique, l'ADN subtélomérique contient des îlots de CpGs, qui peuvent être méthylés par les ADN

méthyltransférases DNMT3a et b (méthylation *de novo*) et DNMT1 (maintenance de la méthylation). Cette méthylation est régulée au cours du développement. L'ADN subtélomérique est hypométhylé dans les cellules germinales, puis méthylé *de novo* dans les cellules somatiques (Riethman, 2008b). Dans un modèle murin, il a été proposé que la méthylation de l'ADN subtélomérique serait un mécanisme secondaire qui permettrait de renforcer le TPE (Gao et al., 2007).

Plusieurs arguments suggèrent que la modification de la chromatine télomérique secondaire au raccourcissement télomérique ou à d'autres stimuli, pourrait modifier l'expression des gènes subtélomériques. Tout d'abord, dans les modèles cellulaires de TPE, il a été montré que l'expression du gène rapporteur était dépendante de la longueur des télomères (Baur et al., 2001). Ensuite, l'expression de gènes subtélomériques a également été étudiée au cours de la sénescence de fibroblastes humains en culture et corrélée à la longueur des télomères correspondants. Bien que 17 des 34 gènes étudiés présentaient une expression différentielle (augmentée ou diminuée), la longueur des télomères ne semblait pas expliquer à elle seule le statut d'expression (Ning et al., 2003). Une autre étude comparant l'expression de 9 gènes subtélomériques dans des fibroblastes de sujets jeunes et de sujets âgés a mis en évidence l'augmentation avec l'âge de l'expression de 3 gènes impliqués dans le remodelage de la chromatine, *EHMT1*, *ZMYND11* et *RASA3* (Hernandez-Caballero et al., 2009). Enfin, l'effet du TPE sur des gènes subtélomériques a été observé en pathologie chromosomique constitutionnelle. Les chromosomes en anneau sont formés par la délétion terminale d'un chromosome et d'une fusion au point de cassure. Certains d'entre eux peuvent néanmoins dériver d'une fusion télomère-télomère, qui pourrait aboutir à un TPE sur des gènes localisés dans les régions subtélomériques, comme dans un cas d'anneau du chromosome 20 avec conservation télomérique où la répression des gènes *CHRN4* et *KCNQ2* responsable d'une épilepsie sévère a été rapportée (Zou et al., 2006).

Les séquences subtélomériques peuvent également se comporter comme une barrière anti-TPE. Chez la levure, différents éléments subtélomériques peuvent s'opposer au TPE (Pryde & Louis, 1999). Chez l'Homme, les gènes d'un même subtélomère ne sont pas réprimés de façon identique (Ning et al., 2003). Il a été suggéré que la méthylation différentielle des îlots de CpGs subtélomériques ainsi que les sites CTCF qui leur sont adjacents pourraient contribuer à délimiter des frontières hétérochromatine-euchromatine, régulant l'expression différentielle des gènes au sein du subtélomère (Riethman, 2008b).

Enfin, les modifications épigénétiques des subtélocères pourraient influencer d'autres aspects de la biologie des télomères comme leur réplication ou leur recombinaison (Gonzalo et al., 2006; Riethman, 2008b).

1.2.3.3. Le mécanisme ALT : “*Alternative Lengthening of telomere*”

Le mécanisme ALT est responsable de la stabilisation des télomères dans 10 à 15 % des tumeurs (Cesare & Reddel, 2010), notamment celles d'origine mésenchymateuse. Sur le plan phénotypique, les cellules ALT sont caractérisées par des télomères de taille très hétérogène de 2 à plus de 50 kb (20 kb en moyenne) (Bryan et al., 1995). Elles contiennent des corps nucléaires APBs (“*ALT associated PML bodies*”) où colocalisent les télomères et qui contiennent les protéines PML, SP100 ainsi que des protéines impliquées dans la réplication, la recombinaison et la réparation de l'ADN (Nittis et al., 2008). Elles sont caractérisées par la présence d'ADN télomérique extra-chromosomique (cercles télomériques double brin), un fort taux de recombinaison télomérique et une augmentation du taux d'échange entre chromatides soeurs télomériques (T-SCE). Le mécanisme ALT repose sur la recombinaison homologue (**Figure 11**). Dunham et collaborateurs ont montré, en suivant la propagation d'un télomère étiqueté (“*taggé*”), que les séquences d'ADN sont recopiées d'un télomère à un autre dans les cellules ALT humaines (Dunham et al., 2000). Le processus de recombinaison s'opérerait au niveau télomérique par un T-SCE inégal, par la recombinaison avec un chromosome adjacent ou par la recombinaison dans la boucle-t (Cesare & Reddel, 2010). L'appariement pourrait également avoir lieu au niveau des régions subtélomériques. Cela a été uniquement démontré chez la levure mais n'a pas été exclu formellement chez l'Homme (Lundblad & Blackburn, 1993; Teng & Zakian, 1999).

Par ailleurs, la méthylation des subtélocères pourrait réguler le processus ALT. En effet, dans un modèle de cellules ES murines déficientes en ADN-méthyltransférase, il a été observé une déméthylation des régions subtélomériques sans modification des marques hétérochromatiniennes des histones, une augmentation de la taille des télomères et une augmentation de la recombinaison homologue entre les séquences télomériques (Gonzalo et al., 2006). Cependant, une étude plus récente suggère que le phénomène de déméthylation subtélomérique, bien que souvent associé aux cellules ALT, ne soit pas nécessaire à ce processus (Tilman et al., 2009).

Figure 11 : Maintenance des télomères par le mécanisme ALT

(d'après Mefford & Trask, 2002)

a : Voie 1: recombinaison homologue entre télomères.

b : Voie 2: recombinaison homologue entre régions subtélomériques.

Il reste encore beaucoup à découvrir sur les fonctions des régions subtélomériques. Leur étude dans les cellules humaines est difficile du fait de leur extrême diversité allélique. Cependant, une séquence subtélomérique à l'extrémité du bras long du chromosome 4 a attiré l'attention car elle est impliquée dans la dystrophie musculaire facio-scapulo-humérale (FSHD), une pathologie dont le mécanisme, plutôt inhabituel, fait intervenir la contraction d'une séquence répétée appelée *D4Z4*.

1.3. Exemple d'un élément subtélomérique impliqué en pathologie humaine : *D4Z4* et la myopathie facio-scapulo-humérale

1.3.1. *D4Z4*, un macrosatellite subtélomérique

D4Z4 associe le double mystère d'être à la fois une séquence subtélomérique et un macrosatellite. Les macrosatellites font partie des séquences satellites qui sont des séquences d'ADN répétées en tandem, c'est-à-dire immédiatement adjacentes et dans le même sens que la précédente (Tremblay et al., 2010). Le plus souvent, le nombre de ces répétitions en tandem est variable d'un individu à l'autre (polymorphe), c'est pourquoi ces séquences sont aussi appelées VNTR ("Variable Number of Tandem Repeats"). Les macrosatellites font partie des plus grands VNTR et sont définis par la taille d'une unité supérieure à 1 kb. Ces séquences sont souvent spécifiques d'un ou deux chromosomes. Peu d'entre elles ont été caractérisées jusqu'à présent : *D4Z4* (4q35, 10q26), *DXZ4* (Xq23q24) (Chadwick, 2008), *RS447* (4p16.1, 8p23), *TAF-11-like* (5p15.1), *SST1* (4q28.3, 19q13.12), *PRR20* (13q21.1) et *ZAV* (9q32) (Tremblay et al., 2010). Certaines de ces séquences contiennent une phase ouverte de lecture (ORF, "Open Reading Frame") transcrite et peuvent présenter une instabilité mitotique et/ou méiotique (Tremblay et al., 2010). Leur rôle n'est pas bien connu à ce jour. *D4Z4* est sans doute le macrosatellite le plus étudié car il est associé à la FSHD.

D4Z4 est une séquence macrosatellite localisée dans la région subtélomérique du bras long du chromosome 4 en 4q35. Elle est composée de 11 à 150 répétitions d'éléments délimités par les sites de restrictions *KpnI* et dont la taille est de 3303 pb (**Figure 12c**). Elle est très riche en GC (73%) et contient 290 dinucléotides CpG (Lee et al., 1995). Chaque unité contient un motif *LSau* (1-340), une séquence *hhspm3* (1313-1780) et un double domaine homeoboîte contenu dans un cadre ouvert de lecture de 1173 pb, *DUX4* (1792-3063) (Hewitt et al., 1994; Winokur et al., 1994; Lee et al., 1995; Gabriels et al., 1999). *LSau* est un élément répété associé aux régions d'hétérochromatine et *hhspm3* est un duplicon riche en GC. Le cadre ouvert de lecture est précédé d'un promoteur (boîte TACAA) situé 149 pb en amont,

possède un intron facultatif mais aucun site de poly-adénylation dans l'élément *D4Z4* même (Gabriels et al., 1999; Dixit et al., 2007).

Plusieurs gènes ont été identifiés dans la région subtélomérique 4q35 en amont de *D4Z4* : *FRG2* (à 37 kb de *D4Z4*), *TUBB4Q* (à 80 kb, pseudogène de la superfamille des β -tubulines), *FRG1* (à 125 kb) et *ANTI* (à 4,8 Mb, adénine nucleotide translocation 1, composant de la mitochondrie) (van Deutekom et al., 1996; van Geel et al., 2000) (**Figure 12b**). La région distale à *D4Z4* est séparée de 5 à 14 kb des répétitions télomériques et ne contient pas de gène (Bengtsson et al., 1994). Elle est caractérisée par un polymorphisme qui permet de distinguer différents allèles, 4qA et 4qB. L'allèle 4qA comporte une séquence pLAM de 260 pb qui correspond à une répétition *D4Z4* divergente, des répétitions de type β -satellite de 68 pb sur une distance de 6,2 kb, suivies des séquences subtélomériques (répétitions (TTAGGG)_n interstitielles et région d'homologie avec l'extrémité 4pter) et le télomère. L'allèle 4qB se termine par les 570 premières pb d'une unité *D4Z4* suivi des séquences subtélomériques et télomériques (Lemmers et al., 2002; van Geel et al., 2002). La fréquence des deux allèles est identique dans la population générale.

Les répétitions *D4Z4* ne sont pas seulement présentes à l'extrémité 4q35 mais aussi au niveau de la région d'hétérochromatine du bras long du chromosome 1 (1q12), des bras courts des chromosomes acrocentriques (13p12, 14p12, 15p12, 21p12, 22p12) et du subtélomère du bras long du chromosome 10 (10q26) (Wijmenga et al., 1992; Hewitt et al., 1994) (**Figure 12a**). Les régions 4qter et 10qter possèdent 99% d'homologie qui s'étend sur une région de plus de 200 kb (Deidda et al., 1996; van Geel et al., 2002). Les deux chromosomes peuvent être distingués par un site de restriction *BlnI* uniquement présent dans la région 10qter (Deidda et al., 1996; Lemmers et al., 2001). Il existe des échanges de répétitions entre les régions 10qter et 4qter dans 20 à 30 % de la population générale (Rossi et al., 2007), à l'origine de séquences hybrides 4q/10q. Ainsi, l'analyse de la population HAPMAP a révélé l'existence de 17 haplotypes 4qter et 8 haplotypes 10qter. Lemmers et collaborateurs (Lemmers et al., 2010b) ont élaboré un modèle évolutif où l'allèle 4qA serait le plus ancien. Seuls quatre transferts de séquence au cours de l'évolution auraient suffi à générer les haplotypes actuels. Ces transferts auraient eu lieu avant la migration d'Afrique. Depuis, les allèles 4q et 10q évolueraient indépendamment (Rossi et al., 2007; Lemmers et al., 2010b). Les séquences hybrides correspondraient à des structures intermédiaires de transition 4q/10q.

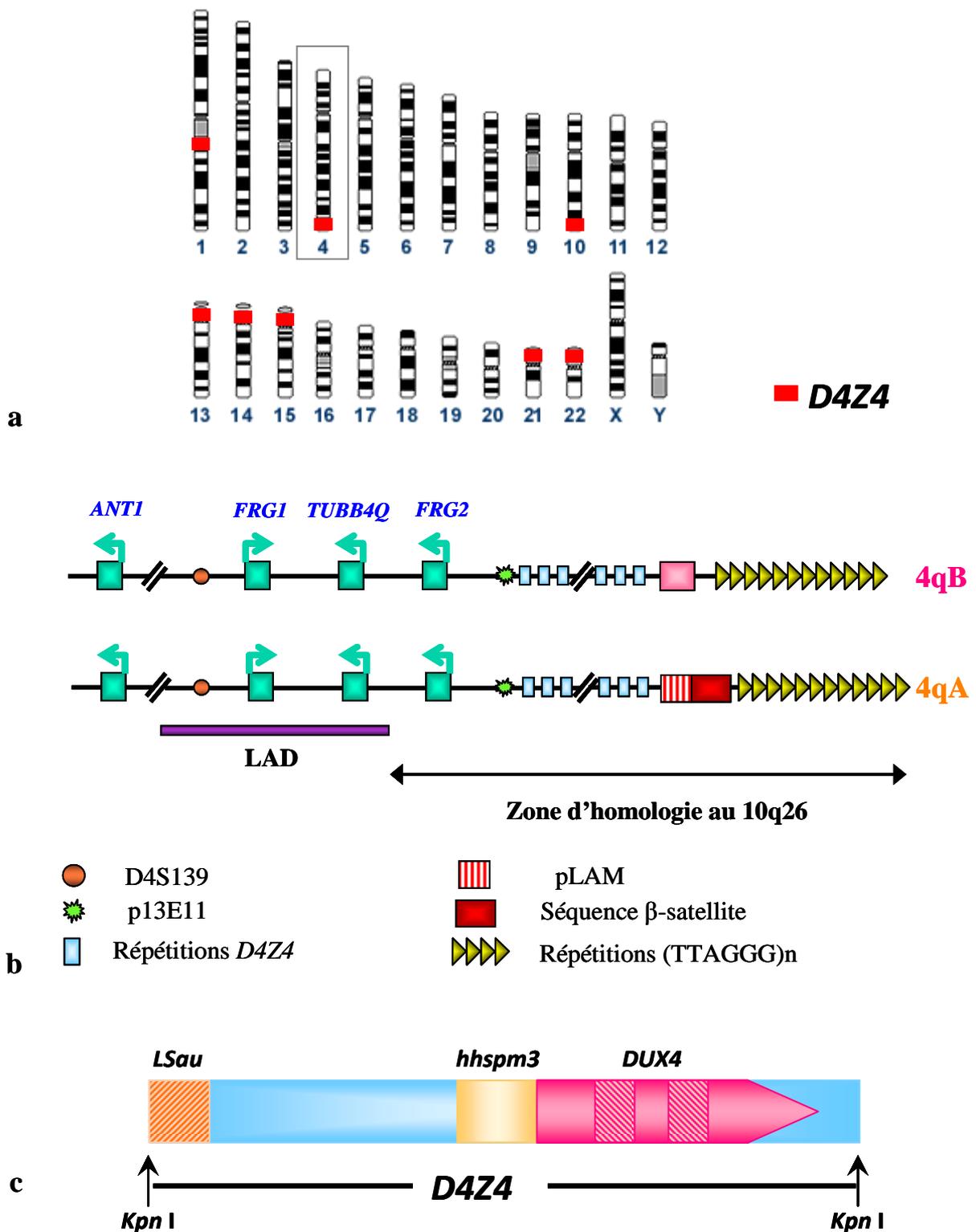


Figure 12: Le macrosatellite *D4Z4*

a : Répartition des répétitions *D4Z4* dans le génome : régions subtélocériques 4q35 et 10q26, région d'hétérochromatine du bras long du chromosome 1 et bras courts des chromosomes acrocentriques.

b : Organisation de la région subtélocérique 4q35. L'allèle 4qA se distingue de l'allèle 4qB par la présence de la séquence pLAM et de répétitions β -satellites. Distance des gènes par rapport à *D4Z4*: *ANTI*: 4,8 Mb; *FRG1* : 125 kb; *TUBB4Q* : 80 kb; *FRG2* : 37 kb; *D4S139* : 215 kb. Le dernier domaine associé à la lamina (LAD) de l'extrémité 4qter est représenté. (Le schéma n'est pas à l'échelle)

c : Structure d'une répétition *D4Z4*. Chaque élément de 3303 pb est défini par deux sites *KpnI*. Distance par rapport au site *KpnI*: *LSau*: 1-340; *hhspm3* 1313-1780; *DUX4*: 1792-3063. Les deux homéoboîtes sont en rose hachuré.

1.3.2. La myopathie facio-scapulo-humérale, aspects cliniques et biologiques

1.3.2.1. Clinique

La myopathie facio-scapulo-humérale (FSHD) (MIM#158900), anciennement appelée myopathie de Landouzy-Déjérine du nom des médecins français qui en firent la première description à la fin du XIX^e siècle, est une dystrophie musculaire de transmission autosomique dominante. Elle est la plus fréquente des dystrophies musculaires avec une prévalence de 7/100 000 (<http://www.orpha.net/consor/cgi-lin/index.php>). Environ 30% des cas sont sporadiques.

La FSHD est caractérisée cliniquement par une faiblesse musculaire progressive impliquant la face, la ceinture scapulaire, les membres supérieurs, les membres inférieurs et enfin la hanche (Padberg et al., 1991; Tawil et al., 1998). L'atteinte musculaire asymétrique est fréquente et caractéristique de la FSHD. L'atteinte des muscles de la face se manifeste par une difficulté à siffler, à fermer fortement les paupières ou à sourire. Les muscles extra-oculaires et bulbaires sont épargnés. L'atteinte scapulaire attire davantage l'attention des patients. Elle se traduit par des difficultés à lever les bras, un décollement scapulaire par atteinte du muscle trapèze et une atrophie pectorale pouvant entraîner un pectus excavatum. Au niveau des membres supérieurs, le deltoïde est touché tardivement. L'atteinte du biceps et du triceps entraîne une atrophie du bras épargnant les muscles de l'avant-bras, formant un aspect de "bras de Popeye". L'atteinte abdominale peut être précoce et se manifeste par une protubérance abdominale associée à une lordose lombaire. Le signe de Beever (ascension de l'ombilic lors de l'antéflexion de la tête) traduit un affaiblissement sélectif des muscles abdominaux inférieurs. Au niveau des membres inférieurs, l'atteinte des muscles extenseurs du pied se manifeste par une difficulté à lever le pied. Une faiblesse de la ceinture pelvienne survient plus tardivement. La fonction respiratoire est habituellement normale mais une insuffisance respiratoire restrictive est décrite dans 1% des cas (Wohlgemuth et al., 2004).

Des manifestations extra-musculaires ont été également décrites. Une vasculopathie rétinienne (télangiectasies, micro-anévrysmes) est notée chez 40 à 60 % des patients mais

n'entraîne pas d'atteinte de la vision en général (Padberg et al., 1995). Dans 60% des cas, une surdité sensorielle est mise en évidence (Brouwer et al., 1991; Padberg et al., 1995). Plus rarement, une prédisposition aux arythmies auriculaires, le plus souvent asymptomatique, à été décrite (5% des cas) (Laforet et al., 1998; Galetta et al., 2005; Trevisan et al., 2006). Une atteinte du système nerveux central a été rapportée dans des formes infantiles présentant retard mental et convulsions (Funakoshi et al., 1998). Enfin douleurs et fatigue sont des symptômes fréquemment décrits par les patients (Kalkman et al., 2005).

Les examens complémentaires montrent des signes souvent aspécifiques comme un taux de CPK (Créatine PhosphoKinase) normal à modérément augmenté et des signes de myopathie non spécifiques à l'EMG (ElectroMyoGramme). La biopsie musculaire met en évidence des signes myopathiques non spécifiques, associés éventuellement à un infiltrat inflammatoire et permet surtout d'éliminer d'autres causes de myopathie (Tawil & Van Der Maarel, 2006). Elle n'est pratiquée que rarement.

1.3.2.2. Evolution

La FSHD est une maladie dont l'expressivité intra et inter-familiale est extrêmement variable. Typiquement la maladie débute chez l'adulte jeune au cours de la deuxième décennie. La pénétrance est élevée mais dépend de l'âge et du sexe. Elle est plus importante chez l'homme que chez la femme (Zatz et al., 1998; Tonini et al., 2004). Quatre-vingt-dix pourcent des sujets atteints ont des signes à l'âge de 20 ans. Certains cas de forme infantile sévère ont été décrits alors que certains individus peuvent rester asymptomatiques toute leur vie. L'évolution est lente et continue même si certains patients peuvent présenter une évolution par poussées. L'espérance de vie n'est pas modifiée mais la maladie peut entraîner un handicap important et 20% des sujets sont en chaise roulante après 50 ans. A l'heure actuelle, le traitement est uniquement symptomatique.

1.3.2.3. Biologie moléculaire

Le locus de la FSHD a été identifié dans les années 1990s par analyse de liaison dans la région subtélomérique du bras long du chromosome 4 en 4q35 au niveau de la répétition macrosatellite *D4Z4* (Wijmenga et al., 1990; Sarfarazi et al., 1992; Wijmenga et al., 1992).

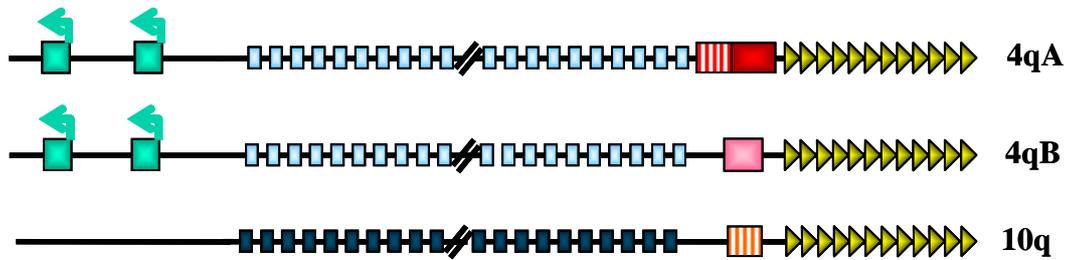
L'anomalie moléculaire responsable est une contraction du nombre de répétitions *D4Z4*. Dans la population générale, le nombre de copies est polymorphe et varie de 11 à 150. Chez 90 à 95 % des patients FSHD, ce nombre est diminué de 1 à 10 unités (**Figure 13**) (Lunt, 1998) et désigne la FSHD1. Plus le nombre de répétitions est faible, plus l'âge de début est précoce et plus la maladie est sévère (Lunt et al., 1995; Zatz et al., 1995). Cependant, il faut souligner quelques particularités.

La FSHD a un fort taux de néomutations d'origine mitotique, aboutissant à un mosaïcisme somatique. Ce mosaïcisme a été mis en évidence chez des parents sains d'enfants porteurs de mutations mais aussi chez des patients *de novo* (van der Maarel et al., 2000). Dans ce cas, les patients de sexe masculin sont plus sévèrement atteints que les patients de sexe féminin (Lemmers et al., 2004a). Le mécanisme de ces réarrangements mitotiques n'est pas clairement établi. Il s'agirait d'une conversion génique impliquant préférentiellement des chromatides sœurs (Lemmers et al., 2004b).

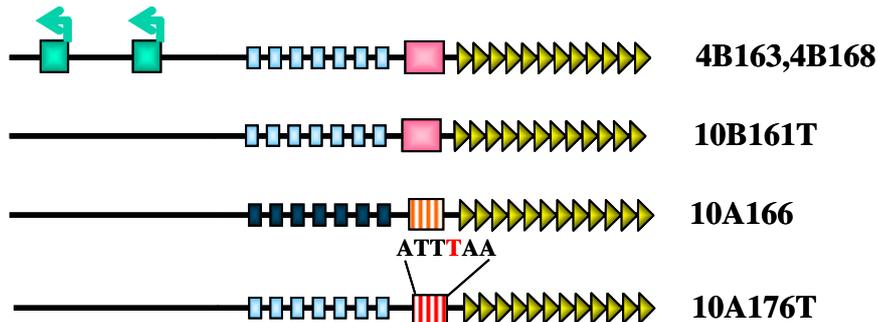
La présence d'au moins une répétition *D4Z4* est nécessaire à l'apparition de la maladie puisque les patients porteurs d'une délétion 4qter n'ont pas de symptômes FSHD (Tupler et al., 1996). Cependant, des délétions proximales emportant *D4Z4* et une région proximale de 40 à 60 kb incluant le locus p13E11 ont été décrites dans 3% des cas (Lemmers et al., 2003; Deak et al., 2007).

La FSHD est spécifique du chromosome 4 puisqu'aucun cas de FSHD n'a été associé au chromosome 10 (en dehors d'une très rare exception sur laquelle nous reviendrons), bien que les répétitions *D4Z4* soient aussi présentes en 10q26. La FSHD survient uniquement sur l'allèle 4qA malgré l'égale fréquence des allèles 4qA et 4qB (Lemmers et al., 2002; Rossi et al., 2007) (**Figure 13a**). L'allèle 4qA possède trois haplotypes principaux 4A161, 4A163 et 4A166. Seules les contractions de *D4Z4* survenant sur l'allèle 4A161, ainsi que sur quelques haplotypes rares (4A161L, 4A159, 4A168) sont associées à la FSHD, ce qui définit ces allèles comme permissifs pour la FSHD. Les contractions de *D4Z4* sur d'autres haplotypes (4A166, 10A166, 10A176T, 4B163) n'auraient pas de conséquence pathologique (Lemmers et al., 2007; Lemmers et al., 2010a). Les allèles permissifs se caractérisent par une séquence identique dans la dernière répétition de *D4Z4* et la région pLAM adjacente qui contient un site de poly-adénylation (ATTAAA) (**Figure 13c**). D'ailleurs, des cas de FSHD secondaires à des contractions survenant sur des allèles hybrides (4q/10q) ont été rapportés. Ces allèles partageaient la même région terminale, y compris chez un patient, où cette région terminale a été transloquée sur l'extrémité du chromosome 10q (Lemmers et al., 2010a).

a : Dans la population générale : 11 à >100 répétitions *D4Z4*



b : Contractions de *D4Z4* (1 à 10 répétitions) sur des haplotypes non permissifs : pas de FSHD



c : Contractions de *D4Z4* (1 à 10 répétitions) sur des haplotypes permissifs : FSHD

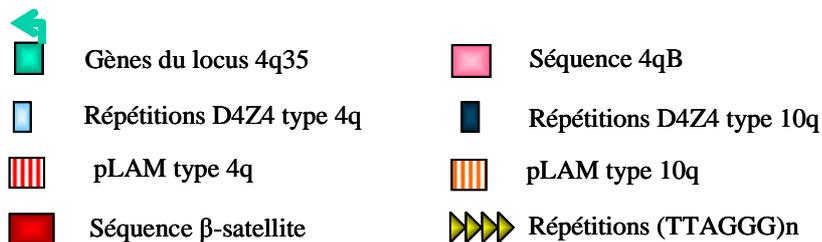
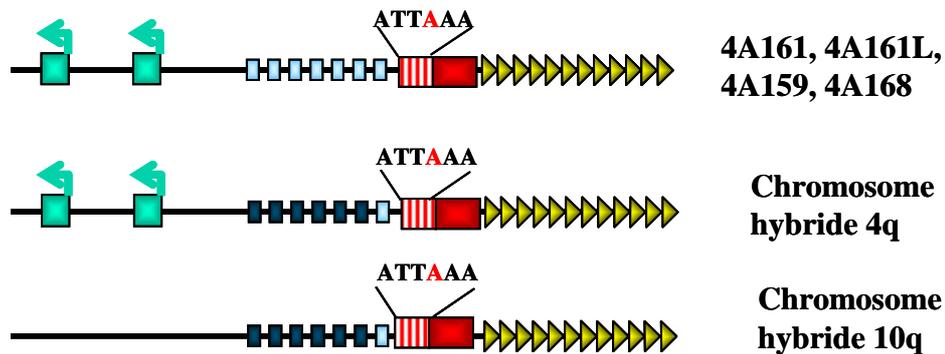


Figure 13 : Mécanisme moléculaire de la FSHD (d'après Lemmers et al., 2010a)

a : Dans la population générale, les répétitions *D4Z4* sont présentes de 11 à plus de 150 copies. Les allèles 4qA et 4qB sont également représentés.

b : La contraction des répétitions *D4Z4* survenant sur un haplotype non permissif n'entraîne pas l'apparition de la maladie. Ces haplotypes ne possèdent pas de signal de poly-adénylation.

c : La contraction des répétitions *D4Z4* survenant sur un haplotype permissif entraîne l'apparition de la maladie. Ces haplotypes partagent des séquences identiques au niveau de la dernière répétition de *D4Z4* et de la région pLAM, notamment un signal de poly-adénylation (ATTAAA).

Les haplotypes non permissifs ne possèdent pas la région pLAM ou une région pLAM dont la séquence est altérée, comme par exemple l'invalidation du signal de poly-adénylation (ATTTAA) par la présence d'un SNP ("Single Nucleotide Polymorphism") sur l'haplotype 10A176T (Lemmers et al., 2010a) (**Figure 13a**).

Enfin, 5% des patients ne présentent pas de contraction de *D4Z4*, définissant la FSHD de type 2 (FSHD2). Toutefois, le phénotype est identique et les patients sont porteurs d'au moins un allèle 4A161 (Krasnianski et al., 2003; de Greef et al., 2010).

1.3.2.4. Diagnostic moléculaire

Le diagnostic moléculaire a recours à l'électrophorèse en champ pulsé (PFGE, "Pulsed Field Gel Electrophoresis") et au Southern-blot et utilise une sonde unique immédiatement adjacente à *D4Z4*, p13E11 (**Figure 12b**). Cette sonde révèle un fragment *EcoRI* de 42 à 300 kb incluant les répétitions *D4Z4* normales et de 10 à 38 kb chez les patients porteurs d'une contraction. Un site de restriction *BlnI* sur les répétitions 10q permet de distinguer l'allèle 4q (résistant) de l'allèle 10q (sensible) (Lemmers et al., 2001). Toutefois, ce diagnostic est souvent d'interprétation délicate. De nouvelles méthodes sont mises au point pour l'améliorer comme le peignage moléculaire qui permet de distinguer les différents allèles 4qA, 4qB, 10q, et leur taille par FISH sur fibres de d'ADN étirées (Nguyen et al., 2010).

1.3.3. Les hypothèses physiopathologiques de la FSHD

Comme la contraction de *D4Z4* ne semble pas perturber la structure d'un gène spécifique, il a été rapidement formulé l'hypothèse que la physiopathologie de la FSHD était liée à un mécanisme épigénétique complexe, notamment une modification de la conformation de la chromatine au locus 4q35 conduisant à une régulation inappropriée de gènes cibles en *cis* et/ou en *trans* (van der Maarel & Frants, 2005; de Greef et al., 2008). Différentes pistes ont été explorées ces dernières années, mais la FSHD n'a pas encore livré tous ses secrets.

1.3.3.1. La méthylation de l'ADN

Une première étude analysant simultanément les chromosomes 4 et 10 avait mis en évidence une hyperméthylation des répétitions *D4Z4* à la fois chez les patients FSHD et les contrôles (Tsien et al., 2001). Plus tard, une autre étude a montré une hypométhylation significative des dinucléotides CpG de *D4Z4* chez des patients FSHD1 comparés aux contrôles (van Overveld et al., 2003). L'hypométhylation est observée chez les patients FSHD1 porteurs d'une contraction de *D4Z4* sur l'allèle 4A161 (de Greef et al., 2009), mais aussi chez les individus porteurs de contraction sur les haplotypes 4q non permissifs ou sur le chromosome 10q. L'état de méthylation de *D4Z4* semble donc dépendre du nombre de répétitions, avec un seuil à partir duquel la méthylation chute fortement. En effet, l'hypométhylation est plus profonde dans un groupe de patients avec une contraction importante (10-20 kb) que dans un groupe de patients avec une contraction moyenne (20-31 kb), dont les symptômes étaient également plus variables. Cette observation a suggérée une corrélation entre la clinique, le nombre de répétitions et le niveau de méthylation (van Overveld et al., 2005). Par ailleurs, les individus atteints de FSHD2 (sans contraction de *D4Z4*) présentent aussi une hypométhylation sur les chromosomes 4q et 10q (van Overveld et al., 2003; de Greef et al., 2009). Dans tous les cas, l'hypométhylation est limitée à *D4Z4* et ne s'étend pas aux régions adjacentes.

L'hypométhylation apparaît donc comme un mécanisme commun mais son lien avec la pathogenèse reste à démontrer.

1.3.3.2. Chromatine et histones

Une première étude de la chromatine au niveau de la région 4q35 avait conclu que *D4Z4* possédait des propriétés d'euchromatine silencieuse plutôt que d'hétérochromatine constitutive (Jiang et al., 2003). Elle avait montré une acétylation de l'histone H4 au niveau de la séquence p13E11 et dans *D4Z4*, ainsi qu'une hyperacétylation des histones des gènes *FRG1* et *ANTI*, sans différence entre les sujets contrôles et les sujets FSHD. Une étude ultérieure confortait ce résultat en recherchant par hybridation *in situ* sur des myoblastes contrôles et FSHD la présence de marque d'hétérochromatine constitutive au niveau de *D4Z4* : triméthylation des résidus H3K9, coloration intense en DAPI, colocalisation avec les foyers d'hétérochromatine constitutive, colocalisation avec la protéine HP1 α , réplication

tardive (Yang et al., 2004). Aucune de ces marques n'était présente au niveau de *D4Z4*, ni chez les patients, ni chez les contrôles. Cependant cette étude a utilisé une sonde localisée de 300 à 800 kb en amont de *D4Z4* pour des problèmes de spécificité (Yang et al., 2004).

Plus récemment, une étude a mis en évidence des remaniements spécifiques des histones chez les patients FSHD (Zeng et al., 2009). Chez les sujets sains, *D4Z4* contient à la fois des marques hétérochromatiniennes, caractérisées par la triméthylation des résidus H3K9 et H3K27, et des marques euchromatiniennes, caractérisées par la diméthylation des résidus H3K4 et l'acétylation de H3. Chez les patients FSHD1 et 2, une perte de la triméthylation des résidus H3K9 à la fois sur les chromosomes 4q et 10q est observée. Par ailleurs, les résidus H3K9 triméthylés permettent normalement de recruter sur *D4Z4* les protéines HP1 γ et des cohésines. Ces protéines favoriseraient l'interaction de *D4Z4* avec d'autres régions du génome, la propagation d'hétérochromatine et la répression de gènes cibles. Ce mécanisme serait aboli dans la FSHD secondairement à la perte de la triméthylation des résidus H3K9 (**Figure 14a**).

1.3.3.3. Complexe répresseur

Un complexe répresseur composé de YY1 (répresseur transcriptionnel), HMGB2 (protéine d'architecture de la chromatine) et nucléoline (protéine de liaison à l'ARN) est capable d'interagir avec *D4Z4* *in vitro* et *in vivo*. Il pourrait médier la répression transcriptionnelle des gènes en 4q35 (**Figure 14b**). En effet, Gabellini et collaborateurs (Gabellini et al., 2002) ont mis en évidence, dans les myoblastes FSHD, une surexpression des gènes en amont de *D4Z4* : *FRG1*, *FRG2* et *ANTI1*. De plus, un modèle de souris surexprimant le gène *FRG1* sous le contrôle d'un promoteur actine spécifique du muscle développe une dystrophie musculaire (Gabellini et al., 2006). Cependant, ce modèle est très contesté puisque plusieurs études n'ont pas confirmé la surexpression des gènes 4q35 proximaux chez les patients FSHD (Jiang et al., 2003; Winokur et al., 2003; Celegato et al., 2006; Osborne et al., 2007) et que les souris présentent un niveau d'expression de *FRG1* quarante fois supérieur aux contrôles.

1.3.3.4. Le modèle de boucle chromatinienne

Un site de fixation à la matrice nucléaire S/MAR (“*Scaffold / Matrix Attachment Region*”), qui permet l’attachement de boucle de chromatine, a été identifié 4 kb en amont de *D4Z4* (Petrov et al., 2006). Dans les myoblastes contrôles, l’association du S/MAR à la matrice nucléaire est très forte et permet de délimiter deux boucles distinctes, l’une contenant *D4Z4*, l’autre les gènes en amont, *FRG1* et *FRG2*. Dans les myoblastes FSHD, l’attachement du S/MAR porté par l’allèle contracté à la matrice nucléaire est plus faible et une seule boucle de chromatine est observée. Un système cellulaire (vecteur pGL3, gène rapporteur luciférase), a démontré que la région 5’ de *D4Z4* se comportait comme un activateur transcriptionnel (“*enhancer*”) qui était capable d’activer la transcription d’un gène rapporteur sous contrôle du promoteur de *FRG1* et que la région S/MAR avait des propriétés de barrière d’activateur (“*enhancer blocking*”) (Petrov et al., 2008).

Un modèle a alors été proposé dans lequel deux boucles de chromatine seraient présentes chez les sujets sains, séparant *D4Z4* des gènes proximaux et réprimant l’activité activatrice de *D4Z4*. Dans la FSHD, la délocalisation du S/MAR de la matrice nucléaire entraînerait la formation d’une seule boucle contenant à la fois *D4Z4* et les gènes proximaux et contribuerait ainsi à la surexpression de *FRG1* et *FRG2* (**Figure 14c**).

Figure 14: Hypothèses physiopathologiques de la FSHD

a : Modèle de propagation de chromatine (d'après van der Maarel & Frants, 2005). Sur l'allèle 4q35 normal, les répétitions *D4Z4* (triangles noirs) et les séquences voisines portent des marques hétérochromatiniennes et les gènes sont réprimés (rectangles rouges). La contraction de *D4Z4* s'accompagne d'un relâchement local de la chromatine (perte de H3K9me3, hypométhylation de l'ADN) dépendant de la distance (dégradé bleu) qui entraîne la dérégulation des gènes en amont (rectangles verts).

b: Complexe répresseur YY1-HMGB2-nucléoline (d'après Gabellini et al., 2002). Sur le chromosome 4qter normal, la liaison de complexes multiprotéiques répresseurs aux répétitions *D4Z4* permet de réprimer l'expression des gènes de la région 4q35. Dans la FSHD, la contraction de *D4Z4* ne permet plus la liaison de complexes répresseurs en quantité suffisante pour assurer la répression des gènes en 4q35. Ces derniers sont alors exprimés de façon inappropriée.

c : Modèle de boucle de chromatine (d'après Petrov et al., 2008). Sur l'allèle 4qter normal, le S/MAR est attaché à la matrice nucléaire et définit deux boucles de chromatine, séparant les gènes de la région 4q35 de l'influence activatrice des répétitions *D4Z4*. Lors de la contraction de *D4Z4*, le S/MAR perd son ancrage à la matrice nucléaire générant une seule boucle chromatinienne contenant *D4Z4* et les gènes de la région 4q35, dont la transcription est alors activée.

Figure 14 : Hypothèses physiopathologiques de la FSHD (suite)

d : Expression de *DUX4* (d'après Mahadevan, 2010). L'allèle 4qter normal possède une chromatine condensée empêchant la transcription de *DUX4*. Une contraction de *D4Z4* survenant au niveau d'un allèle non permissif, dépourvu de signal de poly-adénylation, ne permet pas la transcription de *DUX4*. Une contraction de *D4Z4* survenant au niveau d'un allèle permissif, possédant un signal de poly-adénylation dans la séquence pLAM, est associée à la transcription de *DUX4*, qui serait à l'origine de la toxicité cellulaire.

e : Modèle de l'organisation nucléaire (d'après van der Maarel & Frants, 2005). Les interactions de la région 4qter (flèches noires) avec la lamina nucléaire, qui est associée à la chromatine et des facteurs de transcription (cercles violets), est perturbée dans la FSHD. Cette perturbation conduirait à une redistribution de la chromatine et des facteurs transcriptionnels au niveau de la région 4qter mais aussi dans d'autres régions du génome (carré bleu).

f : Modèle d'insulateur (d'après Ottaviani et al., 2009a). Dans les cellules normales, la multimérisation de *D4Z4* empêche la liaison de CTCF et inhibe l'activité insulatrice. L'expression des gènes cibles, dans la région 4q35 ou dans des régions plus distantes, serait alors réprimée. Dans les cellules FSHD, *D4Z4* se comporte comme un insulateur protégeant l'expression des gènes cibles des influences répressives. Cette activité est dépendante de CTCF et des lamines de type A.

1.3.3.5. A la recherche de gènes candidats

Récemment, différentes méthodes ont été mises en œuvre afin de trouver des gènes cibles ou des régions partenaires impliqués dans la FSHD.

Xu et collaborateurs ont réalisé la cartographie des sites hypersensibles à la DNase I (DH) dans une région de 4 Mb en 4q35 dans des myoblastes contrôles et FSHD (Xu et al., 2009). Les sites DH sont habituellement associés à la chromatine non-nucléosomale et à divers éléments régulateurs. Au niveau de la région 4q35, certains sites DH sont localisés dans ou à proximité de promoteurs de gènes connus comme *FRG1*, *CYP4V2* (famille cytochrome P450), *FAM149A* (protéine hypothétique) ou encore *FAT1*, un gène codant pour une protéine de membrane de type cadhérine impliquée dans différents processus développementaux, et dont le profil DH spécifique aux myoblastes en fait un candidat intéressant. D'autres sites DH sont présents dans des déserts génétiques et peuvent parfois être associés à des séquences courtes répétées de type STR ("*Short Tandem Repeats*"). Enfin, le site DH272 est un candidat intéressant car il est préférentiellement présent dans les cellules FSHD, à proximité de différents transcrits et à cheval sur un site CTCF.

D'autres équipes ont étudié l'expression des gènes dans les cellules FSHD soit par une approche globale comme une puce d'expression type *Affymetrix* (Dixit et al., 2007; Arashiro et al., 2009) soit par une analyse plus ciblée comme le RNAPol-ChIP (immunoprécipitation de chromatine avec des anticorps anti-ARN polymérase II) (Alexiadis et al., 2007), le RNA-DNA FISH (Masny et al., 2010) ou la combinaison de western-blot et de PCR quantitative (Klooster et al., 2009). Les résultats sont assez disparates mais tendent à confirmer l'absence de dérégulation des gènes *FRG1*, *FRG2*, *ANT1* (Alexiadis et al., 2007; Klooster et al., 2009; Masny et al., 2010). Arashiro et collaborateurs ont mis en évidence une dérégulation des gènes *ANKRD37* et *F11* ("*GPI-linked proteins*") chez les sujets atteints de FSHD et une dérégulation des gènes des chimiokines chez les sujets porteurs asymptomatiques de mutations (Arashiro et al., 2009). Par ailleurs, ils suggèrent un rôle pour les miRNA, souvent impliqués dans le mécanisme d'hétérochromatinisation, ainsi que pour les gènes *CXCL9*, *CXCL10*, *CXCL11* localisés en 4q21, qui possèdent un patron d'expression caractéristique chez les patients FSHD.

1.3.3.6. *DUX4* : Enfin la solution ?

Le rôle de *DUX4* a longtemps été controversé. *DUX4* est une ORF conservée chez les rongeurs et les afrothériens (Clapp et al., 2007). Dixit et collaborateurs ont mis en évidence, dans les muscles des patients FSHD, une transcription de *DUX4* associée à une augmentation du facteur de transcription à homéodomaine *PITX1* (Dixit et al., 2007). Ils ont également montré que la protéine *DUX4* peut agir comme activateur transcriptionnel de *PITX1*. Une autre étude a montré l'existence de transcrits partiels et de petits ARN (miRNA) provenant de *DUX4* (Snider et al., 2009).

Plus récemment, un mécanisme d'expression de *DUX4* a été proposé. L'ARN *DUX4* est transcrit à partir de la dernière unité *D4Z4*, et utilise une région 3'UTR, contenant deux introns facultatifs et un site de polyadénylation, qui sont uniquement présents dans la séquence pLAM du chromosome 4, sur l'allèle 4qA (Dixit et al., 2007). Les haplotypes permissifs (4A161, A161L, A159, A168) contiennent un signal de poly-adénylation canonique (ATTAAA) alors que les allèles non permissifs en sont dépourvus (Lemmers et al., 2010a) (**Figure 13b et c**). En conséquence, l'expression d'un transcrit *DUX4* stable de pleine longueur nécessite une copie *D4Z4* adjacente à un signal poly-A permissif, quel que soit le chromosome.

L'expression de *DUX4* est régulée au cours du développement. En effet, la forme longue de *DUX4* est exprimée dans les cellules germinales testiculaires et dans les cellules IPS ("Induced Pluripotent Stem Cells"). Cette expression est perdue au cours de la différenciation et, dans les myoblastes différenciés, seul un faible taux d'une forme courte de *DUX4* persiste. La FSHD correspondrait à une dérégulation de *DUX4* au cours du développement. Un transcrit *DUX4* long a été mis en évidence dans les biopsies musculaires de 5/10 patients FSHD alors qu'il est absent chez les contrôles. De plus, dans les cellules IPS et les myoblastes FSHD différenciés, la protéine *DUX4* longue serait très fortement exprimée dans environ 1/1000 noyaux et serait délétère (Snider et al., 2010).

En effet, la surexpression de *DUX4* est associée à des signes de dystrophie musculaire chez le poisson zèbre et la souris et des anomalies de gastrulation chez le xénope (Bosnakovski et al., 2009; Wallace et al., 2010). Dans un modèle murin, elle entraînerait au niveau cellulaire l'apoptose par un mécanisme dépendant de p53 (Bosnakovski et al., 2009; Wallace et al., 2010), interférerait avec les myogènes MyoD et MYF5 et augmenterait la sensibilité au stress oxydatif (Bosnakovski et al., 2008; Bosnakovski et al., 2009), ce qui est compatible avec les anomalies observées sur les myotubes des patients FSHD (Barro et al.,

2010). De même, la surexpression de *DUX4* dans les cellules de rhabdomyosarcome humain active les caspases 3/7, entraîne la mort cellulaire et altère la distribution de l'émerine à la périphérie nucléaire (Kowaljew et al., 2007).

Un modèle de physiopathologie de la FSHD intégrant *DUX4* a été proposé (**Figure 14d**). Les cellules indifférenciées exprimeraient *DUX4* à partir de la dernière répétition *D4Z4* grâce au signal poly-adénylation dans pLAM. Au cours de la différenciation, l'augmentation de la triméthylation des résidus H3K9 réprimerait l'expression de *DUX4*, ne laissant persister qu'une faible transcription générant une forme courte non toxique par épissage alternatif. Dans la FSHD, la diminution de triméthylation des résidus H3K9 ne permettrait pas la transition vers la forme courte. L'expression de la protéine DUX4 longue persisterait et serait responsable de la toxicité musculaire (Snider et al., 2010).

1.3.3.7. *D4Z4* et périphérie nucléaire

Alors que la plupart des télomères ont tendance à être localisés à l'intérieur du noyau, le télomère 4qter est localisé à la périphérie nucléaire (myoblastes, muscles, fibroblastes et lymphocytes) (Masny et al., 2004; Tam et al., 2004). Le télomère 4q35 reste périphérique qu'il s'agisse d'individus contrôles ou de patients FSHD et aucune différence entre l'allèle normal et contracté n'a été mise en évidence.

Par ailleurs, la localisation de 4qter à la périphérie est conservée chez les primates (Bodega et al., 2007). Elle n'est pas influencée par l'environnement puisque le télomère 4qter est périphérique dans les hybrides somatiques ou dans un cas de translocation t(X;4)(p21;q35) où l'extrémité 4qter permet d'amener une séquence du chromosome X à la périphérie nucléaire (Tam et al., 2004). Masny et collaborateurs ont montré qu'une séquence située 215 kb en amont de *D4Z4* et incluant le microsatellite *D4S139* est plus périphérique que *D4Z4* (Masny et al., 2004). Cette région serait responsable de la localisation périphérique du télomère 4qter et non pas *D4Z4* directement.

Enfin, la localisation périphérique est dépendante des lamines de type A puisqu'elle est abolie dans les cellules déficientes en lamines (Masny et al., 2004). D'ailleurs la région 4qter contient un LAD (Guelen et al., 2008).

Ces résultats sont en faveur du rôle de la périphérie nucléaire dans la FSHD et suggèrent l'hypothèse d'un réseau d'interaction différent à la périphérie nucléaire entre l'allèle normal et l'allèle contracté (**Figure 14e**).

1.3.3.8. FSHD : Gain de fonction de CTCF ?

Dans le but d'étudier les mécanismes épigénétiques impliqués dans la FSHD, notre équipe a mis au point un modèle cellulaire permettant de recréer l'environnement génomique 4q35 présent chez les patients FSHD et les sujets sains. Chaque construction comprend un gène rapporteur *eGFP* associé à une ou plusieurs répétitions *D4Z4* (1 à 12) et des répétitions télomériques (graine de télomère) (§2.2.1.). Ce modèle permet d'analyser l'influence de *D4Z4* sur l'expression du gène rapporteur et a permis de démontrer que *D4Z4* se comporte comme un insulateur dépendant de CTCF et des lamines de type A (Ottaviani et al., 2009a).

En effet, une répétition *D4Z4* se comporte comme un insulateur transcriptionnel empêchant la communication activateur transcriptionnel-promoteur ("*enhancer blocking*") et protégeant aussi contre l'effet de position chromosomique (CPE) (barrière de chromatine) (construction C1X), mais pas ou peu contre le TPE (construction T1X). L'activité insulatrice est présente dans une région de 432 pb (382-814), appelée insulateur proximal, qui protège à la fois contre le CPE et le TPE (constructions C1X Δ B1-3 et T1X Δ B1-3). Par ailleurs, il existe une région distale de 623 pb (2269-2892) se comportant davantage comme un répresseur transcriptionnel, et dont la délétion permet de protéger contre le CPE et le TPE (constructions C1X Δ E et T1X Δ E) (**Figure 15**).

Deux sites de liaison CTCF chevauchant ont été mis en évidence dans la région insulatrice proximale (468-481 et 476-489). La protéine CTCF est capable de se lier à *D4Z4* au niveau de cet insulateur proximal *in vitro* et *in vivo* (retard de migration sur gel, ChIP). Elle est nécessaire à l'activité d'insulation puisque son invalidation par siRNA abolit la fonction insulatrice. Les lamines de type A interagissent également avec *D4Z4* au niveau de l'insulateur proximal et participent à son activité insulatrice (siRNA). La multimérisation de *D4Z4* abolit la protection contre l'effet de position et la liaison à CTCF et s'accompagne d'une augmentation de triméthylation des résidus H3K9. Enfin, un enrichissement en CTCF au niveau de *D4Z4* a été observé dans les myoblastes des patients FSHD mais n'est pas détectable chez les individus non atteints.

En conclusion, la contraction de *D4Z4* permettrait la liaison de CTCF et le changement de fonction biologique de *D4Z4* qui passerait d'un rôle de répresseur à celui d'insulateur, protégeant l'expression de gènes cibles (**Figure 14f**).

Figure 15 : Cartographie des éléments régulateurs de D4Z4

(d'après Ottaviani et al., 2009a)

a : Représentation schématique d'un élément *D4Z4* de la position 1 à 3303 par rapport aux deux sites *KpnI* (K). Les différentes régions de *D4Z4* sont indiquées : *LSau* (position 1-340), région A (position 869-1071), *hhspm3* (position 1313-1780), *DUX4* (position 1792-3063). Les différents sites de restriction utilisés pour le clonage des sous-fragments de *D4Z4* sont indiqués (B : *BamHI* ; B1 : *BlpI* ; F : *FseI* ; E : *EheI*).

b : Les différents fragments obtenus après digestion de *D4Z4* sont clonés en aval du gène rapporteur *eGFP* (constructions "C") ou entre le gène rapporteur et la graine de télomère (constructions "T"). Après transfection du plasmide linéarisé dans des cellules C33-A, le pourcentage de cellules exprimant la GFP est suivi par cytométrie de flux à intervalles de temps réguliers. L'histogramme représente la valeur moyenne du pourcentage des cellules exprimant la GFP de J18 à J29 (% eGFP positive cells), lorsque l'expression de du gène rapporteur atteint le plateau (déviations standard indiquées par la barre d'erreur). Les fragments $\Delta B2-3$ (position 1-382), $\Delta B1-2$ (position 814-1381) et ΔF (position 1549-3303) n'inhibent ni le TPE, ni le CPE. Les fragments $\Delta B1$ (position 1-1381), $\Delta B1-3$ (position 382-814) et ΔE (délétion d'un fragment distal de 623 pb en position 2269-2892) protègent du CPE et du TPE. Les astérisques indiquent les valeurs statistiquement significatives par rapport aux vecteurs contrôles (pCMV et T) (Test t de Student) (* p<0,001; ** p<0,005; *** p<0,05).

Les hypothèses physiopathologiques de la FSHD sont nombreuses. Elles explorent probablement les différentes facettes d'un même processus et ne s'excluent pas les unes des autres. Elles font apparaître que la FSHD met en jeu des modifications de la chromatine associées à un remodelage de l'architecture au locus 4q35, impliquant entre autre boucle de chromatine, insulateur et localisation nucléaire.

2. OBJECTIFS DE L'ETUDE, MATERIELS ET METHODES

2.1. Objectifs de l'étude

Les régions subtélomériques, transition entre les télomères et les séquences spécifiques des chromosomes, sont des régions extrêmement polymorphes et dynamiques dont les fonctions demeurent encore obscures. Du fait de leur proximité avec les télomères, elles sont susceptibles de subir leur influence directe, par exemple en terme de structure chromatinienne. Réciproquement, elles pourraient moduler les fonctions télomériques. Cependant, peu de données sont encore disponibles, en particulier chez l'Homme.

D4Z4 est un élément subtélomérique particulièrement étudié du fait de son implication dans la myopathie facio-scapulo-humérale. Cette pathologie constitue un excellent modèle pour étudier la régulation épigénétique et chromatinienne au niveau des régions télomériques et subtélomériques. Récemment, notre équipe a démontré la fonction insulatrice de *D4Z4* et notamment l'effet barrière contre l'effet de position télomérique d'une séquence proximale au sein de *D4Z4* (Ottaviani et al., 2009a).

Nous avons souhaité poursuivre, avec ce travail, la caractérisation des propriétés fonctionnelles de *D4Z4* en terme de régulation télomérique, en nous intéressant à trois domaines : la compartimentalisation nucléaire, la réplication et l'effet de position télomérique. Tout d'abord, étant donné la localisation nucléaire périphérique du télomère 4qter (Masny et al., 2004; Tam et al., 2004), nous avons étudié les interactions de *D4Z4* avec la périphérie nucléaire et ses conséquences sur l'organisation tridimensionnelle des télomères dans le noyau. Ensuite, nous avons étudié la cinétique de réplication des télomères dans les cellules humaines et l'influence des séquences subtélomériques sur celle-ci. Enfin, nous avons approfondi la caractérisation de la fonction insulatrice de *D4Z4*. Ce travail nous a amené à élargir cette étude à d'autres éléments susceptibles d'interférer avec le TPE, comme la

séquence *D4S139* localisée en 4q35 à 215 kb en amont de *D4Z4* ou encore les transcrits télomériques TERRAs.

Une grande partie de ce travail a été réalisé à partir de modèles cellulaires dont nous décrirons le principe dans le paragraphe suivant. Nous détaillerons également le principe de l'étude de l'organisation nucléaire par 3D-immuno-FISH. Les autres techniques utilisées sont décrites dans les publications présentées dans le chapitre "Résultats" ou en annexe.

2.2. Matériels et méthodes

2.2.1. Modèle cellulaire

Le modèle cellulaire utilisé a été décrit par Koering et collaborateurs (Koering et al., 2002) et Ottaviani et collaborateurs. (Ottaviani et al., 2009a) (**Figure 16a**).

Le vecteur comporte une cassette de sélection correspondant à la fusion du gène de résistance à l'hygromycine au gène suicide de la Thymidine kinase du virus herpes simplex de type I (HSV1), placée sous le contrôle d'un promoteur fort (CMV) et un gène rapporteur exprimant la protéine fluorescente *eGFP* sous le contrôle de ce même promoteur fort. Les constructions T portent une graine de télomère c'est-à-dire des répétitions (TTAGGG)_n organisées en tandem sur une distance de 1,2 kb. Afin d'étudier le rôle insulateur des séquences à analyser, celles-ci sont insérées entre le gène rapporteur et la graine de télomère. Nous avons ainsi testé le rôle d'une à plusieurs répétitions de *D4Z4*, de sous-régions de *D4Z4* ou d'autres séquences d'intérêt qui seront détaillées ultérieurement.

Après linéarisation des vecteurs circulaires au niveau d'un site de restriction *BstXI* localisé en aval des répétitions télomériques, les vecteurs sont transfectés dans les cellules humaines de carcinome cervical C33-A dans des conditions optimisées afin de n'avoir qu'un seul insert par cellule. L'insert s'intègre soit au hasard (construction CMV), soit au niveau d'un télomère formé de *nov*o (construction T) par le processus de "fragmentation télomérique" au cours duquel l'intégration du plasmide induit une cassure chromosomique suivie de l'extension de la séquence télomérique par la télomérase et la perte du fragment chromosomique distal natif (Farr et al., 1992). Trois jours après transfection, des populations ayant intégré stablement le transgène sont sélectionnées en présence d'hygromycine B. La localisation des inserts est vérifiée par hybridation *in situ* en fluorescence sur des étalements de chromosomes métaphasiques. Dans les cellules C33-A, l'efficacité de transfection est importante et le taux de fragmentation télomérique est élevé (85-94 %) (Ottaviani et al., 2009b) (**Table S1**). Aucun site préférentiel d'intégration n'a été mis en évidence (Ottaviani et al., 2009b) (**Figure S2**).

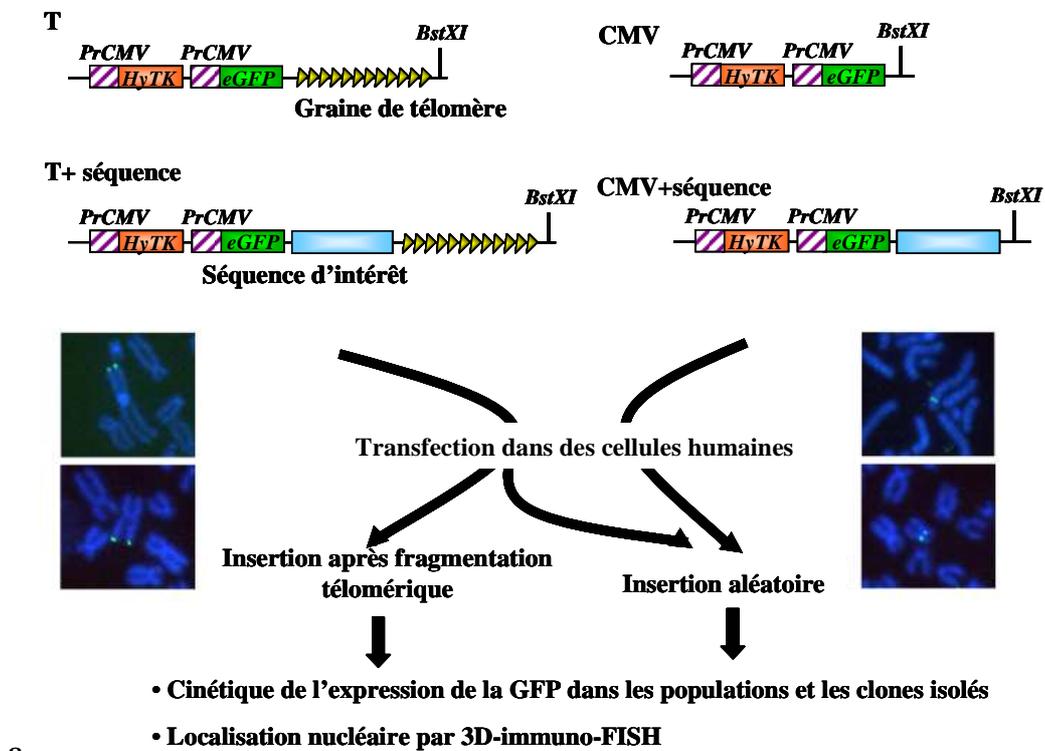


Figure 16 : Matériels et Méthodes

a : Principe du modèle cellulaire utilisé : Les vecteurs comportent un gène de résistance à l'hygromycine fusionné au gène de la thymidine kinase du virus HSV1 (*HyTK*) et un gène rapporteur *eGFP*, placé sous le contrôle de promoteurs forts CMV (*PrCMV*) (construction CMV). La construction T porte une graine de télomère (flèches jaunes). La séquence d'intérêt est insérée en aval du gène rapporteur (construction CMV+séquence) ou entre le gène rapporteur et la graine de télomère (construction T+séquence). Les vecteurs circulaires sont linéarisés au niveau d'un site *BstXI*, localisé en aval de la graine de télomère.

b : Principes de l'analyse de la localisation nucléaire par 3D-immuno-FISH (Ottaviani et al., 2009b). 1 à 4 : Après acquisition en microscopie confocale du signal lamine B (1) et de la sonde ADN (3), un objet tridimensionnel est créé à l'aide du logiciel Imaris (2 et 4, respectivement). 5 : L'objet correspondant à la lamina nucléaire (L) est réduit (1) jusqu'à atteindre le signal FISH. Le rapport des volumes V_L/V_L est alors obtenu de la façon suivante : $V_L/V_L = (d/D)^3$ où D correspond à la distance du centre de l'objet au signal périphérique non réduit et d à la distance du centre de l'objet au signal périphérique réduit.

Le pourcentage de cellules exprimant le gène rapporteur *eGFP* et le niveau d'expression de l'eGFP sont suivis par cytométrie de flux (FACs). La localisation nucléaire de l'insert est déterminée par la technique de 3D-immuno-FISH, décrite dans le paragraphe suivant.

2.2.2. Etude de l'organisation nucléaire en 3D-immuno-FISH

La localisation nucléaire a été déterminée par une technique couplant la détection par immunofluorescence, l'hybridation *in situ* en fluorescence et la reconstitution tridimensionnelle des images obtenues en microscopie confocale (3D-immuno-FISH).

Les cellules, cultivées pendant 24 heures sur des lames en téflon, subissent un prétraitement qui permet de préserver leur architecture nucléaire. Elles sont hybridées soit avec des sondes ADN commerciales (4qter, 10qter, Vysis), soit avec des sondes correspondant au vecteur pCMV ou à des séquences BAC/PAC ("*Phage Artificial Chromosome*") marquées en digoxigénine par déplacement de coupure ("*nick-translation*"). La périphérie du noyau est marquée par immunofluorescence avec un anticorps anti-lamine B (Santa-Cruz). Les images sont acquises sur un microscope confocal (Zeiss LSM510).

L'évaluation de la position radiale du signal est déterminée de la façon suivante (**Figure 16b**). Les noyaux de forme elliptique régulière sont étudiés séparément grâce au logiciel Imaris. Pour chaque noyau, un objet 3D est créé. La limite de la périphérie du noyau (100%) est donnée par la limite extérieure du signal lamine B. La distribution des signaux FISH, correspondant à la séquence d'intérêt, dans le volume nucléaire et le volume ratio sont calculés du centre (0%) vers la périphérie de la sphère après avoir réduit le signal périphérique (limite extérieure du signal lamine B) jusqu'à ce qu'il chevauche le signal FISH, tout en gardant le même centre que celui défini par le cube dans lequel l'objet est inscrit. Il faut noter que le signal Lamine B occupe environ 18% de la périphérie nucléaire.

3. RESULTATS

3.1. Interaction de *D4Z4* avec la périphérie nucléaire

Article 1 : *Identification of a perinuclear positioning element in human subtelomeres that requires A-type lamins and CTCF.* A. Ottaviani, C. Schluth-Bolard et al., EMBO J 2009; 28(16): 2428-2436.

3.1.1. Introduction

La lamina nucléaire est associée à la localisation périphérique de certaines régions chromosomiques et participe ainsi à la compartimentalisation nucléaire, qui joue un rôle important dans la régulation génétique et épigénétique des segments chromosomiques.

La localisation des télomères dans le noyau n'est pas aléatoire, elle varie en fonction des espèces et des types cellulaires. Dans les lymphocytes humains, les télomères sont localisés vers l'intérieur du noyau, mais il existe quelques exceptions comme le télomère 4qter qui est localisé en périphérie. Cependant, peu d'éléments sont connus sur les séquences d'ADN qui pourraient diriger des régions chromosomiques à l'enveloppe nucléaire. Certains subtélomères permettraient d'antagoniser la localisation interne des télomères comme le suggère la présence de LAD à différents subtélomères.

Nous avons exploré les propriétés de *D4Z4*, un macrosatellite subtélomérique de 3,3 kb localisé à l'extrémité du bras long du chromosome 4 (4qter), impliqué dans la FSHD et précédemment décrit comme un insulateur dépendant de CTCF et des lamines de type A, afin de déterminer si cette séquence peut se comporter comme un élément régulateur *en cis* qui permet de diriger les chromosomes vers l'enveloppe nucléaire.

(NB : la référence aux figures des articles joints sera indiqué en italique).

3.1.2. Matériels et méthodes

Les cellules humaines C33-A ont été transfectées avec des constructions contenant 1 à 8 copies de *D4Z4* ou des sous-régions de *D4Z4* (**Figures 1A et 1B**). Des populations stables ont été sélectionnées en présence d'hygromycine B. Le pourcentage de cellules exprimant le gène rapporteur *eGFP* et le niveau de GFP ont été suivis par cytométrie de flux (FACs) à intervalle de temps réguliers. La localisation nucléaire a été déterminée par 3D-immuno-FISH.

3.1.3. Résultats

1) *D4Z4* permet de diriger un subtéломère à la périphérie du noyau

Afin d'évaluer le rôle de *D4Z4* dans la localisation nucléaire des télomères, une construction contenant un élément *D4Z4* entre un gène rapporteur et la graine de télomère (T1X) a été générée et transfectée dans les cellules C33-A (**Figure 1B**).

L'analyse de la localisation nucléaire par 3D-immuno-FISH montre que le télomère 4qter natif est localisé à la périphérie du noyau (**Figures 1C, 1D, 1E**). La construction T1X, comportant une copie de *D4Z4*, est localisée à la périphérie nucléaire alors que la construction contrôle T est localisée de manière plus interne (**Figures 1C, 1D, 1E**). Cette propriété est intrinsèque à *D4Z4* et ne correspond pas à un effet commun aux insulateurs. En effet, une construction dans laquelle la séquence insulatrice 5'HS4 a été insérée entre le gène rapporteur et la graine de télomère (construction T-5'HS4) est localisée vers l'intérieur du noyau, comme la construction contrôle T (**Figures 1B, 1E**).

Afin d'identifier des sous-domaines de *D4Z4* responsables de cette localisation périphérique, des constructions comportant différentes sous-régions de *D4Z4* ont été analysées en 3D-immuno-FISH (**Figure 2**). Quatre constructions (T1X Δ B1, T1X Δ B1-3, T1X Δ E et Tld) sont capables de récapituler à la fois la localisation périphérique et l'activité insulatrice précédemment décrite. Nous avons ainsi identifié un fragment minimal de 80 pb (construction Tld) contenu dans la séquence insulatrice proximale de 432 pb, possédant une activité insulatrice et suffisante pour adresser un télomère à la périphérie (**Figure S3**).

D4Z4 est donc capable de diriger une région subtélomérique à la périphérie du noyau. Cette activité est dépendante, au moins en partie, d'une séquence proximale de 80 pb.

2) Les protéines CTCF et Lamines A contribuent à la localisation périphérique de *D4Z4*

Comme les sous-régions de *D4Z4* permettant la localisation périphérique d'un télomère sont associées à l'activité d'insulation transcriptionnelle de *D4Z4* qui est dépendante des interactions avec CTCF et les lamines de type A, nous avons testé le rôle de ces protéines dans l'adressage périphérique par *D4Z4* en les invalidant par la technique d'interférence à l'ARN (transfection de siRNA) (**Figure S4**). L'invalidation des gènes codant pour CTCF ou les lamines de type A est accompagnée d'une relocalisation du télomère associé à une répétition *D4Z4* (construction T1X) de la périphérie vers l'intérieur du noyau (**Figures 3A et 3B**). Cet effet est spécifique de *D4Z4* puisque la déplétion de CTCF ne perturbe pas l'architecture nucléaire en général. En effet, la localisation des régions 10qter et 4qter natives et celle du territoire du chromosome X ne sont pas modifiées dans les cellules déplétées en CTCF (**Figures 3A, 3C, 3D**).

La localisation d'un télomère à la périphérie nucléaire par *D4Z4* est donc dépendante de CTCF et des lamines de type A.

3) La multimérisation de *D4Z4* entraîne une relocalisation vers l'intérieur du noyau

Nous avons précédemment démontré que la liaison de CTCF à *D4Z4* était perdue lors de la multimérisation de *D4Z4*. Comme la localisation périphérique d'un télomère par *D4Z4* est dépendante de CTCF, nous avons exploré l'effet de la multimérisation de *D4Z4* sur la localisation nucléaire des télomères. L'introduction de 8 copies de *D4Z4* à proximité du télomère (construction T8X) entraîne une relocalisation du télomère vers l'intérieur du noyau (**Figures 4 et S5**). La liaison de CTCF et des lamines de type A au niveau de la séquence insulatrice proximale de *D4Z4* a été étudiée par immunoprécipitation de chromatine (ChIP) à l'aide d'anticorps spécifiques dans les constructions T1X et T8X (1 et 8 copies de *D4Z4*, respectivement). L'enrichissement relatif de l'ADN immunoprécipité par rapport à l'ADN contrôle a été déterminé par PCR quantitative (qPCR) à l'aide d'amorces spécifiques et normalisé sur l'amplification du promoteur de *H4*. De façon concomitante à la perte de la localisation périphérique, les interactions avec CTCF et les lamines de type A sont perdues lors de la multimérisation de *D4Z4* (construction T8X) (**Figure S6**), renforçant l'hypothèse que CTCF interviendrait dans le positionnement nucléaire périphérique.

3.1.4. Discussion

Cette étude montre que *D4Z4* agit *en cis* pour positionner un télomère à la périphérie. Cette activité est dépendante de CTCF et des lamines de type A. Elle est perdue avec la multimérisation de *D4Z4*. Le fait que la localisation du télomère 4qter natif n'est pas affectée par la déplétion de CTCF dans nos modèles cellulaires et qu'il reste périphérique chez les patients FSHD laisse supposer l'existence de deux voies d'ancrage à la périphérie nucléaire (**Figures 5A, 5B**). L'une est indépendante de CTCF et ferait appel à une séquence en amont de *D4Z4*. L'autre est dépendante de CTCF et s'observerait lors de la contraction de *D4Z4*. L'environnement périphérique serait différent entre l'allèle normal et celui porteur d'une contraction, entraînant l'expression différentielle de gènes et l'apparition de la maladie.

D4Z4 possède à la fois une activité d'insulation et de localisation nucléaire périphérique. Il pourrait définir une classe d'éléments insulateurs dépendant de CTCF, qui coopèrent avec les lamines de type A pour localiser une région chromosomique à la périphérie. Ainsi, cette étude renforce l'hypothèse que CTCF jouerait un rôle régulateur global de l'architecture du génome. D'ailleurs un motif consensus de 30 pb capable de lier CTCF *in vitro* a été identifié dans la région proximale de 80 pb de *D4Z4* et au niveau de six autres localisations chromosomiques (**Figure S8, Tableau S2**).

Enfin, les régions subtélomériques sont capables de réguler la localisation des télomères. Chez la levure, elles contrecarrent la tendance des télomères à être à la périphérie pour les positionner vers l'intérieur du noyau. L'effet inverse est observé chez l'Homme où les séquences subtélomériques permettent d'amener un télomère de l'intérieur vers la périphérie (**Figures 5C, 5D**). Dans les deux espèces, cette activité est couplée à l'insulation transcriptionnelle.

En conclusion, ce travail identifie pour la première fois un élément subtélomérique humain qui positionne un télomère à la périphérie nucléaire et suggère un lien entre localisation nucléaire et insulation transcriptionnelle.

3.1.5. Résultats complémentaires

1) *Le motif consensus de 30 pb porteur d'un site CTCF n'est pas suffisant pour localiser un télomère à la périphérie*

Le macrosatellite *D4Z4* possède à la fois la propriété de localiser un télomère à la périphérie nucléaire et la propriété d'insulation transcriptionnelle. Ces deux activités sont récapitulées par une séquence de 80 pb contenant un site CTCF localisée dans la région proximale de *D4Z4*. L'alignement de cette séquence sur le génome humain a permis de définir un motif consensus de 30 pb englobant le site CTCF (**Figure 17d**) et présent au niveau de 6 loci différents qui ne sont pas associés à des répétitions *D4Z4*, sur les chromosomes 5, 8, 12 (bras court et bras long), 17 et 18 (**Figure 17a**). Hormis le chromosome 18 où il est localisé à l'intérieur du bras long, le motif de 30 pb est présent à l'extrémité des chromosomes, à une distance allant de 1,1 Mb à 7,2 Mb du télomère. Ces résultats suggèrent un rôle de ce motif dans l'organisation architecturale du génome.

Tout d'abord, nous avons vérifié si les régions chromosomiques portant le motif de 30 pb étaient localisées à la périphérie du noyau, comme le locus 4q35. Des BACs correspondant aux 6 loci ont été sélectionnés (UCSC, Mars 2006) (CTB-129O4 en 5q35.3 ; CTD-2335L23 en 8p23.1 ; RP11-253E3 en 12p13.3 ; RP11-349K16 en 12q24.3 ; RP11-637C24 en 17q25.3 et RP11-730I16 en 18q12.2) et marqués en digoxigénine par déplacement de coupure ("*nick-translation*"). Après vérification de leur localisation par FISH sur chromosomes métaphasiques, ces sondes ont permis d'étudier la localisation nucléaire de chaque locus par 3D-immuno-FISH. La région subtélomérique 4q35 a été étudiée avec une sonde commerciale (TelVysion 4q SpectrumOrange, Vysis). Un nombre minimum de 50 noyaux a été étudié pour chaque sonde. La région 4q35 ainsi que le locus en 12q24.3 ont une localisation périphérique alors que les autres loci sont localisés vers l'intérieur du noyau (5q35.3, 8p23.1, 12p13.3, 17q25.3, 18q12.2) (**Figure 17b**). Les régions chromosomiques portant le motif consensus de 30 pb ne sont donc pas associées de façon systématique à la périphérie nucléaire.

Les propriétés supposées du motif de 30 pb pouvant être masquées par l'influence des séquences génomiques environnantes, nous avons étudié sa capacité à localiser un télomère à la périphérie nucléaire dans un modèle cellulaire en réalisant une construction dans laquelle la séquence consensus de 30 pb est insérée entre le gène rapporteur *eGFP* et la graine de télomère (construction T+30pb) (**Figure 17c**).

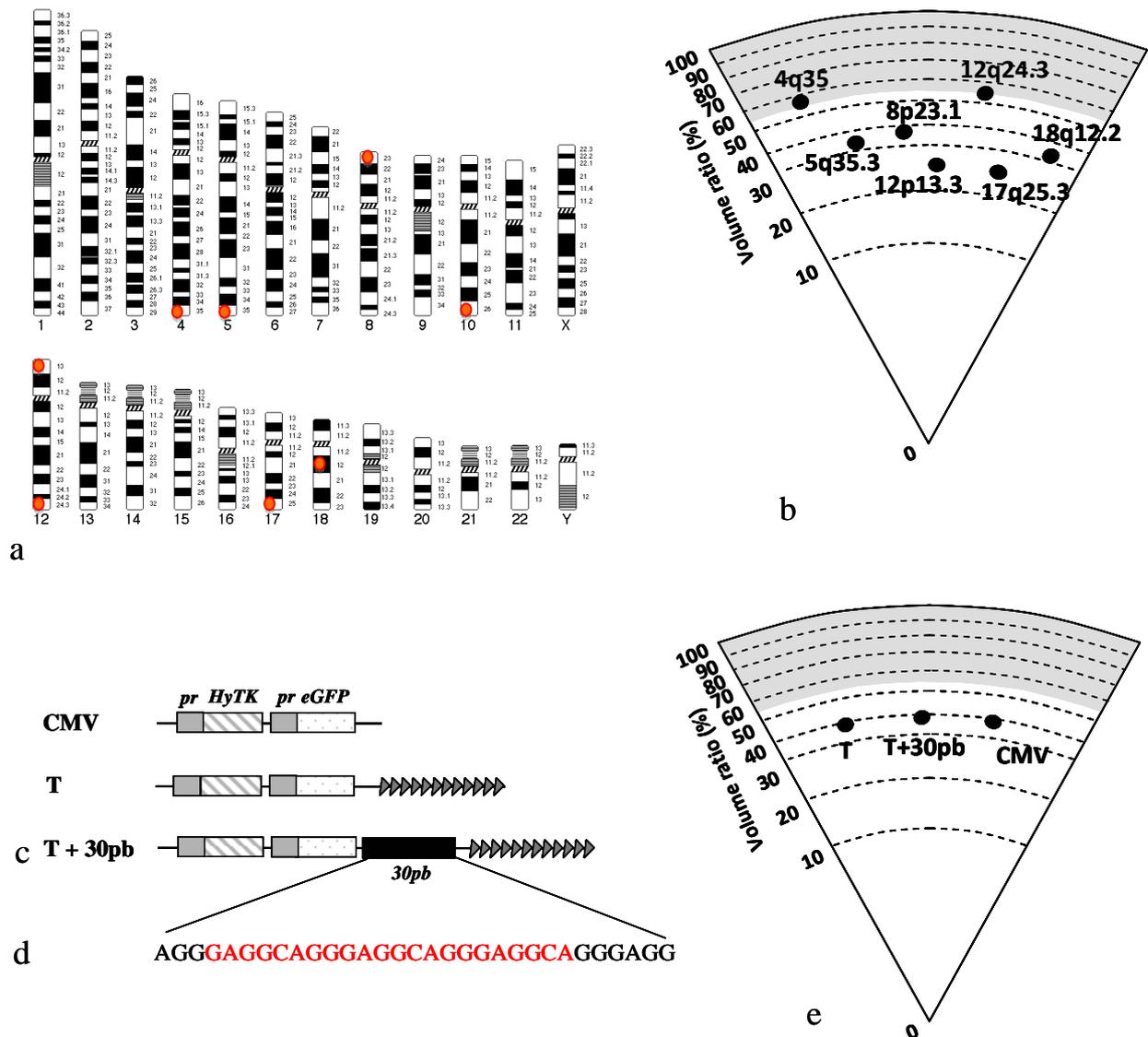


Figure 17 : Rôle de la séquence consensus de 30 pb dans la localisation nucléaire périphérique des chromosomes

a : Localisation chromosomique de la séquence consensus de 30 pb dans le génome. La séquence représentée par un point rouge est localisée au niveau des régions 4q35, 5q35.3, 8p23.1, 10q26.3, 12p13.3, 12q24.3, 17q25.3 et 18q12.2.

b : Localisation nucléaire des régions chromosomiques portant la séquence consensus de 30 pb. La distribution des signaux FISH, après 3D-immuno-FISH, est représentée du centre (0%) vers la périphérie du noyau (100%). Le signal lamine B est représenté par une zone grisée et occupe environ 18% de la périphérie du noyau. Seules les régions 4q35 et 12q24.3 sont périphériques.

c : Représentation schématisée des constructions. Le vecteur comprend une cassette de sélection (*HyTK*) et un gène rapporteur *eGFP*, sous le contrôle de promoteurs forts (*pr*) (construction CMV). La construction T porte une graine de télomère (flèches). Le motif de 30 pb a été inséré entre le gène rapporteur et la graine de télomère (construction T+30pb).

d : Séquence du motif consensus de 30 pb. Le site CTCF est représenté en rouge.

e : Localisation nucléaire de la construction T + 30 pb, déterminée par 3D-immuno-FISH. Le télomère associé au motif de 30 pb occupe une position nucléaire interne.

Après linéarisation, la construction a été transfectée dans les cellules C33-A et la localisation nucléaire du néotélomère associé au motif de 30 pb a été étudiée par 3D-immunofluorescence (FISH) à l'aide d'une sonde complémentaire au vecteur. La construction T+30pb, comportant le motif consensus, a la même localisation nucléaire que la construction contrôle T, qui en est dépourvue, c'est-à-dire à l'intérieur du noyau (**Figure 17e**).

Par conséquent, la séquence de 30 pb seule n'est pas suffisante pour assurer la fonction de localisation nucléaire périphérique. Des séquences adjacentes pourraient être nécessaires à l'adressage périphérique, soit en stabilisant la liaison de CTCF soit en favorisant la liaison de facteurs associés, comme par exemple les lamines de type A. Ces hypothèses pourraient être testées sur les modèles cellulaires par immunoprécipitation de chromatine (ChIP) ou par des tests de liaison *in vitro* de type retard de migration sur gel (EMSA).

2) *La séquence D4S139 ne permet pas de localiser un télomère à la périphérie nucléaire alors qu'elle est périphérique sur le chromosome 4 natif.*

Le fait que le télomère 4qter reste localisé à la périphérie nucléaire chez les patients FSHD suggère l'existence d'une voie d'ancrage de l'extrémité du bras long du chromosome 4 à la périphérie nucléaire indépendante de *D4Z4* qui ferait appel à une séquence en amont de ce macrosatellite. Nous avons tenté d'identifier de telles séquences. Nous avons choisi comme candidat la séquence *D4S139*, un marqueur STS ("*Sequence Tagged Site*") de 150 pb localisé 215 kb en amont de *D4Z4* en 4q35.2 (chr4 : 191005550-191005792) (UCSC Mars 2006) (**Figure 12b**) d'après trois arguments. Tout d'abord, cette région a été décrite par Masny et collaborateurs comme occupant une position plus périphérique dans le noyau que *D4Z4* (Masny et al., 2004). Ensuite, elle appartient au dernier LAD du bras long du chromosome 4 (190367795-191169887, taille 802093 pb) (UCSC Mars 2006) (Guelen et al., 2008), alors que *D4Z4* n'en fait pas partie. Enfin, *D4S139* a été identifiée comme un site hypersensible à la DNaseI de la région 4q35 (région DH13) et pourrait être associée à des éléments régulateurs. (Xu et al., 2009).

Afin d'étudier la capacité de la séquence *D4S139* à localiser un télomère à la périphérie nucléaire, nous avons inséré, entre le gène rapporteur *eGFP* et la graine de télomère, cette séquence de 150 pb (construction T+D4S-150) ou une séquence plus large, de 450 pb, englobant *D4S139* (construction T+D4S-450) (**Figure 18a**).

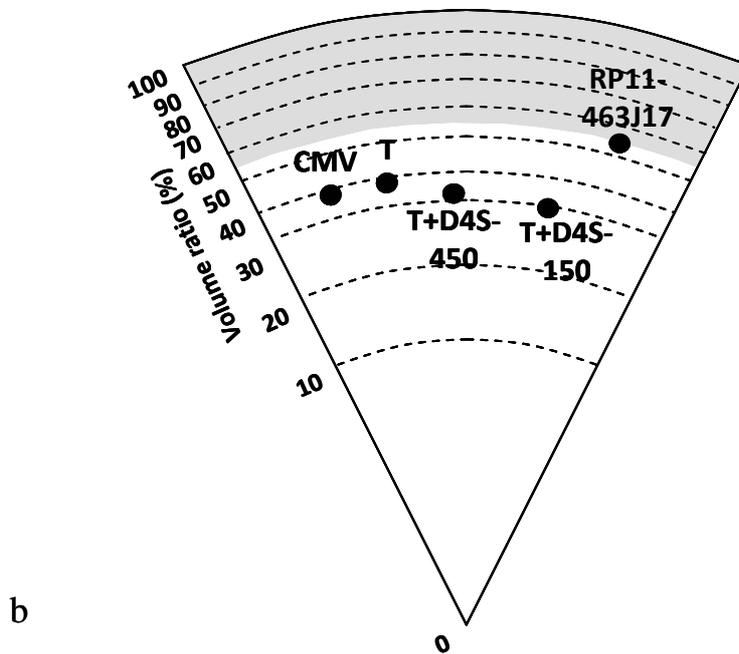
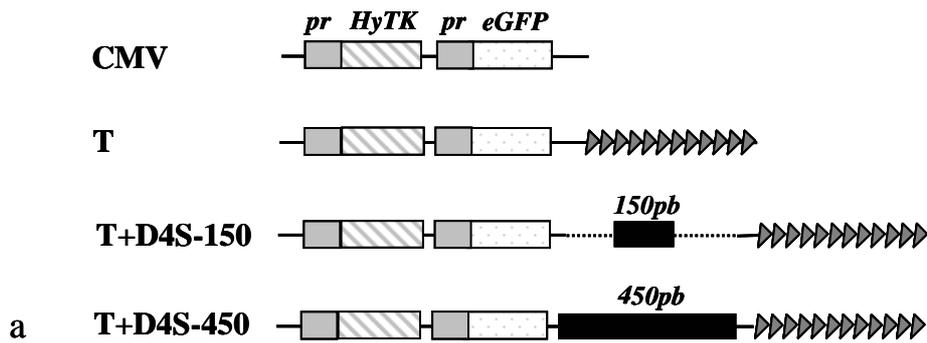


Figure 18 : Etude du rôle de la séquence *D4S139* dans la localisation nucléaire périphérique de la région 4qter

a : Représentation schématique des constructions. Le vecteur comprend une cassette de sélection (*HyTK*) et un gène rapporteur *eGFP*, sous le contrôle de promoteurs forts (*pr*) (construction CMV). La construction T porte une graine de télomère (flèches). La séquence du marqueur STS *D4S139* de 150 pb a été insérée entre le gène rapporteur et la graine de télomère (construction T+D4S-150). Une séquence plus large, de 450 pb, englobant *D4S139* a été insérée dans une autre construction (construction T+D4S-450).

b : Localisation nucléaire des constructions comprenant la séquence *D4S139* (constructions T+D4S-150 et T+D4S-450) et du locus *D4S139* natif en 4q35 (BAC RP11-463J17). La distribution des signaux FISH, après 3D-immuno-FISH, est représenté du centre (0%) vers la périphérie du noyau (100%). Le signal lamine B est représentée par une zone grisée et occupe environ 18% de la périphérie du noyau. Seul le locus *D4S139* natif est périphérique.

Nous avons étudié en 3D-immuno-FISH la localisation nucléaire de ces constructions à l'aide d'une sonde complémentaire au vecteur ainsi que celle du locus *D4S139* sur le chromosome 4 natif grâce à la sonde BAC RP11-463J17 (AF25032.4) (UCSC Mars 2006). Ce BAC de 245 708 pb englobe *D4S139* mais ne contient pas *D4Z4*. Alors que le signal du BAC RP11-463J17 correspondant au locus *D4S139* naturel est périphérique, les signaux des constructions T+D4S-150 et T+D4S-450, portant la séquence *D4S139*, sont localisés vers l'intérieur du noyau (**Figure 18b**).

La séquence *D4S139* ne semble donc pas être directement responsable de la localisation périphérique du télomère 4qter. L'ancrage périphérique de l'extrémité 4qter ferait intervenir d'autres séquences de la région 4q35, isolément ou en coopération, comme d'autres sites CTCF présents dans cette région par exemple. Cette hypothèse est actuellement testée.

3.2. Rôles des séquences subtéломériques dans la régulation de la réplication des télomères

Article n°2 : *Replication timing of human telomeres is chromosome arm-specific, influenced by subtelomeric structures and connected to nuclear localization*. N. Arnoult, C. Schluth-Bolard et al. Plos Genet 2010 ; 6(4) : e1000920.

3.2.1. Introduction

Le contrôle de la réplication des télomères chez les mammifères est encore peu connu. Chez l'Homme, la réplication des télomères se déroule tout au long de la phase S. Une étude récente chez le cerf muntjac a montré que chaque télomère se réplique à un moment précis et que la réplication des bras courts et des bras longs est asynchrone, suggérant que les origines de réplication localisées dans les subtélomères sont soumises à un contrôle spécifique. Le but du présent travail est d'étudier dans les cellules humaines la cinétique de réplication des télomères et d'identifier certains facteurs qui la contrôlent.

3.2.2. Matériels et méthodes

La cinétique de réplication des télomères a été établie grâce à la technique de "*Replicative Detargeting FISH*" (ReDFISH). Les cellules sont synchronisées par un double blocage à l'aide de thymidine et d'aphidicoline. Après la levée du blocage, les cellules sont cultivées en présence d'analogues nucléotidiques, bromodéoxy-uridine et bromodéoxycytidine (BrdU/BrdC) par période d'une heure, couvrant toute la phase S. Le BrdU/ BrdC est incorporé dans les chromatides sœurs en cours de réplication au moment de l'exposition. Les cellules sont ensuite arrêtées en mitose. Le traitement des chromosomes par le colorant de Hoechst et par les UV entraîne la destruction des chromatides qui ont incorporé le BrdU/

BrdC. Ainsi, les télomères répliqués pendant l'exposition deviennent simple brin et peuvent être reconnus soit par une sonde riche en C (signal rouge), soit par une sonde riche en G (signal vert). Les télomères non répliqués ont un signal mixte (jaune) (**Figure 1A, 1B, 1D**). L'analyse des métaphases après chaque période d'exposition au BrdU/BrdC permet de déterminer, par cette méthode, le pourcentage de télomères qui se répliquent à chaque période de la phase S.

3.2.3. Résultats

1) *Les télomères humains se répliquent à des moments préférentiels de la phase S*

La cinétique de réplication des télomères a été déterminée dans les cellules humaines IMR90 (fibroblastes pulmonaires fœtaux) par la technique de ReDFISH. D'une façon générale, les télomères se répliquent tout au long de la phase S avec un pic vers la 4^e heure. La réplication complète d'un télomère individuel est achevée en moins d'une heure. De plus, chaque extrémité télomérique se réplique préférentiellement durant une fenêtre déterminée (précoce/tardive) et la réplication des bras courts et des bras longs est indépendante (**Figure 1**). Par ailleurs, une faible corrélation entre la réplication des télomères et la cinétique de réplication du dernier BAC spécifique du bras chromosomique correspondant a été mise en évidence (**Figure S1**), suggérant un certain degré de coordination entre les télomères et les régions subtélomériques distales.

2) *La longueur télomérique n'influence pas la réplication des télomères humains*

Afin de déterminer les facteurs de régulation de la cinétique de réplication des télomères, nous avons étudié l'influence de la longueur des télomères et de l'expression de la télomérase.

La technique de FISH quantitative (Q-FISH, mesure de l'intensité relative de la fluorescence des télomères) a été couplée à la FISH avec des sondes spécifiques de polymorphismes subtélomériques et la ReDFISH pour déterminer la longueur et la cinétique de réplication de chaque télomère d'une paire de chromosomes homologues donnée. Il a été ainsi observé que des télomères de même longueur présentaient des cinétiques de réplication différentes et que des télomères de longueurs différentes pouvaient avoir des cinétiques de

réplication similaires (**Figure 2**). Il n'existe pas de corrélation entre la longueur des télomères et leur cinétique de réplication.

La stabilisation et l'homogénéisation de la longueur des télomères ont été obtenues par la surexpression des gènes *hTERT* et *TIN2* dans les cellules IMR90 (cellules IMR90-hTERT-TIN2). Bien que la durée de la phase S soit légèrement plus courte, la cinétique de réplication relative n'est pas modifiée (**Figure 3**), renforçant l'observation précédente que la taille des télomères n'a pas d'impact apparent sur leur cinétique de réplication.

3) Le patron de réplication est conservé entre les chromosomes homologues et entre les cellules d'individus différents

Les régions subtélomériques étant extrêmement polymorphes, elles pourraient modifier le départ des origines de réplication ou la progression de la fourche répliative. La caractérisation des polymorphismes segmentaires subtélomériques par FISH à l'aide de sondes spécifiques montre que les extrémités chromosomiques homologues (par exemple les deux allèles 1qter) se répliquent dans la même fenêtre malgré la présence de polymorphismes différents (**Figure 2B**). A l'inverse, les extrémités hétérologues portant le même polymorphisme (par exemple 1p et 7p) ont une cinétique de réplication différente. Ainsi, les polymorphismes subtélomériques ne rendent pas compte des patrons de réplication différentiels.

Par ailleurs, afin d'étudier le comportement des télomères dans les cellules d'un individu différent, la cinétique de réplication des cellules HCA2 (fibroblastes cutanés) surexprimant la sous-unité catalytique de la télomérase hTERT (cellules HCA2-hTERT) a été déterminée. La durée de la phase S et le profil global de réplication des télomères des cellules HCA2-hTERT sont identiques à ceux des cellules IMR90-hTERT-TIN2 (**Figure 4A**). De plus, le temps de réplication moyen d'un télomère donné ne diffère pas entre les deux types cellulaires, y compris pour les extrémités télomériques connues pour porter des variations segmentaires subtélomériques dans les deux lignées cellulaires (**Figures 4B, 4C, 4D**).

Ainsi, la cinétique de réplication ne semble pas être influencée par la présence de polymorphismes segmentaires subtélomériques de grande taille.

4) Les télomères des bras longs des gonosomes ont une réplication tardive

L'hétérochromatine, transcriptionnellement inactive, a tendance à se répliquer plus tardivement, dans la deuxième partie de la phase S. Dans les cellules XX, la réplication des chromosomes X est asynchrone du fait du phénomène d'inactivation de l'un des deux

chromosomes X. Dans les cellules IMR90 (XX), les télomères des bras courts des chromosomes X (Xp) se répliquent de façon synchrone au milieu de la phase S alors que la réplication des télomères des bras longs (Xq) présente une distribution bimodale. Un des télomères Xq se réplique en milieu de phase S et l'autre se réplique en fin de phase S. Ce dernier correspondrait au bras long du chromosome X inactif (**Figure S2**).

Dans les cellules HCA2 (XY), les bras longs des gonosomes Xq/Yq se répliquent de façon plus tardive que les bras courts Xp/Yp. Le profil de réplication du télomère du bras long du chromosome Y pourrait être influencé par la région d'hétérochromatine constitutive du bras long du chromosome Y.

5) *Les télomères porteurs de séquences β -satellites se répliquent tardivement*

Les extrémités 4qter, 10qter et les bras courts des chromosomes acrocentriques font partie des télomères de réplication tardive (**Figure 5**). Ils ont en commun le fait de porter des répétitions de type β -satellite et des répétitions *D4Z4*, ce qui suggère un rôle des répétitions satellites subtélomériques dans le contrôle de la réplication des télomères. Par ailleurs, ces extrémités chromosomiques sont associées à une structure nucléaire spécifique, nucléole ou lamina nucléaire. Ainsi, la localisation nucléaire pourrait également influencer la cinétique de réplication télomérique. Nous avons alors exploré le profil de réplication et la localisation subnucléaire des télomères dans des modèles cellulaires où la composition subtélomérique associée à un télomère néoformé est connue.

6) *La localisation nucléaire périphérique est associée à une réplication plus tardive*

Les cellules humaines C33-A ont été transfectées avec les constructions suivantes : une construction contrôle comportant un gène rapporteur *eGFP* à proximité d'une graine de télomère (construction T), des constructions dans lesquelles une séquence d'intérêt a été insérée entre le gène rapporteur *eGFP* et la graine de télomère : 1 et 8 répétitions *D4Z4* (constructions T1X et T8X, respectivement), une séquence *D4Z4* associée à 4 éléments β -satellites (construction T1X- β -sat) ou 4 éléments β -satellites (construction T- β -sat) (**Figure 6A**). La cinétique de réplication a été étudiée par ReDFISH et la localisation nucléaire par 3D-immuno-FISH.

Les constructions comportant un élément *D4Z4*, T1X et T1X- β -sat, confèrent au télomère néoformé à la fois une localisation périphérique et une réplication plus tardive que les cellules contrôles (construction T) ou les cellules portant 8 copies de *D4Z4* (construction

T8X) (**Figure 6B à 6E**), suggérant que la localisation subnucléaire des télomères influence leur profil de réplication.

Par ailleurs, les télomères néoformés associés à 4 répétitions β -satellites (construction T- β -sat) se répliquent plus tardivement, indépendamment de leur position nucléaire, ce qui suggère un rôle propre des séquences β -satellites dans la régulation de la réplication des télomères.

7) *Les télomères natifs de réplication tardive et non associés à D4Z4 sont localisés à la périphérie nucléaire*

Les résultats précédents suggérant un rôle de la localisation nucléaire dans la réplication télomérique, nous avons étudié la localisation nucléaire de télomères natifs de réplication plutôt précoce (1pter, 5pter, 12pter, 17qter) ou de réplication plutôt tardive (2pter, 3pter, 4qter, 6qter, 12qter) par 3D-immuno-FISH. Les télomères de réplication tardive présentent une localisation nucléaire périphérique (2pter, 3pter, 4qter, 6qter, 12qter) alors que les télomères de réplication précoce (1pter, 5pter, 12pter, 17qter) sont localisés vers l'intérieur du noyau (**Figure 7**). Ainsi, nos résultats suggèrent pour la première fois l'existence d'une corrélation entre le profil de réplication des télomères et leur localisation nucléaire.

3.2.4. Discussion

L'étude du profil de réplication des télomères dans des cellules diploïdes humaines montre que la réplication des télomères se déroule tout au long de la phase S. Cependant, chaque extrémité télomérique possède une cinétique propre avec une fenêtre de réplication préférentielle au cours de la phase S, qui est conservée entre chromosomes homologues et entre les individus.

Le profil de réplication serait déterminé par des éléments spécifiques à chaque chromosome, comme le suggère la corrélation entre la réplication d'un télomère et celle du dernier BAC spécifique du bras chromosomique correspondant. La longueur des télomères et les polymorphismes segmentaires subtélomériques ne semblent pas influencer de façon notable la réplication télomérique. Celle-ci peut cependant être régulée par des séquences de type satellite. Par exemple, les répétitions β -satellites subtélomériques sont associées à une réplication tardive du télomère. La répétition *D4Z4* permet de diriger le télomère qui lui est

adjacent à la périphérie du noyau, ce qui s'accompagne d'un profil de réplication plus tardif. De plus, les télomères de réplication tardive présentent une localisation nucléaire périphérique. Ainsi, notre étude suggère pour la première fois que la localisation nucléaire participerait à la régulation de la réplication des télomères. Des interactions entre la lamina nucléaire et des régions non-téломériques de réplication tardive ont d'ailleurs été décrites (Guelen et al., 2008; Hiratani et al., 2008).

En conclusion, la régulation de la réplication des télomères serait dépendante de particularités spécifiques à chaque chromosome, dont les séquences subtéломériques.

3.3. Influence des séquences subtélomériques sur l'effet de position télomérique

Article n°3 : *D4Z4 as a prototype of CTCF and lamins-dependant insulator in human cells.*

A. Ottaviani, C. Schluth-Bolard et al. *Nucleus* 2010 ; 1(1) : 30-36.

L'effet de position télomérique (TPE) est défini comme le silence transcriptionnel réversible de gènes situés à proximité des télomères. Les mécanismes mis en œuvre dans le TPE chez l'Homme sont encore mal connus. Des études chez la levure ont montré l'importance des séquences subtélomériques dans la régulation du TPE.

Récemment, nous avons montré que le macrosatellite *D4Z4* se comportait comme un insulateur avec les propriétés d'“*enhancer-blocking*” et de barrière de chromatine (Ottaviani et al., 2009a). L'activité barrière contre l'effet de position est récapitulée par une séquence proximale de 432 pb et est dépendante de CTCF et des lamines de type A. Cet insulateur proximal contient une séquence de 80 pb, englobant un site CTCF, qui possède à la fois l'activité barrière de chromatine et la capacité d'adresser un télomère à la périphérie du noyau, par un mécanisme dépendant de CTCF et des lamines de type A (Ottaviani et al., 2009b).

Ce travail suggère que *D4Z4* est un nouveau type d'insulateur dépendant de CTCF et adresse de nouvelles questions comme celle de la nécessité de l'adressage des insulateurs vers un compartiment nucléaire compétent pour leurs fonctions, l'implication des structures nucléaires dans l'activité barrière de chromatine, la capacité de CTCF à coopérer avec d'autres protéines pour former des boucles de chromatine qui permettraient de modeler le génome en différents domaines.

Nous avons poursuivi l'étude des mécanismes d'insulation liés à *D4Z4*, en continuant la caractérisation du rôle du site CTCF proximal, mais aussi en recherchant d'autres sites pouvant coopérer avec celui-ci, soit dans *D4Z4*, soit dans la région 4q35. Ce travail nous a permis d'identifier d'autres séquences permettant de contrer le TPE et d'explorer le rôle des transcrits télomériques dans ce processus.

3.3.1. Rôle du site insulateur proximal de *D4Z4*

1) La séquence consensus de 30 pb a un effet protecteur contre le TPE

Le macrosatellite *D4Z4* possède à la fois la propriété de localiser un télomère à la périphérie nucléaire et la propriété d'insulation transcriptionnelle. Ces deux activités sont récapitulées par une séquence proximale de 80 pb contenant un site CTCF et dans laquelle un motif consensus de 30 pb englobant ce site CTCF a été identifié (Ottaviani et al., 2009b).

Nous avons précédemment montré que cette séquence de 30 pb était insuffisante à localiser un télomère à la périphérie nucléaire (§3.1.5.). Nous avons alors étudié son effet sur le TPE en suivant l'expression du gène rapporteur *eGFP* par cytométrie de flux (FACs) dans des cellules humaines C33-A transfectées avec les constructions précédemment établies c'est-à-dire une construction contrôle comportant le gène rapporteur *eGFP* et la graine de télomère (construction T) et une construction où le motif de 30 pb a été inséré entre le gène rapporteur *eGFP* et la graine de télomère (construction T+30pb) (**Figure 19a**). Nous avons observé une augmentation significative ($p < 0,001$, Test t de Student) du pourcentage de cellules exprimant le gène *eGFP* dans les cellules ayant intégré la construction T+30bp par rapport aux cellules contrôles (construction T).

Le motif de 30 pb possède donc une activité insulatrice. Toutefois, l'activité barrière contre le TPE du motif consensus de 30 pb semble plus faible que celle décrite pour la séquence insulatrice proximale de 80 pb et n'est pas associée à la propriété de diriger un télomère vers la lamina nucléaire. Ainsi, l'action barrière contre le TPE de l'insulateur proximal de *D4Z4* pourrait comporter deux composantes, l'une dépendante du site CTCF et l'autre dépendante des interactions avec la périphérie nucléaire. Nos résultats ne permettent pas de préciser si CTCF et les lamines de type A appartiennent au même complexe au niveau de *D4Z4*, ni si la liaison de *D4Z4* avec les lamines de type A est directe ou non. Cette hypothèse pourrait être explorée en étudiant les interactions de CTCF et des lamines de type A avec le motif de 30 pb par immunoprécipitation de chromatine (ChIP), l'effet de l'inactivation de ces deux protéines (par interférence à l'ARN) sur le TPE et en caractérisant les interactions *D4Z4* – lamina nucléaire.

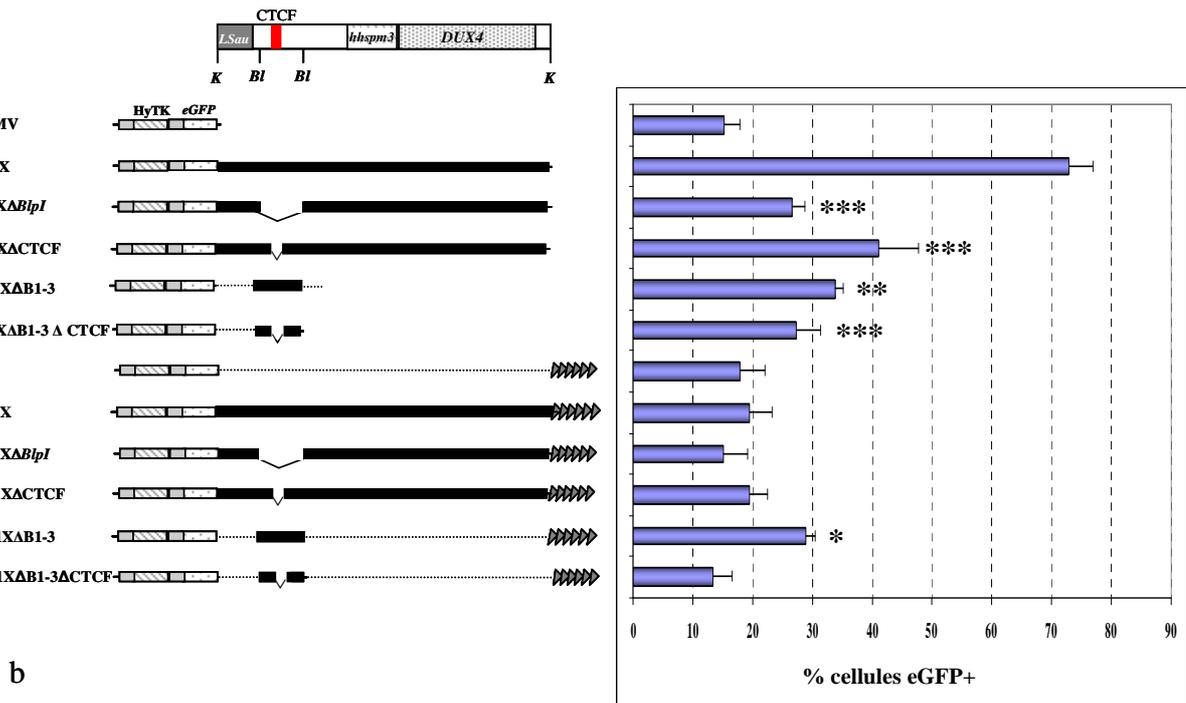
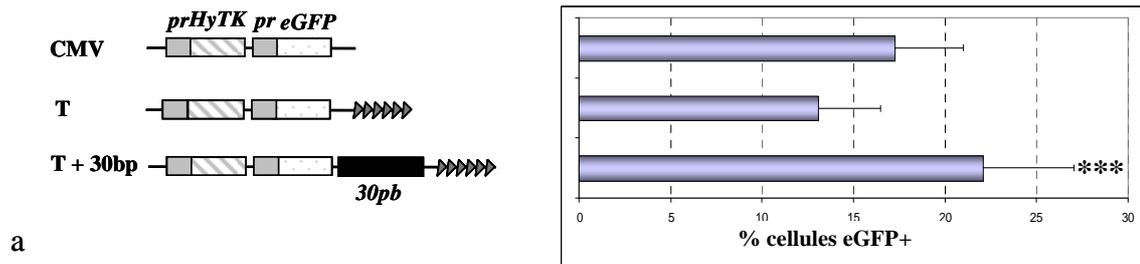


Figure 19 : Rôle du site CTCF de l'insulateur proximal de *D4Z4* sur le TPE

A gauche, représentation schématique des constructions. Le vecteur comprend une cassette de sélection (*HyTK*) et un gène rapporteur *eGFP*, sous le contrôle de promoteurs forts (*pr*) (construction CMV). La construction T porte une graine de télomère (flèches).

A droite, histogrammes représentant la moyenne du pourcentage de cellules exprimant le gène rapporteur *eGFP* (% cellules eGFP+) durant la période où l'expression du gène rapporteur a atteint un plateau stable (entre J16 et J53) (barre d'erreur +/- DS). L'astérisque indique les valeurs statistiquement significatives par rapport au contrôle (construction C ou T) (Test t de Student) * $p < 0,05$; ** $p < 0,005$; *** $p < 0,001$.

a : Le motif consensus de 30 pb a été inséré entre le gène rapporteur et la graine de télomère (construction T+30pb). Une augmentation significative de l'expression du gène rapporteur est observée.

b : Constructions présentant des délétions du site CTCF proximal de *D4Z4*. La séquence *D4Z4* est représentée schématiquement de la position 1 à 3303 par rapport au site *KpnI* (K). Le site CTCF (468-481) proximal est représenté en rouge. Les autres régions de *D4Z4* sont: *LSau* (1-340), *hhspm3* (1313-1780) et *DUX4* (1792-3063). Les sites de restriction *BlpI* (Bl) ont été utilisés pour le clonage et encadrent la séquence insulatrice proximale de 432 pb (382-814). La composition des constructions est détaillée dans le texte.

2) Le site CTCF proximal de *D4Z4* est nécessaire à la protection contre le TPE

Nous avons poursuivi la caractérisation du site CTCF proximal de *D4Z4* dans son activité barrière de chromatine en générant une série de constructions comportant une délétion de ce site CTCF capable de s'intégrer aléatoirement (constructions C) ou en position télomérique (constructions T) (**Figure 19b**). Les constructions comportent une répétition *D4Z4* entière (constructions C1X, T1X), une répétition *D4Z4* comportant une délétion de la séquence insulatrice proximale de 432 pb, localisée entre deux sites de restrictions *BlnI* (position 382-814) (constructions C1X Δ *BlnI* et T1X Δ *BlnI*), une répétition *D4Z4* comportant une délétion du site CTCF proximal (position 468-481) (constructions C1X Δ CTCF et T1X Δ CTCF), la séquence insulatrice proximale de 432 pb (position 382-814) (constructions C1X Δ B1-3 et T1X Δ B1-3) et la séquence insulatrice proximale comportant une délétion du site CTCF (constructions C1X Δ B1-3 Δ CTCF et T1X Δ B1-3 Δ CTCF). Nous avons étudiés l'expression du gène rapporteur *eGFP* par cytométrie de flux (FACs).

Au niveau des constructions intégrées aléatoirement, l'expression de la GFP est augmentée de façon significative dans toutes les constructions, y compris celles comportant une délétion du site CTCF, par rapport aux cellules contrôles (construction CMV), traduisant une activité insulatrice des différentes séquences étudiées. Cependant, cet effet insulateur est plus faible dans les cellules porteuses d'une délétion emportant le site CTCF (constructions C1X Δ *BlnI*, C1X Δ CTCF et C1X Δ B1-3 Δ CTCF) par rapport aux cellules ayant intégré une copie *D4Z4* complète ou la séquence insulatrice proximale intacte (constructions C1X et C1X Δ B1-3, respectivement), comme le montre un niveau d'expression du gène rapporteur *eGFP* diminué. Le site CTCF proximal participe donc à la fonction barrière contre l'effet de position chromosomique (CPE) mais n'est pas indispensable à celle-ci. Ce site CTCF proximal pourrait coopérer avec d'autres séquences présentes au sein de l'élément *D4Z4*.

Au niveau des constructions intégrées en position télomérique, la construction T1X (une répétition *D4Z4*) ne montre pas ou peu d'effet protecteur contre le TPE, comme précédemment décrit. La délétion du site CTCF dans l'élément *D4Z4* (constructions T1X Δ *BlnI* et T1X Δ CTCF) ne modifie pas de façon notable le niveau d'expression du gène rapporteur *eGFP*. Les cellules porteuses de la construction avec la délétion du site CTCF dans l'insulateur proximal (T1X Δ B1-3 Δ CTCF) présentent un niveau de GFP comparable à celui des cellules contrôles (construction T), alors que le niveau de GFP est augmenté dans les cellules T1X Δ B1-3, ayant intégré la séquence insulatrice proximale intacte. Ainsi, l'effet anti-TPE observé avec l'insulateur proximal (construction T1X Δ B1-3) disparaît avec la délétion du site CTCF (construction T1X Δ B1-3 Δ CTCF). Ces résultats suggèrent que le site CTCF est

indispensable à l'effet de protection contre le TPE et que les mécanismes mis en œuvre dans la protection contre le TPE et le CPE ne sont pas identiques.

Par conséquent, ces résultats laissent supposer que le contexte nucléaire et chromatinien dans lequel *D4Z4* exerce ses fonctions a une importance.

3.3.2. Caractérisation d'un site CTCF dans la partie distale de *D4Z4*

L'étude de la séquence insulatrice proximale de *D4Z4* suggère que l'activité barrière de chromatine nécessite la coopération avec d'autres séquences présentes au sein de *D4Z4*.

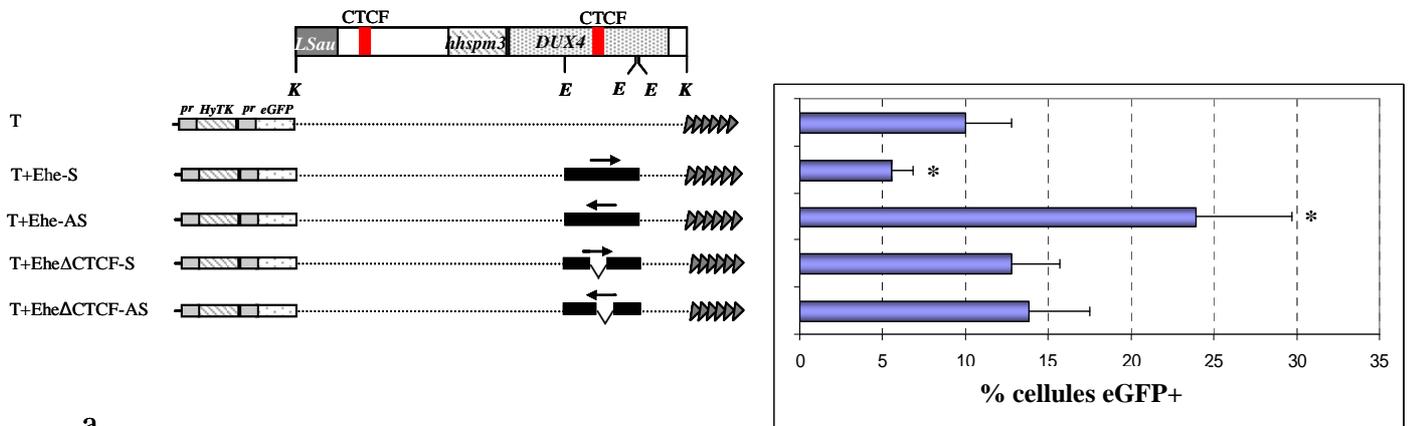
Précédemment, nous avons mis en évidence, dans la partie distale de *D4Z4*, une séquence de 623 pb (encadrée de deux sites de restriction *EheI*) (position 2269-2892) se comportant comme un répresseur transcriptionnel (construction T1XΔF) et dont la délétion permettait de protéger contre le TPE (construction T1XΔE) (Ottaviani et al., 2009a) (**Figure 15**). Cette séquence pourrait donc interférer avec la fonction de l'insulateur proximal. L'analyse *in silico* de cette séquence a révélé la présence d'un site CTCF probable (position 2828-2839, CCGCCTCCGCGCGG). La liaison de CTCF à ce site distal a été confirmée *in vitro* par retard de migration sur gel (EMSA) en incubant des extraits nucléaires issus de cellules C33-A soit avec des oligonucléotides marqués correspondant à cette séquence distale de liaison à CTCF soit avec des oligonucléotides non marqués correspondant à des sites de liaison de CTCF connus (**Figure 2A**) (Ottaviani et al., 2010). Toutefois, aucun enrichissement significatif en CTCF n'a été détecté *in vivo* par immunoprécipitation de chromatine (ChIP) avec des anticorps spécifiques au niveau de cette séquence. L'étude de la localisation nucléaire par 3D-immuno-FISH d'une construction dans laquelle la partie distale de *D4Z4* (position 1549-3303), contenant le site CTCF probable, a été insérée entre le gène rapporteur *eGFP* et la graine de télomère (construction T1XΔF), montre que ce fragment est capable de diriger un télomère à la périphérie nucléaire, même en l'absence de la partie proximale de *D4Z4*. Ainsi, la partie distale de *D4Z4* contribue également à la localisation périphérique d'un télomère (**Figure 2B**) (Ottaviani et al., 2010).

Afin d'étudier la contribution de ce site CTCF distal dans la fonction barrière de *D4Z4*, nous avons réalisé une série de constructions en insérant entre le gène rapporteur et la graine de télomère la région distale de 623 pb contenant le site CTCF distal (position 2269-2892) en orientation sens et anti-sens (T+Ehe-S, T+Ehe-AS, respectivement) et cette même

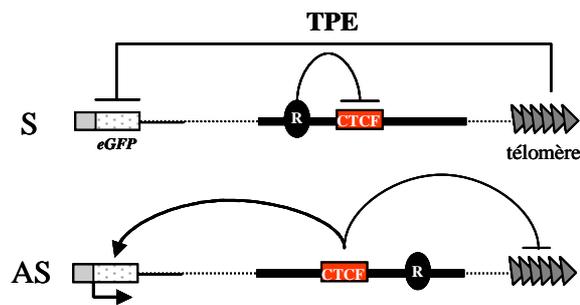
région où la séquence correspondant au site CTCF est délétée, en orientation sens et anti-sens (T+Ehe Δ CTCF-S, T+Ehe Δ CTCF-AS, respectivement). Ces constructions ont été transfectées dans les cellules humaines C33-A et le pourcentage de cellules exprimant le gène rapporteur *eGFP* a été suivi par cytométrie de flux (FACs) sur des populations stables sélectionnées en présence d'hygromycine B (**Figure 20a**).

Nous avons observé que l'insertion de la séquence distale de 623 pb de *D4Z4* est responsable d'une diminution de l'expression du gène rapporteur *eGFP* par rapport aux cellules contrôles (construction T) quand elle est intégrée en orientation sens (construction T+Ehe-S) et qu'elle est responsable d'une augmentation de l'expression du gène rapporteur *eGFP* par rapport aux cellules contrôles (construction T) quand elle est intégrée en orientation anti-sens (construction T+Ehe-AS). L'effet répresseur transcriptionnel (sens) et l'effet activateur transcriptionnel (anti-sens) sont tous les deux dépendants de CTCF, puisque les cellules transfectées avec les constructions portant la délétion de la séquence distale de liaison à CTCF (constructions T+Ehe Δ CTCF-S, T+Ehe Δ CTCF-AS) présentent un niveau d'expression du gène rapporteur *eGFP* semblable à celui des cellules contrôles (construction T) (**Figure 20a**).

Cette observation suggère ainsi l'existence d'une séquence répressive située en amont du site CTCF distal, qui renforcerait l'action du TPE, par exemple en empêchant l'activité du site CTCF 3', comme cela a été précédemment suggéré chez la levure (Defossez et al., 2005). Lorsque la séquence est en orientation anti-sens, le site CTCF est alors entre le gène rapporteur et le répresseur, et serait en mesure de protéger l'expression du gène rapporteur (**Figure 20b**). Cette hypothèse devra être explorée.



a



b

Figure 20 : Caractérisation du site CTCF distal de *D4Z4*

a : A gauche, la séquence *D4Z4* est représentée schématiquement de la position 1 à 3303 par rapport au site *KpnI* (K). Les sites CTCF (468-481 et 2828-2839) sont représentés en rouge. Les autres régions de *D4Z4* sont: *LSau* (1-340), *hhspm3* (1313-1780) et *DUX4* (1792-3063). Le vecteur comprend une cassette de sélection (*HyTK*) et un gène rapporteur *eGFP*, sous le contrôle de promoteurs forts (*pr*) et une graine de télomère (flèches grises) (construction T). La séquence distale de 623 pb délimitée par 2 sites *EheI* (E) (2269-2892) a été insérée entre le gène rapporteur et la graine de télomère : en orientation sens (construction T+Ehe-S) et anti-sens (construction T+Ehe-AS). De même, la séquence distale portant une délétion du site CTCF a été insérée en orientation sens (T+EheΔCTCF-S) et anti-sens (T+EheΔCTCF-AS).

A droite, histogramme représentant la moyenne du pourcentage de cellules exprimant le gène rapporteur *eGFP* (% cellules eGFP+) durant la période où l'expression du gène rapporteur a atteint un plateau stable (entre J21 et J56) (barre d'erreur +/- DS). L'astérisque indique les valeurs statistiquement significatives par rapport aux cellules contrôles (construction T) (Test t de Student) ($p < 0,001$).

b : Proposition d'un modèle expliquant le comportement de la séquence distale de 623 pb de *D4Z4* et du site CTCF distal. En orientation sens (S), le répresseur (R) renforce l'action du TPE sur le gène rapporteur (*eGFP*), en empêchant l'action de CTCF. En orientation anti-sens (AS), le site CTCF est placé entre le répresseur et le gène rapporteur et peut protéger l'expression de ce dernier.

3.3.3. Identification de séquences pouvant interférer avec l'effet de position télomérique

1) *La séquence D4S139 présente un effet protecteur contre le TPE qui dépend de son orientation*

Le macrosatellite *D4Z4* contient une séquence capable d'interférer avec le TPE. Cependant, l'activité insulatrice de *D4Z4* pourrait nécessiter la coopération avec d'autres séquences de la région 4q35. Afin d'identifier d'autres éléments subtélomériques capables d'interférer avec le TPE, nous avons testé l'activité de la séquence *D4S139*, que nous avons précédemment étudiée pour son rôle éventuel dans la localisation périphérique des télomères (§ 3.1.). Ce marqueur STS ("*Sequence Tagged Site*") de 150 pb situé 215 kb en amont de *D4Z4* est localisé à la périphérie nucléaire sur le chromosome 4 (Masny et al., 2004), fait partie d'un LAD (Guelen et al., 2008) et correspond à une région sensible à la DNase I (Xu et al., 2009). Elle pourrait ainsi contenir des éléments régulateurs et influencer l'expression de certains gènes dans la région 4q35 et/ou contribuer au phénotype FSHD.

Nous avons inséré, entre le gène rapporteur *eGFP* et le télomère, la séquence *D4S139* de 150 pb ou une séquence élargie de 450 pb contenant le marqueur *D4S139* en orientation sens (T+D4S-150-S et T+D4S-450-S, respectivement) et anti-sens (T+D4S-150-AS et T+D4S-450-AS, respectivement) (**Figure 21b**). Ces constructions ont été transfectées dans les cellules humaines C33-A et le pourcentage de cellules exprimant le gène rapporteur *eGFP* a été suivi par cytométrie de flux (FACs) sur des populations stables sélectionnées par l'hygromycine B.

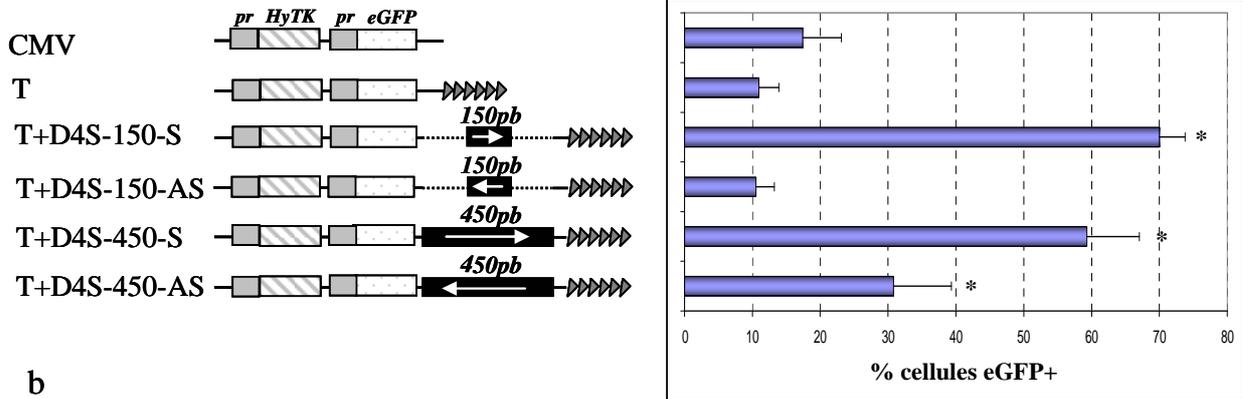
Nous avons observé une augmentation significative de l'expression du gène rapporteur *eGFP* dans les cellules portant les constructions T+D4S-150-S, T+D4S-450-S et T+D4S-450-AS par rapport aux cellules contrôles (construction T). Cette augmentation est beaucoup plus importante dans les constructions où la séquence *D4S139* est en orientation sens (constructions T+D4S-150-S et T+D4S-450-S) (**Figure 21b**). La séquence *D4S139* possède donc la capacité d'interférer avec le TPE et cette propriété dépend de son orientation de 5' vers 3'.

```

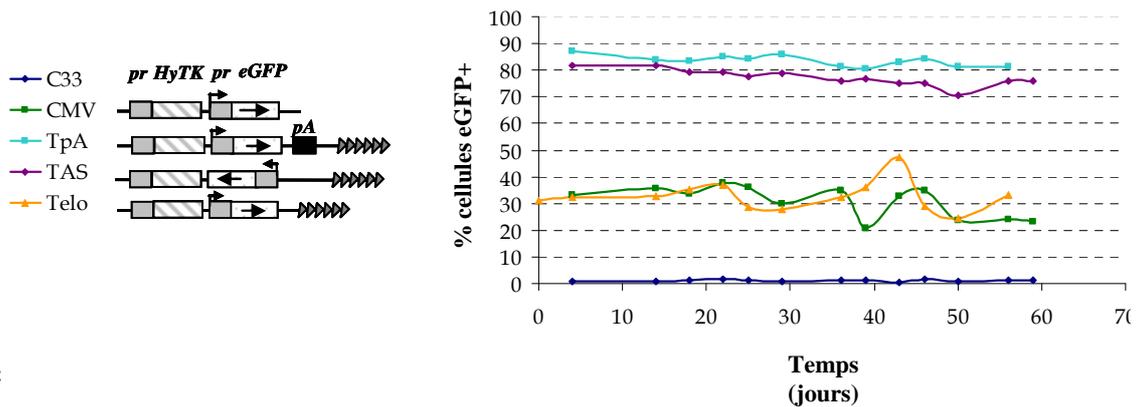
5'aaaatgtgcctctggaatcaaatgatttaaagtctaa
tttttacacttttagtaAATAAAactagtaagaaatgta
ctcttctgtggtctttaaaactcctcttagcaagctattt
ttataaataggctgtcctgactaatgattcttgg 3'

```

a



b



c

Figure 21 : Influence de la séquence *D4S139* sur le TPE

a : Séquence nucléotidique du marqueur STS *D4S139* (UCSC Mars 2006) : présence d'un signal de poly-adénylation canonique (en rouge).

b : A gauche, représentation schématique des constructions. Le vecteur comprend une cassette de sélection (*HyTK*) et un gène rapporteur *eGFP*, sous le contrôle de promoteurs forts (*pr*) (construction CMV). La construction T porte une graine de télomère (flèches). La séquence du marqueur STS *D4S139* de 150 pb a été insérée entre le gène rapporteur et la graine de télomère en orientation sens et antisens (constructions T+D4S-150-S et T+D4S-150-AS). Une séquence plus large, de 450 pb, englobant *D4S139* a été insérée dans une autre construction en orientation sens et anti-sens (constructions T+D4S-450-S et T+D4S-450-AS). A droite, histogramme représentant la moyenne du pourcentage de cellules exprimant le gène rapporteur *eGFP* (% cellules eGFP+) durant la période où l'expression du gène rapporteur a atteint un plateau stable (entre J14 et J49) (barre d'erreur +/- DS). L'astérisque indique les valeurs statistiquement significatives par rapport aux cellules contrôles (construction T) (Test t de Student) ($p < 0,001$).

c : Effet du signal poly-A sur le TPE. A gauche, représentation schématique des constructions. Constructions contrôles CMV et T (cf b); TpA: signal de poly-adénylation inséré entre le gène rapporteur et le télomère; TAS: gène rapporteur orienté en anti-sens. A droite, courbe représentant le suivi du pourcentage de cellules exprimant le gène rapporteur *eGFP* (% cellules eGFP+).

Afin de comprendre le mécanisme sous-jacent, nous avons analysé la séquence du marqueur *D4S139* et nous avons identifié un signal canonique de poly-adénylation (AATAAA) uniquement présent dans l'orientation 5' vers 3' (**Figure 21a**). Nous avons alors inséré un signal de poly-adénylation entre le gène rapporteur et le télomère pour vérifier si ce motif pouvait rendre compte l'effet anti-TPE à lui seul.

2) *La présence d'un signal de polyadénylation fort entre le gène rapporteur et le télomère a un effet anti-TPE.*

Afin de vérifier l'impact du signal de poly-adénylation sur le TPE, nous avons introduit un signal de poly-adénylation fort entre le gène rapporteur *eGFP* et la graine de télomère (construction TpA). Nous avons également réalisé une construction dans laquelle le gène rapporteur *eGFP* et son promoteur ont été insérés en orientation anti-sens (construction TAS), le gène rapporteur étant transcrit du télomère vers le centromère dans cette construction. Ces constructions ont été transfectées dans les cellules humaines C33-A et le pourcentage de cellules exprimant le gène rapporteur *eGFP* a été suivi par cytométrie de flux (FACs) (**Figure 21c**).

Les cellules ayant intégré la construction TpA expriment très fortement le gène rapporteur *eGFP* comparé aux cellules contrôles (construction T) où celui-ci est réprimé par le TPE. Ainsi, la présence d'un signal de poly-adénylation fort permet de contrer les effets du TPE et de maintenir une expression stable et élevée du gène rapporteur.

Les cellules portant la construction TAS (gène rapporteur *eGFP* en orientation anti-sens) ont également un niveau d'expression du gène rapporteur *eGFP* élevé. Il semble donc que le TPE soit dépendant du sens de transcription du gène rapporteur. Il ne s'exerce que lorsque le gène *eGFP* est transcrit du subtélomère vers le télomère. Cette orientation est identique à celle des transcrits télomériques TERRAs, qui sont transcrits à partir du brin C, et dont les sites d'initiation de la transcription sont localisés dans les régions subtélomériques (§1.1.4.2.).

Ces deux observations suggèrent que, dans les cellules ayant intégré la construction T où le TPE est observé, la transcription du gène rapporteur se poursuit dans les répétitions télomériques adjacentes. Ce phénomène serait absent dans les cellules où le gène rapporteur n'est pas réprimé (constructions TpA et TAS), soit du fait de la présence d'un signal d'arrêt de la transcription, soit du fait de l'orientation du gène rapporteur. Le mécanisme du TPE

pourrait donc faire intervenir des transcrits télomériques TERRAs. Nous avons exploré cette hypothèse dans la suite de ce travail.

3.3.4. Exploration du rôle des transcrits télomériques TERRAs dans la régulation du TPE

L'observation que le TPE ne s'exerce que lorsque le gène rapporteur est transcrit dans la même orientation que les TERRAs, associée au fait qu'un signal de poly-adénylation fort abolit le TPE, a conduit à l'hypothèse selon laquelle le TPE pourrait être lié à la transcription des répétitions télomériques.

1) *Mise en évidence d'un transcrit GFP-TERRA*

Afin de caractériser le transcrit *GFP* et de tester l'hypothèse d'une transcription à travers le télomère néo-formé, nous avons choisi une double approche par northern-blot et par transcription inverse suivie de PCR quantitative (RT-qPCR).

L'ARN des cellules T et TpA a été extrait, séparé sur gel dénaturant et transféré sur une membrane en nylon. Trois sondes ont été utilisées pour l'hybridation : une sonde Telo (CCC_nUAA) complémentaire aux répétitions TERRAs (UUAGGG), une sonde GFP complémentaire au transcrit du gène rapporteur et une sonde complémentaire à l'ARNr 18S permettant de normaliser le signal (**Figure 22**). La sonde Telo permet de détecter les ARNs TERRAs totaux, sous la forme d'une traînée allant de 200 bases à plus de 10 kb. Aucune différence en taille ni en quantité n'est notée entre les deux populations cellulaires, T et TpA. Dans les cellules TpA, la sonde GFP montre une forte bande de taille inférieure à 1 kb correspondant à la taille attendue pour le transcrit *GFP*. Dans les cellules T, cette bande est faiblement détectable. Cependant, la présence d'une bande de plus grande taille et d'intensité faible est notée, suggérant l'existence de transcrits hybrides *GFP-TERRAs*, de taille variable.

Une approche complémentaire par RT-qPCR a été entreprise car l'hétérogénéité de taille des transcrits *GFP-TERRAs* pourrait limiter leur détection par northern-blot. La transcription inverse a été réalisée soit avec des amorces poly-dT, complémentaires à tous les transcrits poly-adénylés (*GFP* et une sous-fraction des TERRAs) (Porro et al., 2010) (RT-dT), soit une amorce TR, complémentaire aux répétitions télomériques et spécifique des transcrits

TERRAs (RT-TR). La qPCR permet d'amplifier le transcrit *GFP* grâce à des amorces spécifiques. Les résultats ont été normalisés sur l'amplification du transcrit β -actine. La RT-qPCR ainsi réalisée permet d'amplifier trois populations de transcrits : le transcrit *GFP* poly-adénylé, le transcrit *GFP-TERRA* poly-adénylé et le transcrit *GFP-TERRA* non poly-adénylé (**Figure 23a**). L'amplification des transcrits poly-adénylés (RT-dT) met en évidence un fort taux de transcrits *GFP* dans les cellules TpA alors que celui-ci est indétectable dans les cellules T (**Figure 23b**). Les transcrits *GFP-TERRAs*, sont détectés grâce à l'amorce TR (RT-TR) à la fois dans les cellules T et TpA. Le taux de transcrit *GFP-TERRAs* est très faible comparé à celui des transcrits *GFP* poly-adénylés. Par ailleurs ce taux est cinq fois plus élevé dans les cellules T que dans les cellules TpA.

En conclusion, ces deux approches mettent en évidence l'existence d'un transcrit hybride *GFP-TERRA* comprenant les régions subtélomériques correspondant au gène rapporteur et les régions télomériques. Ce transcrit est présent en quantité plus importante dans les cellules T, où le gène rapporteur est soumis au TPE, comparé aux cellules TpA. Nous nous sommes ensuite interrogés sur la façon dont ce transcrit pouvait participer au mécanisme du TPE. Pour cela, nous avons exploré la stabilité du transcrit *GFP* et l'état de la chromatine au locus *eGFP*.

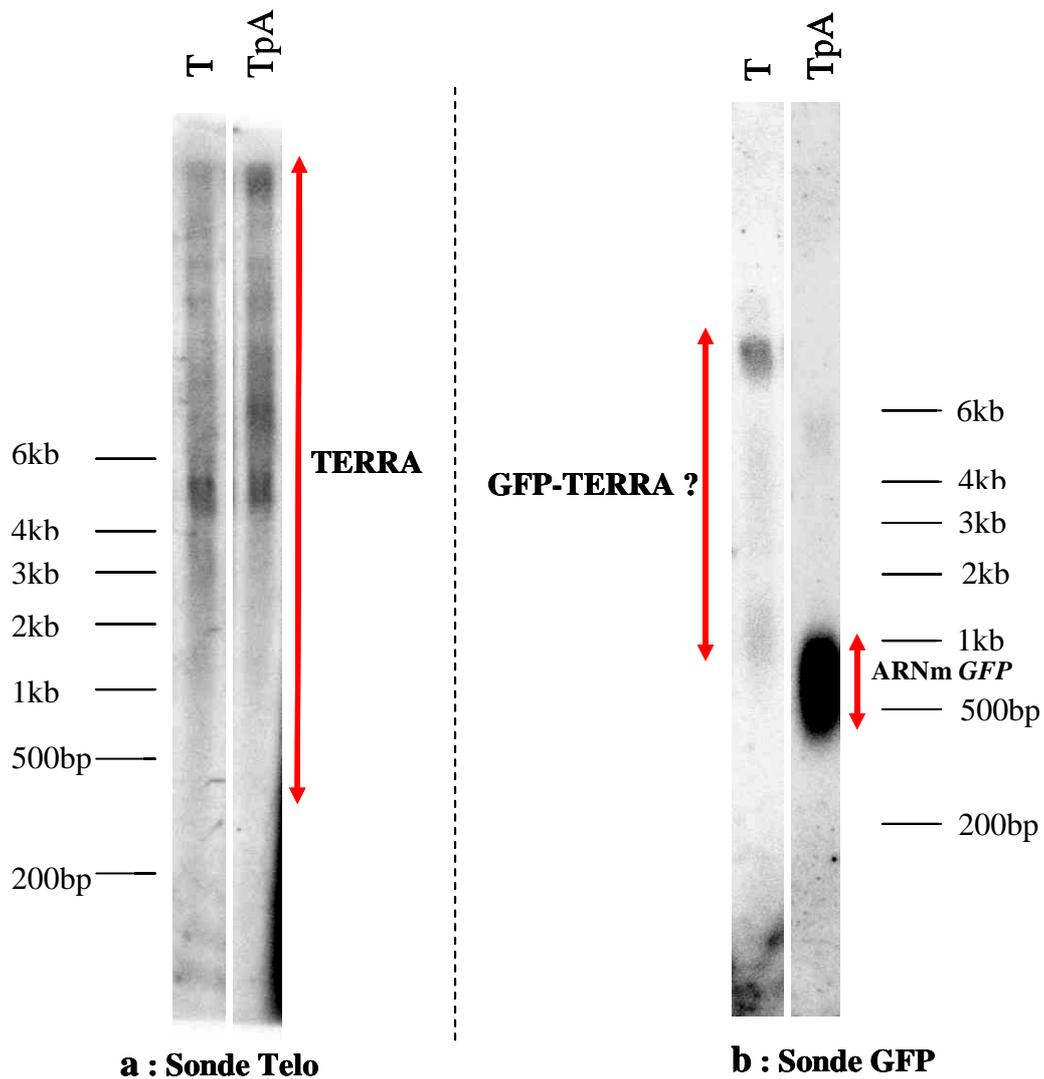
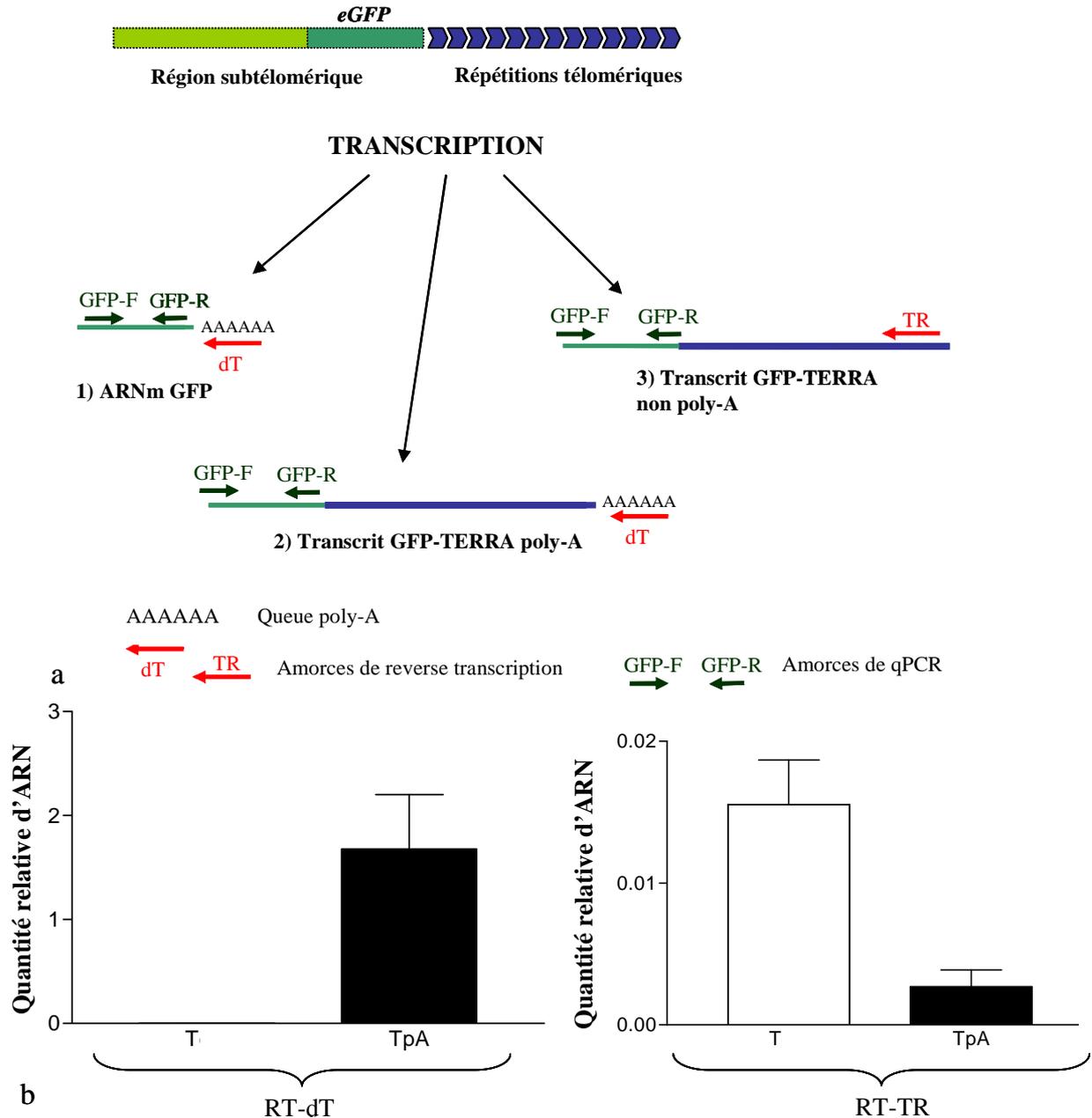


Figure 22 : Mise en évidence d'un transcrit *GFP-TERRA* par northern-blot

L'ARN des cellules T et TpA a été analysé par northern blot.

a : hybridation de la sonde Telo, complémentaire aux répétitions télomériques et montrant les TERRAs totaux (70h d'exposition)

b : hybridation de la sonde GFP, complémentaire au transcrit du gène rapporteur (27h d'exposition). La présence d'une traînée suggère l'existence de transcrits hybrides *GFP-TERRA*s dans les cellules T.



2) La présence du signal poly-A affecte la stabilité du transcrit *GFP*.

Afin de vérifier si l'augmentation d'expression du gène rapporteur eGFP dans les cellules TpA est secondaire à une stabilisation du transcrit *GFP* par le signal de polyadénylation, nous avons évalué la demi-vie de ces ARNs dans les cellules CMV, T et TpA.

Les cellules ont été traitées avec de l'actinomycine D (5µg/mL concentration finale), une drogue qui inhibe l'ensemble de la transcription, et les ARNs ont été récoltés au temps 0 et après 1, 2, 4, 6 et 8 heures d'incubation. Un microgramme d'ARN total a été transféré par la technique de slot-blot sur une membrane et hybridé successivement avec les sondes GFP, Telo (complémentaire aux répétitions TERRAs), β -actine et 18S. Après exposition, l'intensité des taches a été mesurée et normalisée par rapport à l'intensité du signal 18S. L'expérience a été répétée 4 fois de façon indépendante.

La demi-vie des TERRAs est de 5,2 heures dans les cellules TpA et de 6 heures dans les cellules CMV et T (**Figure 24**). Ces résultats sont compatibles avec ce qui a déjà été rapporté (Porro et al., 2010). Dans les trois populations cellulaires, la demi-vie de l'ARNm β -actine est stable. La demi-vie des transcrits GFP est supérieure à 8 h. Cependant la courbe de décroissance est plus importante dans les cellules T que dans les cellules CMV et TpA. La demi-vie évaluée à partir de la droite de régression est de 11,5 h pour les cellules T, 15,5 h pour les cellules TpA et 13,9 h pour les cellules CMV.

En conséquence, la stabilité des transcrits *GFP* est diminuée dans les cellules T par rapport aux cellules TpA. Ceci pourrait être la conséquence d'une dégradation plus rapide des transcrits GFP dans les cellules T, mais aussi refléter la participation des ARNs *GFP*-TERRAs au processus d'hétérochromatinisation par l'intermédiaire de leur dégradation en petits ARNs participant à la voie d'interférence à l'ARN. Cette dernière hypothèse a été explorée, en partie, par l'analyse de la chromatine au locus *eGFP*.

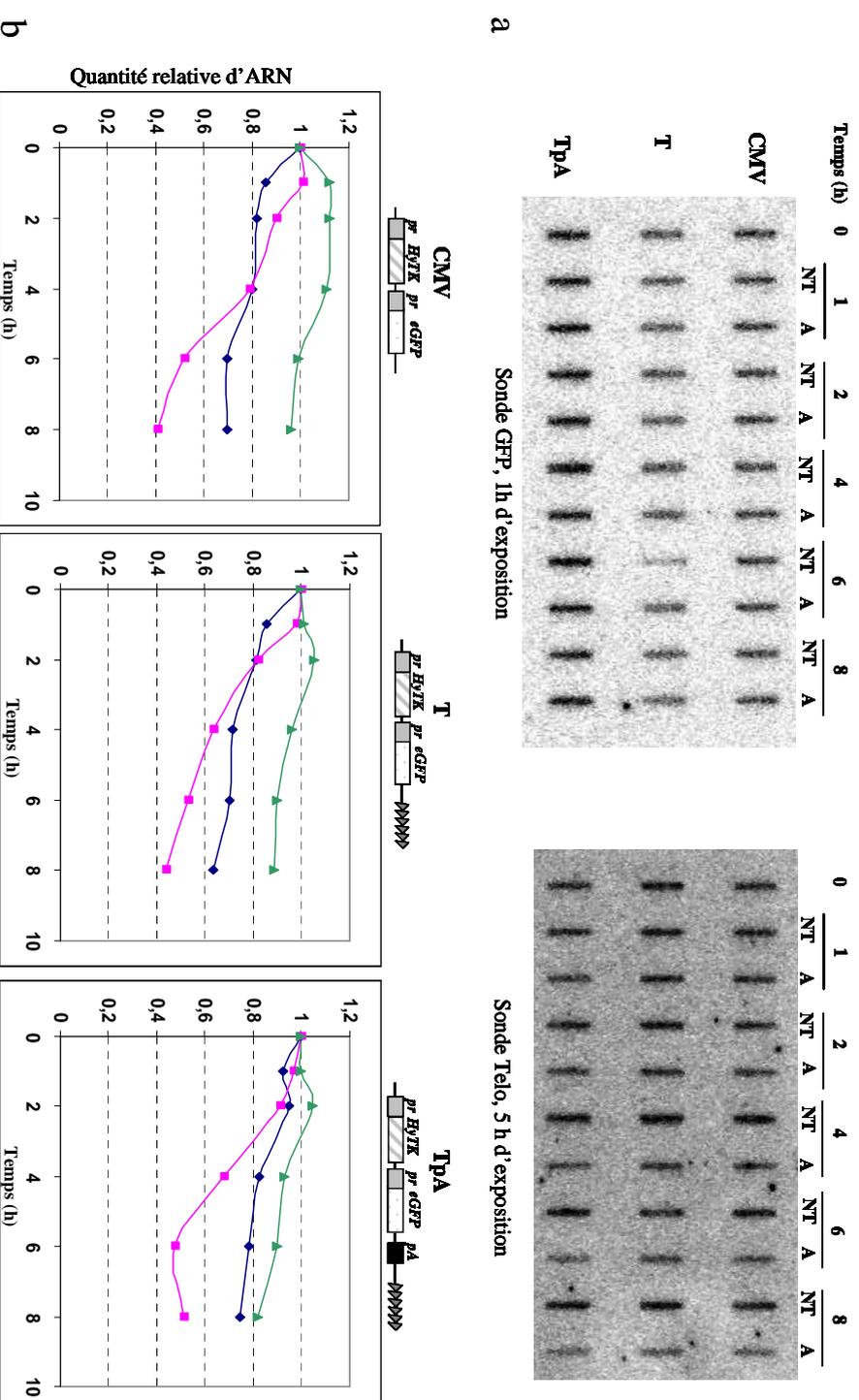


Figure 24 : Le signal de poly-adenylation n'affecte pas la stabilité des transcrits GFP.

La demi-vie des ARNs a été évaluée après traitements des cellules CMV, T et TpA par l'actinomycine D, transfert de l'ARN sur une membrane par slot-blot et hybridation des sondes GFP, Telo, β -actine et 18S.

a : Membrane de slot-blot hybridée avec la sonde GFP et la sonde Telo (détection des TERRAs totaux). Le temps (en heure) d'exposition est indiqué au dessus des membranes. NT : exposition à un milieu contrôle ; A : exposition à l'actinomycine D.

b : Courbes de demi-vie pour les ARNs GFP, TERRAs totaux et β -actine, dans les cellules CMV, T et TpA. Les constructions sont rappelées au-dessus des graphiques (voir figure 21 pour détail). La quantité relative d'ARN a été déterminée après normalisation par rapport à l'intensité du signal 18S. L'ARNm β -actine est stable dans les trois populations cellulaires alors que la demi-vie des TERRAs est de 5,2 h dans les cellules TpA et de 6h dans les cellules CMV et T. La demi-vie du transcrit GFP est de 13,9 h dans les cellules CMV, 11,5 h dans les cellules T et 15,5 h dans les cellules TpA.

3) *La présence du transcrit GFP-TERRA est associée à des modifications de la chromatine au locus GFP.*

Nous avons étudié le profil de modifications post-traductionnelles des histones au niveau du locus *eGFP* et des télomères totaux dans les cellules T et TpA afin de déterminer si l'expression d'un transcrit *GFP-TERRA* s'accompagne de modifications spécifiques de la chromatine.

L'immunoprécipitation de chromatine (ChIP) a été réalisée selon Magdinier et collaborateurs (Magdinier et al., 2004). ADN et protéines ont été fixés *in vivo* en présence de formaldéhyde. Les complexes nucléoprotéiques ont été cassés par sonication et immunoprécipités en présence d'anticorps spécifiques (H3 total, H3K9me3, H3K9Ac, H4K20me3). L'ADN des complexes immuns a été élué. Le locus *eGFP* a été étudié par qPCR avec des amorces spécifiques du gène *eGFP*. Les résultats ont été normalisés par rapport à l'amplification du promoteur du gène codant pour l'histone H4.

Afin d'étudier les modifications chromatiniennes associées aux télomères totaux, la technique de slot-blot a été réalisée. Pour cela, 100 ng d'ADN issus d'immunoprécipitation de chromatine ont été transférés sur une membrane de nylon et hybridés successivement avec la sonde Telo, complémentaire aux répétitions télomériques et marquée radioactivement, et une sonde complémentaire aux répétitions Alu, permettant de normaliser les résultats sur l'ADN génomique total. L'intensité des taches après exposition a été mesurée et normalisée par rapport à l'intensité du signal Alu. L'ensemble des résultats est exprimé en enrichissement par rapport à l'ADN total après traitement à la formaldéhyde et cassure par sonication (*input*). L'expérience a été reproduite trois fois pour les cellules T.

L'analyse de la chromatine au niveau des télomères totaux dans les cellules T et TpA (**Figure 25a**) montre des marques de chromatine répressive, caractérisées par un enrichissement en résidus H4K20 triméthylés, en résidus H3K9 triméthylés et un faible taux d'acétylation des résidus H3K9, conformément à ce qui a été décrit. Le locus *eGFP* présente un profil similaire (**Figure 25b**). Cependant, l'enrichissement relatif en résidus H4K20 triméthylés est plus important dans les cellules T (14 fois) que dans les cellules TpA (10 fois) alors que l'enrichissement relatif en résidus H3K9 triméthylés est plus faible dans les cellules T (1,6 fois dans les cellules T ; 6 fois dans les cellules TpA). Ces résultats doivent cependant être confirmés.

Il semble donc que les cellules T se distinguent par une faible méthylation des résidus H3K9 et une augmentation de résidus H4K20 triméthylés au locus *eGFP* par rapport aux

cellules TpA. Ces modifications pourraient être liées à l'expression du transcrit *GFP-TERRA*. En effet, il a été suggéré un rôle des TERRAs dans la régulation de la chromatine télomérique (Schoeftner & Blasco, 2008).

4) *Discussion et perspectives*

Nous avons mis en évidence la présence d'un transcrit englobant le gène *eGFP* subtélomérique et les répétitions télomériques que nous avons appelé *GFP-TERRA*. La présence de ce transcrit est associée à une faible expression du gène *eGFP*, reflétant le TPE. La présence d'un signal de poly-adénylation fort entre le gène rapporteur et les répétitions télomériques s'accompagne d'une augmentation de l'expression de la GFP et d'une diminution du taux de transcrits hybrides *GFP-TERRA*. Les mécanismes par lesquels le transcrit *GFP-TERRA* participerait au TPE sont encore obscurs. Il semble affecter la stabilité de l'ARN *eGFP* et s'accompagne de modifications d'histones spécifiques (diminution de la méthylation des résidus H3K9, augmentation de résidus H4K20 triméthylés). L'exploration de ce transcrit hybride et de son rôle dans la régulation du TPE se poursuit par l'étude de la chromatine (ChIP), la localisation nucléaire du transcrit (RNA-FISH) et la recherche de voies capables de réguler sa synthèse.

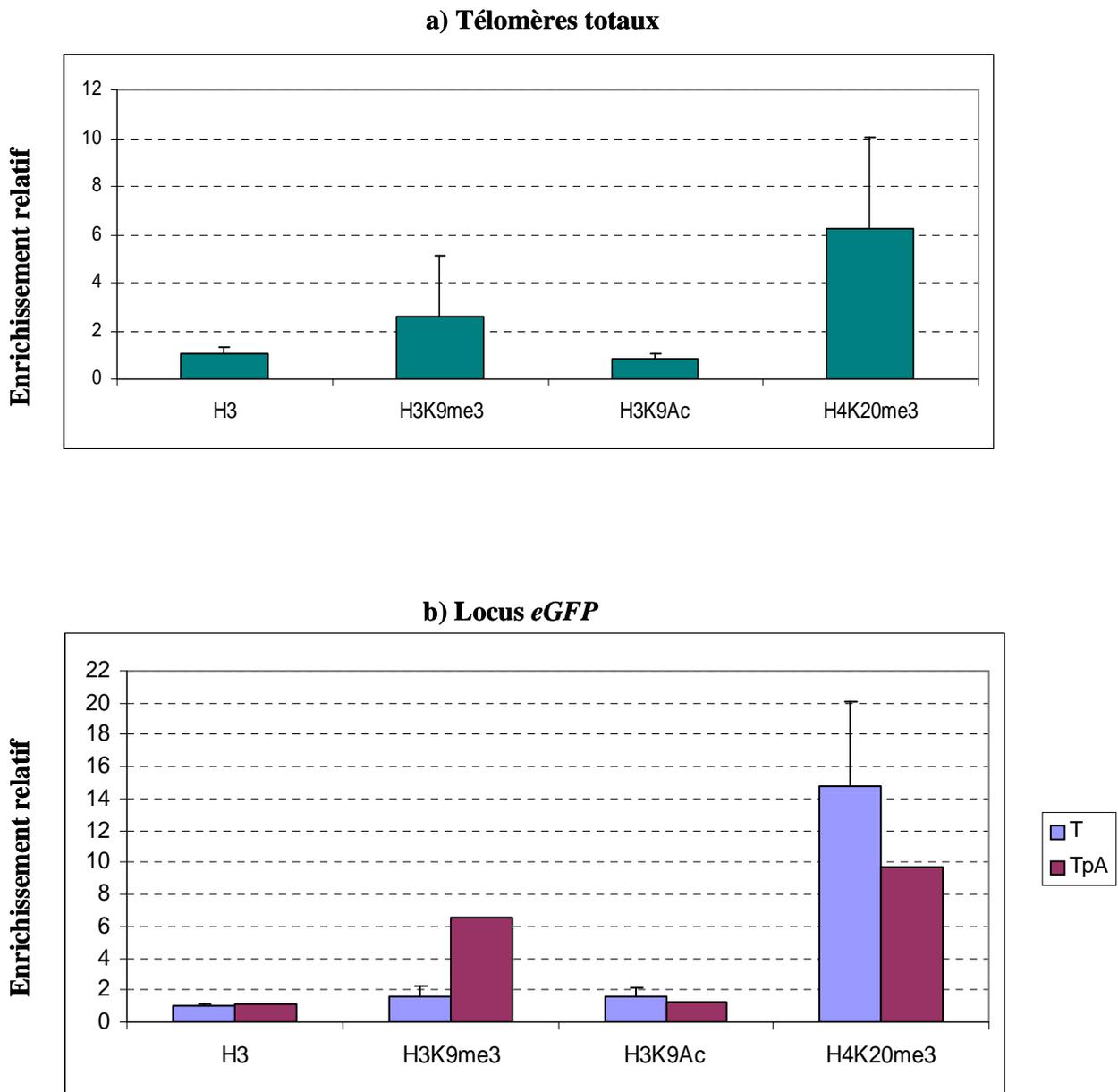


Figure 25 : Etude de la chromatine au niveau des télomères et du locus *eGFP*

Les marques de chromatine suivantes ont été étudiées par ChIP dans les cellules T et TpA: H3: Histone H3 total; H3K9me3: résidu H3K9 triméthylé; H3K9Ac: résidu H3K9 acétylé; H4K20me3: résidu H4K20 triméthylé. Les résultats sont exprimés en enrichissement relatif par rapport à l'ADN *input* (barre d'erreur +/- DS)

a) Etude des télomères totaux: L'ADN est déposé sur membrane par slot-blot et hybridé avec la sonde Telo (TTAGGG)_n (normalisation par rapport à l'intensité du signal Alu).

b) Etude du locus *eGFP* par amplification du gène *eGFP* par qPCR avec des amorces spécifiques (normalisation par rapport à l'amplification du promoteur du gène de l'histone H4).

4. DISCUSSION ET PERSPECTIVES

Au cours de ce travail, nous avons montré que certaines séquences subtélomériques peuvent influencer certaines propriétés biologiques des télomères comme leur localisation nucléaire, leur réplication ou l'effet de position télomérique (TPE). Nous avons plus particulièrement étudié *D4Z4*, un macrosatellite qui est capable de localiser un télomère à la périphérie du noyau et qui contient un élément insulateur dans sa partie proximale pouvant s'opposer au TPE. Cette séquence est par ailleurs impliquée dans la myopathie facio-scapulo-humérale.

4.1. Influence de la lamina nucléaire sur les fonctions télomériques

La séquence macrosatellite subtélomérique *D4Z4* possède la propriété de diriger un télomère à la périphérie du noyau de façon dépendante de CTCF et des lamines de type A. Nos résultats suggèrent que d'autres séquences subtélomériques, non encore identifiées, sont également susceptibles d'influencer la localisation nucléaire des télomères. La périphérie nucléaire est un environnement qui regroupe différentes protéines composant la lamina nucléaire, la membrane nucléaire interne et les complexes de pore nucléaire ainsi que des protéines associées à la chromatine. Elle est ainsi particulièrement propice à de nombreuses interactions.

L'importance de la lamina nucléaire dans la fonction des télomères a été démontrée par l'étude de mutations du gène *LMNA* soit présentes chez des patients atteints de laminopathies, soit induites dans des modèles cellulaires ou animaux. Dans les cellules de patients atteints de progeria, l'attrition télomérique est plus rapide (Allsopp et al., 1992; Huang et al., 2008) et s'accompagne de modification de l'hétérochromatine télomérique avec

une diminution des résidus H3K9 triméthylés et une augmentation des résidus H4K20 triméthylés. D'autres études ont mis en évidence qu'une altération de la lamina contribuait à la sénescence, entraînait une hypermobilité intranucléaire des télomères (De Vos et al., 2010) et une délocalisation des télomères soit vers la périphérie (Raz et al., 2008; Taimen et al., 2009) soit vers les corps PML (Uhlirova et al., 2010). Un modèle de souris déficiente en lamines de type A (*Lmna*^{-/-}) présente des caractéristiques semblables (Gonzalez-Suarez et al., 2009). La perte des lamines de type A est responsable d'un raccourcissement télomérique, d'une relocalisation des télomères vers la périphérie, d'anomalies de la structure chromatinienne télomérique (triméthylation des résidus H3K9 inchangée, perte de résidus H4K20 triméthylés), d'une diminution de la protéine Rb (voie de régulation de Suv4h-20, qui est responsable de la triméthylation de H4K20), d'une diminution des TERRAs et d'une augmentation de l'instabilité génomique (pertes télomériques, cassures chromosomiques, augmentation des foci enrichis en γ -H2Ax TIFs). Enfin, le dysfonctionnement du NHEJ et une déstabilisation de 53BP1, protéine impliquée dans la réparation de l'ADN, ont été observés. Les lamines de type A jouent donc un rôle essentiel dans la maintenance de la longueur, de la distribution nucléaire et de la stabilité des télomères.

Toutefois, nos résultats suggèrent que la relocalisation périphérique des télomères affecte d'autres fonctions telles que la réplication ou le TPE. De façon intéressante, la lamina nucléaire est impliquée dans d'autres processus nucléaires comme l'organisation de la chromatine interphasique, la réplication de l'ADN, la transcription, la prolifération cellulaire, la différenciation et la réparation de l'ADN (Dechat et al., 2008), suggérant plusieurs niveaux de régulation entre lamina et télomères.

Par exemple, la périphérie nucléaire étant un compartiment de régulation transcriptionnelle, trois équipes ont développé des systèmes de relocalisation basé sur la reconnaissance d'un site LacO attaché à un gène rapporteur par un élément lacI fusionné à une protéine de la membrane nucléaire interne (Lap2 β , Δ -EMD, LMNB1) dans des cellules de mammifères (HT1080, NIH3T3, U2OS) afin d'étudier les conséquences transcriptionnelles de la relocalisation des gènes à la périphérie du noyau (Finlan et al., 2008; Kumaran & Spector, 2008; Reddy et al., 2008). La relocalisation d'un locus à la périphérie du noyau nécessite le passage par une mitose, c'est à dire le désassemblage et la reconstruction de la membrane nucléaire (Kumaran & Spector, 2008; Reddy et al., 2008). Deux modèles ont permis de montrer une répression transcriptionnelle, mais pas un silence complet, du gène rapporteur et aussi de certains gènes endogènes au site d'insertion du transgène et donc autour de la zone d'ancrage à la périphérie (Finlan et al., 2008; Kumaran & Spector, 2008). Cette répression est

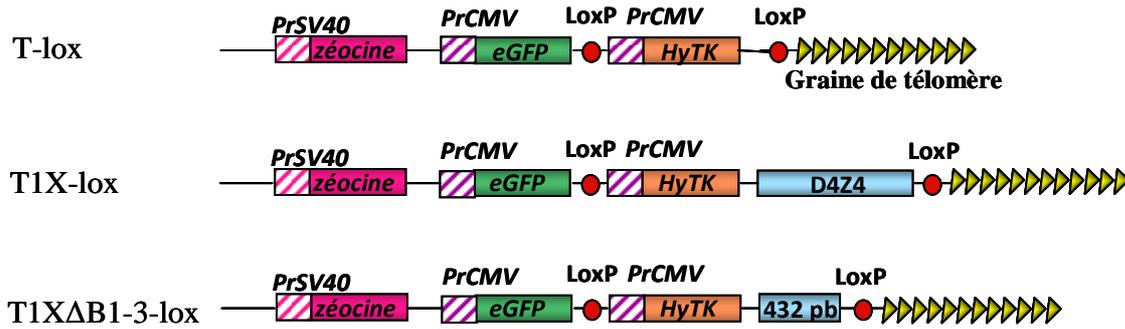
réversible lors de la levée de l'interaction avec la périphérie nucléaire et par le traitement des cellules avec la TSA (un inhibiteur des histones désacétylases), suggérant que le mécanisme ferait intervenir l'hypo-acétylation des histones (Finlan et al., 2008). Kumaran et Spector n'ont pas mis en évidence de différence entre l'activité transcriptionnelle du gène rapporteur périphérique et celle du gène rapporteur non périphérique (Kumaran & Spector, 2008). Bien que basés sur le même principe, les modèles diffèrent par le type cellulaire, les gènes rapporteurs et les promoteurs associés, la protéine de membrane nucléaire utilisée, le système d'induction, le nombre et le lieu d'intégration du gène rapporteur, ce qui peut expliquer les résultats non concordants. Toutefois, ces études montrent que la localisation périphérique d'un gène peut être associée à sa répression transcriptionnelle mais n'est pas obligatoirement incompatible avec son expression.

La lamina nucléaire participe également à la régulation de la réplication chez les mammifères. Chez le hamster chinois, la réplication tardive du locus silencieux de la β -globine est établie en phase G1 en même temps que son repositionnement vers l'hétérochromatine périphérique (Li et al., 2001). De même, au niveau du locus soumis à empreinte *Igf2/H19* chez la souris, l'allèle de réplication tardive présente une localisation périphérique alors que l'allèle de réplication précoce est plus central (Gribnau et al., 2003). Enfin dans les cellules ES en différenciation, la modification de la cinétique de réplication est corrélée au changement transcriptionnel et à la position nucléaire (Hiratani et al., 2008).

La périphérie nucléaire joue un rôle au cours de la différenciation (Joffe et al., 2010), notamment par les interactions avec différentes voies de signalisation comme celles impliquant les MAP kinases, la voie Wnt et β -caténine, le TGF β et Notch (Andres & Gonzalez, 2009). Enfin, la périphérie nucléaire pourrait intervenir dans la régulation de la prolifération et du cycle cellulaire. Il existe des interactions entre les lamines de type A et la protéine Rb hypophosphorylée via la protéine associée à la lamina Lap2 α (Markiewicz et al., 2002) et des interactions entre GCL (Germ Cell Less) et Lap2 β -emerin (Nili et al., 2001; Holaska et al., 2003). Les protéines Rb et GCL sont des répresseurs transcriptionnels agissant par la voie E2F, un facteur de transcription associé à la progression du cycle cellulaire. En conclusion, les rôles de la périphérie nucléaire sont nombreux et encore en pleine exploration (Taddei et al., 2004; Dechat et al., 2008; Andres & Gonzalez, 2009; Joffe et al., 2010).

Afin de préciser le rôle de la périphérie nucléaire sur les fonctions télomériques, nous avons construit une version conditionnelle des constructions T, T1X et T1X Δ B1-3 où soit l'élément *D4Z4* complet soit la séquence insulatrice proximale de 432 pb sont encadrés de deux sites LoxP (T-lox, T1X-lox et T1X Δ B1-3-lox) (**Figure 26**). Après transfection et

intégration de la construction en position télomérique, des clones seront sélectionnés et caractérisés cytogénétiquement par hybridation *in situ* en fluorescence sur métaphase (FISH). Puis, les cellules seront infectées avec un adénovirus exprimant la recombinaise Cre. L'induction de la recombinaison Cre-Lox permettra d'exciser l'élément *D4Z4* de la construction et le repositionnement du gène rapporteur de la périphérie vers l'intérieur du noyau sera étudié par 3D-immuno-FISH. Nous avons déjà établi une collection de 39 clones dans les cellules C33-A (13 clones T-lox ; 14 clones T1X-lox et 12 clones T1X Δ B1-3-lox) qui sont en cours de caractérisation cytogénétique. Ce modèle permettra d'étudier le rôle de la compartimentalisation nucléaire dans la régulation de la chromatine au niveau des éléments *D4Z4*, des régions voisines de la construction et du télomère. Contrairement aux modèles mentionnés précédemment, ce modèle présente l'avantage d'utiliser une séquence naturellement positionnée de manière périphérique. Nous étudierons plus particulièrement la progression du cycle cellulaire, l'expression du gène rapporteur *eGFP*, les modifications post-traductionnelles des histones, la méthylation de l'ADN au site d'intégration ainsi que les conséquences sur la structure télomérique (longueur du télomère, la liaison aux protéines associées). Ce travail apportera des éléments sur la régulation de régions chromosomiques spécifiques par des interactions avec l'enveloppe nucléaire.



1) Transfection dans des cellules humaines C33-A

Caractérisation des clones:

Site d'insertion (FISH sur chromosomes métaphasiques)

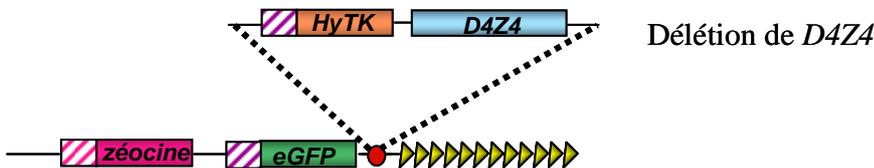
Localisation nucléaire (3D-immuno-FISH)

Expression de la GFP (FACs)



2) Infection des cellules par un virus exprimant la Cre-recombinase

Induction de la recombinaison Cre-LoxP



3) Suivi du repositionnement à l'intérieur du noyau

- Cycle cellulaire
- Modifications post-traductionnelles des histones
- Méthylation de l'ADN
- Facteurs de régulation
- Structure des télomères

Figure 26 : Modèle cellulaire conditionnel

Le vecteur contient un gène de résistance à la zéocine sous le contrôle d'un promoteur SV40 (*PrSV40*), un gène rapporteur *eGFP* sous contrôle d'un promoteur fort (*PrCMV*), deux sites LoxP (point rouge) qui encadrent une cassette de sélection correspondant à la fusion du gène de résistance à l'hygromycine avec le gène suicide de la Thymidine Kinase du virus HSV1 (*HyTK*) sous contrôle d'un promoteur fort (*PrCMV*) et des répétitions télomériques (construction T-lox). La séquence *D4Z4* complète ou celle correspondant à l'insulateur proximal (432 bp) est insérée entre les deux sites LoxP (constructions T1X-lox et T1XΔB1-3-lox). Les cellules transfectées avec ces constructions sont résistantes à l'hygromycine et à la zéocine et sensibles au ganciclovir. La recombinaison Cre-Lox entraîne la délétion de la cassette *HyTK* et de la séquence d'intérêt. Les cellules qui ont subi ce processus sont sélectionnées par leur résistance au ganciclovir.

4.2. Régulation de l'effet de position télomérique chez l'Homme

Bien que des études sur des modèles cellulaires aient suggéré un rôle de l'acétylation des histones, des protéines HP1, de la longueur des télomères, de la distance des gènes par rapport au télomère ou encore de la méthylation de l'ADN subtélomérique (Baur et al., 2001; Koering et al., 2002; Gao et al., 2007; Kulkarni et al., 2010), les mécanismes du TPE chez l'Homme sont encore obscurs. La difficulté à étudier le TPE de télomères naturels sur des gènes endogènes ainsi que l'observation que des gènes adjacents sur le même subtélomère ne présentaient pas la même activité transcriptionnelle (Ning et al., 2003) ont conduit à l'hypothèse que des éléments subtélomériques étaient capables de réguler le TPE, comme cela a été démontré chez la levure (Pryde & Louis, 1999).

Nous avons identifié, dans nos modèles cellulaires, des séquences pouvant contrer le TPE : la présence de *D4Z4* ou d'un signal de poly-adénylation entre une séquence codante et un télomère. L'effet anti-TPE de *D4Z4* est localisé au niveau d'une séquence insulatrice proximale contenant un site de liaison pour la protéine CTCF et dont l'action est dépendante des interactions avec CTCF et les lamines de type A. Le mécanisme d'action du signal de poly-adénylation doit encore être précisé, mais il serait responsable, dans nos modèles cellulaires, de la diminution d'un transcrit englobant le gène rapporteur subtélomérique et les répétitions (TTAGGG)_n télomériques (transcrit TERRA). Nous discuterons des rôles que pourraient jouer les insulateurs et CTCF ainsi que les TERRAs dans la régulation du TPE.

4.2.1. Quel rôle pour les TERRAs dans le TPE ?

Les TERRAs appartiennent au groupe des longs ARNs non codants (lncRNA) qui sont définis comme des ARNs de plus de 200 bases, dont l'ORF est inférieure à 100 amino-acides en longueur (Lipovich et al., 2010). L'abondance de ces ARNs a été mise en évidence lors de l'analyse du transcriptome de la souris. Selon une estimation récente, il existe environ 23 000 lncRNAs dans le génome humain (Carninci & Hayashizaki, 2007). Quatre-vingt-dix pourcent

du génome serait transcrit, réalisant un réseau complexe de transcrits, une même séquence d'ADN pouvant générer des ARNs codants et non-codants, en orientation sens et anti-sens. Les lncRNAs sont extrêmement variés, de même que leurs fonctions et leurs mécanismes d'action. Ils interviennent notamment dans la modification de la chromatine, la régulation de la transcription, la régulation post-transcriptionnelle et la compartimentalisation subnucléaire (Carninci & Hayashizaki, 2007; Wilusz et al., 2009; Lipovich et al., 2010).

L'un des lncRNAs les plus anciennement connus est l'ARN *Xist* qui intervient dans la mise en place de l'inactivation du chromosome X, mécanisme de compensation du dosage génique chez les mammifères. L'inactivation de l'X est sous le contrôle du centre d'inactivation de l'X (XIC), localisé sur le bras long du chromosome X, qui contient de nombreux lncRNAs nécessaires au bon déroulement de l'inactivation : *Xist*, son anti-sens *Tsix*, *Xite* et *DXPas34* qui régulent *Tsix* et l'ARN *Repeat A* (RepA), dans la région 5' de *Xist* (Lee, 2009). Au cours de la différenciation, l'ARN *Xist* est stabilisé sur le futur chromosome X inactif et l'ARN RepA recrute le complexe polycomb PRC2, permettant la couverture du chromosome X par *Xist* et la méthylation des histones (Zhao et al., 2008). Ces premières étapes sont suivies par le silence transcriptionnel du chromosome X inactif. L'expression de *Tsix* est limitée au chromosome X actif et contrôle la répression de *Xist*.

Les lncRNAs sont également impliqués dans le phénomène d'empreinte génomique parentale, un mécanisme qui permet l'expression différentielle d'un gène en fonction de son origine parentale. Par exemple, la région 11p15.5, impliquée dans le syndrome de Beckwith-Wiedemann (MIM#130650), possède deux domaines soumis à empreinte, *IGF2/H19* et *CDKN1C/KCNQ1OT1/KCNQ1*. Les ARNs *H19* et *KCNQ1OT1* participent à la mise en place de l'empreinte dans cette région (Bartolomei et al., 1991; Cai & Cullen, 2007; Pandey et al., 2008).

La régulation transcriptionnelle par les lncRNAs fait appel à différents mécanismes. Dans l'interférence transcriptionnelle, la transcription d'un lncRNA perturbe celle des gènes adjacents. Ce mécanisme a été observé au locus humain *DHFR*, où un lncRNA en amont du gène *DHFR* est transcrit dans le même sens que celui-ci. Sa transcription s'étend sur le promoteur de *DHFR*, induisant la dissociation de la machinerie transcriptionnelle (Martianov et al., 2007). Les lncRNAs peuvent réguler les facteurs de transcription de façon directe (séquestration dans le cytoplasme, activation ou répression pour les ARNs nucléaires) ou indirecte en interagissant avec des protéines co-régulatrices. Par exemple, le lncRNA *Evf-2* forme un complexe avec la protéine *Dlx2* qui peut se comporter alors comme un activateur, régulant la transcription des gènes *Dlx5/6*, candidats dans le syndrome mains-pieds fendus

(MIM %183600) (Feng et al., 2006). Certains lncRNAs peuvent agir en trans, comme *HOTAIR*, transcrit en orientation anti-sens à partir du locus *HOXC* et réprimant la transcription au locus *HOXD* par le recrutement du complexe polycomb PRC2 et l'induction d'une chromatine répressive (Rinn et al., 2007).

La régulation post-transcriptionnelle affecte le métabolisme des ARNs. Par exemple, un lncRNA issu de la transcription en orientation antisens du locus *Zeb2* est capable de former un duplex d'ARN avec l'ARNm *Zeb2*, masquant ainsi un site de splicing. Ce mécanisme contribue à la régulation de l'expression de la protéine *Zeb2* au cours du développement (Beltran et al., 2008). Les lncRNAs peuvent être les précurseurs de petits ARNs qui participent au processus d'ARN interférence (Fejes-Toth et al., 2009). Enfin les lncRNAs peuvent structurer certains sous-compartiments nucléaires comme les paraspeckles (compartiment de stockage d'ARNs) (Sasaki et al., 2009).

Etant donné les modes d'action des lncRNAs, les TERRAs pourraient intervenir de différentes manières dans le silence transcriptionnel lié au TPE. La transcription des TERRAs à travers les régions subtélomériques pourrait affecter la transcription des gènes les plus proches des télomères (*WASH*, *DDX11L*) par un mécanisme d'interférence transcriptionnelle. Il a été également suggéré que les TERRAs participent au maintien de l'hétérochromatine télomérique (Schoeftner & Blasco, 2008; Deng et al., 2009), mais ils pourraient aussi contribuer à la propagation de l'hétérochromatine aux régions subtélomériques sur une distance plus ou moins grande en recrutant des facteurs de modification de la chromatine ou de la méthylation de l'ADN. Par exemple, les TERRAs sont capables d'interagir avec ORC1 (Deng et al., 2009), une protéine nécessaire à la liaison de HP1 à la chromatine et à l'organisation de l'hétérochromatine.

La poursuite de l'étude de nos modèles cellulaires permettra de préciser l'implication des TERRAs dans le TPE. Nous poursuivrons plus particulièrement la caractérisation de l'hétérochromatine télomérique dans différents contextes (invalidation de facteurs de régulation par exemple). Nous étudierons également le statut de méthylation de l'ADN au niveau du gène rapporteur et des régions adjacentes et la présence de régulateurs de la chromatine (modifications post-traductionnelles d'histones, présence de variants d'histones). Le processus d'interférence transcriptionnelle pourrait être approché en déterminant l'origine le départ de la transcription du transcrit GFP-TERRA ainsi qu'en étudiant la présence de la machinerie transcriptionnelle.

4.2.2. Rôle des insulateurs et de CTCF dans la régulation du TPE

Le macrosatellite *D4Z4* se comporte comme un insulateur, à la fois de type “*enhancer-blocking*” et de type barrière de chromatine (Ottaviani et al., 2009a). L’activité protectrice contre le TPE est assurée par une séquence proximale de 432 pb et dépendante de CTCF et des lamines de type A. La caractérisation ultérieure de cette séquence proximale a mis en évidence la présence d’un site CTCF contenu dans une séquence de 30 pb qui est suffisante pour protéger contre le TPE. Cependant, cette activité n’est pas complète suggérant l’implication de séquences adjacentes soit pour stabiliser la liaison de CTCF, soit pour recruter d’autres facteurs ou favoriser des interactions entre séquences distantes. Ainsi, *D4Z4* est le premier élément insulateur décrit chez l’Homme conférant une protection contre le TPE, d’une façon dépendante des lamines de type A et de CTCF. Chez le poulet, il a déjà été montré que l’insulateur 5’HS4 était capable de s’opposer au TPE (Rincon-Arango et al., 2007). Le rôle de CTCF n’a pas été testé dans ce cas, probablement car il participe au niveau de ce site à l’activité de type “*enhancer-blocking*” et non à celle de type barrière de chromatine.

Les sites CTCF sont associés à des éléments frontières qui définissent les limites entre différents domaines. Ils sont fréquemment retrouvés à la jonction entre euchromatine et hétérochromatine (Cuddapah et al., 2009; Witcher & Emerson, 2009) ou les bordures des LADs (Guelen et al., 2008). Nos travaux et des travaux récents suggèrent également un rôle de CTCF dans la régulation des séquences répétées. En effet, le macrosatellite *DXZ4*, localisé sur le chromosome X, a été caractérisé comme un élément de 3 kb, ne présentant aucune homologie de séquence avec *D4Z4*, mais riche en dinucléotides CpGs (62%), présent en 50 à 100 copies et portant un site CTCF situé à proximité d’un promoteur bidirectionnel (Chadwick, 2008). Le comportement chromatinien de *DXZ4* diffère en fonction de sa position sur le chromosome X actif (Xa) ou inactif (Xi). Sur le chromosome Xa, l’élément *DXZ4* est dans une région d’hétérochromatine caractérisée par une forte méthylation des CpGs, la triméthylation des résidus H3K9 et la présence de la protéine HP1 γ . Cependant, il est le siège d’une activité transcriptionnelle forte générant des petits ARNs, qui participeraient à la mise en place de l’hétérochromatine par un mécanisme d’ARN interférence. Sur l’Xi, les éléments *DXZ4* sont dans une région de chromatine plus relâchée (hypométhylation des îlots de CpGs, diméthylation des résidus H3K4) entourée de deux régions d’hétérochromatine, l’une constitutive et l’autre facultative. La liaison de CTCF à *DXZ4* est spécifique du chromosome Xi. Elle est associée à une augmentation des transcrits longs, suggérant la répression de

l'hétérochromatinisation par la voie d'ARN interférence. Ainsi, le site CTCF d'un macrosatellite peut contribuer à isoler une région d'euchromatine au sein de l'hétérochromatine.

Les régions subtélomériques subterminales sont enrichies en sites CTCF (Riethman, 2008b), certains d'entre eux pourraient, comme l'insulateur proximal de *D4Z4*, préserver l'expression de gènes en s'opposant localement à la diffusion de l'hétérochromatine depuis le télomère. Ceci pourrait expliquer les difficultés à étudier le TPE au niveau des gènes endogènes (Ofir et al., 1999; Ning et al., 2003; Hernandez-Caballero et al., 2009). De plus, comme cela a été démontré chez la levure, la répression transcriptionnelle par effet de position peut ne pas être continue mais dépendre de mécanismes régulateurs complexes encore non identifiés, comme le positionnement nucléaire, la cinétique de réplication ou les marques chromatiniennes.

Nous avons évoqué ici le rôle des insulateurs et des lncRNAs, deux éléments qui participent à de nombreux processus épigénétiques, parfois de façon conjointe comme dans l'inactivation du chromosome X ou l'empreinte génomique parentale (Bell & Felsenfeld, 2000; Hark et al., 2000; Gabory et al., 2009). Ils pourraient également coopérer dans la régulation du TPE chez l'Homme. Cependant, la dissection des mécanismes du TPE chez l'Homme n'en est qu'à ses débuts et nécessite encore de nombreuses investigations.

4.3. Les pathologies subtélomériques constitutionnelles

4.3.1. La FSHD : un modèle de pathologie subtélomérique, une nouvelle laminopathie, ou... ?

Bien que de nombreux gènes morbides soient présents dans le dernier mégabase des extrémités chromosomiques, la FSHD se distingue des autres pathologies constitutionnelles subtélomériques. Tout d'abord, la région 4qter est extrêmement pauvre en gènes comparé aux autres régions subtélomériques et elle est une des rares régions subtélomériques dont le dernier Mb contient un LAD (Guelen et al., 2008). Ensuite, le locus responsable de la FSHD, *D4Z4*, est le locus morbide le plus proche d'un télomère (60 kb). Enfin, alors que l'expansion de courtes séquences nucléotidiques est à l'origine de nombreuses maladies génétiques, la FSHD implique un élément répété peu connu, un macrosatellite, dans un mécanisme peu courant, la contraction de répétitions. Malgré l'identification du locus *D4Z4* il y a une vingtaine d'années, la physiopathologie de la FSHD n'est toujours pas complètement élucidée.

Nos résultats sont en faveur d'un gain de fonction de CTCF secondaire à la contraction de *D4Z4* entraînant une modification des interactions de la région 4q35 avec la lamina nucléaire. Il a d'ailleurs été démontré l'existence d'interactions entre CTCF et les lamines de type A, bien que celles-ci ne soient pas complètement caractérisées (Yusufzai et al., 2004). Parmi les différentes laminopathies (**Tableau I**), la myopathie d'Emery-Dreyfuss (EDMD) présente des similitudes avec la FSHD. D'ailleurs, certains patients sont parfois diagnostiqués comme FSHD au début de l'évolution de la maladie. L'EDMD est caractérisée par des contractures musculaires précoces (coudes, chevilles, extenseurs de la nuque) débutant dans l'enfance, une faiblesse musculaire progressive de distribution huméro-péronéale/scapulo-péronéale apparaissant de la première à la troisième décennie et enfin une atteinte cardiaque, avec des troubles du rythme et de la conduction et une cardiomyopathie. Il existe une grande variabilité inter et intra-familiale dans la sévérité et l'évolution de la

maladie. L'EDMD est de transmission récessive liée à l'X par des mutations de l'émerine (gène *EMD* 61%) et de *FHL1* (10%) ou autosomique dominante (AD) par mutation du gène *LMNA* (45% des formes AD). Les mutations du gène *EMD* sont le plus souvent responsables d'une absence complète d'expression de l'émerine ou de la production d'une protéine non fonctionnelle alors que les mutations du gène *LMNA* sont de type faux-sens (85%) générant une protéine mutante normalement exprimée mais avec un acide aminé modifié (Bonne et al., 1993). Comme les protéines émerine et lamines de type A ont de nombreuses interactions avec la chromatine et la matrice nucléaire, il a été proposé que l'EDMD serait liée à une dérégulation de voies de signalisation dans le muscle, induisant des anomalies de prolifération et/ou de différenciation. L'analyse du transcriptome de patients EDMD (par mutations *EMD* ou *LMNA*) a mis en évidence une anomalie de la voie de signalisation Lamine-Rb-MyoD lors de la transition myoblastes/myotubes (Bakay et al., 2006). La même étude montrait que le transcriptome des patients FSHD étaient proche de celui des patients EDMD. Une autre étude a mis en évidence chez les patients FSHD la dérégulation de gènes impliqués dans le contrôle du cycle cellulaire, la prolifération et la différenciation musculaire (Winokur et al., 2003). Certains d'entre eux sont impliqués dans la réponse au stress oxydatif, d'autres sont des cibles directes de MyoD, renforçant l'hypothèse d'une voie physiopathologique commune entre EDMD et FSHD.

Plus récemment, un modèle impliquant le transcrite *DUX4* a été proposé (§1.3.3.6. et **Figure 14d**) (Lemmers et al., 2010a; Snider et al., 2010). Ce modèle permettrait d'expliquer la spécificité allélique de la FSHD et la nécessité d'avoir au moins une répétition *D4Z4* résiduelle, et de relier la physiopathologie de FSHD1 et FSHD2. En résumé, au cours de la différenciation, l'expression de la forme longue de *DUX4* à partir de la dernière répétition *D4Z4* est normalement réprimée. Dans la FSHD, le relâchement de la chromatine ne permettrait pas cette répression et la persistance de l'expression d'une forme longue de *DUX4* serait à l'origine de la maladie (Snider et al., 2010). Ce transcrite faiblement détectable serait produit dans une petite proportion de cellules et conduirait à la production d'une faible quantité de protéine DUX4 toxique, susceptible de diffuser d'un noyau à l'autre au sein de la fibre musculaire. La toxicité de DUX4 serait secondaire à l'interaction avec la voie de myogenèse MyoD (Bosnakovski et al., 2008; Bosnakovski et al., 2009; Wallace et al., 2010), l'activation du facteur transcriptionnel PITX1 (Dixit et al., 2007) ou l'induction de l'apoptose (Kowaljew et al., 2007). Cependant, la plupart de ces données proviennent de modèle de surexpression très forte de DUX4. Or, seulement 1% cellules musculaires FSHD en culture exprime un fort taux de DUX4 et seule la moitié des patients testés semble présenter le

transcrit correspondant à la forme longue de DUX4. De plus, l'expression de la protéine n'a encore jamais été évaluée au niveau tissulaire chez les patients. Il est donc licite de s'interroger si l'expression de DUX4 chez les patients peut rendre compte du phénotype et si cette surexpression est véritablement le modèle "unificateur" proposé par ses auteurs pour expliquer la pathologie, d'autant plus que le lien avec la réduction du nombre de répétitions *D4Z4* n'a pas été évaluée dans ce modèle (Lemmers et al., 2010a). Enfin, ce modèle ne tient pas compte des signes extra-musculaires de la FSHD (rétinopathie, surdité), qui sont relativement fréquent, même s'ils sont peu invalidant par rapport à la faiblesse musculaire. En effet, l'expression de MyoD est limitée au muscle et PITX1 est impliqué dans le développement des membres inférieurs (responsable de pieds bots congénitaux, MIM#119800).

L'observation, dans nos modèles cellulaires, que la présence d'un signal de polyadénylation fort est associé à une diminution des transcrits GFP-TERRAs pourrait suggérer l'hypothèse d'un modèle où la pathogénie proviendrait non pas (seulement) de la présence anormale de la protéine DUX4, mais aussi de la diminution des transcrits TERRAs dans la région 4q35 (**Figure 27**). Cette diminution des TERRAs serait responsable d'une stabilisation de l'état chromatinien relâché qui permettrait de recruter au niveau de *D4Z4* des facteurs comme la protéine CTCF et les lamines de type A. Ces dernières, du fait de leur important réseau d'interaction, pourraient contribuer non seulement aux signes musculaires mais aussi extra-musculaires. Avant de pouvoir s'aventurer dans une telle hypothèse, il faudrait accumuler davantage de connaissances sur la régulation de la chromatine télomérique et subtélomérique par les TERRAs, la régulation transcriptionnelle des TERRAs, identifier et caractériser la population des TERRAs en 4q35 et le réseau d'interaction de la région 4q35 avec la lamina nucléaire et en particulier les lamines de type A. Les modèles cellulaires que nous avons développé pourraient y contribuer.

La FSHD est donc une pathologie subtélomérique tout à fait particulière, dont la physiopathologie n'est pas totalement éclaircie. Cependant, l'ensemble des travaux qu'elle a suscités a permis d'acquérir de réelle connaissance sur les macrosatellites et les régions subtélomériques, qui pourront sans doute être appliquées à d'autres maladies impliquant des gènes subtélomériques.

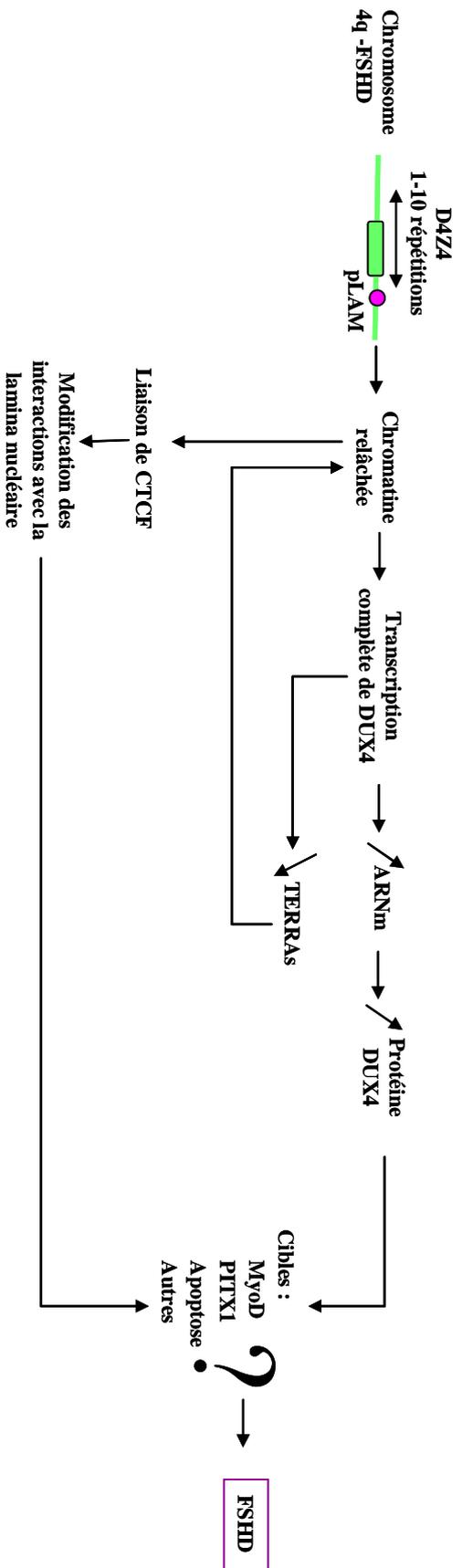


Figure 27 : Proposition d'une hypothèse physiopathologique de la FSHD

La présence d'un signal de polyadénylation (allèle 4qA161) et d'un état chromatinien relâché (du fait de la contraction de *D4Z4*) permettent la transcription complète de l'ARN DUX4. Il en résulterait une augmentation de la protéine DUX4 et une diminution des TERRAS. Celle-ci permettrait de stabiliser l'état chromatinien et de recruter différents facteurs au niveau de *D4Z4* comme CTCF et les Lamines A.

4.3.2. Les syndromes subtélomériques

En pathologie constitutionnelle, la notion de désordre subtélomérique implique le plus souvent la dernière bande (ou sous-bande) du chromosome, c'est-à-dire plusieurs mégabases. L'observation du dernier mégabase de chaque extrémité chromosomique montre qu'il contient de nombreux gènes impliqués dans des pathologies constitutionnelles mendéliennes, dominantes et récessives (**Tableau II**). Par ailleurs, les régions subtélomériques peuvent être l'objet de remaniements chromosomiques responsables de retard mental et/ou de malformations congénitales. Le retard mental touche environ 3% de la population générale (Tyson et al., 2004) et 2,5% des cas sont secondaires à une perte et/ou un gain de matériel des régions subtélomériques (Ravnan et al., 2006; Ballif et al., 2007). Les réarrangements sont de taille variable mais 40% d'entre eux sont supérieurs à 5 Mb (Ballif et al., 2007) et emportent donc en plus des régions proprement subtélomériques de larges régions d'ADN spécifiques, responsables du phénotype. On peut observer des délétions terminales, interstitielles ou complexes, *de novo* ou héritées (Adeyinka et al., 2005). L'étude des délétions terminales n'a identifié aucun point de cassure récurrent, malgré la présence de séquences répétées. Ces délétions seraient donc secondaires à des cassures chromosomiques aléatoires (Ledbetter & Martin, 2007). Certains réarrangements subtélomériques sont corrélés à un phénotype identifiable. Nous donnerons quelques exemples de ces syndromes subtélomériques et de leur mécanisme physiopathologique.

La monosomie 1p36 est la plus fréquente des délétions subtélomériques (1/5000). Elle est caractérisée par un retard mental et de développement, une surdité, des convulsions, une hypotonie, des cardiopathies et une dysmorphie faciale caractéristique (Gajecka et al., 2007). Deux tiers des réarrangements *de novo* sont apparemment de simples délétions terminales. Le tiers restant correspond à des anomalies complexes (délétion et duplication inversée, large duplication et triplication avec petite délétion, délétion interstitielle). Certaines équipes ont tenté de démontrer que le syndrome 1p36 était un syndrome des gènes contigus et ont proposé des gènes candidats pour les convulsions et les caractéristiques faciales (Heilstedt et al., 2003a; Heilstedt et al., 2003b). Cependant, il n'existe pas de corrélation entre la taille de la délétion et le nombre de signes cliniques observés (Gajecka et al., 2007). De plus, aucun point de cassure commun, ni d'intervalle délété commun entre les patients n'a été identifié. Enfin, Redon et collaborateurs ont étudiés six patients en CGH-array avec une puce *tiling-path*.

Deux des 6 patients présentaient un phénotype très proche mais pas de délétion chevauchante (Redon et al., 2005). Il a donc été proposé que la monosomie 1p36 serait secondaire à un effet de position plutôt qu'à une haplo-insuffisance de gènes contigus.

La microdélétion 22qter est caractérisée par une hypotonie néonatale sévère, un retard global de développement, un comportement autistique, une croissance normale ou accélérée, un langage absent ou sévèrement retardé et une dysmorphie faciale mineure (Balikova et al., 2007). La délétion est habituellement terminale bien que des cas de délétions interstitielles soient décrits. La taille varie de 130 kb à 9 Mb (Wilson et al., 2003). Cependant, il y a peu de corrélation entre la taille de la délétion et la sévérité du phénotype (Koolen et al., 2005). Il a été démontré que des patients avaient en commun une délétion de *SHANK3*, un gène codant pour une protéine de structure dans les synapses excitatrices (Wilson et al., 2003; Bonaglia et al., 2006). De plus, des patients présentant une disruption ou des mutations ponctuelles dans ce gène ont un phénotype identique à celui de la délétion 22qter (Bonaglia et al., 2001; Durand et al., 2007). Ainsi, il apparaît que *SHANK3* est le gène majeur responsable du phénotype neurocomportemental du syndrome de microdélétion 22qter.

La délétion 3pter est un syndrome plus controversé. Ce syndrome est caractérisé par un retard de croissance, une microcéphalie, une dysmorphie faciale (ptosis, pointe du nez large, micrognathie), un retard mental, des malformations cardiaques, une surdité et une polydactylie (Cargile et al., 2002; Fernandez et al., 2004). Des gènes candidats pour le retard mental, les cardiopathies ou la surdité ont été identifiés (Green et al., 2000; Fernandez et al., 2004; Dijkhuizen et al., 2006; McCullough et al., 2007; Shuib et al., 2009). Cependant, certains patients sont très peu, voire pas symptomatiques, pouvant faire parfois discuter l'implication réelle de ces gènes dans le phénotype (Knight et al., 1995; Takagishi et al., 2006; Hoo & Shrimpton, 2008; Pohjola et al., 2010).

Chromosome	LAD	Gène	Distance/télomère (kb)	Maladie	Transmission
2qter	-	D2HGDH	594,245	Acidurie Hydroxyglutarique	AR
3qter	-	RPL35A	334,709	Anémie de Blackfan-Diamond	AD
4pter	-	PDE6B	609,373	Rétinite pigmentaire/ cécité nocturne congénitale	AR/AD
4qter	+	IDUA	970,785	Mucopolysaccharidose de type Ih/s	AR
4qter	+	D4Z4	60	Myopathie facio-scapulo-humérale	AD
5pter	-	SDHA	271,356	Syndrome de Leigh/ déficit en complexe II de la chaîne respiratoire mitochondriale	AR
5qter	+	FLT4	848,636	Lymphoedème héréditaire	AD
6qter	-	TBP	176,12	Ataxie spinocérébelleuse 17	AD
7pter	-	FAM20C	288,052	Syndrome de Raine	AR
8qter	-	SLC39A4	661,745	Acrodermatite entéropathique	AR
8qter	-	RECQL4	560,818	Syndromes de Rapadilino/ Baller-Gerold/ Rothmund-Thomson	AR
9pter	+	DOCK8	263,048	Retard mental	AD
9pter	+	KANK1	494,703	Quadriplégie spastique	AD
9pter	+	DMRT1	831,69	Réversion sexuelle	AD
9qter	-	EHMT1	422,853	Syndrome délétion 9qter	AD
9qter	-	NELF	799,677	Hypogonadisme hypogonadotrope	AD
11pter	-	TALDO1	737,432	Déficit en transaldolase	AR
11pter	-	SLC25A22	780,475	Epilepsie myoclonique néonatale avec suppression-burst	AR
11pter	-	PNPLA2	808,902	Myopathie et anomalie de stockage des lipides	AR
11pter	-	CD151	822,952	Néphropathie, épidermolyse bulleuse prétibiale et surdité	AR
11pter	-	HRAS	522,242	Syndrome de Costello	AD
11qter	-	ACAD8	811,428	Déficit en isobutyrylCoA déshydrogénase	AR
12pter	-	WNK1	732,486	Neuropathie héréditaire sensorielle et autonome type II/ Pseudohypaldostéronisme lic	AD/AD
13qter	-	GRK1	769,007	Maladie d'Oguchi 2	AR
16pter	-	AXIN1	277,441	Anomalie de duplication caudale	AD
16pter	-	LMF1	843,635	Déficit combiné en lipase	AR
16pter	-	HBA1/HBA2	162,846	α -thalassémie	AD
16qter	-	FANCA	416,688	Anémie de Fanconi	AR
16qter	-	PGN	675,579	Paraplégie spastique 7	AR
17qter	-	ZNF750	383,522	Dermatite séborrhéique et éléments psoriasiformes	AD
18qter	+	CTDP1	501,658	Caracte congénitale, dysmorphie faciale et neuropathie	AR
19pter	-	ELA2	803,291	Neutropénie congénitale	AD
19pter	-	CFD	810,665	Déficit en facteur D du complément	AD
19pter	-	KISS1R	868,342	Hypogonadisme hypogonadotrope/ puberté précoce centrale	AD
20pter	+	RSPO4	887,096	Anonychie congénitale	AR
20qter	-	CHRNA4	972,825	Epilepsie nocturne frontale	AD
20qter	-	KCNQ2	861,527	Epilepsie néonatale bénigne	AD
20qter	-	SOX18	282,541	Syndrome hypotrichose, lymphoedème, telangiectasie	AD/AR
21qter	-	FTCD	544,414	Déficit en glutamate forminotransférase	AR
21qter	-	PCNT2	254,213	Nanisme microcéphalique ostéodysplasique primordial	AR
21qter	-	COL6A1	694,932	Myopathie de Bethlem/ Dystrophie musculaire congénitale d'Ullrich	AD/AR
21qter	-	COL6A2	567,132	Myopathie de Bethlem/ Dystrophie musculaire congénitale d'Ullrich	AD/AR
22qter	-	MLC1	825,524	Leucoencéphalopathie magalencéphalique avec kystes sous-corticaux	AR
22qter	-	SCO2	380,532	Cardioencéphalomyopathie infantile fatale	AR
22qter	-	ARSA	277,959	Leucodystrophie métachromatique	AR
22qter	-	SHANK3	172,925	Syndrome de microdélétion 22q13.3	AD
Xpter	-	SHOX	505,079	Petite taille familiale/ Dyschondrostéose de Leri-Weill/ Dysplasie mésomélique de Langer	AD/AD/AR

Tableau II : Gènes morbides présents dans le dernier mégabase des extrémités chromosomiques (d'après UCSC Mars 2006)

AD : autosomique dominant ; AR : autosomique récessif

Ces exemples illustrent que la variabilité phénotypique n'est pas toujours dépendante de la taille de la délétion, à la façon d'un syndrome des gènes contigus. La présence de CNVs (Variation du nombre de copies) sur l'allèle non délété ou au niveau d'autres régions du génome peut être évoquée. Cependant, le problème de la variabilité phénotypique des microdélétions subtélomériques n'a pas encore été appréhendé sous l'angle de l'influence du télomère et notamment du TPE. La délétion peut placer à proximité du télomère des gènes normalement éloignés de celui-ci. De plus, l'absence des séquences subtélomériques strictes ne permet plus de faire barrière entre le télomère et les gènes. Ceux-ci pourraient donc subir l'influence directe du TPE. Sur l'allèle sain, l'expression des gènes pourrait être influencée par les différents haplotypes subtélomériques permettant d'atténuer plus ou moins le TPE. La dissection des mécanismes du TPE chez l'Homme permettra sans doute de progresser dans ce domaine.

5. CONCLUSION

En conclusion, ce travail a permis de caractériser l'implication de séquences subtélomériques, et notamment celle du macrosatellite *D4Z4*, dans la régulation de certaines des propriétés biologiques des télomères, comme leur compartimentalisation nucléaire, leur réplication ou l'effet de position télomérique. Il suggère l'existence d'interactions réciproques entre télomères et subtélomères qui feraient appel à des facteurs de régulations de la chromatine et de l'architecture du génome comme les insulateurs et la protéine CTCF, les lamines de type A ou les transcrits télomériques TERRAs. Ces différents facteurs pourraient interagir au sein d'un réseau d'interactions complexes, qui restent encore à élucider.

De plus, ce travail apporte de nouveaux éléments dans la compréhension des mécanismes physiopathologiques de la myopathie facio-scapulo-humérale et ouvre des perspectives dans la physiopathologie d'autres maladies constitutionnelles liées aux subtélomères.

ANNEXES

1. Matériels et méthodes complémentaires

1.1 Extraction d'ARN

Les cellules en culture ont été récoltées après traitement par la trypsine, centrifugées pendant 5 min à 1000 tour/min et rincées dans du PBS1X. L'ARN total a été extrait avec le kit RNeasy Mini Kit (Qiagen), selon le protocole indiqué. La qualité des ARNs a été vérifiée par électrophorèse sur gel d'agarose et les ARNs ont été dosés par le spectrophotomètre Nanodrop.

1.2. Marquage de sonde par amorçage aléatoire (“*random-priming*”)

Cinquante nanogrammes d'ADN préalablement dénaturés à 100°C pendant 5 min ont été incubés en présence de 0,2 nmol de nucléotides dATP, dTTP et dGTP, 5 UI du fragment Klenow (3'-5' exo-) de l'ADN polymérase I (New England Biolabs), un mélange d'hexanucléotides aléatoires (Hexanucleotide Mix, Roche) et 0,4 nmol de nucléotide radiomarqué α -dCTP³², dans un volume final de 20 μ L, pendant 1h à 37°C. L'excès de α -dCTP³² est éliminé par purification à l'aide d'une colonne G50 (GE Healthcare). La sonde est dénaturée 5 min à 100°C avant l'hybridation.

1.3. Northern-Blot

Dix microgrammes d'ARNs ont été déposés sur un gel dénaturant (agarose 1,3% ; MOPS1X, formaldéhyde 2,2 M) et séparés par électrophorèse (30 V) sur la nuit. Le gel a été rincé avec de l'eau, puis incubé dans une solution NaOH 0,05M / NaCl 1,5 M pendant 20 min, puis une solution Tris pH7,4 0,5 M / NaCl 1,5 M pendant 20 min et enfin dans du tampon SSC10X pendant 10 min. Les échantillons ont été transférés sur une membrane en nylon Hybond N+ (GE Healthcare) dans un tampon SSC10X. La membrane a été traitée par les UV (0,120J/cm²). L'hybridation avec des sondes radioactives a été réalisée dans du tampon Church à 55°C sur la nuit. Les lavages post-hybridation comprenaient deux lavages de 20 min dans une solution SSC2X/ 0,1% SDS à 55°, complété par un lavage supplémentaire dans une solution SSC0,2X / 0,1% SDS si nécessaire. Les membranes ont été exposées sur un écran PhosphorScreen pendant des temps variables et lues avec l'appareil Fuji fla-5100.

1.4. Transcription inverse

Un microgramme d'ARN a été soumis à la réaction de transcription inverse avec le kit SuperScript III First-Strand Synthesis System (Invitrogen), d'après le protocole du fabricant, optimisé pour la synthèse de séquences riches en GC. Les amorces utilisées sont soit l'oligonucléotide oligo(dT) soit l'oligonucléotide TR (5' TGCTCCGTGCATCTGGCATCCCCTAACCCCTAACCCCTAACCCCTAA 3'). Les échantillons ont été traités par la RNase H près la synthèse d'ADNc. Les ADNc ont ensuite été amplifiés par PCR quantitative.

1.5. Etude de la demi-vie des ARNs

Les cellules à 70 % de confluence ont été traitées soit avec de l'actinomycine D (Sigma) à une concentration finale de 5 µg/mL soit avec un milieu contrôle sans actinomycine D. Les cellules ont été récoltées au temps 0, puis après 1, 2, 4, 6 et 8 heures d'exposition. L'ARN a été extrait selon la méthode décrite ci-dessus. Les échantillons ont été transférés sur une membrane de nylon Hybond N+ (GE Healthcare) par slot-blot selon la technique suivante. Un microgramme d'ARN est incubé dans une solution dénaturante (66 % formamide / 8 % formaldéhyde / MOPS1X) à 65°C pendant 5 min, puis dilué dans du tampon de transfert (SSC10X final). Après le passage des échantillons, la membrane est traitée aux UV (0,120J/cm²) et pré-hybridée pendant 1h à 55°C dans du tampon Church. L'hybridation avec des sondes radiomarquées et les lavages post-hybridation sont identiques à ce qui est décrit pour le Northern Blot.

1.6. Immunoprécipitation de chromatine (ChIP)

L'immunoprécipitation de chromatine a été réalisée selon Magdinier et collaborateurs (Magdinier et al., 2004). ADN et protéines ont été fixés *in vivo* en présence de formaldéhyde. Les complexes ADN-protéines ont été cassés par sonication, de façon à obtenir des complexes nucléoprotéiques de 400 à 600 pb (Bioruptor, Diagenode) et immunoprécipités en présence d'anticorps spécifiques : H3 total (Abcam) H3K9me3, H3K9Ac, H4K20me3 (Millipore). Un échantillon d'ADN a été conservé avant immunoprécipitation (Input). L'ADN des complexes immuns a été récolté à l'aide billes immuno-affines (Sigma), élué, traité à la protéinase K (Roche) et précipité par l'éthanol. Les échantillons d'ADN ont été dosés par le spectrophotomètre Nanodrop. Ils ont été soit amplifiés par PCR quantitative, soit déposés sur une membrane de nylon Hybond N+ (GE Healthcare) par slot-blot.

1.7. Slot-blot des échantillons d'ADN issus de ChIP

Cent nanogrammes d'ADN sont incubés dans le tampon de dénaturation (0,5 M NaOH / 0,3 M NaCl) pendant 10 min à température ambiante, puis dilués dans un tampon de dilution (0,1XSSC / 0,125M NaOH) et déposés sur la membrane. Après le passage des échantillons, la membrane est rincée dans un tampon de neutralisation (0,5M Tris-HCl pH7,5 / 0,5M NaCl) et traitée aux UV (0,120J/cm²). La membrane est pré-hybridée pendant 1 h à 55°C dans un tampon de préhybridation (SSC6X / 0,5% SDS/ 1% lait). L'hybridation avec les sondes radiomarquées a été réalisée dans un tampon SSC6X / 0,1% SDS/ 1% lait à 55°C pendant la nuit. Les lavages post-hybridation à 55°C comprennent un lavage de 10 min dans un tampon SSC2X, un lavage de 30 min dans un tampon SSC2X / SDS 1%, et un à deux lavages de 30 min dans un tampon SSC0,2X / SDS 1%. Les membranes ont été exposées sur un écran PhosphorScreen pendant des temps variables et lues avec l'appareil Fuji fla-5100.

1.8. Amplification par PCR quantitative (qPCR)

Les qPCR ont été réalisées sur le thermocycleur LightCycler (Roche) à l'aide du kit Light Cycler FastStart DNA Master Sybr Green I (Roche). Un microlitre d'ADNc ou 10 ng d'ADN ont été amplifiés par des amorces spécifiques dans un volume final de 20 µL selon le programme suivant : dénaturation 95°C 8 min, 45 cycles comprenant 20 sec de dénaturation, 8 sec d'hybridation et 12 sec d'élongation. Les séquences des amorces sont les suivantes :

promoteur du gène codant pour l'histone H4 (F : 5' TGTGTGATGGGAAGATGGGA 3') (R : 5' GAGCAAAATAGCGAGACCCC 3'); β -actine (F : 5' TCCCTGGAGAAGAGCTACGA 3') (R : 5' AGCACTGTGTTGGCGTACAG 3'); GFP (F : 5' AGGAGCGCACCATCTTC 3') (R : 5' GATGTTGTGGCGGATCTTG 3'). Des courbes de dilution ont été réalisées pour chaque amplification afin de déterminer l'efficacité d'amplification.

2. Publication de revue et de synthèse

2009 C. Schluth-Bolard, A. Ottaviani, A. Bah, A. Boussouar, E. Gilson, F. Magdinier.
Dynamics and plasticity of chromosome ends : consequences in human pathology.
Atlas Genet Cytogenet Oncol Haematol June 2009.
<http://atlasgeneticsoncology.org/Deep/SubTelomereID20025.html>

3. Publication d'un chapitre de livre

C. Schluth-Bolard, A. Ottaviani, E. Gilson, F. Magdinier. Chromosomal position effects & gene variegation: impact in pathologies. In Handbook of Epigenetics: the new molecular and medical genetics. Trygve Tollefsbol. 2011: 77-105.

BIBLIOGRAPHIE

- Adeyinka A, Adams SA, Lorentz CP, Van Dyke DL, Jalal SM. 2005. Subtelomere deletions and translocations are frequently familial. *Am J Med Genet A* 135:28-35.
- Akhtar A, Gasser SM. 2007. The nuclear envelope and transcriptional control. *Nat Rev Genet* 8:507-517.
- Alexiadis V, Ballestas ME, Sanchez C, Winokur S, Vedanarayanan V, Warren M, Ehrlich M. 2007. RNAPol-ChIP analysis of transcription from FSHD-linked tandem repeats and satellite DNA. *Biochimica et biophysica acta* 1769:29-40.
- Allsopp RC, Vaziri H, Patterson C, Goldstein S, Younglai EV, Futcher AB, Greider CW, Harley CB. 1992. Telomere length predicts replicative capacity of human fibroblasts. *Proceedings of the National Academy of Sciences of the United States of America* 89:10114-10118.
- Ambrosini A, Paul S, Hu S, Riethman H. 2007. Human subtelomeric duplicon structure and organization. *Genome biology* 8:R151.
- Andres V, Gonzalez JM. 2009. Role of A-type lamins in signaling, transcription, and chromatin organization. *The Journal of cell biology* 187:945-957.
- Aparicio OM, Billington BL, Gottschling DE. 1991. Modifiers of position effect are shared between telomeric and silent mating-type loci in *S. cerevisiae*. *Cell* 66:1279-1287.
- Arashiro P, Eisenberg I, Kho AT, Cerqueira AM, Canovas M, Silva HC, Pavanello RC, Verjovski-Almeida S, Kunkel LM, Zatz M. 2009. Transcriptional regulation differs in affected facioscapulohumeral muscular dystrophy patients compared to asymptomatic related carriers. *Proceedings of the National Academy of Sciences of the United States of America* 106:6220-6225.
- Armanios M. 2009. Syndromes of telomere shortening. *Annual review of genomics and human genetics* 10:45-61.
- Azzalin CM, Reichenbach P, Khoriauli L, Giulotto E, Lingner J. 2007. Telomeric repeat containing RNA and RNA surveillance factors at mammalian chromosome ends. *Science (New York, NY)* 318:798-801.
- Bakay M, Wang Z, Melcon G, Schiltz L, Xuan J, Zhao P, Sartorelli V, Seo J, Pegoraro E, Angelini C, Shneiderman B, Escolar D, Chen YW, Winokur ST, Pachman LM, Fan C, Mandler R, Nevo Y, Gordon E, Zhu Y, Dong Y, Wang Y, Hoffman EP. 2006. Nuclear envelope dystrophies show a transcriptional fingerprint suggesting disruption of Rb-MyoD pathways in muscle regeneration. *Brain* 129:996-1013.
- Balikova I, Menten B, de Ravel T, Le Caignec C, Thienpont B, Urbina M, Doco-Fenzy M, de Rademaeker M, Mortier G, Kooy F, van den Ende J, Devriendt K, Fryns JP, Speleman F, Vermeesch JR. 2007. Subtelomeric imbalances in phenotypically normal individuals. *Human mutation* 28:958-967.
- Ballif BC, Sulpizio SG, Lloyd RM, Minier SL, Theisen A, Bejjani BA, Shaffer LG. 2007. The clinical utility of enhanced subtelomeric coverage in array CGH. *Am J Med Genet A* 143A:1850-1857.
- Barro M, Carnac G, Flavier S, Mercier J, Vassetzky Y, Laoudj-Chenivresse D. 2010. Myoblasts from affected and non-affected FSHD muscles exhibit morphological differentiation defects. *Journal of cellular and molecular medicine* 14:275-289.
- Bartolomei MS, Zemel S, Tilghman SM. 1991. Parental imprinting of the mouse H19 gene. *Nature* 351:153-155.

- Baumann P, Cech TR. 2001. Pot1, the putative telomere end-binding protein in fission yeast and humans. *Science (New York, NY)* 292:1171-1175.
- Baur JA, Zou Y, Shay JW, Wright WE. 2001. Telomere position effect in human cells. *Science (New York, NY)* 292:2075-2077.
- Bayne RA, Broccoli D, Taggart MH, Thomson EJ, Farr CJ, Cooke HJ. 1994. Sandwiching of a gene within 12 kb of a functional telomere and alpha satellite does not result in silencing. *Human molecular genetics* 3:539-546.
- Bell AC, Felsenfeld G. 2000. Methylation of a CTCF-dependent boundary controls imprinted expression of the Igf2 gene. *Nature* 405:482-485.
- Bell AC, West AG, Felsenfeld G. 1999. The protein CTCF is required for the enhancer blocking activity of vertebrate insulators. *Cell* 98:387-396.
- Beltran M, Puig I, Pena C, Garcia JM, Alvarez AB, Pena R, Bonilla F, de Herreros AG. 2008. A natural antisense transcript regulates Zeb2/Sip1 gene expression during Snail1-induced epithelial-mesenchymal transition. *Genes & development* 22:756-769.
- Benetti R, Garcia-Cao M, Blasco MA. 2007. Telomere length regulates the epigenetic status of mammalian telomeres and subtelomeres. *Nature genetics* 39:243-250.
- Benetti R, Schoeftner S, Munoz P, Blasco MA. 2008. Role of TRF2 in the assembly of telomeric chromatin. *Cell cycle (Georgetown, Tex)* 7:3461-3468.
- Bengtsson U, Altherr MR, Wasmuth JJ, Winokur ST. 1994. High resolution fluorescence in situ hybridization to linearly extended DNA visually maps a tandem repeat associated with facioscapulohumeral muscular dystrophy immediately adjacent to the telomere of 4q. *Human molecular genetics* 3:1801-1805.
- Bennett L, Palucka AK, Arce E, Cantrell V, Borvak J, Banchereau J, Pascual V. 2003. Interferon and granulopoiesis signatures in systemic lupus erythematosus blood. *The Journal of experimental medicine* 197:711-723.
- Bilaud T, Brun C, Ancelin K, Koering CE, Laroche T, Gilson E. 1997. Telomeric localization of TRF2, a novel human telobox protein. *Nature genetics* 17:236-239.
- Blackburn EH. 1984. The molecular structure of centromeres and telomeres. *Annual review of biochemistry* 53:163-194.
- Blackburn EH, Gall JG. 1978. A tandemly repeated sequence at the termini of the extrachromosomal ribosomal RNA genes in Tetrahymena. *Journal of molecular biology* 120:33-53.
- Blasco MA, Lee HW, Hande MP, Samper E, Lansdorp PM, DePinho RA, Greider CW. 1997. Telomere shortening and tumor formation by mouse cells lacking telomerase RNA. *Cell* 91:25-34.
- Bodega B, Cardone MF, Muller S, Neusser M, Orzan F, Rossi E, Battaglioli E, Marozzi A, Riva P, Rocchi M, Meneveri R, Ginelli E. 2007. Evolutionary genomic remodelling of the human 4q subtelomere (4q35.2). *BMC evolutionary biology* 7:39.
- Bonaglia MC, Giorda R, Borgatti R, Felisari G, Gagliardi C, Selicorni A, Zuffardi O. 2001. Disruption of the ProSAP2 gene in a t(12;22)(q24.1;q13.3) is associated with the 22q13.3 deletion syndrome. *American journal of human genetics* 69:261-268.
- Bonaglia MC, Giorda R, Mani E, Aceti G, Anderlid BM, Baroncini A, Pramparo T, Zuffardi O. 2006. Identification of a recurrent breakpoint within the SHANK3 gene in the 22q13.3 deletion syndrome. *Journal of medical genetics* 43:822-828.
- Bonne G, Leturcq F, Ben Yaou R. 1993. Emery-Dreifuss Muscular Dystrophy.
- Bosnakovski D, Daughters RS, Xu Z, Slack JM, Kyba M. 2009. Biphasic myopathic phenotype of mouse DUX, an ORF within conserved FSHD-related repeats. *PloS one* 4:e7003.
- Bosnakovski D, Xu Z, Gang EJ, Galindo CL, Liu M, Simsek T, Garner HR, Agha-Mohammadi S, Tassin A, Coppee F, Belayew A, Perlingeiro RR, Kyba M. 2008. An

- isogenetic myoblast expression screen identifies DUX4-mediated FSHD-associated molecular pathologies. *The EMBO journal* 27:2766-2779.
- Boulton SJ, Jackson SP. 1998. Components of the Ku-dependent non-homologous end-joining pathway are involved in telomeric length maintenance and telomeric silencing. *The EMBO journal* 17:1819-1828.
- Boveri T. 1909. Blastomerenkerne von *Ascaris megalocphala* und die Theorie der Chromosomenindividualität. *Arch Zellforsch* 3:181-268.
- Boyle S, Gilchrist S, Bridger JM, Mahy NL, Ellis JA, Bickmore WA. 2001. The spatial organization of human chromosomes within the nuclei of normal and emerin-mutant cells. *Human molecular genetics* 10:211-219.
- Branco MR, Pombo A. 2006. Intermingling of chromosome territories in interphase suggests role in translocations and transcription-dependent associations. *PLoS biology* 4:e138.
- Brouwer OF, Padberg GW, Ruys CJ, Brand R, de Laat JA, Grote JJ. 1991. Hearing loss in facioscapulohumeral muscular dystrophy. *Neurology* 41:1878-1881.
- Brown WR, MacKinnon PJ, Villasante A, Spurr N, Buckle VJ, Dobson MJ. 1990. Structure and polymorphism of human telomere-associated DNA. *Cell* 63:119-132.
- Bryan TM, Englezou A, Dalla-Pozza L, Dunham MA, Reddel RR. 1997. Evidence for an alternative mechanism for maintaining telomere length in human tumors and tumor-derived cell lines. *Nature medicine* 3:1271-1274.
- Bryan TM, Englezou A, Gupta J, Bacchetti S, Reddel RR. 1995. Telomere elongation in immortal human cells without detectable telomerase activity. *The EMBO journal* 14:4240-4248.
- Cai X, Cullen BR. 2007. The imprinted H19 noncoding RNA is a primary microRNA precursor. *RNA (New York, NY)* 13:313-316.
- Cargile CB, Goh DL, Goodman BK, Chen XN, Korenberg JR, Semenza GL, Thomas GH. 2002. Molecular cytogenetic characterization of a subtle interstitial del(3)(p25.3p26.2) in a patient with deletion 3p syndrome. *American journal of medical genetics* 109:133-138.
- Carninci P, Hayashizaki Y. 2007. Noncoding RNA transcription beyond annotated genes. *Current opinion in genetics & development* 17:139-144.
- Caslini C, Connelly JA, Serna A, Broccoli D, Hess JL. 2009. MLL associates with telomeres and regulates telomeric repeat-containing RNA transcription. *Molecular and cellular biology* 29:4519-4526.
- Celegato B, Capitanio D, Pescatori M, Romualdi C, Pacchioni B, Cagnin S, Vigano A, Colantoni L, Begum S, Ricci E, Wait R, Lanfranchi G, Gelfi C. 2006. Parallel protein and transcript profiles of FSHD patient muscles correlate to the D4Z4 arrangement and reveal a common impairment of slow to fast fibre differentiation and a general deregulation of MyoD-dependent genes. *Proteomics* 6:5303-5321.
- Celli GB, de Lange T. 2005. DNA processing is not required for ATM-mediated telomere damage response after TRF2 deletion. *Nature cell biology* 7:712-718.
- Celli GB, Denchi EL, de Lange T. 2006. Ku70 stimulates fusion of dysfunctional telomeres yet protects chromosome ends from homologous recombination. *Nature cell biology* 8:885-890.
- Cesare AJ, Reddel RR. 2010. Alternative lengthening of telomeres: models, mechanisms and implications. *Nat Rev Genet* 11:319-330.
- Chadwick BP. 2008. DXZ4 chromatin adopts an opposing conformation to that of the surrounding chromosome and acquires a novel inactive X-specific role involving CTCF and antisense transcripts. *Genome research* 18:1259-1269.

- Charron MJ, Michels CA. 1988. The naturally occurring alleles of MAL1 in *Saccharomyces* species evolved by various mutagenic processes including chromosomal rearrangement. *Genetics* 120:83-93.
- Chong L, van Steensel B, Broccoli D, Erdjument-Bromage H, Hanish J, Tempst P, de Lange T. 1995. A human telomeric protein. *Science (New York, NY)* 270:1663-1667.
- Chubb JR, Bickmore WA. 2003. Considering nuclear compartmentalization in the light of nuclear dynamics. *Cell* 112:403-406.
- Chubb JR, Boyle S, Perry P, Bickmore WA. 2002. Chromatin motion is constrained by association with nuclear compartments in human cells. *Curr Biol* 12:439-445.
- Ciccone R, Mattina T, Giorda R, Bonaglia MC, Rocchi M, Pramparo T, Zuffardi O. 2006. Inversion polymorphisms and non-contiguous terminal deletions: the cause and the (unpredicted) effect of our genome architecture. *Journal of medical genetics* 43:e19.
- Clapp J, Mitchell LM, Bolland DJ, Fantes J, Corcoran AE, Scotting PJ, Armour JA, Hewitt JE. 2007. Evolutionary conservation of a coding function for D4Z4, the tandem DNA repeat mutated in facioscapulohumeral muscular dystrophy. *American journal of human genetics* 81:264-279.
- Cohen SB, Graham ME, Lovrecz GO, Bache N, Robinson PJ, Reddel RR. 2007. Protein composition of catalytically active human telomerase from immortal cells. *Science (New York, NY)* 315:1850-1853.
- Costa V, Casamassimi A, Roberto R, Gianfrancesco F, Matarazzo MR, D'Urso M, D'Esposito M, Rocchi M, Ciccodicola A. 2009. DDX11L: a novel transcript family emerging from human subtelomeric regions. *BMC genomics* 10:250.
- Cremer M, Grasser F, Lanctot C, Muller S, Neusser M, Zinner R, Solovei I, Cremer T. 2008. Multicolor 3D fluorescence in situ hybridization for imaging interphase chromosomes. *Methods in molecular biology (Clifton, NJ)* 463:205-239.
- Cremer T, Cremer C. 2001. Chromosome territories, nuclear architecture and gene regulation in mammalian cells. *Nat Rev Genet* 2:292-301.
- Cremer T, Cremer M. 2010. Chromosome territories. *Cold Spring Harbor perspectives in biology* 2:a003889.
- Cremer T, Cremer M, Dietzel S, Muller S, Solovei I, Fakan S. 2006. Chromosome territories-- a functional nuclear landscape. *Current opinion in cell biology* 18:307-316.
- Croft JA, Bridger JM, Boyle S, Perry P, Teague P, Bickmore WA. 1999. Differences in the localization and morphology of chromosomes in the human nucleus. *The Journal of cell biology* 145:1119-1131.
- Cuddapah S, Jothi R, Schones DE, Roh TY, Cui K, Zhao K. 2009. Global analysis of the insulator binding protein CTCF in chromatin barrier regions reveals demarcation of active and repressive domains. *Genome research* 19:24-32.
- d'Adda di Fagagna F, Reaper PM, Clay-Farrace L, Fiegler H, Carr P, Von Zglinicki T, Saretzki G, Carter NP, Jackson SP. 2003. A DNA damage checkpoint response in telomere-initiated senescence. *Nature* 426:194-198.
- de Greef JC, Frants RR, van der Maarel SM. 2008. Epigenetic mechanisms of facioscapulohumeral muscular dystrophy. *Mutation research* 647:94-102.
- de Greef JC, Lemmers RJ, Camano P, Day JW, Sacconi S, Dunand M, van Engelen BG, Kiuru-Enari S, Padberg GW, Rosa AL, Desnuelle C, Spuler S, Tarnopolsky M, Venance SL, Frants RR, van der Maarel SM, Tawil R. 2010. Clinical features of facioscapulohumeral muscular dystrophy 2. *Neurology* 75:1548-1554.
- de Greef JC, Lemmers RJ, van Engelen BG, Sacconi S, Venance SL, Frants RR, Tawil R, van der Maarel SM. 2009. Common epigenetic changes of D4Z4 in contraction-dependent and contraction-independent FSHD. *Human mutation* 30:1449-1459.

- de Lange T, Shiue L, Myers RM, Cox DR, Naylor SL, Killery AM, Varmus HE. 1990. Structure and variability of human chromosome ends. *Molecular and cellular biology* 10:518-527.
- De Vos WH, Houben F, Hoebe RA, Hennekam R, van Engelen B, Manders EM, Ramaekers FC, Broers JL, Van Oostveldt P. 2010. Increased plasticity of the nuclear envelope and hypermobility of telomeres due to the loss of A-type lamins. *Biochimica et biophysica acta* 1800:448-458.
- Deak KL, Lemmers RJ, Stajich JM, Klooster R, Tawil R, Frants RR, Speer MC, van der Maarel SM, Gilbert JR. 2007. Genotype-phenotype study in an FSHD family with a proximal deletion encompassing p13E-11 and D4Z4. *Neurology* 68:578-582.
- Dechat T, Pflieger K, Sengupta K, Shimi T, Shumaker DK, Solimando L, Goldman RD. 2008. Nuclear lamins: major factors in the structural organization and function of the nucleus and chromatin. *Genes & development* 22:832-853.
- Defossez PA, Kelly KF, Filion GJ, Perez-Torrado R, Magdinier F, Menoni H, Nordgaard CL, Daniel JM, Gilson E. 2005. The human enhancer blocker CTC-binding factor interacts with the transcription factor Kaiso. *The Journal of biological chemistry* 280:43017-43023.
- Deghani H, Dellaire G, Bazett-Jones DP. 2005. Organization of chromatin in the interphase mammalian cell. *Micron* 36:95-108.
- Deidda G, Cacurri S, Piazzo N, Felicetti L. 1996. Direct detection of 4q35 rearrangements implicated in facioscapulohumeral muscular dystrophy (FSHD). *Journal of medical genetics* 33:361-365.
- Denchi EL, de Lange T. 2007. Protection of telomeres through independent control of ATM and ATR by TRF2 and POT1. *Nature* 448:1068-1071.
- Deng Z, Campbell AE, Lieberman PM. 2010. TERRA, CpG methylation and telomere heterochromatin: lessons from ICF syndrome cells. *Cell cycle (Georgetown, Tex)* 9:69-74.
- Deng Z, Norseen J, Wiedmer A, Riethman H, Lieberman PM. 2009. TERRA RNA binding to TRF2 facilitates heterochromatin formation and ORC recruitment at telomeres. *Molecular cell* 35:403-413.
- Deniaud E, Bickmore WA. 2009. Transcription and the nuclear periphery: edge of darkness? *Current opinion in genetics & development* 19:187-191.
- Dietzel S, Jauch A, Kienle D, Qu G, Holtgreve-Grez H, Eils R, Munkel C, Bittner M, Meltzer PS, Trent JM, Cremer T. 1998. Separate and variably shaped chromosome arm domains are disclosed by chromosome arm painting in human cell nuclei. *Chromosome Res* 6:25-33.
- Dijkhuizen T, van Essen T, van der Vlies P, Verheij JB, Sikkema-Raddatz B, van der Veen AY, Gerssen-Schoorl KB, Buys CH, Kok K. 2006. FISH and array-CGH analysis of a complex chromosome 3 aberration suggests that loss of CNTN4 and CRBN contributes to mental retardation in 3pter deletions. *Am J Med Genet A* 140:2482-2487.
- Dixit M, Ansseau E, Tassin A, Winokur S, Shi R, Qian H, Sauvage S, Matteotti C, van Acker AM, Leo O, Figlewicz D, Barro M, Laoudj-Chenivresse D, Belayew A, Coppee F, Chen YW. 2007. DUX4, a candidate gene of facioscapulohumeral muscular dystrophy, encodes a transcriptional activator of PITX1. *Proceedings of the National Academy of Sciences of the United States of America* 104:18157-18162.
- Doheny JG, Mottus R, Grigliatti TA. 2008. Telomeric position effect--a third silencing mechanism in eukaryotes. *PloS one* 3:e3864.
- Dunham MA, Neumann AA, Fasching CL, Reddel RR. 2000. Telomere maintenance by recombination in human cells. *Nature genetics* 26:447-450.

- Durand CM, Betancur C, Boeckers TM, Bockmann J, Chaste P, Fauchereau F, Nygren G, Rastam M, Gillberg IC, Anckarsater H, Sponheim E, Goubran-Botros H, Delorme R, Chabane N, Mouren-Simeoni MC, de Mas P, Bieth E, Roge B, Heron D, Burglen L, Gillberg C, Leboyer M, Bourgeron T. 2007. Mutations in the gene encoding the synaptic scaffolding protein SHANK3 are associated with autism spectrum disorders. *Nature genetics* 39:25-27.
- Farr CJ, Stevanovic M, Thomson EJ, Goodfellow PN, Cooke HJ. 1992. Telomere-associated chromosome fragmentation: applications in genome manipulation and analysis. *Nature genetics* 2:275-282.
- Fejes-Toth K, Sotirova V, Sachidanandam R, Assaf G, Hannon GJ, Kapranov P, Foissac S, Willingham AT, Dutttagupta R, Dumais E, Gingeras TR. 2009. Post-transcriptional processing generates a diversity of 5'-modified long and short RNAs. *Nature* 457:1028-1032.
- Fellermann K, Stange DE, Schaeffeler E, Schmalzl H, Wehkamp J, Bevins CL, Reinisch W, Teml A, Schwab M, Lichter P, Radlwimmer B, Stange EF. 2006. A chromosome 8 gene-cluster polymorphism with low human beta-defensin 2 gene copy number predisposes to Crohn disease of the colon. *American journal of human genetics* 79:439-448.
- Feng J, Bi C, Clark BS, Mady R, Shah P, Kohtz JD. 2006. The Evf-2 noncoding RNA is transcribed from the Dlx-5/6 ultraconserved region and functions as a Dlx-2 transcriptional coactivator. *Genes & development* 20:1470-1484.
- Fernandez T, Morgan T, Davis N, Klin A, Morris A, Farhi A, Lifton RP, State MW. 2004. Disruption of contactin 4 (CNTN4) results in developmental delay and other features of 3p deletion syndrome. *American journal of human genetics* 74:1286-1293.
- Feuerhahn S, Iglesias N, Panza A, Porro A, Lingner J. 2010. TERRA biogenesis, turnover and implications for function. *FEBS letters* 584:3812-3818.
- Filesi I, Cacchione S, De Santis P, Rossetti L, Savino M. 2000. The main role of the sequence-dependent DNA elasticity in determining the free energy of nucleosome formation on telomeric DNAs. *Biophysical chemistry* 83:223-237.
- Finlan LE, Sproul D, Thomson I, Boyle S, Kerr E, Perry P, Ylstra B, Chubb JR, Bickmore WA. 2008. Recruitment to the nuclear periphery can alter expression of genes in human cells. *PLoS genetics* 4:e1000039.
- Flynn RL, Centore RC, O'Sullivan RJ, Rai R, Tse A, Songyang Z, Chang S, Karlseder J, Zou L. 2011. TERRA and hnRNPA1 orchestrate an RPA-to-POT1 switch on telomeric single-stranded DNA. *Nature*.
- Freitas-Junior LH, Bottius E, Pirrit LA, Deitsch KW, Scheidig C, Guinet F, Nehrbass U, Wellems TE, Scherf A. 2000. Frequent ectopic recombination of virulence factor genes in telomeric chromosome clusters of *P. falciparum*. *Nature* 407:1018-1022.
- Fu D, Collins K. 2007. Purification of human telomerase complexes identifies factors involved in telomerase biogenesis and telomere length regulation. *Molecular cell* 28:773-785.
- Funakoshi M, Goto K, Arahata K. 1998. Epilepsy and mental retardation in a subset of early onset 4q35-facioscapulohumeral muscular dystrophy. *Neurology* 50:1791-1794.
- Gabellini D, D'Antona G, Moggio M, Prella A, Zecca C, Adami R, Angeletti B, Ciscato P, Pellegrino MA, Bottinelli R, Green MR, Tupler R. 2006. Facioscapulohumeral muscular dystrophy in mice overexpressing FRG1. *Nature* 439:973-977.
- Gabellini D, Green MR, Tupler R. 2002. Inappropriate gene activation in FSHD: a repressor complex binds a chromosomal repeat deleted in dystrophic muscle. *Cell* 110:339-348.
- Gabory A, Ripoche MA, Le Digarcher A, Watrin F, Ziyyat A, Forne T, Jammes H, Ainscough JF, Surani MA, Journot L, Dandolo L. 2009. H19 acts as a trans regulator

- of the imprinted gene network controlling growth in mice. *Development (Cambridge, England)* 136:3413-3421.
- Gabriels J, Beckers MC, Ding H, De Vriese A, Plaisance S, van der Maarel SM, Padberg GW, Frants RR, Hewitt JE, Collen D, Belayew A. 1999. Nucleotide sequence of the partially deleted D4Z4 locus in a patient with FSHD identifies a putative gene within each 3.3 kb element. *Gene* 236:25-32.
- Gajecka M, Mackay KL, Shaffer LG. 2007. Monosomy 1p36 deletion syndrome. *Am J Med Genet C Semin Med Genet* 145C:346-356.
- Galetta F, Franzoni F, Sposito R, Plantinga Y, Femia FR, Galluzzi F, Rocchi A, Santoro G, Siciliano G. 2005. Subclinical cardiac involvement in patients with facioscapulohumeral muscular dystrophy. *Neuromuscul Disord* 15:403-408.
- Galiova G, Bartova E, Raska I, Krejci J, Kozubek S. 2008. Chromatin changes induced by lamin A/C deficiency and the histone deacetylase inhibitor trichostatin A. *European journal of cell biology* 87:291-303.
- Ganz T. 2003. Defensins: antimicrobial peptides of innate immunity. *Nat Rev Immunol* 3:710-720.
- Gao Q, Reynolds GE, Innes L, Pedram M, Jones E, Junabi M, Gao DW, Ricoul M, Sabatier L, Van Brocklin H, Franc BL, Murnane JP. 2007. Telomeric transgenes are silenced in adult mouse tissues and embryo fibroblasts but are expressed in embryonic stem cells. *Stem cells (Dayton, Ohio)* 25:3085-3092.
- Garcia-Cao M, O'Sullivan R, Peters AH, Jenuwein T, Blasco MA. 2004. Epigenetic regulation of telomere length in mammalian cells by the Suv39h1 and Suv39h2 histone methyltransferases. *Nature genetics* 36:94-99.
- Gaszner M, Felsenfeld G. 2006. Insulators: exploiting transcriptional and epigenetic mechanisms. *Nat Rev Genet* 7:703-713.
- Gehring WJ, Klemenz R, Weber U, Kloter U. 1984. Functional analysis of the white gene of *Drosophila* by P-factor-mediated transformation. *The EMBO journal* 3:2077-2085.
- Giglio S, Broman KW, Matsumoto N, Calvari V, Gimelli G, Neumann T, Ohashi H, Voullaire L, Larizza D, Giorda R, Weber JL, Ledbetter DH, Zuffardi O. 2001. Olfactory receptor-gene clusters, genomic-inversion polymorphisms, and common chromosome rearrangements. *American journal of human genetics* 68:874-883.
- Gilbert N, Boyle S, Fiegler H, Woodfine K, Carter NP, Bickmore WA. 2004. Chromatin architecture of the human genome: gene-rich domains are enriched in open chromatin fibers. *Cell* 118:555-566.
- Gilbert N, Ramsahoye B. 2005. The relationship between chromatin structure and transcriptional activity in mammalian genomes. *Briefings in functional genomics & proteomics* 4:129-142.
- Gilson E, Segal-Bendirdjian E. 2010. The telomere story or the triumph of an open-minded research. *Biochimie* 92:321-326.
- Glusman G, Yanai I, Rubin I, Lancet D. 2001. The complete human olfactory subgenome. *Genome research* 11:685-702.
- Goetze S, Mateos-Langerak J, Gierman HJ, de Leeuw W, Giromus O, Indemans MH, Koster J, Ondrej V, Versteeg R, van Driel R. 2007a. The three-dimensional structure of human interphase chromosomes is related to the transcriptome map. *Molecular and cellular biology* 27:4475-4487.
- Goetze S, Mateos-Langerak J, van Driel R. 2007b. Three-dimensional genome organization in interphase and its relation to genome function. *Seminars in cell & developmental biology* 18:707-714.
- Goldman RD, Shumaker DK, Erdos MR, Eriksson M, Goldman AE, Gordon LB, Gruenbaum Y, Khuon S, Mendez M, Varga R, Collins FS. 2004. Accumulation of mutant lamin A

- causes progressive changes in nuclear architecture in Hutchinson-Gilford progeria syndrome. *Proceedings of the National Academy of Sciences of the United States of America* 101:8963-8968.
- Gonzalez-Suarez I, Redwood AB, Perkins SM, Vermolen B, Lichtensztejn D, Grotzky DA, Morgado-Palacin L, Gapud EJ, Sleckman BP, Sullivan T, Sage J, Stewart CL, Mai S, Gonzalo S. 2009. Novel roles for A-type lamins in telomere biology and the DNA damage response pathway. *The EMBO journal* 28:2414-2427.
- Gonzalo S, Garcia-Cao M, Fraga MF, Schotta G, Peters AH, Cotter SE, Eguia R, Dean DC, Esteller M, Jenuwein T, Blasco MA. 2005. Role of the RB1 family in stabilizing histone methylation at constitutive heterochromatin. *Nature cell biology* 7:420-428.
- Gonzalo S, Jaco I, Fraga MF, Chen T, Li E, Esteller M, Blasco MA. 2006. DNA methyltransferases control telomere length and telomere recombination in mammalian cells. *Nature cell biology* 8:416-424.
- Gotta M, Laroche T, Formenton A, Maillet L, Scherthan H, Gasser SM. 1996. The clustering of telomeres and colocalization with Rap1, Sir3, and Sir4 proteins in wild-type *Saccharomyces cerevisiae*. *The Journal of cell biology* 134:1349-1363.
- Gottschling DE, Aparicio OM, Billington BL, Zakian VA. 1990. Position effect at *S. cerevisiae* telomeres: reversible repression of Pol II transcription. *Cell* 63:751-762.
- Grasser F, Neusser M, Fiegler H, Thormeyer T, Cremer M, Carter NP, Cremer T, Muller S. 2008. Replication-timing-correlated spatial chromatin arrangements in cancer and in primate interphase nuclei. *Journal of cell science* 121:1876-1886.
- Green EK, Priestley MD, Waters J, Maliszewska C, Latif F, Maher ER. 2000. Detailed mapping of a congenital heart disease gene in chromosome 3p25. *Journal of medical genetics* 37:581-587.
- Greider CW, Blackburn EH. 1985. Identification of a specific telomere terminal transferase activity in *Tetrahymena* extracts. *Cell* 43:405-413.
- Gribnau J, Hochedlinger K, Hata K, Li E, Jaenisch R. 2003. Asynchronous replication timing of imprinted loci is independent of DNA methylation, but consistent with differential subnuclear localization. *Genes & development* 17:759-773.
- Griffith JD, Comeau L, Rosenfield S, Stansel RM, Bianchi A, Moss H, de Lange T. 1999. Mammalian telomeres end in a large duplex loop. *Cell* 97:503-514.
- Guelen L, Pagie L, Brassat E, Meuleman W, Faza MB, Talhout W, Eussen BH, de Klein A, Wessels L, de Laat W, van Steensel B. 2008. Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. *Nature* 453:948-951.
- Hara E, Tsurui H, Shinozaki A, Nakada S, Oda K. 1991. Cooperative effect of antisense-Rb and antisense-p53 oligomers on the extension of life span in human diploid fibroblasts, TIG-1. *Biochemical and biophysical research communications* 179:528-534.
- Hark AT, Schoenherr CJ, Katz DJ, Ingram RS, Levorse JM, Tilghman SM. 2000. CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus. *Nature* 405:486-489.
- Harrington L, Zhou W, McPhail T, Oulton R, Yeung DS, Mar V, Bass MB, Robinson MO. 1997. Human telomerase contains evolutionarily conserved catalytic and structural subunits. *Genes & development* 11:3109-3115.
- Hartlerode AJ, Scully R. 2009. Mechanisms of double-strand break repair in somatic mammalian cells. *The Biochemical journal* 423:157-168.
- Hayflick L. 1965. The Limited in Vitro Lifetime of Human Diploid Cell Strains. *Experimental cell research* 37:614-636.

- Heilstedt HA, Ballif BC, Howard LA, Kashork CD, Shaffer LG. 2003a. Population data suggest that deletions of 1p36 are a relatively common chromosome abnormality. *Clinical genetics* 64:310-316.
- Heilstedt HA, Ballif BC, Howard LA, Lewis RA, Stal S, Kashork CD, Bacino CA, Shapira SK, Shaffer LG. 2003b. Physical map of 1p36, placement of breakpoints in monosomy 1p36, and clinical characterization of the syndrome. *American journal of human genetics* 72:1200-1212.
- Hepperger C, Mannes A, Merz J, Peters J, Dietzel S. 2008. Three-dimensional positioning of genes in mouse cell nuclei. *Chromosoma* 117:535-551.
- Hernandez-Caballero E, Herrera-Gonzalez NE, Salamanca-Gomez F, Arenas-Aranda DJ. 2009. Role of telomere length in subtelomeric gene expression and its possible relation to cellular senescence. *BMB reports* 42:747-751.
- Herrera E, Samper E, Blasco MA. 1999. Telomere shortening in mTR^{-/-} embryos is associated with failure to close the neural tube. *The EMBO journal* 18:1172-1181.
- Hewitt JE, Lyle R, Clark LN, Valleley EM, Wright TJ, Wijmenga C, van Deutekom JC, Francis F, Sharpe PT, Hofker M, et al. 1994. Analysis of the tandem repeat locus D4Z4 associated with facioscapulohumeral muscular dystrophy. *Human molecular genetics* 3:1287-1295.
- Hiratani I, Ryba T, Itoh M, Yokochi T, Schwaiger M, Chang CW, Lyou Y, Townes TM, Schubeler D, Gilbert DM. 2008. Global reorganization of replication domains during embryonic stem cell differentiation. *PLoS biology* 6:e245.
- Ho CY, Murnane JP, Yeung AK, Ng HK, Lo AW. 2008. Telomeres acquire distinct heterochromatin characteristics during siRNA-induced RNA interference in mouse cells. *Curr Biol* 18:183-187.
- Hockemeyer D, Palm W, Else T, Daniels JP, Takai KK, Ye JZ, Keegan CE, de Lange T, Hammer GD. 2007. Telomere protection by mammalian Pot1 requires interaction with Tpp1. *Nature structural & molecular biology* 14:754-761.
- Hockemeyer D, Sfeir AJ, Shay JW, Wright WE, de Lange T. 2005. POT1 protects telomeres from a transient DNA damage response and determines how human chromosomes end. *The EMBO journal* 24:2667-2678.
- Holaska JM, Lee KK, Kowalski AK, Wilson KL. 2003. Transcriptional repressor germ cell-less (GCL) and barrier to autointegration factor (BAF) compete for binding to emerlin in vitro. *The Journal of biological chemistry* 278:6969-6975.
- Hollox EJ, Huffmeier U, Zeeuwen PL, Palla R, Lascorz J, Rodijk-Olthuis D, van de Kerkhof PC, Traupe H, de Jongh G, den Heijer M, Reis A, Armour JA, Schalkwijk J. 2008. Psoriasis is associated with increased beta-defensin genomic copy number. *Nature genetics* 40:23-25.
- Hoo JJ, Shrimpton AE. 2008. Distal 3p deletion is not necessarily associated with dysmorphic features or psychomotor delay. *Am J Med Genet A* 146A:538.
- Hsu HL, Gilley D, Blackburn EH, Chen DJ. 1999. Ku is associated with the telomere in mammals. *Proceedings of the National Academy of Sciences of the United States of America* 96:12454-12458.
- Hsu HL, Gilley D, Galande SA, Hande MP, Allen B, Kim SH, Li GC, Campisi J, Kohwi-Shigematsu T, Chen DJ. 2000. Ku acts in a unique way at the mammalian telomere to prevent end joining. *Genes & development* 14:2807-2812.
- Huang S, Risques RA, Martin GM, Rabinovitch PS, Oshima J. 2008. Accelerated telomere shortening and replicative senescence in human fibroblasts overexpressing mutant and wild-type lamin A. *Experimental cell research* 314:82-91.

- Huffman KE, Levene SD, Tesmer VM, Shay JW, Wright WE. 2000. Telomere shortening is proportional to the size of the G-rich telomeric 3'-overhang. *The Journal of biological chemistry* 275:19719-19722.
- Ishii T, Onda H, Tanigawa A, Ohshima S, Fujiwara H, Mima T, Katada Y, Deguchi H, Suemura M, Miyake T, Miyatake K, Kawase I, Zhao H, Tomiyama Y, Saeki Y, Nojima H. 2005. Isolation and expression profiling of genes upregulated in the peripheral blood cells of systemic lupus erythematosus patients. *DNA Res* 12:429-439.
- Ishikawa F, Naito T. 1999. Why do we have linear chromosomes? A matter of Adam and Eve. *Mutation research* 434:99-107.
- Jackson DA, Pombo A. 1998. Replicon clusters are stable units of chromosome structure: evidence that nuclear organization contributes to the efficient activation and propagation of S phase in human cells. *The Journal of cell biology* 140:1285-1295.
- Jain D, Cooper JP. 2010. Telomeric strategies: means to an end. *Annual review of genetics* 44:243-269.
- Jiang G, Yang F, van Overveld PG, Vedanarayanan V, van der Maarel S, Ehrlich M. 2003. Testing the position-effect variegation hypothesis for facioscapulohumeral muscular dystrophy by analysis of histone modification and gene expression in subtelomeric 4q. *Human molecular genetics* 12:2909-2921.
- Joffe B, Leonhardt H, Solovei I. 2010. Differentiation and large scale spatial organization of the genome. *Current opinion in genetics & development* 20:562-569.
- Kalkman JS, Schillings ML, van der Werf SP, Padberg GW, Zwarts MJ, van Engelen BG, Bleijenberg G. 2005. Experienced fatigue in facioscapulohumeral dystrophy, myotonic dystrophy, and HMSN-I. *Journal of neurology, neurosurgery, and psychiatry* 76:1406-1409.
- Kalverda B, Pickersgill H, Shloma VV, Fornerod M. 2010. Nucleoporins directly stimulate expression of developmental and cell-cycle genes inside the nucleoplasm. *Cell* 140:360-371.
- Karlseder J, Broccoli D, Dai Y, Hardy S, de Lange T. 1999. p53- and ATM-dependent apoptosis induced by telomeres lacking TRF2. *Science (New York, NY)* 283:1321-1325.
- Kim NW, Piatyszek MA, Prowse KR, Harley CB, West MD, Ho PL, Coviello GM, Wright WE, Weinrich SL, Shay JW. 1994. Specific association of human telomerase activity with immortal cells and cancer. *Science (New York, NY)* 266:2011-2015.
- Kind J, van Steensel B. 2010. Genome-nuclear lamina interactions and gene regulation. *Current opinion in cell biology* 22:320-325.
- Kipling D, Cooke HJ. 1990. Hypervariable ultra-long telomeres in mice. *Nature* 347:400-402.
- Kishi S, Zhou XZ, Ziv Y, Khoo C, Hill DE, Shiloh Y, Lu KP. 2001. Telomeric protein Pin2/TRF1 as an important ATM target in response to double strand DNA breaks. *The Journal of biological chemistry* 276:29282-29291.
- Klooster R, Straasheijm K, Shah B, Sowden J, Frants R, Thornton C, Tawil R, van der Maarel S. 2009. Comprehensive expression analysis of FSHD candidate genes at the mRNA and protein level. *Eur J Hum Genet* 17:1615-1624.
- Knight LA, Yong MH, Tan M, Ng IS. 1995. Del(3) (p25.3) without phenotypic effect. *Journal of medical genetics* 32:994-995.
- Koering CE, Pollice A, Zibella MP, Bauwens S, Puisieux A, Brunori M, Brun C, Martins L, Sabatier L, Pulitzer JF, Gilson E. 2002. Human telomeric position effect is determined by chromosomal context and telomeric chromatin integrity. *EMBO reports* 3:1055-1061.
- Koolen DA, Reardon W, Rosser EM, Lacombe D, Hurst JA, Law CJ, Bongers EM, van Ravenswaaij-Arts CM, Leisink MA, van Kessel AG, Veltman JA, de Vries BB. 2005. Molecular characterisation of patients with subtelomeric 22q abnormalities using

- chromosome specific array-based comparative genomic hybridisation. *Eur J Hum Genet* 13:1019-1024.
- Kowaljow V, Marcowycz A, Anseau E, Conde CB, Sauvage S, Matteotti C, Arias C, Corona ED, Nunez NG, Leo O, Wattiez R, Figlewicz D, Laoudj-Chenivresse D, Belayew A, Coppee F, Rosa AL. 2007. The DUX4 gene at the FSHD1A locus encodes a pro-apoptotic protein. *Neuromuscul Disord* 17:611-623.
- Krasnianski M, Neudecker S, Eger K, Jakubiczka S, Zierz S. 2003. Typical facioscapulohumeral dystrophy phenotype in patients without FSHD 4q35 deletion. *Journal of neurology* 250:1084-1087.
- Kulkarni A, Zschenker O, Reynolds G, Miller D, Murnane JP. 2010. Effect of telomere proximity on telomere position effect, chromosome healing, and sensitivity to DNA double-strand breaks in a human tumor cell line. *Molecular and cellular biology* 30:578-589.
- Kumaran RI, Spector DL. 2008. A genetic locus targeted to the nuclear periphery in living cells maintains its transcriptional competence. *The Journal of cell biology* 180:51-65.
- Kupper K, Kolbl A, Biener D, Dittrich S, von Hase J, Thormeyer T, Fiegler H, Carter NP, Speicher MR, Cremer T, Cremer M. 2007. Radial chromatin positioning is shaped by local gene density, not by gene expression. *Chromosoma* 116:285-306.
- Kurukuti S, Tiwari VK, Tavoosidana G, Pugacheva E, Murrell A, Zhao Z, Lobanenko V, Reik W, Ohlsson R. 2006. CTCF binding at the H19 imprinting control region mediates maternally inherited higher-order chromatin conformation to restrict enhancer access to Igf2. *Proceedings of the National Academy of Sciences of the United States of America* 103:10684-10689.
- Kyrion G, Boakye KA, Lustig AJ. 1992. C-terminal truncation of RAP1 results in the deregulation of telomere size, stability, and function in *Saccharomyces cerevisiae*. *Molecular and cellular biology* 12:5159-5173.
- Laforet P, de Toma C, Eymard B, Becane HM, Jeanpierre M, Fardeau M, Duboc D. 1998. Cardiac involvement in genetically confirmed facioscapulohumeral muscular dystrophy. *Neurology* 51:1454-1456.
- Laud PR, Multani AS, Bailey SM, Wu L, Ma J, Kingsley C, Lebel M, Pathak S, DePinho RA, Chang S. 2005. Elevated telomere-telomere recombination in WRN-deficient, telomere dysfunctional cells promotes escape from senescence and engagement of the ALT pathway. *Genes & development* 19:2560-2570.
- Ledbetter DH, Martin CL. 2007. Cryptic telomere imbalance: a 15-year update. *Am J Med Genet C Semin Med Genet* 145C:327-334.
- Lee HW, Blasco MA, Gottlieb GJ, Horner JW, 2nd, Greider CW, DePinho RA. 1998. Essential role of mouse telomerase in highly proliferative organs. *Nature* 392:569-574.
- Lee JH, Goto K, Matsuda C, Arahata K. 1995. Characterization of a tandemly repeated 3.3-kb KpnI unit in the facioscapulohumeral muscular dystrophy (FSHD) gene region on chromosome 4q35. *Muscle & nerve* 2:S6-13.
- Lee JT. 2009. Lessons from X-chromosome inactivation: long ncRNA as guides and tethers to the epigenome. *Genes & development* 23:1831-1842.
- Lemmers RJ, de Kievit P, Sandkuijl L, Padberg GW, van Ommen GJ, Frants RR, van der Maarel SM. 2002. Facioscapulohumeral muscular dystrophy is uniquely associated with one of the two variants of the 4q subtelomere. *Nature genetics* 32:235-236.
- Lemmers RJ, Osborn M, Haaf T, Rogers M, Frants RR, Padberg GW, Cooper DN, van der Maarel SM, Upadhyaya M. 2003. D4F104S1 deletion in facioscapulohumeral muscular dystrophy: phenotype, size, and detection. *Neurology* 61:178-183.
- Lemmers RJ, van der Vliet PJ, Klooster R, Sacconi S, Camano P, Dauwerse JG, Snider L, Straasheijm KR, van Ommen GJ, Padberg GW, Miller DG, Tapscott SJ, Tawil R,

- Frants RR, van der Maarel SM. 2010a. A unifying genetic model for facioscapulohumeral muscular dystrophy. *Science (New York, NY)* 329:1650-1653.
- Lemmers RJ, van der Vliet PJ, van der Gaag KJ, Zuniga S, Frants RR, de Knijff P, van der Maarel SM. 2010b. Worldwide population analysis of the 4q and 10q subtelomeres identifies only four discrete interchromosomal sequence transfers in human evolution. *American journal of human genetics* 86:364-377.
- Lemmers RJ, van der Wielen MJ, Bakker E, Padberg GW, Frants RR, van der Maarel SM. 2004a. Somatic mosaicism in FSHD often goes undetected. *Annals of neurology* 55:845-850.
- Lemmers RJ, Van Overveld PG, Sandkuijl LA, Vrieling H, Padberg GW, Frants RR, van der Maarel SM. 2004b. Mechanism and timing of mitotic rearrangements in the subtelomeric D4Z4 repeat involved in facioscapulohumeral muscular dystrophy. *American journal of human genetics* 75:44-53.
- Lemmers RJ, Wohlgemuth M, van der Gaag KJ, van der Vliet PJ, van Teijlingen CM, de Knijff P, Padberg GW, Frants RR, van der Maarel SM. 2007. Specific sequence variations within the 4q35 region are associated with facioscapulohumeral muscular dystrophy. *American journal of human genetics* 81:884-894.
- Lemmers RJL, de Kievit P, van Geel M, van der Wielen MJ, Bakker E, Padberg GW, Frants RR, van der Maarel SM. 2001. Complete allele information in the diagnosis of facioscapulohumeral muscular dystrophy by triple DNA analysis. *Annals of neurology* 50:816-819.
- Lenain C, Bauwens S, Amiard S, Brunori M, Giraud-Panis MJ, Gilson E. 2006. The Apollo 5' exonuclease functions together with TRF2 to protect telomeres from DNA repair. *Curr Biol* 16:1303-1310.
- Li B, Oestreich S, de Lange T. 2000. Identification of human Rap1: implications for telomere evolution. *Cell* 101:471-483.
- Li F, Chen J, Izumi M, Butler MC, Keezer SM, Gilbert DM. 2001. The replication timing program of the Chinese hamster beta-globin locus is established coincident with its repositioning near peripheral heterochromatin in early G1 phase. *The Journal of cell biology* 154:283-292.
- Linardopoulou EV, Parghi SS, Friedman C, Osborn GE, Parkhurst SM, Trask BJ. 2007. Human subtelomeric WASH genes encode a new subclass of the WASP family. *PLoS genetics* 3:e237.
- Linardopoulou EV, Williams EM, Fan Y, Friedman C, Young JM, Trask BJ. 2005. Human subtelomeres are hot spots of interchromosomal recombination and segmental duplication. *Nature* 437:94-100.
- Lipovich L, Johnson R, Lin CY. 2010. MacroRNA underdogs in a microRNA world: evolutionary, regulatory, and biomedical significance of mammalian long non-protein-coding RNA. *Biochimica et biophysica acta* 1799:597-615.
- Liu D, O'Connor MS, Qin J, Songyang Z. 2004. Telosome, a mammalian telomere-associated complex formed by multiple telomeric proteins. *The Journal of biological chemistry* 279:51338-51342.
- Liu L, Blasco M, Trimarchi J, Keefe D. 2002. An essential role for functional telomeres in mouse germ cells during fertilization and early development. *Developmental biology* 249:74-84.
- Luderus ME, van Steensel B, Chong L, Sibon OC, Cremers FF, de Lange T. 1996. Structure, subnuclear distribution, and nuclear matrix association of the mammalian telomeric complex. *The Journal of cell biology* 135:867-881.

- Luke B, Panza A, Redon S, Iglesias N, Li Z, Lingner J. 2008. The Rat1p 5' to 3' exonuclease degrades telomeric repeat-containing RNA and promotes telomere elongation in *Saccharomyces cerevisiae*. *Molecular cell* 32:465-477.
- Lundblad V, Blackburn EH. 1993. An alternative pathway for yeast telomere maintenance rescues est1- senescence. *Cell* 73:347-360.
- Lundblad V, Szostak JW. 1989. A mutant with a defect in telomere elongation leads to senescence in yeast. *Cell* 57:633-643.
- Lunt PW. 1998. 44th ENMC International Workshop: Facioscapulohumeral Muscular Dystrophy: Molecular Studies 19-21 July 1996, Naarden, The Netherlands. *Neuromuscul Disord* 8:126-130.
- Lunt PW, Jardine PE, Koch M, Maynard J, Osborn M, Williams M, Harper PS, Upadhyaya M. 1995. Phenotypic-genotypic correlation will assist genetic counseling in 4q35-facioscapulohumeral muscular dystrophy. *Muscle & nerve* 2:S103-109.
- Luo L, Gassman KL, Petell LM, Wilson CL, Bewersdorf J, Shopland LS. 2009. The nuclear periphery of embryonic stem cells is a transcriptionally permissive and repressive compartment. *Journal of cell science* 122:3729-3737.
- Ma H, Samarabandu J, Devdhar RS, Acharya R, Cheng PC, Meng C, Berezney R. 1998. Spatial and temporal dynamics of DNA replication sites in mammalian cells. *The Journal of cell biology* 143:1415-1425.
- Magdinier F, Yusufzai TM, Felsenfeld G. 2004. Both CTCF-dependent and -independent insulators are found between the mouse T cell receptor alpha and Dad1 genes. *The Journal of biological chemistry* 279:25381-25389.
- Mahadevan MS. 2010. Genetics. Exposing a DUX tale. *Science (New York, NY)* 329:1607-1608.
- Mahy NL, Perry PE, Bickmore WA. 2002a. Gene density and transcription influence the localization of chromatin outside of chromosome territories detectable by FISH. *The Journal of cell biology* 159:753-763.
- Mahy NL, Perry PE, Gilchrist S, Baldock RA, Bickmore WA. 2002b. Spatial organization of active and inactive genes and noncoding DNA within chromosome territories. *The Journal of cell biology* 157:579-589.
- Makarov VL, Hirose Y, Langmore JP. 1997. Long G tails at both ends of human chromosomes suggest a C strand degradation mechanism for telomere shortening. *Cell* 88:657-666.
- Markiewicz E, Dechat T, Foisner R, Quinlan RA, Hutchison CJ. 2002. Lamin A/C binding protein LAP2alpha is required for nuclear anchorage of retinoblastoma protein. *Molecular biology of the cell* 13:4401-4413.
- Martianov I, Ramadass A, Serra Barros A, Chow N, Akoulitchev A. 2007. Repression of the human dihydrofolate reductase gene by a non-coding interfering transcript. *Nature* 445:666-670.
- Masny PS, Bengtsson U, Chung SA, Martin JH, van Engelen B, van der Maarel SM, Winokur ST. 2004. Localization of 4q35.2 to the nuclear periphery: is FSHD a nuclear envelope disease? *Human molecular genetics* 13:1857-1871.
- Masny PS, Chan OY, de Greef JC, Bengtsson U, Ehrlich M, Tawil R, Lock LF, Hewitt JE, Stocksdales J, Martin JH, van der Maarel SM, Winokur ST. 2010. Analysis of allele-specific RNA transcription in FSHD by RNA-DNA FISH in single myonuclei. *Eur J Hum Genet* 18:448-456.
- McClintock B. 1941. The stability of broken ends of chromosomes in *Zea mays*. *Genetics* 26:234-282.
- McCullough BJ, Adams JC, Shilling DJ, Feeney MP, Sie KC, Tempel BL. 2007. 3p-- syndrome defines a hearing loss locus in 3p25.3. *Hearing research* 224:51-60.

- McElligott R, Wellinger RJ. 1997. The terminal DNA structure of mammalian chromosomes. *The EMBO journal* 16:3705-3714.
- Mefford HC, Trask BJ. 2002. The complex structure and dynamic evolution of human subtelomeres. *Nat Rev Genet* 3:91-102.
- Menon BB, Sarma NJ, Pasula S, Deminoff SJ, Willis KA, Barbara KE, Andrews B, Santangelo GM. 2005. Reverse recruitment: the Nup84 nuclear pore subcomplex mediates Rap1/Gcr1/Gcr2 transcriptional activation. *Proceedings of the National Academy of Sciences of the United States of America* 102:5749-5754.
- Meyerson M, Counter CM, Eaton EN, Ellisen LW, Steiner P, Caddle SD, Ziaugra L, Beijersbergen RL, Davidoff MJ, Liu Q, Bacchetti S, Haber DA, Weinberg RA. 1997. hEST2, the putative human telomerase catalytic subunit gene, is up-regulated in tumor cells and during immortalization. *Cell* 90:785-795.
- Millard TH, Sharp SJ, Machesky LM. 2004. Signalling to actin assembly via the WASP (Wiskott-Aldrich syndrome protein)-family proteins and the Arp2/3 complex. *The Biochemical journal* 380:1-17.
- Miyake Y, Nakamura M, Nabetani A, Shimamura S, Tamura M, Yonehara S, Saito M, Ishikawa F. 2009. RPA-like mammalian Ctc1-Stn1-Ten1 complex binds to single-stranded DNA and protects telomeres independently of the Pot1 pathway. *Molecular cell* 36:193-206.
- Mostoslavsky R, Chua KF, Lombard DB, Pang WW, Fischer MR, Gellon L, Liu P, Mostoslavsky G, Franco S, Murphy MM, Mills KD, Patel P, Hsu JT, Hong AL, Ford E, Cheng HL, Kennedy C, Nunez N, Bronson R, Frendewey D, Auerbach W, Valenzuela D, Karow M, Hottiger MO, Hursting S, Barrett JC, Guarente L, Mulligan R, Demple B, Yancopoulos GD, Alt FW. 2006. Genomic instability and aging-like phenotype in the absence of mammalian SIRT6. *Cell* 124:315-329.
- Moyzis RK, Buckingham JM, Cram LS, Dani M, Deaven LL, Jones MD, Meyne J, Ratliff RL, Wu JR. 1988. A highly conserved repetitive DNA sequence, (TTAGGG)_n, present at the telomeres of human chromosomes. *Proceedings of the National Academy of Sciences of the United States of America* 85:6622-6626.
- Muller HJ. 1930. Types of visible variations induced by X-ray in *Drosophila*. *J Genet* 22:299-334.
- Muller HJ. 1932. Further studies on the nature and causes of gene mutations. *Proc Sixth Int Congr Genet* 1:213-255.
- Muller HJ. 1938. The remaking of chromosomes. *Collect Net* 13:181-195.
- Munkel C, Eils R, Dietzel S, Zink D, Mehring C, Wedemann G, Cremer T, Langowski J. 1999. Compartmentalization of interphase chromosomes observed in simulation and experiment. *Journal of molecular biology* 285:1053-1065.
- Murmann AE, Gao J, Encinosa M, Gautier M, Peter ME, Eils R, Lichter P, Rowley JD. 2005. Local gene density predicts the spatial position of genetic loci in the interphase nucleus. *Experimental cell research* 311:14-26.
- Nagele RG, Velasco AQ, Anderson WJ, McMahon DJ, Thomson Z, Fazekas J, Wind K, Lee H. 2001. Telomere associations in interphase nuclei: possible role in maintenance of interphase chromosome topology. *Journal of cell science* 114:377-388.
- Nakamura TM, Morin GB, Chapman KB, Weinrich SL, Andrews WH, Lingner J, Harley CB, Cech TR. 1997. Telomerase catalytic subunit homologs from fission yeast and human. *Science (New York, NY)* 277:955-959.
- Naumov G, Turakainen H, Naumova E, Aho S, Korhola M. 1990. A new family of polymorphic genes in *Saccharomyces cerevisiae*: alpha-galactosidase genes MEL1-MEL7. *Mol Gen Genet* 224:119-128.

- Nergadze SG, Farnung BO, Wischnewski H, Khoriantse L, Vitelli V, Chawla R, Giulotto E, Azzalin CM. 2009. CpG-island promoters drive transcription of human telomeres. *RNA (New York, NY)* 15:2186-2194.
- Nguyen K, Walfaren P, Chaix C, Vo Van C, Bernard R, Attarian S, Pouget J, Vannier A, Dufrane N, Bensimon A, Levy N. 2010. Visualisation directe des haplotypes 4qter et 10qter par peignage moléculaire, un nouveau test diagnostique de la dystrophie facio-scapulo-humérale. *Medecine Science* 26:13.
- Nickerson J. 2001. Experimental observations of a nuclear matrix. *Journal of cell science* 114:463-474.
- Niimura Y, Nei M. 2007. Extensive gains and losses of olfactory receptor genes in mammalian evolution. *PloS one* 2:e708.
- Nikitina T, Woodcock CL. 2004. Closed chromatin loops at the ends of chromosomes. *The Journal of cell biology* 166:161-165.
- Nili E, Cojocaru GS, Kalma Y, Ginsberg D, Copeland NG, Gilbert DJ, Jenkins NA, Berger R, Shaklai S, Amariglio N, Brok-Simoni F, Simon AJ, Rechavi G. 2001. Nuclear membrane protein LAP2beta mediates transcriptional repression alone and together with its binding partner GCL (germ-cell-less). *Journal of cell science* 114:3297-3307.
- Ning Y, Xu JF, Li Y, Chavez L, Riethman HC, Lansdorp PM, Weng NP. 2003. Telomere length and the expression of natural telomeric genes in human fibroblasts. *Human molecular genetics* 12:1329-1336.
- Nittis T, Guittat L, Stewart SA. 2008. Alternative lengthening of telomeres (ALT) and chromatin: is there a connection? *Biochimie* 90:5-12.
- O'Connor MS, Safari A, Liu D, Qin J, Songyang Z. 2004. The human Rap1 protein complex and modulation of telomere length. *The Journal of biological chemistry* 279:28585-28591.
- Ofir R, Wong AC, McDermid HE, Skorecki KL, Selig S. 1999. Position effect of human telomeric repeats on replication timing. *Proceedings of the National Academy of Sciences of the United States of America* 96:11434-11439.
- Olovnikov AM. 1973. A theory of marginotomy. The incomplete copying of template margin in enzymic synthesis of polynucleotides and biological significance of the phenomenon. *Journal of theoretical biology* 41:181-190.
- Opresko PL, von Kobbe C, Laine JP, Harrigan J, Hickson ID, Bohr VA. 2002. Telomere-binding protein TRF2 binds to and stimulates the Werner and Bloom syndrome helicases. *The Journal of biological chemistry* 277:41110-41119.
- Osborne RJ, Welle S, Venance SL, Thornton CA, Tawil R. 2007. Expression profile of FSHD supports a link between retinal vasculopathy and muscular dystrophy. *Neurology* 68:569-577.
- Osterhage JL, Friedman KL. 2009. Chromosome end maintenance by telomerase. *The Journal of biological chemistry* 284:16061-16065.
- Ottaviani A, Gilson E, Magdinier F. 2008. Telomeric position effect: from the yeast paradigm to human pathologies? *Biochimie* 90:93-107.
- Ottaviani A, Rival-Gervier S, Boussouar A, Foerster AM, Rondier D, Sacconi S, Desnuelle C, Gilson E, Magdinier F. 2009a. The D4Z4 macrosatellite repeat acts as a CTCF and A-type lamins-dependent insulator in facio-scapulo-humeral dystrophy. *PLoS genetics* 5:e1000394.
- Ottaviani A, Schluth-Bolard C, Gilson E, Magdinier F. 2010. D4Z4 as a prototype of CTCF and lamins-dependent insulator in human cells. *Nucleus (Austin, Tex)* 1:30-36.
- Ottaviani A, Schluth-Bolard C, Rival-Gervier S, Boussouar A, Rondier D, Foerster AM, Morere J, Bauwens S, Gazzo S, Callet-Bauchu E, Gilson E, Magdinier F. 2009b.

- Identification of a perinuclear positioning element in human subtelomeres that requires A-type lamins and CTCF. *The EMBO journal* 28:2428-2436.
- Padberg GW, Brouwer OF, de Keizer RJ, Dijkman G, Wijmenga C, Grote JJ, Frants RR. 1995. On the significance of retinal vascular disease and hearing loss in facioscapulohumeral muscular dystrophy. *Muscle & nerve* 2:S73-80.
- Padberg GW, Lunt PW, Koch M, Fardeau M. 1991. Diagnostic criteria for facioscapulohumeral muscular dystrophy. *Neuromuscul Disord* 1:231-234.
- Palm W, de Lange T. 2008. How shelterin protects mammalian telomeres. *Annual review of genetics* 42:301-334.
- Pandey RR, Mondal T, Mohammad F, Enroth S, Redrup L, Komorowski J, Nagano T, Mancini-Dinardo D, Kanduri C. 2008. Kcnq1ot1 antisense noncoding RNA mediates lineage-specific transcriptional silencing through chromatin-level regulation. *Molecular cell* 32:232-246.
- Parseghian MH, Newcomb RL, Hamkalo BA. 2001. Distribution of somatic H1 subtypes is non-random on active vs. inactive chromatin II: distribution in human adult fibroblasts. *Journal of cellular biochemistry* 83:643-659.
- Pedram M, Sprung CN, Gao Q, Lo AW, Reynolds GE, Murnane JP. 2006. Telomere position effect and silencing of transgenes near telomeres in the mouse. *Molecular and cellular biology* 26:1865-1878.
- Petrov A, Allinne J, Pirozhkova I, Laoudj D, Lipinski M, Vassetzky YS. 2008. A nuclear matrix attachment site in the 4q35 locus has an enhancer-blocking activity in vivo: implications for the facio-scapulo-humeral dystrophy. *Genome research* 18:39-45.
- Petrov A, Pirozhkova I, Carnac G, Laoudj D, Lipinski M, Vassetzky YS. 2006. Chromatin loop domain organization within the 4q35 locus in facioscapulohumeral dystrophy patients versus normal human myoblasts. *Proceedings of the National Academy of Sciences of the United States of America* 103:6982-6987.
- Phillips JE, Corces VG. 2009. CTCF: master weaver of the genome. *Cell* 137:1194-1211.
- Pisano S, Galati A, Cacchione S. 2008. Telomeric nucleosomes: forgotten players at chromosome ends. *Cell Mol Life Sci* 65:3553-3563.
- Pisano S, Leoni D, Galati A, Rhodes D, Savino M, Cacchione S. 2010. The human telomeric protein hTRF1 induces telomere-specific nucleosome mobility. *Nucleic acids research* 38:2247-2255.
- Pohjola P, de Leeuw N, Penttinen M, Kaariainen H. 2010. Terminal 3p deletions in two families--correlation between molecular karyotype and phenotype. *Am J Med Genet A* 152A:441-446.
- Porro A, Feuerhahn S, Reichenbach P, Lingner J. 2010. Molecular dissection of telomeric repeat-containing RNA biogenesis unveils the presence of distinct and multiple regulatory pathways. *Molecular and cellular biology* 30:4808-4817.
- Pryde FE, Louis EJ. 1999. Limitations of silencing at native yeast telomeres. *The EMBO journal* 18:2538-2550.
- Ravnan JB, Tepperberg JH, Papenhausen P, Lamb AN, Hedrick J, Eash D, Ledbetter DH, Martin CL. 2006. Subtelomere FISH analysis of 11 688 cases: an evaluation of the frequency and pattern of subtelomere rearrangements in individuals with developmental disabilities. *Journal of medical genetics* 43:478-489.
- Raz V, Vermolen BJ, Garini Y, Onderwater JJ, Mommaas-Kienhuis MA, Koster AJ, Young IT, Tanke H, Dirks RW. 2008. The nuclear lamina promotes telomere aggregation and centromere peripheral localization during senescence of human mesenchymal stem cells. *Journal of cell science* 121:4018-4028.
- Reddy KL, Zullo JM, Bertolino E, Singh H. 2008. Transcriptional repression mediated by repositioning of genes to the nuclear lamina. *Nature* 452:243-247.

- Redon R, Ishikawa S, Fitch KR, Feuk L, Perry GH, Andrews TD, Fiegler H, Shapero MH, Carson AR, Chen W, Cho EK, Dallaire S, Freeman JL, Gonzalez JR, Gratacos M, Huang J, Kalaitzopoulos D, Komura D, MacDonald JR, Marshall CR, Mei R, Montgomery L, Nishimura K, Okamura K, Shen F, Somerville MJ, Tchinda J, Valsesia A, Woodwark C, Yang F, Zhang J, Zerjal T, Zhang J, Armengol L, Conrad DF, Estivill X, Tyler-Smith C, Carter NP, Aburatani H, Lee C, Jones KW, Scherer SW, Hurles ME. 2006. Global variation in copy number in the human genome. *Nature* 444:444-454.
- Redon R, Rio M, Gregory SG, Cooper RA, Fiegler H, Sanlaville D, Banerjee R, Scott C, Carr P, Langford C, Cormier-Daire V, Munnich A, Carter NP, Colleaux L. 2005. Tiling path resolution mapping of constitutional 1p36 deletions by array-CGH: contiguous gene deletion or "deletion with positional effect" syndrome? *Journal of medical genetics* 42:166-171.
- Redon S, Reichenbach P, Lingner J. 2010. The non-coding RNA TERRA is a natural ligand and direct inhibitor of human telomerase. *Nucleic acids research* 38:5797-5806.
- Riethman H. 2008a. Human subtelomeric copy number variations. *Cytogenetic and genome research* 123:244-252.
- Riethman H. 2008b. Human telomere structure and biology. *Annual review of genomics and human genetics* 9:1-19.
- Riethman H, Ambrosini A, Castaneda C, Finklestein J, Hu XL, Mudunuri U, Paul S, Wei J. 2004. Mapping and initial analysis of human subtelomeric sequence assemblies. *Genome research* 14:18-28.
- Riethman HC, Xiang Z, Paul S, Morse E, Hu XL, Flint J, Chi HC, Grady DL, Moyzis RK. 2001. Integration of telomere sequences with the draft human genome sequence. *Nature* 409:948-951.
- Rincon-Arano H, Furlan-Magaril M, Recillas-Targa F. 2007. Protection against telomeric position effects by the chicken cHS4 beta-globin insulator. *Proceedings of the National Academy of Sciences of the United States of America* 104:14044-14049.
- Rinn JL, Kertesz M, Wang JK, Squazzo SL, Xu X, Bruggmann SA, Goodnough LH, Helms JA, Farnham PJ, Segal E, Chang HY. 2007. Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. *Cell* 129:1311-1323.
- Rossi M, Ricci E, Colantoni L, Galluzzi G, Frusciante R, Tonali PA, Felicetti L. 2007. The Facioscapulohumeral muscular dystrophy region on 4qter and the homologous locus on 10qter evolved independently under different evolutionary pressure. *BMC medical genetics* 8:8.
- Rouquette J, Genoud C, Vazquez-Nin GH, Kraus B, Cremer T, Fakan S. 2009. Revealing the high-resolution three-dimensional network of chromatin and interchromatin space: a novel electron-microscopic approach to reconstructing nuclear architecture. *Chromosome Res* 17:801-810.
- Ruault M, Dubarry M, Taddei A. 2008. Re-positioning genes to the nuclear envelope in mammalian cells: impact on transcription. *Trends Genet* 24:574-581.
- Rudenko G, Van der Ploeg LH. 1989. Transcription of telomere repeats in protozoa. *The EMBO journal* 8:2633-2638.
- Sachs RK, van den Engh G, Trask B, Yokota H, Hearst JE. 1995. A random-walk/giant-loop model for interphase chromosomes. *Proceedings of the National Academy of Sciences of the United States of America* 92:2710-2714.
- Sarfaraizi M, Wijmenga C, Upadhyaya M, Weiffenbach B, Hyser C, Mathews K, Murray J, Gilbert J, Pericak-Vance M, Lunt P, et al. 1992. Regional mapping of

- facioscapulohumeral muscular dystrophy gene on 4q35: combined analysis of an international consortium. *American journal of human genetics* 51:396-403.
- Sasaki YT, Ideue T, Sano M, Mituyama T, Hirose T. 2009. MENepsilon/beta noncoding RNAs are essential for structural integrity of nuclear paraspeckles. *Proceedings of the National Academy of Sciences of the United States of America* 106:2525-2530.
- Schneider R, Grosschedl R. 2007. Dynamics and interplay of nuclear architecture, genome organization, and gene expression. *Genes & development* 21:3027-3043.
- Schoeftner S, Blasco MA. 2008. Developmentally regulated transcription of mammalian telomeres by DNA-dependent RNA polymerase II. *Nature cell biology* 10:228-236.
- Schotta G, Lachner M, Sarma K, Ebert A, Sengupta R, Reuter G, Reinberg D, Jenuwein T. 2004. A silencing pathway to induce H3-K9 and H4-K20 trimethylation at constitutive heterochromatin. *Genes & development* 18:1251-1262.
- Sfeir A, Kabir S, van Overbeek M, Celli GB, de Lange T. 2010. Loss of Rap1 induces telomere recombination in the absence of NHEJ or a DNA damage signal. *Science (New York, NY)* 327:1657-1661.
- Shiloh Y. 2003. ATM and related protein kinases: safeguarding genome integrity. *Nature reviews* 3:155-168.
- Shimi T, Pfliegerhaer K, Kojima S, Pack CG, Solovei I, Goldman AE, Adam SA, Shumaker DK, Kinjo M, Cremer T, Goldman RD. 2008. The A- and B-type nuclear lamin networks: microdomains involved in chromatin organization and transcription. *Genes & development* 22:3409-3421.
- Shuib S, McMullan D, Rattenberry E, Barber RM, Rahman F, Zatyka M, Chapman C, Macdonald F, Latif F, Davison V, Maher ER. 2009. Microarray based analysis of 3p25-p26 deletions (3p- syndrome). *Am J Med Genet A* 149A:2099-2105.
- Siderakis M, Tarsounas M. 2007. Telomere regulation and function during meiosis. *Chromosome Res* 15:667-679.
- Smogorzewska A, de Lange T. 2004. Regulation of telomerase by telomeric proteins. *Annual review of biochemistry* 73:177-208.
- Snider L, Asawachaicharn A, Tyler AE, Geng LN, Petek LM, Maves L, Miller DG, Lemmers RJ, Winokur ST, Tawil R, van der Maarel SM, Filippova GN, Tapscott SJ. 2009. RNA transcripts, miRNA-sized fragments and proteins produced from D4Z4 units: new candidates for the pathophysiology of facioscapulohumeral dystrophy. *Human molecular genetics* 18:2414-2430.
- Snider L, Geng LN, Lemmers RJ, Kyba M, Ware CB, Nelson AM, Tawil R, Filippova GN, van der Maarel SM, Tapscott SJ, Miller DG. 2010. Facioscapulohumeral dystrophy: incomplete suppression of a retrotransposed gene. *PLoS genetics* 6:e1001181.
- Solovei I, Gaginskaya ER, Macgregor HC. 1994. The arrangement and transcription of telomere DNA sequences at the ends of lampbrush chromosomes of birds. *Chromosome Res* 2:460-470.
- Sprung CN, Sabatier L, Murnane JP. 1996. Effect of telomere length on telomeric gene expression. *Nucleic acids research* 24:4336-4340.
- Stansel RM, de Lange T, Griffith JD. 2001. T-loop assembly in vitro involves binding of TRF2 near the 3' telomeric overhang. *The EMBO journal* 20:5532-5540.
- Starr DA. 2009. A nuclear-envelope bridge positions nuclei and moves chromosomes. *Journal of cell science* 122:577-586.
- Strahl BD, Allis CD. 2000. The language of covalent histone modifications. *Nature* 403:41-45.
- Sturtevant A. 1925. The effect of unequal crossing over at the bar locus in *Drosophila*. *Genetics* 10:117-147.

- Taddei A, Hediger F, Neumann FR, Gasser SM. 2004. The function of nuclear architecture: a genetic approach. *Annual review of genetics* 38:305-345.
- Taimen P, Pfliegerhaa K, Shimi T, Moller D, Ben-Harush K, Erdos MR, Adam SA, Herrmann H, Medalia O, Collins FS, Goldman AE, Goldman RD. 2009. A progeria mutation reveals functions for lamin A in nuclear assembly, architecture, and chromosome organization. *Proceedings of the National Academy of Sciences of the United States of America* 106:20788-20793.
- Takagishi J, Rauhen KA, Drumheller T, Kousseff B, Sutcliffe M. 2006. Chromosome 3p25 deletion in mother and daughter with minimal phenotypic effect. *Am J Med Genet A* 140:1587-1593.
- Takai H, Smogorzewska A, de Lange T. 2003. DNA damage foci at dysfunctional telomeres. *Curr Biol* 13:1549-1556.
- Takenawa T, Suetsugu S. 2007. The WASP-WAVE protein network: connecting the membrane to the cytoskeleton. *Nat Rev Mol Cell Biol* 8:37-48.
- Tam R, Smith KP, Lawrence JB. 2004. The 4q subtelomere harboring the FSHD locus is specifically anchored with peripheral heterochromatin unlike most human telomeres. *The Journal of cell biology* 167:269-279.
- Tawil R, Figlewicz DA, Griggs RC, Weiffenbach B. 1998. Facioscapulohumeral dystrophy: a distinct regional myopathy with a novel molecular pathogenesis. FSH Consortium. *Annals of neurology* 43:279-282.
- Tawil R, Van Der Maarel SM. 2006. Facioscapulohumeral muscular dystrophy. *Muscle & nerve* 34:1-15.
- Teng SC, Zakian VA. 1999. Telomere-telomere recombination is an efficient bypass pathway for telomere maintenance in *Saccharomyces cerevisiae*. *Molecular and cellular biology* 19:8083-8093.
- Tennen RI, Chua KF. 2011. Chromatin regulation and genome maintenance by mammalian SIRT6. *Trends in biochemical sciences* 36:39-46.
- Tilman G, Loriot A, Van Beneden A, Arnoult N, Londono-Vallejo JA, De Smet C, Decottignies A. 2009. Subtelomeric DNA hypomethylation is not required for telomeric sister chromatid exchanges in ALT cells. *Oncogene* 28:1682-1693.
- Tomita K, Cooper JP. 2007. The telomere bouquet controls the meiotic spindle. *Cell* 130:113-126.
- Tommerup H, Dousmanis A, de Lange T. 1994. Unusual chromatin in human telomeres. *Molecular and cellular biology* 14:5777-5785.
- Tonini MM, Passos-Bueno MR, Cerqueira A, Matioli SR, Pavanello R, Zatz M. 2004. Asymptomatic carriers and gender differences in facioscapulohumeral muscular dystrophy (FSHD). *Neuromuscul Disord* 14:33-38.
- Tremblay DC, Alexander G, Jr., Moseley S, Chadwick BP. 2010. Expression, tandem repeat copy number variation and stability of four macrosatellite arrays in the human genome. *BMC genomics* 11:632.
- Trevisan CP, Pastorello E, Armani M, Angelini C, Nante G, Tomelleri G, Tonin P, Mongini T, Palmucci L, Galluzzi G, Tupler RG, Barchitta A. 2006. Facioscapulohumeral muscular dystrophy and occurrence of heart arrhythmia. *European neurology* 56:1-5.
- Tsien F, Sun B, Hopkins NE, Vedanarayanan V, Figlewicz D, Winokur S, Ehrlich M. 2001. Methylation of the FSHD syndrome-linked subtelomeric repeat in normal and FSHD cell cultures and tissues. *Molecular genetics and metabolism* 74:322-331.
- Tupler R, Berardinelli A, Barbierato L, Frants R, Hewitt JE, Lanzi G, Maraschio P, Tiepolo L. 1996. Monosomy of distal 4q does not cause facioscapulohumeral muscular dystrophy. *Journal of medical genetics* 33:366-370.

- Tyson C, McGillivray B, Chijiwa C, Rajcan-Separovic E. 2004. Elucidation of a cryptic interstitial 7q31.3 deletion in a patient with a language disorder and mild mental retardation by array-CGH. *Am J Med Genet A* 129A:254-260.
- Uhlirova R, Horakova AH, Galiova G, Legartova S, Matula P, Fojtova M, Varecha M, Amrichova J, Vondracek J, Kozubek S, Bartova E. 2010. SUV39h- and A-type lamin-dependent telomere nuclear rearrangement. *Journal of cellular biochemistry* 109:915-926.
- van der Maarel SM, Deidda G, Lemmers RJ, van Overveld PG, van der Wielen M, Hewitt JE, Sandkuijl L, Bakker B, van Ommen GJ, Padberg GW, Frants RR. 2000. De novo facioscapulohumeral muscular dystrophy: frequent somatic mosaicism, sex-dependent phenotype, and the role of mitotic transchromosomal repeat interaction between chromosomes 4 and 10. *American journal of human genetics* 66:26-35.
- van der Maarel SM, Frants RR. 2005. The D4Z4 repeat-mediated pathogenesis of facioscapulohumeral muscular dystrophy. *American journal of human genetics* 76:375-386.
- van Deutekom JC, Lemmers RJ, Grewal PK, van Geel M, Romberg S, Dauwerse HG, Wright TJ, Padberg GW, Hofker MH, Hewitt JE, Frants RR. 1996. Identification of the first gene (FRG1) from the FSHD region on human chromosome 4q35. *Human molecular genetics* 5:581-590.
- van Geel M, Dickson MC, Beck AF, Bolland DJ, Frants RR, van der Maarel SM, de Jong PJ, Hewitt JE. 2002. Genomic analysis of human chromosome 10q and 4q telomeres suggests a common origin. *Genomics* 79:210-217.
- van Geel M, van Deutekom JC, van Staalduinen A, Lemmers RJ, Dickson MC, Hofker MH, Padberg GW, Hewitt JE, de Jong PJ, Frants RR. 2000. Identification of a novel beta-tubulin subfamily with one member (TUBB4Q) located near the telomere of chromosome region 4q35. *Cytogenetics and cell genetics* 88:316-321.
- van Overveld PG, Enthoven L, Ricci E, Rossi M, Felicetti L, Jeanpierre M, Winokur ST, Frants RR, Padberg GW, van der Maarel SM. 2005. Variable hypomethylation of D4Z4 in facioscapulohumeral muscular dystrophy. *Annals of neurology* 58:569-576.
- van Overveld PG, Lemmers RJ, Sandkuijl LA, Enthoven L, Winokur ST, Bakels F, Padberg GW, van Ommen GJ, Frants RR, van der Maarel SM. 2003. Hypomethylation of D4Z4 in 4q-linked and non-4q-linked facioscapulohumeral muscular dystrophy. *Nature genetics* 35:315-317.
- van Steensel B, de Lange T. 1997. Control of telomere length by the human telomeric protein TRF1. *Nature* 385:740-743.
- van Steensel B, Smogorzewska A, de Lange T. 1998. TRF2 protects human telomeres from end-to-end fusions. *Cell* 92:401-413.
- Venteicher AS, Abreu EB, Meng Z, McCann KE, Terns RM, Veenstra TD, Terns MP, Artandi SE. 2009. A human telomerase holoenzyme protein required for Cajal body localization and telomere synthesis. *Science (New York, NY)* 323:644-648.
- Venteicher AS, Meng Z, Mason PJ, Veenstra TD, Artandi SE. 2008. Identification of ATPases pontin and reptin as telomerase components essential for holoenzyme assembly. *Cell* 132:945-957.
- Verschure PJ, van Der Kraan I, Manders EM, van Driel R. 1999. Spatial relationship between transcription sites and chromosome territories. *The Journal of cell biology* 147:13-24.
- Volpi EV, Chevret E, Jones T, Vatcheva R, Williamson J, Beck S, Campbell RD, Goldsworthy M, Powis SH, Ragoussis J, Trowsdale J, Sheer D. 2000. Large-scale chromatin organization of the major histocompatibility complex and other regions of human chromosome 6 and its response to interferon in interphase nuclei. *Journal of cell science* 113 (Pt 9):1565-1576.

- Wallace LM, Garwick SE, Mei W, Belayew A, Coppee F, Ladner KJ, Guttridge D, Yang J, Harper SQ. 2010. DUX4, a candidate gene for facioscapulohumeral muscular dystrophy, causes p53-dependent myopathy in vivo. *Annals of neurology*.
- Walter J, Schermelleh L, Cremer M, Tashiro S, Cremer T. 2003. Chromosome order in HeLa cells changes during mitosis and early G1, but is stably maintained during subsequent interphase stages. *The Journal of cell biology* 160:685-697.
- Wang RC, Smogorzewska A, de Lange T. 2004. Homologous recombination generates T-loop-sized deletions at human telomeres. *Cell* 119:355-368.
- Watson JD. 1972. Origin of concatemeric T7 DNA. *Nature: New biology* 239:197-201.
- Weierich C, Brero A, Stein S, von Hase J, Cremer C, Cremer T, Solovei I. 2003. Three-dimensional arrangements of centromeres and telomeres in nuclei of human and murine lymphocytes. *Chromosome Res* 11:485-502.
- Weinrich SL, Pruzan R, Ma L, Ouellette M, Tesmer VM, Holt SE, Bodnar AG, Lichtsteiner S, Kim NW, Trager JB, Taylor RD, Carlos R, Andrews WH, Wright WE, Shay JW, Harley CB, Morin GB. 1997. Reconstitution of human telomerase with the template RNA component hTR and the catalytic protein subunit hTRT. *Nature genetics* 17:498-502.
- West AG, Gaszner M, Felsenfeld G. 2002. Insulators: many functions, many mechanisms. *Genes & development* 16:271-288.
- Wijmenga C, Frants RR, Brouwer OF, Moerer P, Weber JL, Padberg GW. 1990. Location of facioscapulohumeral muscular dystrophy gene on chromosome 4. *Lancet* 336:651-653.
- Wijmenga C, Hewitt JE, Sandkuijl LA, Clark LN, Wright TJ, Dauwerse HG, Gruter AM, Hofker MH, Moerer P, Williamson R, et al. 1992. Chromosome 4q DNA rearrangements associated with facioscapulohumeral muscular dystrophy. *Nature genetics* 2:26-30.
- Wilkie AO, Higgs DR, Rack KA, Buckle VJ, Spurr NK, Fischel-Ghodsian N, Ceccherini I, Brown WR, Harris PC. 1991. Stable length polymorphism of up to 260 kb at the tip of the short arm of human chromosome 16. *Cell* 64:595-606.
- Wilson HL, Wong AC, Shaw SR, Tse WY, Stapleton GA, Phelan MC, Hu S, Marshall J, McDermid HE. 2003. Molecular characterisation of the 22q13 deletion syndrome supports the role of haploinsufficiency of SHANK3/PROSAP2 in the major neurological symptoms. *Journal of medical genetics* 40:575-584.
- Wilusz JE, Sunwoo H, Spector DL. 2009. Long noncoding RNAs: functional surprises from the RNA world. *Genes & development* 23:1494-1504.
- Winokur ST, Bengtsson U, Feddersen J, Mathews KD, Weiffenbach B, Bailey H, Markovich RP, Murray JC, Wasmuth JJ, Altherr MR, et al. 1994. The DNA rearrangement associated with facioscapulohumeral muscular dystrophy involves a heterochromatin-associated repetitive element: implications for a role of chromatin structure in the pathogenesis of the disease. *Chromosome Res* 2:225-234.
- Winokur ST, Chen YW, Masny PS, Martin JH, Ehmsen JT, Tapscott SJ, van der Maarel SM, Hayashi Y, Flanigan KM. 2003. Expression profiling of FSHD muscle supports a defect in specific stages of myogenic differentiation. *Human molecular genetics* 12:2895-2907.
- Witcher M, Emerson BM. 2009. Epigenetic silencing of the p16(INK4a) tumor suppressor is associated with loss of CTCF binding and a chromatin boundary. *Molecular cell* 34:271-284.
- Wohlgemuth M, van der Kooi EL, van Kesteren RG, van der Maarel SM, Padberg GW. 2004. Ventilatory support in facioscapulohumeral muscular dystrophy. *Neurology* 63:176-178.

- Worman HJ, Fong LG, Muchir A, Young SG. 2009. Laminopathies and the long strange trip from basic cell biology to therapy. *The Journal of clinical investigation* 119:1825-1836.
- Wright WE, Piatyszek MA, Rainey WE, Byrd W, Shay JW. 1996. Telomerase activity in human germline and embryonic tissues and cells. *Developmental genetics* 18:173-179.
- Xu X, Tsumagari K, Sowden J, Tawil R, Boyle AP, Song L, Furey TS, Crawford GE, Ehrlich M. 2009. DNaseI hypersensitivity at gene-poor, FSH dystrophy-linked 4q35.2. *Nucleic acids research* 37:7381-7393.
- Xu Y, Suzuki Y, Ito K, Komiyama M. 2010. Telomeric repeat-containing RNA structure in living cells. *Proceedings of the National Academy of Sciences of the United States of America* 107:14579-14584.
- Yamaguchi H, Condeelis J. 2007. Regulation of the actin cytoskeleton in cancer cell migration and invasion. *Biochimica et biophysica acta* 1773:642-652.
- Yang F, Shao C, Vedanarayanan V, Ehrlich M. 2004. Cytogenetic and immuno-FISH analysis of the 4q subtelomeric region, which is associated with facioscapulohumeral muscular dystrophy. *Chromosoma* 112:350-359.
- Ye JZ, de Lange T. 2004. TIN2 is a tankyrase 1 PARP modulator in the TRF1 telomere length control complex. *Nature genetics* 36:618-623.
- Ye JZ, Hockemeyer D, Krutchinsky AN, Loayza D, Hooper SM, Chait BT, de Lange T. 2004. POT1-interacting protein PIP1: a telomere length regulator that recruits POT1 to the TIN2/TRF1 complex. *Genes & development* 18:1649-1654.
- Yehezkel S, Segev Y, Viegas-Pequignot E, Skorecki K, Selig S. 2008. Hypomethylation of subtelomeric regions in ICF syndrome is associated with abnormally short telomeres and enhanced transcription from telomeric regions. *Human molecular genetics* 17:2776-2789.
- Yusufzai TM, Tagami H, Nakatani Y, Felsenfeld G. 2004. CTCF tethers an insulator to subnuclear sites, suggesting shared insulator mechanisms across species. *Molecular cell* 13:291-298.
- Zalensky AO, Allen MJ, Kobayashi A, Zalenskaya IA, Balhorn R, Bradbury EM. 1995. Well-defined genome architecture in the human sperm nucleus. *Chromosoma* 103:577-590.
- Zatz M, Marie SK, Cerqueira A, Vainzof M, Pavanello RC, Passos-Bueno MR. 1998. The facioscapulohumeral muscular dystrophy (FSHD1) gene affects males more severely and more frequently than females. *American journal of medical genetics* 77:155-161.
- Zatz M, Marie SK, Passos-Bueno MR, Vainzof M, Campiotto S, Cerqueira A, Wijmenga C, Padberg G, Frants R. 1995. High proportion of new mutations and possible anticipation in Brazilian facioscapulohumeral muscular dystrophy families. *American journal of human genetics* 56:99-105.
- Zeng W, de Greef JC, Chen YY, Chien R, Kong X, Gregson HC, Winokur ST, Pyle A, Robertson KD, Schmiesing JA, Kimonis VE, Balog J, Frants RR, Ball AR, Jr., Lock LF, Donovan PJ, van der Maarel SM, Yokomori K. 2009. Specific loss of histone H3 lysine 9 trimethylation and HP1gamma/cohesin binding at D4Z4 repeats is associated with facioscapulohumeral dystrophy (FSHD). *PLoS genetics* 5:e1000559.
- Zhang LF, Ogawa Y, Ahn JY, Namekawa SH, Silva SS, Lee JT. 2009. Telomeric RNAs mark sex chromosomes in stem cells. *Genetics* 182:685-698.
- Zhao J, Sun BK, Erwin JA, Song JJ, Lee JT. 2008. Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome. *Science (New York, NY)* 322:750-756.
- Zhu XD, Kuster B, Mann M, Petrini JH, de Lange T. 2000. Cell-cycle-regulated association of RAD50/MRE11/NBS1 with TRF2 and human telomeres. *Nature genetics* 25:347-352.

- Zhu XD, Niedernhofer L, Kuster B, Mann M, Hoeijmakers JH, de Lange T. 2003. ERCC1/XPF removes the 3' overhang from uncapped telomeres and represses formation of telomeric DNA-containing double minute chromosomes. *Molecular cell* 12:1489-1498.
- Zink D, Amaral MD, Englmann A, Lang S, Clarke LA, Rudolph C, Alt F, Luther K, Braz C, Sadoni N, Rosenecker J, Schindelbauer D. 2004. Transcription-dependent spatial arrangements of CFTR and adjacent genes in human cell nuclei. *The Journal of cell biology* 166:815-825.
- Zou YS, Van Dyke DL, Thorland EC, Chhabra HS, Michels VV, Keefe JG, Lega MA, Feely MA, Uphoff TS, Jalal SM. 2006. Mosaic ring 20 with no detectable deletion by FISH analysis: Characteristic seizure disorder and literature review. *Am J Med Genet A* 140:1696-1706.

Sites internet :

OMIM : <http://www.ncbi.nlm.nih.gov/omim/>

Orphanet : <http://www.orpha.net/consor/cgi-lin/index.php>

UCSC : <http://genome.ucsc.edu/>