



**HAL**  
open science

# Caractérisation génétique de la race de mouton Awassi du Liban en utilisant comme marqueurs des rétrovirus endogènes et l'ADN mitochondrial

Jeanne El Hage

► **To cite this version:**

Jeanne El Hage. Caractérisation génétique de la race de mouton Awassi du Liban en utilisant comme marqueurs des rétrovirus endogènes et l'ADN mitochondrial. Génétique animale. Université Paris sciences et lettres, 2017. Français. NNT : 2017PSLEP068 . tel-02107066

**HAL Id: tel-02107066**

**<https://theses.hal.science/tel-02107066>**

Submitted on 23 Apr 2019

**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.

# THÈSE DE DOCTORAT

de l'Université de recherche Paris Sciences et Lettres  
PSL Research University

Préparée à l'École Pratique des Hautes Études

Caractérisation génétique de la race de mouton Awassi du Liban en  
utilisant comme marqueurs des rétrovirus endogènes et l'ADN  
mitochondrial

École doctorale de l'EPHE – ED 472

Spécialité : Biodiversité, génétique et évolution

Soutenue par :

**Jeanne EL HAGE**

le 19 décembre 2017

Dirigée par :

**Frédéric ARNAUD**

**Alain ABI RIZK**

COMPOSITION DU JURY :

M. Michel AFRAM  
Institut de Recherche Agronomique Libanais  
Président du jury

M. François POMPANON  
Université Grenoble Alpes  
Rapporteur

M. Jean-Nicolas VOLFF  
Ecole Normale Supérieure de Lyon  
Rapporteur

M. Frédéric BOYER  
Université Grenoble Alpes  
Examineur

M. Frédéric ARNAUD  
UMR754 UCBL/INRA/EPHE  
Directeur de thèse

M. Alain ABI RIZK  
Université Saint-Esprit de Kaslik  
Codirecteur de thèse



École Pratique  
des Hautes Études





A mes parents, et à ma famille.

## REMERCIEMENTS

Les mots me manquent en évoquant la personne qui m'a présentée l'opportunité de faire cette thèse : Professeur Christophe TERZIAN, merci pour cette chance qui a changé le parcours de ma vie et m'a donné le goût de la génétique des populations, une thématique qui m'était tout à fait nouvelle. Je ne vous oublierai jamais.

Un grand merci témoignant toute ma reconnaissance est adressé à mon directeur, Dr Frederick ARNAUD, pour l'expérience enrichissante et pleine d'intérêt qu'il m'a fait vivre durant ma thèse et pour le temps qu'il m'a consacré, pour sa disponibilité et son enthousiasme : Merci pour toutes vos initiatives personnelles pour régler mes problèmes administratifs et m'aider à continuer ma thèse au moment où moi-même j'ai renoncé d'y croire. Vous m'avez dit que « vous le devez à Christophe » et maintenant je vous dis : cette thèse je vous la dois.

Je remercie également Dr Alain Abi Rizk, mon co-directeur de thèse. Merci pour vos conseils et votre aide. Merci pour votre disponibilité et écoute. Vous avez supporté mes plaintes avec une patience rare. Merci pour votre optimisme qui me remontait le moral.

J'adresse mes vifs remerciements au président et aux membres du jury : Pr Michel AFRAM, Pr François POMPANON, Pr Jean-Nicolas VOLFF et Dr Frédéric BOYER d'avoir accepté de juger ce travail.

Un grand merci pour mon Directeur Général, Pr Michel Afram, qui ne cesse de ménager aucun effort pour soutenir toute initiative dans le but d'améliorer la qualité et le niveau de travail à l'IRAL. Il a offert toutes les facilités afin que ce projet arrive à ses fins. Pour lui, j'exprime toute ma gratitude et ma reconnaissance pour son support moral et financier.

Je tiens à remercier toutes les personnes qui par leur implication dans ce projet, m'ont été d'un grand secours, tout au long de ces trois années et sans leur collaboration ce travail n'aurait pas abouti :

Tous les membres de l'UMR754 UCBL/INRA/EPHE, surtout Dr Fabienne ARCHER, directrice de l'unité, pour son soutien. Je remercie bien spécialement l'équipe de « Biologie des phlébovirus » : Barbara VIGINIER, Maxime RATINIER, Saw-See HONG, Sophie DESLOIRE, Marie-Pierre CONFORT pour leurs supports et leurs aides pendant mes séjours à l'unité. Merci aussi à Dr Caroline LEROUX, Barbara GINEYS et Christine DOLMAZON qui m'ont aidé dans les manipes de JSRV.

Mes collègues à l'Institut de Recherche Agronomique Libanais : Merci à Wassim, Elvis et Carine, Chantal, Ziad, Nada, Raed et Joyce pour votre amitié et votre support. Merci pour

votre présence à mes côtés et pour votre patience et écoute. Vous m'avez grandement aidé à avancer, ne pas relâcher et repartir dans la bonne humeur.

Merci à ma famille qui m'a écouté me plaindre et râler.... Et bien sûr, merci à ma sœur et collègue, Rima : elle était à côté de moi dans mes moments les plus bas ....Merci et désolée, je sais que j'étais « un peu » insupportable ces trois dernières années ; Je vous aime !!!

Merci !

## RÉSUMÉ

La domestication des bétails représente une étape importante dans l'histoire de l'humanité. Le mouton était l'un des premiers animaux à être domestiqués dans le croissant fertile. Ces événements de domestication, probablement initiés au début du Néolithique, ont génétiquement construit les races contemporaines du Moyen-Orient mais aussi du monde entier. L'élevage de moutons, principalement mouton de la race Awassi, représente une activité économique essentielle du Liban ; cependant, jusqu'à présent, il n'existe que très peu de données génétiques sur cette race. De nos jours, les outils moléculaires disponibles nous permettent de définir en détail la diversité génétique des populations de moutons et de retracer leur histoire évolutive. Par conséquent, l'objectif principal de mon projet de thèse était de caractériser génétiquement la race Awassi du Liban. Pour cette étude, 277 échantillons d'ADN génomique prélevés des moutons Awassi du Liban (n = 254) et de la Syrie (n = 23) ont été analysés. Au début, nous avons utilisé cinq rétrovirus endogènes (rétrovirus endogène de moutons de Jaagsiekte-enJSRV) qui sont polymorphiques par insertion dans les génomes du mouton domestique (enJSRV-18, -7, -15, -16 et -22) et ont été précédemment considérés comme très informatifs principalement pour distinguer génétiquement les moutons primitifs des races plus modernes (c.-à-d. le dernier issu de l'épisode migratoire impliquant des moutons avec des traits de production améliorés). En utilisant cette approche, nos résultats montrent une prédominance du type R2 (enJSRV-18 seulement) confirmant que le mouton Awassi du Liban est une race moderne. Comme prévu, le rétrotype R4 (à la fois enJSRV-18 et enJSRV-7), une caractéristique commune des populations de moutons du bassin méditerranéen, se trouve également dans le génome des moutons d'Awassi du Liban et plus accentué dans les troupeaux Syriens. Il est intéressant de noter que les populations de moutons d'Awassi situés dans le nord-est du Liban et ayant ainsi un accès plus restreint à la mer Méditerranée que les autres populations (c'est-à-dire en raison de la chaîne de montagne centrale qui coupe le pays sur deux), présentent une faible fréquence de R4. Bien que l'origine des animaux utilisés pour établir les troupeaux analysés au cours de cette étude soit inconnue, nos résultats fournissent également certaines preuves que le mode d'élevage (ouvert ou fermé) peut influencer les rérotypes observés et en particulier le R4. De manière surprenante, au cours de cette étude, nous avons également dévoilé la présence de soi-disant "Solo-LTR" (c'est-à-dire généré par une recombinaison homologue) pour un autre enJSRV (enJSRV-6) qui prédomine dans deux troupeaux d'une région particulière du Liban (Nabatieh). Et comme approche complémentaire, deux marqueurs mitochondriaux ont été utilisés, le cytochrome b

(*Cyt-b*) et D-Loop, pour étudier l'origine maternelle de cette race et sa relation phylogénétique au sein de la famille *Ovis aries*. Dans notre étude, le *Cyt-b* se révèle plus discriminant que le D-Loop. Des mouton d'Awassi analysé, quatre haplogroupes (HPG) du Moyen-Orient ont été trouvés avec l'analyse du *Cyt-b* : HPG A, B, C et E, ce dernier étant peu fréquent. De même, l'analyse de la super-séquence, alignement *Cyt-b*\_D-Loop, a permis l'identification de l'HPG D, un HPG extrêmement rare et limité jusqu'à présent aux moutons à queue grasse tel que l'Awassi. Enfin, une expansion passée de la population est observée pour les HPG A, B et C (mais pas pour HPG E) avec les distributions incompatibles et des tests de neutralité négatifs significatifs. Dans l'ensemble, les résultats obtenus au cours de cette étude fournissent une caractérisation génétique complète ainsi que quelques idées sur la structure phylogéographique des populations de moutons de la race Awassi au Liban.

Mots-clés : Domestication, Mouton, rétrovirus endogènes, ADN mitochondrial.



## ABSTRACT

Livestock domestication represents a milestone in the history of mankind. Sheep was one of the first animals to be domesticated in the Fertile Crescent. These domestication events, probably initiated in the early Neolithic, have genetically built the contemporary races of the Middle East but also of the whole world. Sheep farming, mainly sheep of Awassi breed, represents an essential economic activity of Lebanon; however, so far, only very few genetic data exist on this breed. Nowadays, the molecular tools available allow us to define in details the genetic diversity of sheep populations and to trace their evolutionary history. Hence, the main objective of my PhD project was to genetically characterize the Awassi breed of Lebanon. For this study, 277 genomic DNA samples collected from Awassi sheep of Lebanon (n=254) and Syria (n=23) were analyzed. Initially, we used five endogenous retroviruses (endogenous Jaagsiekte sheep retrovirus-enJSRV) that are insertionally polymorphic within the genomes of domestic sheep (enJSRV-18, -7, -15, -16 and -22) and have been previously shown to be very informative mainly to genetically distinguish between primitive sheep from more modern breeds (i.e. the latter originating from the migratory episode involving sheep with improved production traits). Using this approach, our results show a predominance of the R2 retrotype (enJSRV-18 only) confirming that the Awassi sheep of Lebanon is a modern breed. As expected, the R4 retrotype (both enJSRV-18 and enJSRV-7), a common feature of the sheep populations present within the Mediterranean area, is also found in the Awassi sheep of Lebanon and to more extend in those of Syria. Interesting, the populations of Awassi sheep located in the northeast of Lebanon and thus having a more restricted access to the Mediterranean Sea than the other populations (i.e. due to the central mountain chain cutting the country in two) present R4 weaklier. Even though the origin of the animals used to establish the herds analyzed during this study is unknown, our results also provide some evidences that the mode of rearing (open or closed) may influence the observed retrotypes and in particular R4. Surprisingly, during this study, we also unveiled the presence of so-called “Solo-LTR” (i.e. generated by homologous recombination) for another enJSRV (enJSRV-6) that are predominant in two herds of a particular region of Lebanon (Nabatieh). As a complementary approach, two mitochondrial markers were used, the cytochrome b (*Cyt-b*) and D-Loop, to investigate the maternal origin of this breed and its phylogenetic relationship within the *Ovis aries* family. In our study, the *Cyt-b* turns out to be more discriminative than the D-Loop. From the Awassi sheep analyzed, four haplogroups (HPGs) of the Middle-East were found with *Cyt-b* analysis: HPG A, B, C and E, the latter being the least frequent. Also,

the super-sequence analysis, Cyt-b\_D-Loop alignment, allowed the identification of HPG D, an extremely rare HPG, limited till now to fat-tailed sheep such as Awassi. Finally, a past population expansion is observed for the HPG A, B and C (but not for HPG E) with mismatch distributions and significant negative neutrality tests. Overall, the results obtained during this study provide a comprehensive genetic characterization as well as some insights into the phylogeographic structure of the sheep populations of the Awassi breed in Lebanon.

**Keywords :** Domestication, sheep, endogenous retroviruses, mitochondrial DNA.

Ce travail a été réalisé à l'Institut de Recherche Agronomique libanais- IRAL, Département de Santé animal, à l'UMR754 UCBL/INRA/EPHE : Infections Virales et Pathologie Comparée (IVPC) et à l'Université Saint-Esprit de Kaslik –USEK.



Institut de Recherche Agronomique libanais – IRAL  
Département de Santé animale  
Fonar – El Metn- Liban



UMR754 UCBL/INRA/EPHE : Infections Virales et  
Pathologie Comparée (IVPC)  
50 Avenue Tony Garnier, 69366 Lyon CEDEX 07



Université Saint-Esprit de Kaslik -USEK  
B.P. 446 - Jounieh - Liban



## TABLE DES MATIERES

REMERCIEMENTS .....	iv
RÉSUMÉ.....	vi
ABSTRACT .....	viii
LISTE DES FIGURES .....	5
LISTE DES TABLEAUX .....	8
LISTE DES ABREVIATIONS .....	9
<b>Chapitre 1 : La Domestication des Moutons .....</b>	<b>12</b>
1.1. Introduction .....	13
1.2. La domestication du mouton .....	14
1.2.1. Domestication et migration successives .....	14
1.2.2. Sélection et changement morphologiques au cours de la domestication du mouton .....	16
1.2.3. Classification du genre Ovis.....	18
1.3. La race AWASSI.....	20
1.3.1. Origine et Évolution .....	20
1.3.2. L’agriculture et la race Awassi au Liban.....	21
1.4. Problématique.....	25
<b>Chapitre 2 : Les Rétrovirus .....</b>	<b>26</b>
2.1. Les Rétrovirus.....	27
2.1.1. Historique .....	27
2.1.2. Taxonomie & Classification.....	28
2.1.3. Les rétrovirus endogènes .....	30
2.1.4. Organisation génomique des rétrovirus.....	32
2.2. Cycle de réplication des rétrovirus .....	36
2.2.1. Les phases précoces.....	37
2.2.2. Les phases tardives .....	40
2.3. L'adénocarcinome pulmonaire ovin et le Jaagsiekte Sheep Rétrovirus.....	42
2.3.1. Historique .....	42
2.3.2. Adénocarcinome pulmonaire.....	42
2.3.3. Le rétrovirus exogène : Jaagsiekte Sheep Retrovirus.....	45
2.4. Les rétrovirus endogènes de type JSRV .....	48
2.4.1. Interactions des enJSRVs avec le virus exogène JSRV et leur hôte .....	48
2.4.2. Coadaptation entre enJSRVs/JSRV et le mouton au cours de l'évolution .....	50

2.4.3. L’histoire de la domestication du mouton révélée par les rétrovirus endogènes enJSRVs. ....	53
<b>Chapitre 3 : Les ADNs Mitochondriaux .....</b>	<b>57</b>
3.1. Introduction .....	58
3.2. Les mitochondries.....	58
3.2.1. Origine des mitochondries.....	59
3.2.2. Structure et fonction des mitochondries .....	60
3.2.3. L’ADN mitochondrial .....	61
3.2.4. Héritage de l’ADN mitochondrial .....	64
3.2.5. Homoplasmie et Hétéroplasmie .....	65
3.2.6. Mutations de l’ADNmt.....	66
3.3. L’ADN mitochondrial comme marqueur génétique.....	67
3.4. L’ADNmt et la diversité génétique des ovins .....	68
3.4.1. Cytochrome B.....	68
3.4.2. D-Loop .....	73
3.5. L’ADN mitochondrial et les haplogroupes .....	78
<b>Chapitre 4 : Matériel et Méthodes .....</b>	<b>80</b>
4.1. Echantillons : Nature et origine .....	81
4.2. Extraction et dosage de l’ADN génomique .....	82
4.3. Amplification des rétrovirus endogènes .....	82
4.4. Amplification des gènes mitochondriaux .....	87
4.4.1. Le cytochrome b .....	87
4.4.2. La D-Loop .....	88
4.5. Purification et séquençage .....	89
4.6. Analyses statistiques.....	89
<b>Chapitre 5 : Résultats .....</b>	<b>92</b>
5.1. Analyse génétique des populations de mouton de la race Awassi au Liban via les rétrovirus enJSRVs. ....	93
5.2. Analyse génétique des populations de mouton de la race Awassi au Liban avec l’ADN mitochondrial.....	104
5.2.1. Le cytochrome-b : Cyt-b .....	104
5.2.2. D-Loop .....	111

Chapitre 6 : Discussion .....	118
6.1. Les rétrovirus endogènes enJSRVs sont-ils des bons marqueurs génétiques pour l'analyse de populations d'une même race et d'une région géographique donnée ?.....	119
6.2. L'invité surprise, le solo-LTR !.....	121
6.3. Des enJSRVs oui mais où est le JSRV ? .....	122
6.4. L'ADN mitochondrial et son pouvoir discriminant .....	123
6.5. Le Liban est-il un centre de domestication du mouton ?.....	124
ANNEXE 1 : Analyse en composantes principales sur les insertions des enJSRVs.....	126
ANNEXE 2 : Alignement des séquences des haplotypes obtenus à partir du Cyt-b .....	127
ANNEXE 3 : Alignement des séquences des haplotypes obtenus à partir de la D-Loop .	128
RÉFÉRENCES BIBLIOGRAPHIQUES .....	129
Publications .....	146

## LISTE DES FIGURES

Figure 1-1 : Voies de domestication. ....	14
Figure 1-2: A) Chronologie d'évènements de domestication du mouton. B) Voies de migration vers l'Asie, l'Afrique et l'Europe.. ....	16
Figure 1-3: A. Moutons à queue grasse sur un bol en pierre de la période d'Uruk d'Ur III (3000 Av. J.-C.). B. Moutons à queue grasse. Gravures rupestres de l'Arabie centrale, deuxième ou au début du premier millénaire avant JC.. ....	17
Figure 1-4: La Taxonomie du genre Ovis selon Geist. ....	19
Figure 1-5: La répartition mondiale des sept races de mouton sauvage. ....	20
Figure 1-6: les systèmes d'élevage de bétail au Liban. ....	23
Figure 1-7: Un bélier avec des brebis et un agneau de la race Awassi à Sanine - Mont Liban	24
Figure 2-1 : Morphologie des rétrovirus de type A-D observée par microscope électronique	29
Figure 2-2: Structure des deux principales classes d'éléments transposables.. ....	31
Figure 2-3 : Classification des ERVs selon leur degré de parenté avec un rétrovirus exogène Dendrogramme neighbor joining (NJ) des sept genres de rétrovirus .....	32
Figure 2-4: Structure génomique des Retroviridae. Les rétrovirus peuvent être de structure simple ou complexe.....	33
Figure 2-5: Représentation schématique d'un provirus et des transcrits correspondants. ....	34
Figure 2-6: Les étapes précoces et tardives du cycle rétroviral. ....	37
Figure 2-7: Processus de transcription inverse des rétrovirus.....	39
Figure 2-8: Schéma résumant les principales étapes de l'intégration de l'ADN rétroviral dans le génome .....	40
Figure 2-9: Structure d'une particule virale mature, enveloppée et infectieuse.....	41
Figure 2-10: Schéma représentant les types cellulaires présents au niveau des bronchioles et des alvéoles. ....	44
Figure 2-11:Organisation génomique de l'ARN viral de JSRV .....	45
Figure 2-12: Modèle hypothétique de l'interaction entre l'Env de JSRV avec son récepteur la hyaluronidase 2 (Hyal2) à la membrane plasmique.. ....	46
Figure 2-13: Voies de signalisation impliquées dans la transformation par l'Env du JSRV ...	47
Figure 2-14: Organisation génomique des provirus enJSRV.....	50
Figure 2-15 : Histoire évolutive des enJSRVs et leur hôte. ....	52
Figure 2-16 La domestication et les vagues successives de migration des moutons selon les rétrovirus endogènes .....	54



Figure 3-1: Morphologies adoptées par les mitochondries dans les différents tissus vues par microscopie électronique.....	60
Figure 3-2 : Structure d'une mitochondrie .....	61
Figure 3-3: Organisation de l'ADNmt.....	62
Figure 3-4: A) Chaîne respiratoire/phosphorylation oxydative. B) Structure du complexe Complexe III (cytochrome bc1). .....	70
Figure 3-5: Arbres phylogénétiques à partir des séquences du cytochrome b. ....	72
Figure 3-6: Présentation schématique de la structure de la D-Loop .....	73
Figure 3-7: Structure de la D-Loop chez les ovins.....	74
Figure 3-8: Modes de répllication de l'ADNmt. ....	77
Figure 4-1: Distribution des échantillons sur le territoire Libanais .....	82
Figure 4-2: Représentation schématique des PCRs avec les amorces utilisées. ....	83
Figure 4-3: Le cycle thermique de l'amplification des rétrovirus endogènes.....	84
Figure 5-1: A) Schémas représentant les trois types de PCR . B) Gels représentatifs des résultats des PCR obtenus pour chaque enJSRV testé sur des échantillons provenant de différentes régions du Liban.....	94
Figure 5-2: Cartes de densité représentant la fréquence et la répartition des enJSRVs polymorphiques dans les troupeaux de mouton de la race Awassi au Liban. ....	95
Figure 5-3 : Alignement de la séquence Solo-LTR d'enJSRV-6.....	96
Figure 5-4:A) Tableau récapitulatif des résultats des PCR d'enJSRV-6 en termes de zygote. B) Cartes de densités des trois cas de figure pour enJSRV-6 . ....	97
Figure 5-5: Fréquence et distribution géographique des rétrotypes R2 et R4.....	98
Figure 5-6: Fréquence et distribution géographique des rétrotypes R2 et RA. ....	99
Figure 5-7: Fréquence et distribution géographique des rétrotypes R2 et RB. ....	100
Figure 5-8: Fréquence et distribution géographique des rétrotypes R2 et R6.....	101
Figure 5-9: Fréquence et distribution géographique de l'ensemble des rétrotypes.....	102
Figure 5-10: Fréquence et distribution géographique des différents types de zygote retrouvés pour enJSRV-6.....	103
Figure 5-11: Arbre phylogénétique des séquences de <i>Cyt-b</i> .....	105
Figure 5-12. Arbre phylogénétique NJ des haplotypes basés sur les séquences du <i>Cyt-b</i> . ....	107
Figure 5-13: La « mismatch distributions », pour chaque HPG identifié au sein des échantillons de mouton d'Awassi libanais. ....	109
Figure 5-14 : Fréquence et répartition géographique des HPGs selon le <i>Cyt-b</i> .....	111
Figure 5-15: Arbre phylogénétique selon NJ des séquences de la D-Loop .....	112

Figure 5-16 : Distribution incompatible des HPGs A et B des séquences D-Lopp . . . . .	113
Figure 5-17 : Fréquence et répartition géographique des HPGs selon la D-Loop . . . . .	114
Figure 5-18 : Cinq lignées maternelles illustrées avec l'arbre Neighbor-Joining . . . . .	115

## LISTE DES TABLEAUX

Tableau 1-1: Distribution mondiale des races ovines domestiques ( <i>O. aries</i> ) .....	20
Tableau 2-1: classification au sein de la famille des Retroviridae selon l'ICTV .....	29
Tableau 4-1: Nombre et nature des échantillons .....	81
Tableau 4-2 : Le PCR mixe pour les rétrovirus endogènes.....	84
Tableau 4-3 : Amorces des rétrovirus endogènes. ....	85
Tableau 4-4: Le PCR mixe pour le cytochrome b.....	87
Tableau 4-5: Le PCR mixe de la D-Loop .....	88
Tableau 5-1: Variations phylogénétiques des haplogroupes mitochondriaux des moutons de la race Awassi au Liban .....	106
Tableau 5-2 : Les nombres moyens de différence nucléotidique ( $K_{xy}$ ) et de substitutions nucléotidiques par site ( $D_{xy}$ ) entre les populations des différents haplogroupes .....	108
Tableau 5-3 : Valeurs statistiques des tests de neutralité et du nombre moyen de la différence nucléotidique $K$ pour chaque haplogroupe. ....	110
Tableau 5-4: Variations phylogénétiques des haplogroupes mitochondriaux (D-Loop) de la race Awassi au Liban .....	113
Tableau 5-5 : Valeurs statistiques des tests de neutralité et du nombre moyen de la différence nucléotidique $K$ pour chaque haplogroupe.....	114
Tableau 5-6 : Les nombres moyens de différence nucléotidique ( $K_{xy}$ ) et de substitutions nucléotidiques par site ( $D_{xy}$ ) entre les populations des différents haplogroupes des moutons Awassi libanais avec la séquence « super gène » : Cytb-DLoop. ....	116
Tableau 5-7 : les variations phylogéniques des séquences supergènes.....	117
Tableau 6-1 : Diversités génétiques observées dans les trois types de séquences mitochondriales analysées.....	124

## LISTE DES ABREVIATIONS

2DAGE	Electrophorèse en gel d'agarose bidimensionnelle
A	Adénine
AD	Anno Domini
ADNc	ADN complémentaire
ADNmt	l'ADN mitochondrial
ADNnu	ADN nucléaire
ADNsss	Minus-strand strong-stop DNA
Av. J.-C.	Avant Jésus-Christ
ALV	Virus de leucose aviaire / Avian leucosis virus
ARNr	ARN ribosomique
ARNt	ARN de transfert
ASLV	Virus aviaire du groupe sarcome-leucose
ATP	Adénosine triphosphate
BASC	Cellules souches broncho-alvéolaires
BLV	Bovine leukaemia virus
BP	Before Present
Brin H	Heavy strand
Brin L	Light strand
C	Cytosine
CA	Capside
CR	Control Region
CSB	Conserved Sequence Block
CT	Queue cytoplasmique
<i>cyt-b</i>	Cytochrome B
DD <sub>35E</sub>	Aspartate (D), aspartate (D), glutamate (E) catalytic triad
D-Loop	Displacement Loop
dUTPase	Déoxyuridine triphosphatase
Dxy	Average number of nucleotide substitutions per site between populations
enJSRV	Endogenous Jaagsiekte sheep retrovirus
env	enveloppe
ERV	rétrovirus endogènes
ET	éléments transposables
FADH2	Flavine adénine dinucleotide
FAO	Food and Agriculture Organization
G	Guanine
<i>gag</i>	<i>group specific antigen</i>
GPI	Glycophosphatidylinositol
Hd	Diversité haplotypique
HFV	Foamy virus humain / Human foamy virus
HH-CC	Histidine-histidine-Cystidine-cystidine
HIV-1	Human immunodeficiency virus-1
HNF3 $\beta$	Hepatocyte Nuclear Factor 3 $\beta$
HPG	Haplogroupe

HSP	Heavy strand promoter
HVR	Hyper Variable regions
Hyal2	Hyaluronidase 2
ICTV	Comité international sur la taxonomie des virus
IN	Intégrase
PBS	Partitioned Bremer Support
JSRV	Jaagsiekte sheep retrovirus
LSP	Light Strand promoter
LTR	Long Terminal Repeat
MA	Matrice
MLV	Murine leukaemia virus
MMTV	Mouse Mammary Tumour Virus
M-PMV	Virus du singe Mason-Pfizer
MPR	Mitochondrial RNA-Processing enzyme
MSSS	Minus-strand-strong stop DNA
MVV	Maedi-Visna Virus
NC	Nucléocapside
NF-1	Nuclear Factor 1
NHP	Nonhuman primates
NUMTS	Nuclear mitochondrial DNA segment
OH	Origine de replication du brin lourd
OIE	Organisation mondiale de la santé animale
OL	Origine de replication du brin léger
OPA	Adénocarcinome pulmonaire ovin
ORFs	Cadres ouverts de lectures
OXPHOS	Phosphorylation oxydative
PBS	Primer Binding site
PCR	Polymerase Chain Reaction
PI3K	Phosphatidylinositol 3-kinase
PIB	Produit intérieur brut
PIC	PreIntegration Complex
pol	polymerase
PPT	Poly Purine Tract
PR	Protéase
Pro	Protease
PSSS	Plus-strand strong-stop DNA
R	Arginine (one-letter amino acid code)
RH	RNase H
RSV	Virus du sarcome de Rous
RT	Reverse transcriptase
s/s/yr	substitution/site/year
SP-1	Specific Protein 1
SP-A	Protéine A du surfactant
SP-C	Protéine C du surfactant
SU	Protéine de surface

T	Thymine
TAS	Termination-associated Sequence
TAS-BP	TAS Binding protein
TBE	Tris-borate EDTA
TLR 4	Toll-Like receptor 4
TM	sous-unité dite transmembranaire
UTR	UnTranslated Region)
VR	Régions variables
w/v	weight / volume
YXXM	Y est une tyrosine, X un acide aminé, M est une méthionine
$\pi$	Diversité nucléotidique

# Chapitre 1 : *La Domestication des Moutons*

---

## 1.1. Introduction

Le passage de la cueillette et de la chasse à l'agriculture et l'élevage est un seuil important dans l'histoire de l'humanité. La domestication et les économies agricoles basées sur cette transition, sont associées à une restructuration radicale des sociétés humaines, à des altérations mondiales de la biodiversité et à des changements importants dans la topographie de la Terre et son atmosphère [1]. Toutes les définitions de la domestication reconnaissent que la domestication implique une relation entre les hommes d'une part et la population d'animaux visée d'autre part [2]. En 2012, Zeder décrit trois scénarios pour la domestication soit trois voies majeures que les animaux ont suivies pour être domestiqués :

a) La voie commensale :

Cette voie est celle des animaux qui se nourrissent de déchets autour des habitats humains ou des animaux qui s'en prennent à d'autres animaux attirés par des environnements anthropiques. La domestication de ces animaux commence en établissant une relation commensale avec l'homme - une relation dans laquelle un partenaire bénéficie et l'autre ne récolte que peu ou pas de bienfaits. Mais au bout d'un certain temps, cette association crée un lien social ou économique entre les deux parties aboutissant à la domestication. (Figure 1-1A). Comme exemple, on peut citer le chien, le chat, les cochons et les oiseaux dont les poules.

b) La voie de proie :

C'est la voie de domestication du bétail. Ces animaux étaient les principales espèces chassées par l'homme pour leur viande. Pour rendre l'approvisionnement en proie et par suite en nourriture disponible, l'homme a initié cette domestication. Donc il est passé de la chasse à la gestion de troupeaux (Figure 1-1B). C'est le cas des bovins, des moutons ou bien encore des chèvres.

c) La voie dirigée :

Cette dernière voie de domestication est un processus plus intentionnel et dirigé. Il s'agit d'un processus initié aussi par l'homme dans un but bien précis tel que le transport (Figure 1-1C). C'est, par exemple, la voie de domestication suivie pour le Cheval, l'âne, et le chameau.



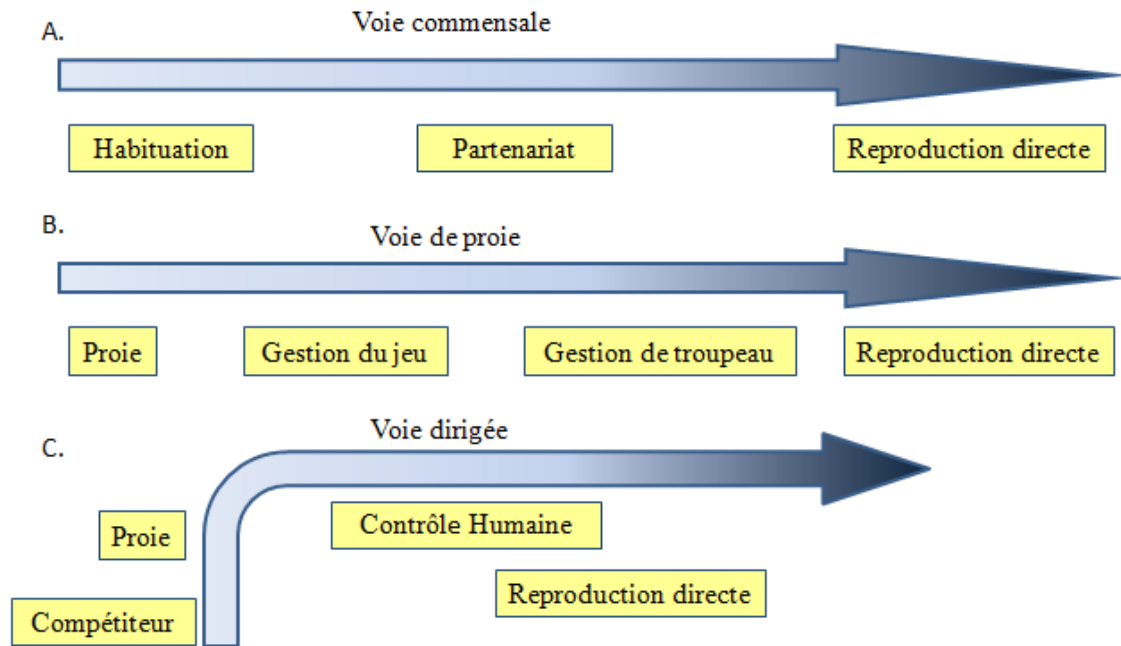


Figure 1-1 : Voies de domestication. A. Voie commensale, B. Voie de proie, C. Voie dirigée. Voir le texte pour plus de détails [3]

## 1.2. La domestication du mouton

La domestication des animaux fut une étape clef de l'humanité et dans l'établissement des fondements de notre société actuelle. Le processus de domestication se situe en Orient entre l'apparition de l'agriculture et de la céramique à l'époque néolithique entre -12000 et -6000 ans avant Jésus-Christ (Figure 1-2A). A cette période, certains groupes d'individus sont donc passés du statut de chasseurs-cueilleurs à celui d'agriculteurs-éleveurs [1, 4]. Le mouton fut l'une des premières espèces domestiques après le chien et la chèvre et à peu près au même moment que le porc [5, 6]. Plusieurs hypothèses s'affrontent en ce qui concerne le lieu de domestication du mouton ainsi que les migrations qui ont suivies vers l'Europe, l'Afrique et l'Asie (Figure 1-2B).

### 1.2.1. Domestication et migration successives

D'après les données archéologiques, la domestication du mouton semble s'être déroulée au sein d'un foyer unique dans une vaste région du Moyen Orient correspondant au croissant fertile de l'ancienne Mésopotamie (Iran, Irak et Afghanistan de nos jours) [7] mais à partir d'évènements multiples [8]. Le premier site archéologique datant de  $10\,870 \pm 300$  ans avant notre ère et mettant en évidence le contrôle humain sur la domestication et l'élevage d'ovins se situe dans la région de « Zawi Chemi Shanidar » dans le Nord-Est de l'Iran. Cette

découverte suggère donc que l'homme a commencé à domestiquer le mouton dans la région de l'Asie occidentale au cours des huitième et septième millénaires Av. J.-C. Dans un premier temps, le mouton fut domestiqué pour sa viande puis au cours de l'histoire pour ces produits dérivés tels que le lait ou la laine [6, 9, 10]. De nombreuses études génétiques tentant de dévoiler l'ancêtre sauvage exact du mouton domestique ont été réalisées. Les espèces sauvages ayant été domestiquées appartiennent au genre *Ovis*. Au sein de ce groupe, le mouflon oriental ou d'Asie (*Ovis orientalis*) vivant actuellement en Turquie, Arménie et Azerbaïdjan semble être l'indiscutable ancêtre du mouton domestique (*Ovis aries*) [11-13]. Le mouflon d'Europe (*Ovis musimon*) semble, quant à lui, être issu des premières migrations de moutons domestiques retournés à la vie sauvage. Il est principalement retrouvé en Corse et en Sardaigne. Ce premier évènement de domestication fut suivi de migrations rapides vers le Nord-Ouest en Europe et vers l'est en Asie et Extrême-Orient et l'Afrique [14]. Suite à la domestication de ces espèces sauvages dites « primitives », un deuxième évènement de domestication et de migration a eu lieu plus tardivement, il y a environ 5000 ans, remplaçant les populations de mouton issues de la première vague sur le continent européen, africain et asiatique (Figure 1-2A). La migration vers l'Europe s'est effectuée en suivant probablement deux voies principales, l'une terrestre à travers la vallée du Danube et l'autre maritime, le long de la côte méditerranéenne (Figure 1-2B) [1, 15]. La première apparition des restes de moutons domestiques dans la partie occidentale de l'Europe méditerranéenne, datant d'environ 5400 av. J.-C., reflète une propagation rapide par la mer [16, 17]. Cependant, la mer Méditerranée a également joué un rôle clé dans l'histoire du bétail à l'époque post-néolithique, lorsque des peuples comme les Phéniciens, les Grecs, les Romains et les Berbères ont probablement introduit de nouvelles espèces d'animaux et de nouvelles races de bétail en Europe. Au fil du temps, ces animaux ont ensuite pu être sélectionnés et améliorés localement la relation génétique étroite entre les races de moutons domestiques *Ovis aries* présentes aujourd'hui en Europe, en Afrique et au Proche-Orient.

D'autre part, une étude phylogénétique récente montre que les moutons d'Asie de l'Est sont génétiquement plus divers que les moutons communs actuellement présents en Europe [18]. Les auteurs de cette étude suggèrent que ce riche héritage est le produit de deux vagues successives de migration des animaux domestiqués entre le Moyen-Orient et l'Asie de l'Est. En effet, l'analyse de 150 races de plusieurs pays confirme que les moutons domestiqués ont d'abord émergé dans le Croissant Fertile, il y a environ 8.000 à 11.000 ans, pour ensuite se diriger à l'Est, lors de la première vague de migration, vers la Chine et la Mongolie, par la Route de la Soie ; un ensemble de routes commerciales qui s'étendent sur près de 6500 km et

qui a facilitée le commerce et les migrations humaines entre l'Asie et l'Europe pendant des milliers d'années. [18]. De même, ils ont pu mettre en évidence que les éleveurs, dans ce qui est maintenant le nord de la Chine et de la Mongolie, ont développé leurs propres races. Lors d'une deuxième vague de migration, il y a environ 5000 ans, ces animaux remontèrent vers l'ouest le long de la Route de la Soie, où le commerce fréquent des femelles reproductrices ou des brebis leur permettait de se mêler à la descendance de leurs ancêtres pour produire des races encore plus distinctes. Cette étude montre que les lignées génétiques de moutons asiatiques modernes ont été façonnées par des milliers d'années de commerce et d'élevage se déplaçant d'abord d'ouest en Est, puis de retour, d'Est en Ouest (Figure 1-2B), ce qui a créé la diversité unique de ces moutons d'aujourd'hui (avis de Olivier Hanotte).

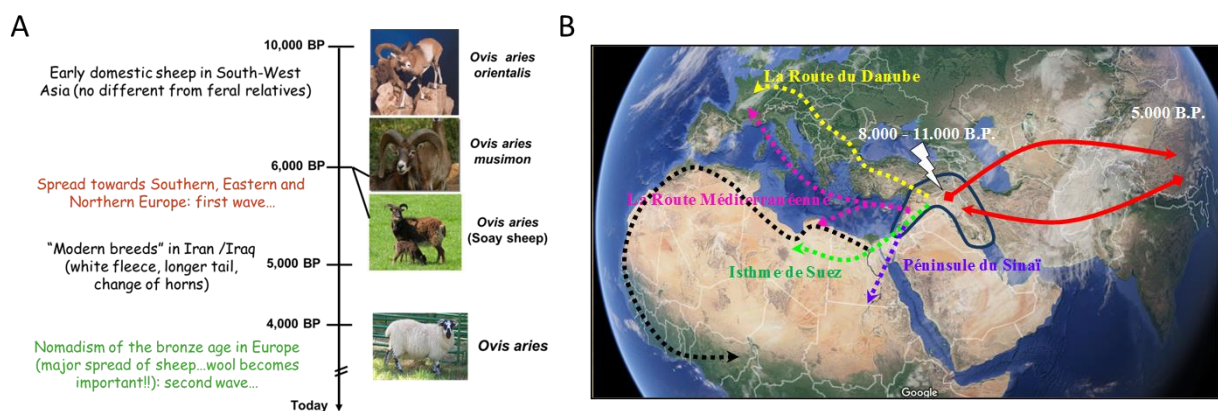


Figure 1-2: A) Chronologie d'événements de domestication du mouton. B) Voies de migration vers l'Asie, l'Afrique et l'Europe. Notez les deux événements de domestication du mouton (10000BP et 5000BP) qui ont été tous deux suivis de migrations vers l'Europe, l'Afrique et l'Asie.

### 1.2.2. Sélection et changement morphologiques au cours de la domestication du mouton

En cours du deuxième événement de domestication et de ses migrations successives, un processus de sélection du mouton a eu lieu et à entraîner des changements morphologiques de l'espèce résultant aux races dites « modernes » de nos jours. Ces changements morphologiques sont à la base des études archéologiques grâce à l'analyse des ossements de ces animaux. On note notamment une réduction de la taille des animaux domestiques par rapport aux espèces sauvages ainsi qu'une diminution de l'âge des animaux abattus qui sont en général plus vieux lorsqu'ils résultent d'une activité de chasse. Ces changements dans la taille du corps sont souvent un marqueur de la domestication [19]. Les changements morphologiques observés lors de l'analyse des restes d'ovins provenant de différents sites néolithiques (sixième et cinquième millénaire) montrent une absence de cornes chez les femelles et un raccourcissement des os des membres. Il est probable que le pelage a également changé par rapport aux animaux sauvages à la suite de mutations et d'une sélection

artificielle, peut-être dès le sixième millénaire. Le manteau extérieur du mouton sauvage est raide, velu et couvre une sous-couche mince de laine qui ne pousse que pendant l'hiver. Chez les moutons domestiqués, la toison est entièrement composée de sous-couches laineuses qui se développent toute l'année [14]. La plupart de ces caractéristiques de domestication des moutons, telles que l'altération de la forme de la corne, les brebis « hornless », l'engorgement (grossissement) et la longueur de la queue, et la toison laineuse et blanche étaient déjà retrouvées en Asie occidentale dès 3000 Av. J.-C. En effet, tous ces caractères sont visibles sur des représentations picturales de Mésopotamie et sont également décrits dans les textes babyloniens.

D'après les données archéologiques et génétiques, on distingue deux groupes de moutons ayant une origine ancestrale commune mais une histoire génétique différente [20] :

- Les moutons dits à queue fine qui sont les premiers moutons arrivés en Afrique, il y a 7500 et 7000 ans par l'Isthme de Suez et /ou par le sud de la Péninsule du Sinaï [21]
- Les moutons dits à queue grasse (dont fait partie la race Awassi qui est l'objet de ma thèse) arrivés en Afrique par le nord-est du continent et la corne de l'Afrique [20]. Plusieurs évidences archéologiques et bibliques révèlent la présence de moutons à queue grasse dans les régions du bassin méditerranéen il y a au moins 6.000 Av. J.-C [5] (Figure 1-3).

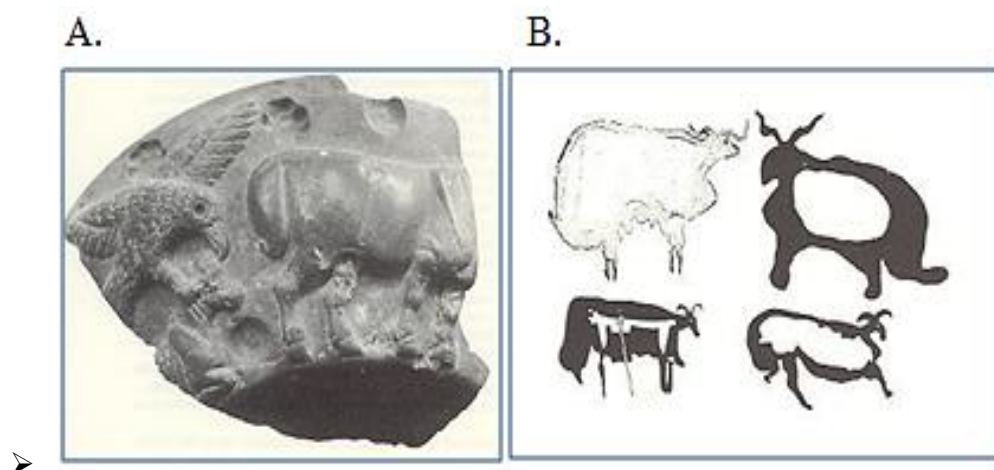


Figure 1-3: A. Moutons à queue grasse sur un bol en pierre de la période d'Uruk d'Ur III (3000 Av. J.-C.). (Le Musée Métropolitain de l'Art) B. Moutons à queue grasse. Gravures rupestres de l'Arabie centrale, deuxième ou au début du premier millénaire avant JC. [22].

Les principaux changements dans les traits morphologiques tels que l'allongement de la queue et l'accumulation de graisse sont considérés comme des changements majeurs qui ont suivi la domestication. Ce trait morphologique de queue grasse peut avoir été acquis longtemps après la domestication des races dites à queues longues et fines de l'Asie. Les premiers animaux à

queue grasse représentés avaient une queue courte et large et la sélection de ce trait tend vers l'augmentation de graisse [23]. En effet, ce dépôt de graisse dans la queue représente une accumulation de réserve, similaire aux bosses du chameau et du dromadaire [5]. Ce phénotype « queue grasse » a évolué dans les Steppes et les déserts caractérisés par de longues périodes de sécheresse et de raréfaction de nourritures pour les animaux. La queue grasse semble ainsi originaire des steppes d'Asie Centrale/Mésopotamie, où plusieurs gravures rupestres de moutons à queue grasse sont représentées (Figure 1-3). Dans les steppes et les déserts, c'est-à-dire, là où le peuple de ces régions était dépourvu d'autres animaux producteurs de graisse, les moutons à queue grasse était particulièrement sélectionnés pour l'élevage [5].

### 1.2.3. Classification du genre *Ovis*

Les Chèvres et les moutons appartiennent tous les deux à la sous-famille des *Caprinae* de la famille des *Bovidae* dans le sous-ordre Ruminantia de l'ordre Artiodactyla [24-26]. Ce sont des ruminants typiques à sabots clochés de taille relativement petite. Le genre *Ovis* est l'un des genres des mammifères le plus complexe de point de vue son évolution et sa systématique.

La classification du genre *Ovis* (*Ovis* Linnaeus, 1758 – sheep) n'est pas claire et controversée. Néanmoins six espèces de moutons sauvages et une espèce de moutons domestiques sont actuellement décrites (Figure 1-4) [27, 28] :

1. *Ovis orientalis* (*O.o.*) : Mouflon, avec une sous espèce en Asie (*O.o.laristanica*) et une en Europe (*O.o.musimon*) ;
2. *Ovis ammon* : Argali, vivant en Asie centrale, au Tibet en Chine et en Mongolie ;
3. *Ovis vignei* : Urial. *Ovis orientalis* est quelque fois considérée synonyme d'*O. vignei* (*Ovis orientalis vignei*) vivant en Asie mineur ;
4. *Ovis canadensis* : Bighorn, au Canada et aux Etats-Unis d'Amérique ;
5. *Ovis nivicola* : le mouflon des neiges de Sibérie ;
6. *Ovis dalli* : Mouflon de Dall (Thinhorn) de l'Alaska ;
7. *Ovis aries* : Moutons domestiques (races modernes) de nos jours.

Toutes ces espèces *Ovis* sont entièrement inter-fertiles et peuvent donc être considérées comme monotypiques.

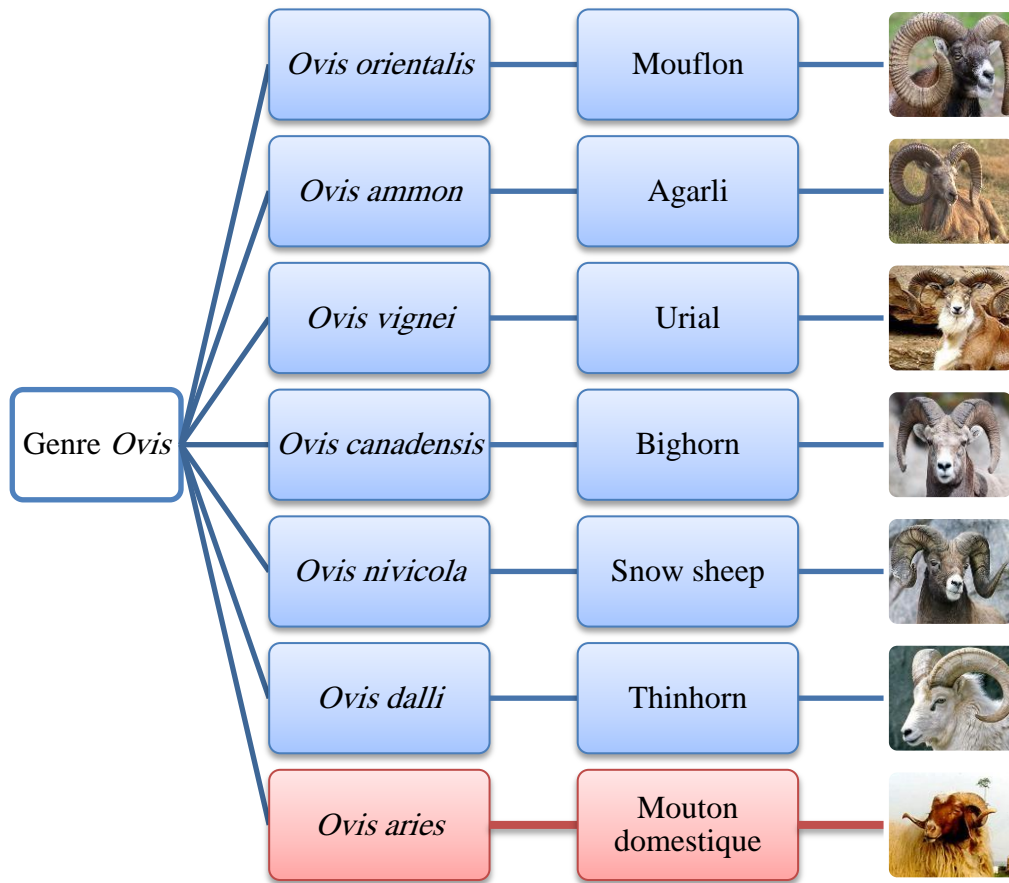


Figure 1-4: La Taxonomie du genre *Ovis* selon Geist 1991. Voir détails dans le texte.

Pour les distinguer des espèces sauvages, tous les moutons domestiques sont maintenant classés comme *Ovis aries*. Ce groupe renferme 2502 races distribuées sur tous les continents dont presque la moitié de ces races se trouve en Europe (FAO, DAD-IS Feb. 2017) (Tableau 1-1).

Tableau 1-1: Distribution mondiale des races ovines domestiques (*O. aries*)

Continent	Nombre de Race
Afrique	277
Asie pacifique	349
Europe	1267
Amérique Latine & Les Caraïbes	278
Proche Orient	253
Amérique du Nord	78
<b>Total</b>	<b>2502</b>

Cependant les moutons sauvages vivent tous dans l'hémisphère nord de la planète comme présenté par Rezaei et ses collègues [13].

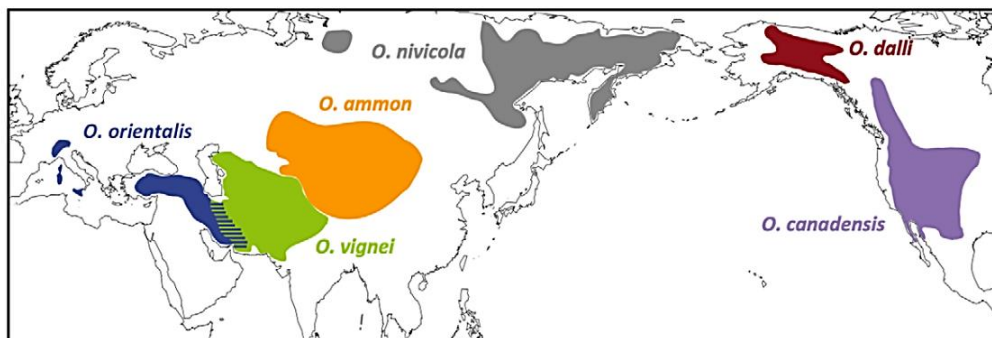


Figure 1-5: La répartition mondiale des sept races de mouton sauvage [13].

### 1.3. La race AWASSI

#### 1.3.1. Origine et Évolution

Le nom Awassi est attribué à la tribu El Awas qui vivait entre le Tigre et l'Euphrate. En arabe littéraire, awas est le terme pour le vêtement de chameau rouge et blanc ou pour un mouton blanc [29]. Awassi est la race de mouton domestique, donc appartenant à *O. aries*, la plus répandue dans le sud-ouest de l'Asie. C'est la race dominante en Irak, le mouton le plus important de la République arabe syrienne et la seule race indigène d'ovins au Liban, en Jordanie et en Israël [5, 30]. En Turquie, les Awassi représentent 3.5% de la population ovine [31] et sont localisés dans le sud de l'Anatolie dans une bande frontalière avec la Syrie. Les propriétés physiques et physiologiques de la race Awassi semblent être très proches du prototype à partir duquel les moutons à queue grasse de l'Asie, de l'Afrique et de l'Europe sont issus. Ainsi la majorité des races ovines présentes en Chypre, en Afrique du Nord, en Turquie

et en Iran sont semblables à la race Awassi. De même, la race Ronderib Afrikander de l'Afrique du Sud et le mouton mongol de l'Asie orientale sont très similaires à cette race en présentant une queue grasse [32]. L'origine de la race Awassi n'est actuellement pas très claire mais, comme mentionné précédemment, certains chercheurs suggèrent que les moutons à queue grasse ont évolué dans les steppes de la Syrie et de l'Arabie où les conditions climatiques favorisaient le développement de réserves de graisse [5, 33].

### **1.3.2. L'agriculture et la race Awassi au Liban**

#### **1.3.2.1. Agriculture et géographie du Liban**

Bien que le rôle de l'agriculture à l'échelle économique nationale du Liban soit relativement en déclin, l'agriculture joue toujours un rôle important dans l'économie rurale et a un impact significatif sur le niveau de vie. En 2011-2012, l'agriculture représentait 4% du produit intérieur brut (PIB) du Liban, 15% du total des transactions commerciales en 2012 et employait 6% de la population active (2009) avec néanmoins de grandes variations selon les régions (jusqu'à 25%) [33]. Le Liban est situé sur les rives orientales de la mer Méditerranée avec une superficie totale de 10452 km<sup>2</sup>. Le pays est composé de deux chaînes de montagnes parallèles séparées par la vallée de la Bekaa. Cette variation de la topographie se traduit par des zones de systèmes agricoles diversifiés. Quatre zones principales sont distinguées :

#### **1. Systèmes agricoles intensifs dans la plaine côtière fertile jusqu'à 200m :**

La plaine d'Akkar se trouve près de la frontière nord-ouest avec la Syrie. Elle abrite une large gamme de cultures irriguées, y compris la pomme de terre, les céréales, les fruits, la vigne et les légumes dans les champs ouverts ou dans les serres. Cependant, les zones agricoles au Sud, comme à Saïda et à Tyre se caractérisent par les plantations de bananiers et des citronniers. La production animale gagne en importance, surtout dans les systèmes mixtes où elle est combinée à l'agriculture (Agropastoralismes) avec un mode d'élevage extensif. Les quelques écosystèmes naturels encore épargnés des activités humaines sont les écosystèmes côtiers, marins et riverains.

#### **2. Systèmes de cultures extensives pluviales :**

Sur les terrasses de montagne, la culture pluviale domine principalement au-dessous de 1200 m d'altitude sur les pentes de la chaîne des montagnes ouest du Liban ainsi que dans les hautes altitudes de Baalbek Hermel et de la Bekaa. A cette altitude, les moutons sont rares surtout sur le versant ouest de la chaîne de montagne Mont Liban.



Dans les plaines de Bekaa et de Baalbek-Hermel, les céréales prédominent, ainsi que le pâturage extensif des petits ruminants surtout les moutons où le mode transhumant nomade est le plus répandu.

### **3. Les plantations d'arbres fruitiers sur les pentes de montagne :**

L'abondance des sources d'eau sur les pentes de montagne permet la culture d'arbres fruitiers irrigués, ainsi que de légumes d'été. Cette zone est une voie naturelle pour les petits ruminants vers les pâturages estivaux des hautes altitudes, surtout les chèvres. L'élevage est traditionnel et à petite échelle pour les moutons ; La topologie géographique des montagnes ne convient pas trop à la race Awassi dont la faiblesse de ses pattes l'empêche de grimper. Les écosystèmes naturels comprennent des forêts variées et uniques, combinant des chênes, des pins, des cèdres, des sapins et des genévriers, soit dans des peuplements forestiers purs ou mixtes.

### **4. Systèmes de cultures irriguées intensives dans la vallée de la Bekaa :**

Cette plaine constitue la principale zone agricole et périmètre irrigué du pays. Il comprend une mosaïque de systèmes agricoles, y compris les cultures irriguées annuelles et pérennes, et la pêche. La production animale est également importante dans la Bekaa ; La volaille et les fermes laitières sont en vigueur, principalement dans la Bekaa centrale. Les moutons et les chèvres dominent dans les parties Nord et Est.

En ce qui concerne le bétail et les petits ruminants, les systèmes transhumants nomades semi-extensifs dominent à Baalbek-Hermel, alors que le système semi-extensif est pratiqué dans la Bekaa, l'Akar et le Nord (Figure 1-6). Au Sud-Ouest, Saïda et Tyre, les systèmes agropastoralismes sont les rependus où on voit souvent les troupeaux de moutons dans les plantations et les champs de cultures. Les productions laitières intensives se retrouvent dans tous les systèmes agricoles susmentionnés à des intensités différentes. Malgré l'absence des terrains de pâturages, plus de 15800 éleveurs procèdent à ce type d'élevage intensif qui représente 9% de la main d'œuvre agricole libanaise.

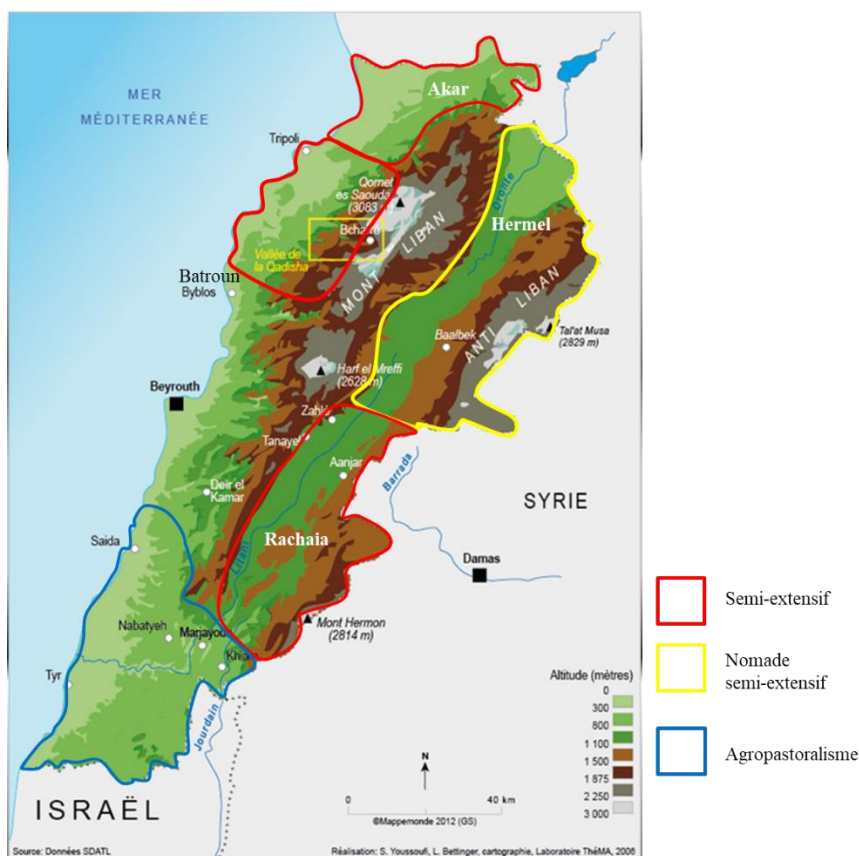


Figure 1-6: les systèmes d'élevage de bétail au Liban.

### 1.3.2.2. Importance de la race Awassi au Liban

Les ovins Awassi ont été toujours connus comme étant des composants essentiels de la vie rurale des différents pays méditerranéens ainsi qu'au Liban et en Syrie. Outre son lait et sa laine, les populations du Moyen-Orient en général, et du Liban en particulier, sont de grands consommateurs de viandes rouges des petits ruminants provenant en majorité de l'élevage de la race Awassi [34]. Vu que la consommation de ses produits rentre dans la tradition de ces pays, cette race ovine joue un rôle indispensable dans la vie des agriculteurs et dans leur subsistance. La production de la race Awassi par les éleveurs, en particulier au Liban, est aujourd'hui favorisée par une demande croissante et une hausse des prix de par la qualité de sa viande [35]. En plus de son rôle économique, l'élevage des Awassi a un rôle écologique, environnemental et culturel au sein de ces pays [5, 36]. Les offrandes de bétail, surtout le mouton, font partie des traditions religieuses et culturelles de certaines communautés et régions au Liban. La race Awassi est l'offrande de choix durant la fête de l'Adha chez les musulmans, et lors des réceptions de grandes personnalités politiques, sociales ou familiales. De plus, la laine de cette race est aussi très recherchée pour fabriquer des couettes et des matelas ; tandis que la peau est utilisée comme tapis ou décorations. En 2010, le nombre

d'Awassi présent au Liban était estimé à environ 270 000 têtes tenues par 4150 éleveurs, donc une moyenne de 65 têtes par éleveur, notons que le quart de ces éleveurs ne possèdent aucun terrain de pâturages et détient 36% du cheptel du pays [37] et donc optent pour un système d'élevage extensif ou de transhumance. Plus récemment, en 2014, le nombre d'Awassi au Liban a atteint 450 000 têtes avec l'arrivée de troupeaux provenant de la Syrie en guerre ces dernières années [33].

### 1.3.2.3. Caractéristiques de la race Awassi

#### a. Traits physiques de la race

La race Awassi a été décrite pour la première fois par Mason en 1967 et plus tard par Epstein en 1985. Le corps du mouton Awassi est recouvert de longue laine de couleur blanche crémeuse (Figure 1-7). La tête est longue et étroite de couleur brune ou noire avec un profil de front convexe. Les béliers ont généralement des cornes qui sont tordus vers l'arrière avec des pointes dirigées vers l'extérieur, tandis que les brebis sont dépourvues de cornes. Les oreilles sont de tailles moyennes et souvent pendantes. Le cou est relativement long, fort chez les béliers, moins robuste chez les brebis. Le corps est principalement étroit et les jambes sont de longueur moyenne et large. La queue de la race Awassi est caractéristique des moutons à queue grasse se terminant au-dessus des jarrets [36].



Figure 1-7: Un bélier avec des brebis et un agneau de la race Awassi à Sanine - Mont Liban (Miguel Salome)

### *b. Traits physiologiques de la race*

La race Awassi est une race très rustique, très calme, utilisée pour une gamme de produits tels que le lait, la viande et la laine. Cette race est connue pour être la plus grande race de traite après la race Friesian (Friesian East) [38]. Certaines études ont décrit l'Awassi comme l'une des races de brebis laitières les plus productives avec une moyenne de 200-300 litres de lait en 158 jours de lactation [39, 40] De plus et comme mentionné précédemment, sa viande est connue principalement par sa forte saveur probablement due à la présence d'une quantité importante d'acides gras. La race Awassi possède également des caractéristiques très souhaitables en ce qui concerne l'endurance aux fluctuations nutritionnelles, la résistance aux maladies et aux parasites ainsi que la tolérance aux températures extrêmes [41, 42]. En 1985, Epstein a qualifié l'Awassi de « *mouton audacieux qui s'adapte parfaitement aux conditions difficiles de son environnement* ». La race s'est étendue de son lieu d'origine en Asie du Sud-Ouest à tous les continents, y compris en Amérique du Sud. C'est d'ailleurs la seule race de mouton d'origine non européenne qui a une distribution aussi large [36].

## **1.4. Problématique**

La race Awassi est l'une des plus anciennes races de moutons déjà référencée avant les temps bibliques. Aujourd'hui, l'élevage des moutons de cette race joue un rôle économique et social très important dans les pays du Moyen-Orient et particulièrement au Liban. En raison de sa situation géographique, le Liban constitue un lieu clé de la domestication du mouton et une étape incontournable dans toutes les routes commerciales ancestrales : phéniciennes, romaines, grecques ou bien encore de l'Égypte antique sans oublier la route de soie qui reliait l'Asie de l'Est à l'Asie de l'Ouest. Jusqu'à présent et malgré l'importance historique et contemporaine de cette race, aucune étude ne s'est attachée à caractériser les populations de mouton de la race Awassi du Liban ; ainsi il apparaît important de mieux comprendre l'origine, l'évolution et la diversité de cette race. **L'objectif principal de ma thèse a donc été de caractériser génétiquement la race Awassi du liban. Pour atteindre cet objectif, j'ai utilisé deux types marqueurs génétiques : i) les rétrovirus endogènes de type Jaagsiekte Sheep Rétrovirus (enJSRV) et ii) l'ADN mitochondrial (cytochrome B et D-Loop).**

En parallèle de ma thèse, j'ai continué mon travail au sein du laboratoire de référence nationale pour le diagnostic des maladies animales à l'Institut de Recherche agronomique Libanais- IRAL. Ces travaux ont abouti à trois publications jointes à la fin de ce manuscrit.

## Chapitre 2 : *Les Rétrovirus*

---

## 2.1. Les Rétrovirus

### 2.1.1. Historique

Le développement de la biologie au cours du siècle dernier est passé du niveau organismique au niveau moléculaire. Cette large tendance historique a été également suivie par la virologie et la rétrovirologie [43]. Les rétrovirus ont été découverts au début du 20<sup>ème</sup> siècle. La première description de ce qui s'est avéré plus tard être une maladie rétrovirale, l'anémie infectieuse équine, remonte à 1904 [44]. De même, deux autres études consacrées aux maladies néoplasiques chez les poulets ont marqué cette découverte. En 1908, l'équipe de deux médecin-vétérinaires danois, Vilhelm Ellermann et Oluf Bang, a montré que la leucose du poulet, une forme de leucémie et de lymphome, était d'origine virale. En 1911, Peyton Rous au Rockefeller Institute de New York a rapporté la transmission cellulaire d'un sarcome chez les poulets [45, 46]. Les agents découverts par Ellermann et Bang sont maintenant connus sous le terme collectif de virus de leucose aviaire ou ALV tandis que le virus isolé par Rous porte son nom, le virus du sarcome de Rous (RSV). L'ensemble de ces virus constituent le genre aviaire de type C, souvent désigné sous le nom de virus aviaire du groupe sarcome-leucose (ASLV) [43]. La découverte de tumeurs induites par un virus a été ensuite étendue aux mammifères, lorsque John Bittner [47] et Ludwik Gross [48] ont démontré que les rétrovirus peuvent causer une maladie néoplasique chez la souris. Au cours des deux décennies suivantes, de nombreux virus de ce type responsables de maladies néoplasiques chez la souris, le chat, le bétail et le singe ont été identifiés. Le nombre de tumeurs induites par des virus chez les volailles a également considérablement augmenté; au début des années 1930, environ 20 isolats viraux causant des tumeurs histologiquement distinctes ont été décrits [49] et la liste des nouveaux rétrovirus aviaires ne cesse de croître [50]. Plusieurs virus isolés, pendant cette période, sont d'ailleurs devenus des systèmes modèles importants, étudiés activement aux niveaux cellulaire et moléculaire à ce jour. Dans les cinq premières décennies après la découverte de ces virus, l'infectiosité et l'oncogénicité ont été généralement évaluées directement chez l'animal hôte. Une première étape vers la quantification virale directe a été réalisée par inoculation du RSV sur la membrane chorioallantoïde de l'embryon de poulet [51]. Ce n'est qu'en 1958, quand Howard Temin et Harry Rubin ont appliqués le potentiel de la culture cellulaire à la rétrovirologie et que des titrages sensibles et quantitatifs ont été mis en place pour doser l'infectiosité et de la tumorigénicité de ces virus. Mais l'évènement qui fait date pour la rétrovirologie n'est autre que la découverte en 1970 de la transcriptase inverse (ou reverse transcriptase « RT ») codée par ces virus et qui permet de synthétiser une

molécule d'ADN double brin à partir du génome d'ARN viral simple brin [52, 53]. Cette étape clé et caractéristique du cycle de ces virus est à l'origine de leur nom "rétrovirus" [54]. Depuis leur découverte, on sait aujourd'hui que les rétrovirus infectent de nombreuses espèces de vertébrés et invertébrés, y compris les poissons, les oiseaux, les reptiles et les mammifères, et sont impliqués dans une grande variété de maladies, comme les cancers (leucémies, carcinomes mammaires), les immunodéficiences et l'arthrite [43]. L'exception à cette règle, sont les Foamy virus du genre *Spumavirus* qui sont considérés non pathogènes pour les animaux naturellement infectés tels que les primates sauvages (Nonhuman primates- NHPs), les félins, les bovins, les petits ruminants et les équines [55]. Ces virus zoonotiques infectent également l'homme de manière persistante mais pour tous les cas étudiés, l'infection semble asymptomatique [56].

### **2.1.2. Taxonomie & Classification**

Les rétrovirus appartiennent à la famille des *Retroviridae*. Ce sont des virus à ARN enveloppés définis par des dénominateurs taxonomiques communs qui incluent la structure génomique et les propriétés répliquatives [57]. Les rétrovirus sont caractérisés par leur stratégie de réplication. En effet, ils ont la capacité de transcrire leur génome d'ARN en ADN bicaténaire qui s'insère ensuite de manière stable et permanente dans le génome de l'hôte infecté. Cet ADN intégré est alors appelé provirus. Comme mentionné précédemment, l'étape de transcription inverse est réalisée par la transcriptase inverse tandis que l'étape d'intégration de l'ADN viral est médiée par une autre enzyme virale appelée l'intégrase (IN) [43]. Les rétrovirus ont été initialement classés en quatre groupes, nommés de type A à D, en fonction de la morphologie et de la position de la capsid des particules virales observées par microscopie électronique (Figure 2-1). Les virus de type A sont définis comme des particules cytoplasmiques non enveloppées. Elles sont observables uniquement à l'intérieur des cellules et proviendraient de rétrovirus endogènes. Par ailleurs, cette structure est semblable à celle des virions immatures [58]. Les virus de type B, tels que le Jaagsiekte Sheep Retrovirus (JSRV) et le Mouse Mammary Tumor virus (MMTV), sont enveloppées, extracellulaires et présentent un core excentré [59-61]. Inversement, les virus de type C contiennent un core central, ALV étant le prototype [62-64]. Enfin, les virus de type D possèdent un core en forme de barre et sont légèrement plus grands (120nm). Ils comprennent le virus du singe Mason Pfizer Monkey virus (M-PMV) [58].

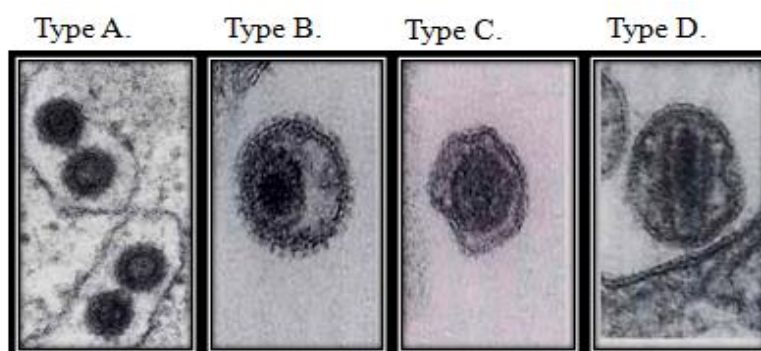


Figure 2-1 : Morphologie des rétrovirus de type A-D observée par microscope électronique [43]

Depuis 2002, le Comité international sur la taxonomie des virus (ICTV) a reclassé les rétrovirus en deux sous-familles, les *Orthoretrovirinae* et les *Spumaretrovirinae*. La majorité des rétrovirus sont inclus dans les *Orthoretrovirinae* divisés en six genres (Alpharetrovirus, Betaretroviruses, Gammaretroviruses, Deltaretroviruses, Epsilonretrovirus, Lentivirus) basés sur la divergence au niveau du génome viral (définie par la similarité de la séquence de la transcriptase inverse et la complexité du génome), le site d'assemblage des virions et leur morphologie. La famille des *Spumaretrovirinae* ne comprend qu'un seul genre, les Spumavirus, dont le prototype est le foamy virus humain (HFV) (Tableau 2-1).

Tableau 2-1: classification au sein de la famille des *Retroviridae* selon l'ICTV

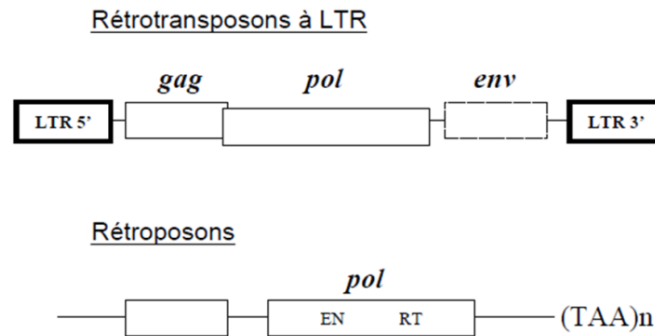
Sous-famille	Genre	Prototype	Génome	Morphologie
Orthoretrovirinae	Alpharetrovirus	Avian leucosis virus (ALV)	Simple	C-type
	Betaretrovirus	Mouse mammary Tumour virus (MMTV)	Simple	B-type D-type
	Gammaretrovirus	Murine leukaemia virus (MLV)	Simple	C-type
	Deltaretrovirus	Bovine leukaemia virus (BLV)	Complex	C-type
	Epsilonretrovirus	Walleye dermal sarcoma virus (WDSV)	Simple	C-type
	Lentivirus	Human immunodeficiency virus-1 (HIV-1)	Complex	C-type
Spumaretrovirinae	Spumavirus	Human foamy virus (HFV)	Complex	D-type



### 2.1.3. Les rétrovirus endogènes

Les rétrovirus sont également différenciés en deux catégories dites des rétrovirus « exogènes » et « endogènes ». Les rétrovirus exogènes sont transmis horizontalement d'un hôte infecté à un autre non infecté comme tout autre virus [65]. Parfois, les rétrovirus peuvent infecter les cellules germinales de l'hôte conduisant à l'intégration de leur génome viral dans celui de ces cellules. Ces provirus intégrés sont ensuite transmis verticalement à la descendance en suivant les lois mendéliennes [66]. On parle alors de « rétrovirus endogènes » [67] qui, s'ils ne sont pas délétères pour leur hôte ou s'ils sont bénéfiques pour celui-ci, se fixent et se maintiennent dans le génome de l'espèce considérée au cours de l'évolution. Par ailleurs, certains rétrovirus récemment endogénéisés sont encore complets et intacts conservant ainsi leur capacité à produire des virus infectieux [68, 69]. Ils peuvent donc être transmis verticalement en tant que provirus mais aussi potentiellement par infection horizontale en compétition ou coexistence avec leur homologue exogène. Avec le temps, certains ERVs délétères ou neutres pour leur hôte accumulent des mutations, des délétions et deviennent inactifs au cours du temps tandis que d'autres provirus ou certaines de leurs séquences se fixent dans le génome *via* une sélection positive lorsque ceux-ci sont bénéfiques pour leur hôte [70]. D'un point de vue génétique, les ERVs sont des éléments transposables (ETs). Les éléments transposables sont des séquences moyennement répétées des génomes qui ont la propriété de changer de localisation chromosomique. Une grande diversité d'éléments est observée, la classification s'effectue selon l'intermédiaire de transposition qui est soit ARN (classe I) soit ADN (classe II), et ces classes sont organisées selon la structure génomique des éléments (Figure 2-2) [71]. La classe I regroupe ce que l'on nomme les rétroéléments dont font partie les ERVs du fait d'une transposition *via* un intermédiaire ARN. Ces ETs sont des constituants importants des génomes puisqu'on estime aujourd'hui leur représentation à plus de 40% pour les génomes de mammifères [72], 50% pour le génome du maïs [73, 74] et entre 15% et 22% pour le génome de *Drosophila melanogaster* [75, 76]. Ils contribuent grandement à la diversité génétique et l'évolution d'un grand nombre d'organismes [77].

## Les éléments de classe I: Les rétrotransposons



## Les éléments de classe II: Les transposons

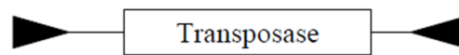


Figure 2-2: Structure des deux principales classes d'éléments transposables. Les éléments de classe I et II transposent respectivement via un intermédiaire ARN ou ADN. Les rétrotransposons à LTR sont composés de LTR « Long Terminal Repeat » encadrant un, deux ou trois cadres ouverts de lecture. Le gène *gag* code pour la matrice (MA), capsid (CA) et nucléocapside (NC). Le gène *pol* des rétrotransposons code une EN (endonucléase) et une RT (transcriptase inverse). Les transposons codent une transposase encadrée d'ITR (Inverted Terminal Repeat). Le gène *env* donne la glycoprotéine de surface (SU) et la glycoprotéine transmembranaire (TM).

L'importance biologique des ETs et des ERVs en particulier a été longtemps discutée et ils étaient considérés comme des ADN non fonctionnels des génomes « junk DNA » [78]. Cependant, au cours des dernières décennies, plusieurs études ont démontré les rôles bénéfiques majeurs des ERVs pour leurs hôtes. En effet, certains de ces ERVs ou séquences issues d'ERVs assurent des fonctions fondamentales comme par exemple lors du processus de placentation chez les mammifères [79] et/ou en étant capables de bloquer le cycle de réplication des virus exogènes homologues et donc en jouant un rôle protecteur pour leur hôte [68, 80, 81]. Jusqu'à présent, les ERVs ont été trouvés dans le génome de tous les vertébrés [82]. Par exemple, le séquençage et l'analyse du génome humain a révélé que les ERVs représentent une partie importante de celui-ci (8%, avec environ 98 000 éléments et fragments) [74, 78]. Ces ERVs sont classés en deux catégories: « anciens » et « modernes » selon le moment de leur intégration dans le génome de leur hôte [83]. Les ERVs anciens se sont insérés dans le génome avant la spéciation et de ce fait sont présents dans chaque individu au même emplacement génomique des espèces génétiquement apparentées. Les ERVs modernes sont, quant à eux, insérés de manière polymorphique dans le génome d'individus de différentes espèces [82, 84]. Contrairement aux ERVs « anciens », la plupart des ERVs « modernes » coexistent avec leur virus homologue exogène, ce qui a conduit à

l'hypothèse que le processus d'endogénéisation est l'une des étapes qui contribue à l'extinction des rétrovirus infectieux transmis horizontalement [85]. Les ERVs ne sont pas inclus dans le système de classification officiel, mais ils sont généralement divisés en trois classes selon leur degré de parenté avec un rétrovirus exogène (Figure 2-3) [86] :

Classe I : endogènes similaires aux gamma- et epsilon-rétrovirus.

Class II : endogènes similaires aux beta-, delta-, alpha-rétrovirus et aux lentivirus.

Class III : endogènes similaires aux spumavirus.

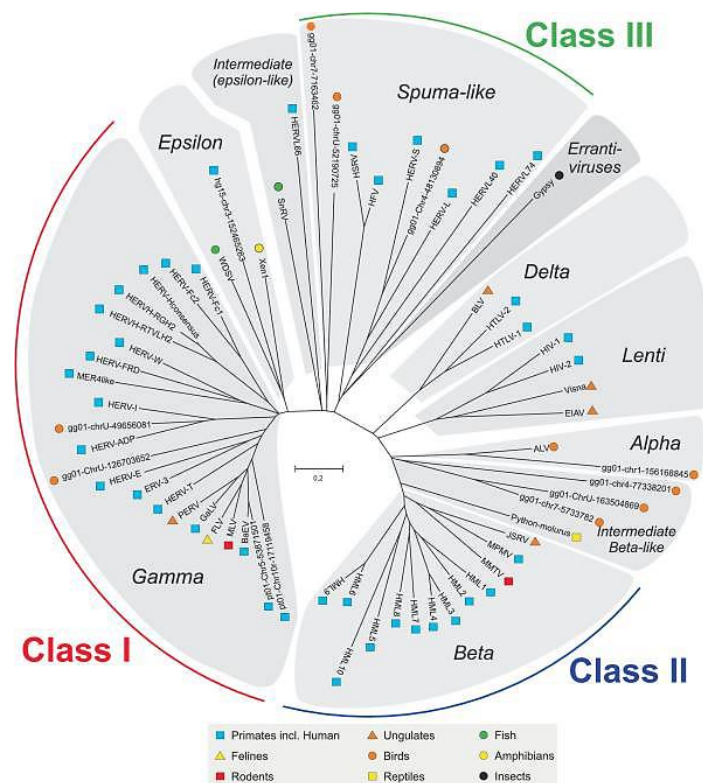


Figure 2-3 : Classification des ERVs selon leur degré de parenté avec un rétrovirus exogène Dendrogramme neighbor joining (NJ) des sept genres de rétrovirus : alpha, bêta, gamma, delta, epsilon, lenti et spuma. Les différentes classes de rétrovirus endogènes sont indiquées en périphérie et les espèces hôtes sont représentées par des symboles à côté de chaque unité taxonomique. [86].

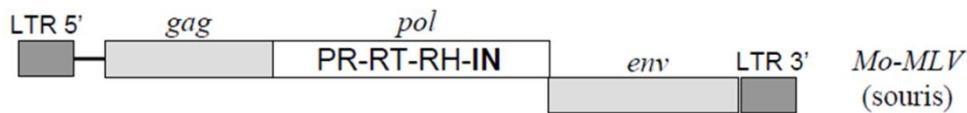
### 2.1.4. Organisation génomique des rétrovirus.

Les rétrovirus sont des virus enveloppés de 80 à 100 nm de diamètre. Leur génome est diploïde et composé de deux brins monocaténares d'ARN non-segmentés, de polarité positive, dont chacun est d'environ 7 à 12 kb [43]. La forme provirale intégrée des rétrovirus présentent aux extrémités deux régions répétées appelées LTR (Long Terminal Repeat), lesquelles encadrent les séquences non codantes dites UTR (UnTranslated Region) et les phases ouvertes de lecture (ORFs) appelées *gag*, *pol* et *env* des rétrovirus [63]. On distingue

deux catégories de rétrovirus selon l'organisation de leur génome : les virus dits simples et complexes (Figure 2-4) [87].

### *Retroviridae*

#### Structure simple



#### Structure complexe

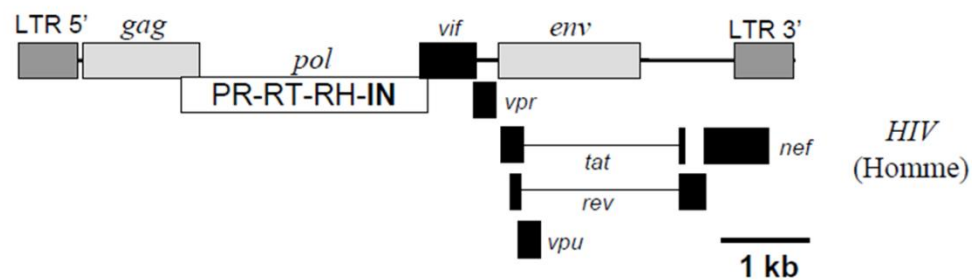


Figure 2-4: Structure génomique des Retroviridae. Les rétrovirus peuvent être de structure simple ou complexe. La structure simple est illustrée par le Mo-MLV (virus de la leucémie murine de Moloney) renfermant trois cadres de lecture : gag, pol et env. Les rétrovirus de structure complexe renferment en plus des cadres de lecture ouverts codant pour des protéines accessoires. Cette dernière structure est illustrée par le rétrovirus HIV humain.

Les rétrovirus simples ne comportent que les ORFs essentiels à leur réplication (*gag*, *pol*, et *env*) tandis que les rétrovirus complexes possèdent des cadres ouverts de lecture supplémentaires codant pour des protéines accessoires. La Figure 2-4 présente la structure de deux rétrovirus types : le virus de la leucémie murine de Moloney (Mo-MLV) qui a une structure dite simple et le virus de l'immunodéficience humaine (HIV) qui a une structure dite complexe du fait qu'il code des protéines accessoires impliquées dans la régulation des gènes viraux ou cellulaires. Ces protéines virales accessoires sont en général impliquées dans l'échappement de ces virus à la réponse immunitaire de l'hôte [63, 88].

#### 2.1.4.1. Le gène *gag*

Le gène *gag* (*group specific antigen*) code pour une polyprotéine structurale impliquée dans la formation des particules virales. Cette polyprotéine Gag est clivée par la protéase virale (PR), codée par le virus en sous-unités matures appelées matrice (MA), capsid (CA) et nucléocapsid (NC), qui s'assemblent pour former la capsid de la particule virale (Figure 2-9). D'autre part, Gag est une protéine qui présente une affinité pour les acides nucléiques et intervient dans l'encapsidation des ARNs génomiques dans les particules virales néoformées.

#### 2.1.4.2. Le gène *pol*

Le gène *pol* des rétrovirus code pour une polyprotéine donnant naissance à des protéines enzymatiques telles qu'une protéase (PR), la transcriptase inverse (RT) avec un domaine RNase H [89] et une intégrase (IN), indispensables au cycle de réplication. Les protéines matures sont issues du clivage du précurseur Gag-Pol ou Gag-Pro-Pol. La maturation a lieu au moment de l'assemblage des particules dans le cytoplasme et même plus tard après la sortie de la cellule selon les virus considérés. La protéase est responsable de la maturation de la polyprotéine Gag en MA, CA, NC ainsi que de Pol en PR, RT/RH et IN. L'activité rétrotranscriptase permet la synthèse de l'ADN à partir de l'ARN génomique matrice, tandis que la RNase H détruit cette matrice. L'intégrase permet, quant à elle, l'intégration du provirus dans le génome de l'hôte. Elle contient trois domaines fonctionnels essentiels aux *Retroviridae*. En N-terminal, un motif HH-CC (Paires d'histidine et de cystine qui sont conservées parmi les intégrases rétrovirales), similaire à un doigt de zinc, serait responsable de la formation d'un complexe stable entre l'intégrase et les LTR de la séquence provirale. Cette région semble également promouvoir la multimérisation de l'intégrase du HIV-1 sous forme d'un tétramère actif et son activité enzymatique [90]. La partie centrale, appelée « core », contient le motif DDE (DD<sub>35</sub>E : Aspartate (D), Aspartate (D), Glutamate (E)) qui est la partie catalytique de l'enzyme (CCD : Catalitic Core Domain). La partie C-terminale présente de nombreuses charges positives dont un motif conservé GPY/F qui lierait l'ADN génomique cible [91].

#### 2.1.4.3. Le gène *env*

Chez les rétrovirus, la polyprotéine Env est indispensable à l'étape d'infection des rétrovirus et est synthétisée à partir d'un ARNm épissé, dit subgénomique (Figure 2-5).

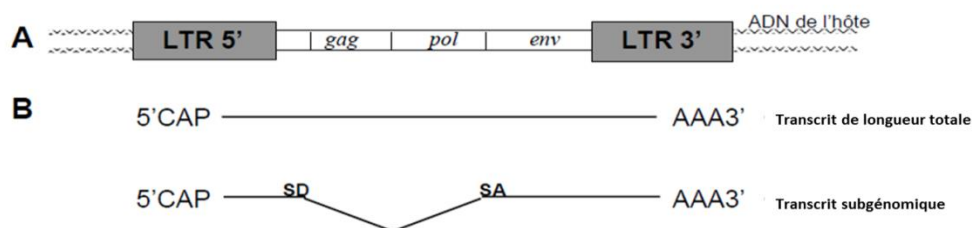


Figure 2-5: Représentation schématique d'un provirus et des transcrits correspondants. (A) Schéma d'un provirus avec trois cadres de lecture et les LTR 5' et 3'. (B) Structure des messagers proviraux (CAP : coiffe en 5' ; SD et SA: sites donneur et accepteur d'épissage ; AAA: queue polyadénylée). Le transcript subgénomique permet la synthèse de la protéine Env.

Env est une protéine à ancrage membranaire possédant un peptide signal en N-terminal. En effet, le précurseur de la protéine Env est synthétisé dans le lumen du réticulum endoplasmique grâce à ce peptide signal. Comme Gag et Pol, elle est clivée, mais cette fois, par des protéases cellulaires. Le site consensus de coupure par la protéase de type furine est RX(K/R)R (R: Arginine; K: Lysine). Les deux sous-unités matures restent liées par des liaisons non covalentes ou par des ponts disulfures. La sous-unité N-terminale donne naissance à une protéine de surface (SU), très glycosylée, qui est responsable de l'interaction avec des facteurs d'attachement ou des récepteurs présents à la surface des cellules cibles. La sous-unité C-terminal est dite transmembranaire (TM) et renferme un domaine d'ancrage à la membrane et un domaine hydrophobe responsable de la fusion membranaire entre l'enveloppe du virus et la membrane plasmique de la cellule cible au moment de l'entrée virale, première étape indispensable d'une infection virale.

#### 2.1.4.4. Les régions non codantes

- **Les LTR (Long Terminal Repeat)**

Les LTRs des rétrovirus contiennent la plupart des éléments de contrôle transcriptionnel tels que le promoteur et des séquences régulatrices nécessaires à son expression. Les LTRs sont de longues régions répétées, localisées aux extrémités de la forme provirale. Chaque LTR se subdivise en trois parties : U3, R et U5 qui ne présentent pas forcément d'homologie entre elles. La région R se définit en 5' par le site d'initiation de transcription et en 3' par le signal d'arrêt de transcription. Cette région centrale R est donc la seule qui soit répétée dans l'ARNm viral après la transcription du provirus. Les régions U3 et U5 se répartissent de part et d'autre de la région R. C'est au cours de la synthèse de l'ADN proviral que sont générés les LTR complets du provirus (Figure 2-7). Chacune de ces régions a une taille variable selon les provirus, mais le plus souvent la région centrale R est plus petite. Il est intéressant de noter que les LTRs d'ERVs peuvent réguler la transcription de gènes avoisinants comme décrit pour les gènes de l'amylase et de la pléiotrophine chez l'homme [92, 93].

- **La région PBS (Primer Binding Site)**

La région PBS est localisée dans la région 5' non-traduite des rétrovirus entre le LTR 5' et le gène *gag*. Cette séquence est complémentaire sur 18 nucléotides à l'extrémité 3' d'un ARN de transfert différent selon le rétrovirus considéré (ARNt) et qui permet l'amorçage de la transcription inverse (Figure 2-7).

- **La région PPT (Poly Purine Tract)**

La région PPT est la plus petite : 9 nucléotides composés d'adénines et de guanines. La région PPT est localisée dans la région 3' non-traduite des rétrovirus entre le gène *env* et le LTR 3'. Elle permet l'initiation de la synthèse du second brin d'ADN, dit le brin « plus » au cours de la transcription inverse (Figure 2-7).

- **La région 5' non-traduite (5' UTR « 5' UnTranslated Region »)**

Cette région est localisée entre le LTR 5' et le gène *gag*. Elle est composée du PBS, de la séquence Psi ( $\psi$ ) qui permet l'encapsidation des ARN génomiques dans les particules virales grâce à la protéine Gag, et souvent des sites donneurs d'épissage utilisés lors de la maturation de certains ARN viraux. La région 5' non traduite est impliquée dans des étapes cruciales de la mobilisation des séquences provirales telles que la synthèse de l'ADN proviral, la dimérisation de l'ARN et l'encapsidation [63, 94]. C'est une région qui peut également renfermer de nombreux signaux comme des séquences régulatrices de la transcription des ARN viraux [95-99].

## 2.2. Cycle de réplication des rétrovirus

La marque distinctive des rétrovirus réside dans la stratégie de réplication qui comprend comme étape essentielle la transcription inverse de l'ARN génomique viral en ADN double brin linéaire puis l'intégration de cet ADN dans le génome de la cellule [43]. On distingue deux phases au cours du cycle des rétrovirus, les étapes précoces (pré-intégration) et les étapes tardives (post-intégration) (Figure 2-6).

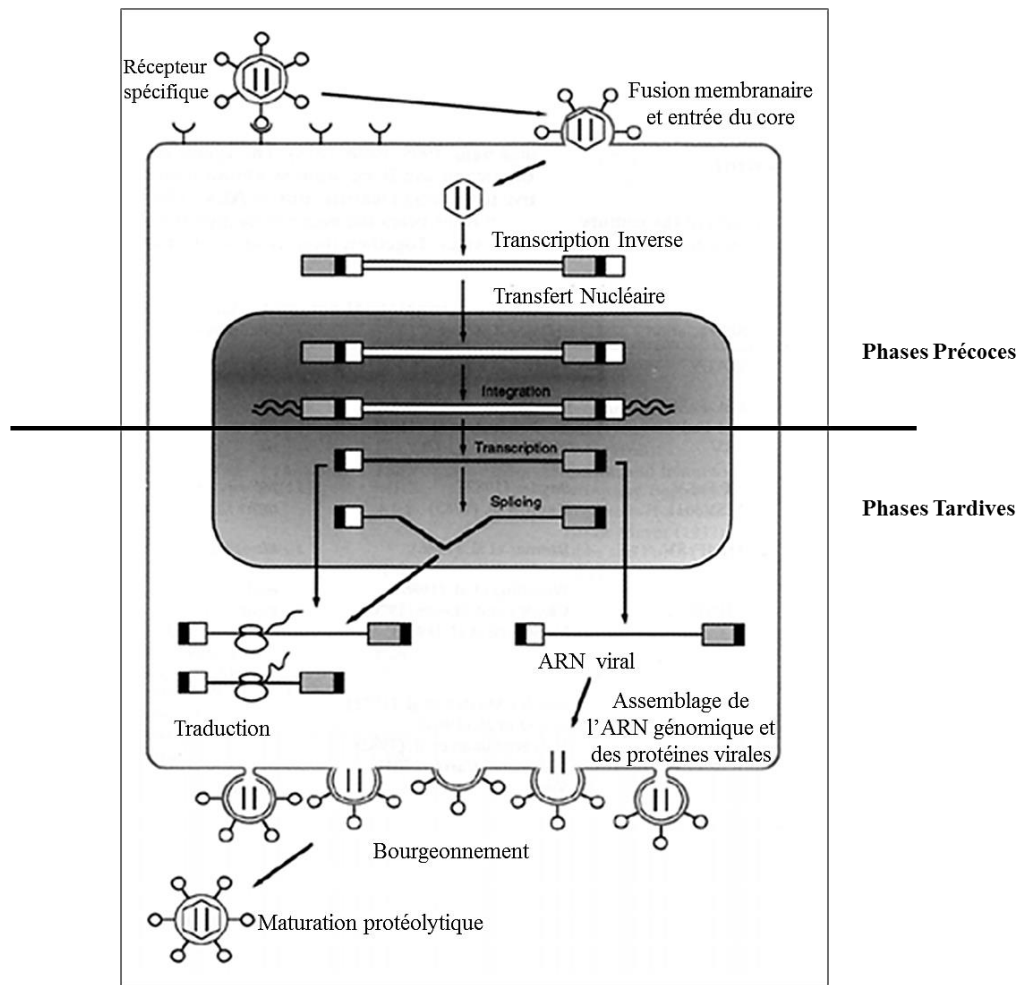


Figure 2-6: Les étapes précoces et tardives du cycle rétroviral d'après Coffin *et al.*,1997.

### 2.2.1. Les phases précoces

- **Infection de la cellule hôte**

Chez les rétrovirus, le virion est formé de deux molécules d'ARN à polarité positive entourées par les protéines de la capsid, le tout étant enveloppé par les protéines d'enveloppe. L'entrée des virus enveloppés dans la cellule cible nécessite la fusion de la membrane plasmidique de la cellule cible et de la membrane virale dans un processus catalysé par la protéine d'enveloppe. L'interaction est réalisée entre la protéine d'enveloppe de surface (SU) et un récepteur situé à la surface des cellules [100]. Puis, les membranes fusionnent grâce au domaine hydrophobe de la sous-unité TM de la protéine Env [101]. Une fois les membranes fusionnées, le core du virion est libéré dans le cytoplasme de la cellule.



- **Le désassemblage et la transcription inverse de l'ARN viral**

Suite à sa libération dans la cellule hôte, le core viral formé par les sous-unités de la protéine Gag (MA, CA, NC) et contenant les enzymes virales et l'ARN génomique se désassemble dans le cytoplasme [102]. Ce processus de désassemblage n'est pas encore très clair mais on pense qu'il est nécessaire pour que la transcription inverse de l'ARN viral soit complète [103]. Le processus de transcription inverse donne naissance à une molécule d'ADN double brin complémentaire à l'ARN génomique et encadrée par les deux LTRs (Figure 2-7). La synthèse de l'ADNc, brin négatif, commence au niveau du LTR5'. Elle est amorcée par un ARNt spécifique et complémentaire à la région PBS sur une vingtaine de bases [104, 105]. La transcription inverse aboutit à la synthèse des régions R et U5 présentes dans le LTR 5'. Ce court ADNc est appelé MSSS « Minus-strand-strong stop DNA ». L'activité RNase H dégrade l'ARN génomique hybridé à l'ADNsss. Ce MSSS est alors transféré au LTR 3' dont il est complémentaire au niveau de la région R, cette étape est appelée saut de brin. La transcription inverse permet alors la synthèse complète du brin négatif. L'ARN génomique est alors dégradé sauf au niveau de la région PPT qui servira d'amorce à la synthèse du brin positif. La synthèse de ce brin est ensuite stoppée une fois la région U5 atteinte. Cette portion du brin nouvellement synthétisé est appelée PSSS, « Plus-strand strong-stop DNA ». L'ARNt est ensuite dégradé libérant la région PBS du brin moins. Un second transfert de brin permet au PSSS de s'hybrider au niveau de la séquence complémentaire au PBS. La synthèse des deux brins, positif et négatif, peut alors être achevée. Une molécule d'ADN double brin (ADN proviral) encadrée de LTR est synthétisée. Les LTR 5' et 3' sont régénérés à partir d'une seule séquence de l'ARN viral, ce qui a pour conséquence de générer des LTR à chaque extrémité qui sont strictement identiques avant l'étape d'intégration dans le génome. Cette étape conduit à la formation d'une copie linéaire double brin extra-chromosomique du rétrovirus. La molécule d'ADN ainsi obtenue pénètre alors dans le noyau pour s'intégrer dans l'ADN de la cellule hôte.

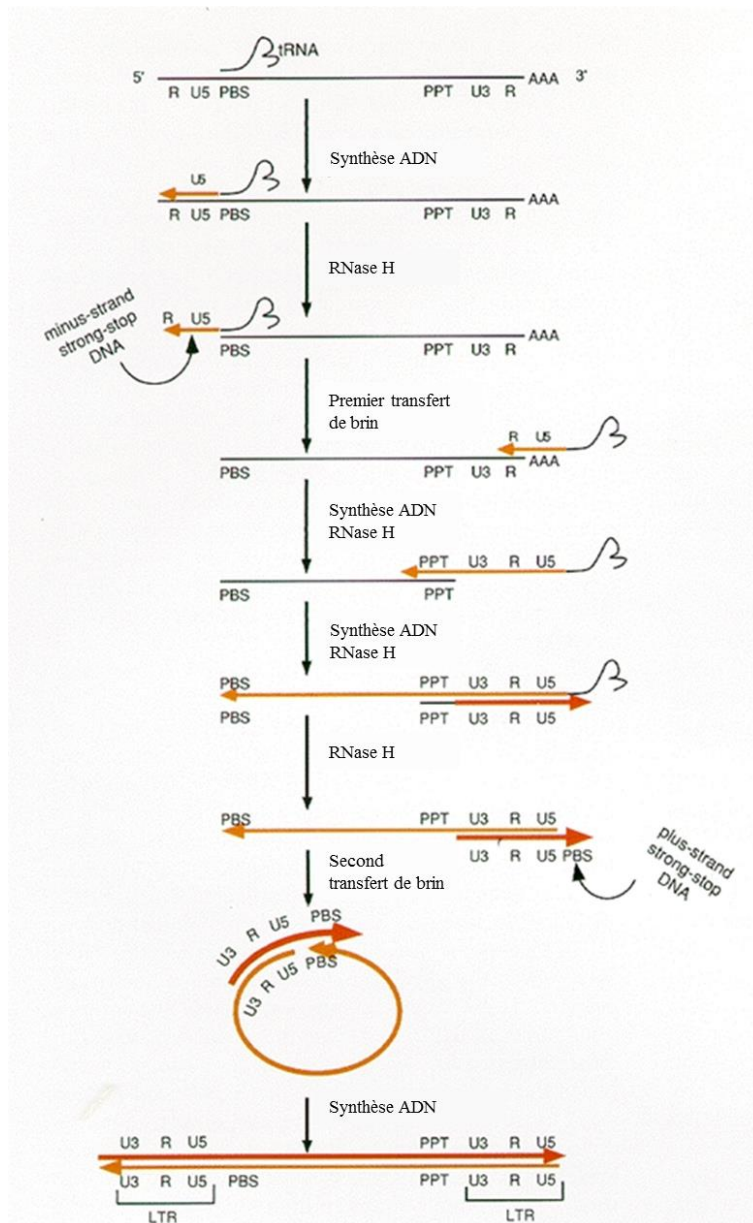


Figure 2-7: Processus de transcription inverse des rétrovirus. La transcription inverse est médiée par la RT et permet la synthèse d'ADN double brin à partir de l'ARN génomique des rétrovirus. Abbreviations: PBS, «Poly Binding Site »; PPT, «Poly PurineTract »; LTR, «Long Terminal Repeat »; MSSS, «Minus-strand-strong stop DNA »et PSSS, «Plus-strand strong-stop DNA »[63].

#### ▪ **Intégration de l'ADN rétroviral dans le génome cellulaire**

Après la transcription inverse, l'intégration du provirus dans le génome de l'hôte est réalisée grâce à une enzyme virale appelée Intégrase (IN) [106]. L'ADN proviral parvient dans le noyau sous la forme d'un complexe nucléoprotéique appelé complexe de préintégration (PIC = PreIntegration Complex). Le PIC contient l'ADN double brin néosynthétisé, l'intégrase virale (IN) ainsi que des facteurs cellulaires nécessaires à l'intégration. La plupart des rétrovirus, à l'exception de HIV et quelques Lentivirus, traversent la membrane nucléaire pendant la division cellulaire en phase M. [106]. Après le ciblage du PIC sur le génome de

l'hôte et la reconnaissance du site d'insertion, la réaction d'intégration du provirus ADN peut avoir lieu. Dans un premier temps, habituellement dans le cytoplasme, la protéine IN clive l'ADN viral en libérant deux (ou, rarement, trois) nucléotides à chaque extrémité 3' libérant ainsi des groupes 3'-OH (Figure 2-8). Lorsque ce complexe arrive dans le noyau, l'IN clive l'ADN cible en générant une coupure décalée variant suivant le rétrovirus considéré. Les groupes 3'OH vont alors se lier aux groupes 5'-P à chaque extrémité de l'ADN hôte [106]. Le provirus s'intègre ainsi dans le génome de l'hôte et les extrémités sont réparées par les enzymes cellulaires créant ainsi une duplication de la séquence cible située aux bornes du provirus intégré et donc spécifique de chaque provirus inséré dans le génome à une localisation chromosomique précise.

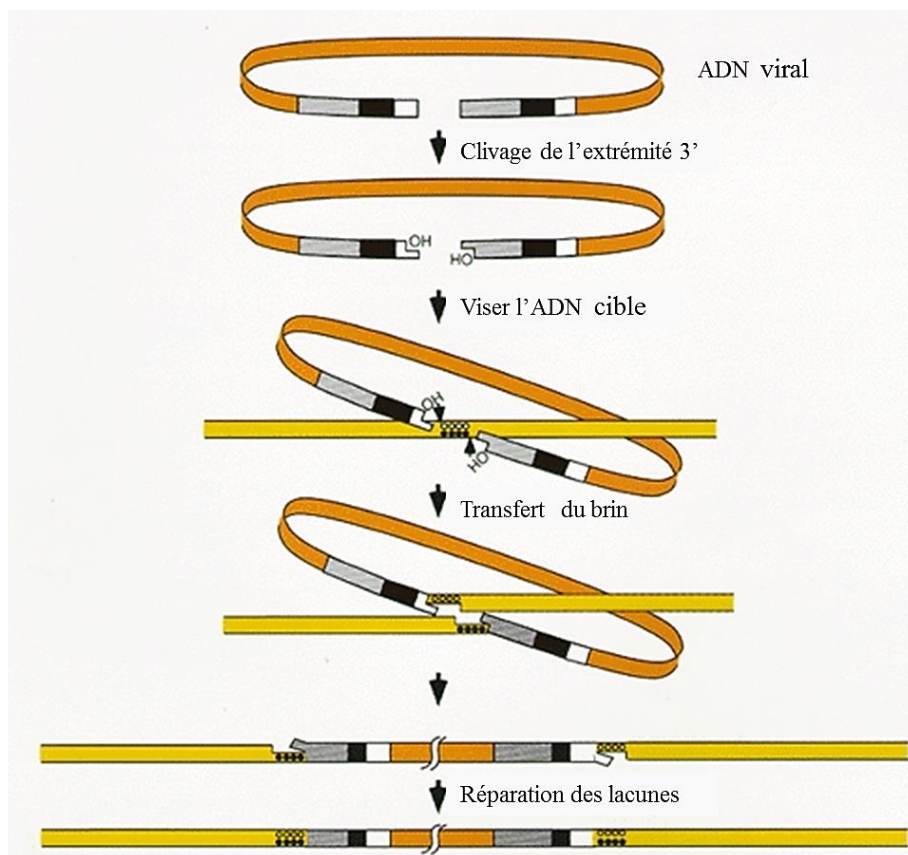


Figure 2-8: Schéma résumant les principales étapes de l'intégration de l'ADN rétroviral dans le génome [106]

### 2.2.2. Les phases tardives

Une fois intégré, l'ADN proviral peut ensuite être transcrit par la machinerie transcriptionnelle de la cellule hôte. La transcription est initiée par l'ARN polymérase de type II à partir du LTR 5' et se termine dans le LTR 3'. Deux transcrits majeurs sont alors observés, le premier est un transcrit de longueur totale (non épissé) correspondant à l'ARN génomique et le second, épissé pour *gag* et *pol*, est dit subgénomique. C'est à partir de ce

dernier que sera synthétisée la polyprotéine Env (Figure 2-5). Les ARN viraux sont ensuite transportés vers le cytoplasme où ils seront traduits. La traduction du gène *gag* s'effectue à partir du messageur de longueur totale. Le précurseur Gag-Pol ou Gag-Pro-Pol permet la synthèse de la polyprotéine Pol. La traduction du gène *pol* s'effectue grâce à un « frame-shift » (décalage de la phase de lecture), par suppression de la terminaison de traduction, ou bien encore *via* un messageur épissé spécifique de *pol* comme pour les spumavirus [107]. Tous ces mécanismes conduisent à l'obtention des polyprotéines Gag et Pol pour permettre un assemblage correct des particules. Les particules virales peuvent être formées par bourgeonnement au niveau de la membrane plasmique ou dans des structures internes de type endosomes [108]. La particule sort de la cellule par bourgeonnement en se recouvrant de la membrane cellulaire, probablement par l'intermédiaire de la reconnaissance des sous-unités MA de la protéine Gag et TM de la protéine Env. Les précurseurs Gag, Gag-Pro-Pol et Gag-Pol sont ensuite clivés pour générer les différentes sous-unités matures et fonctionnelles et former ainsi une particule virale infectieuse (Figure 2-9) [70].

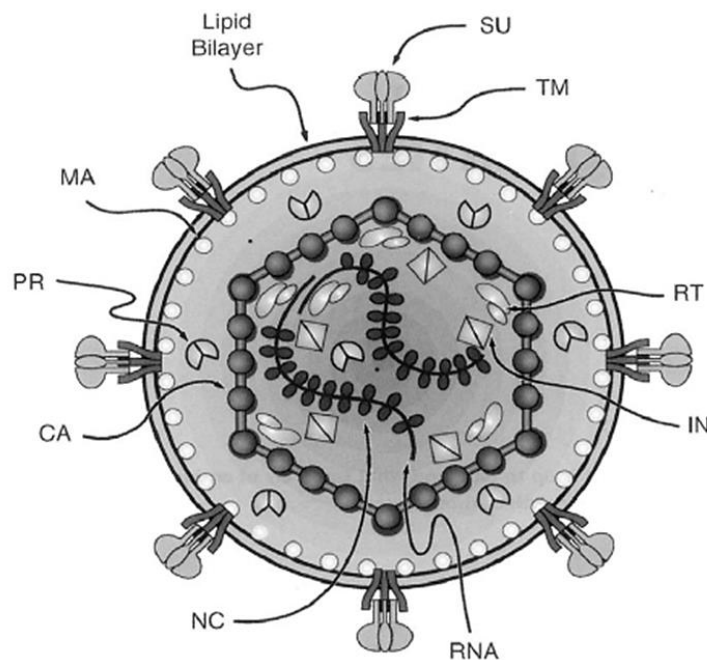


Figure 2-9: Structure d'une particule virale mature, enveloppée et infectieuse. L'enveloppe virale est formée d'une bicouche lipidique d'origine cellulaire et dans laquelle viennent s'insérer les protéines Env. Les sous-unités transmembranaires (TM) et de surface (SU) sont reliées par des ponts disulfures. Les protéines codées par le gène *gag*, la matrice (MA), la capsid (CA) et la nucléocapsid (NC), sont localisées à l'intérieur de cette particule. La protéase (PR), la rétrotranscriptase (RT) et l'intégrase (IN), codées par le gène *pol*, sont également incorporés dans la particule [63].

## **2.3. L'adénocarcinome pulmonaire ovin et le Jaagsiekte Sheep Rétrovirus**

### **2.3.1. Historique**

Le Jaagsiekte Sheep Rétrovirus (JSRV) est l'agent causal de l'adénocarcinome pulmonaire ovin (OPA) également connu sous le nom d'adénomatose pulmonaire ovine, ou bien encore de carcinome pulmonaire ovine. Il est unique parmi les rétrovirus car il est le seul connu à ce jour comme capable de provoquer un cancer des poumons contagieux chez les moutons [109, 110]. "Jaagsiekte" est le nom donné à la maladie en Afrique de Sud et dérive de mot "jaag" qui signifie chassant et "Siekte" pour maladie ; les moutons malades présentant des signes de détresse respiratoire [111]. L'OPA a été décrite la première fois en Afrique de sud au début du 19<sup>ème</sup> siècle [112, 113]. Les premiers rapports sur cette maladie décrivent que l'agent infectieux de l'OPA était un virus à ARN semblable à d'autres rétrovirus de type B et de type D [114, 115]. Des études suivantes ont révélé que JSRV est exclusivement et systématiquement présent dans les tissus tumoraux et les sécrétions pulmonaires des animaux infectés, et qu'expérimentalement ce virus induit un OPA chez le mouton et la chèvre [112, 116]. Depuis, cette maladie s'est propagée à travers le monde, en Europe, en Afrique, en Asie et en Amérique et affecte aussi bien les moutons domestiques que les moutons sauvages [113, 117, 118], plus rarement les chèvres sous forme subclinique mais pas, jusqu'à présent, les bovins ou les autres animaux [119]. Cette maladie a été éradiquée en Islande après son introduction en 1933 par des béliers Russian Karakul importés d'Allemagne qui ont introduit non seulement l'OPA, mais aussi le Maedi-Visna Virus (MVV) et la Paratuberculosis qui ont entraîné la mort de plus de 60% des troupeaux de l'île [113]. Cette maladie est endémique au Pérou, en Europe et en Afrique du Sud alors qu'elle est absente de la large population de moutons d'Australie et de Nouvelle-Zélande.

### **2.3.2. Adénocarcinome pulmonaire.**

L'OPA est une tumeur pulmonaire contagieuse des moutons et, plus rarement des chèvres. C'est la tumeur pulmonaire la plus fréquente chez les moutons et on la retrouve dans de nombreux pays à travers le monde [120]. Cette maladie figurait sur l'ancienne liste B de l'OIE (Organisation mondiale de la santé animale) contenant les maladies transmissibles qui sont considérées comme ayant un impact significatif en termes socio-économique, de santé publique et/ou sur le commerce international d'animaux et de leurs produits dérivés. En effet, l'impact économique de l'OPA peut être très significatif : 20-25% du troupeau peut être perdu suite à une première infection au virus avec des pertes annuelles de 1 à 5%. Éradiquer cette

maladie d'un troupeau est difficile à cause de l'absence d'un test de diagnostic pouvant détecter le JSRV à des stades précliniques [121]. Dans des conditions naturelles, la plupart des cas d'OPA apparaissent chez les animaux âgés de 1 à 4 ans, bien que tous les âges soient susceptibles dans des contextes expérimentaux. La période d'incubation varie entre 6 à 8 mois dans les cas naturels et de 5 à 12 mois chez les moutons adultes et les agneaux infectés expérimentalement. L'apparition de signes cliniques se produit très rapidement lorsque les agneaux nouveaux nés sont infectés expérimentalement (de 4-6 jours jusqu'à 3-6 semaines) [115, 122, 123]. Des données récentes montrent que le colostrum, le lait et les sécrétions pulmonaires provenant d'animaux infectés sont des réservoirs majeurs de JSRV et représentent, en plus de la transmission intra-utérine de la mère au fœtus [119, 124], des voies majeures permettant la transmission de JSRV à des animaux non infectés [122, 125].

Les symptômes cliniques apparaissent une fois que les lésions sont suffisamment importantes pour interférer avec la fonction pulmonaire. Tout d'abord, les moutons semblent moins actifs, ont l'air épuisé, sont fébriles et perdent du poids. Ceci est suivi par une tachypnée, des mouvements intenses de la paroi abdominale et de l'accumulation de fluides dans les voies aériennes respiratoires. Ce fluide séro-muqueux est déversé à travers les narines et peut provoquer une toux spasmodique. Les animaux finiront par mourir après des durées variables en fonction de la présence d'infections pulmonaires secondaires [119]. Chez les chèvres, JSRV est aussi capable d'induire expérimentalement un cancer de poumon mais la réplication virale de ce virus dans les cellules de chèvres est beaucoup plus restreinte que chez le mouton [126]. De manière remarquable, les animaux infectés par JSRV ne développent pas d'anticorps contre le virus [127], très probablement en raison de la présence de rétrovirus endogènes très similaires à JSRV, appelés enJSRVs qui sont insérés dans le génome et jouent un rôle fondamental *in vivo* dans le processus de placentation chez le mouton [128]. Chez les animaux naturellement infectés, JSRV est principalement détecté dans les leucocytes et les organes lymphoïdes périphériques plutôt que dans les poumons, et seule une minorité de moutons infectés par JSRV développent une OPA [124].

L'OPA peut se manifester sous deux formes pathologiques : classique et atypique. Dans l'OPA classique, les tumeurs apparaissent principalement dans les parties cranio-ventrales de tous les lobes pulmonaires comme des lésions diffuses ou nodulaires. Elles sont habituellement multiples et ont tendance à se regrouper, ce qui donne une apparence glandulaire diffuse. Les métastases sont rares mais parfois peuvent se produire dans les ganglions lymphatiques régionaux et les organes extrathoraciques [119]. Dans la forme atypique, elles ont tendance à être plus nodulaires et localisées principalement dans le lobe

diaphragmatique. Elles sont de consistance très dures et bien délimitées du tissu environnant. Les deux formes peuvent être présentes au sein d'un même troupeau ou d'un seul individuel sans qu'il y ait une différence au niveau moléculaire en termes de souche virale de JSRV [119]. Les cellules tumorales dérivent des pneumocytes de type II ou des cellules Club (anciennement cellules de Clara) (Figure 2-10), compte tenu de leur ultrastructure et de l'expression de marqueurs spécifiques tels que la protéine A du surfactant (SP-A) et C (SP-C) pour les pneumocytes de type II et la protéine appelée CC-10 spécifique des cellules Club [118].

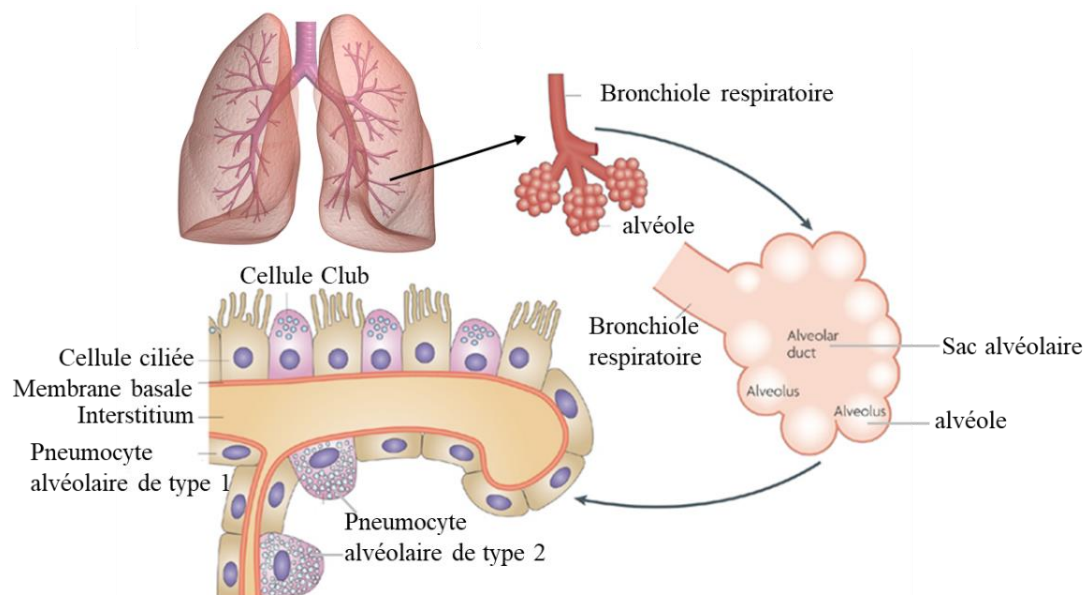


Figure 2-10: Schéma représentant les types cellulaires présents au niveau des bronchioles et des alvéoles. [129]

Les deux types de cellules ont des propriétés sécrétoires ce qui explique la grande quantité importante de sécrétions qui s'accumulent dans les poumons des moutons les plus touchés par la maladie [130, 131]. Histologiquement, les lésions sont caractérisées par la prolifération néoplasique des cellules épithéliales des parois alvéolaires et bronchiolaires formant des foyers qui compriment les tissus environnants. L'OPA ressemble au carcinome broncho-alvéolaire humain et constitue un modèle pour le cancer des poumons chez les humains [109, 132, 133]. Une étude récemment publiée, a montré que JSRV peut infecter les cellules humaines mais ne semble jouer aucun rôle dans l'induction de cancer du poumon chez l'homme [134].

### 2.3.3. Le rétrovirus exogène : Jaagsiekte Sheep Retrovirus

JSRV est membre du genre au sein de la famille des *Retroviridae* et de la sous-famille *Orthoretroviranae*. JSRV est phylogénétiquement lié à M-PMV et MMTV et est considéré comme un rétrovirus de type D. JSRV a l'organisation génomique classique d'un rétrovirus simple avec les gènes *gag*, *pol* et *env* flanqués de chaque côté par les longues régions terminales répétées en 5' et 3' (LTR) (Figure 2-11) [135]. Dans le génome de JSRV, il y a un cadre de lecture ouvert supplémentaire appelé *orf-X* dont le rôle et la fonction reste encore inconnu et qui chevauche l'extrémité 3' du gène *pol* [109, 136]. Ce cadre de lecture est conservé et intact pour toutes les souches de JSRV identifiées à ce jour suggérant un rôle essentiel dans la réplication de ce virus *in vivo*. D'autre part, une étude *in vitro* a montré que *orf-X* n'est pas nécessaire pour le processus de transformation cellulaire [137, 138]. A noter que le gène *pro* des Betarétrovirus code pour une deuxième enzyme (autre que la protéase), la déoxyuridine triphosphatase (dUTPase). Cette enzyme est essentielle pour la réplication de JSRV dans les monocytes/macrophages des organes lymphoïdes ainsi que dans les cellules Club et les pneumocytes de type II chez le mouton [113].

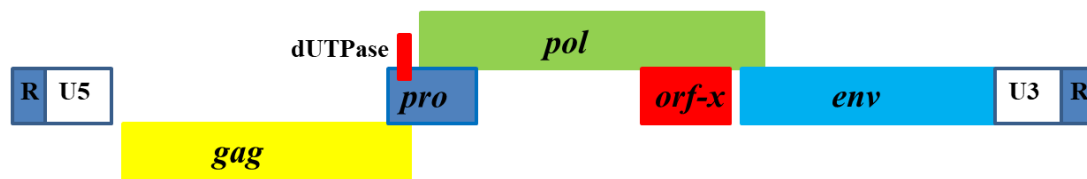


Figure 2-11: Organisation génomique de l'ARN viral de JSRV

#### 2.3.3.1. Tumorigénèse induite par JSRV

JSRV induit une transformation cellulaire par un mécanisme différent de la grande majorité des rétrovirus oncogènes puisque c'est sa glycoprotéine Env qui agit comme une oncoprotéine dominante [138-142]. L'Env de JSRV est une protéine transmembranaire de type I ayant environ 620 acides aminés et composée d'un domaine de surface (SU) et d'un domaine transmembranaire (TM) liés par des liaisons de disulfure (Figure 2-12) [109, 136, 143]. Cette protéine virale est présente à la surface des particules virales et est indispensable au processus d'entrée virale de JSRV dans les cellules en interagissant avec son récepteur cellulaire (Hyal2) [144]. D'autre part, l'expression de l'Env de JSRV est nécessaire et suffisante pour induire la transformation d'une variété de lignées cellulaires *in vitro* [138, 145-147]. De même, des études chez la souris et le mouton ont confirmé que la glycoprotéine Env est capable d'induire



seule un adénocarcinome pulmonaire *in vivo* [141, 142, 148]. Ainsi, à la fois *in vitro* et *in vivo*, le JSRV Env fonctionne comme une oncoprotéine puissante.

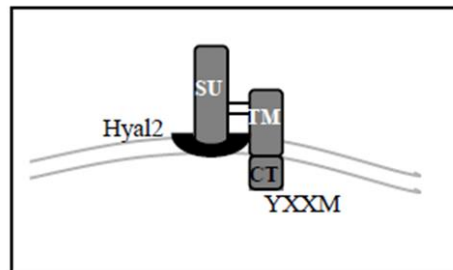


Figure 2-12: Modèle hypothétique de l'interaction entre l'Env de JSRV avec son récepteur la hyaluronidase 2 (Hyal2) à la membrane plasmique. L'Env de JSRV est composé de deux sous-unités : une sous-unité de surface (SU) et une sous-unité transmembranaire (TM). La queue cytoplasmique (CT) du domaine TM contient un motif YXXM, un site putatif de liaison PI3K / p85 [149].

Les mécanismes de la transformation cellulaire induite par l'Env de JSRV ne sont pas complètement compris [123] et des mécanismes récepteurs-indépendants et récepteurs-dépendants ont été proposés [150]. En effet, plusieurs voies de signalisation cellulaire ont été décrites comme étant impliquées dans la transformation par l'Env JSRV *in vitro* (Figure 2-13) :

1. La voie de de PI3K-Akt (phosphatidylinositol 3-kinase - Protein Kinase B), bien connue dans la prolifération et la survie cellulaire, la transformation oncogénique et le développement du cancer [151, 152].
2. La voie de MEK/ERK (MAPK: MEK/ Mitogen- activated protein kinase). Cette voie est impliquée dans la croissance cellulaire, la différenciation, la survie, et la transformation.
3. De multiples autres mécanismes ou voies ont été également décrits comme potentiellement impliqués dans la transformation des cellules par Env tel que la voie de signalisation Hyal2-RON [153], TLR 4 (Toll-Like receptor 4) [150] ou bien encore Src [154].

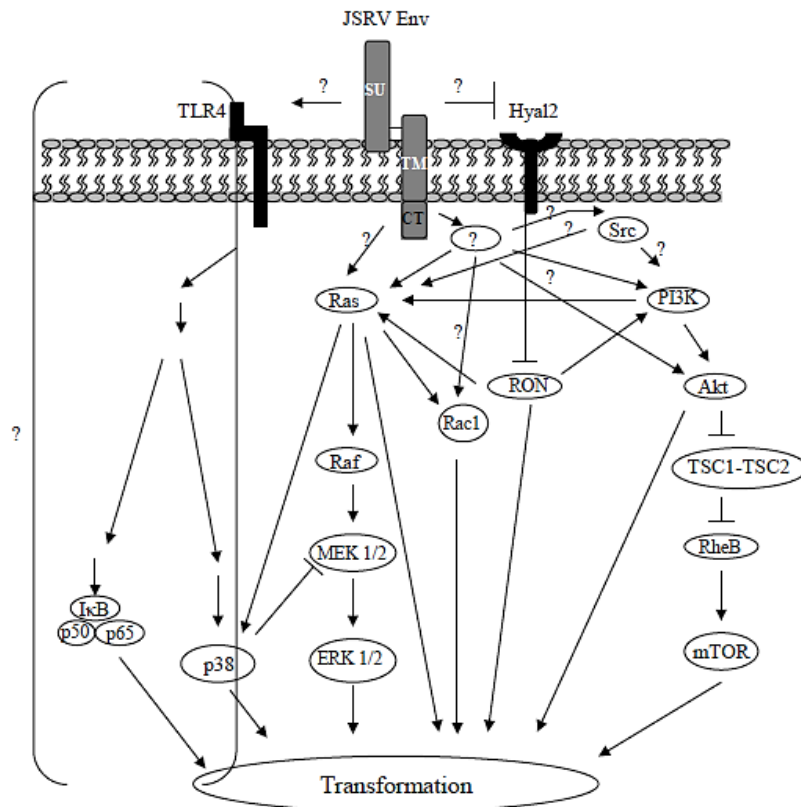


Figure 2-13: Voies de signalisation impliquées dans la transformation par l'Env du JSRV [150] Selon le type de lignée cellulaire étudiée, plusieurs voies de signalisation ont été décrites comme étant impliquées dans des cellules transformées par JSRV; Les voies PI3K/Akt, Ras-MEK-ERK et Hyal2-RON. La signalisation à travers la voie TLR4-NFκB est également représentée car l'Env de JSRV interagit avec TLR4 [150] même si l'activation de cette voie par Env n'a pas encore été décrite. Une quatrième voie; la signalisation impliquant Rac1, a été récemment identifiée, mais on ne sait encore que peu de chose à ce sujet [155] À l'exception de la voie Hyal2-RON, les mécanismes par lesquels l'Env de JSRV engage les voies Ras-MEK-ERK et PI3K/Akt ne sont pas connus. Les interactions possibles, ou encore les facteurs identifiés qui pourraient être impliqués dans la transformation médiée par l'Env sont indiqués par des points d'interrogation (?).

### 2.3.3.2. Tropisme de JSRV

La plupart des rétrovirus, à l'exception des lentivirus, nécessitent une réplication active des cellules afin de permettre au complexe de préintégration virale de passer à travers la membrane nucléaire [106]. Les cellules transformées dans le cas de l'OPA sont dérivées des cellules épithéliales différenciées des poumons, les pneumocytes de type II et les cellules Club, qui ont un indice répliatif très faible chez les animaux adultes et qui, par conséquent, ne sont pas facilement infectées par JSRV [142]. Cependant, ces cellules se divisent plus rapidement chez les jeunes agneaux ou chez les adultes pour réparer des lésions de l'épithélium pulmonaire. Dans les deux cas, ces cellules en mitose seront plus permissives à l'infection par JSRV [142] et donc plus susceptibles à la transformation et à l'apparition d'OPA [156]. Des études ont d'ailleurs été proposées que des cellules souches broncho-alvéolaires (BASC) pourraient être la cible réelle de la réplication du JSRV dans les poumons

chez le mouton [157, 158]. En général, Env et les LTR sont les principaux déterminants du tropisme rétroviral (respectivement du tropisme d'entrée et d'expression). Comme mentionné précédemment, l'entrée des rétrovirus dans les cellules s'effectue par la reconnaissance entre Env et un récepteur situé à la surface de la membrane plasmique de la cellule. Ces récepteurs cellulaires déterminent ainsi le tropisme cellulaire et la nature de l'hôte pouvant être infecté par un virus donné. JSRV et sa forme endogène enJSRV utilisent tous les deux le même récepteur cellulaire, HyaL2 (Hyaluronoglucosaminidase 2) [147, 159, 160]. HyaL2 dégrade les acides hyaluroniques qui sont des composants de type glycosaminoglycane essentiels au maintien de l'intégrité de la matrice extracellulaire [161]. HyaL2 est une protéine ancrée à la membrane plasmique grâce à un glycosylphosphatidylinositol (GPI) et ayant une faible activité hyaluronidase. HyaL2 est le seul membre de la famille des hyaluronidases que JSRV utilise pour l'entrée virale [160] ce qui suggère que l'activité hyaluronidase *per se* n'est pas requise pour la fonction du récepteur viral et l'entrée du virus dans les cellules [162, 163]. La protéine HyaL2 est exprimée ubiquitairement à la surface de toutes les cellules ovines et caprines et, par conséquent, n'est donc pas à l'origine du tropisme restreint de JSRV pour certaines cellules du mouton. En revanche, l'expression de JSRV est probablement préférentielle dans les pneumocytes de type II et les cellules Club du fait de l'expression dans ces cellules du facteur de transcription HNF-3 $\beta$  (Hepatocyte Nuclear Factor 3) [164, 165]. En effet, le LTR 5' de JSRV possède deux sites de fixation pour ce facteur (HNF-3 beta) dans sa région régulatrice du promoteur viral. De plus, il contient également d'autres facteurs de transcription tel que le facteur NF-1 (Nuclear Factor 1) et le facteur SP-1 (Specific Protein 1) [164]

## 2.4. Les rétrovirus endogènes de type JSRV

### 2.4.1. Interactions des enJSRVs avec le virus exogène JSRV et leur hôte

Le mouton est un modèle unique pour étudier l'interaction entre les rétrovirus et leur hôte. En effet, JSRV coexiste avec des rétrovirus endogènes très similaires, enJSRVs, qui sont insérés dans le génome du mouton. Ces enJSRVs présentent 85-89% d'identité de séquence nucléotidique avec JSRV. Les principales différences se situent au niveau de la région U3 des LTRs ainsi que dans trois régions dites variables et appelées VR1, 2 et 3 (VR1-2-3) qui sont situées dans les protéines Gag et Env [166-168]. VR1 et VR2 résident dans la MA de la protéine Gag et sont distantes de 50 résidus tandis que VR3 est localisée dans Env et comprend les 67 derniers acides aminés de la région TM. Une étape essentielle du cycle de

réplication des rétrovirus est l'intégration de leur génome viral dans l'ADN génomique de leur hôte. Au cours de l'évolution, des rétrovirus exogènes transmis horizontalement peuvent occasionnellement infecter la lignée germinale de leur hôte et devenir ainsi des rétrovirus endogènes qui se transmettent verticalement de génération en génération suivant les règles de l'hérédité mendélienne. Faisant partie intégrante du génome, certaines fonctions de rétrovirus endogènes ont été sélectionnées et conservés au cours de l'évolution, jouant ainsi un rôle important pour leurs hôtes. Les rétrovirus endogènes enJSRVs et le mouton sont un très bon exemple de cette association symbiotique. En effet, ils sont indispensables au bon déroulement de la grossesse chez le mouton [79] et jouent ou auraient joué un rôle de protection contre les infections par le virus pathogène JSRV par interférence avec l'étape précoce d'entrée virale de JSRV par compétition de leur Env avec celle de JSRV pour l'interaction avec le récepteur Hyal2 [128] ou en bloquant les phases tardives de son cycle de réplication [169]. Concernant ce dernier, un rétrovirus endogène appelé enJS56A1 agit comme un facteur de restriction virale capable de bloquer les phases tardives du cycle de réplication de JSRV [81]. Ce rétrovirus endogène est défectif et incapable de produire des particules virales en culture de cellules. De plus, lorsqu'il est co-exprimé avec le rétrovirus exogène JSRV, il agit comme un dominant négatif et empêche ce virus pathogène d'exporter et de produire des particules virales. Un résidu Tryptophane en position 21 (W21) de la protéine Gag d'enJS56A1 est responsable du blocage des particules virales de JSRV dans le cytoplasme (JSRV ayant un résidu Arginine en position 21). Deux études ont décrit le mécanisme de restriction virale induit par enJS56A1 sur JSRV [61, 170]. Tout d'abord, il a été montré que les protéines Gag d'enJS56A1 et de JSRV sont co-localisées et interagissent dans des cellules exprimant ces deux protéines. Une étape importante du cycle de réplication des Betarétrovirus et de JSRV est le transport des protéines Gag au Centre Organisateur des Microtubules (COMT) proche du noyau de la cellule où l'assemblage cytoplasmique des particules virales a lieu. Cependant, et contrairement aux protéines Gag de JSRV, celles d'enJS56A1 ne s'accumulent pas au COMT et conservent une localisation dispersée dans la cellule au cours du temps. En présence d'enJS56A1, les protéines Gag de JSRV ne sont plus capables de se localiser au COMT et s'accumulent sous forme d'amas dispersés dans le cytoplasme. Ces résultats indiquent que la protéine Gag d'enJS56A1 agit comme un dominant négatif qui bloque le trafic intracellulaire « normal » des protéines Gag de JSRV et induit leur dégradation intracellulaire. En effet, les protéines Gag d'enJS56A1 (du fait du W21) se comportent donc comme des protéines non structurées ou mal formées qui s'accumulent dans la cellule et sont partiellement dégradées par le protéasome.

## 2.4.2. Coadaptation entre enJSRVs/JSRV et le mouton au cours de l'évolution

Afin d'identifier et de caractériser les enJSRVs présents au sein du génome du mouton, une banque d'ADN génomique a été criblée par PCR et hybridation pour la présence du gène *env* [68]. Suite à ce crible, 27 provirus enJSRVs ont été identifiés par séquençage grâce à leur site unique d'intégration dans le génome et leurs différentes structures génomiques (Figure 2-14).

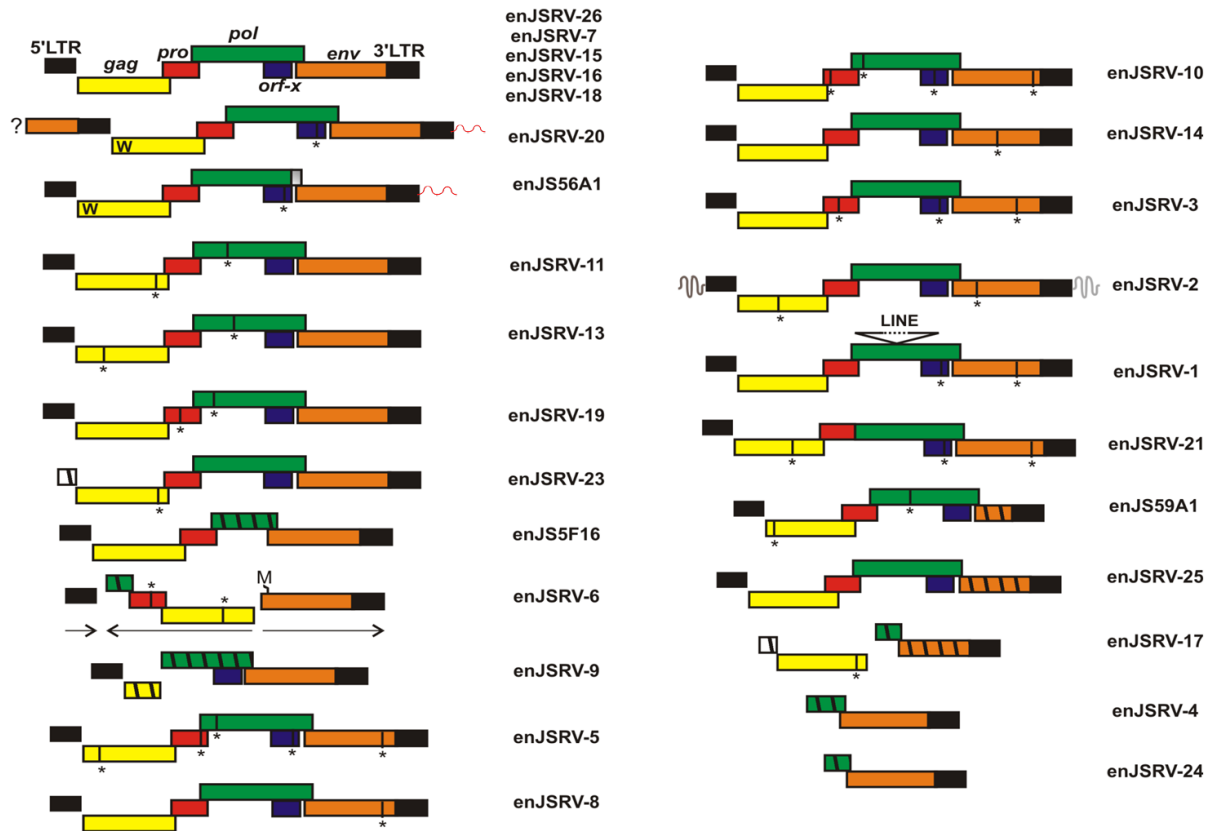


Figure 2-14: Organisation génomique des provirus enJSRV. Cinq provirus ont une organisation génomique typique du JSRV exogène (en haut). Le " W " dans le cadre de lecture gag d'enJS56A1 et enJSRV-20 indique la substitution R21W présente au niveau de ces deux provirus transdominants. Toutes les séquences génomiques flanquant les provirus enJSRV contiennent une duplication de six paires de bases qui est la caractéristique de l'intégration rétrovirale. Les seules exceptions sont représentées par enJSRV-20, qui contient une partie d'un gène *env* (indiquée par une boîte orange et un point d'interrogation) avant le 5' LTR et enJSRV-2, qui ne contient pas les mêmes séquences flanquant les LTR. Le locus enJSRV-6 possède une méthionine supplémentaire (M) dans Env en plus du codon de départ canonique présent dans JSRV et d'autres loci enJSRV. De plus, dans enJSRV-6, gag et pro sont en sens inverse par rapport aux LTR 5' et 3' et env (indiqué par des flèches horizontales). EnJSRV-1 présente un élément LINE dans la région de codage *pol*. Les codons stop prématurés sont représentés par une ligne verticale et un astérisque (\*). Les grandes délétions dans les génomes proviraux sont indiquées par des boîtes hachurées. [68]

La majorité de ces enJSRVs possède dans leurs génomes des mutations non-sens, des délétions ou recombinaisons accumulées au cours de l'évolution. Cependant, cinq enJSRVs possèdent une structure génomique intacte ce qui les rend potentiellement compétents pour réaliser un cycle de réplication viral complet. De plus, deux rétrovirus endogènes, enJSRV-20 et enJS56A1 possèdent le résidu tryptophane en position 21 (W21) de leurs protéines Gag qui,

comme décrit précédemment, leur permet de bloquer les étapes tardives du cycle de réplication de JSRV [61, 81, 170].

Les *Caprinae* forment une sous-famille des bovidés qui comprend notamment les genres *Capra* (chèvres) et *Ovis* (moutons). Le genre *Ovis* est constitué du mouton domestique *Ovis aries* et d'ovins dit sauvages tels que *Ovis ammon*, *Ovis canadensis* ou bien encore *Ovis dalli*. La connaissance des dates approximatives de spéciation de chaque espèce des *Caprinae* a permis d'estimer la date d'insertion des enJSRVs dans le génome en criblant spécifiquement par PCR la présence ou non de chaque enJSRVs dans l'ADN génomique de ces différentes espèces [68]. Cette étude montre que les premières intégrations de ces rétrovirus dans la lignée germinale de leur hôte ont eu lieu il y a cinq à sept millions d'année, avant la séparation des groupes caprins et ovins, ont continué au cours de l'évolution du mouton et sont peut-être même encore en cours de nos jours (Figure 2-15). En effet, certains enJSRVs sont présents uniquement chez les ovins, soient dans le génome de toutes les espèces du genre *Ovis* testées (ovins domestiques et sauvages) comme enJS56A1, soient restreints aux moutons domestiques (*Ovis aries*). De plus, les cinq enJSRVs ayant une structure génomique intacte sont intégrés uniquement dans le génome des moutons domestiques et de manière polymorphique, c'est à dire, qu'ils sont présents seulement chez certains moutons domestiques. Cette observation suggère une intégration dans la lignée germinale et une « endogénisation » de ces rétrovirus très récente au cours de la domestication du mouton.

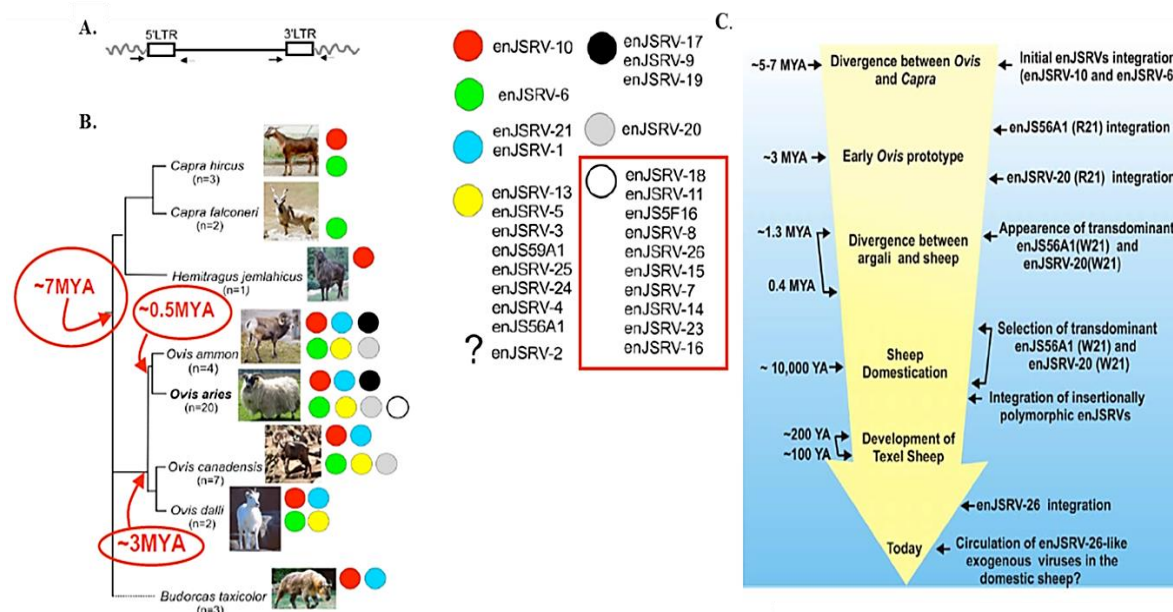


Figure 2-15 : Histoire évolutive des enJSRVs et leur hôte [68]. A. Représentation schématique des PCRs utilisées pour amplifier spécifiquement chaque provirus enJSRV. B. Arbre phylogénétique simplifié de la sous-famille des *Caprinae* avec une estimation des dates de spéciation de chaque espèce. Les cercles rouges indiquent les dates prédictives des événements d'intégration des enJSRVs dans le génome et ayant eu lieu avant les différentes spéciations. C. Diagramme illustrant les événements clés de l'histoire évolutive des enJSRVs avec des dates estimées au cours de l'évolution du mouton domestique et des *Caprinae*.

Les rétrovirus endogènes, enJS56A1 et enJSRV-20, agissant comme facteur de restriction contre JSRV sont présents chez les ovins sauvages et domestiques. De manière intéressante, un résidu Arginine est présent (comme chez tous les rétrovirus compétents pour le cycle de réplication viral) au sein de la protéine Gag de ces deux rétrovirus endogènes dans les espèces dites « sauvages » tel qu'*Ovis canadensis* et *Ovis dalli*. En revanche, un acide aminé Tryptophane (W21) est présent dans la protéine Gag de ces deux rétrovirus endogènes chez tous les moutons domestiques issus des différentes races testées. De plus, il a été montré que cet endogène protecteur enJS56A1 a été amplifié au sein d'une même région chromosomique en plusieurs copies, résultant aujourd'hui à un nombre de copies enJS56A1-like variable dans le génome des moutons domestiques de nos jours mais en tous plus important que dans celui des moutons sauvages [171]. Cette « fixation » et cette amplification des enJSRVs ayant un rôle protecteur contre le virus exogène et pathogène JSRV au sein du génome des moutons domestiques suggèrent que la mutation faux-sens changeant une arginine en tryptophane dans la protéine Gag a été sélectionnée positivement au cours de l'évolution et de la domestication du mouton. En effet, le simple processus de domestication et la proximité des animaux ont certainement pu favoriser les infections virales de type JSRV et induire la sélection d'animaux plus résistants ayant ces rétrovirus endogènes « protecteurs ». D'autre part, deux autres rétrovirus endogènes ayant une structure génomique intacte sont présents dans le génome

d'une seule race de mouton domestique (Texel) et seulement chez certains animaux de cette race. Par conséquent, l'endogénéisation de ces rétrovirus a eu lieu après l'établissement de cette race (il y a moins de 200 ans). De manière intéressante, l'un d'entre eux, enJSRV-26 n'a pu être détecté que chez un seul animal (celui à l'origine de la banque d'ADN génomique) sur 157 animaux de la même race testés (Texel) incluant trois de la même descendance que l'animal possédant enJSRV-26 et 62 animaux provenant de 18 races différentes de mouton. *In vitro*, en culture de cellules, les cinq rétrovirus endogènes ayant une structure génomique intacte produisent et exportent en dehors de la cellule des particules virales avec une protéine d'enveloppe capable d'utiliser le récepteur membranaire nécessaire à l'entrée du virus dans la cellule. Lorsqu'enJS56A1 ou enJSRV-20 sont co-transfectés avec ces cinq rétrovirus endogènes, quatre sur les cinq sont incapables d'exporter leurs particules virales hors de la cellule. Cependant un de ces rétrovirus, enJSRV-26, est capable d'échapper au mécanisme de restriction virale et de produire des particules virales. Ces résultats et cette étude montrent que le plus récent de ces rétrovirus endogènes, enJSRV-26, a évolué afin d'échapper au mécanisme d'interférence induit par d'autres rétrovirus endogènes du génome (enJS56A1 et enJSRV-20) [68]. Une étude suivante révélera le mécanisme moléculaire de l'échappement d'enJSRV-26 à la restriction induite par enJS56A1 qui est dû à la substitution d'un seul acide aminé dans le peptide signal de l'Env d'enJSRV-26 [171].

#### **2.4.3. L'histoire de la domestication du mouton révélée par les rétrovirus endogènes enJSRVs.**

Comme mentionné dans le chapitre 1, la domestication des animaux fut une étape clé de l'humanité et dans l'établissement des fondements de notre société actuelle. Néanmoins, beaucoup d'incertitudes subsistent encore sur l'origine de domestication et les vagues de migrations successives. Comme décrit précédemment, cinq enJSRVs se sont insérés dans le génome du mouton domestique (*Ovis aries*) que très récemment à l'échelle de l'évolution. En effet, la présence de ces rétrovirus endogènes est variable dans la population c'est-à-dire qu'ils sont présents dans le génome de certains animaux et non dans d'autres. Ces rétrovirus endogènes constituent par conséquent des marqueurs génétiques très puissants pour établir de manière claire des liens génétiques entre différents animaux et révéler l'histoire et les événements qui ont conduit au mouton domestique de nos jours. Un travail mené par Bernardo Chessa et ses collègues a révélé, grâce à l'utilisation de ces enJSRVs comme marqueurs génétiques, certaines des étapes clés de la domestication du mouton et des migrations consécutives en Eurasie et Afrique [172] (Figure 2-16).



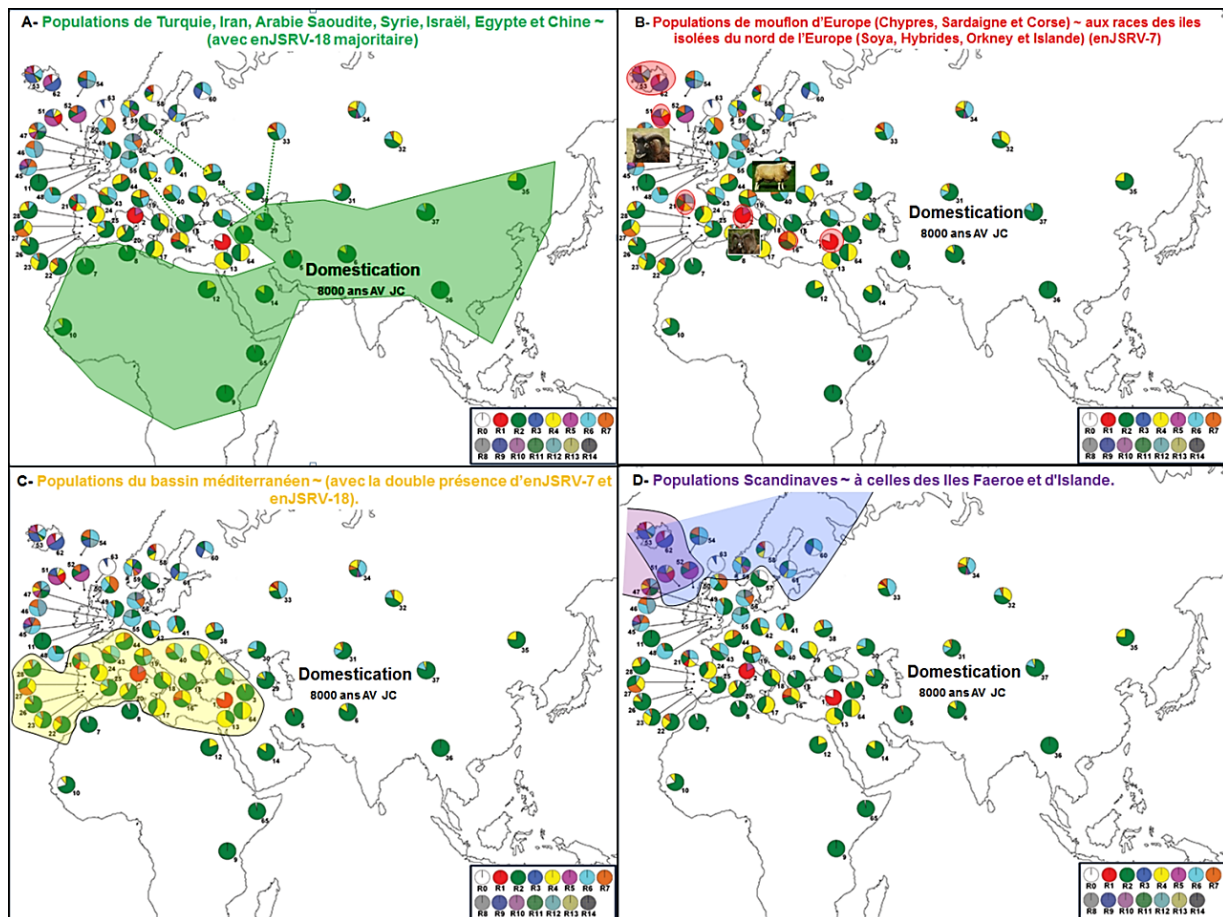


Figure 2-16 La domestication et les vagues successives de migration des moutons selon les rétrovirus endogènes [172]. A. Races de moutons domestiques dites modernes (issues de la deuxième vague de migration) qui sont caractérisées par la présence d'enJSRV-18 (Rétrotype R2). B. Races de mouton dites primitives (issues de la première vague de migration) qui sont caractérisées par la présence d'enJSRV-7 (Rétrotype R1). Notez qu'enJSRV-7 est plus ancien dans le génome du mouton qu'enJSRV-18. Il s'agit ici de populations de mouflon (premiers moutons domestiqués puis retournés à la vie sauvage) qui sont retrouvés dans des îles isolées en Europe du Sud et du Nord. C. les populations du bassin méditerranéen se caractérisent par la double présence d'enJSRV-7 et d'enJSRV-18 (Rétrotype R4). D- Les populations Scandinaves présentent un rétrotype comparable à celles des Îles Faeroe et d'Islande (Rétrotypes R5 et R9) caractérisés par la présence d'enJSRV-22 et R0.

Pour cette étude, une collection de 1362 ADN génomiques provenant d'animaux de 133 races de mouton domestique et d'espèces sauvages proches a été établie. Les races et espèces sélectionnées sont pour la grande majorité d'origine géographique connue et non soumises aux programmes de croisements commerciaux intensifs. Chaque enJSRVs est intégré à un site particulier du génome, permettant son criblage spécifique par PCR. Chaque ADN génomique a donc été testé pour la présence ou non de ces cinq enJSRVs. L'absence ou la présence d'un ou plusieurs des cinq rétrovirus endogènes ainsi que leur fréquence au sein de chaque population a permis l'établissement de « rétrotype » permettant une analyse statistique et globale des données. Deux sur les cinq enJSRVs testés sont particulièrement informatifs. Le rétrovirus endogène, enJSRV-18, est celui retrouvé le plus fréquemment dans le génome

des animaux testés (85%). Malgré sa large distribution au sein des populations, cet enJSRV est absent chez le mouflon d'Asie, d'Europe ainsi que chez les animaux du groupe *Ovis vignei* (mouflon vivant de l'Iran à l'Inde occidentale et séparé du mouton domestique il y a 800000 ans). Un seul rétrovirus endogène, enJSRV-7, est retrouvé majoritairement dans le génome d'*Ovis vignei*, du mouflon d'Asie et d'Europe. Ce rétrovirus endogène est par conséquent le plus ancien de ces rétrovirus endogènes étant apparu dans le génome avant le processus de domestication. Ainsi, cette étude montre que les populations méditerranéennes de mouflon d'Europe (Chypres, Sardaigne et Corse) et les races de mouton issues des îles isolées du nord de l'Europe (Soya, Hybrides, Orkney et Islande) ont des rétrotypes très similaires (par la présence d'enJSRV-7) et nettement séparés des autres races de moutons retrouvés sur le continent européen (dans lesquelles enJSRV-18 est majoritaire) (Figure 2-16A et B). Ces populations représentent donc des reliques des premières migrations de mouton domestique encore présent chez les mouflons de ces îles de la Méditerranée et les populations de mouton domestique à la périphérie nord de l'Europe vivant dans des îles isolées. Ces races de mouton ont de plus des traits morphologiques d'espèces « primitives » par rapport à celle dites « modernes » retrouvées en Europe, Afrique, Asie et représentent probablement des populations de mouton domestique retournées à la vie sauvage. Les populations de mouton de Turquie, Iran, Arabie Saoudite, Syrie, Israël, Égypte et Chine présentent toutes un rétrotype similaire (avec une présence majoritaire d'enJSRV-18) en accordance avec les données archéologiques supportant l'hypothèse d'un centre de domestication unique au Moyen-Orient et d'un lien direct des migrations de population entre les régions du Moyen Orient, d'Afrique et d'Asie (Figure 2-16A). D'autre part, un rétrotype commun (avec la double présence d'enJSRV-7 et enJSRV-18) est observé pour les populations de mouton du bassin méditerranéen et du sud de l'Europe supportant les données obtenues précédemment *via* l'analyse des ADN mitochondriaux [7] et montrant l'importance des échanges commerciaux maritimes de la mer méditerranée dans les mouvements de mouton (Figure 2-16C). Les populations Scandinaves présentent également un rétrotype comparable à celles des Îles Faeroe et d'Islande appuyant les données historiques de mouvement maritime de populations de mouton avec les navigateurs du Nord (Vikings) il y a environ 1200 ans (Figure 2-16D). Enfin, cette étude apporte également des arguments génétiques à l'origine anecdotique de certaines races de mouton. Par exemple, une des dix races des îles britanniques analysée et nommée Jacob possède un rétrotype totalement différent des autres populations britanniques et très similaire à celles d'Afrique et du Moyen Orient (avec seulement enJSRV-18). L'origine de cette race de mouton britannique est inconnue, pour certains cette race a une origine mésopotamienne et

du Moyen Orient, pour d'autres elle fut importée par les vikings. Cette race a été nommée en référence à l'histoire biblique de Jacob qui pris « *tous les moutons tachetés ou mouchetés du troupeau de son beau-père Luban en gage de salaire* » (Genèses 30 : 25-43, probablement une des premières références de croisement sélectif de bétail). Cette étude supporte donc un lien direct entre ces moutons Jacob et le moyen Orient ou l'Afrique. Enfin, un lien génétique entre les moutons de la race Soya (île située au nord de l'Europe) et les mouflons méditerranéens et asiatiques semble exclure l'origine nordique et « viking » de cette race si longtemps débattue.

**Au cours de mon travail de thèse, afin d'étudier et de caractériser génétique les populations de mouton de la race de mouton Awassi au Liban, j'ai donc utilisé ces enJSRVs insérés de manière polymorphique dans le génome des moutons domestiques comme marqueurs génétiques.**

## Chapitre 3 : *Les ADNs Mitochondriaux*

---

### 3.1. Introduction

L'ADN mitochondrial (ADNmt) est de loin l'outil moléculaire le plus utilisé pour étudier la domestication [173]. L'ADNmt peut également révéler de récents processus démographiques touchant une population, par exemple, son expansion démographique ou même son histoire plus complexe. Chez les mammifères, l'ADNmt est presque exclusivement héréditaire de la mère, haploïde, il ne subit pas de recombinaison et la structure et l'organisation de ses gènes sont bien conservées [174]. L'ADNmt est donc fréquemment utilisé afin de mener des analyses phylogénétiques à plusieurs niveaux taxonomiques, au sein d'une même espèce jusqu'à des ordres de mammifères. Chez les bétails, il a par exemple été utilisé pour étudier la variation génétique et le degré de filiation entre des animaux domestiques de nos jours et leur ancêtre présumé devenu aujourd'hui sauvage. [2, 173]

### 3.2. Les mitochondries

Les mitochondries sont des organites cellulaires trouvés dans le cytoplasme. Ces organites cytoplasmiques essentiels sont entourés de deux bicouches lipidiques [175] et se trouvent dans toutes les cellules nucléées. Les mitochondries sont les principaux générateurs d'énergie de la cellule en contribuant à la production d'ATP (adénosine triphosphate) intracellulaire *via* la chaîne respiratoire et le cycle de Krebs. Elles sont responsables de la génération de 90% de l'ATP cellulaire par le processus de la Phosphorylation oxydative [176]. Ainsi, elles participent à la vie et à la mort de la cellule. Les mitochondries sont les seuls organites dans les cellules animales possédant leur propre ADN. Chez les mammifères, les mitochondries sont transmises à la progéniture que par la mère [177-179]. La possibilité d'héritage paternel et de recombinaison de l'ADNmt a été suggérée mais cela reste controversé [180-182]. Au sein des mitochondries, l'ADNmt se présente sous la forme d'une molécule circulaire fermée d'environ 16600 paires de bases [176, 183, 184]. Contrairement à l'ADN du génome nucléaire, l'ADNmt évolue à un rythme plus rapide [185] avec des régions présentant des taux de variations nucléotidiques différents, ce qui rend cette molécule idéale pour réaliser des comparaisons intra et inter-espèces par des analyses phylogénétiques au niveau « microévolutif » [186].

### 3.2.1. Origine des mitochondries

Selon la théorie endosymbiotique, les mitochondries des eucaryotes semblent être le résultat d'une endosymbiose entre une cellule bactérienne oxydante et une cellule glycolytique primitive [187, 188]. En 1998, l'analyse phylogénétique des gènes mitochondriaux codant pour les protéines et les ARNs ribosomiques (ARNr) suggère que l'ancêtre des mitochondries est une  $\alpha$ -protéobactéries [189]. Par contre, l'hôte de l'événement endosymbiotique s'est avéré plus difficile à déterminer [190-192]. Une seconde analyse des arbres phylogénétiques, construites à partir des gènes d'ARNr et en utilisant des modèles plus appropriés en tenant compte de l'hétérogénéité de la composition des séquences (les séquences utilisées pour la construction de l'arbre), soutient une autre hypothèse dite hypothèse de l'eocyte [193]. L'hypothèse de l'eocyte propose les *Crenarchaea* comme étant les hôtes archaebactériens de cette endosymbiose. Les comparaisons des séquences phylogénétiques entre les gènes eucaryotes dérivés de l'endosymbionte et les eubactéries datent l'événement endosymbiotique à environ 1-2 milliards d'années [194, 195]. Les génomes nucléaires (ADNn) et mitochondriaux (ADNmt) dans les eucaryotes modernes sont très diversifiés et, dans la plupart des cas, varient largement par rapport à ceux des hôtes ancestraux ou de l'endosymbionte du fait de la dérive génétique, de l'évolution réductrice (perte de gènes ou d'introns lors de l'évolution) et du transfert de gènes des mitochondries au noyau (ADNn) [196, 197]. En raison de l'héritage uniparental de l'ADNmt et du goulot d'étranglement créé lors de l'embryogenèse, (réduction du nombre de copies d'ADNmt au cours de l'embryogenèse conduisant à une homogénéisation de la composition en ADNmt dans les générations suivantes [198]), la transmission de l'ADNmt partage de nombreuses caractéristiques avec celle des génomes de petits organismes asexués intracellulaires obligatoires (bactéries, parasites, etc.) [199]. En effet, l'évolution du génome mitochondrial est fortement affectée par la dérive génétique, le Cliquet de Muller (accumulation irréversible de séquences défavorables au sein d'une lignée se reproduisant par multiplication asexuée) et la redondance des gènes dans le génome de l'hôte [199]. L'évolution réductrice de l'ADNmt chez les animaux semble être arrêtée puisque tous les ADNmt d'animaux contiennent les mêmes 37 gènes, à quelques exceptions près : les nématodes, les bivalves et les cnidaires [200]. Cependant, l'analyse de l'ADNn humain révèle des fragments de séquence comparables à celles de l'ADNmt (connus sous le nom de séquences mitochondriales nucléaires, *numts*) continuellement intégrés dans le génome nucléaire mais sans qu'il y ait de redondance ou d'élimination des gènes codants de l'ADNmt [201].

### 3.2.2. Structure et fonction des mitochondries

Comme mentionné précédemment, les mitochondries contiennent leur propre ADN qui est transmis verticalement de la mère à la descendance. A l'intérieur de la cellule, l'ADN extrachromosomique est uniquement retrouvé dans les mitochondries (à l'exception des chloroplastes chez les plantes). Cependant, les protéines mitochondriales sont codées à la fois par l'ADNmt et l'ADNn. Les mitochondries adoptent différentes formes selon le type cellulaire, le tissu et les exigences métaboliques de celui-ci (Figure 3-1) [197, 202].

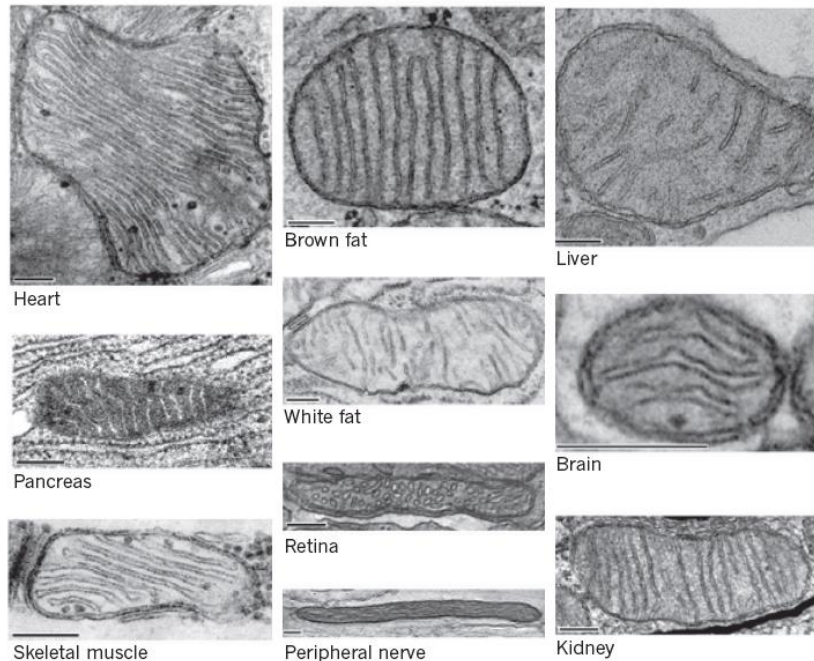


Figure 3-1: Morphologies adoptées par les mitochondries dans les différents tissus vues par microscopie électronique (échelle 200 nm) [197]

Les deux morphologies prédominantes se reflètent dans le nom grec de l'organite - « mitos » pour fil et le « chondros » pour grain. Les mitochondries se trouvent dans toutes les cellules nucléées. Elles ont généralement une taille de 1 à 10 $\mu$ m de long et de 0,5 à 1 $\mu$ m de large et sont délimitées par deux membranes. La membrane externe lisse contient une grande quantité d'une protéine membranaire (la porine) qui forme des canaux aqueux permettant à de petites molécules ayant un poids moléculaire maximal de 5000 Daltons de se diffuser à travers la membrane. La membrane interne est repliée sous forme de structures tubulaires ou lamellaires appelées crêtes (Figure 3-2). Le noyau de la mitochondrie est la matrice qui abrite l'ADNmt.

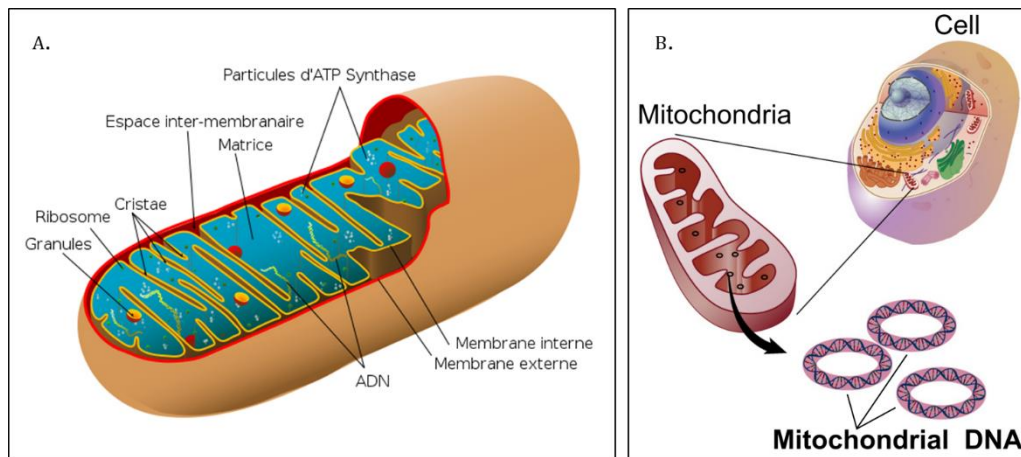


Figure 3-2 : Structure d'une mitochondrie : **A.** Compartiments membranaires de la mitochondrie. La membrane externe sépare les mitochondries du cytoplasme. Elle entoure la membrane interne, qui sépare l'espace inter-membranaire de la matrice centrale. La membrane interne est différenciée en la membrane limite interne et les crêtes. Les crêtes s'étendent plus ou moins profondément dans la matrice et sont les principaux sites de conversion d'énergie mitochondriale. Le gradient de protons entre l'espace inter-membranaire (pH 7,2-7,4) et la matrice (pH 7,9-8) entraîne la production d'ATP par l'ATP synthase dans les membranes des crêtes [203]. **B.** ADN mitochondrial circulaire présent dans la matrice [204]

Les mitochondries sont intimement impliquées dans l'homéostasie cellulaire et sont essentielles pour la synthèse d'énergie cellulaire sous forme d'ATP. En outre, elles sont importantes pour la thermogénèse, l'homéostasie du calcium et du fer, la signalisation de différentes voies intracellulaires [205] ou bien encore l'apoptose [206], ainsi que pour diverses voies métaboliques intermédiaires comme le métabolisme des acides aminés, des lipides, du cholestérol, des stéroïdes et des nucléotides par phosphorylation oxydative (OXPHOS) [207, 208].

### 3.2.3. L'ADN mitochondrial

Comme l'ADNmt est une molécule relativement petite, abondante et facile à isoler, il a été la cible de choix pour les premiers projets de séquençage de génome. Ainsi, la séquence nucléotidique de l'ADNmt d'un grand nombre d'espèces est déjà connue [184]. En effet, l'ADNmt est présent en grande quantité dans les cellules. Une cellule somatique contient deux copies d'ADNn alors qu'elle renferme des centaines voire des milliers de copies d'ADNmt [209-213]. La séquence nucléotidique de l'ADNmt humain a été la première séquence complète documentée d'un génome mitochondrial en 1981 [214] et révisée plus tard par Andrews et collègues en 1999 [215]. En 1998, la séquence complète de l'ADNmt du mouton (16616 nucléotides) a été déterminée par l'équipe Hiendleder [216]. La structure et l'organisation des gènes de l'ADNmt sont relativement bien conservées entre les animaux [217] aussi divers que l'homme [184], la souris [218], la grenouille [219] et la drosophile [219] (Figure 3-3).



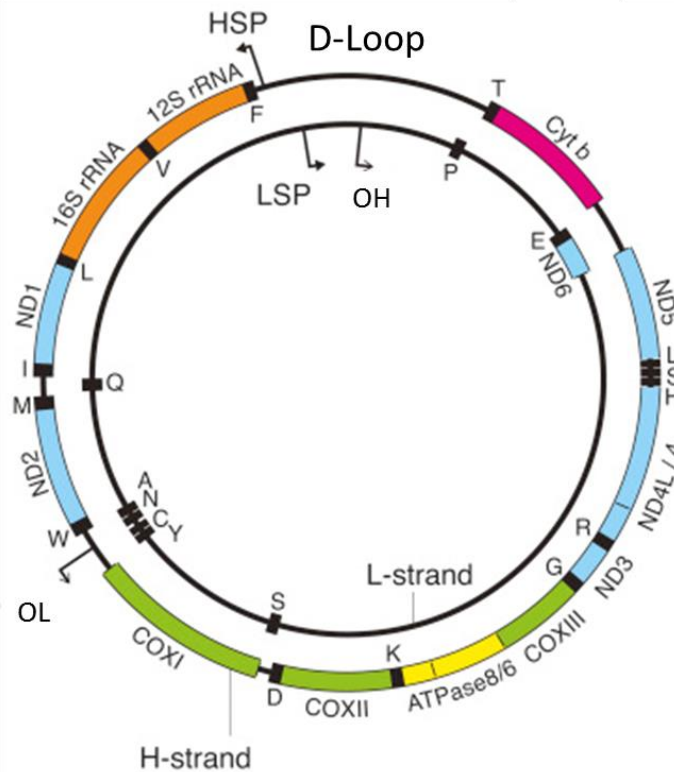


Figure 3-3: Organisation de l'ADNmt. OH et OL, origines de réplication du brin lourd et du brin léger respectivement ; ND1–ND6, sous-unités 1–6 de la NADH déshydrogénase (ETC complexe I) ; Cox1–Cox3, sous-unités 1–3 du cytochrome oxydase (ETC complexe IV); ATP6 and ATP8, sous-unités 6 et 8 de l'ATPase mitochondriale (complexe V); Cyt b, cytochrome b (complexe III) et la D-Loop. [220].

L'ADNmt est un petit ADN circulaire double brin très compact. Les deux brins sont différents d'un point de vue de leur composition nucléotidique. De ce fait, en se basant sur leur densité ils sont nommés brin lourd (Heavy = H) avec une forte teneur en Guanine (G) et Thymine (T) et brin léger (Light = L) possédant une forte teneur en Adénine (A) et Cytosine (C) [221]. La majorité des gènes ne renferment pas d'introns et les séquences intergéniques sont soit absentes ou réduites à quelques bases [222]. D'autre part, les ARNt et ARNr sont petits en taille et certains gènes codant des protéines se chevauchent [217]. La composition de l'ADNmt en base est de 33.68% A, 27.36% T, 25.86% C, et 13.10% ; avec un pourcentage de (A+T) (61.04%) plus élevé que (G+C) (38.96%) [223]. Les codes génétiques utilisés pour l'ADNmt diffèrent du code génétique universel par un usage différent des codons pour un certain nombre d'espèces [224]. Par exemple, le codon TGA qui est habituellement un codon stop dans l'ADNn, code pour le tryptophane dans l'ADNmt des vertébrés. De même, AGR (R peut être un A ou G) qui est un codon stop dans l'ADNmt des vertébrés, code pour la sérine chez les échinodermes (ADNmt) et l'arginine chez les levures (ADNmt). (Table 3-1).

Table 3-1: Comparaison du code génétique universel par rapport aux codes génétiques utilisés pour l'ADNmt

<i>Organisme</i>	<i>Codon</i>	<i>Standard</i>	<i>Variation</i>
<i>Vertébrés</i>	AGA, AGG	Arginine	Stop
	AUA	Isoleucine	Méthionine
	UGA	Stop	Tryptophane
<i>Invertébrés</i>	AGA, AGG	Arginine	Serine
	AUA	Isoleucine	Méthionine
	UGA	Stop	Tryptophane
<i>Levure</i>	AUA	Isoleucine	Méthionine
	UGA	Stop	Tryptophane
	CUR (R=A/U/G/C)	Leucine	Thréonine

L'ADNmt de tous les vertébrés renferme 37 gènes codant pour treize protéines qui composent les cinq complexes de la chaîne respiratoire à l'exception du complexe « II » qui est uniquement codées par des gènes nucléaires [214, 225, 226]. Sept gènes de l'ADNmt codent pour les sous-unités du complexe « I » (ND1, 2, 3, 4L, 4, 5, 6), un pour la sous-unité du complexe « III » (cyt b), trois pour les sous-unités du complexe « IV » (COI, II, III) et deux pour les sous-unités du complexe « V » (ATP 6 et 8) [227]. De plus, ces gènes de l'ADNmt codent également pour deux ARN ribosomiaux (ARNr), 12S et 16S, et vingt-deux ARN de transfert (ARNt) nécessaires pour la synthèse intra-mitochondriale de protéines [214]. Les gènes codant pour 12 de ces protéines, les 2 ARNr et 15 des 22 ARNt sont présents sur le brin lourd de l'ADNmt alors que le brin léger code pour la treizième protéine (ND6) et les 8 autres ARNt [184] (Figure 3-3 et Table 3-2).

Table 3-2: Gènes trouvés dans les génomes mitochondriaux des animaux : 13 protéines, 2 ARNr, et 22 ARNt [200]

<b>Protéines</b>	<b>Gènes</b>	<b>Synonymes</b>	<b>Brin</b>
Sous-unités I, II, III du cytochrome oxydase	<i>COI, COII, COIII</i>	<i>cox1, cox2, cox3</i>	H
Apoenzyme du cytochrome b	<i>Cyt-b</i>	<i>cob</i>	H
Sous-unités 1-6, 4L de la NADH déshydrogénase	<i>ND1-6, 4L</i>	<i>nad1-6, 4L</i>	ND6 (L) H
Sous-unités 6, 8 de l'ATP synthétase	<i>A6, A8 ou ATP6, ATP8</i>	<i>atp6, atp8</i>	H
Grande sous-unité de l'ARN ribosomique	<i>lARNr</i>	<i>rnl</i>	H
Petite sous-unité de l'ARN ribosomique	<i>sARNr</i>	<i>rns</i>	H
18 ARNs de transfert chacun spécifiant un seul acide aminé.	La lettre correspondante du code de l'acide aminé	<i>trnX</i>	L= ARNt Glu, Gln, Ala, Asn, Cys, Tyr, Pro <sup>1</sup> Les autres sur le brin H
2 ARNs de transfert spécifiant la leucine	Différencié par le codon reconnu, <i>L(CUN)</i> et <i>L(UUR)</i>		H
2 ARNs de transfert spécifiant la sérine	Différencié par le codon reconnu, <i>S(AGN)</i> et <i>S(UCN)</i>		L= UCN H= AGN

En plus des régions codantes, l'ADNmt renferme une région très variable appelée la région de contrôle (CR) ou D-Loop. Dans la littérature, les termes CR et D-Loop sont fréquemment utilisés comme des synonymes [202, 223, 229, 230]. Dans ce manuscrit, on a adopté le terme D-Loop pour décrire cette région contrôlant la réplication et la transcription de l'ADNmt [231].

### 3.2.4. Héritage de l'ADN mitochondrial

L'héritage maternel de l'ADNmt chez les vertèbres est largement accepté comme le modèle standard de la transmission de l'ADNmt [232-234]. La destruction ou l'élimination enzymatique de l'ADNmt du spermatozoïde dans les oocytes a été expliquée par plusieurs mécanismes ayant lieu avant, pendant ou après la fécondation [235]. Par exemple chez la langouste, le spermatozoïde ne contient pas d'ADNmt [236]. Dans le cas des *Ascidia*, seul le noyau du spermatozoïde passe dans l'œuf alors que le reste est à l'extérieure [237]. Un autre mécanisme, décrit chez les oursins [238] et les souris [239] est la destruction des

<sup>1</sup> Glu=Glutamate, Gln=Glutamine, Ala=Alanine, Asn=Asparagine, Cys=Cystéine, Tyr=Tyrosine, Pro=Proline.

mitochondries des spermatozoïdes par l'oocyte après fécondation par des processus protéolytiques et par oxydation [240]. L'héritage maternel de l'ADNmt peut aussi être tout simplement le résultat de la dilution de l'ADNmt paternel dans l'oocyte [184] du fait du nombre élevé de mitochondries dans l'oocyte (>105) [210, 241] par rapport à celui d'un spermatozoïde (50-70) [242]. Toutefois, un certain nombre d'études ont contesté ce modèle en démontrant un héritage paternel de l'ADNmt chez différents animaux vertèbres et invertébrés comme la drosophile [228], l'abeille [243], les moutons [244], la souris [245], les oiseaux (*Parus major*) [246], les moules [247], et l'homme [248]. Généralement, même si la fécondation implique un mélange de l'ADNmt maternel et paternel, la contribution de l'ADNmt paternel au réservoir du zygote semble certes exister mais reste *a priori* relativement faible voire inexistante [249].

### 3.2.5. Homoplasmie et Hétéroplasmie

Les cellules nucléées des mammifères contiennent de multiples copies d'ADNmt. Quand toutes ces molécules d'ADNmt sont identiques on parle d'homoplasmie, un phénomène renforcé par le goulot d'étranglement (ou bottleneck) qui réduit la diversité des ADNmt transmis à la génération suivante par deux processus pré et post-fécondation.

- Celui en pré-fécondation, se produit durant l'ovogenèse, quand le nombre de mitochondries est considérablement réduit dans l'ovocyte avant sa maturation.
- Celui en post-fécondation se produit entre la formation du zygote et les stages embryonnaires du blastocyste, durant lesquelles la division cellulaire est intense tandis que la prolifération mitochondriale est bloquée ; ce mécanisme entraîne donc mécaniquement une diminution du nombre de mitochondries par cellule [235].

Toutefois, certaines cellules peuvent renfermer un mélange d'ADNmt avec des génotypes différents, ainsi appelé hétéroplasmie [250]. Les variants hétéroplasmiques sont aussi héréditaires, mais le niveau d'hétéroplasmie peut varier considérablement entre une mère et ses enfants. Ce changement dans les niveaux d'hétéroplasmie s'explique apparemment par l'hypothèse de bottleneck mitochondrial [251]. Théoriquement, le goulot d'étranglement fournit un mécanisme évolutif permettant l'élimination des variantes d'ADNmt potentiellement néfastes d'une population, protégeant ainsi les espèces de l'effondrement mutationnel prédit par le Cliquet de Muller (accumulation irréversible de séquences défavorables) [252, 253]. L'hétéroplasmie présente donc un intérêt tout particulier dans les études de génétique des populations vu qu'il peut s'agir d'une évidence directe de l'apparition et de la transmission de nouveaux variant d'ADNmt au sein d'une espèce donnée [254].

Différents haplotypes mitochondriaux peuvent donc coexister dans un même zygote ce qui empêche ou réduit les phénomènes de recombinaison entre ADNmt hétérologues. Le mode de transmission et l'héritage de l'ADNmt ainsi que ce relativement faible taux d'échange génétique entre des haplotypes mitochondriaux différents représentent donc des caractéristiques très utiles pour les études phylogénétiques ayant pour but de retracer les origines et les mouvements géographiques de populations [186].

### 3.2.6. Mutations de l'ADNmt

En raison d'un taux de mutation relativement élevé, l'ADNmt a été largement utilisé comme marqueur phylogénétique. En 1979, Brown et ses collègues ont étudié l'ADNmt provenant de quatre espèces de primates supérieurs (Babouin de Guinée, Rhesus Macaque, Guenon et Humain) pour déterminer l'horloge moléculaire des mitochondries et ont obtenu une valeur de 0,02 substitutions par paire de base par million d'années [255]. Bien que la méthode utilisée pour cette étude ne soit pas nécessairement généralisable aux autres espèces, cette valeur est considérée comme celle de référence en l'absence d'autres données pertinentes chez les vertébrés. Chez l'homme, ce taux est estimé à  $1.7 \times 10^{-8}$  (s/s/yr) (substitution/site/year) [256] tandis que chez les animaux du genre *Ovis* il est de  $2.51 \times 10^{-8}$  (s/s/yr) [7]. Ce taux n'est cependant pas uniforme tout le long de l'ADNmt avec un taux faible au niveau des régions codantes et relativement plus élevé dans la D-Loop. D'autre part, le taux de mutation au niveau de l'ADNmt est généralement beaucoup plus élevé que celui du génome nucléaire [185, 255]. De ce fait, le taux d'évolution du génome mitochondrial dépasse celui de la copie unique du génome nucléaire d'environ un facteur 10. De plus, de par le principe du cliquet de Muller, les génomes non recombinants comme les mitochondries accumulent des mutations (délétions) plus rapidement que ceux qui le sont (génomes nucléaires) [253, 257]. Cette forte mutabilité de l'ADNmt s'explique par l'action de radicaux libres oxygénés qui sont générés et libérés lors de la respiration sur l'ADNmt non protégé par des histones et ayant recourt qu'à un faible nombre de mécanismes efficaces de réparation d'ADN [184]. La majorité de ces mutations dans l'ADNmt sont des substitutions et/ou des délétions de nucléotides. Les délétions sont souvent la cause de certaines maladies héréditaires alors que beaucoup de substitutions semblent être plutôt neutres [258]. Ces variations de séquences nucléotidiques compliquent d'ailleurs l'identification des mutations liées à des pathologies de celles distinguant simplement le polymorphisme entre deux individus [256, 259].

### 3.3. L'ADN mitochondrial comme marqueur génétique

L'ADNmt ne représente qu'une petite fraction de la taille du génome de l'organisme, mais il a été, et de loin, le marqueur le plus populaire pour étudier la diversité moléculaire chez les animaux au cours des quatre dernières décennies [213]. L'analyse du polymorphisme de l'ADNmt est parfaitement adaptée à l'étude de l'origine et de l'évolution des espèces. Les raisons de l'utilisation de l'ADNmt comme marqueur génétique de choix pour ces études sont bien connues. Expérimentalement, l'ADNmt est relativement facile à amplifier vu qu'il est présent en plusieurs copies dans la cellule. De plus, les gènes mitochondriaux sont fortement conservés chez les animaux, avec très peu de duplication, pas d'introns et des régions intergéniques très courtes [260]. L'ADNmt est très variable dans les populations naturelles en raison de son taux de mutation élevé, ce qui peut générer un indicateur sur l'histoire des populations à courts intervalles de temps étant transmis verticalement par la mère à la descendance. En revanche, les régions variables (e.g. la D-Loop) sont flanquées de régions très conservées (e.g. ADN ribosomique) permettant de concevoir des amorces et de réaliser des PCR afin de les séquencer. Les seules questions techniques associées aux marqueurs mitochondriaux proviennent de l'amplification par PCR des gènes qui ont été insérés dans le génome nucléaire (*numts*) chez certaines espèces [261]. De toute évidence, l'ADNmt est la solution la plus commode et la plus économique lorsqu'une nouvelle espèce doit être génétiquement explorée dans la nature [213]. En plus des cotés techniques et pratiques, l'utilisation de l'ADNmt comme marqueur repose sur des raisons plus fondamentales [262, 263]. En effet, L'ADNmt possède un certain nombre de propriétés biologiques spécifiques qui en font un marqueur approprié de la biodiversité moléculaire. Premièrement, en raison de son mode d'héritage non mendélien, la molécule d'ADNmt est considérée comme un locus unique et non-recombinant d'où le fait que tous les gènes mitochondriaux partagent une généalogie commune [264]. Cela simplifie considérablement la représentation et l'analyse des variations des données au sein d'une espèce. Deuxièmement, l'ADNmt est supposé évoluer de façon presque neutre. Étant impliqué dans les fonctions métaboliques de base (respiration), les gènes mitochondriaux codants sont considérés moins susceptibles que d'autres gènes à subir des mutations non-synonymes. Enfin, et dans la même idée, le taux d'évolution de l'ADNmt est souvent considéré comme « très précis » ; en l'absence de mutations issues de sélection positive, seules les mutations neutres s'accumulent dans le temps, de sorte que les niveaux de divergence de l'ADNmt reflètent plus réellement les temps de divergence. Combinant ces trois caractéristiques, l'ADNmt apparaît donc vraiment comme un témoin idéal de l'histoire

des populations et des espèces [213]. Cependant, du fait que l'ADNmt est hérité par la mère, il ne faut pas oublier que les phylogénies et les structures de population dérivées des données de l'ADNmt peuvent ne pas refléter celles du génome nucléaire puisque, contrairement aux femelles, les males ne transmettent pas le patrimoine génétique de leurs organelles mais uniquement celui de leur génome nucléaire [265] ou bien encore du fait de l'introgession (i.e. transfert de gènes d'une espèce à une autre) [266]. De plus, les marqueurs de l'ADNmt peuvent subir des réversions (i.e. des sites ayant subi une substitution nucléotidique reviennent à leur état d'origine), des substitutions parallèles (i.e. plusieurs mutations se produisent au même site de manière indépendante chez deux lignées différentes) et des taux d'hétérogénéité ou de mutation « hotspot » (i.e. grandes différences dans le taux auquel certains sites subissent une mutation par rapport à d'autres sites dans une même région) [267].

### **3.4. L'ADNmt et la diversité génétique des ovins**

Les marqueurs de l'ADNmt ont aussi été largement utilisés pour étudier la diversité génétique des races de mouton dans le monde entier. Deux régions de l'ADNmt sont préférentiellement utilisées pour ces analyses, cytochrome b (*Cyt-b*) et la D-Loop. Le *Cyt-b* et la D-Loop sont des marqueurs moléculaires permettant d'inférer une relation phylogénétique à différents niveaux, explorer la relation entre les moutons domestiques d'une part et entre le mouton domestique et le mouton sauvage d'autre part [7]. Ces deux marqueurs permettent la classification des populations en haplogroupes. Le mot haplogroupe vient de mot haploïde, haploûs en grec qui veut dire "unique, simple" et du mot groupe. Un haplogroupe (HPG), un groupe de séquences monophylétiques, représente un ensemble d'individus qui partage un ancêtre commun et il est constitué d'haplotypes. Un haplotype est un groupe de séquences définies par des combinaisons de sites polymorphiques sur l'ADNmt.

#### **3.4.1. Cytochrome B**

Le *Cyt-b* est couramment utilisé afin de déterminer les relations phylogénétiques entre les organismes, en raison de sa grande variabilité de séquence. Il est considéré comme le plus utile pour déterminer les relations au sein des familles et des genres. Les études comparatives impliquant le *Cyt-b* ont abouti à de nouveaux schémas de classification et ont été utilisés pour attribuer des espèces nouvellement décrites à un genre ainsi que pour approfondir la compréhension des relations évolutives. Le gène mitochondrial *Cyt-b* est aussi très utilisé dans les études systématiques pour résoudre les divergences au niveau taxonomique.

### 3.4.1.1. Structure du gène Cytochrome b

Le gène *Cyt-b* a une longueur de 1137pb chez *Ovis aries*. [223]<sup>2</sup> et est codé par le brin H riche en guanine [214, 223]. Le gène *Cyt-b* code pour une seule protéine, le cytochrome b (Cyt-b). Il est transcrit sous forme polycistronique et fait partie de l'ARNm du brin H [214, 222, 268]. *Cyt-b* présente en 5' une région de 4pb non codante qui est suivie par le codon d'initiation de traduction AUG [214, 223] et se terminant par un codon stop AGA [223]. *Cyt-b* est encadré par deux gènes codant pour les deux ARNts : tRNAGlu et tRNAThr. *Cyt-b* est très conservé au cours de l'évolution et mais présente néanmoins une séquence spécifique pour chaque espèce.

### 3.4.1.2. La protéine cytochrome b

La protéine Cyt-b intervient dans le processus de phosphorylation oxydative, un ensemble de réactions qui *via* l'oxydation du NADH (nicotinamide adénine dinucléotide) et du FADH<sub>2</sub> (Flavine adénine dinucleotide) génère de l'énergie (ATP) grâce à la formation d'un gradient de protons. Ce processus est effectué par la chaîne respiratoire qui est localisée dans la membrane interne mitochondriale et qui est composée de 5 complexes I à V (Figure 3-4A) :

- Complexe I : NADH-coenzyme Q oxydoréductase,
- Complexe II : succinate-coenzyme Q oxydoréductase,
- Complexe III : coenzyme Q-cytochrome c oxydoréductase,
- Complexe IV : cytochrome c oxydase.
- Complexe V : ATP synthase F<sub>0</sub> et F<sub>1</sub>

Cyt-B appartient au complexe III qui est un dimère formé de deux monomères identiques. Chacun de ces monomères renferme 11 sous-unités différentes désignées par sous-unités I-XI [269]. Ces sous-unités comprennent les protéines de base I et II, le cytochrome b (sous-unité III), le cytochrome c<sub>1</sub> (sous-unité IV), la protéine Rieske fer-soufre (sous-unité V) et plusieurs autres protéines plus petites (Figure 3-4B). Les sous-unités III, IV et V sont les principales protéines impliquées dans les réactions d'oxydo-réduction tandis que les protéines de base I et II sont essentielles pour l'assemblage et l'association correcte des centres redox dans le Complexe III [270-273]. Le complexe III se situe à l'intérieur de la membrane interne mitochondriale et il est le troisième complexe enzymatique à intervenir dans la chaîne de transport d'électrons au cours de la phosphorylation oxydative mitochondriale. Il catalyse l'oxydation d'une molécule d'ubiquinol Q<sub>10</sub>H<sub>2</sub> et la réduction de deux molécules de cytochrome c et transporte les protons de la matrice vers l'espace intermembranaire *via* un

---

<sup>2</sup> GenBank Accession Number: KJ954145



ensemble de réaction appelé le cycle Q (Figure 3-4). La protéine Cyt-b est très conservée entre les espèces, hydrophobe, elle contient environ 400 acides aminés formant huit domaines transmembranaires. Cyt-b se lie de manière non-covalente à deux groupes hémiques qui diffèrent par leur potentiel d'oxydoréduction et leur spectre d'absorption d'où leur nomenclature b-562 (haut potentiel=Cyt-b<sub>H</sub>) et b-566 (faible potentiel=Cyt-b<sub>L</sub>) [274-276]. La protéine Cyt-b est la sous-unité III du complexe III (coenzyme Q-cytochrome c réductase encore appelé complexe cytochrome bc<sub>1</sub>) de la chaîne respiratoire mitochondriale [277]. Cyt-b est par ailleurs la seule sous-unité protéique de ce complexe codée par l'ADNmt.

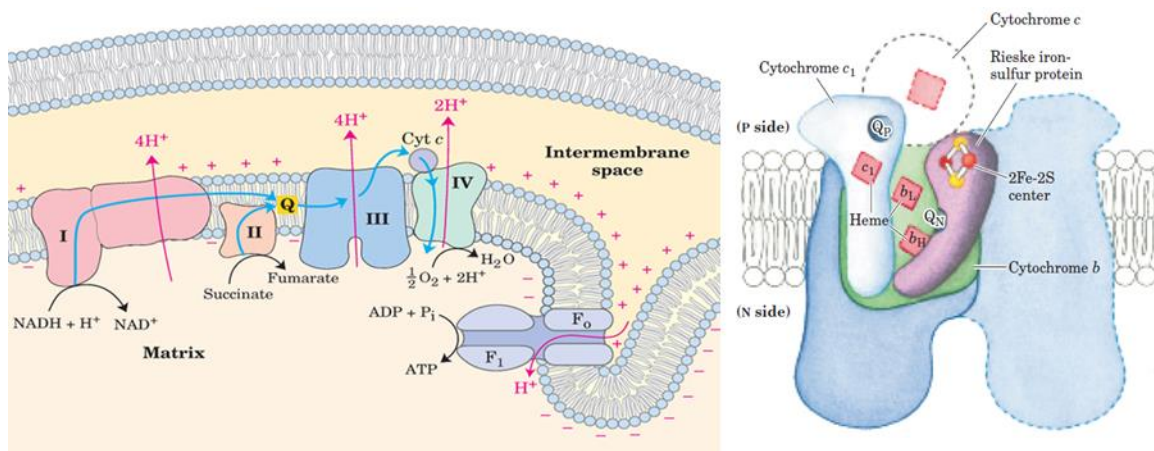


Figure 3-4: A) Chaîne respiratoire/phosphorylation oxydative : les électrons du NADH et d'autres substrats oxydables passent à travers une chaîne de transporteurs (complexes de la chaîne respiratoire) disposés asymétriquement dans la membrane interne. Le flux d'électrons (représenté par les flèches bleues) s'accompagne d'un transfert de protons à travers la membrane interne vers l'espace intermembranaire, produisant un gradient chimique (pH alcalin dans la matrice) et un gradient électrique (la membrane chargée négativement du côté de la matrice). La membrane mitochondriale interne est imperméable aux protons ; les protons ne peuvent entrer dans la matrice que par des canaux spécifiques (F<sub>0</sub>). La force motrice des protons qui les conduit dans la matrice fournit l'énergie pour la synthèse de l'ATP, catalysée par le complexe F<sub>1</sub> associé à F<sub>0</sub>. B) Structure du complexe Complexe III (cytochrome bc<sub>1</sub>) [269]. P site= espace intermembranaire, N site= Matrice. Cytochrome c<sub>1</sub> et la protéine Rieske-Fe-S se projettent à la surface P et peuvent interagir avec le cytochrome c dans l'espace intermembranaire. Le complexe présente deux sites distincts de liaison à l'ubiquinone, Q<sub>N</sub> et Q<sub>P</sub>. La structure dimérique est essentielle à la fonction du Complexe III. L'interface entre les monomères forme deux poches, chacune contenant un site Q<sub>P</sub> à partir d'un monomère et un site Q<sub>N</sub> de l'autre. Les intermédiaires d'ubiquinone se déplacent dans ces poches.

### 3.4.1.3. Le cytochrome b et la phylogénie

Comme déjà mentionné, *Cyt-b* est utilisée comme marqueur génétique bien que son taux de mutation soit inférieur à celui de la D-Loop [229, 278]. Le *Cyt-b* évolue lentement en termes de substitutions non synonymes néanmoins son taux d'évolution pour les substitutions synonymes ou silencieuses est relativement rapide [279]. En effet, *Cyt-b* est soumis à de fortes contraintes en termes d'évolution car certaines parties du gène et de la protéine résultante doivent être impérativement conservées pour des raisons fonctionnelles [280]. La

plupart des séquences variables de *Cyt-b* sont situées dans les régions codantes pour les domaines transmembranaires ou aux extrémités amino et carboxy-terminal de la protéine Cyt-b [279]. Cependant, *Cyt-b* est considéré comme suffisamment variable pour répondre à des questions au niveau de population et suffisamment conservé pour clarifier les relations phylogénétiques plus profondes au niveau des familles et des genres. Les études comparatives impliquant *Cyt-b* ont entraîné de nouveaux schémas de classification [281, 282]. La diversité de *Cyt-b* est aussi utilisée pour déterminer l'horloge moléculaire de l'ADNmt, pour estimer le temps de divergence ou le lien de parenté d'une espèce ou une population donnée avec des ancêtres proches ou présumés. Plusieurs études ont utilisé cette région pour établir des arbres phylogénétiques chez les ovins [7, 13]. A la recherche d'événements de domestication non détectés, Meadows et ses collègues [7] ont étudiés 9 races de moutons (197 échantillons) de la Turquie et d'Israël ; deux pays faisant partie du croissant fertile, lieu de domestication des moutons. La majorité des échantillons se regroupent dans les trois haplogroupes déjà connus A, B, C tandis que six se distinguent des autres par la présence de trois haplotypes constituant deux nouveaux haplogroupes : un de ses haplotypes correspond à l'haplogroupe D et deux à l'haplogroupe E. Ces deux nouveaux haplogroupes constituent un argument supplémentaire que d'autres événements de domestication ont eu lieu. De plus, une analyse phylogénétique des séquences de *Cyt-b* a aussi permis une comparaison directe entre les haplogroupes des moutons domestiques et sauvages. Comme présenté sur la Figure 3-5A, les HPGs forment des branches proches de celles de *Ovis orientalis*, *O. vignei* et *O. ammon* tout en présentant une diversité de séquence assez élevée. Globalement, cette étude a révélé un haut niveau de diversité nucléotidique des moutons analysés menant à la découverte de deux nouvelles origines maternelles et confirmant ainsi le rôle central de ces régions (Turquie et Israël) au cours de la domestication du mouton.

Par contre, dans une thèse publiée en ligne [283], l'analyse du *Cyt-b* a pu démontrer clairement l'implication de *O. orientalis gmelini* dans le processus de domestication.

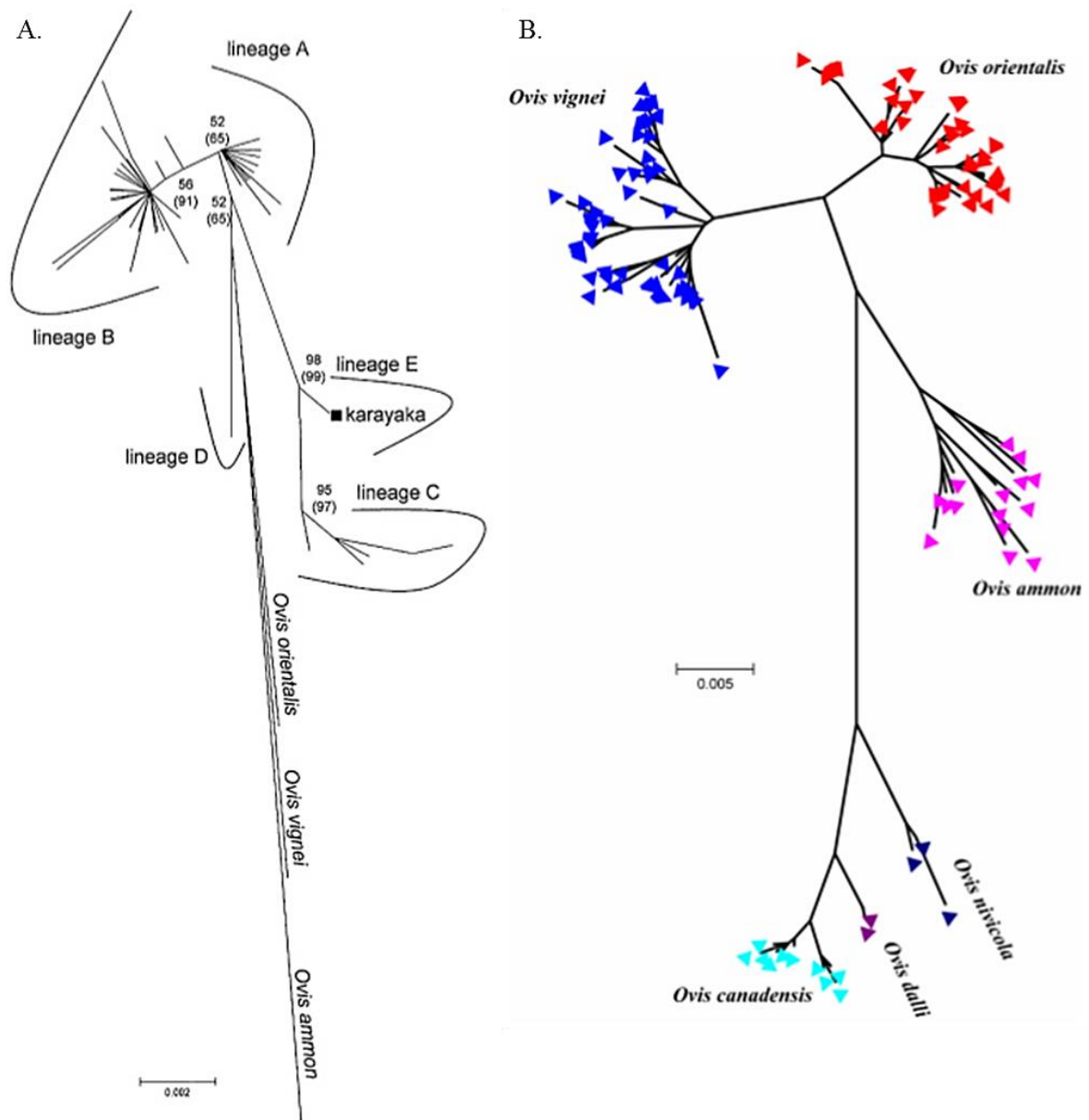


Figure 3-5: Arbres phylogénétiques à partir des séquences du cytochrome b. A. Arbre selon Neighbor-joining montrant différentes lignées génétiques (A-E) de moutons domestiques (*O. aries*) par rapport à des moutons sauvages [7]. B. Arbre phylogénétique montrant la diversité génétique au sein de différentes espèces sauvages de mouton : *Ovis canadensis* ; *Ovis dalli* ; *Ovis nivicola* ; *Ovis ammon* ; *Ovis orientalis* ; *Ovis vignei* [13].

Plusieurs classifications ont été proposées pour le genre *Ovis* en fonction de la morphologie des animaux, de leur distribution géographique [284-287] ou bien encore selon leur nombre de chromosome [288]. En 2010, Rezaei et ses collègues ont étudié la classification et l'évolution de genre *Ovis* en se basant sur l'étude des séquences du *Cyt-b* des moutons sauvages [13]. Des échantillons de fèces (=290 échantillons) représentant la plupart des espèces et sous-espèces décrites dans le genre *Ovis* ont été collectés de 43 régions différentes en Eurasie et en Amérique du Nord. L'étude phylogénétique basée sur *Cyt-b* de ces moutons a bien clarifié la classification des moutons sauvages en 6 espèces différentes rejetant ainsi

l'hypothèse qu'*Ovis orientalis* (mouflon asiatique) et *O. musimon* (mouflon européen) soient deux espèces distinctes (Figure 3-5B).

### 3.4.2. D-Loop

#### 3.4.2.1. Structure

Au niveau de l'ADNmt, il existe seulement deux régions non-codantes. La première région de 30pb renferme l'origine de réplication du brin léger ( $O_L$ ) située entre les gènes ARNt asparagine et ARNt cystéine [183]. La deuxième est la région D-Loop qui est la principale région régulatrice du génome mitochondrial puisqu'elle contient l'origine de la réplication du brin lourd ( $O_H$ ) ainsi que des promoteurs pour la transcription du brin H (HSP), HSP1 et HSP2 (Heavy Strand promotor), et du brin L, LSP (Light Strand Promotor) (Figure 3-6). Dans les cellules de vertébrés métaboliquement actives, l'ADNmt double-brin forme dans cette région une boucle de déplacement (Displacement loop, d'où le nom D-Loop donné à cette région) qui est une structure triple-brin au niveau de laquelle un petit brin d'acide nucléique appelé ADN 7S d'environ 0.5kb et complémentaire du brin léger (L-strand) déplace le brin lourd correspondant (H-strand) (Figure 3-6) [183, 289].

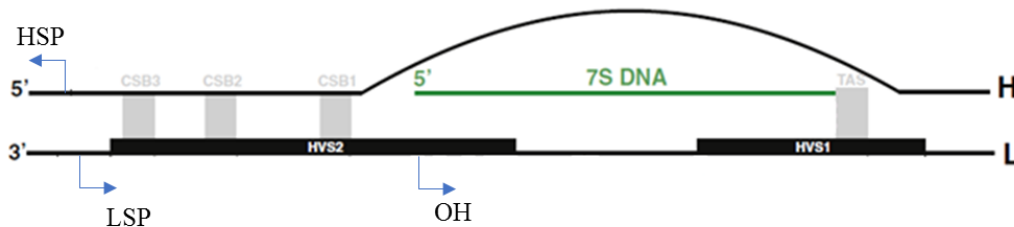


Figure 3-6: Présentation schématique de la structure de la D-Loop [289]

Chez *Ovis aries*, la D-loop est de 1180pb et se situe entre les nucléotides 15436-16615 de l'ADNmt [223]. Elle est limitée aux extrémités 5' et 3' par, respectivement, les gènes de l'ARNt proline (Pro-ARNt) et de l'ARNt phénylamine (Phe-ARNt) [290]. Ainsi, la D-Loop peut être divisée en trois domaines appelés L (Left), R [291] et C (Central) (Figure 3-7) [292]. Le domaine L, situé à l'extrémité 5', s'étend du gène pro-ARNt jusqu'à la séquence de terminaison associée TAS (Termination-Associated Sequence) de 15pb, conservées chez les vertébrés et retrouvées à l'extrémité 3' de la D-loop, elle est impliquée dans la terminaison de la réplication [291, 293]. Il contient également deux séquences palindromiques de 75pb qui sont impliquées en tant que site de reconnaissance dans la terminaison de la synthèse du brin H. Le domaine R, adjacent au gène Phe-ARNt, contient les deux promoteurs (HSP et LSP) et l'origine de la réplication ( $O_H$ ), ce qui en fait la partie fonctionnelle la plus importante de la D-

Loop. Il contient également des blocs de séquences conservées appelé Conserved Sequence Block (CSB) [294] localisés entre le promoteur LSP et l'origine de la réplication du brin lourd ( $O_H$ ). Ces séquences sont conservées chez tous les vertébrés et participent à la formation d'une amorce ARN nécessaire à la réplication de l'ADNmt [289]. En effet, trois de ces séquences, CSB I, CSB II et CSB III sont conservées entre l'ADNmt humain, de rat et de souris avec une  $O_H$  presque toujours située près de CSB I [290]. Néanmoins, le CSB I est habituellement associé à l'initiation de la synthèse de l'ADN de la D-Loop au niveau de  $O_H$  [295]. CSB II joue un rôle dans la terminaison de la transcription [183, 296] tandis que le rôle de CSBIII n'est pas encore bien déterminé. Enfin, le domaine central (central CSB) (Figure 3-7), situé entre les domaines L et R, est le plus conservé et est celui qui est utilisé comme horloge moléculaire pour estimer les temps de divergence des différentes espèces [230].

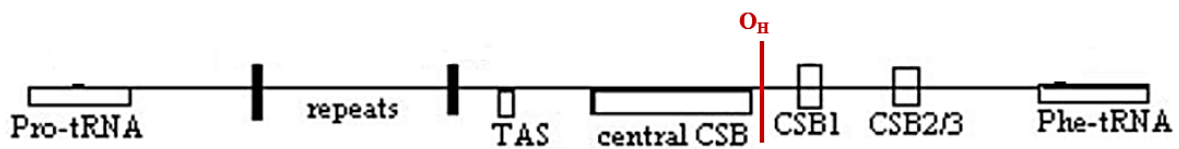


Figure 3-7: Structure de la D-Loop chez les ovins. CSB (Conserved Sequence Block) : Bloc de séquence conservée ; TAS (termination-associated sequence) : Séquence de terminaison associée ; Repeats : Région de séquences répétées [244] 93, 399–403).

En général, les séquences non codantes comme la D-Loop présentent des niveaux élevés de variation par rapport aux séquences codantes telles que le gène *Cyt-b* [297]. En effet, la D-loop est la région la plus variable dans l'ADNmt. Elle contient deux régions hypervariables : HVR-I qui chevauche la séquence TAS et HVR-II couvrant notamment les CSBI, II et III (Figure 3-6) [176, 184]. Le taux de mutation des deux régions hypervariables (HVR-I, HVR-II) de la D-loop est estimé à 100 à 200 fois plus élevé que celui de l'ADNn [298]. En particulier, il a été suggéré qu'une région poly-C de la D-Loop appelée D310 au niveau du CBSII est plus susceptible aux dommages oxydatifs et aux attaques électrophiles et donc aux mutations par rapport aux autres régions de l'ADNmt [299]. Les mutations résultantes font que cette région de l'ADNmt est intéressante pour les analyses phylogénétiques.

### 3.4.2.2. Fonction de la D-Loop

#### a. La transcription de l'ADNmt

Les deux brins de l'ADNmt sont transcrits en ARN de façon bidirectionnelle et indépendante [188]. Cette transcription nécessite à la fois une ARN-polymérase ADN-dépendante de type T3/T7 [188] et des facteurs de transcription (mTFA, mTFB) ainsi que les promoteurs

chevauchants du brin H (HSP1 et HSP2) et le promoteur du brin L (LSP) tous les trois localisés dans la D-Loop. La transcription à partir de H1 est la plus fréquente et permet la synthèse des deux ARN ribosomiques (12S et 16S) et des deux ARNt : Phe-ARNt et Val-ARNt. Le promoteur H2 permet quant à lui d'initier la transcription de l'ARN polycistronique codant pour 12 protéines et 12 ARNt du brin H. Cette deuxième transcription est 20x moins fréquente que la première. Enfin, le brin L est transcrit en un seul ARN polycistronique pour coder les 8 autres ARNt et l'ARNm de ND6 (Table 3-2).

### *b. La Réplication*

La réplication de l'ADNmt s'effectue dans la matrice mitochondriale, indépendamment de la phase du cycle cellulaire et de la réplication de l'ADNn [300]. Dans la littérature, trois modèles de réplication de l'ADN mitochondrial ont été décrits :

#### *Modèle de déplacement asynchrone (Strand-displacement)*

Dans ce modèle, la réplication de l'ADNmt débute dans la D-Loop par la synthèse du brin H à partir d' $O_H$  jusqu'à ce que  $O_L$  soit exposée pour permettre la synthèse du brin L dans la direction opposée (Figure 3-8.A) [214, 301]. L'initiation de la réplication du brin H nécessite des amorces d'ARN courtes issues des transcrits du promoteur LSP. Par conséquent, la réplication de l'ADNmt chez les mammifères est liée et dépendante de la transcription mitochondriale [302]. De plus, et comme décrit précédemment, d'autres séquences de la D-loop participent au processus de réplication, les CSB et la TAS. Le processus d'initiation de la réplication peut être décrit en quatre étapes [301, 303]:

- 1) Un fragment ARN est transcrit à partir de LSP par l'ARN polymérase en présence du facteur d'activation de la transcription (mTFA).
- 2) Cet ARN nouvellement synthétisé (précurseur de l'amorce) reste hybridé à une région de l'ADNmt en amont de l' $O_H$  et contenant les CSB en formant une structure D-Loop stable.
- 3) Ce fragment d'ARN est ensuite clivé par la RNase MPR (Mitochondrial RNA-Processing enzyme) générant ainsi une amorce mature pour l'initiation de la réplication.
- 4) L'ADN polymérase mitochondriale initie la réplication du brin H par extension à partir de cette amorce ARN.

Après l'initiation de la réplication du brin H, deux cas peuvent se produire : i) le cas le plus fréquent, la synthèse du brin H est arrêté autour des séquences TAS [304] en présence de la protéine TAS-BP (protéine liant cette séquence de terminaison) ou ii) plus rarement et en

absence de TAS-BP (ou faible concentration), la réplication continue sur toute la longueur du génome mitochondrial. Chez les mammifères et chez les vertébrés en général,  $O_L$  est située dans une petite région non codante loin d' $O_H$  localisée dans la D-Loop. Une fois initiée, la réplication du brin L passe sur toute la longueur du brin [305, 306]. Cette réplication asymétrique des deux brins permet à la synthèse d'être continue. La durée du cycle de réplication a été estimée à environ 1h dû à une polymérisation assez lente de 270 nt/min [307]. Une fois la synthèse des deux brins terminés, les molécules filles de l'ADNmt doivent être séparées, les amorces ARN situées aux deux origines de réplication  $O_H$  et  $O_L$  sont ensuite éliminées par la RNase H1 [220, 308]. L'ADNmt néosynthétisé est lié puis réparé à chaque extrémité pour reformer un ADN circulaire [309]. En résumé, cette réplication est unidirectionnelle, continue, asymétrique, et asynchrone [310].

#### *Modèle de déplacement couplé.*

Depuis l'an 2000, un autre modèle de réplication d'ADNmt a été discuté dans la littérature et remet en question la validité du modèle de déplacement asynchrone. Dans ce modèle, la réplication des deux brins H et L de l'ADNmt se fait en même temps et à partir d'un même site d'initiation situé à proximité de  $O_H$  (différent d' $O_H$  et de  $O_L$ ) de façon bidirectionnelle (Figure 3-8.B) [311].

#### *Modèle RITOLS*

Ce modèle appelé **RITOLS** (**R**NA **I**ncorporated **T**hrough**O**ut the **L**agging **S**trand) est au modèle alternatif à celui du déplacement couplé et implique un processus de réplication dans lequel le brin H est d'abord transcrit en un intermédiaire ARN qui sert de matrice pour la synthèse de l'ADN (Figure 3-8B). Dans ce modèle, le mécanisme de réplication du brin L est similaire à celui décrit précédemment dans le modèle asynchrone [312, 313].

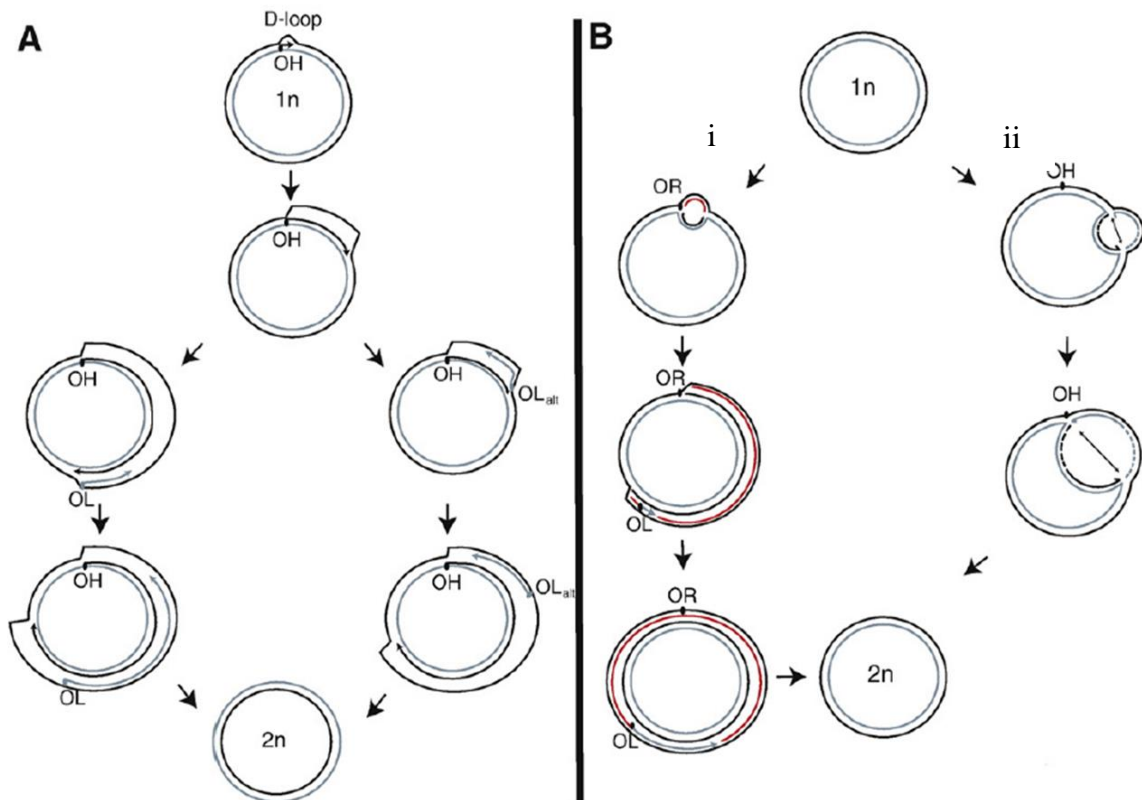


Figure 3-8: Modes de réplication de l'ADNmt [308] (A) Mode de déplacement asynchrone : la réplication du brin lourd (H) démarre en  $O_H$  et se déroule de manière unidirectionnelle (flèche noire). Lorsque l'origine de réplication du brin léger (L) ( $O_L/O_{Lalt}$ ) est exposée, sa synthèse commence dans la direction opposée (flèche grise). (B) le modèle de réplication RITOLS (i) et le déplacement couplé (ii). i) En mode RITOLS, la réplication du brin H démarre en  $O_R$  (Origine de réplication) et se déroule de manière unidirectionnelle (flèche noire). Le brin H est tout d'abord transcrit en ARN (ligne rouge) puis remplacé par de l'ADN (flèche grise). La synthèse du brin L est plus tardive et se fait selon le même principe que celui précédemment décrit pour le mode de déplacement asynchrone. ii) Dans le mode de déplacement couplé, l'initiation de la synthèse des deux brins se produit de manière synchronisée et bidirectionnelle à partir d'origine multiple dans une large zone en aval d' $O_H$ . Ceci est suivi par la progression de deux fourches de réplication qui s'arrêtent une fois arrivées à  $O_H$ .

### 3.4.2.3. La D-Loop et la phylogénie

L'ADNmt, et en particulier la D-Loop, présente une grande variation de séquence nucléotidique au niveau intra-espèce ce qui en fait un marqueur moléculaire très intéressant et largement utilisé pour étudier l'histoire évolutive de nombreux animaux [173]. En effet, cette région D-Loop permet d'identifier un nombre élevé d'haplogroupes et d'haplotypes au sein d'une même espèce. Dans une étude récente, des séquences complètes d'ADNmt ont été analysées afin d'étudier la relation des différents haplogroupes entre des populations de moutons domestiques et sauvages [314]. L'analyse du mitogénome mitochondrial a clairement rejeté l'hypothèse que l'Argali et l'Urial peuvent être les potentiels ancêtres des moutons modernes, mais n'a pas pu confirmer le statut du mouflon Asiatique par manque de caractérisation de son mitogénome. Cette étude a également permis d'évaluer la capacité de différentes séquences mitochondriales à discriminer différents haplogroupes (HPG). Les



échantillons des moutons domestiques sélectionnés correspondent aux cinq haplogroupes (HPG A, HPG B, HPG C, HPG D et HPG E) en se basant sur les études phylogénétiques du *Cyt-b* et de la D-Loop. Chaque HPG est représenté par deux animaux portant l'haplotype le plus fréquent. Les échantillons des moutons sauvages proviennent d'un seul animal *O. ammon* (Kazakhstan), de deux *O. musimon* (Allemagne) et de trois *O. vignei* (Kazakhstan). D'où au total, seize génomes mitochondriaux séquencés et analysés. L'utilité des différentes séquences du mitogénomes pour l'inférence phylogénétique a donc été examinée à l'aide de la subdivision des données, de la construction des arbres et des indices PBS (Partitioned Bremer Support). Ces tests permettent de mesurer la contribution de chaque série de séquences dans le support lors de la génération de l'arbre phylogénétique avec le mitogénome. Après élimination des régions de séquences répétées, la D-Loop affiche la plus grande valeur de PBS avec ou sans l'inclusion des mitogénomes d'animaux sauvages. Ces résultats indiquent donc que la D-Loop est en effet le marqueur le plus utile et le plus puissant pour élucider l'origine maternelle des moutons.

### 3.5. L'ADN mitochondrial et les haplogroupes

Comme décrit précédemment, l'ADNmt est sujet à la dérive génétique. Au cours de la domestication du mouton et des migrations successives de populations, certains haplogroupes ou haplotypes disparaissent ou au contraire apparaissent ce qui permet de générer des cartes phylogéographiques permettant de déterminer la diversité génétique de populations dans un lieu géographique donné [315]. Les moutons domestiques d'aujourd'hui sont répertoriés selon cinq haplogroupes mitochondriaux distincts, identifiés et nommés : HPG A, HPG B, HPG C, HPG D ou HPG E [7, 316]. Les HPG A et HPG B sont les haplogroupes dominants, respectivement, en Europe et en Asie [16, 315]. Chronologiquement, HPG A et HPG B ont tout d'abord été observés chez les moutons de Nouvelle-Zélande en 1996 [317]. Ensuite en 1998, Hiendleder et ses collègues ont confirmé la présence de ces deux haplogroupes en séquençant la D-Loop d'échantillons provenant de moutons de Russie (la race Gizarr), du Kazakhstan (les races Edilbey, Astrachan) et de l'Allemagne (les races Merinolandscha, East Friesian, Rhoenschaf) [216]. Le troisième haplogroupe, HPG C, a été détecté à une fréquence relativement élevée dans le sud-ouest de l'Asie [12], en Turquie [8], en Chine [318, 319] et au Portugal [17]. Vu leur large distribution et leur forte fréquence, ces trois haplogroupes, sont considérés comme les haplogroupes majeurs. Les deux autres haplogroupes : HPG D et HPG E sont considérés comme des haplogroupes rares, car ils ne sont présents qu'au Moyen-Orient

et en Asie, et retrouvés chez très peu d'animaux [7, 16]. Par exemple, HPG D a été observé dans le nord du Caucase dans la race Karachai [16] et dans deux individus de race Morkaraman [7] et dans deux individus (l'un Akkaraman et l'autre Norduz) [320] en Turquie et un individu de la race Linzhou au Tibet [321]. Le dernier haplogroupe, HPG E a été observé en Israël dans la race Awassi, en Turquie dans la race Tuj, Karayaka, Dağlıç, Akkaraman, Karagül, Karayaka, Morkaraman et en Chine dans les races mongoles et la race Hu [7, 8, 318, 320]. Cette répartition géographique différente des haplogroupes de moutons en Eurasie a probablement été façonnée par les événements de domestication, de sélection et de migrations successives au cours de l'évolution de ces animaux. La présence d'haplogroupes mitochondriaux distincts est généralement interprétée comme étant le résultat d'événements différents de domestication [322]. Les haplogroupes ne sont pas nécessairement représentatifs de tous les événements indépendants de domestication ayant eu lieu mais ils représentent plus, en l'état actuel de nos connaissances, le nombre minimal d'évènement de domestication susceptible d'avoir façonné l'évolution du mouton moderne et révélant d'un processus complexe impliquant des populations ancestrales diversifiées. Les supposés ancêtres sauvages du mouton domestique moderne, tel que le mouflon asiatique, l'Argali et l'Urial, ont été le sujet de plusieurs études. Ces études ont conclu que l'*O. orientalis* (Mouflon Asiatique) est l'ancêtre le plus probable de l'*O. aries*.

Aujourd'hui, les données sur les haplogroupes et les séquences obtenues par différents laboratoires chez les moutons peuvent être combinées pour réaliser des analyses conjointes. En effet, il existe une quantité suffisante de données sur l'ADNmt provenant de différentes races de mouton dans des endroits géographiques auxquelles on peut accéder librement. Par conséquent, même si des données de séquences générées proviennent d'une région limitée, comme c'est le cas dans mon étude, elles peuvent être analysées avec l'ensemble des autres données disponibles en ligne et presque à l'échelle mondiale [323].

**Le Liban fait partie de la région du centre de domestication du mouton, et la race Awassi est prédominante et extrêmement importante dans ce pays. Néanmoins, nous ne disposons actuellement d'aucune information génétique sur les populations de moutons appartenant à cette race au Liban. En complément de l'approche « rétrovirus endogènes », j'ai donc utilisé des marqueurs mitochondriaux pour mon étude, à savoir le *Cyt-b* et la D-loop, afin de caractériser génétiquement ces populations de mouton libanais et d'étudier leur relation phylogénétique à différents niveaux au sein de l'espèce *Ovis aries***

## Chapitre 4 : *Matériel et Méthodes*

---

## 4.1. Echantillons : Nature et origine

Les échantillons étudiés sont du sang complet (avec héparine) collecté des troupeaux de moutons libanais en 2015 et en 2016 (Figure 4-1). Il s'agit de 254 échantillons de sang du mouton domestiques de la race Awassi. En plus des échantillons Libanais, on a prélevé 23 échantillons de 2 troupeaux de moutons Awassi de la Syrie (Tableau 4-1). La majorité de ces animaux sont asymptomatiques avec seulement 14 présentant des troubles respiratoires

Tableau 4-1: Nombre et nature des échantillons

<i>Département</i>	<i>N° Troupeau</i>	<i>N° Echantillons</i>	<i>Elevage</i>
Hermel	6 (H-1 → H-6)	24	Ouvert ; Transhumant nomade semi-extensif
Baalbak	6 (B-1 → B-6)	18	Ouvert ; Transhumant nomade semi-extensif
	B-7	7	Fermé ; transhumant ; extensif
Beqaa Awssat	7 (BK-1 → BK-7)	25	Ouvert ; sédentaire ; semi-Extensif
	BK-8	3	Fermé ; transhumant ; extensif
Rachaïa	5	24	Fermé ; transhumant ; semi-extensif
Kobayât	1	24	Commercial
Batroun	5	27	Fermé ; sédentaire ; semi-extensif
Mont Liban	ML-C	17	Fermé ; semi-extensif ; sédentaire
	ML-F	7	Commercial
	ML-K	4	Commercial
Jezzine	1	10	Fermé ; sédentaire ; intensif ; Vieux
Saida	4	29	Agropastoralismes ; extensif ; jeune
Nabatieh	2	10	Agropastoralisme ; extensif ; jeune
Tyre	4	20	Ouverts ; Agropastoralismes ; extensif ; Vieux
	T-3	5	Fermé ; sédentaire ; semi-extensif
<b>Total</b>	<b>47</b>	<b>254</b>	
ICARDA	1	17	Fermé ; Expérimental
Syrie- AL Qusayr	1	6	Fermé ; Semi-nomade transfrontalier ; extensif
<b>TOTAL</b>	<b>49</b>	<b>277</b>	

Des échantillons de 5 ml de sangs ont été prélevés par la veine jugulaire des animaux après désinfection du site dans des tubes contenant de l'EDTA. Les prélèvements ont été aussitôt acheminés sur glace au laboratoire puis conservés à  $-20^{\circ}\text{C}$  jusqu'à utilisation.

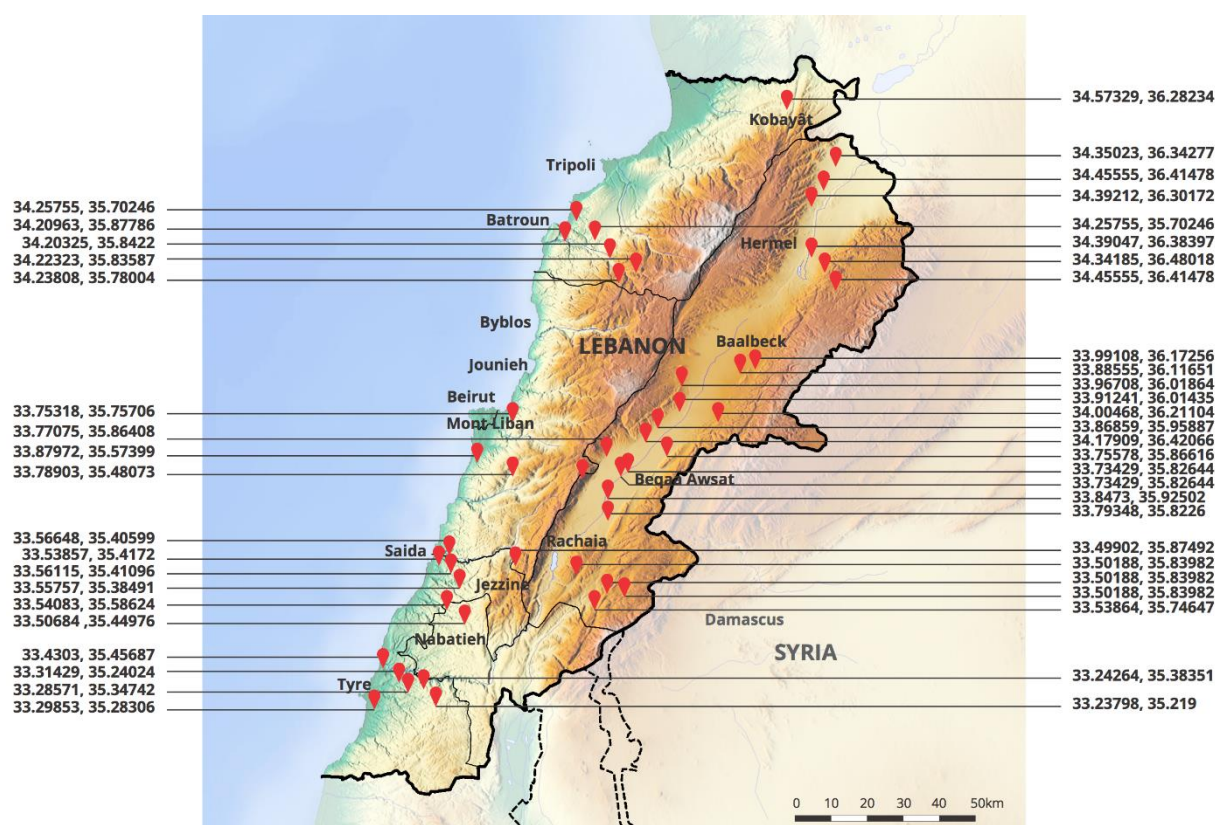


Figure 4-1: Distribution des échantillons sur le territoire Libanais

## 4.2. Extraction et dosage de l'ADN génomique

L'ADN génomique est extrait du sang avec héparine en utilisant le kit commercial « GenElute™ Blood Genomic DNA Kit » (Sigma-Aldrich®) en suivant les instructions du fabricant. Ce kit utilise le principe de la chromatographie sur colonne avec un prétraitement de l'échantillon par la protéinase k et la RNase A.

Avant de procéder à son amplification, Un dosage de l'ADN a été effectué Le spectrophotomètre NanoDrop ND-1000 (ThermoScientific, Wilmington, DE, USA) afin d'ajuster sa concentration à 10-20ng /μL pour tous les échantillons. L'extrait d'ADN (2 μL) est déposé sur un piédestal mesurant l'absorbance à 260 nm.

## 4.3. Amplification des rétrovirus endogènes

L'ADN génomique ainsi extrait est utilisé comme matrice pour la recherche des rétrovirus endogènes chez les moutons. Cinq enJSRVs ayant une structure génomique intacte et sont intégrés uniquement dans le génome des moutons domestiques de manière polymorphique : enJSRV-7 ; enJSRV-15 ; enJSRV-16 ; enJSRV-18 ; enJSRV-22 (enJS5F16) ont été étudiés.

Ces rétrovirus sont considérés comme des marqueurs génétiques très puissants afin d'établir des liens génétiques entre différents animaux et déduire l'histoire et les événements d'évolution du mouton domestique de nos jours. En plus de ces cinq rétrovirus endogènes, 2 autres sont aussi testés, enJSRV-13 et enJSRV-6.

Pour chaque rétrovirus endogène, cinq régions vont être amplifiées [68, 172] avec des couples d'amorces différentes. Pour l'amplification du provirus deux paires d'amorces sont utilisées : 5'FlankF/ProvR et Prov/3'FlankR (Figure 4-2A). L'amorce 5'FlankF est une amorce directe, complémentaire de la région flanquante du provirus et adjacente au 5'LTR. L'amorce ProvR est une amorce inverse, complémentaire d'une région au niveau du gène *gag*. L'amorce ProvF a été utilisée comme amorce directe, conçue au niveau de l'extrémité 3' du gène *env* du provirus. L'amorce inverse 3'FlankR est conçue dans une région flanquante qui est adjacente à 3'LTR. Les Solo LTRs sont détectés avec trois couples d'amorces : 5'FlankF/LTRR, LTRF/3'FlankR alors que le locus vide avec un le couple d'amorce 5'FlankF/3'FlankR. Les amorces LTRF (directe) et LTRR (inverse) sont désignées au niveau des LTR en 5' et en 3' respectivement.

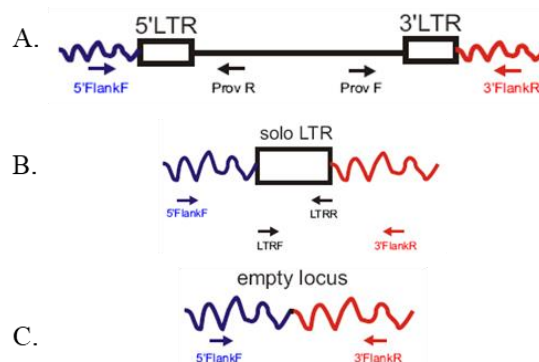


Figure 4-2: Représentation schématique des PCR avec les amorces utilisées. A. stratégie des PCR pour la détection des polymorphismes d'insertion. B et C. Stratégies de PCR pour la détection des solo LTRs et le locus vide

Les réactions ont été réalisés dans un volume final de 50  $\mu$ L, contenant 5  $\mu$ L de matrice ADN, 1x Taq Buffer (Tris-Cl, KCl,  $(\text{NH}_4)_2\text{SO}_4$ , 1.5 mM  $\text{MgCl}_2$  ; pH 8.7), 1x de la solution Q, 200 $\mu$ M de chaque dNTP, 2.5U de la HotStar Taq ADN Polymérase (« HotStar Taq® DNA Polymerase », Qiagen, Allemagne) et 0.2 $\mu$ M de chaque amorce. Le pré-mix de la PCR de tous les couples d'amorce est préparé selon le Tableau 4-2.

Tableau 4-2 : Le PCR mixe pour les rétrovirus endogènes.

	Stock Solution	Concentration finale	Vol/rx (µl)
Eau			32.5
Taq Buffer	10x	1x	5
Q-Solution	5x	1x	5
dNTP	10mM	200µM	1
Enzyme	5U/µl	2.5	0.5
Primer F	20µM	0.2µM	0.5
Primer R	20µM	0.2µM	0.5
Total Volume			45µl
DNA			5µl (50-100ng)

Les conditions des PCRs des différentes amorces sont représentées dans la Figure 4-3 avec la température d'hybridation pour chaque couple d'amorce mentionnée dans le Tableau 4-3.

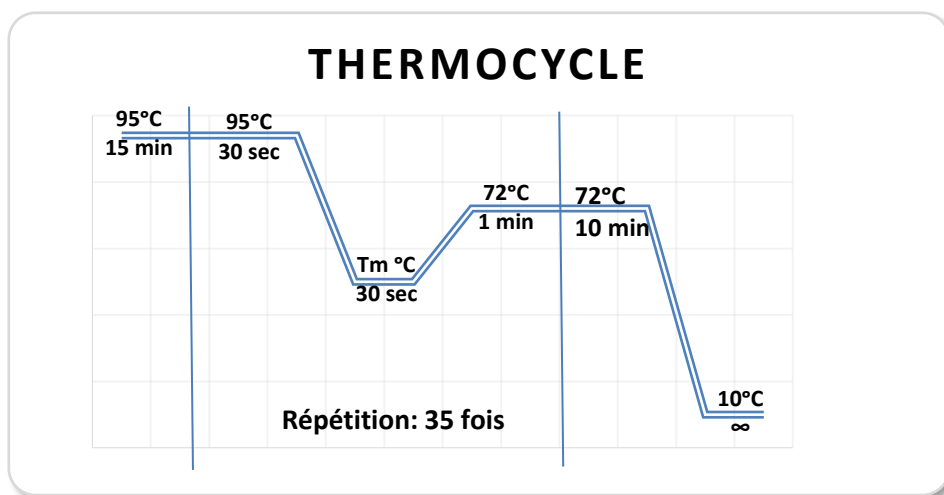


Figure 4-3: Le cycle thermique de l'amplification des rétrovirus endogènes.

Tableau 4-3 : Amorces des rétrovirus endogènes selon Arnaud *et al.*,2007 et Chessa *et al.* 2009.

Locus	Amorces	Séquence 5'-3'	Fragment	T° Hybridation °C	Produit (≈Pb)
<b>enJSRV-13</b>	5'FlankF	5'-GCATTTCATTCGTACCGCTACCA	Provirus 5'	61°C	700
	Prov R	5'-GTAGTGGCGAGGAGAACTGTCGAGC			
	Prov F	5'-AGCCTTCATTCACTGTGGCGA	Provirus 3'	61	800
	3'Flank R	5'-AGGTGTCCTTACTTCAAGCCTGC			
<b>enJSRV-18</b>	5'Flank F	5'-GGGAAGATTCGTTTCTTAGGCGCTC	Provirus 5'	61	500-600
	Prov R	5'-GTAGTGGCGAGGAGAACTGTCGAGC			
	Prov F	5'-AGCCTTCATTCACTGTGGCGA	Provirus 3'	61	700-800
	3'Flank R	5'-CAAGTGCCAGAGCCCAGAGCCA			
<b>enJSRV-7</b>	5'Flank F	5'-TGTGCACACGTGGTGGGAGTC	Provirus 5'	61	600-700
	Prov R	5'-GTAGTGGCGAGGAGAACTGTCGAGC			
	Prov F	5'-AGCCTTCATTCACTGTGGCGA	Provirus 3'	61	800-900
	3'Flank R1	5'-AGGAACTCCAGGGTCGTGCCA			
	5'Flank F	5'-TGTGCACACGTGGTGGGAGTC	Solo LTR 5'	61	600
	LTR R	5'-GCCAGCACAAGCAAGAGTCG			
	LTR F	5'-GCCAGCACAAGCAAGAGTCG	Solo LTR 3'	61	600-700
	3'Flank R1	5'-AGGAACTCCAGGGTCGTGCCA			
	5'Flank F	5'-TGTGCACACGTGGTGGGAGTC	Empty Locus	61	300
	3'Flank R2	5'-ACTCGAGAGGAAGCACGAGGGTC			
<b>enJSRV-15</b>	5'Flank F1	5'-CTTCTAGTACAGAATATAGTGGG	Provirus 5'	59	700-800
	Prov R	5'-GTAGTGGCGAGGAGAACTGTCGAGC			
	Prov F	5'-AGCCTTCATTCACTGTGGCGA	Provirus 3'	59	700-800
	3'Flank R1	5'-TCACTGTTGCTTCCTGACCAGG			
	5'Flank F1	5'-CTTCTAGTACAGAATATAGTGGG	Solo LTR 5'	59	600-700
	LTR R	5'-GCCAGCACAAGCAAGAGTCG			
	LTR F	5'-GCCAGCACAAGCAAGAGTCG	Solo LTR 3'	59	500
	3'Flank R1	5'-TCACTGTTGCTTCCTGACCAGG			
	5'Flank F2	5'-GTGGGAAGAAATTGCTTGGTACAC	Empty Locus	59	400
	3'Flank R2	5'-CCATAAGGGTGGTATCGTCTGCA			



<b>enJSRV-16</b>	5'Flank F	5'-TGCTCAGTTTCCCAGGTGCCCCA	Provirus 5'	61	600-700
	Prov R	5'-GTAGTGGCGAGGAGAAGTGTGCGAGC			
	Prov F	5'-AGCCTTCATTCCTGTGGCGA	Provirus 3'	61	700-800
	3'Flank R1	5'-GGCCAGTTGACATCTGCCAGGA			
	5'Flank F	5'-TGCTCAGTTTCCCAGGTGCCCCA	Solo LTR 5'	61	600
	LTR R	5'-GCCAGCACAAGCAAGAGTCG			
	LTR F	5'-GCCAGCACAAGCAAGAGTCG	Solo LTR 3'	61	600-700
	3'Flank R1	5'-GGCCAGTTGACATCTGCCAGGA			
	5'Flank F	5'-TGCTCAGTTTCCCAGGTGCCCCA	Empty Locus	61	300
	3'Flank R2	5'-GTGCAAGAGCTAGAGCTGGAAGG			
<b>enJSRV-22</b>	5'Flank F	5'-GGATAAGCTACACTATAAAACCAAAG	Provirus 5'	57	700-800
	Prov R	5'-GTAGTGGCGAGGAGAAGTGTGCGAGC			
	Prov F	5'-AGCCTTCATTCCTGTGGCGA	Provirus 3'	57	800-900
	3'Flank R	5'-CCATATGTAGGGATTGGGGGGTG			
	5'Flank F	5'-GGATAAGCTACACTATAAAACCAAAG	Solo LTR 5'	61	600-700
	LTR R	5'-AGAGTCACACCTGCACAGGGAG			
	LTR F	5'-GCCAGCACAAGCAAGAGTCG	Solo LTR 3'	61	700-800
	3'Flank R	5'-CCATATGTAGGGATTGGGGGGTG			
	5'Flank F	5'-GGATAAGCTACACTATAAAACCAAAG	Empty Locus	57	400
	3'Flank R	5'-CCATATGTAGGGATTGGGGGGTG			
<b>enJSRV-6</b>	5'Flank F	5'-CCAGTTCAGAAAGGGAAGGGAG	Provirus 5'	61	900-1000
	Prov R	5'-AGCCCTACAACCTGGGTGCCCA			
	Prov F	5'-AGCCTTCATTCCTGTGGCGA	Provirus 3'	61	800
	3'Flank R	5'-CAGGGGAATAACTGGTGCTACCT			
	5'Flank F	5'-CCAGTTCAGAAAGGGAAGGGAG	Solo LTR 5'	61	750
	LTR R	5'-GCCAGCACAAGCAAGAGTCG			
	LTR F	5'-GCCAGCACAAGCAAGAGTCG	Solo LTR 3'	61	750
	3'Flank R	5'-CAGGGGAATAACTGGTGCTACCT			
	5'Flank F	5'-CCAGTTCAGAAAGGGAAGGGAG	Empty Locus	61	302
	3'Flank R	5'-CAGGGGAATAACTGGTGCTACCT			

## 4.4. Amplification des gènes mitochondriaux

Afin d'étudier l'origine maternelle de la race Awassi, deux fragments de l'ADN mitochondrial ont été analysés ; le cytochrome b et la D-Loop.

### 4.4.1. Le cytochrome b

Les amorces utilisées pour amplifier le fragment de 1272 bp du gène cyt-b sont désignées à partir du génome mitochondrial complet (AF010406 positions 14,078–15,349) [324].

CytB-F : 5'-GTCATCATCATTCTCACATGGAATC-3'

CytB-R : 5'-CTCCTTCTCTGGTTTACAAGACCAG-3'

Les réactions se font dans un volume final de 50 µL, contenant 5 µL de matrice ADN, 1x Taq Buffer (Tris-Cl, KCl, (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 1.5 mM MgCl<sub>2</sub> ; pH 8.7), 1x de la solution Q, 200µM de chaque dNTP, 1U de la Taq ADN Polymérase (« Taq PCR Core Kit », Qiagen, Allemagne) et 0.5µM de chaque amorce. Le pré-mix de tous les couples d'amorce est préparé selon le Tableau 4-4.

Tableau 4-4: Le PCR mixe pour le cytochrome b.

	Stock Solution	Concentration finale	Vol/rx (µl)
Eau			31.3
Taq Buffer	10x	1x	5
Q-solution	5x	0.5x	5
dNTP	10mM	0.2mM	1
Enzyme	5U/µl	1U	0.2
CytB-F	20mM	0.5mM	1.25
CytB-R	20mM	0.5mM	1.25
		Total Volume	45µl
		DNA	5µl (50-100µg)

Le programme de PCR consistait en une dénaturation initiale à 95°C pour 5 minutes ; suivie de 40 cycles associant une dénaturation à 94°C pour 30 secondes, une hybridation à 55°C pour 30 secondes, une Elongation à 72°C pour 1 minute ; et enfin une élongation finale de 10 minute à 72°C.

#### 4.4.2. La D-Loop

Les amorces de la D-Loop ont été désignées à partir du génome mitochondrial complet (NC\_001941 position 15957- 16620) pour amplifier un fragment de 664pb [325].

D-Loop F: 5'-GTTTCACTGAAGCATGTAGGG-3'

D-Loop R: 5'-GTATTGAGGGCGGGATAAAT-3'

Les réactions se font dans un volume final de 50  $\mu$ L, contenant 5  $\mu$ L de matrice ADN, 1x Taq Buffer (Tris-Cl, KCl,  $(\text{NH}_4)_2\text{SO}_4$ , 1.5 mM  $\text{MgCl}_2$  ; pH 8.7), 1x de la solution Q, 400 $\mu$ M de chaque dNTP, 2.5U de la HotStar Taq ADN Polymérase (« HotStar Taq® DNA Polymerase », Qiagen, Allemagne) et 0.2 $\mu$ M de chaque amorce. Le volume de  $\text{MgCl}_2$  est ajusté pour avoir une concentration finale de 2.5mM. Le pré-mix de tous les couples d'amorce est préparé selon Tableau 4-5.

Tableau 4-5: Le PCR mixe de la D-Loop

	Stock Solution	Concentration finale	Vol/rx ( $\mu$ l)
water			35.3
Taq Buffer	10x	1x	5
$\text{MgCl}_2$	25mM	2.5mM	2
dNTP	10mM	0.4mM	2
Enzyme	5U/ $\mu$ l	2.5U	0.5
D-Loop F	20 $\mu$ M	0.2 $\mu$ M	0.5
D-Loop R	20 $\mu$ M	0.2 $\mu$ M	0.5
		Volume Total	45 $\mu$ l
		DNA	5 $\mu$ l (50-100ng)

Les PCRs se sont déroulées sous les conditions suivantes : une dénaturation initiale à 95°C pour 15 minutes ; suivie de 35 cycles : Dénaturation à 94°C pour 30 secondes – Hybridation à 58°C pour 40 secondes – Elongation à 72°C pour 1 minute ; et enfin une élongation finale de 10 minute à 72°C.

Les amplicons résultants de toutes les réactions de PCR ont ensuite été séparés par migration sur gel d'agarose 1% (m/v) dans du TBE 1X (0.1M Tris-base, 0.1M acide borique, 2mM EDTA, pH 8.3) pendant 25 minutes à 100 Volts. Puis, ils sont révélés par fluorescence avec le bromure d'éthidium selon les recommandations du fournisseur.

## 4.5. Purification et séquençage

Les produits PCR sont ensuite purifiés en utilisant le kit « GenElute™ PCR Clean-Up kit » en suivant les recommandations du fabricant (Sigma-Aldrich®). L'ADN est élué avec un volume de 40µl de la solution d'éluion fournie dans le kit. Les ADNs sont conservés à -20°C jusqu'au séquençage.

Les produits de PCR purifiés sont séquencés en utilisant les amorces directes et les amorces inverses selon la méthode Sanger [326]. Les ADNs du *Cyt-b* sont séquencés à l'aide du BigDye® Terminator v1.1 Cycle Sequencing Kit (Applied Biosystems, Foster City, CA) dans des conditions standard par l'unité de génétique médicale – Université saint Joseph-Liban. Ces produits sont ensuite soumis à une électrophorèse sur un système de séquençage d'analyse génétique ABI 3130 (Applied Biosystems, Foster City, CA, États-Unis). Les électrogrammes sont analysés à l'aide du logiciel d'analyse séquentielle version 5.2 (Applied Biosystems, Foster City, CA, USA). Alors que le séquençage des ADNs de la D-Loop a été réalisé par la société GACT Biotech à Lyon.

On a effectué la création des séquences consensus avec le logiciel BioEdit version 7.1.9.

## 4.6. Analyses statistiques

### Répartition des rétrotypes et des haplogroupes

Les analyses statistiques des rétrotypes et des HPGs ont été réalisées avec le logiciel NCSS9 par un test non paramétrique de Kruskal-Wallis. La comparaison a été faite entre les pourcentages de différents troupeaux pour le rétrotype R4 et les différents types de zygote observés pour enJSRV-6 ainsi que les pourcentages des différents troupeaux pour un HPG donné. Les différences observées sont considérées comme significatives lorsque la p-value est strictement inférieure à 0,05 ( $p < 0,05$ ).

### Les Analyses phylogénétiques

- 1) Les séquences ont été alignées avec Muscle, seule la partie centrale de l'alignement a été conservée. L'arbre est construit avec l'algorithme Neighbor-Joining, la distance utilisée est le nombre de substitutions (sans correction). L'arbre a été enraciné avec la méthode du « midpoint rooting »
- 2) L'analyse polymorphique des séquences de l'ADN mitochondrial est calculée par le logiciel DnaSP version 6.0.73 :

- a) La diversité nucléotidique ( $\pi$ ) est le nombre moyen des différences nucléotidiques par site entre deux séquences, calculée selon l'équation 10.5 [327] :

$$\pi = \sum_{ij} x_i x_j \pi_{ij}$$

Où  $x_i$  et  $y_i$  sont les fréquences respectives des séquences  $i$  et  $j$ , alors que  $\pi_{ij}$  est le nombre de différence nucléotidique par site nucléotidique entre les séquences  $i$  et  $j$ .

- b) La diversité haplotypique ( $Hd$ ) est calculée selon l'équation 8.4 [327]

$$Hd = \frac{n}{n-1} (1 - \sum x_i^2)$$

Où  $x_i$  est la fréquence d'un haplotype dans une population et  $n$  est le nombre des

- 3) La divergence entre les populations est analysée par le logiciel DnaSP version 6.0.73 en absence d'indels avec deux paramètres : Le nombre moyen de la différence nucléotidique entre les populations ( $K_{xy}$ ), équation 10.3 [327]

$$K_{xy} = \frac{1}{n_x n_y} \sum_{i=1}^{n_x} \sum_{j=1}^{n_y} k_{ij}$$

Où dans deux population,  $x$  et  $y$  de taille  $n_x$  et  $n_y$  respectivement,  $k_{ij}$  mesure le nombre différence nucléotidique entre le  $i^{\text{ème}}$  haplotype de  $x$  et le  $j^{\text{ème}}$  haplotype de  $y$ .

et le nombre moyen des substitutions nucléotidiques par site entre les population ( $D_{xy}$ ) équation 10.20 [327].

$$D_{xy} = \sum_{ij} x_i y_j d_{ij}$$

Où dans deux population,  $x$  et  $y$ ,  $d_{ij}$  mesure le nombre de différence nucléotidique entre le  $i^{\text{ème}}$  haplotype de  $x$  et le  $j^{\text{ème}}$  haplotype de  $y$ .

Ces paramètres sont utilisés pour estimer la quantité de divergence des séquences d'ADN entre les populations.

- 4) Estimation des paramètres démographiques des populations

Le calcul de ces paramètres permet de nous renseigner sur les changements démographiques des populations comme par exemple si une population a subi une expansion ou bien un goulot d'étranglement. Pour cela, nous avons généré les courbes de distribution incompatible (mismatch distribution) pour visualiser la signature d'une expansion par DnaSP6 avec le modèle « population growth-decline ». La forme de la distribution donne des informations sur l'histoire de la population en particulier sur son expansion [328] : une distribution régulière en forme de cloche indique une soudaine expansion de population alors qu'une distribution multimodale, irrégulière indique une

taille constante de la population. Les tests de neutralité statistiques de Fu,  $F_s$ , équation 1 [329] et de Tajima, D, équation 38 [330] sont aussi réalisés : ces tests comparent la diversité observée d'une population à la diversité attendue selon la théorie de neutralité. La p-value de  $F_s$  est calculée par le programme de simulation coalescente (1000 simulations). Le seuil de signification de D est 0,05 ( $p < 0,05$ ) et alors que celui de  $F_s$  est 0.02 ( $p < 0.02$ ).

#### 5) Pouvoir discriminant des deux marqueurs mitochondriaux

Le pouvoir discriminant (D) est la probabilité moyenne de différencier deux individus.

Le calcul de D du Cyt-b et de la D-Loop a été effectué selon la formule suivante [331] :

$$D = 1 - \frac{1}{N(N-1)} \sum_{j=1}^s x_j (x_j - 1)$$

Où N est le nombre de séquences testées, s est le nombre des différents haplogroupes obtenus, et  $x_j$  est le nombre de séquences correspondant au haplogroupe  $j$  et en supposant que les séquences sont classées exclusivement dans un HPG donné. Une valeur de D=1 indique que la méthode de typage a pu distinguer chaque haplotype dans une population donnée, et inversement une valeur de D=0 indique que toutes les séquences sont identiques.

## Chapitre 5 : *Résultats*

---

## 5.1. Analyse génétique des populations de mouton de la race Awassi au Liban *via* les rétrovirus enJSRVs.

Comme décrit précédemment en introduction, certains enJSRVs sont insérés de manière polymorphique dans le génome des moutons domestiques de nos jours. Dans une étude précédente, l'utilisation de six de ces enJSRVs (enJSRV-18, enJSRV-7, enJSRV-8, enJSRV-15, enJSRV-16 et enJSRV-22 aussi appelé enJS5F16) comme marqueurs génétiques a permis de révéler certaines des étapes clés de la domestication du mouton et des migrations consécutives en Eurasie et Afrique [172]. Cette étude a été réalisée sur une collection de 1362 ADN génomiques provenant d'animaux de 133 races de mouton domestique et d'espèces sauvages proches d'Europe, d'Asie et d'Afrique.

Dans le cadre de ma thèse, ces enJSRVs insérés de manière polymorphique (à l'exception d'enJSRV-8 qui est restreint à l'Europe du Nord) ont été utilisés pour essayer d'établir des liens génétiques entre les différentes populations de mouton Awassi du Liban. De plus, similairement aux études précédentes, l'enJSRV-6 a été utilisé comme un contrôle de qualité d'ADN extrait vu qu'il était « endogénisé » avant la séparation des groupes Caprins et Ovins et donc n'étant pas inséré de manière polymorphique dans le génome des moutons domestiques [172, 332].

Une fois intégré dans le génome, les ERVs peuvent subir des recombinaisons homologues entre les LTR 5' et 3' flanquant le provirus ce qui aboutit à l'excision d'une partie de l'ADN proviral et à la formation de ce qu'on nomme un « solo-LTR » composé uniquement de la séquence complète d'un LTR encadré par le site d'intégration dupliqué à chaque extrémité (Figure 5-1Ab). Ainsi, pour révéler la présence ou l'absence de chacun de ces enJSRVs ou de solo-LTR trois types de PCRs ont été réalisées (Figure 5-1A): i) des PCRs dites « Provirus » permettant d'amplifier spécifiquement l'extrémité 5' et 3' (à partir du gène *gag* en 5' et *env* en 3') de chacun de ces enJSRVs insérées à un locus particulier dans le génome grâce à des amorces situées dans le génome ovin de part et d'autre du site d'insertion de chaque provirus (Figure 2.1A-a); ii) des PCRs dites « solo-LTR » permettant d'amplifier l'extrémités 5' et 3' des provirus ou de solo-LTR (Figure 2.1A-b) et iii) des PCRs dites « locus vide » permettant d'amplifier soit les loci génomiques sans provirus (environ 300-400pb) soit les solo LTR (environ 600-800pb) grâce à des amorces situées dans le génome de part et d'autre des sites d'insertion de chaque provirus (Figure 2.1A-c).



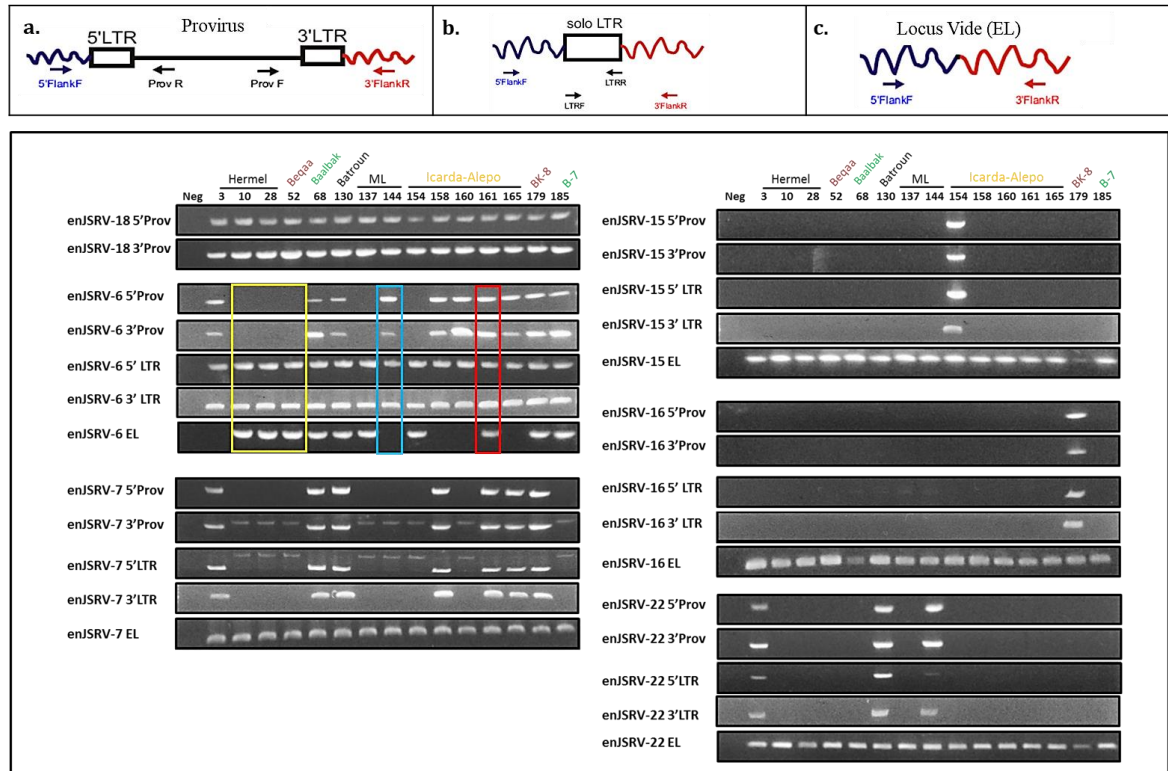


Figure 5-1: A) Schémas représentant les trois types de PCR : a) provirus, b) solo-LTR et c) locus vide réalisés pour la détection de chaque enJSRV : enJSRV-18, -6, -7, -15, -16 et -22 (voir le texte pour plus de détails). B) Gels représentatifs des résultats des PCR obtenus pour chaque enJSRV testé sur des échantillons provenant de différentes régions du Liban. « Neg » correspond au contrôle négatif de PCR réalisé sans ADN génomique. Les gels encadrés correspondent aux différents types de zygote observés pour enJSRV-6.

Ces analyses PCRs ont été réalisées sur tous les 277 échantillons d'ADN collectés et ces résultats sont illustrés en Figure 5-1B avec des échantillons représentatifs de l'étude réalisée. On observe une corrélation parfaite des échantillons positifs pour les PCRs 5' et 3' « provirus » et 5' et 3' « solo-LTR » pour tous les enJSRVs (à l'exception d'enJSRV-6 qui sera discuté plus tard) indiquant la présence d'un provirus complet dans ces génomes. Pour ces cinq enJSRVs, on note également l'amplification pour tous les échantillons du locus vide (vérifié par séquençage) et indiquant l'hétérozygotie d'insertion de ces provirus.

Les 277 échantillons ont été testés pour les cinq enJSRVs (-18, -7, -15, -16 et -22) et des cartes de densité ont été réalisées afin de déterminer la fréquence et la répartition de chaque enJSRV dans les animaux testés au Liban. Comme présenté sur la figure 2.2, enJSRV-18 est présent dans tous les échantillons testés (100%) tandis que 38 échantillons (~15%) renferment l'enJSRV-7 (Figure 5-2A et B). La présence d'enJSRV-18 dans le génome ovin est caractéristique des moutons dits « modernes » c'est-à-dire issues du deuxième évènement de domestication tandis que les moutons dits « primitifs », issus de la première

vague de domestication (Urial, Mouflon méditerranéen et mouflon asiatique) sont caractérisés par l'absence d'enJSRV-18 et la présence d'enJSRV-7 [172].

**D'une part ces données confirment que les moutons de race Awassi du Liban sont des moutons « modernes » issus de la deuxième vague de migration.** D'autre part, on peut noter qu'enJSRV-7 est plus présent dans le génome des populations de mouton située à l'Ouest et au Sud du Liban. La double présence d'enJSRV-18 et enJSRV-7 est caractéristique des moutons domestiques présents au pourtour du bassin méditerranéen (défini comme un rétrotype méditerranéen). **Ce rétrotype méditerranéen est bien présent dans le génome des populations des moutons Awassi du Liban**, ce pays étant méditerranéen et fut un lieu de passage incontournable des routes commerciales de l'antiquité passant par la mer méditerranée. Les autres enJSRVs sont retrouvés à une fréquence plus faible dans le génome des animaux testés : l'enJSRV-16 dans 27 animaux (10,6%), enJSRV-15 dans 17 animaux (6,7%) et enJSRV-22 dans 14 animaux (5,5%). enJSRV-15 semble être plus fréquemment retrouvé au Nord-Ouest, enJSRV-16 au sud et enJSRV-22 au Nord-Est (Figure 5-2C-E).

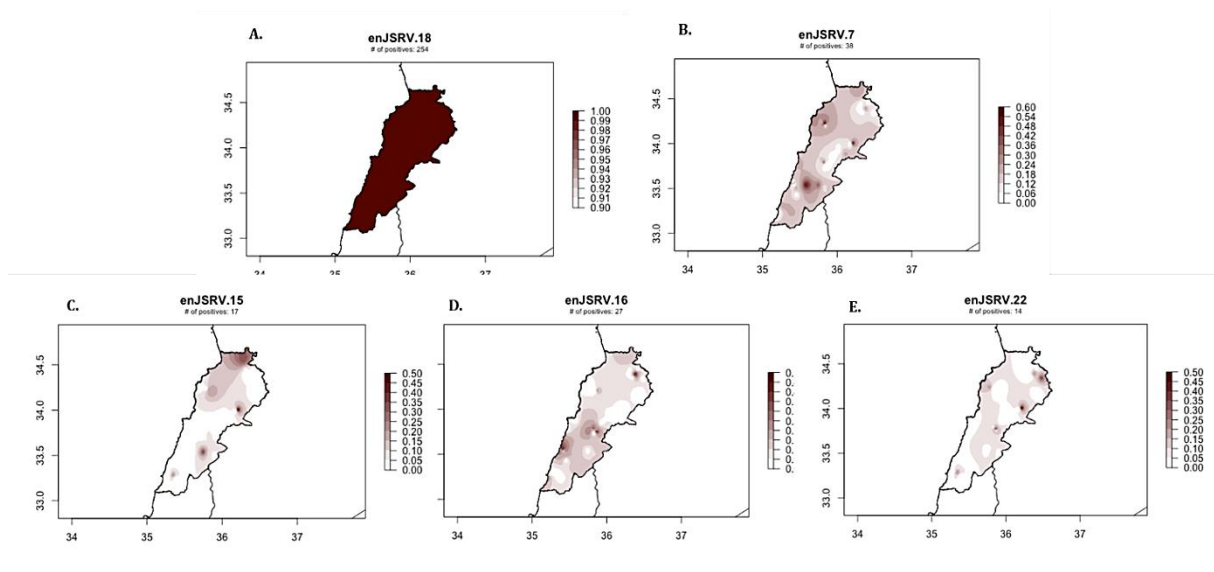


Figure 5-2: Cartes de densité représentant la fréquence et la répartition des enJSRVs polymorphiques dans les troupeaux de mouton de la race Awassi au Liban. A) enJSRV-18 ; B) enJSRV-7 ; C) enJSRV-15 ; D) enJSRV-16 et D) enJSRV-22

Concernant enJSRV-6 (utilisé comme contrôle de qualité), certains échantillons (15 échantillons) apparaissent positifs pour les PCRs 5' et 3' LTR mais négatif pour les PCRs 5' et 3' provirus (Figure 5-1B et Figure 5-2A). De plus, certains échantillons sont positifs pour la PCR « locus vide » (échantillons positifs ou non pour le provirus) ; mais avec un produit PCR de 750pb au lieu de 302bp qui correspond pour un locus vide suggérant la

présence de solo-LTR d'enJSRV-6 dans certains échantillons ce qui a pu être vérifiée par séquençage (Figure 5-3)

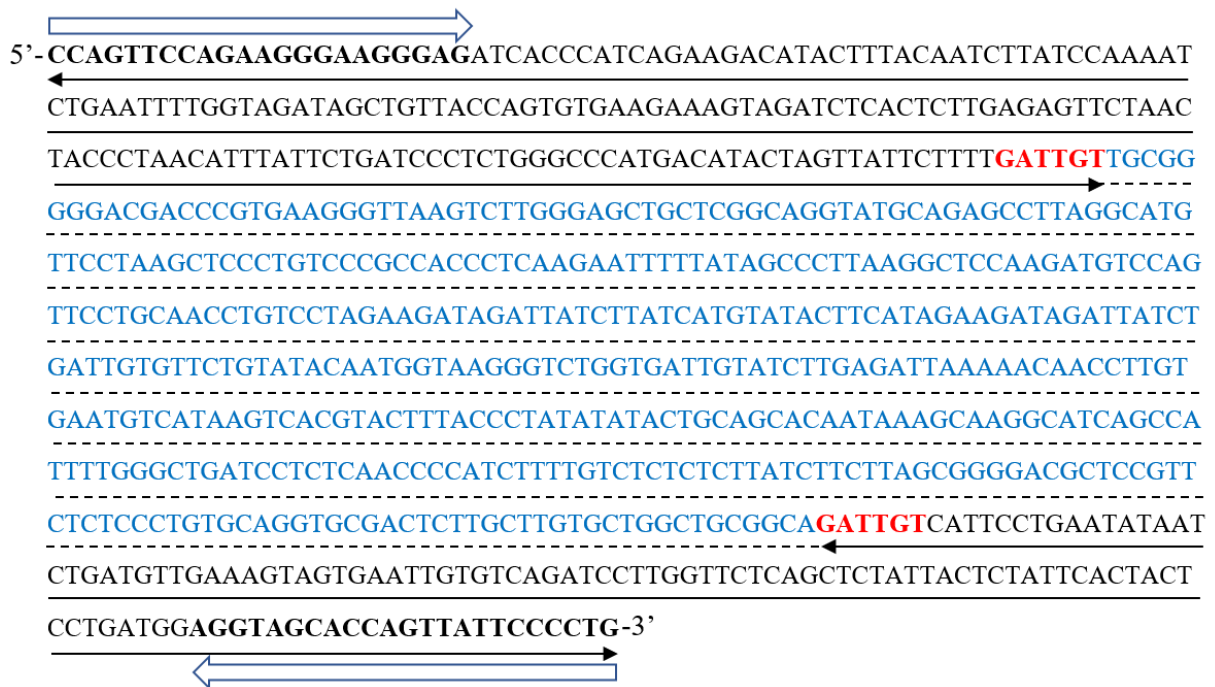


Figure 5-3 : Alignement de la séquence Solo-LTR d'enJSRV-6 dont la taille est 750pb : Les amorces, directe et inverse, sont représentées en caractères gras et par une flèche. Les séquences génomiques flanquantes (en noir) du Solo-LTR en 5' et 3' de taille 189pb et 119pb respectivement (T=308pb) limitant un LTR (bleu) de 442pb. GTTGT : site génomique d'intégration dupliqué à chaque extrémité de 6pb (rouge).

enJSRV-6 est un des plus anciens enJSRVs testés pour cette étude et nous n'avons jamais amplifié *via* la PCR « locus vide » de produit correspondant au locus vide (302pb). Ces résultats montrent donc qu'enJSRV-6 est ou a été inséré dans tous les ADNs génomiques testés et que pour certains, des recombinaisons ont eu lieu après son intégration pour former des solo-LTR. Les données de ces différentes PCRs pour enJSRV-6 permettent de définir trois cas de figure quant à la zygote d'enJSRV-6 (Figure 5-4A) : homozygote pour le provirus, hétérozygote (provirus et solo-LTR) ou homozygote pour le solo-LTR que j'appellerai dans la suite du manuscrit et pour plus de simplicité « Homo/Prov », « Hétéro/Prov-LTR » et « Homo/LTR ». Ainsi, des cartes de densité ont également été réalisées pour ces trois cas de figure (Figure 5-4B).

A.

enJSRV-6	Homozygote Provirus	Hétérozygote Provirus/Solo LTR	Homozygote Solo LTR
Provirus	Positif	Positif	Négatif
Solo LTR	Positif	Positif	Positif
Locus vide	Négatif	Positif	Positif
Nombre d'échantillons	126	136	15

B.

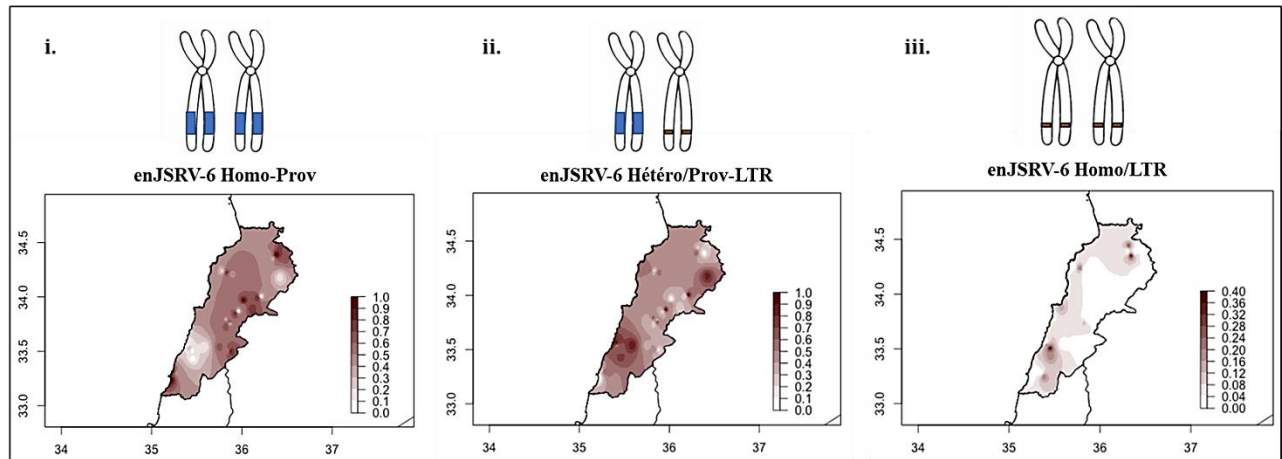


Figure 5-4:A) Tableau récapitulatif des résultats des PCR d'enJSRV-6 en termes de zygote. B) Cartes de densités des trois cas de figure pour enJSRV-6 : i. homozygotes pour le provirus « Homo-Prov », n=120 ; ii. Hétérozygotes (provirus et solo LTR) « Hétéro/Prov-LTR », n=121 ; iii. Homozygote pour le solo LTR « Homo/LTR », n=13.

Globalement, la fréquence des « Homo/Prov » et celle des « Hétéro/Prov-LTR » sont équivalentes (respectivement 47,24% et 47,64%) tandis que les Homo/LTR sont retrouvés à une fréquence beaucoup plus faible dans le génome des moutons Awassi libanais (5,12%). Néanmoins, en termes de répartition, on peut observer une plus forte présence des Hétéro/Prov-LTR et des Homo/LTR au Sud-Ouest et Nord-Est du Liban (Figure 5-4A. ii et iii) ce qui corrèle avec une fréquence plus faible des Homo/Prov dans cette région (Figure 5-4B-i). Le fait que les Hétéro/Prov-LTR et Homo/LTR se retrouvent à une fréquence plus forte dans les mêmes régions a du sens si l'on assume que les Homo/LTR sont issus du croisement de deux parents étant *a minima* Hétéro/Prov-LTR.

Dans le but d'analyser en détail la distribution de ces enJSRVs dans le génome des populations d'Awassi, les rétrotypes (combinaison d'enJSRV-18 avec chacun des autres enJSRVs) ont été déterminés en adoptant la nomenclature déjà publiée [172] (R2= enJSRV-18, R4= enJSRV-18+enJSRV-7, R6= enJSRV-18+enJSRV-22). Dans ce travail, deux autres rétrotypes non référencés par Chessa et collègues ont également été analysés et nommés comme suit : RA= enJSRV-18+enJSRV-15 et RB= enJSRV-18+enJSRV-16. La

fréquence et répartition du rétrotype R2 avec chacun des autres rétrotypes ont été déterminées par région et par troupeau. Comme présenté figure 2.4, le rétrotype R4 (enJSRV-18 et enJSRV-7) est présent dans environ 15% des échantillons testés avec des pourcentages qui peuvent varier considérablement d'une région à l'autre et d'un troupeau à un autre (de 0% à 50%) (Figure 5-5A et B).

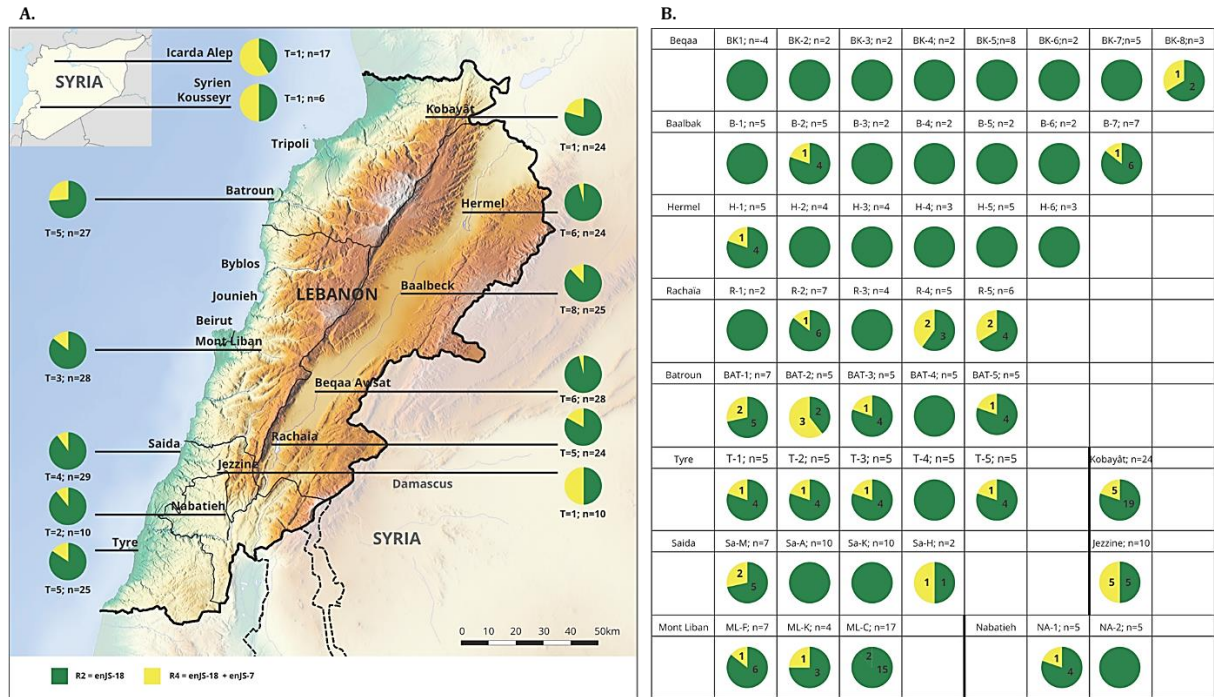


Figure 5-5: Fréquence et distribution géographique des rétrotypes R2 (en vert) et R4 « méditerranéen » (en jaune). A) par région, B) par troupeau. Encadré en rouge, les populations du Nord-Est du Liban (Beqaa, Baalbak et Hermel) présentant un rétrotype R4 plus faible que les autres.

Globalement, on peut observer que ce rétrotype méditerranéen R4 est retrouvé de manière significative plus faiblement dans les ADNs génomiques des troupeaux de moutons situés au Nord-Est du Liban et plus précisément à la Beqaa, à Baalbak et à Hermel tandis qu'il est clairement plus présent et de manière plus homogène entre les différents troupeaux du côté de la méditerranée (côté Ouest) ainsi que dans la région de Rachalaïa située au Sud-Est ( $p=0,005377/p<0,01$ , test non paramétrique Kruskal-Wallis ; Figure 5-5A et B). Comme précédemment décrit, R4 est un rétrotype plus fréquemment retrouvé dans le génome des populations de moutons domestiques situées au pourtour du bassin méditerranéen. **Il est donc intéressant de noter qu'au Liban, ces populations de mouton d'Awassi localisées au Nord-Est et donc ayant un accès plus restreint à la mer méditerranée que les autres populations (i.e. du fait de la chaîne centrale de montagne, le Mont Liban, coupant le pays en deux), présente le rétrotype R4 plus faiblement.** Enfin, on remarque que les deux populations syriennes possèdent une rétrotype R4 beaucoup plus fort que des

populations libanaises de mouton Awassi avec plus de 50% de R4 ( $p=0,000164$ , test non paramétrique Kruskal-Wallis). Les autres rétrotypes sont apparus comme étant un peu moins informatifs que le R4.

Le rétrotype RA (enJSRV-18+enJSRV-15) est absent des populations de mouton de la région de Hermel, et de toute la région qui s'étend de Mont Liban jusqu'au Sud à l'exception d'un échantillon d'ADN génomique à Tyre (Figure 5-6A). Bien qu'il se trouve de façon générale en faible pourcentage sur l'ensemble des échantillons testés, **RA est plus fréquent chez moutons Awassi issus du troupeau de la région de Kobayât. Ces données mettent aussi en évidence un troupeau de la région de Baalbak qui, même s'il ne s'agit que de deux animaux, semble différent des autres avec la présence de RA (B-5).** Néanmoins, les informations obtenues sur ce troupeau ne permettent pas à l'heure actuelle d'expliquer de manière simple cette différence observée. Il pourrait s'agir de la provenance des animaux utilisés pour établir ce troupeau qui peut être différente des autres ou bien encore, étant ouvert, de l'intégration d'animaux différents par rapport aux autres troupeaux de Baalbak.

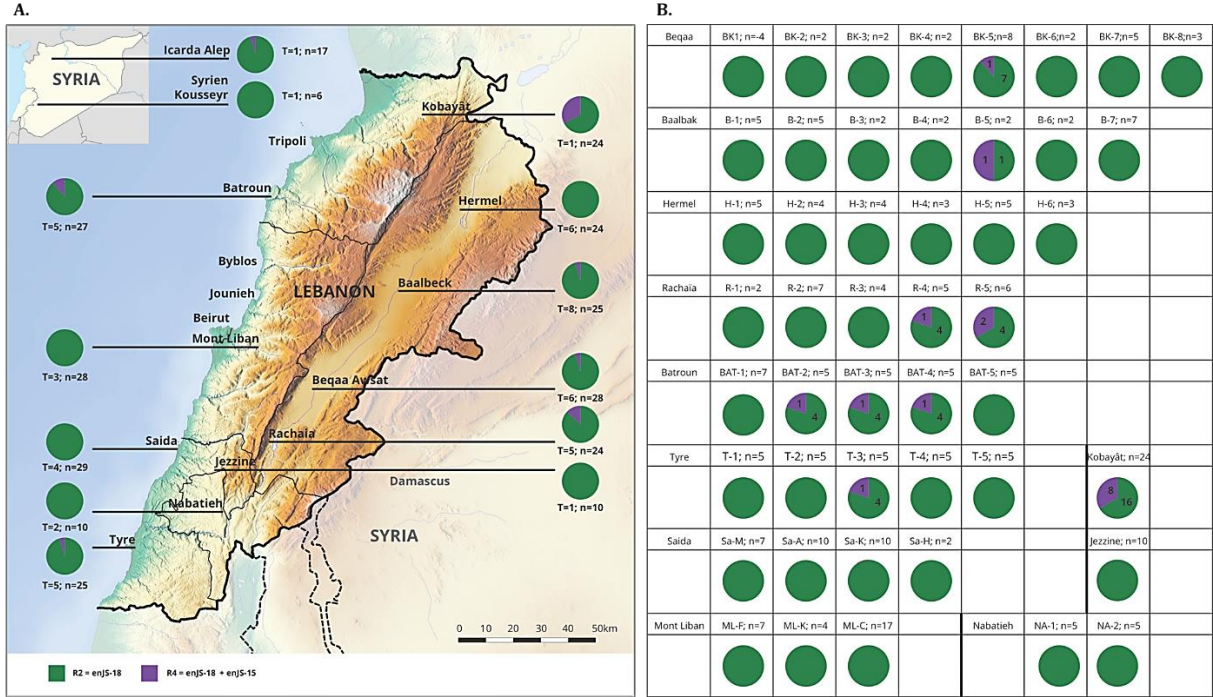


Figure 5-6:Fréquence et distribution géographique des rétrotypes R2 (en vert) et RA (en violet). A) par région, B) par troupeau.

Le rétrotype RB (enJSRV-18+enJSRV-16) est le plus fréquent après R2 et R4 avec une prévalence globale de 10,63% des échantillons libanais. RB est absent à Baalbak et se trouve dans 27,59% des échantillons de Saïda. Il est aussi présent avec une prévalence de 14,29% au Mont Liban et à la Beqaa (Figure 5-7A). La prévalence la plus élevée de RB est détectée dans le troupeau syrien d'Al Qusayr avec une valeur de 66,67%.

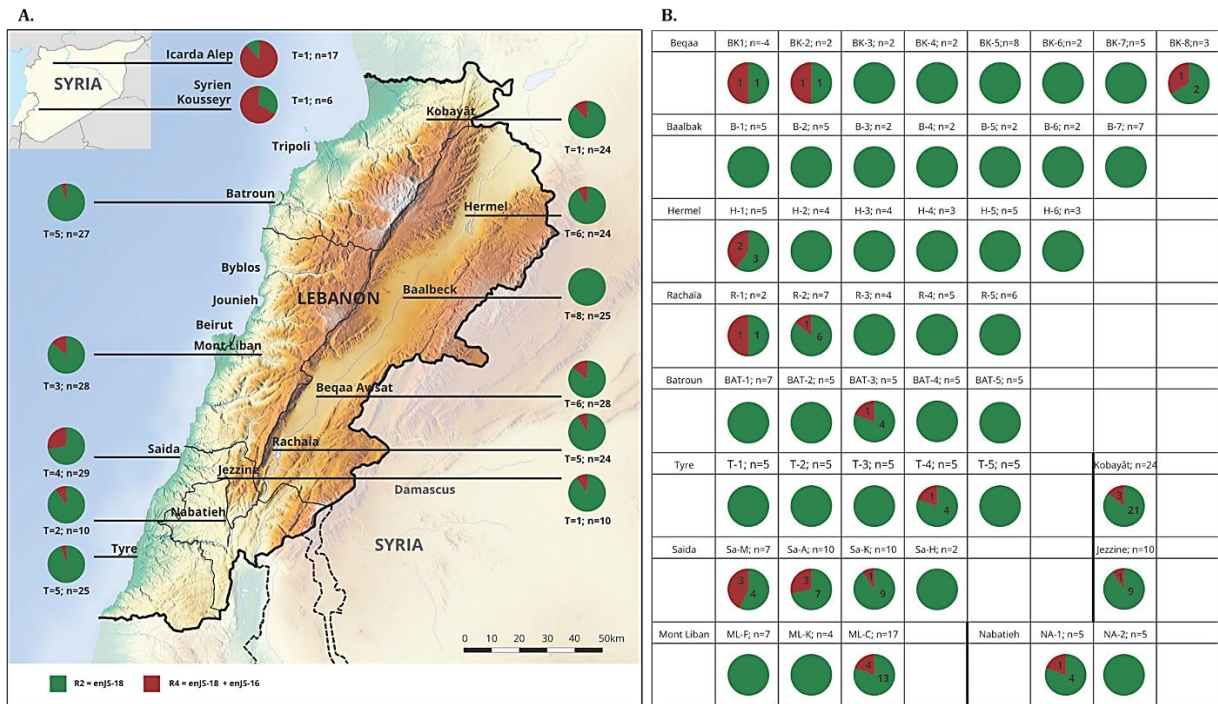


Figure 5-7: Fréquence et distribution géographique des rétrotypes R2 (en vert) et RB (en marron). A) par région, B) par troupeau.

D'après les études précédentes, le rétrotype R6 est spécifique des pays scandinaves et des Îles Britanniques où il est retrouvé à une haute fréquence [172] ainsi que dans le Nord de la Caucase et dans la région de Pamir [332]. Il n'est donc pas étonnant que ce rétrotype soit le plus rare dans les populations libanaises de la race Awassi du Liban avec une prévalence de 5,51%. En effet, seuls 14 échantillons présentent le rétrotype R6 et ce dans 14 troupeaux différents répartis uniformément sur l'ensemble du Liban (Figure 5-8A et B). On peut cependant observer que le troupeau B-5 de la région de Baalbak qui, comme précédemment décrit, diffère des autres troupeaux de la région avec la présence de RA (Figure 5-6) se distingue aussi par la présence de R6 renforçant l'idée que ce troupeau est génétiquement distinct des autres en tout cas d'un point de vue « Rétrotype ».

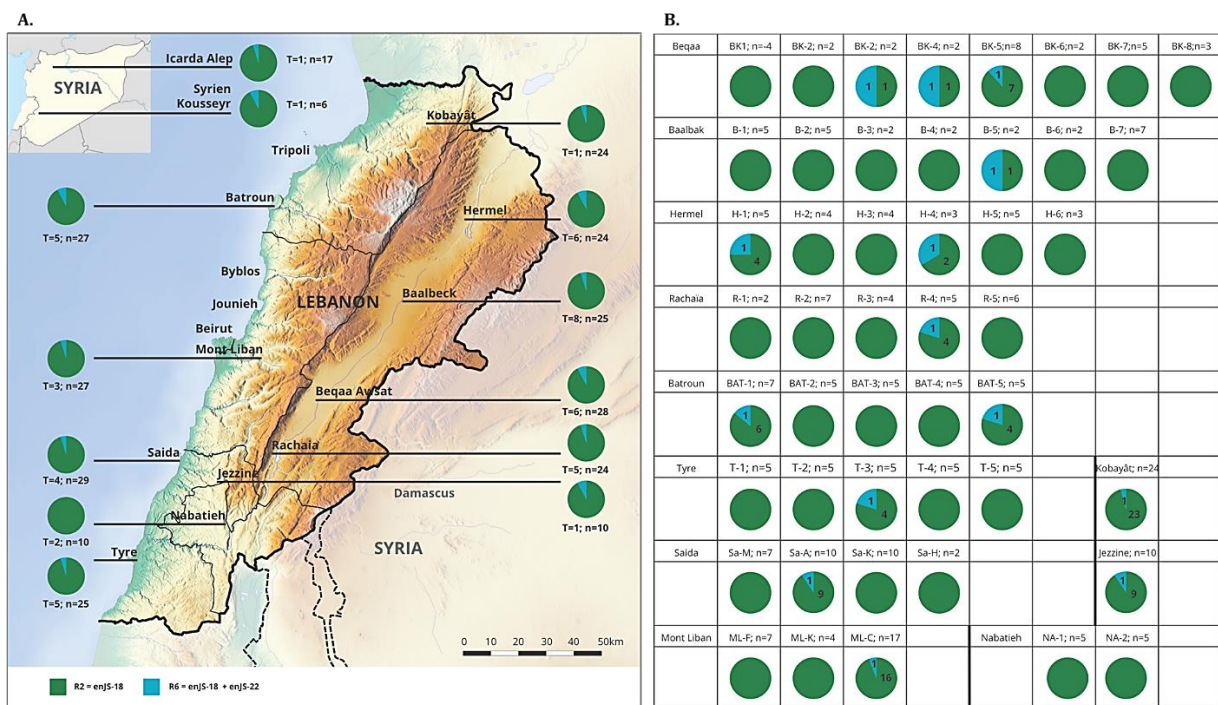


Figure 5-8: Fréquence et distribution géographique des rétrotypes R2 (en vert) et R6 (en bleu). A) par région, B) par troupeau.



L'analyse globale avec l'ensemble des rétrotypes retrouvés dans le génome des moutons Awassi testés a été réalisée et est présentée dans la Figure 5-9.

Le rétrotype R2, défini par la seule présence d'enJSRV-18, prédomine dans toutes les régions (68,11% des échantillons). De même, on observe que le rétrotype R4 (11% des échantillons) est présent dans la plupart des régions à l'exception de deux (Beqaa Al Awssat et Hermel).

Le troisième rétrotype le plus fréquent est RB présent dans 7,87% des échantillons. Les deux rétrotypes, RA (enJSRV-18 et enJSRV15) et R6 (enJSRV-18 et enJSRV22), sont présents dans 4,33% et 3,15% des animaux respectivement. Les autres rétrotypes (issus de combinaison de plus de trois enJSRVs) sont relativement rares avec des pourcentages allant de 1,57% pour RF (enJSRV-18 + enJSRV-7 + enJSRV-16) à 0,39% pour RD (enJSRV-18+ enJSRV-15 + enJSRV-16). RG (enJSRV-18+ enJSRV-7 + enJSRV-16+ enJSRV-22) et RH (enJSRV-18+ enJSRV-16+ enJSRV-22) sont qu'en à eux retrouvés dans un seul animal chacun. Les trois troupeaux présentant le plus de diversité en termes de rétrotype (au moins cinq différents) étant le troupeau de kobayât et les deux troupeaux syriens.

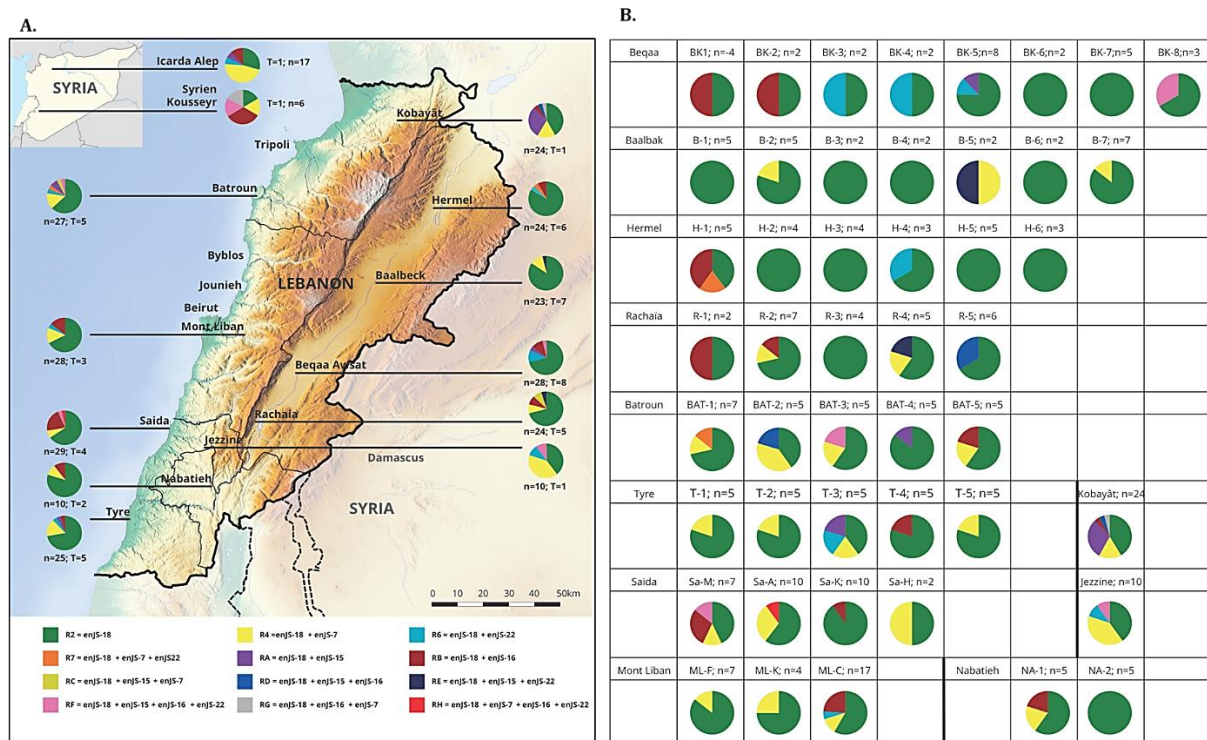


Figure 5-9: Fréquence et distribution géographique de l'ensemble des rétrotypes retrouvés dans le génome des populations d'Awassi au Liban A) par région, B) par troupeau.

Enfin, le même type d'analyse a été effectué pour enJSRV-6 et les trois cas de figure retrouvés pour ce rétrovirus endogène en termes de zygote (Figure 5-10). Comme précédemment observé sur les cartes de densité (Figure 5-4B), on observe une fréquence plus importante des « Hétéro/Prov-LTR » chez les populations de mouton Awassi du Sud-Ouest du Liban (Saida, Jezzine, Nabatieh) et ce de manière homogène dans les différents troupeaux ( $p=0,000578$ , test non paramétrique Kruskal-Wallis ; Figure 5-10A). Ce taux élevé d'Hétéro/Prov-LTR corrèle avec une plus forte proportion d'animaux portant des solo-LTR dans ces trois régions (Saida, Jezzine, Nabatieh) par rapport aux autres ( $p=0,042320$ , test non paramétrique Kruskal-Wallis). Le cas extrême étant les deux troupeaux de Nabatieh ( $n=10$ ) qui sont tous « Hétéro/Prov-LTR » ou « Homo/LTR » mais en aucun cas « Homo/Prov » (Figure 5-10B). Dans la carte de densité (Figure 5-3B), (le point au Nord-Est (Hermel) des solo-LTR n'est pas visible du fait qu'on a plusieurs coordonnées GPS s très proches).

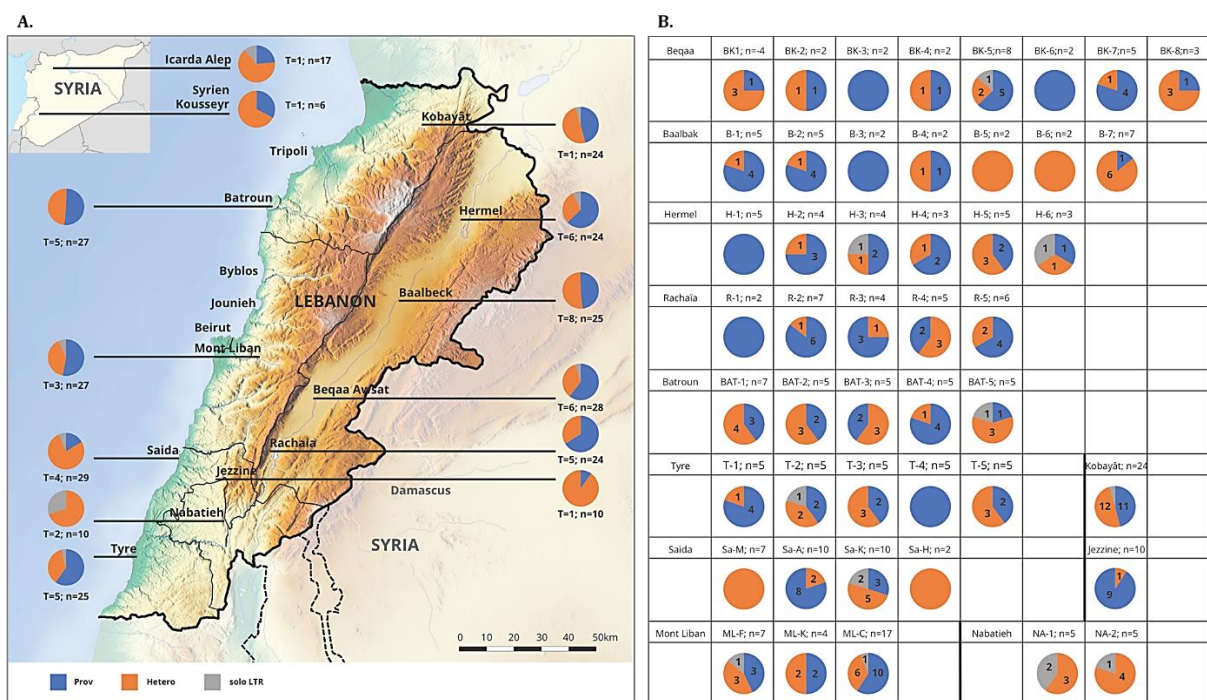


Figure 5-10: Fréquence et distribution géographique des différents types de zygote retrouvés pour enJSRV-6 dans le génome des populations d'Awassi au Liban A) par région, B) par troupeau. « Homo/Prov » (Bleu), « Hétéro/Prov-LTR » (orange) et « Homo/LTR » [333].

## **5.2. Analyse génétique des populations de mouton de la race Awassi au Liban avec l'ADN mitochondrial.**

Dans le but de compléter la caractérisation génétique des moutons Awassi, deux gènes de l'ADN mitochondrial ont été utilisés comme marqueurs génétiques : le cytochrome b et la D-Loop. Dans cette étude, 99 échantillons pour le *Cyt-b* et 117 pour la D-Loop ont été analysés par amplification PCR puis séquençage.

### **5.2.1. Le cytochrome-b : Cyt-b**

La taille des séquences analysées pour *Cyt-b* est de 1034pb (G+C=42,3%) ce qui couvre presque entièrement la séquence du gène *Cyt-b* qui fait 1140pb. A partir de ces séquences, l'identification des haplogroupes a été réalisée en utilisant comme séquences de référence celles de deux études précédentes [13, 320] soit 389 séquences au total correspondant aux cinq HPG identifiés à ce jour. Un arbre phylogénétique selon l'algorithme Neighbor-Joining, basé sur 488 séquences de *Cyt-b* (99 Awassi et 389 séquences de référence) a été construit pour déterminer les haplogroupes mitochondriaux des échantillons des différentes populations d'Awassi analysées (Figure 5-11).

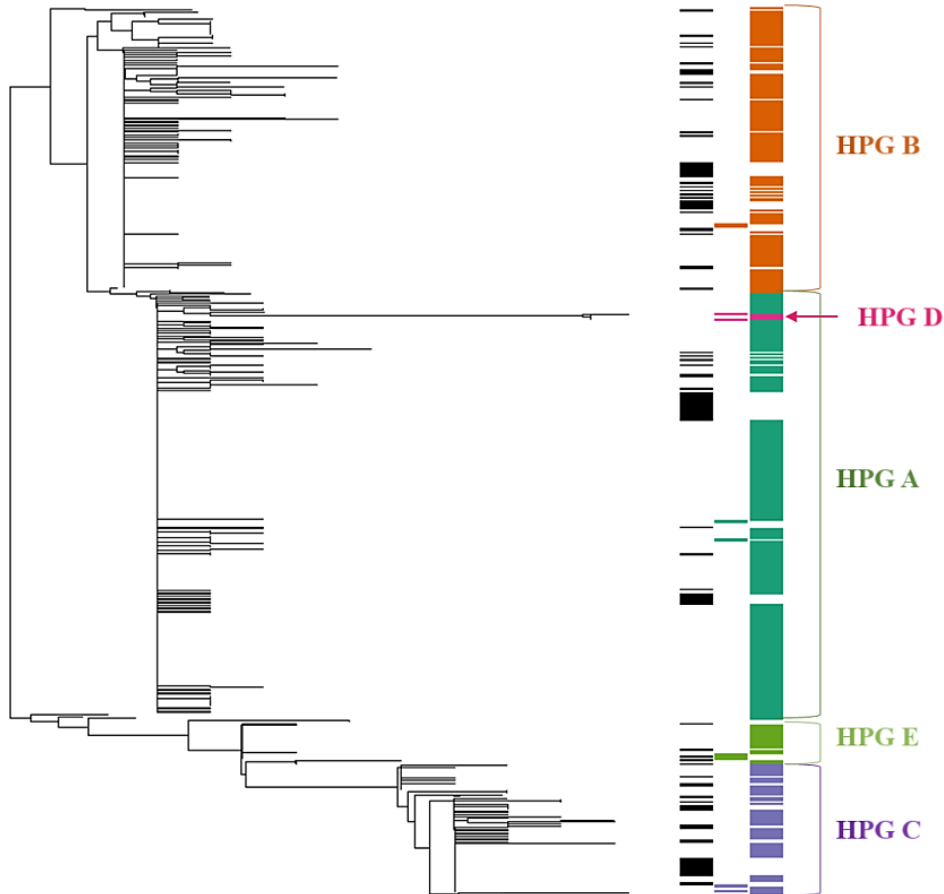


Figure 5-11: Arbre phylogénétique des séquences de *Cyt-b* (n=488) montrant les cinq haplogroupes maternels ovins. Les traits noirs représentent les échantillons d'ADN issus des populations de mouton Awassi du Liban analysées dans notre étude. Les colonnes en couleurs sont les séquences utilisées comme références : Demirci *et al.*, 2013 (colonne du milieu) et Rezaei *et al.*, 2010 (colonne à droite).

L'analyse des séquences *Cyt-b* a permis l'identification de quatre haplogroupes présents dans les populations de mouton de la race Awassi du Liban analysées : HPG A, HPG B, HPG C et HPG E. Le nombre de séquences pour chaque HPG est reparti comme suit : HPG A n= 32, l'HPG B n= 38, l'HPG C n=24 et l'HPG E n=5 avec l'absence de séquences correspondant à l'HPG D.

#### Polymorphisme des séquences du *Cyt-b*

Le polymorphisme des séquences de *Cyt-b* des moutons Awassi libanais a ensuite été analysé (Tableau 5-1). Les 99 séquences renferment 57 sites polymorphiques (S) dont 20 sont phylogénétiquement informatifs et qui ont abouti à la définition de 45 haplotypes (Annexe 2) avec une diversité haplotypique (i.e. La probabilité de sélectionner au hasard 2 séquences différentes) [327] ( $H_d = 0,917 \pm 0,00029$ ) et une diversité nucléotidique (i.e. nombre moyen des différences nucléotidiques par site entre deux séquences choisies au hasard) [334]  $\pi_t = 5,87 \times 10^{-3}$ .

Concernant les HPGs, l'HPG E présente la plus grande diversité nucléotidique ( $\pi_E=1.74 \times 10^{-3}$ ) et haplotypique ( $Hd_E=0,900 \pm 0,026$ ). Ces données indiquent que l'HPG E est l'haplogroupe le plus divers dans les échantillons d'Awassi analysés sachant d'autant plus que la diversité haplotypique ( $Hd$ ) dépend de la taille de la population [328] (i.e. nombre de séquences par haplogroupe) et que le nombre de séquences HPG E est plus faible par rapport aux autres HPGs. Concernant les valeurs de la diversité nucléotidique ( $\pi$ ) (qui ne dépendent pas de la taille de la population), elles sont comparables avec celles précédemment obtenues par Meadows et *al.* 2007 pour les HPG A ( $1,07 \times 10^{-3}$ ), B ( $1,25 \times 10^{-3}$ ) et C ( $1,23 \times 10^{-3}$ ). Par rapport à cette étude, la seule différence se situe au niveau de HPG E où la diversité nucléotidique ( $\pi_E=1,74 \times 10^{-3}$ ) est 3 fois plus grande que celle observée dans les échantillons de Meadows et *al.* Cet étude a décrit huit races de moutons en Turquie (Karakas, Norduz, Morkaraman, Tuj, Cine-Capari, Sakiz, Karya et Karayaka) et une race d'Israël (Improved Awassi appelé Assaf),

En combinant ces données, on peut donc déduire que l'HPG E porte le taux de polymorphisme le plus élevé parmi ceux retrouvés dans cette analyse. **L'Awassi renferme une diversité génétique élevée suggérant un rôle du Liban dans la domestication des moutons.**

Tableau 5-1: Variations phylogénétiques des haplogroupes mitochondriaux des moutons de la race Awassi au Liban

	HPG A	HPG B	HPG C	HPG E
Séquences	32	38	24	5
Sites Variables (S)	15	19	11	4
Diversité Nucléotidique ( $\pi$ )	$1,07 \times 10^{-3}$	$1,25 \times 10^{-3}$	$1,23 \times 10^{-3}$	$1,74 \times 10^{-3}$
Haplotypes	15	16	10	4
Diversité Haplotypique (Hd)	$0,724 \pm 0,0078$	$0,724 \pm 0,0064$	$0,746 \pm 0,0082$	$0,900 \pm 0,026$

#### Divergence entre les populations

Les nombres moyens de différence nucléotidique ( $K_{xy}$ ) et de substitutions nucléotidiques par site ( $D_{xy}$ ) entre les populations d'Awassi appartenant aux différents HPGs ont ensuite été calculés et comparés pour analyser leurs relations phylétiques (Tableau 5-2). Ces paramètres représentent la diversité des séquences entre deux HPGs. Les valeurs de  $K_{xy}$  les

plus faibles sont entre l'HPG A et l'HPG B ( $K_{xy}=3,168$ ) et entre l'HPG C/HPG E ( $K_{xy}=5,33$ ), d'autre part la plus grande valeur de  $K_{xy}$  est observée entre HPG B et HPG C ( $K_{xy}=12,886$ ) divisant ainsi les haplotypes en deux groupes : HPGA/HPGB et HPGC/HPGE. Cette division est clairement visible avec l'arbre phylogénétique construit selon la méthode statistique de Neighbor-Joining (NJ) (Figure 5-12).

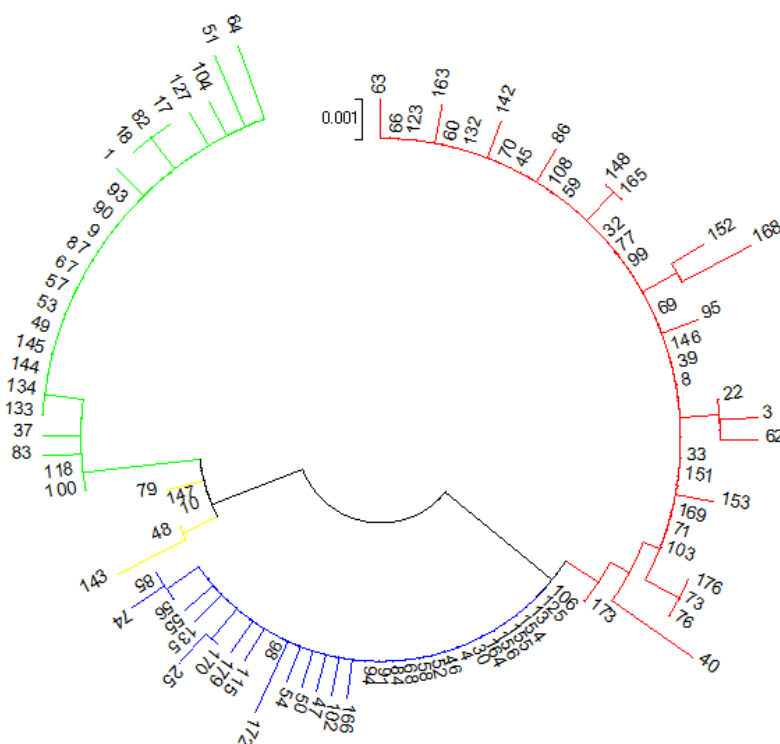


Figure 5-12. Arbre phylogénétique NJ des haplotypes basés sur les séquences du *Cyt-b* des échantillons Awassi : HPG A en bleu, HPG B en rouge, HPG C en vert, et HPG E en jaune. Cet arbre montre clairement la divergence des HPG A/HPG B d'une part des HPG C/HPG E d'autre part.

Les valeurs de  $D_{xy}$  obtenues suivent la même tendance que celles des  $K_{xy}$  avec la valeur de substitution la plus faible (0,306%) pour HPG A/HPG B et celle la plus grande (1,246%) pour HPG B/HPG C indiquant que l'HPG C est le clade le plus divergent. Et comme HPG E est proche de HPG C ( $D_{xy}= 0,516\%$ ), c'est deux HPG divergent des deux autres (HPG A et HPG B) (Figure 5-12). Les nombres moyens de substitution par site,  $D_{xy}$ , entre les HPGs de notre population sont comparables avec ceux obtenus par Meadows 2007 (valeurs entre parenthèse dans le tableau). La divergence des haplotypes domestiques de HPGA/HPGB des haplotypes de HPGC/HPGE est claire. En effet, la distance génétique entre HPG A et HPG B est faible, il s'agit donc peut-être de deux événements de domestication (1 pour A/B et 1 pour C et à priori E) qui ont impliqués des ancêtres sauvages très divergents. Enfin, **vu que les haplogroupes A et B sont les seuls HPGs dans les populations des moutons en**

**Europe et que ces HPGs existent en Asie avec les HPG C et E, il est probable que chaque groupe (A/B et C/E) correspond à une vague de migration différente ou provienne de centre de domestication distinct.**

Tableau 5-2 : Les nombres moyens de différence nucléotidique (Kxy) et de substitutions nucléotidiques par site (Dxy) entre les populations des différents haplogroupes des moutons Awassi libanais. Les valeurs de Dxy entre parenthèse sont extraites de l'étude de Meadows *et al.*, 2007.

	HPG A		HPG B		HPG C	
	Dxy %	Kxy	Dxy %	Kxy	Dxy %	Kxy
<b>HPG B</b>	0,306 (0.31)	3,168				
<b>HPG C</b>	1,054 (1.08)	10,89	1,246 (1.29)	12,886		
<b>HPG E</b>	0,731 (0.70)	7,563	0,921 (0.91)	9,521	0,516 (0.44)	5,33

#### Expansion démographique des haplogroupes

Les séquences de l'ADNmt sont souvent utilisées pour détecter la croissance historique des populations. La signature moléculaire de l'expansion d'une population se traduit par le fait que toutes ses séquences (ADNmt) ont le même âge c'est-à-dire le degré de divergence entre les séquences est semblable. Ceci est dû au fait que durant une croissance rapide de la population, la plupart des nouvelles mutations sont conservées et par conséquent, ces mutations permettent l'estimation de la date de l'initiation de la croissance [2]. Pour une population en expansion, si on dessine la distribution des divergences par paires de séquences, on obtient une courbe régulière en forme de cloche (Distribution normale). Pour une population en déclin ou stationnaire, les mutations sont perdues rapidement et au hasard, ce qui génère une distribution irrégulière ou multimodale [335].

Afin de savoir si les HPGs de la race Awassi du Liban ont subi une expansion de population dans le passé, les « mismatch distributions », aussi appelées la répartition des différences de site nucléotidique observées par paire de séquences (the distribution of the observed pairwise nucleotide site differences), ont été établies pour chaque HPGs (Figure 5-13) : En abscisse la différence nucléotidique entre deux séquences appelée différence deux à deux (pairwise difference) et en ordonnée les fréquences, observées et attendues, de chacune de ces différences nucléotidiques. La courbe de la « mismatch distribution » est en forme de cloche pour les haplogroupes A, B et C mais elle est irrégulière pour l'HPG E.

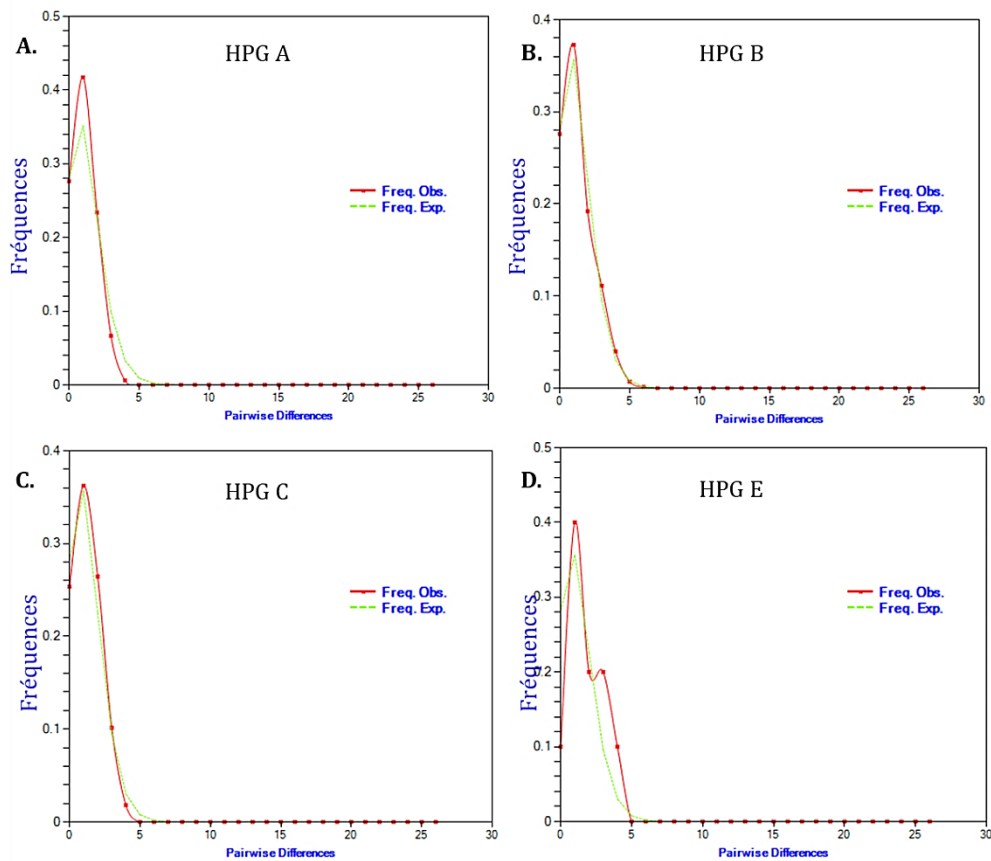


Figure 5-13: La « mismatch distributions », pour chaque HPG identifié au sein des échantillons de mouton d'Awassi libanais. Les graphiques de distribution de l'HPG A (A), HPG B (B), HPG C (C) et HPG E (D) sont représentés en fonction des fréquences observées (en rouge) et en fonction des fréquences attendues (en vert fluo).

Nous avons eu recours à deux tests de neutralité couramment utilisés pour analyser la signification statistique des événements démographiques. Le test de Tajima (D) [330] utilise la fréquence totale des mutations, tandis que le test de Fu ( $F_S$ ) [329] utilise la distribution des allèles ou des haplotypes (Tableau 5-3). La présence d'une expansion de population a été confirmée en analysant l'HPG A ( $F_S=-15,524$ ,  $p<0,001$ ;  $D=-2,335$ ,  $p<0,001$ ) et l'HPG B ( $F_S=-14,691$ ,  $p<0,001$ ;  $D=-2,3686$   $p<0,001$ ) de manière hautement significative ainsi que l'HPG C ( $F_S=-6,322$ ,  $p<0,001$  ;  $D=-1,9288$ ,  $p<0,05$ ), tandis qu'elle ne l'est pas pour l'HPG E ( $F_S=-1,195$ ;  $D=-0,4102$  ;  $p>0,10$ ). Ainsi ces valeurs confirment les résultats obtenus par les graphiques de « mismatch distributions » (Figure 5-13). Les courbes « mismatch distribution » observées supportent le modèle d'un événement d'expansion de population ayant les HPGs A, B et C.



Tableau 5-3 : Valeurs statistiques des tests de neutralité et du nombre moyen de la différence nucléotidique K pour chaque haplogroupe. (\*\*\*) =  $p < 0,001$  ; (\*\*) =  $0,001 < p < 0,01$  ; (\*) =  $p < 0,05$ . P-value est calculée par le programme de simulation coalescente (1000simulations). †= seuil de signification  $p < 0,02$  ; ‡= seuil de signification  $p < 0,05$ .

	Nombre moyen de la différence nucléotidique (K)	Fu's $F_S$ †	Tajima's $D$ ‡
HPG A	1,109	-15,524***	-2,335***
HPG B	1,293	-14,691***	-2,3686**
HPG C	1,268	-6,322***	-1,9288*
HPG E	1,8	-1,195	-0,4102

#### Répartition géographique des HPGs au sein des différentes régions et troupeau de populations de mouton Awassi du Liban

Les séquences *cyt-b* analysées sont issues de 27 troupeaux de moutons d'Awassi libanais et 2 troupeaux syriens. L'HPG B est le plus fréquent des quatre HPGs retrouvés dans les échantillons analysés (38%), alors que HPG E est le moins faiblement représenté (5%). C'est la première fois qu'une telle étude est réalisée au Liban et les résultats obtenus corrélerent avec les études précédentes ayant identifié l'HPG E au Moyen Orient et en Turquie [7, 320]. Comme précédemment réalisé pour les rétrotypes, la répartition des HPGs par région et par troupeau a été analysée. Les HPGs du Liban sont uniformément repartis (Figure 5-14), sans particularité visible liée à une région ou un troupeau donné. On remarque cependant la particularité d'un animal du troupeau B-5 appartenant au HPG E qui se distingue des autres animaux de cette région par le rétrotype (RE=enJSRV18+enJSRV15+enJSRV22). Les autres séquences HPG E ne présentent pas de particularité au niveau des rétrotypes.

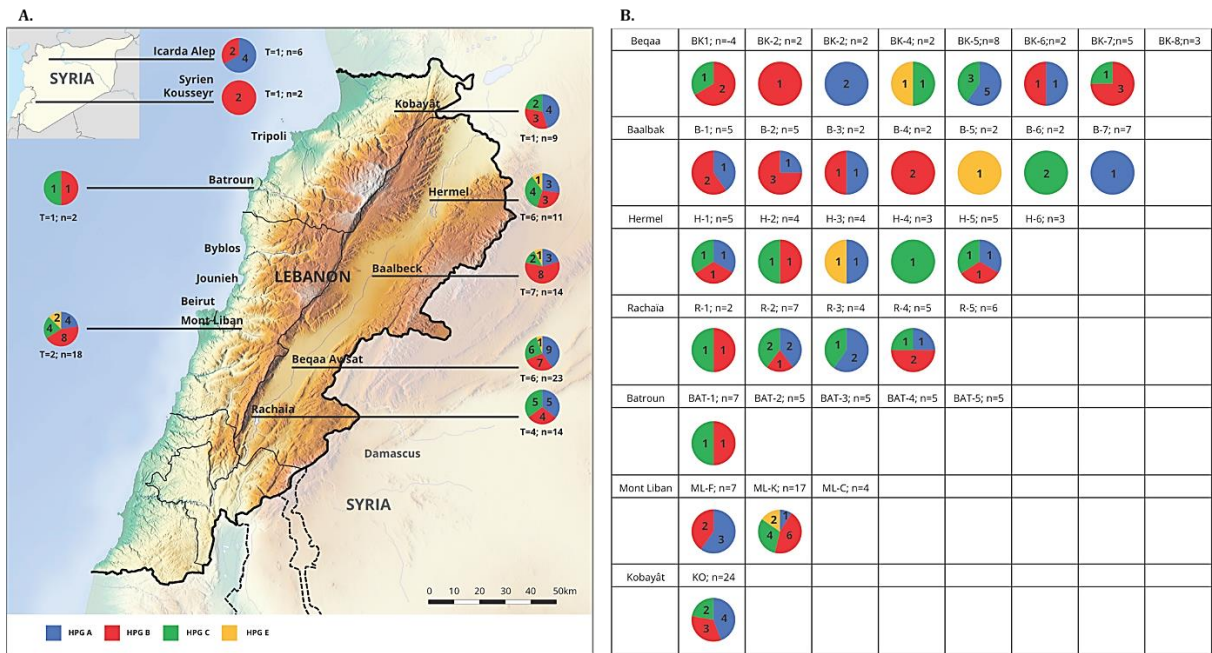


Figure 5-14 : Fréquence et répartition géographique des HPGs des moutons Awassi du Liban. A) en fonction des régions et B. en fonction des troupeaux. Le nombre n dans le tableau représente le nombre d'échantillon prélevé du troupeau alors que le nombre dans les pies représente le nombre d'échantillons dont le Cyt-b a été séquencé.

### 5.2.2. D-Loop

Lors de l'étude de ce marqueur, 8 séquences de la D-Loop ont montrées une hétéroplasmie. Pour confirmer l'absence de *numts* et éliminer toute possibilité de confusion dans la génération des résultats, ces séquences ont été exclues de l'analyse. Comme dans le cas du *Cyt-b*, un arbre Neighbor-Joining basée sur 357 séquences de la D-Loop (508pb) a été généré pour caractériser les HPGs mitochondriaux des échantillons (Figure 5-15). L'identification des HPGs a été réalisée grâce aux séquences de référence extraites de l'étude de Demirci *et al.*, 2013 (240 séquences).

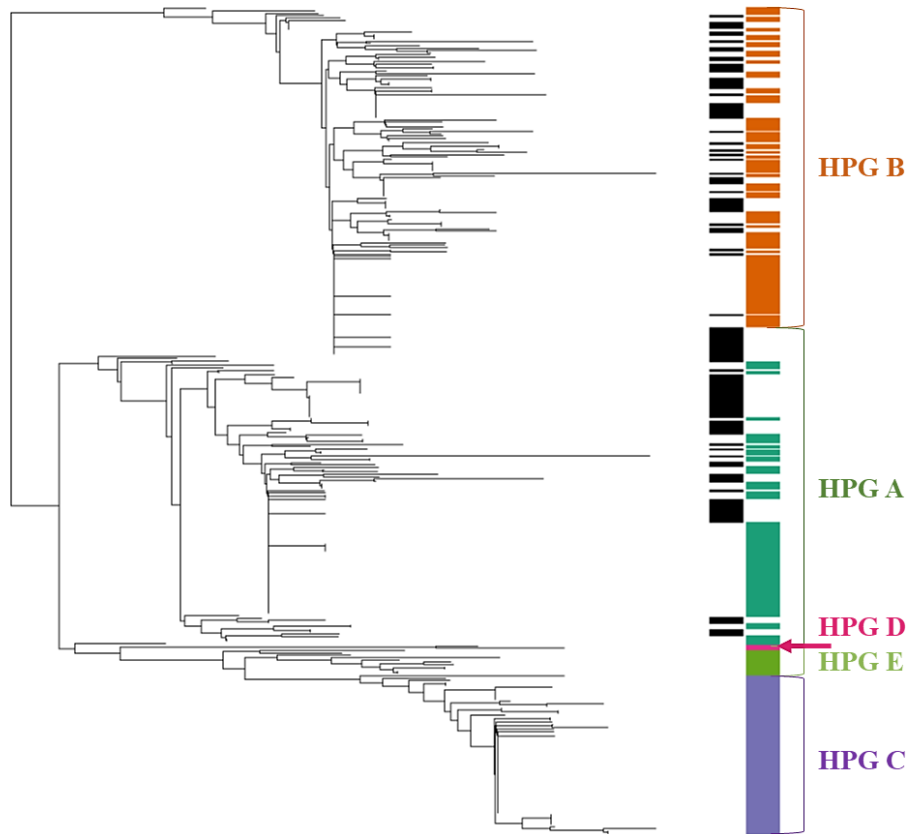


Figure 5-15: Arbre phylogénétique selon NJ des séquences de la D-Loop (n=357) montrant les cinq haplogroupes maternels des moutons. Les traits noirs représentent les échantillons analysés dans notre étude et la colonne colorée représente les séquences utilisées comme référence [320]

Cette analyse distingue seulement deux HPGs dans les populations de mouton de la race Awassi étudiée : HPG A (n=55) et HPG B (n=62).

Le polymorphisme de toutes les séquences a également été calculé (Tableau 5-4). Ces 117 séquences renferment 60 sites polymorphiques dont 34 sont phylogénétiquement informatifs et qui ont abouti à la définition de 66 haplotypes (Annexe 3) avec une diversité haplotypique  $Hd_t = 0,973 \pm 0,0004$  et une diversité nucléotidique  $\pi_t = 15,15 \times 10^{-3}$ . Elles se regroupent donc en deux haplogroupes : HPG A qui présente une diversité nucléotidique plus élevée que HPG B ( $\pi_A = 5,95 \times 10^{-3} > \pi_B = 5,51 \times 10^{-3}$ ) tandis que HPG B présente une diversité haplotypique plus grande ( $Hd_B = 0,957 \pm 0,00022 > Hd_A = 0,933 \pm 0,00038$ ). Ces valeurs diffèrent de celles obtenues pour HPGA et HPGB au cours d'une étude réalisée en Turquie [320] où HPG B est plus polymorphique que HPG A.

Le taux de polymorphisme de HPGs observées par l'étude de la D-Loop ( $S_t=66$  ;  $Hd_t=0,973 \pm 0,0004$  ;  $\pi_t = 15,15 \times 10^{-3}$ ) sont plus élevées que celles du Cyt-b ( $S_t=57$  ;  $Hd_t=0,917 \pm 0,00029$  ;  $\pi_t=5,87 \times 10^{-3}$ ) du fait que la séquence D-Loop renferme deux régions

hypervariables HVR1 et HVR2. Nos résultats corrélient avec ceux obtenus par une étude précédente [324].

Tableau 5-4: Variations phylogénétiques des haplogroupes mitochondriaux (D-Loop) de la race Awassi au Liban

	HPG A	HPG B
Séquences	55	62
Sites Variables (S)	35	39
Diversité Nucléotidique ( $\pi$ )	$5,95 \times 10^{-3}$	$5,51 \times 10^{-3}$
Haplotypes	29	37
Diversité des Haplotypes (Hd)	$0,933 \pm 0,00038$	$0,957 \pm 0,00022$

L'expansion des populations d'HPGA et B a aussi été testée avec les séquences de la D-Loop comme décrit précédemment pour le Cyt-b. Les courbes en cloches des deux haplogroupes montrent une expansion des populations pour les deux HPGs (Figure 5-16).

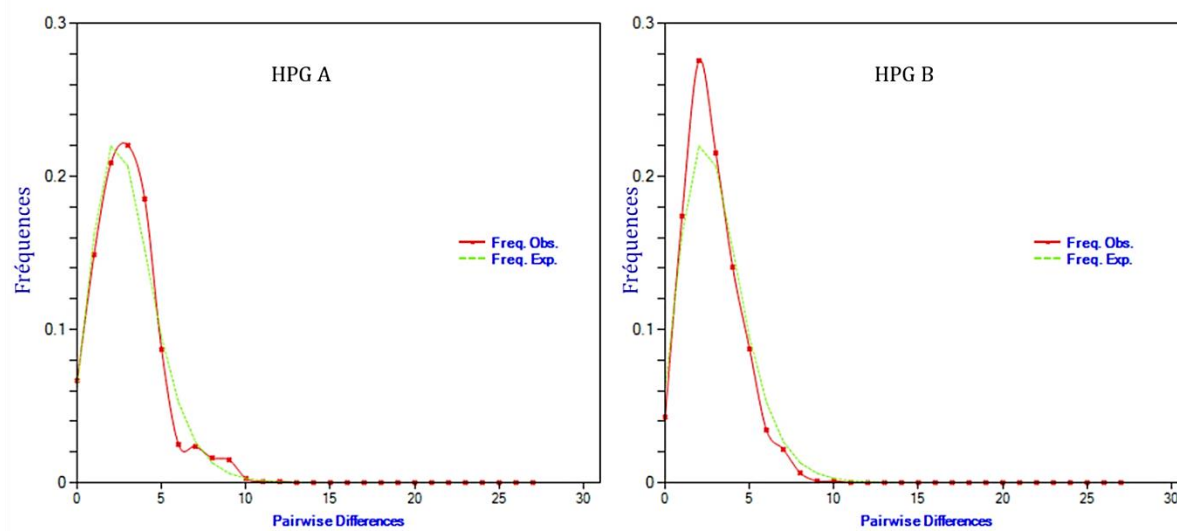


Figure 5-16 : Distribution incompatible des HPGs A et B des séquences D-Loop des populations Awassi du Liban. A) Graphiques montrant la distribution de l'HPG A (A) et de l'HPG B (B).

Cette expansion des populations des deux HPGs est confirmée par les tests statistiques de neutralité Fu,  $F_s$  [329] et de Tajima, D [330] : HPG A ( $F_s = -24,581$  ;  $D = -2,053$  ;  $p < 0,05$ ) et HPG B ( $F_s = -34,274$  ;  $D = -2,25$  ;  $P < 0,01$ ) (Tableau 5-5).

Tableau 5-5 : Valeurs statistiques des tests de neutralité et du nombre moyen de la différence nucléotidique K pour chaque haplogroupe. (\*\*\*) =  $p < 0,001$  ; (\*\*) =  $0,001 < p < 0,01$  ; (\*) =  $p < 0,05$ . P-value est calculée par le programme de simulation coalescente (1000simulations). †= seuil de signification  $p < 0,02$  ; ‡= seuil de signification  $p < 0,05$ .

	Nombre moyen de la différence nucléotidique (K)	Fu's $F_s$ †	Tajima's D
HPG A	3,022	-24,581	-2,053*
HPG B	2,799	-34,274	-2,25**

L'analyse des séquences de la D-Loop n'a pu discriminée que deux HPGs (HPG A et HPG B) qui sont clairement divergents ( $D_{xy}=2,451\%$  ;  $K_{xy}=12,45$ ). Cette valeur de divergence est plus grande que celle obtenue par le *Cyt-b* ( $D_{xy}=0,306\%$  ;  $K_{xy}=3,168$ ). 53% des échantillons se regroupe en HPG B et 47% dans l'HPG A. Aucune différence significative n'est observée au niveau des régions ou au niveau des troupeaux.

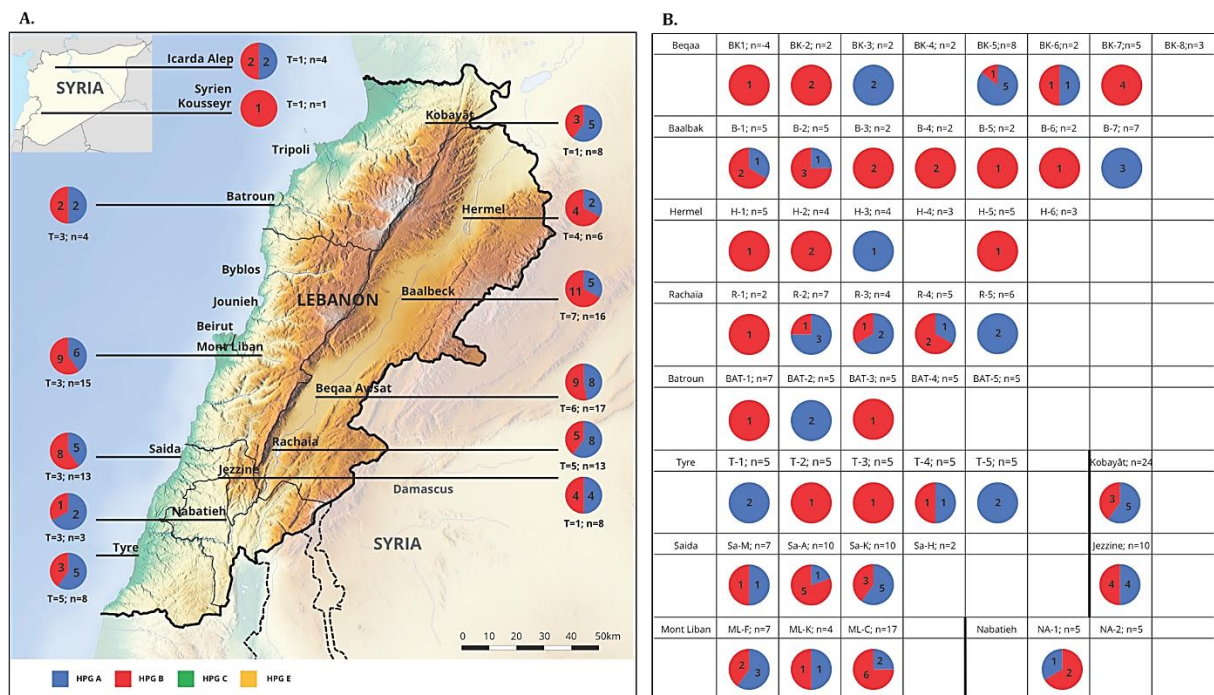


Figure 5-17 : Fréquence et répartition géographique des HPGs des moutons Awassi du Liban. A) en fonction des régions et B. en fonction des troupeaux. Le nombre n dans le tableau représente le nombre d'échantillon prélevé du troupeau alors que le nombre dans les pies représente le nombre d'échantillons dont la D-Loop a été séquencé.

D'après nos résultats, le pouvoir discriminant de la D-Loop ( $D=0,5025$ ) est plus faible que celui du *Cyt-b* ( $D=0,6939$ ). La question est donc posée pour comprendre ce faible pouvoir discriminant des HPGs par la séquence de la D-Loop. En effet, la D-Loop est plus

polymorphique que le *Cyt-b* et de ce fait on s'attendrait à retrouver au moins les mêmes HPGs détectés par le *Cyt-b*. On remarque néanmoins que certains échantillons ont basculé d'un HPG à un autre. Par exemple, quatre séquences *Cyt-b* de l'HPG C sont passées en HPG A par l'étude de la D-Loop, trois séquences d'HPG C en HPG B et même HPG E (un seul individu pour ce dernier). La seule corrélation observée entre le *Cyt-b* et la D-Loop est pour les deux troupeaux syriens, ICARDA avec deux HPGs (A et B) et le troupeau du Kouseyr avec un seul, l'HPG B. Pour confirmer leur appartenance à un HPG donné, on a relié la séquence du *Cyt-B* à celle de la D-Loop pour former ce qu'on appelle le supergène dont la taille est de 1542pb. Meadows et ses collègues ont réalisé une analyse similaire pour tous les échantillons testés et ils ont pu identifier tous les HPGS (A, B, C, D et E) [7]. La construction de l'arbre phylogénétique selon le modèle de Neibghor-Joining a séparé nos 71 « super-séquences » en cinq haplogroupes (Figure 5-18).

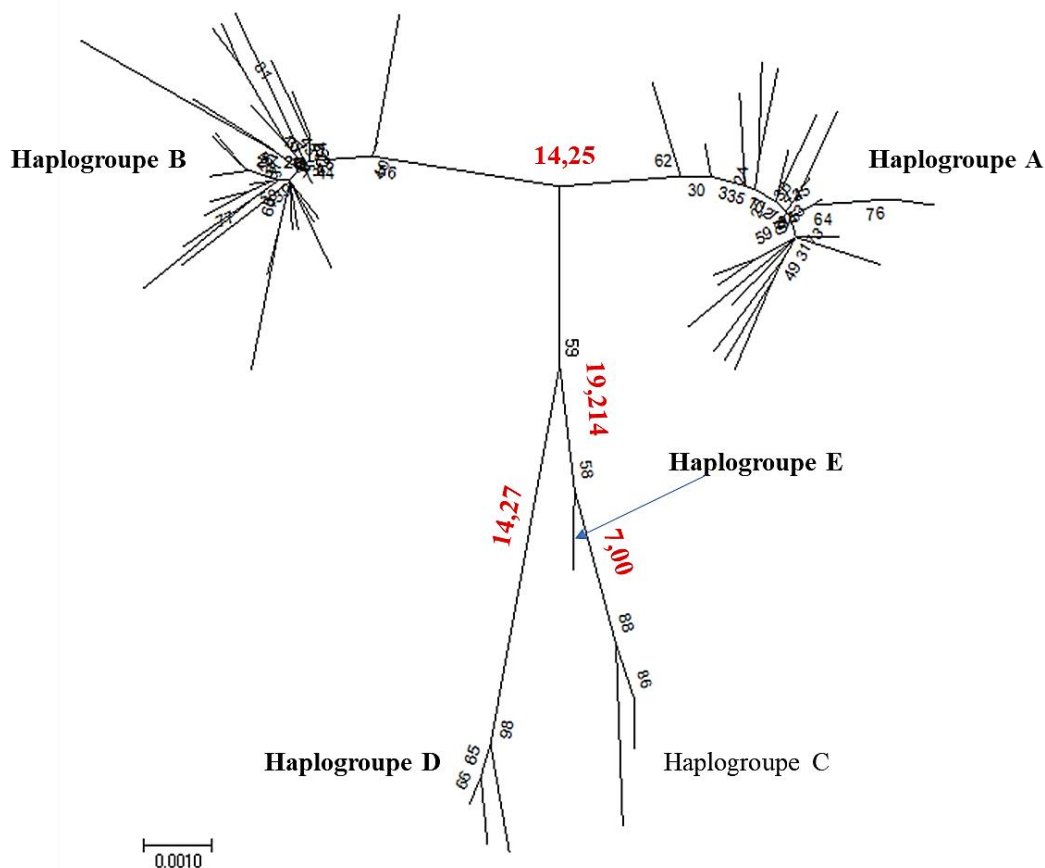


Figure 5-18 : Cinq lignées maternelles illustrées avec l'arbre Neighbor-Joining généré par le test de phylogénie Bootstrap avec 1000 réplifications. Cet arbre montre les 71 séquences *Cyt-b*-D-Loop (1540pb) de la race Awassi au Liban. La divergence entre les haplogroupes (Kxy) est notée sur les branches de l'arbre ainsi que les valeurs des nœuds de support (>50%).

Cette analyse groupée, contrairement aux précédentes (Cytb et D-Loop), permet d'identifier un cinquième HPG (HPG D) (Figure 5-18). En effet, la distance génétique qui sépare HPG D de HPG A ( $K_{xy}=14,27$ ) est égale à celle qui sépare HPG A de HPG B ( $K_{xy}=14,25$ ). De plus, cette distance est plus grande entre HPG D/HPG C ( $K_{xy}=13,82$ ) qu'entre HPG C/E ( $K_{xy}=7$ ) (Tableau 5-6). Enfin, les valeurs des nœuds (bootstrap value) (Figure 5-18B) sont supérieures à 50% et statistiquement significatives pour définir un clade, donc un HPG distinct semblable à l'HPG D retrouvé dans l'étude précédente de Meadows 2007. Ces résultats confirment donc que quatre échantillons appartiennent à un HPG (i.e. HPG D) différent des quatre autres précédemment identifiés par l'analyse du Cyt-b (A, B, C et E).

Tableau 5-6 : Les nombres moyens de différence nucléotidique ( $K_{xy}$ ) et de substitutions nucléotidiques par site ( $D_{xy}$ ) entre les populations des différents haplogroupes des moutons Awassi libanais avec la séquence « super gène » : Cytb-DLoop.

	<b>HPG A</b>		<b>HPG B</b>		<b>HPG C</b>		<b>HPG D</b>	
	$D_{xy}$ %	$K_{xy}$	$D_{xy}$ %	$K_{xy}$	$D_{xy}$ %	$K_{xy}$	$D_{xy}$ %	$K_{xy}$
<b>HPG B</b>	0,924	14,25						
<b>HPG C</b>	1,496	23,07	1,038	16,02				
<b>HPG D</b>	0,925	14,27	1,62	24,99	0,903	13,92		
<b>HPG E</b>	1,246	19,214	1,039	11,629	0,454	7	1,184	18,25

Les deux marqueurs mitochondriaux utilisés pour cette étude semblent donc finalement apporter des informations plus pertinentes lorsqu'on effectue une analyse de type « supergène ». La race Awassi renferme donc, en des proportions différentes, les cinq HPGs maternels identifiées à ce jour chez les ovins. La diversité génétique de ces HPGs est présentée dans le Tableau 5-7. Ce supergène présente une diversité nucléotidique plus grande que celle du *Cyt-b* mais plus faible que celle de la D-Loop.

Tableau 5-7 : les variations phylogéniques des séquences supergènes.

	HPG A	HPG B	HPG C	HPG D	HPG E
Séquences	28	35	3	4	1
Sites Variables (S)	41	43	6	5	
Diversité Nucléotidique ( $\pi$ )	$3,43 \times 10^{-3}$	$2,45 \times 10^{-3}$	$2,59 \times 10^{-3}$	$1,62 \times 10^{-3}$	
Haplotypes	24	28	3	3	1
Diversité Haplotypique (Hd)	$0,987 \pm 0,0002$	$0,98 \pm 0,00017$	$1,00 \pm 0,07407$	$0,833 \pm 0,05$	



## Chapitre 6 : *Discussion*

---

Dans le but de comprendre l'histoire évolutive des moutons, de nombreux marqueurs génétiques ont été utilisés au cours de ces deux dernières décennies : L'ADNmt [7, 8, 17, 173, 216, 324], le chromosome Y [325, 336]; les microsatellites [337, 338], les polymorphismes nucléotidiques (single-nucleotide polymorphism, SNPs) [67] et les rétrovirus endogènes insérés dans l'ADN génomique des moutons [172]. L'unique mouton indigène du Liban est celui de la race Awassi. Plusieurs études ont été réalisées sur cette race dans le but d'améliorer son rendement en lait et sa reproduction mais aucune ne porte sur sa caractérisation génétique et son évolution. Dans ce projet de thèse, nous avons utilisé deux approches génétiques pour caractériser les populations de mouton de cette race au Liban. La première est basée sur les rétrovirus endogène (enJSRV/Rétrotypage) et la deuxième sur les séquences de l'ADNmt : *Cyt-b* et D-Loop.

### **6.1. Les rétrovirus endogènes enJSRVs sont-ils des bons marqueurs génétiques pour l'analyse de populations d'une même race et d'une région géographique donnée ?**

Cinq enJSRVs insérés de manière polymorphique dans le génome des moutons domestiques (*Ovis Aries*) (enJSRV-18, enJSRV-7, enJSRV-15, enJSRV-16, et enJSRV-22) ont été utilisés dans cette étude comme des marqueurs génétiques et ainsi définir, pour chaque animal/troupeau/population de mouton Awassi du Liban, des « Rétrotypes » comme précédemment décrit par Chessa *et al.*, 2009 (à l'exception d'enJSRV-8 qui n'a pas été testé étant restreint à l'Europe du Nord). Néanmoins, dans l'étude de Chessa, l'analyse portait sur des populations de moutons sauvages et domestiques, de races différentes et de localisation géographique très diverse (Asie, Afrique et de toute l'Europe). Au début de cette étude, il n'était pas du tout certain que ces enJSRVs soient suffisamment discriminants et puissent apporter des informations pertinentes sur l'analyse de populations de mouton d'une seule race [339] et d'une région donnée (le Liban). Afin d'exploiter ce marqueur génétique (enJSRVs) au maximum de ses possibilités nous avons défini des rétrotypes supplémentaires (RA-RH) par rapport à l'étude de Chessa qui n'a combiné que quatre enJSRVs pour établir les rétrotypes R0 à R14 : enJSRV-7 + enJSRV-18 + enJSRV-22 + enJSRV-8. L'enJSRV-18 est retrouvé dans tous nos échantillons testés (n=277) corrélant ainsi avec les résultats de Chessa *et al.*, 2009. La présence d'enJSRV-18 (R2) est caractéristique des moutons dits modernes issus de la deuxième vague de migration.

De plus, l'Awassi renferme l'enJSRV-7 qui est plus fréquemment retrouvé dans les populations de mouton situées au pourtour du bassin méditerranéen. Le rétrotype R4 dit méditerranéen, traduisant la double présence d'enJSRV-18 et enJSRV-7, est bien présent dans le génome des Awassi au Liban dont la côte Ouest méditerranéenne s'étend sur 210km avec quinze ports phéniciens. En analysant, ce rétrotype, nous avons remarqué qu'il était significativement plus présent dans les populations Awassi de cette côte (dont Batroun et Tyr qui sont deux régions connues pour leur port phénicien) par rapport à celles situées à l'Est du Liban (Hermel, Baalbak et la Beqaa). Notons qu'administrativement, les trois régions, Hermel, Baalbak et la Beqaa, sont distinctes mais sur le terrain elles constituent une seule zone, c'est une plaine où les gens sont reliés par des liens de parenté, culturels et de ce fait les échanges commerciaux au sein de cette région se font naturellement. Globalement, ce résultat pour le R4 est très intéressant car l'Ouest et l'Est de Liban sont séparées par une haute chaîne de montagne suggérant que, suite aux échanges passés *via* la Mer méditerranéenne de populations de moutons (et dont R4 est une trace génétique), la topologie du Liban a restreint la mobilité animale et les échanges commerciaux Ouest-Est (observé aujourd'hui par une distribution différentielle du rétrotype R4). Il serait d'ailleurs intéressant de voir si dans d'autres pays du bassin méditerranéen ayant des topologies particulières, nous pouvons également observer de telles différences de distribution au niveau de R4. Le troupeau Libanais présentant le plus de diversité en termes de rétrotype (au moins cinq différents) est le troupeau de kobayât. Il est intéressant de noter qu'il s'agit d'un troupeau commercial et donc soumis à de nombreux échanges pouvant expliquer cette diversité rétrotypique. Enfin, nous avons observé au cours de notre analyse que certains troupeaux : H1 (R7), H4 (R6), B-5 (RE) et BK-8 (RF) sont différents « rétrotypiquement » des autres troupeaux de leurs régions (Hermel, Baalbak et Beqaa). Néanmoins, les informations obtenues sur ces troupeaux ne permettent pas à l'heure actuelle d'expliquer de manière simple l'ensemble des différences observées. Il pourrait s'agir de la provenance des animaux utilisés pour établir ce troupeau qui peut être différente des autres ou bien encore, étant ouvert, de l'intégration d'animaux différents par rapport aux autres troupeaux. La seule information intéressante concerne le troupeau BK-8 (RF=enJSRV18+enJSRV-16+enJSRV-7) qui se distingue des autres troupeaux de la région par la présence d'enJSRV-7 et qui contrairement aux autres est fermé ce qui peut expliquer qu'enJSRV-7 ne soit pas partagé par les autres populations de Beqaa. Le troupeau B-5 et en particulier un des deux animaux analysés est très intéressant car il se distingue par son rétrotype (RE) mais aussi par son HPG (HPG E) indiquant très probablement une provenance

extrarégionale de cet animal. Il sera important d'étudier un nombre plus important d'animaux de cette région Baalbak pour voir si RE et HPG E sont retrouvés spécifiquement dans ce troupeau et/ou dans d'autres animaux de B-5 ou s'il s'agit vraiment d'un animal isolé. Les quatre autres échantillons HPG E ne se distinguent pas par leurs rétrotypes : Trois d'entre eux ont un rétrotype R2 et un seul le rétrotype RB.

Globalement, enJSRVs sont de bons marqueurs génétiques qui furent utilisés pour étudier l'histoire de la domestication du mouton *via* une analyse à grande échelle de nombreuses et diverses populations. Au niveau d'une seule race et d'une région géographique donnée, comme dans notre étude, ils sont bien entendu un peu moins informatifs mais fournissent pourtant des informations intéressantes liées à certaines particularités du pays concerné (i.e., topologie/R4) en plus d'établir une caractérisation « Rétrotypique » assez exhaustive des populations de mouton Awassi présentes dans les différentes régions du Liban.

## 6.2. L'invité surprise, le solo-LTR !

L'analyse du provirus enJSRV-6 a montré la présence de solo-LTR pour cet enJSRV dans 54,5% des échantillons testés (soit sous forme homozygote ou hétérozygote avec le provirus complet). Cette forte fréquence dans les populations testées suggère une apparition relativement ancienne de ce solo-LTR qui a pu être transmis de génération en génération au Liban. De manière générale, les anciens ERV existent sous forme de solo-LTR en raison de la recombinaison entre les LTR de 5' et 3' du provirus au cours de l'évolution avec leur hôte [340]. Notons aussi dans notre étude, l'absence de détection des sites « vides » c'est-à-dire sans insertion d'enJSRV-6 (locus vide= 302 pb) ce qui corrèle avec les études menées par l'équipe du Prof. Palmarini indiquant qu'enJSRV-6 s'est inséré dans le génome avant la spéciation des genres *Ovis* et *Capra* et est donc présent dans tous les génomes d'*Ovis aries*.

Les précédentes études publiées n'ont pas observé de Solo-LTR pour enJSRVs et enJSRV-6 en particulier [68, 172]. Cependant, un manuscrit de thèse en ligne [341], données non publiées, rapporte l'absence d'enJSRV-6 dans le génome de certains moutons locaux de Turquie appartenant aux races Karayaka et Awassi (nommée aussi İvesi en Turquie); ce qui semble à première vue en contradiction avec les études menées par l'équipe du Prof. Palmarini. AYANOĞLU et collègues ont séquencé ces locus vides et les séquences obtenues s'alignent parfaitement avec les extrémités 5' et 3' de la séquence d'enJSRV-6 déposée dans GenBank. Les auteurs ont interprété ce résultat comme correspondant aux

séquences génomiques présentent de part et d'autre d'enJSRV-6. Cependant, les séquences génomiques encadrant le provirus n'ont pas été publiées et il s'agit ici que de la séquence provirale. Ainsi, en y regardant de plus près, on s'aperçoit qu'il s'agit en fait des deux LTR 5' et 3' d'enJSRV-6 indiquant la présence d'un solo-LTR dans le génome de ces moutons de Turquie. Le solo-LTR d'enJSRV-6 ne serait donc spécifique des moutons de la race Awassi du Liban, même s'il n'a pas été retrouvé dans les moutons Awassi d'Israël analysés par Chessa *et al.*, (n=17), mais il serait bien présent dans les Awassi de Turquie et *a minima* dans d'autres races de mouton de cette région (Karayaka). Il sera donc intéressant de rechercher la présence ou l'absence de ce solo-LTR d'enJSRV-6 à travers une étude à plus grande échelle ; celui-ci pouvant être très informatif pour établir des liens génétiques et ainsi comprendre l'origine et l'histoire des différentes races locales de moutons du Moyen-Orient et peut-être d'ailleurs. Une telle étude pourrait plus particulièrement nous permettre d'émettre des hypothèses quant à la forte présence de Solo-LTR dans les populations Awassi du Sud-Ouest du Liban. Ce solo-LTR pourrait donc être initialement apparu dans cette région, étant très probablement plus récent que les différents enJSRVs analysés, il ne s'est pas révélé être aussi informatif que le rétrotypage.

### 6.3. Des enJSRVs oui mais où est le JSRV ?

En parallèle de ce travail de thèse, une tentative d'étude épidémiologique du rétrovirus exogène JSRV a été réalisée en amplifiant le gène *env* par hemi-nested PCR à partir d'ADN extraits de sang de moutons Awassi des différentes régions du Liban [342]. Nos résultats ont suggéré une prévalence de 5,5% (15 échantillons sur 273) avec seulement trois animaux présentant des signes cliniques. Ces animaux positifs sont surtout localisés dans la région de Hermel (10 échantillons positifs dont 2 provenant de chèvres) où les services vétérinaires sont rares. Néanmoins, ces données ne sont pas présentées dans la partie « résultats » de fait de l'impossibilité de reproduire ces résultats par PCR et de confirmer par séquençage ; ceci étant probablement dû au fait que la première série d'expériences fut réalisée à partir de sang frais et non congelés, l'amplification par PCR du génome de JSRV à partir du sang étant extrêmement faible et difficile à réaliser.

Sachant que les enJSRVs *via* leur enveloppe peuvent bloquer l'entrée virale de JSRV, il pourrait être intéressant, ayant caractérisé génétiquement les enJSRVs des génomes d'Awassi, de corrélérer ces données avec la prévalence de JSRV dans les différentes régions

du Liban ; les troupeaux de mouton d'Hermel étant par exemple un peu moins divers rétrotypiquement (qualitativement et quantitativement) que ceux d'autres régions. En parallèle, le séquençage du gène Gag d'enJS56A1 et enJSRV-20 (pour détecter le W21 protecteur) ainsi que la détermination du nombre de copie d'enJS56A1-like dans ces génomes pourraient aussi apporter des informations complémentaires quant à la susceptibilité ou la résistance à JSRV de certaines populations d'Awassi du Liban.

#### **6.4. L'ADN mitochondrial et son pouvoir discriminant**

Pour mieux caractériser et essayer de comprendre l'histoire de la race Awassi du Liban, la diversité génétique de ces moutons a été étudiée grâce à l'ADN mitochondrial. Les deux marqueurs mitochondriaux utilisés, le cytochrome b et la D-Loop, semblent avoir un pouvoir discriminant différent au sein de notre population à caractère unique : une seule race dans une même zone géographique. En effet, la séquence complète du *Cyt-b* (1034pb) a distingué nos échantillons (n=99) en quatre HPGs : A (32%), B (38%), C (24%) et E (5%) alors que la séquence de D-Loop (534pb) a pu n'en identifier que deux dans nos échantillons (n=117) : HPG A (47%) et HPG B (53%). Afin d'être plus discriminant, j'ai eu recours à une analyse du « supergène » (Cyt-b + D-Loop) (Multiple sequence alignment). Cette analyse nous a donnés des résultats pour 71 échantillons (i.e. ceux pour lesquels nous avons les deux types de séquences). L'arbre phylogénétique de ce supergène a pu discriminer les cinq HPGs maternels connus à ce jour : HPG A (39,44%), HPG B (49,3%), HPG C (4,23%), HPG D (5,63%) et HPG E (1,41%). Une étude précédente a aussi utilisé l'alignement des séquences multiples pour détecter l'HPG D [7] alors que d'autres équipes ont retrouvé cette HPG en analysant seulement la D-Loop [16, 320, 321]. Jusqu'à présent, l'HPG D n'a jamais été détecté par l'analyse du *Cyt-b* seul.

La répartition des HPGs n'est pas aussi géographiquement représentative que les précédentes analyses (Cyt-b et D-Loop séparément) en l'absence d'échantillon du Sud du Liban. Nous projetons donc d'analyser plus d'échantillons pour couvrir toutes les régions du Liban et ce de façon à avoir aux moins 10 échantillons par région. Cela nous permettra de renforcer nos conclusions et peut-être d'obtenir plus d'informations sur les HPGs de ces populations mais aussi de pouvoir mieux corrélérer ces données avec le rétrotypage. Les analyses de diversité génétique de ces trois séquences mitochondriales (Tableau 6-1) n'a pas fourni des informations claires pour définir des critères de discrimination des HPGs. Néanmoins, le nombre de séquences (n) n'est pas le même pour tous les marqueurs ce qui

peut influencer les valeurs des paramètres étudiés. Ces analyses seront donc refaites lorsque nous aurons plus de données et un nombre de séquences identiques. Cependant, il apparaît que d'après nos analyses, le super gène soit le plus discriminant bien qu'étant basé sur un plus faible nombre de séquence. Théoriquement, l'analyse des HPGs est basée principalement sur une quantité inférée du taux de mutation et des polymorphismes héréditaires. Ainsi, l'analyse des supergènes, qui combine l'information des deux gènes, a permis de résoudre la discordance des résultats obtenus à partir du Cyt-b et de la D-Loop et de créer des HPGs consensus.

Tableau 6-1 : Diversités génétiques observées dans les trois types de séquences mitochondriales analysées.

	Taille pb	Séquences (n)	Sites Variables (S)	Haplotypes	Diversité Haplotypique (Hd)	Diversité Nucléotidique ( $\pi$ )
<b>Cyt-b</b>	1034	99	57	45	0,917±0,00029	5,87x10 <sup>-3</sup>
<b>D-Loop</b>	508	117	60	66	0,973±0,0004	15,15x10 <sup>-3</sup>
<b>Cyt-b &amp; D-Loop</b>	1542	71	90	59	0,993±0,00002	7,35x10 <sup>-3</sup>

La race Awassi et la race Akkaraman en Turquie [320] sont les deux seules races domestiques à renfermer les cinq HPGs maternels. L'HPG D et E sont les plus rares et retrouvés uniquement en Asie. L'HPG D est, quant à lui, limité aux moutons à queue grasse tel que l'Awassi, la race Karachai au Nord du Caucase [16], la race Morkaraman en Turquie [7], la race Akkaraman en Turquie [320], la race Linzhou au Tibet [321] et, en accord avec ces études, HPG D est aussi retrouvé dans les populations d'Awassi du Liban. Nous avons aussi pu prouver que les HPG A, B et C ont subi une expansion de population (Figure 5-13 et Tableau 5-5) alors que le faible nombre des deux autres HPGs (D et E) a rendu cette analyse difficile.

## 6.5. Le Liban est-il un centre de domestication du mouton ?

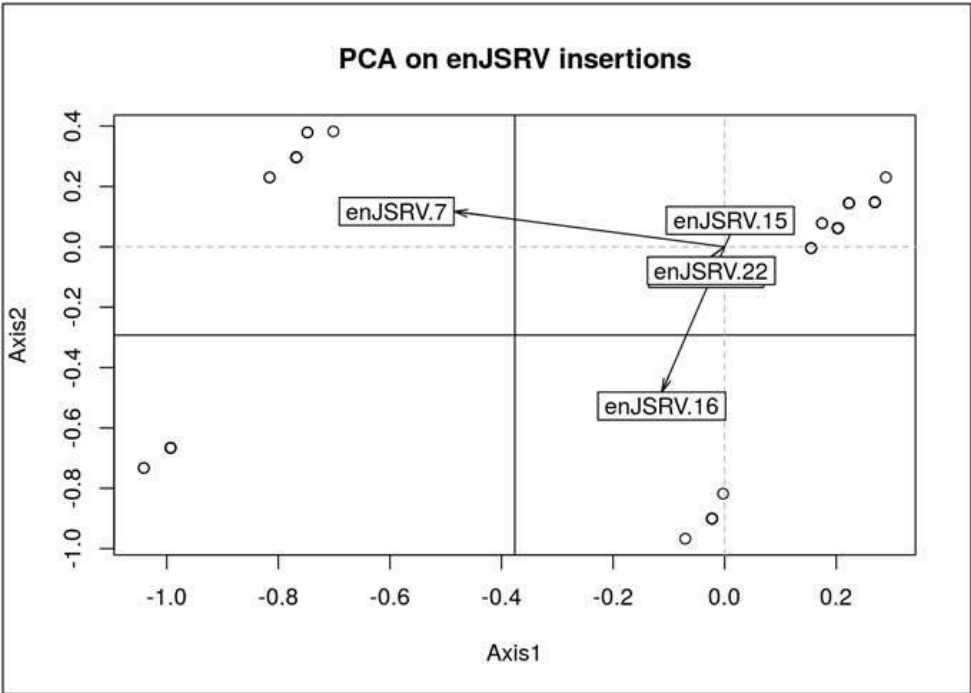
Le centre de domestication des moutons a été situé au Levant (pays bordant la côte orientale de la Mer Méditerranée : en premier lieu le Liban et la Syrie) lors de la révolution agricole néolithique suggérant que le Liban puisse être un centre de domestication des moutons. Globalement, les troupeaux de moutons Awassi du Liban sont très divers que ce soit d'un point rétrotype (par la présence des cinq enJSRVs polymorphiques) et de l'ADN mitochondrial (par la détection des cinq lignées maternelles universellement reconnues).

Nos résultats sont similaires aux études effectuées en Turquie avec l'ADNmt [7, 320, 324] mais divergent de celle réalisée avec les enJSRVs sur des moutons de races natives de Turquie [341] et dans lesquels les enJSRV-15 et enJSRV-16 n'ont pas été détectés. La recherche dans les populations Awassi du Liban du dernier enJSRV polymorphique qui n'a pas été analysé dans cette étude (enJSRV-8) et qui est a priori restreint à l'Europe du Nord permettra peut-être de renforcer ce point. Les populations de mouton du centre de domestication sont caractérisées par une diversité génétique élevée, cette dernière diminue lorsqu'on s'en éloigne (ex : présence uniquement d'HPG A et d'HPG B en Europe); les populations de la race Awassi du Liban respectent donc ce critère de forte diversité pour les deux marqueurs génétiques analysés dans cette étude.

*THE END...*



ANNEXE 1 : Analyse en composantes principales sur les insertions des enJSRVs



L'analyse ACP a regroupé les échantillons selon enJSRV-7 et enJSRV-16

## ANNEXE 2 : Alignement des séquences des haplotypes obtenus à partir du *Cyt-b*

	111
	112222223333344555555566666666777777888888999999000
	34918023347801477573445666713666789911225933458011467122
	896656724791870525544075024383379540247490136512217882906
Consensus	TTGTCTCTCGTGTTCGGTTAACCGCTGTCGTCCATACTCGTCCATTACGTGGACCTCA
H_1	.....CT.....
H_2	.....T.....
H_3	.....C.....
H_4	.....T.....
H_5	A.....
H_6	A.....
H_7	A.....T.....
H_8	.....C.....
H_9	.....G.....
H_10	A.....
H_11	.....T.....
H_12	.....C.....
H_13	T.....C.....
H_14	.....A.....
H_15	.....C.....G.....A.....C.....
H_16	.....C.....G.....
H_17	.....C.....G.....C.....
H_18	.....C.....G.....G.....A.....G.....
H_19	.....C.....G.....G.....C.....
H_20	.....C.....G.....T.....
H_21	A.....C.....G.....
H_22	.....C.....G.....C.....
H_23	.....C.....G.....A.....
H_24	.....C.....G.....A.....
H_25	.....C.....G.....T.....
H_26	.....C.....C.....A.....G.....
H_27	.....C.....G.....C.....
H_28	.....C.....G.....T.....
H_29	.....C.....A.....G.....C.....G.....
H_30	.....C.....
H_31	.....C.....T.....CCG.....C.....T.....G.....A.....T.....C.....
H_32	.....C.....T.....CCG.....C.....T.....G.....A.....T.....
H_33	.....C.....T.....CCG.....C.....T.....G.....C.....A.....T.....
H_34	.....C.....CCG.....C.....T.....G.....AA.....T.....
H_35	C.....CT.....T.....CCG.....C.....T.....G.....A.....T.....
H_36	.....C.....T.....CCG.....C.....T.....C.....G.....A.....A.....T.....
H_37	.....C.....CCG.....T.....C.....T.....G.....A.....T.....
H_38	.....C.....CCG.....C.....T.....G.....A.....T.....
H_39	.....C.....T.....CCG.....C.....T.....G.....G.....A.....T.....
H_40	.....C.....T.....CCG.....GC.....T.....G.....A.....T.....
H_41	.....C.....C.....G.....C.....G.....A.....
H_42	.....C.....C.....G.....C.....A.....G.....A.....
H_43	.....C.....AC.....G.....C.....G.....A.....
H_44	C.....C.....A.....C.....G.....C.....A.....G.....A.....
H_45	.....

### ANNEXE 3 : Aligement des séquences des haplotypes obtenus à partir de la D-Loop

	111111111122222233333344444444444444445555
	12222233445566899022344459900114634589900023445588990000
	345712345709571264569678102543707052131492736767190926583567
Consensus	TTATGCATTTTGGTTTTCCAACATACTCTATCTCATTTCGTTGAGGCTCTATTTATCCTC
H_1	.....G.C.A.....TG.....CT.T.A.....C.....
H_2	.....G.C.A.C..TG.....T.T.....CT.....
H_3	A.....G.C.A.C..TG.....T.T.....C.....
H_4	.....G.C.A.C..TG.....T.T.....C.A..C.....
H_5	.....G.C.A.C..TG.....TCT..T.....C.....
H_6	.....G.A.C.TG.....T.T.....TC.....
H_7	.....G.C.A.C..TG..C.....T.T.....C.....
H_8	.....G.A.C.TG.....T.T..T.....C.....
H_9	.....G.C.A.C..TG.....T.T.....A.C.C.....
H_10	.....G.C.A.C..G..C.....T.T.....C.....
H_11	.....G.A.C.TG..C.....T.T.....C.....
H_12	.....G.C.A.C..TG.T.....G.T.T.....C.C.....
H_13	.....G.C.A.....TG.T.....T.T..A.....C.....
H_14	.....G.C.A.....G.....T.T.....C.....
H_15	.....G.C.....G.....T.T.....A.C.....
H_16	.....G.A.C.TG.....T.T..A.....C.....
H_17	.....G.A.C.TG..C.....T.T.....C.....
H_18	.....G.CC.A.C..TG.....T.T.....C.....
H_19	.....G.C.A.C..TG.....T.T..A.....C.....
H_20	.....G.C.A.C..TG.....T.T..A.....C..G.....
H_21	.....A.G.C.A.C..TG.....T.T.....T.....C.....
H_22	.....G.C.A.C.C.TG.....T.T.....C.....
H_23	.....G.C.A.C..TG..C.....T.T.....C.....C.....
H_24	.....G.C.A.C..TGG.....G.T.T.....A.C.....GG.G..A.....
H_25	.....G.C.A.C..TG..C.....T.T.....C.....C.....
H_26	.....G.C.A.C..TG.....T.T.....C.....G.....
H_27	.....G.C.A.C..TG.....T.T.....A.A.C.....
H_28	.....G.C.A.CC.TG.....T.T.....CG.GA..C.....
H_29	.....C.A.....C.....G.....
H_30	.....T.....
H_31	.....T.....G.....G.....
H_32	.....T.....G.....
H_33	.....T.....A..G.....
H_34	.....G.....
H_35	.....A.....T.....G.....
H_36	.....T..C.....G.....
H_37	.....T.....G.....
H_38	.....T.C.....C.....G.....
H_39	.....T.....G.....
H_40	.....G.....T.....G.....
H_41	.....C.....T.....G.....
H_42	.....T.....T.....G.....
H_43	.....C.T.....G.....
H_44	.....T.....G.....GC.....G.....
H_45	.....T.....GC.....
H_46	.....T.....A.....G.....
H_47	.....T..T.....A..A..G.....
H_48	.....T.....T.....A..G.....
H_49	.....T.....T.....C..G.....
H_50	.....C.....C.....C..G.....
H_51	.....C.....T.....
H_52	.....C.....T.....G..T.....
H_53	.....T.....C..G.....C.....
H_54	.....T.....T.....G.....
H_55	.....TC.....C..G.....
H_56	.....T.....A..G..C.AA.....
H_57	.....T.....G..T.....T.....G.....C.....
H_58	.....T.C.....G.....G.....
H_59	.....T.....G.....C.....
H_60	.....T.C.....C.....C.....
H_61	.....C.....T.....
H_62	.....T..C.T.G.....A.....C.....
H_63	.....T.....G.....A.....
H_64	.....AT.....T.....G.....C.....
H_65	.....T.....T..G.....G.....
H_66	.....G.C.A.C..TG.....T.T.....C.....

## RÉFÉRENCES BIBLIOGRAPHIQUES

1. Zeder, M.A., *Domestication and early agriculture in the Mediterranean Basin: Origins, diffusion, and impact*. Proc Natl Acad Sci U S A, 2008. **105**(33): p. 11597-604.
2. Zeder, M.A., et al., *Documenting domestication: the intersection of genetics and archaeology*. Trends Genet, 2006. **22**(3): p. 139-55.
3. MA, Z., *The domestication of animals*. J. Anthropol. Res, 2012. **68**: p. 161–190.
4. Sue, C., C. James, and ShennanStephen, *The evolution of Neolithic farming from SW Asian origins to NW European limits*. European Journal of Archaeology 2005. **8**(2): p. 137-156.
5. Epstein, H., *The Awassi sheep with special reference to the improved dairy type*. FAO ANIMAL PRODUCTION AND HEALTH PAPER, 1985. **57**.
6. Ryder, M.L., *Sheep and man*. 1983, London-UK: Gerald Duckworth & Co. Ltd. 846.
7. Meadows, J.R., et al., *Five ovine mitochondrial lineages identified from sheep breeds of the near East*. Genetics, 2007. **175**(3): p. 1371-9.
8. Pedrosa, S., et al., *Evidence of three maternal lineages in Near Eastern sheep supporting multiple domestication events*. Proc Biol Sci, 2005. **272**(1577): p. 2211-7.
9. SHERRATT, A., *Plough and pastoralism: aspects of the secondary products revolution*, in *Pattern of the past: studies in honour of David Clarke*. 1981, Cambridge University Press.: Cambridge. p. 261–305.
10. Helmer D, G.L.a.V.E., *The development of the exploitation of products from Capra and Ovis (meat, milk and fleece) from the PPNB to the Early Bronze in the northern Near East (8700 to 2000 BC cal.)*. Anthropozoologica, 2007. **42**: p. 41–69.
11. Hiendleder, S., et al., *Molecular analysis of wild and domestic sheep questions current nomenclature and provides evidence for domestication from two different subspecies*. Proc Biol Sci, 2002. **269**(1494): p. 893-904.
12. Bruford, M. and S. Townsend, *Case studies in the genetics of animal domestication: sheep*, in *Documenting Domestication: New Genetic and Archaeological Paradigms*, D.-W.D. Zeder M, Bradley D, Smith BD, Editor. 2006, California University Press: Berkeley, CA, USA.
13. Rezaei, H.R., et al., *Evolution and taxonomy of the wild species of the genus Ovis (Mammalia, Artiodactyla, Bovidae)*. Mol Phylogenet Evol, 2010. **54**(2): p. 315-26.
14. Clutton-Brock, J., *A natural history of domesticated mammals*. 1999, Cambridge: Cambridge University Press.
15. Dobney, K., Larson, G., *Genetics and animal domestication: new windows on an elusive process*. Journal of Zoology, 2006. **269**(2): p. 261–271.
16. Tapio, M., et al., *Sheep mitochondrial DNA variation in European, Caucasian, and Central Asian areas*. Mol Biol Evol, 2006. **23**(9): p. 1776-83.
17. Pereira, F., et al., *Genetic signatures of a Mediterranean influence in Iberian Peninsula sheep husbandry*. Mol Biol Evol, 2006. **23**(7): p. 1420-6.
18. Lv, F.-H.P., Wei-Feng; Yang, Ji; Zhao, Yong-Xin; Li, Wen-Rong; Liu, Ming-Jun; ..... Kantanen, Juha; Han, Jian-Lin; Li, Meng-Hua, *Mitogenomic Meta-Analysis Identifies Two Phases of Migration in the History of Eastern Eurasian Sheep*. Mol. Biol. Evol., 2015. **32**(10): p. 2515–2533.
19. Davis, J.M. and J.V. Beckett, *Animal husbandry and agricultural improvement: the archaeological evidence from animal bones and teeth*. Rural Hist, 1999. **10**(1): p. 1-17.

20. Muigai, A.W.T. and O. Hanotte, *The Origin of African Sheep: Archaeological and Genetic Perspectives*. Afr Archaeol Rev 2013. **30**: p. 39–50.
21. Gifford-Gonzalez, D. and O. Hanotte, *Domesticating Animals in Africa: Implications of Genetic and Archaeological Findings*. J World Prehist 2011. **24**: p. 1–23.
22. Anati, E., *ARTE RUPESTRE NELLE REGIONI OCCIDENTALI DELLA PENISOLA IBERICA*. 1968, CAPO DI PONTE, VALCAMONICA. 136.
23. Ryder, M., *Domestication, history and breed evolution in sheep*,. 1991, Amsterdam: Elsevier.
24. Zeuner, F., *A History of Domesticated Animals* 1963, London.: Hutchison, C. A.
25. Corbet, G.B., *The Mammals of the Palearctic Region: a taxonomic review*, ed. B.M.N. History). 1978: Cornell University Press, Ithaca.
26. Ryder, M., *Sheep*, in *Evolution of Domesticated Animals*., I. Mason, Editor. 1984, Longman Group Limited; London and New York. p. 63–84.
27. Grubb, P., *Order Artiodactyla*, in *Mammal Species of the World: A Taxonomic and Geographic Reference*, D.E. Wilson, Reeder, D.M., Editor. 2005, Johns Hopkins University Press. p. 267-722.
28. Geist, V., *On the taxonomy of giant sheep (Ovis ammon Linnaeus, 1766)*. Can. J. Zool., 1991. **69**: p. 706–723.
29. Hirsch, S., *Sheep and Goat in palestine*, in *Bulletin of the Palestine Economic Society*. 1933: Palestine.
30. Pritchard, C.J.R., Pennell, A.E., Williams, G.L., *A note on the wool characteristics of sheep at the hofuf Agricultural Research Center*. University College of North wales, 1975.
31. Gürsoy, O., *Small ruminant breeds of Turkey*, in *Characterization of Small Ruminant Breeds in West Asia , North Africa*, L. Iñiguez, Editor. 2005, ICARDA: Aleppo, Syria. p. 239-416.
32. Epstein, H., *Domestic animals of China*. Boerma. Tech. Commun. 1969: Commonw. Bur. Anim. Breed. Genet.
33. FAO, *Lebanon Plan of Action for Resilient Livelihoods 2014 – 2018*. 2014.
34. Hilali, M., et al., *Dietary supplementation with nonconventional feeds from the Middle East: assessing the effects on physicochemical and organoleptic properties of Awassi sheep milk and yogurt*. J Dairy Sci, 2011. **94**(12): p. 5737-49.
35. Hartwell, B.W., Iniguez, L., Mueller, J., Wurzinger, M., Knaus, W. F., *Characterization of Awassi lamb fattening systems: a Syrian case study*. Trop Anim Health Prod, 2010. **42**(7): p. 1573-8.
36. Galal, S., Gürsoy, O ., Shaat, I ., *Awassi sheep as a genetic resource and efforts for their genetic improvement*. Small Ruminant Research, 2008. **23**(79): p. 99-108.
37. Asmar, F.R., *Country Pasture/Forage Resource Profiles*. 2011, FAO: Lebanon.
38. Iñiguez, L., *Small ruminant breeds of Turkey*. Characterization of small ruminant breeds in West Asia, North Africa. ICARDA, 2005. **1**: p. 239-416.
39. Amin, M.Y. and K.-J. Peters, *Awassi Sheep Production and the Development of breeding Program Options in Syria*, in *Prosperity and Poverty in a Globalised World— Challenges for Agricultural Research*. October 11-13, 2006.: Bonn, Germany.
40. Reiad, K., Al-Azzawi, W., Al-Najjar, K., Masri, Y., Salhab, S., Abdo, Z., El-Herek, I., Omed, H., Saatc, M., *Factors Influencing the Milk Production of Awassi Sheep in A Flock With the Selected Lines at the Agricultural Scientific Research Centre in Salamieh/Syria*. Journal of The Faculty of Veterinary Medecine, kafkas university, 2010. **16**(3): p. 425-430.

41. Gürsoy, O., Pekel, E., Özcan, L., Torun, O., Timon, V., *Genetic selection for milk yield in Awassi sheep. I. Reproduction , Lactation*. Doga Turkish Journal of Veterinary , Animal Science, 1992. **16**(3): p. 535-546.
42. Gürsoy, O., Pekel, E., Özcan, L., Torun, O., Timon, V., *Comparisons of production traits of Ceylanpinar Awassi sheep with top producing ewes of national flocks in the GAP area. II. Growth performance , carcass traits*. Doğa Turkish Journal of Veterinary , Animal Science, 1993. **17**: p. 65-72.
43. Vogt, P.K., *Historical Introduction to the General Properties of Retroviruses*, in *Retroviruses*, J.M. Coffin, Hughes, SH., Varmus, HE., Editor. 1997, Cold Spring Harbor Laboratory Press: Cold Spring Harbor-New York.
44. Vallée, H. and H. Carré, *Sur la nature infectieuse de l'anémie du cheval*. C. R. Hebd Seances Acad Sci Ser D Sci Nat, 1904. **139**: p. 331-333.
45. Ellermann, V. and O. Bang, *Experimentelle Leukämie bei Hühnern*. Zentralbl Bakteriol Parasitenkd Infektionskr Hyg Abt Orig, 1908. **46**: p. 595-609.
46. Rous, P., *A Sarcoma of the Fowl Transmissible by an Agent Separable from the Tumor Cells*. J Exp Med, 1911. **13**(4): p. 397-411.
47. Bittner, J.J., *Some Possible Effects of Nursing on the Mammary Gland Tumor Incidence in Mice*. Science, 1936. **84**(2172): p. 162.
48. Gross, L., *Development and serial cellfree passage of a highly potent strain of mouse leukemia virus*. Proc Soc Exp Biol Med, 1957. **94**(4): p. 767-71.
49. Claude, A., Murphy, J.G. , *Transmissible tumors of the fowl*. Physiol. Rev., 1933. **13**: p. 246–275.
50. Payne, L.N., *Biology of avian retroviruses*, in *The retroviridae*, J.A. Levy, Editor. 1992, Plenum Press: New York. p. 299-404.
51. Keogh, E.V., *Ectodermal lesions produced by the virus of Rous sarcoma*. British Journal of experimental pathology, 1938. **19**: p. 1–9.
52. Baltimore, D., *RNA-dependent DNA polymerase in virions of RNA tumour viruses*. Nature, 1970. **226**(5252): p. 1209-11.
53. Temin, H.M. and S. Mizutani, *RNA-dependent DNA polymerase in virions of Rous sarcoma virus*. Nature, 1970. **226**(5252): p. 1211-3.
54. Dalton, A.J., et al., *The case for a family of reverse transcriptase viruses: Retroviridae*. Intervirology, 1974. **4**(4): p. 201-6.
55. Murray, S.M., et al., *Replication in a superficial epithelial cell niche explains the lack of pathogenicity of primate foamy virus infections*. J Virol, 2008. **82**(12): p. 5981-5.
56. Linial, M.L., *Foamy viruses*, in *Fields virology*, B.N. Fields, D.M. Knipe, and P.M. Howley, Editors. 2007, Wolters Kluwer/Lippincott Williams & Wilkins: Philadelphia, PA. p. 2245–2263.
57. Coffin, J.M., *Genetic diversity and evolution of retroviruses*. Curr Top Microbiol Immunol, 1992. **176**: p. 143-64.
58. Vogt, V.M., *Retroviral Virions and Genomes*. 1997.
59. Payne, A.L., D.W. Verwoerd, and H.M. Garnett, *The morphology and morphogenesis of jaagsiekte retrovirus (JSRV)*. Onderstepoort J Vet Res, 1983. **50**(4): p. 317-22.
60. Sfakianos, J.N. and E. Hunter, *M-PMV capsid transport is mediated by Env/Gag interactions at the pericentriolar recycling endosome*. Traffic, 2003. **4**(10): p. 671-80.
61. Murcia, P.R., F. Arnaud, and M. Palmarini, *The transdominant endogenous retrovirus enJS56A1 associates with and blocks intracellular trafficking of Jaagsiekte sheep retrovirus Gag*. J Virol, 2007. **81**(4): p. 1762-72.

62. Yeger, H., V.I. Kalnins, and J.R. Stephenson, *Type-C retrovirus maturation and assembly: post-translational cleavage of the gag-gene coded precursor polypeptide occurs at the cell membrane*. *Virology*, 1978. **89**(1): p. 34-44.
63. Coffin, J.M., S.H. Hughes, and H.E. Varmus, *The Interactions of Retroviruses and their Hosts*. 1997.
64. Swanstrom, R. and J.W. Wills, *Synthesis, Assembly, and Processing of Viral Proteins*. 1997.
65. Rosenberg, N. and P. Jolicoeur, *Retroviral Pathogenesis*, in *Retroviruses*. , J.M. Coffin, S.H. Hughes, and H.E. Varmus, Editors. 1997, Cold Spring Harbor Laboratory Press: New York. p. 475-586.
66. Stoye, J.P., *Endogenous retroviruses: still active after all these years?* *Curr Biol*, 2001. **11**(22): p. R914-6.
67. Kijas, J.W., et al., *Genome-wide analysis of the world's sheep breeds reveals high levels of historic mixture and strong recent selection*. *PLoS Biol*, 2012. **10**(2): p. e1001258.
68. Arnaud, F., et al., *A paradigm for virus-host coevolution: sequential counter-adaptations between endogenous and exogenous retroviruses*. *PLoS Pathog*, 2007. **3**(11): p. e170.
69. Tarlinton, R.E., J. Meers, and P.R. Young, *Retroviral invasion of the koala genome*. *Nature*, 2006. **442**(7098): p. 79-81.
70. Jern, P. and J.M. Coffin, *Effects of retroviruses on host genome function*. *Annu Rev Genet*, 2008. **42**: p. 709-32.
71. Finnegan, D.J., *Eukaryotic transposable elements and genome evolution*. *Trends Genet* 1989. **5**(4): p. 103-107.
72. Smit, A.F., *Interspersed repeats and other mementos of transposable elements in mammalian genomes*. *Curr Opin Genet Dev*, 1999. **9**(6): p. 657-63.
73. SanMiguel, P., et al., *The paleontology of intergene retrotransposons of maize*. *Nat Genet*, 1998. **20**(1): p. 43-5.
74. Lander, E.S., et al., *Initial sequencing and analysis of the human genome*. *Nature*, 2001. **409**(6822): p. 860-921.
75. Capy, P., *Dynamics and evolution of transposable elements*. 1998: Springer.
76. Kapitonov, V.V. and J. Jurka, *Molecular paleontology of transposable elements in the *Drosophila melanogaster* genome*. *Proc Natl Acad Sci U S A*, 2003. **100**(11): p. 6569-74.
77. Biemont, C. and C. Vieira, *Genetics: junk DNA as an evolutionary force*. *Nature*, 2006. **443**(7111): p. 521-4.
78. Bock, M. and J.P. Stoye, *Endogenous retroviruses and the human germline*. *Curr Opin Genet Dev*, 2000. **10**(6): p. 651-5.
79. Dunlap, K.A., M. Palmarini, and T.E. Spencer, *Ovine endogenous betaretroviruses (enJSRVs) and placental morphogenesis*. *Placenta*, 2006. **27 Suppl A**: p. S135-40.
80. Varela, M., et al., *Association of RON tyrosine kinase with the Jaagsiekte sheep retrovirus envelope glycoprotein*. *Virology*, 2006. **350**(2): p. 347-57.
81. Mura, M., et al., *Late viral interference induced by transdominant Gag of an endogenous retrovirus*. *Proc Natl Acad Sci U S A*, 2004. **101**(30): p. 11117-22.
82. Boeke, J.D. and J.P. Stoye, *Retrotransposons, Endogenous Retroviruses, and the Evolution of Retroelements*, in *Retroviruses*, J.H. Coffin, S.H.; Varmus, H.E., Editor. 1997, Cold Spring Harbor Laboratory Press: Plainview, NY. p. 343-436.
83. Coffin, J., *Retroviridae and their replication*, in *Virology*, B.N. Fields, et al, Editor. 1990, Raven Press: New York. p. 1437–1500.

84. Coffin, J.M., *Evolution of retroviruses: fossils in our DNA*. Proc Am Philos Soc, 2004. **148**(3): p. 264-80.
85. Denner, J., *Endogenous retroviruses*, in *Retroviruses: Molecular biology, genomics and pathogenesis*, R. Kurth and N. Bannert, Editors. 2010, Caister Academic Press: Hethersett, UK. p. 35-69.
86. Jern, P., G.O. Sperber, and J. Blomberg, *Use of endogenous retroviral sequences (ERVs) and structural markers for retroviral phylogenetic inference and taxonomy*. Retrovirology, 2005. **2**: p. 50.
87. Coffin, J.M., *Structure and classification of retroviruses*, in *The retroviridae*, J.A. Levy, Editor. 1992, Plenum Press: New York. p. 19-49.
88. Bannert, N., Fiebig, U. and Hohn, O., *Retroviral particles, proteins and genomes*, in *Retroviruses: Molecular Biology, Genomics and Pathogenesis*, R.K.a.N. Bannert, Editor. 2010, Caister Academic Press, U.K. .
89. Lv, F.H., et al., *Mitogenomic Meta-Analysis Identifies Two Phases of Migration in the History of Eastern Eurasian Sheep*. Mol Biol Evol, 2015. **32**(10): p. 2515-33.
90. Zheng, R., T.M. Jenkins, and R. Craigie, *Zinc folds the N-terminal domain of HIV-1 integrase, promotes multimerization, and enhances catalytic activity*. Proc Natl Acad Sci U S A, 1996. **93**(24): p. 13659-64.
91. Nymark-McMahon, M.H. and S.B. Sandmeyer, *Mutations in nonconserved domains of Ty3 integrase affect multiple stages of the Ty3 life cycle*. J Virol, 1999. **73**(1): p. 453-65.
92. Ting, C.N., et al., *Endogenous retroviral sequences are required for tissue-specific expression of a human salivary amylase gene*. Genes Dev, 1992. **6**(8): p. 1457-65.
93. Schulte, A.M., et al., *Human trophoblast and choriocarcinoma expression of the growth factor pleiotrophin attributable to germ-line insertion of an endogenous retrovirus*. Proc Natl Acad Sci U S A, 1996. **93**(25): p. 14759-64.
94. Johansen, L.K. and C.D. Morrow, *Inherent instability of poliovirus genomes containing two internal ribosome entry site (IRES) elements supports a role for the IRES in encapsidation*. J Virol, 2000. **74**(18): p. 8335-42.
95. Cavarec, L., et al., *Molecular cloning and characterization of a transcription factor for the copia retrotransposon with homology to the BTB-containing lola neurogenic factor*. Mol Cell Biol, 1997. **17**(1): p. 482-94.
96. Conte, C., B. Dastugue, and C. Vaury, *Promoter competition as a mechanism of transcriptional interference mediated by retrotransposons*. EMBO J, 2002. **21**(14): p. 3908-16.
97. Matyunina, L.V., I.K. Jordan, and J.F. McDonald, *Naturally occurring variation in copia expression is due to both element (cis) and host (trans) regulatory variation*. Proc Natl Acad Sci U S A, 1996. **93**(14): p. 7097-102.
98. Smith, D.B., et al., *Variation of the hepatitis C virus 5' non-coding region: implications for secondary structure, virus detection and typing*. The International HCV Collaborative Study Group. J Gen Virol, 1995. **76** ( Pt 7): p. 1749-61.
99. Wilson, G.M., M.Z. Vasa, and R.G. Deeley, *Stabilization and cytoskeletal-association of LDL receptor mRNA are mediated by distinct domains in its 3' untranslated region*. J Lipid Res, 1998. **39**(5): p. 1025-32.
100. Dimitrov, D.S., *How do viruses enter cells? The HIV coreceptors teach us a lesson of complexity*. Cell, 1997. **91**(6): p. 721-30.
101. Goff, S.P., *Retroviridae: The retroviruses and their replication*, in *Fields Virology*, D.M. Knipe and P.M. Howley, Editors. 2001, Lippincott Williams & Wilkins: New York, NY, USA. p. 1871-1940.



102. Goff, S.P., *Host factors exploited by retroviruses*. Nat Rev Microbiol, 2007. **5**(4): p. 253-63.
103. Basu, V.P., et al., *Strand transfer events during HIV-1 reverse transcription*. Virus Res, 2008. **134**(1-2): p. 19-38.
104. Telesnitsky, A. and S.P. Goff, *Reverse Transcriptase and the Generation of Retroviral DNA*, in *Retroviruses*, J.M. Coffin, S.H. Hughes, and H.E. Varmus, Editors. 1997: Cold Spring Harbor (NY).
105. Nowak, E., et al., *Structural analysis of monomeric retroviral reverse transcriptase in complex with an RNA/DNA hybrid*. Nucleic Acids Res, 2013. **41**(6): p. 3874-87.
106. Brown, P.O., *Integration*, in *Retroviruses*, J.M. Coffin, S. H. Hughes and H.E. Varmus, Editor. 1997, Cold Spring Harbor Laboratory: New York. p. 161-204.
107. Yu, S.F., et al., *Human foamy virus replication: a pathway distinct from that of retroviruses and hepadnaviruses*. Science, 1996. **271**(5255): p. 1579-82.
108. Demirov, D.G. and E.O. Freed, *Retrovirus budding*. Virus Res, 2004. **106**(2): p. 87-102.
109. Palmarini, M., et al., *Jaagsiekte sheep retrovirus is necessary and sufficient to induce a contagious lung cancer in sheep*. J Virol, 1999. **73**(8): p. 6964-72.
110. Palmarini, M. and H. Fan, *Retrovirus-induced ovine pulmonary adenocarcinoma, an animal model for lung cancer*. J Natl Cancer Inst, 2001. **93**(21): p. 1603-14.
111. Tustin, R.C., *Ovine Jaagsiekte*. Journal of South African Veterinary Medical Association 1969. **40**: p. 3-23.
112. Tustin, R.C., et al., *Experimental transmission of jaagsiekte (ovine pulmonary adenomatosis) to goats*. Onderstepoort J Vet Res, 1988. **55**(1): p. 27-32.
113. York, D.F. and G. Querat, *A history of ovine pulmonary adenocarcinoma (jaagsiekte) and experiments leading to the deduction of the JSRV nucleotide sequence*. Curr Top Microbiol Immunol, 2003. **275**: p. 1-23.
114. Verwoerd, D.W. and E.M. de Villiers, *On the aetiology of Jaagsiekte*. J S Afr Vet Assoc, 1980. **51**(2): p. 71-4.
115. Sharp, J.M., et al., *Rapid transmission of sheep pulmonary adenomatosis (jaagsiekte) in young lambs. Brief report*. Arch Virol, 1983. **78**(1-2): p. 89-95.
116. Palmarini, M., et al., *Jaagsiekte retrovirus establishes a disseminated infection of the lymphoid tissues of sheep affected by pulmonary adenomatosis*. J Gen Virol, 1996. **77** ( Pt 12): p. 2991-8.
117. Griffiths, D.J., H.M. Martineau, and C. Cousens, *Pathology and pathogenesis of ovine pulmonary adenocarcinoma*. J Comp Pathol, 2010. **142**(4): p. 260-83.
118. Sharp, J.M. and J.C. DeMartini, *Natural history of JSRV in sheep*. Curr Top Microbiol Immunol, 2003. **275**: p. 55-79.
119. De las Heras, M., L. Gonzalez, and J.M. Sharp, *Pathology of ovine pulmonary adenocarcinoma*. Current Topics in Microbiology and Immunology, 2003. **275**: p. 25-54.
120. OIE, *OVINE PULMONARY ADENOCARCINOMA*, in *Terrestrial Manual 2014*. 2014.
121. Cousens, C., et al., *Prevalence of ovine pulmonary adenocarcinoma (Jaagsiekte) in a UK slaughterhouse sheep study*. Vet Rec, 2015. **176**(16): p. 413.
122. Cousens, C., et al., *Jaagsiekte sheep retrovirus is present at high concentration in lung fluid produced by ovine pulmonary adenocarcinoma-affected sheep and can survive for several weeks at ambient temperatures*. Res Vet Sci, 2009. **87**(1): p. 154-6.
123. Fan, H., M. Palmarini, and J.C. DeMartini, *Transformation and oncogenesis by jaagsiekte sheep retrovirus*. Curr Top Microbiol Immunol, 2003. **275**: p. 139-77.

124. Caporale, M., et al., *Infection of lung epithelial cells and induction of pulmonary adenocarcinoma is not the most common outcome of naturally occurring JSRV infection during the commercial lifespan of sheep*. *Virology*, 2005. **338**(1): p. 144-53.
125. Grego, E., et al., *Colostrum and milk can transmit jaagsiekte retrovirus to lambs*. *Vet Microbiol*, 2008. **130**(3-4): p. 247-57.
126. Caporale, M., et al., *Host species barriers to Jaagsiekte sheep retrovirus replication and carcinogenesis*. *J Virol*, 2013. **87**(19): p. 10752-62.
127. Ortin, A., et al., *Lack of a specific immune response against a recombinant capsid protein of Jaagsiekte sheep retrovirus in sheep and goats naturally affected by enzootic nasal tumour or sheep pulmonary adenomatosis*. *Vet Immunol Immunopathol*, 1998. **61**(2-4): p. 229-37.
128. Spencer, T.E., et al., *Receptor usage and fetal expression of ovine endogenous betaretroviruses: implications for coevolution of endogenous and exogenous retroviruses*. *J Virol*, 2003. **77**(1): p. 749-53.
129. Sun, S., J.H. Schiller, and A.F. Gazdar, *Lung cancer in never smokers--a different disease*. *Nat Rev Cancer*, 2007. **7**(10): p. 778-90.
130. Perk, K., I. Hod, and T.A. Nobel, *Pulmonary adenomatosis of sheep (jaagsiekte). I. Ultrastructure of the tumor*. *J Natl Cancer Inst*, 1971. **46**(3): p. 525-37.
131. Pott, U., et al., *A new Cys2/His2 zinc finger gene, rKr2, is expressed in differentiated rat oligodendrocytes and encodes a protein with a functional repressor domain*. *J Neurochem*, 1995. **65**(5): p. 1955-66.
132. Palmarini, M., H. Fan, and J.M. Sharp, *Sheep pulmonary adenomatosis: a unique model of retrovirus-associated lung cancer*. *Trends Microbiol*, 1997. **5**(12): p. 478-83.
133. Leroux, C., et al., *Jaagsiekte Sheep Retrovirus (JSRV): from virus to lung cancer in sheep*. *Vet Res*, 2007. **38**(2): p. 211-28.
134. Miller, A.D., et al., *Evidence against a role for jaagsiekte sheep retrovirus in human lung cancer*. *Retrovirology*, 2017. **14**(1): p. 3.
135. Palmarini, M. and H. Fan, *Molecular biology of jaagsiekte sheep retrovirus*. *Curr Top Microbiol Immunol*, 2003. **275**: p. 81-115.
136. York, D.F., et al., *Nucleotide sequence of the jaagsiekte retrovirus, an exogenous and endogenous type D and B retrovirus of sheep and goats*. *J Virol*, 1992. **66**(8): p. 4930-9.
137. Rosati, S., et al., *An accessory open reading frame (orf-x) of jaagsiekte sheep retrovirus is conserved between different virus isolates*. *Virus Res*, 2000. **66**(1): p. 109-16.
138. Maeda, N., et al., *Direct transformation of rodent fibroblasts by jaagsiekte sheep retrovirus DNA*. *Proc Natl Acad Sci U S A*, 2001. **98**(8): p. 4449-54.
139. Varela, M., Spencer, T. E., Palmarini, M., Arnaud, F., *Friendly viruses: the special relationship between endogenous retroviruses and their host*. *Ann N Y Acad Sci*, 2009. **1178**: p. 157-72.
140. DeMartini, J.C., et al., *Retrovirus-induced lung cancer: mechanisms of transformation of alveolar type II epithelial cells*. *Chest*, 2004. **125**(5 Suppl): p. 106S-7S.
141. Wootton, S.K., C.L. Halbert, and A.D. Miller, *Sheep retrovirus structural protein induces lung tumours*. *Nature*, 2005. **434**(7035): p. 904-7.
142. Caporale, M., et al., *Expression of the jaagsiekte sheep retrovirus envelope glycoprotein is sufficient to induce lung tumors in sheep*. *J Virol*, 2006. **80**(16): p. 8030-7.

143. DeMartini, J.C., et al., *Jaagsiekte sheep retrovirus proviral clone JSRV(JS7), derived from the JS7 lung tumor cell line, induces ovine pulmonary carcinoma and is integrated into the surfactant protein A gene.* J Virol, 2001. **75**(9): p. 4239-46.
144. Liu, S.L. and A.D. Miller, *Oncogenic transformation by the jaagsiekte sheep retrovirus envelope protein.* Oncogene, 2007. **26**(6): p. 789-801.
145. Allen, T.E., et al., *The jaagsiekte sheep retrovirus envelope gene induces transformation of the avian fibroblast cell line DF-1 but does not require a conserved SH2 binding domain.* J Gen Virol, 2002. **83**(Pt 11): p. 2733-42.
146. Liu, S.L. and A.D. Miller, *Transformation of madin-darby canine kidney epithelial cells by sheep retrovirus envelope proteins.* J Virol, 2005. **79**(2): p. 927-33.
147. Rai, S.K., et al., *Candidate tumor suppressor HYAL2 is a glycosylphosphatidylinositol (GPI)-anchored cell-surface receptor for jaagsiekte sheep retrovirus, the envelope protein of which mediates oncogenic transformation.* Proc Natl Acad Sci U S A, 2001. **98**(8): p. 4443-8.
148. Chitra, E., et al., *Generation and characterization of JSRV envelope transgenic mice in FVB background.* Virology, 2009. **393**(1): p. 120-6.
149. Hofacre, A. and H. Fan, *Jaagsiekte sheep retrovirus biology and oncogenesis.* Viruses, 2011. **2**(12): p. 2618-48.
150. Hofacre, A. and H. Fan, *Jaagsiekte sheep retrovirus biology and oncogenesis.* Viruses, 2010. **2**(12): p. 2618-48.
151. Cully, M., et al., *Beyond PTEN mutations: the PI3K pathway as an integrator of multiple inputs during tumorigenesis.* Nat Rev Cancer, 2006. **6**(3): p. 184-92.
152. Brunet, A., et al., *Akt promotes cell survival by phosphorylating and inhibiting a Forkhead transcription factor.* Cell, 1999. **96**(6): p. 857-68.
153. Danilkovitch-Miagkova, A., et al., *Hyaluronidase 2 negatively regulates RON receptor tyrosine kinase and mediates transformation of epithelial cells by jaagsiekte sheep retrovirus.* Proc Natl Acad Sci U S A, 2003. **100**(8): p. 4580-5.
154. Hull, S. and H. Fan, *Mutational analysis of the cytoplasmic tail of jaagsiekte sheep retrovirus envelope protein.* J Virol, 2006. **80**(16): p. 8069-80.
155. Maeda, N. and H. Fan, *Signal transduction pathways utilized by enzootic nasal tumor virus (ENTV-1) envelope protein in transformation of rat epithelial cells resemble those used by jaagsiekte sheep retrovirus.* Virus Genes, 2008. **36**(1): p. 147-55.
156. Murgia, C., et al., *Lung adenocarcinoma originates from retrovirus infection of proliferating type 2 pneumocytes during pulmonary post-natal development or tissue repair.* PLoS Pathog, 2011. **7**(3): p. e1002014.
157. Berns, A., *Stem cells for lung cancer?* Cell, 2005. **121**(6): p. 811-3.
158. Kim, C.F., et al., *Identification of bronchioalveolar stem cells in normal lung and lung cancer.* Cell, 2005. **121**(6): p. 823-35.
159. Dirks, C., et al., *Mechanism of cell entry and transformation by enzootic nasal tumor virus.* J Virol, 2002. **76**(5): p. 2141-9.
160. Miller, A.D., *Identification of Hyal2 as the cell-surface receptor for jaagsiekte sheep retrovirus and ovine nasal adenocarcinoma virus.* Curr Top Microbiol Immunol, 2003. **275**: p. 179-99.
161. Lepperdinger, G., J. Mullegger, and G. Kreil, *Hyal2--less active, but more versatile?* Matrix Biol, 2001. **20**(8): p. 509-14.
162. Vigdorovich, V., R.K. Strong, and A.D. Miller, *Expression and characterization of a soluble, active form of the jaagsiekte sheep retrovirus receptor, Hyal2.* J Virol, 2005. **79**(1): p. 79-86.

163. Vigdorovich, V., A.D. Miller, and R.K. Strong, *Ability of hyaluronidase 2 to degrade extracellular hyaluronan is not required for its function as a receptor for jaagsiekte sheep retrovirus*. J Virol, 2007. **81**(7): p. 3124-9.
164. Palmarini, M., et al., *The long terminal repeat of Jaagsiekte sheep retrovirus is preferentially active in differentiated epithelial cells of the lungs*. J Virol, 2000. **74**(13): p. 5776-87.
165. McGee-Estrada, K. and H. Fan, *In vivo and in vitro analysis of factor binding sites in Jaagsiekte sheep retrovirus long terminal repeat enhancer sequences: roles of HNF-3, NF-I, and C/EBP for activity in lung epithelial cells*. J Virol, 2006. **80**(1): p. 332-41.
166. Bai, J., et al., *Sequence comparison of JSRV with endogenous proviruses: envelope genotypes and a novel ORF with similarity to a G-protein-coupled receptor*. Virology, 1999. **258**(2): p. 333-43.
167. Palmarini, M., et al., *Molecular cloning and functional analysis of three type D endogenous retroviruses of sheep reveal a different cell tropism from that of the highly related exogenous jaagsiekte sheep retrovirus*. J Virol, 2000. **74**(17): p. 8065-76.
168. Arnaud, F., et al., *Coevolution of endogenous betaretroviruses of sheep and their host*. Cell Mol Life Sci, 2008. **65**(21): p. 3422-32.
169. Armezzani, A., et al., *"Menage a Trois": the evolutionary interplay between JSRV, enJSRVs and domestic sheep*. Viruses, 2014. **6**(12): p. 4926-45.
170. Arnaud, F., P.R. Murcia, and M. Palmarini, *Mechanisms of late restriction induced by an endogenous retrovirus*. J Virol, 2007. **81**(20): p. 11441-51.
171. Armezzani, A., et al., *The signal peptide of a recently integrated endogenous sheep betaretrovirus envelope plays a major role in eluding gag-mediated late restriction*. J Virol, 2011. **85**(14): p. 7118-28.
172. Chessa, B., et al., *Revealing the history of sheep domestication using retrovirus integrations*. Science, 2009. **324**(5926): p. 532-6.
173. Bruford, M.W., D.G. Bradley, and G. Luikart, *DNA markers reveal the complexity of livestock domestication*. Nat Rev Genet, 2003. **4**(11): p. 900-10.
174. Ladoukakis, E.D. and E. Zouros, *Evolution and inheritance of animal mitochondrial DNA: rules and exceptions*. J Biol Res (Thessalon), 2017. **24**: p. 2.
175. Alberts, B., et al., *Molecular biology of the cell, 3rd edition*. 1994, Garland Publishing. 1294
176. Penta, J.S., et al., *Mitochondrial DNA in human malignancy*. Mutat Res, 2001. **488**(2): p. 119-33.
177. Giles, R.E., et al., *Maternal inheritance of human mitochondrial DNA*. Proc Natl Acad Sci U S A, 1980. **77**(11): p. 6715-9.
178. Hayashi, J.I., et al., *Strictly maternal inheritance of rat mitochondrial DNA*. Biochem Biophys Res Commun, 1978. **83**(3): p. 1032-8.
179. Hutchison, C.A., 3rd, et al., *Maternal inheritance of mammalian mitochondrial DNA*. Nature, 1974. **251**(5475): p. 536-8.
180. Gyllensten, U., et al., *Paternal inheritance of mitochondrial DNA in mice*. Nature, 1991. **352**(6332): p. 255-7.
181. Piganeau, G., M. Gardner, and A. Eyre-Walker, *A broad survey of recombination in animal mitochondria*. Mol Biol Evol, 2004. **21**(12): p. 2319-25.
182. Piganeau, G. and A. Eyre-Walker, *A reanalysis of the indirect evidence for recombination in human mitochondrial DNA*. Heredity (Edinb), 2004. **92**(4): p. 282-8.

183. Bailey, L.J. and A.J. Doherty, *Mitochondrial DNA replication: a PrimPol perspective*. *Biochem Soc Trans*, 2017. **45**(2): p. 513-529.
184. Taanman, J.W., *The mitochondrial genome: structure, transcription, translation and replication*. *Biochim Biophys Acta*, 1999. **1410**(2): p. 103-23.
185. Wilson, A.C., et al., *Mitochondrial DNA and two perspectives on evolutionary genetics*. *Biological Journal of the Linnean Society*, 1985. **26**(4): p. 375-400.
186. Avise, J., et al., *Intraspecific phylogeography - the mitochondrial-DNA bridge between population genetics and systematics*. *Annual Review of Ecology and Systematics*, 1987. **18**: p. 489-522.
187. Margulis, L., *Symbiosis and evolution*. *Sci Am*, 1971. **225**(2): p. 48-57.
188. Gray, M.W., G. Burger, and B.F. Lang, *Mitochondrial evolution*. *Science*, 1999. **283**(5407): p. 1476-81.
189. Gray, M.W., *Rickettsia, typhus and the mitochondrial connection*. *Nature*, 1998. **396**(6707): p. 109-10.
190. Baldauf, S.L., J.D. Palmer, and W.F. Doolittle, *The root of the universal tree and the origin of eukaryotes based on elongation factor phylogeny*. *Proc Natl Acad Sci U S A*, 1996. **93**(15): p. 7749-54.
191. Harris, J.K., et al., *The genetic core of the universal ancestor*. *Genome Res*, 2003. **13**(3): p. 407-12.
192. Woese, C.R., O. Kandler, and M.L. Wheelis, *Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya*. *Proc Natl Acad Sci U S A*, 1990. **87**(12): p. 4576-9.
193. Cox, C.J., et al., *The archaeobacterial origin of eukaryotes*. *Proc Natl Acad Sci U S A*, 2008. **105**(51): p. 20356-61.
194. Feng, D.F., G. Cho, and R.F. Doolittle, *Determining divergence times with a protein clock: update and reevaluation*. *Proc Natl Acad Sci U S A*, 1997. **94**(24): p. 13028-33.
195. Sicheritz-Ponten, T., C.G. Kurland, and S.G. Andersson, *A phylogenetic analysis of the cytochrome b and cytochrome c oxidase I genes supports an origin of mitochondria from within the Rickettsiaceae*. *Biochim Biophys Acta*, 1998. **1365**(3): p. 545-51.
196. Wolf, Y.I. and E.V. Koonin, *Genome reduction as the dominant mode of evolution*. *Bioessays*, 2013. **35**(9): p. 829-37.
197. Vafai, S.B. and V.K. Mootha, *Mitochondrial disorders as windows into an ancient organelle*. *Nature*, 2012. **491**(7424): p. 374-83.
198. Cree, L.M., et al., *A reduction of mitochondrial DNA molecules during embryogenesis explains the rapid segregation of genotypes*. *Nat Genet*, 2008. **40**(2): p. 249-54.
199. Andersson, S.G. and C.G. Kurland, *Reductive evolution of resident genomes*. *Trends Microbiol*, 1998. **6**(7): p. 263-8.
200. Boore, J.L., *Animal mitochondrial genomes*. *Nucleic Acids Res*, 1999. **27**(8): p. 1767-80.
201. Mourier, T., et al., *The Human Genome Project reveals a continuous transfer of large mitochondrial fragments to the nucleus*. *Mol Biol Evol*, 2001. **18**(9): p. 1833-7.
202. Lee, S.R. and J. Han, *Mitochondrial Nucleoid: Shield and Switch of the Mitochondrial Genome*. *Oxid Med Cell Longev*, 2017. **2017**: p. 8060949.
203. <https://fr.wikipedia.org/wiki/Mitochondrie>.
204. Wikipidea, N.i.o.h.N.H.G.R.I.

205. Chandel, N.S., *Evolution of Mitochondria as Signaling Organelles*. Cell Metab, 2015. **22**(2): p. 204-6.
206. Sinha, A., et al., *Improving the vitamin D status of vitamin D deficient adults is associated with improved mitochondrial oxidative function in skeletal muscle*. J Clin Endocrinol Metab, 2013. **98**(3): p. E509-13.
207. Cheng, Z. and M. Ristow, *Mitochondria and metabolic homeostasis*. Antioxid Redox Signal, 2013. **19**(3): p. 240-2.
208. Ahn, C.S. and C.M. Metallo, *Mitochondria as biosynthetic factories for cancer proliferation*. Cancer Metab, 2015. **3**(1): p. 1.
209. Robin, E.D. and R. Wong, *Mitochondrial DNA molecules and virtual number of mitochondria per cell in mammalian cells*. J Cell Physiol, 1988. **136**(3): p. 507-13.
210. Michaels, G.S., W.W. Hauswirth, and P.J. Laipis, *Mitochondrial DNA copy number in bovine oocytes and somatic cells*. Dev Biol, 1982. **94**(1): p. 246-51.
211. Shuster, R.C., A.J. Rubenstein, and D.C. Wallace, *Mitochondrial DNA in anucleate human blood cells*. Biochem Biophys Res Commun, 1988. **155**(3): p. 1360-5.
212. Wiesner, R.J., J.C. Ruegg, and I. Morano, *Counting target molecules by exponential polymerase chain reaction: copy number of mitochondrial DNA in rat tissues*. Biochem Biophys Res Commun, 1992. **183**(2): p. 553-9.
213. Galtier, N., et al., *Mitochondrial DNA as a marker of molecular diversity: a reappraisal*. Mol Ecol, 2009. **18**(22): p. 4541-50.
214. Anderson, S., et al., *Sequence and organization of the human mitochondrial genome*. Nature, 1981. **290**(5806): p. 457-65.
215. Andrews, R.M., et al., *Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA*. Nat Genet, 1999. **23**(2): p. 147.
216. Hiendleder, S., et al., *The complete mitochondrial DNA sequence of the domestic sheep (*Ovis aries*) and comparison with the other major ovine haplotype*. J Mol Evol, 1998. **47**(4): p. 441-8.
217. Wolstenholme, D.R., *Animal mitochondrial DNA: structure and evolution*. Int Rev Cytol, 1992. **141**: p. 173-216.
218. Battey, J. and D.A. Clayton, *The transcription map of mouse mitochondrial DNA*. Cell, 1978. **14**(1): p. 143-56.
219. Dawid, I., et al., in *The Genetic Function of Mitochondrial DNA*, C. Saccone and A. Kroon, Editors. 1976: North-Holland, Amsterdam. p. 3-13.
220. Uhler, J.P. and M. Falkenberg, *Primer removal during mammalian mitochondrial DNA replication*. DNA Repair (Amst), 2015. **34**: p. 28-38.
221. Kasamatsu, H. and J. Vinograd, *Replication of circular DNA in eukaryotic cells*. Annu Rev Biochem, 1974. **43**(0): p. 695-719.
222. Ojala, D., J. Montoya, and G. Attardi, *tRNA punctuation model of RNA processing in human mitochondria*. Nature, 1981. **290**(5806): p. 470-4.
223. Hu, X.D. and L.Z. Gao, *The complete mitochondrial genome of domestic sheep, *Ovis aries**. Mitochondrial DNA A DNA Mapp Seq Anal, 2014. **27**(2): p. 1425-7.
224. Osawa, S., et al., *Recent evidence for evolution of the genetic code*. Microbiol Rev, 1992. **56**(1): p. 229-64.
225. Macreadie, I.G., et al., *Biogenesis of mitochondria: the mitochondrial gene (*aap1*) coding for mitochondrial ATPase subunit 8 in *Saccharomyces cerevisiae**. Nucleic Acids Res, 1983. **11**(13): p. 4435-51.
226. Chomyn, A., et al., *URF6, last unidentified reading frame of human mtDNA, codes for an NADH dehydrogenase subunit*. Science, 1986. **234**(4776): p. 614-8.
227. Li, H., et al., *Physiology and pathophysiology of mitochondrial DNA*. Adv Exp Med Biol, 2012. **942**: p. 39-51.

228. Kondo, R., et al., *Incomplete maternal transmission of mitochondrial DNA in Drosophila*. Genetics, 1990. **126**(3): p. 657-63.
229. Aquadro, C.F. and B.D. Greenberg, *Human mitochondrial DNA variation and evolution: analysis of nucleotide sequences from seven individuals*. Genetics, 1983. **103**(2): p. 287-312.
230. Saccone, C., G. Pesole, and E. Sbisà, *The main regulatory region of mammalian mitochondrial DNA: structure-function model and evolutionary pattern*. J Mol Evol, 1991. **33**(1): p. 83-91.
231. Nesheva, D., *Aspects of ancient mitochondrial DNA analysis in different populations for understanding human evolution*. Balkan J Med Genet, 2014. **17**(1): p. 5-14.
232. Stoneking, M. and H. Soodyall, *Human evolution and the mitochondrial genome*. Curr Opin Genet Dev, 1996. **6**(6): p. 731-6.
233. Wallace, D.C., M.D. Brown, and M.T. Lott, *Mitochondrial DNA variation in human evolution and disease*. Gene, 1999. **238**(1): p. 211-30.
234. Birky, C.W., Jr., *The inheritance of genes in mitochondria and chloroplasts: laws, mechanisms, and models*. Annu Rev Genet, 2001. **35**: p. 125-48.
235. Mishra, P. and D.C. Chan, *Mitochondrial dynamics and inheritance during cell division, development and disease*. Nat Rev Mol Cell Biol, 2014. **15**(10): p. 634-46.
236. Moses, M.J., *Spermiogenesis in the crayfish (Procambarus clarkii) II. Description of stages*. J Biophys Biochem Cytol, 1961. **10**: p. 301-33.
237. Ursprung, H. and E. Schabtach, *Fertilization in tunicates: loss of the paternal mitochondrion prior to sperm entry*. J Exp Zool, 1965. **159**(3): p. 379-83.
238. Anderson, W., *structure and fate of the parental mitochondrion during early embryogenesis of Paracentrotus lividus*. Journal of Ultrastructure Research, 1968. **24**: p. 311-321.
239. Sutovsky, P., et al., *Ubiquitin tag for sperm mitochondria*. Nature, 1999. **402**(6760): p. 371-2.
240. Sato, M. and K. Sato, *Maternal inheritance of mitochondrial DNA by diverse mechanisms to eliminate paternal mitochondrial DNA*. Biochim Biophys Acta, 2013. **1833**(8): p. 1979-84.
241. Piko, L. and L. Matsumoto, *Number of mitochondria and some properties of mitochondrial DNA in the mouse egg*. Dev Biol, 1976. **49**(1): p. 1-10.
242. Hecht, N.B., et al., *Maternal inheritance of the mouse mitochondrial genome is not mediated by a loss or gross alteration of the paternal mitochondrial DNA or by methylation of the oocyte mitochondrial DNA*. Dev Biol, 1984. **102**(2): p. 452-61.
243. Meusel, M.S. and R.F. Moritz, *Transfer of paternal mitochondrial DNA during fertilization of honeybee (Apis mellifera L.) eggs*. Curr Genet, 1993. **24**(6): p. 539-43.
244. Zhao, X., et al., *Further evidence for paternal inheritance of mitochondrial DNA in the sheep (Ovis aries)*. Heredity (Edinb), 2004. **93**(4): p. 399-403.
245. Kaneda, H., et al., *Elimination of paternal mitochondrial DNA in intraspecific crosses during early mouse embryogenesis*. Proc Natl Acad Sci U S A, 1995. **92**(10): p. 4542-6.
246. Kvist, L., et al., *Paternal leakage of mitochondrial DNA in the great tit (Parus major)*. Mol Biol Evol, 2003. **20**(2): p. 243-7.
247. Zouros, E., et al., *Direct evidence for extensive paternal mitochondrial DNA inheritance in the marine mussel Mytilus*. Nature, 1992. **359**(6394): p. 412-4.
248. Schwartz, M. and J. Vissing, *Paternal inheritance of mitochondrial DNA*. N Engl J Med, 2002. **347**(8): p. 576-80.

249. Pakendorf, B. and M. Stoneking, *Mitochondrial DNA and human evolution*. *Annu Rev Genomics Hum Genet*, 2005. **6**: p. 165-83.
250. Larsson, N.G. and D.A. Clayton, *Molecular genetic aspects of human mitochondrial disorders*. *Annu Rev Genet*, 1995. **29**: p. 151-78.
251. Howell, N., et al., *Mitochondrial gene segregation in mammals: is the bottleneck always narrow?* *Hum Genet*, 1992. **90**(1-2): p. 117-20.
252. Bergstrom, C.T. and J. Pritchard, *Germline bottlenecks and the evolutionary maintenance of mitochondrial genomes*. *Genetics*, 1998. **149**(4): p. 2135-46.
253. Felsenstein, J., *The evolutionary advantage of recombination*. *Genetics*, 1974. **78**(2): p. 737-56.
254. McLeod, B.A. and B.N. White, *Tracking mtDNA heteroplasmy through multiple generations in the North Atlantic right whale (*Eubalaena glacialis*)*. *J Hered*, 2009. **101**(2): p. 235-9.
255. Brown, W.M., M. George, Jr., and A.C. Wilson, *Rapid evolution of animal mitochondrial DNA*. *Proc Natl Acad Sci U S A*, 1979. **76**(4): p. 1967-71.
256. Ingman, M., et al., *Mitochondrial genome variation and the origin of modern humans*. *Nature*, 2000. **408**(6813): p. 708-13.
257. Muller, H.J., *The Relation of Recombination to Mutational Advance*. *Mutat Res*, 1964. **106**: p. 2-9.
258. Taylor, R.W. and D.M. Turnbull, *Mitochondrial DNA mutations in human disease*. *Nat Rev Genet*, 2005. **6**(5): p. 389-402.
259. Krings, M., et al., *Neandertal DNA sequences and the origin of modern humans*. *Cell*, 1997. **90**(1): p. 19-30.
260. Gissi, C., et al., *Lineage-specific evolutionary rate in mammalian mtDNA*. *Mol Biol Evol*, 2000. **17**(7): p. 1022-31.
261. Bensasson, D., et al., *Mitochondrial pseudogenes: evolution's misplaced witnesses*. *Trends Ecol Evol*, 2001. **16**(6): p. 314-321.
262. Ballard, J.W. and M.C. Whitlock, *The incomplete natural history of mitochondria*. *Mol Ecol*, 2004. **13**(4): p. 729-44.
263. Ballard, J.W.O. and D.M. Rand, *The population biology of mitochondrial DNA and its phylogenetic implications*. *Annual Review Of Ecology Evolution And Systematics*, 2005. **36**: p. 621-642.
264. Avise, J.C., *Molecular Markers, Natural History, and Evolution*. 1994., New York, NY.: Springer US. 511.
265. Birky, C.W., Jr., P. Fuerst, and T. Maruyama, *Organelle gene diversity under migration, mutation, and drift: equilibrium expectations, approach to equilibrium, effects of heteroplasmic cells, and comparison to nuclear genes*. *Genetics*, 1989. **121**(3): p. 613-27.
266. Leroy, G., et al., *How do introgression events shape the partitioning of diversity among breeds: a case study in sheep*. *Genet Sel Evol*, 2015. **47**: p. 48.
267. Liu, C.Y., et al., *Mitochondrial DNA mutation and depletion increase the susceptibility of human cells to apoptosis*. *Ann N Y Acad Sci*, 2004. **1011**: p. 133-45.
268. Attardi, G., et al., *Identification and mapping of human mitochondrial genes*. *Cytogenet Cell Genet*, 1982. **32**(1-4): p. 85-98.
269. Nelson, D.L.a.C., M.M., *Lehninger Principles of Biochemistry (3rd Edition)*. 2000.
270. Linke, P. and H. Weiss, *Reconstitution of ubiquinol-cytochrome-c reductase from *Neurospora* mitochondria with regard to subunits I and II*. *Methods Enzymol*, 1986. **126**: p. 201-10.



271. Crivellone, M.D., M.A. Wu, and A. Tzagoloff, *Assembly of the mitochondrial membrane system. Analysis of structural mutants of the yeast coenzyme QH<sub>2</sub>-cytochrome c reductase complex*. J Biol Chem, 1988. **263**(28): p. 14323-33.
272. Oudshoorn, P., et al., *Subunit II of yeast QH<sub>2</sub>:cytochrome-c oxidoreductase. Nucleotide sequence of the gene and features of the protein*. Eur J Biochem, 1987. **163**(1): p. 97-103.
273. Tzagoloff, A., M.A. Wu, and M. Crivellone, *Assembly of the mitochondrial membrane system. Characterization of COR1, the structural gene for the 44-kilodalton core protein of yeast coenzyme QH<sub>2</sub>-cytochrome c reductase*. J Biol Chem, 1986. **261**(36): p. 17163-9.
274. Esposti, M.D., et al., *Mitochondrial cytochrome b: evolution and structure of the protein*. Biochim Biophys Acta, 1993. **1143**(3): p. 243-71.
275. Wikstrom, M. and K. Krab, *The semiquinone cycle. A hypothesis of electron transfer and proton translocation in cytochrome bc-type complexes*. J Bioenerg Biomembr, 1986. **18**(3): p. 181-93.
276. Saraste, M., *Location of haem-binding sites in the mitochondrial cytochrome b*. FEBS Lett, 1984. **166**(2): p. 367-72.
277. Trumppower, B.L., *Cytochrome bc<sub>1</sub> complexes of microorganisms*. Microbiol Rev, 1990. **54**(2): p. 101-29.
278. Vigilant, L., et al., *Mitochondrial DNA sequences in single hairs from a southern African population*. Proc Natl Acad Sci U S A, 1989. **86**(23): p. 9350-4.
279. Irwin, D.M., T.D. Kocher, and A.C. Wilson, *Evolution of the cytochrome b gene of mammals*. J Mol Evol, 1991. **32**(2): p. 128-44.
280. Meyer, A., *Shortcomings Of The Cytochrome-b Gene As A Molecular Marker*. Trends In Ecology & Evolution, 1994. **9**: p. 278-280.
281. Castresana, J., *Cytochrome b phylogeny and the taxonomy of great apes and mammals*. Mol Biol Evol, 2001. **18**(4): p. 465-71.
282. Seddigh, S. and M. Darabi, *Functional, structural, and phylogenetic analysis of mitochondrial cytochrome b (cytb) in insects*. Mitochondrial DNA A DNA Mapp Seq Anal, 2017: p. 1-17.
283. REZAEI, H.R., *Phylogénie moléculaire du Genre Ovis (Mouton et Mouflons), Implications pour la Conservation du Genre et pour l'Origine de l'Espèce Domestique*, in *Ecologie, Environnement*. 2007, Université de Grenoble. p. 175.
284. Tsalkin, V.I., *European and Asian mountain sheep (Gornje barany evropy i asii)*. Moskovskoe Obshchest. Iss. Privody., Moscow Society of Naturalists Zool. Sect.(Moskovskoe Obshchest. Isp. Prirody) Materials for the recognition of the fauna and flora of the USSR 27.Moscow., 1951.
285. Haltenorth, T., *Klassifikation der Säugethiere: artiodactyla*. Handbuch der Zoologie? Walter de Gruyter, Berlin Germany., 1963. **8**(32).
286. Valdez, R., *The wild sheep of the world*. 1982
287. Wilson, D. and D. Reeder, *Mammal species of the world: a taxonomic and geographic reference*. Second edn. . 1993, Washington, D.C.: Smithsonian Institution Press.
288. Nadler, C.F., et al., *Cytogenetic differentiation, geographic distribution and domestication of palaeartic sheep (Ovis)*. Zeitschrift für Säugetierkunde, 1973. **38**: p. 109-125.
289. Nicholls, T.J. and M. Minczuk, *In D-loop: 40 years of mitochondrial 7S DNA*. Exp Gerontol, 2014. **56**: p. 175-81.

290. Walberg, M.W. and D.A. Clayton, *Sequence and properties of the human KB cell and mouse L cell D-loop regions of mitochondrial DNA*. Nucleic Acids Res, 1981. **9**(20): p. 5411-21.
291. Doda, J.N., C.T. Wright, and D.A. Clayton, *Elongation of displacement-loop strands in human and mouse mitochondrial DNA is arrested near specific template sequences*. Proc Natl Acad Sci U S A, 1981. **78**(10): p. 6116-20.
292. Saccone, C., M. Attimonelli, and E. Sbisà, *Structural elements highly preserved during the evolution of the D-loop-containing region in vertebrate mitochondrial DNA*. J Mol Evol, 1987. **26**(3): p. 205-11.
293. Brown, T.A. and D.A. Clayton, *Release of replication termination controls mitochondrial DNA copy number after depletion with 2',3'-dideoxycytidine*. Nucleic Acids Res, 2002. **30**(9): p. 2004-10.
294. Zardoya, R., et al., *Nucleotide sequence of the sheep mitochondrial DNA D-loop and its flanking tRNA genes*. Curr Genet, 1995. **28**(1): p. 94-6.
295. Clayton, D.A., *Transcription and replication of mitochondrial DNA*. Hum Reprod, 2000. **15 Suppl 2**: p. 11-7.
296. Pham, X.H., et al., *Conserved sequence box II directs transcription termination and primer formation in mitochondria*. J Biol Chem, 2006. **281**(34): p. 24647-52.
297. Brown, J.R., A.T. Beckenbach, and M.J. Smith, *Intraspecific DNA sequence variation of the mitochondrial control region of white sturgeon (Acipenser transmontanus)*. Mol Biol Evol, 1993. **10**(2): p. 326-41.
298. Sharawat, S.K., et al., *Mitochondrial D-loop variations in paediatric acute myeloid leukaemia: a potential prognostic marker*. Br J Haematol, 2010. **149**(3): p. 391-8.
299. Mambo, E., et al., *Electrophile and oxidant damage of mitochondrial DNA leading to rapid evolution of homoplasmic mutations*. Proc Natl Acad Sci U S A, 2003. **100**(4): p. 1838-43.
300. Bogenhagen, D. and D.A. Clayton, *Mouse L cell mitochondrial DNA molecules are selected randomly for replication throughout the cell cycle*. Cell, 1977. **11**(4): p. 719-27.
301. Shadel, G.S. and D.A. Clayton, *Mitochondrial DNA maintenance in vertebrates*. Annu Rev Biochem, 1997. **66**: p. 409-35.
302. Clayton, D.A., *Transcription and replication of animal mitochondrial DNAs*. Int Rev Cytol, 1992. **141**: p. 217-32.
303. Lee, D.Y. and D.A. Clayton, *Properties of a primer RNA-DNA hybrid at the mouse mitochondrial DNA leading-strand origin of replication*. J Biol Chem, 1996. **271**(39): p. 24262-9.
304. Clayton, D.A., *Replication of animal mitochondrial DNA*. Cell, 1982. **28**(4): p. 693-705.
305. Wong, T.W. and D.A. Clayton, *In vitro replication of human mitochondrial DNA: accurate initiation at the origin of light-strand synthesis*. Cell, 1985. **42**(3): p. 951-8.
306. Hixson, J.E., T.W. Wong, and D.A. Clayton, *Both the conserved stem-loop and divergent 5'-flanking sequences are required for initiation at the human mitochondrial origin of light-strand DNA replication*. J Biol Chem, 1986. **261**(5): p. 2384-90.
307. Graves, S.W., A.A. Johnson, and K.A. Johnson, *Expression, purification, and initial kinetic characterization of the large subunit of the human mitochondrial DNA polymerase*. Biochemistry, 1998. **37**(17): p. 6050-8.
308. Wanrooij, S. and M. Falkenberg, *The human mitochondrial replication fork in health and disease*. Biochim Biophys Acta, 2010. **1797**(8): p. 1378-88.


309. Lecrenier, N. and F. Foury, *New features of mitochondrial DNA replication system in yeast and man*. *Gene*, 2000. **246**(1-2): p. 37-48.
310. McKinney, E.A. and M.T. Oliveira, *Replicating animal mitochondrial DNA*. *Genet Mol Biol*, 2013. **36**(3): p. 308-15.
311. Holt, I.J., H.E. Lorimer, and H.T. Jacobs, *Coupled leading- and lagging-strand synthesis of mammalian mitochondrial DNA*. *Cell*, 2000. **100**(5): p. 515-24.
312. Yang, M.Y., et al., *Biased incorporation of ribonucleotides on the mitochondrial L-strand accounts for apparent strand-asymmetric DNA replication*. *Cell*, 2002. **111**(4): p. 495-505.
313. Yasukawa, T., et al., *Replication of vertebrate mitochondrial DNA entails transient ribonucleotide incorporation throughout the lagging strand*. *EMBO J*, 2006. **25**(22): p. 5358-71.
314. Meadows, J.R., S. Hiendleder, and J.W. Kijas, *Haplogroup relationships between domestic and wild sheep resolved using a mitogenome panel*. *Heredity (Edinb)*, 2010. **106**(4): p. 700-6.
315. MW, B. and S. Townsend, *Mitochondrial DNA diversity in modern sheep: implications for domestication*, in *Documenting domestication. New genetics and archaeological paradigms*, Z. MA, et al., Editors. 2006, University of California Press, Ltd: Berkeley, CA. p. 306-316.
316. Pedrosa, S., et al., *Mitochondrial diversity and the origin of Iberian sheep*. *Genet Sel Evol*, 2007. **39**(1): p. 91-103.
317. Wood, N.J. and S.H. Phua, *Variation in the control region sequence of the sheep mitochondrial genome*. *Anim Genet*, 1996. **27**(1): p. 25-33.
318. Guo, J., et al., *A novel maternal lineage revealed in sheep (*Ovis aries*)*. *Anim Genet*, 2005. **36**(4): p. 331-6.
319. Chen, S.Y., et al., *Origin, genetic diversity, and population structure of Chinese domestic sheep*. *Gene*, 2006. **376**(2): p. 216-23.
320. Demirci, S., et al., *Mitochondrial DNA diversity of modern, ancient and wild sheep (*Ovis gmelinii anatolica*) from Turkey: new insights on the evolutionary history of sheep*. *PLoS One*, 2013. **8**(12): p. e81952.
321. Liu, J., et al., *Genetic Diversity and Phylogenetic Evolution of Tibetan Sheep Based on mtDNA D-Loop Sequences*. *PLoS One*, 2016. **11**(7): p. e0159308.
322. Luikart, G., et al., *Multiple maternal origins and weak phylogeographic structure in domestic goats*. *Proc Natl Acad Sci U S A*, 2001. **98**(10): p. 5927-32.
323. Olivieri, C., et al., *Phylogenetic position of a copper age sheep (*Ovis aries*) mitochondrial DNA*. *PLoS One*, 2012. **7**(3): p. e33792.
324. Meadows, J.R., et al., *Mitochondrial sequence reveals high levels of gene flow between breeds of domestic sheep from Asia and Europe*. *J Hered*, 2005. **96**(5): p. 494-501.
325. Niemi, M., et al., *Mitochondrial DNA and Y-chromosomal diversity in ancient populations of domestic sheep (*Ovis aries*) in Finland: comparison with contemporary sheep breeds*. *Genet Sel Evol*, 2013. **45**: p. 2.
326. Sanger, F., et al., *Use of DNA polymerase I primed by a synthetic oligonucleotide to determine a nucleotide sequence in phage fl DNA*. *Proc Natl Acad Sci U S A*, 1973. **70**(4): p. 1209-13.
327. Nei, M., *Molecular Evolutionary Genetics*. 1987: New York, NY: Columbia University Press.
328. Jobling, M.A., Hurles, M. E. and Tyler-Smith, C. , *Human Evolutionary Genetics: origins, peoples and disease*. 2004: London/New York: Garland Science Publishing.

329. Fu, Y.X., *Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection*. Genetics, 1997. **147**(2): p. 915-25.
330. Tajima, F., *Statistical method for testing the neutral mutation hypothesis by DNA polymorphism*. Genetics, 1989. **123**(3): p. 585-95.
331. Hunter, P.R., *Reproducibility and indices of discriminatory power of microbial typing methods*. J Clin Microbiol, 1990. **28**(9): p. 1903-5.
332. Schroeder, O., et al., *Endogenous Retroviral Insertions Indicate a Secondary Introduction of Domestic Sheep Lineages to the Caucasus and Central Asia between the Bronze and Iron Age*. Genes (Basel), 2017. **8**(6).
333. Tapio, M., et al., *Native breeds demonstrate high contributions to the molecular variation in northern European sheep*. Mol Ecol, 2005. **14**(13): p. 3951-63.
334. Nei, M. and W.H. Li, *Mathematical model for studying genetic variation in terms of restriction endonucleases*. Proc Natl Acad Sci U S A, 1979. **76**(10): p. 5269-73.
335. Rogers, A.R. and H. Harpending, *Population growth makes waves in the distribution of pairwise genetic differences*. Mol Biol Evol, 1992. **9**(3): p. 552-69.
336. Meadows, J.R., et al., *Globally dispersed Y chromosomal haplotypes in wild and domestic sheep*. Anim Genet, 2006. **37**(5): p. 444-53.
337. Uzun, M., et al., *Genetic relationships among Turkish sheep*. Genet Sel Evol, 2006. **38**(5): p. 513-24.
338. Peter, C., et al., *Genetic diversity and subdivision of 57 European and Middle-Eastern sheep breeds*. Anim Genet, 2007. **38**(1): p. 37-44.
339. <http://www.ansi.okstate.edu/breeds/sheep/awassi/index.htm>.
340. Mager, D.L. and J.P. Stoye, *Mammalian Endogenous Retroviruses*. Microbiol Spectr, 2015. **3**(1): p. MDNA3-0009-2014.
341. AYANOĞLU, İ.C., *REVEALING THE HISTORY OF ANATOLIAN SHEEP DOMESTICATION BY USING RETROVIRUS INTEGRATIONS*, in *Biology*. 2013, THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES OF MIDDLE EAST TECHNICAL UNIVERSITY. p. 133.
342. Archer, F., et al., *Alveolar type II cells isolated from pulmonary adenocarcinoma: a model for JSRV expression in vitro*. Am J Respir Cell Mol Biol, 2007. **36**(5): p. 534-40.

## Publications

1. Ibrahim E, Sirawan A, El-Bazzal B, **El Hage J**, Abi Said M, Zaraket H, Kandeil A, Ali MA, Kayali G. **2016**. Complete genome sequence of the first H5N1 avian influenza virus isolated from chickens in Lebanon in 2016. *Genome Announc* 4(5):e01062-16. doi:10.1128/genomeA.01062-16.
2. Ali El Romeh, Bianca Zecchin, Alice Fusaro, Elias Ibrahim, Bassel El Bazzal, **Jeanne El Hage**, Adelaide Milani, Gianpiero Zamperin, and Isabella Monne. **2017**. Highly Pathogenic Avian Influenza H5N1 Clade 2.3.2.1c Virus in Lebanon, 2016. *Avian Diseases*, 61(2):271-273. doi.org/10.1637/11544-113016-Case.1
3. Alain Abi-Rizk, Tony Kanaan and **Jeanne El Hage**. **2017**. Seroprevalence of Schmallerberg virus and other Simbu group viruses among the Lebanese sheep. *Open Veterinary Journal*, Vol. 7(3): 290-293. DOI:10.4314/ovj.v7i3.15

# Complete Genome Sequence of the First H5N1 Avian Influenza Virus Isolated from Chickens in Lebanon in 2016

Elias Ibrahim,<sup>a</sup> Abeer Sirawan,<sup>a</sup> Bassel El-Bazzal,<sup>a</sup> Jeanne El Hage,<sup>b</sup> Mounir Abi Said,<sup>c</sup>  Hassan Zaraket,<sup>d</sup> Ahmed Kandeil,<sup>e</sup> Mohamed A. Ali,<sup>e</sup>  Ghazi Kayali<sup>f,g</sup>

Ministry of Agriculture, Beirut, Lebanon<sup>a</sup>; Lebanese Agricultural Research Institute, Animal Health Laboratory, Fanar, Lebanon<sup>b</sup>; Lebanese University, El Fanar, Lebanon<sup>c</sup>; American University of Beirut, Beirut, Lebanon<sup>d</sup>; Center of Scientific Excellence for Influenza Viruses, National Research Centre, Giza, Egypt<sup>e</sup>; Department of Epidemiology, Human Genetics, and Environmental Sciences, University of Texas Health Sciences Center, Houston, Texas, USA<sup>f</sup>; Human Link, Hazmieh, Lebanon<sup>g</sup>

**We generated the full genome of a highly pathogenic H5N1 avian influenza virus that caused an outbreak on a chicken farm in Lebanon in April 2016. Analysis revealed that the virus belonged to clade 2.3.2.1c that recently caused outbreaks in West Africa and the United Arab Emirates.**

Received 9 August 2016 Accepted 26 August 2016 Published 20 October 2016

**Citation** Ibrahim E, Sirawan A, El-Bazzal B, El Hage J, Abi Said M, Zaraket H, Kandeil A, Ali MA, Kayali G. 2016. Complete genome sequence of the first H5N1 avian influenza virus isolated from chickens in Lebanon in 2016. *Genome Announc* 4(5):e01062-16. doi:10.1128/genomeA.01062-16.

**Copyright** © 2016 Ibrahim et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Ghazi Kayali, ghazi@human-link.org.

Highly pathogenic avian influenza (HPAI) H5N1 viruses were first detected in 1997 (1). Then, they spread widely and evolved rapidly into 10 distinct clades (0 to 9). Clades 2.1.3.2, 2.2.1.2, 2.3.2.1, 2.3.4.4, and 7.2 were detected in 2014 and continue to circulate globally (2). Clade 2.2 viruses were first detected in the Middle East in 2005 causing sporadic outbreaks in several countries and becoming enzootic in Egypt (3). No outbreaks were detected in Lebanon until April 2016, when a chicken farm in the Beqaa region had high mortality with symptoms resembling HPAI infection. The Lebanese Ministry of Agriculture detected HPAI H5 infection and successfully controlled the outbreak through culling and monitoring of neighboring farms.

An A/chicken/Lebanon/157/2016(H5N1) virus was isolated from infected chickens. Viral RNA was extracted using a QIAamp viral RNA minikit (Qiagen, Germany) according to the manufacturer's protocol. The first-strand cDNA was synthesized using Superscript III reverse transcriptase (Invitrogen, CA) and Uni-12 primer (5' AGCRAAAGCAGG3') per manufacturer's protocol. Genes were amplified using the Phusion master mix kit (Thermo, MA) and universal primers (4). Amplified products were sequenced at Macrogen (Macrogen, South Korea) and were assembled using SeqMan (DNASTAR, WI).

The eight genome segments encoded PB2, PB1, PB1-F2, PA, PA-X, HA, NP, NA, M1, M2, NS1, and NS2 viral proteins with lengths 759, 757, 57, 716, 252, 567, 498, 449, 252, 97, 225, and 121 amino acids (aa), respectively. The HA cleavage motif sequence of H5N1 isolates was RERRRKR\*GLF, which is the signature of HPAI viruses. Analysis of the NXT/S motif (X can be any amino acid except proline) revealed that the isolate had seven potential glycosylation sites, at positions 10 (NNST), 23 (NVT), 140 (NSS), 165 (NNT), 286 (NSS), 483 (NGS), and 542 (NGS) (H5 numbering) within the HA molecule. The virus had Q222 and G224 at the receptor binding site, suggesting that it favors avian-like receptors (5). However, it also contained aa N and A at sites 94 and 133 instead of D and S, respectively, which is believed to

enhance binding to  $\alpha$ 2-6 receptors (6). The analysis of the NA gene revealed a 20 aa deletion in the NA stalk region (positions 49 to 68). The virus had an NA substitution V149A that potentially reduced susceptibility to zanamivir (7). Mutations E627K and D701N in the PB2 that facilitate the adaptation of avian viruses to mammals, and increase transmission and/or pathogenicity (8, 9), were not recorded. The BLASTn and phylogenetic analysis of all eight segments revealed that the virus was closely related to the recent clade 2.3.2.1c viruses that circulated in West Africa (Ivory Coast, Niger, and Ghana) and India with high sequence homologies (98 to 99%). This clade has been previously detected in the United Arab Emirates in wild bird species (10). The emergence of this H5N1 clade in the Middle East is of both veterinary and human public health concern and requires surveillance at the human-animal interface.

**Accession number(s).** The complete genome sequence of the A/chicken/Lebanon/157/2016(H5N1) virus was deposited in GenBank under the accession numbers [KX644138](https://ncbi.nlm.nih.gov/nucl/KX644138) to [KX644145](https://ncbi.nlm.nih.gov/nucl/KX644145).

## ACKNOWLEDGMENTS

We express gratitude to the field personnel who collected the samples and controlled the outbreak. We also express gratitude to the laboratory personnel who conducted the laboratory analysis in Lebanon and Egypt. This work was funded by an internal grant from Human Link, Lebanon.

## FUNDING INFORMATION

This work, including the efforts of Ghazi Kayali, was funded by Human Link (IG001).

## REFERENCES

- Shortridge KF, Zhou NN, Guan Y, Gao P, Ito T, Kawaoka Y, Kodihalli S, Krauss S, Markwell D, Murti KG, Norwood M, Senne D, Sims L, Takada A, Webster RG. 1998. Characterization of avian H5N1 influenza viruses from poultry in Hong Kong. *Virology* 252:331–342. [http://dx.doi.org/10.1006/viro.1998.9488](https://doi.org/10.1006/viro.1998.9488).
- Smith GJ, Donis RO, World Health Organization/World Organisation for Animal Health, Agriculture Organization HEWG. 2015. Nomenclature

- updates resulting from the evolution of avian influenza A(H5) virus clades 2.1.3.2a, 2.2.1, and 2.3.4 during 2013–2014. *Influenza Other Respir Viruses* 9:271–276. <http://dx.doi.org/10.1111/irv.12324>.
3. World Health Organization. 14 July 2014. H5N1 highly pathogenic avian influenza: Timeline of major events. [http://www.who.int/influenza/human\\_animal\\_interface/h5n1\\_avian\\_influenza\\_update20140714.pdf?ua=1&ua=1](http://www.who.int/influenza/human_animal_interface/h5n1_avian_influenza_update20140714.pdf?ua=1&ua=1).
  4. Hoffmann E, Stech J, Guan Y, Webster RG, Perez DR. 2001. Universal primer set for the full-length amplification of all influenza A viruses. *Arch Virol* 146:2275–2289. <http://dx.doi.org/10.1007/s007050170002>.
  5. Yamada S, Suzuki Y, Suzuki T, Le MQ, Nidom CA, Sakai-Tagawa Y, Muramoto Y, Ito M, Kiso M, Horimoto T, Shinya K, Sawada T, Kiso M, Usui T, Murata T, Lin Y, Hay A, Haire LF, Stevens DJ, Russell RJ. 2006. Haemagglutinin mutations responsible for the binding of H5N1 influenza A viruses to human-type receptors. *Nature* 444:378–382. <http://dx.doi.org/10.1038/nature05264>.
  6. Tang DJ, Lam YM, Siu YL, Lam CH, Chu SL, Peiris JS, Buchy P, Nal B, Bruzzone R. 2012. A single residue substitution in the receptor-binding domain of H5N1 hemagglutinin is critical for packaging into pseudotyped lentiviral particles. *PLoS One* 7:e43596. <http://dx.doi.org/10.1371/journal.pone.0043596>.
  7. Naughtin M, Dyason JC, Mardy S, Sorn S, von Itzstein M, Buchy P. 2011. Neuraminidase inhibitor sensitivity and receptor-binding specificity of Cambodian clade 1 highly pathogenic H5N1 influenza virus. *Antimicrob Agents Chemother* 55:2004–2010. <http://dx.doi.org/10.1128/AAC.01773-10>.
  8. Gao Y, Zhang Y, Shinya K, Deng G, Jiang Y, Li Z, Guan Y, Tian G, Li Y, Shi J, Liu L, Zeng X, Bu Z, Xia X, Kawaoka Y, Chen H. 2009. Identification of amino acids in HA and PB2 critical for the transmission of H5N1 avian influenza viruses in a mammalian host. *PLoS Pathog* 5:e1000709. <http://dx.doi.org/10.1371/journal.ppat.1000709>.
  9. Shinya K, Hamm S, Hatta M, Ito H, Ito T, Kawaoka Y. 2004. PB2 amino acid at position 627 affects replicative efficiency, but not cell tropism, of Hong Kong H5N1 influenza A viruses in mice. *Virology* 320:258–266. <http://dx.doi.org/10.1016/j.virol.2003.11.030>.
  10. Naguib MM, Kinne J, Chen H, Chan KH, Joseph S, Wong PC, Woo PC, Wernery R, Beer M, Wernery U, Harder TC. 2015. Outbreaks of highly pathogenic avian influenza H5N1 clade 2.3.2.1c in hunting falcons and kept wild birds in Dubai implicate intercontinental virus spread. *J Gen Virol* 96:3212–3222. <http://dx.doi.org/10.1099/jgv.0.000274>.

## **Highly Pathogenic Avian Influenza H5N1 Clade 2.3.2.1c Virus in Lebanon, 2016**

Author(s): Ali El Romeh, Bianca Zecchin, Alice Fusaro, Elias Ibrahim, Bassel El Bazzal, Jeanne El Hage, Adelaide Milani, Gianpiero Zamperin, and Isabella Monne

Source: Avian Diseases, 61(2):271-273.

Published By: American Association of Avian Pathologists

<https://doi.org/10.1637/11544-113016-Case.1>

URL: <http://www.bioone.org/doi/full/10.1637/11544-113016-Case.1>

---

BioOne ([www.bioone.org](http://www.bioone.org)) is a nonprofit, online aggregation of core research in the biological, ecological, and environmental sciences. BioOne provides a sustainable online platform for over 170 journals and books published by nonprofit societies, associations, museums, institutions, and presses.

Your use of this PDF, the BioOne Web site, and all posted and associated content indicates your acceptance of BioOne's Terms of Use, available at [www.bioone.org/page/terms\\_of\\_use](http://www.bioone.org/page/terms_of_use).

Usage of BioOne content is strictly limited to personal, educational, and non-commercial use. Commercial inquiries or rights and permissions requests should be directed to the individual publisher as copyright holder.



## Case Report—

**Highly Pathogenic Avian Influenza H5N1 Clade 2.3.2.1c Virus in Lebanon, 2016**

Ali El Romeh,<sup>A</sup> Bianca Zecchin,<sup>BD</sup> Alice Fusaro,<sup>B</sup> Elias Ibrahim,<sup>A</sup> Bassel El Bazzal,<sup>A</sup> Jeanne El Hage,<sup>C</sup>  
 Adelaide Milani,<sup>B</sup> Gianpiero Zamperin,<sup>B</sup> and Isabella Monne<sup>B</sup>

<sup>A</sup>Ministry of Agriculture, Animal Resources Directorate, Bir Hassan, Embassies Street, Beirut, Lebanon

<sup>B</sup>Istituto Zooprofilattico Sperimentale delle Venezie, Viale dell'Università, 10, 35020, Legnaro, Padova, Italy

<sup>C</sup>Lebanese Agricultural Research Institute, Fanar, Main Road, 90-1965, Jdeidet El-Metin Fanar, Lebanon

Received 5 December 2016; Accepted 27 February 2017; Published ahead of print 9 March 2017; Published June 2017

**SUMMARY.** We report the phylogenetic analysis of the first outbreak of H5N1 highly pathogenic avian influenza virus detected in Lebanon from poultry in April 2016. Our whole-genome sequencing analysis revealed that the Lebanese H5N1 virus belongs to genetic clade 2.3.2.1c and clusters with viruses from Europe and West Africa.

**RESUMEN.** *Reporte de caso-* Virus de la influenza aviar altamente patógena H5N1 Clado 2.3.2.1c en el Líbano, en el año 2016. Se presenta el análisis filogenético del primer brote del virus de influenza aviar altamente patógena H5N1 detectado en avicultura en el Líbano en abril del 2016. El análisis de secuenciación de todo el genoma reveló que el virus H5N1 libanés pertenece al clado genético 2.3.2.1c y se agrupó con virus de Europa y África Occidental.

**Key words:** influenza A virus; H5N1 subtype; Lebanon; phylogeny.

**Abbreviations:** HPAI = highly pathogenic avian influenza; RT-PCR = reverse transcriptase-PCR

On April 22, 2016 a highly pathogenic avian influenza (HPAI) H5N1 virus was detected for the first time in Lebanon. The virus had infected two chicken layer farms that were very close to each other and located less than 10 km from the Syrian border, in one of Lebanon's most important farming regions, the Bekaa valley (11). The first outbreak was detected in Nabichit in a poultry farm of 80,000 animals. A sudden increase in deaths among layer hens was reported and systemic pathomorphological alterations were identified in the affected animals. An overall mortality rate of 25% and a case fatality rate of 100% were observed before the application of the depopulation measures. The quarantine and a monitored containment zone were established around the primary outbreak and a surveillance plan was set up. The measures applied to control the outbreak included immediate stamping out, disinfection of the infected herds and disposal of the animal products, carcasses, and waste. Twenty days after the first confirmation, a second outbreak was notified within the protection zone in Sariin Tehta. In this case the mortality rate for the 126,000 chickens was 15.87% and the case fatality rate was 100% (11). Overall, out of a total of 206,000 susceptible animals, 40,000 died and 166,000 were destroyed. The infection was efficiently controlled thanks to the forthwith activation of a contingency plan for HPAI H5N1 implemented by the Ministry of Public Health and the Ministry of Agriculture in collaboration with the army and local security forces.

In 2015 several cases of HPAI of the H5N1 subtype had been detected in poultry in the surrounding area, including Israel, the Palestinian territories and Turkey (10). These viruses belonged to two distinct clades: 2.2.1.2, found in Israel and the Palestinian territories, and 2.3.2.1c, identified in Turkey (8). At the end of 2015 Iraq had also been affected by an extensive HPAI H5N1 epizootic,

and as of July 2016, the country was still reporting cases of H5N1 infections (1). On the other hand, Syria, which lies in the north and eastern border of Lebanon, has never reported HPAI H5N1 outbreaks.

The genetic information on these outbreaks is fragmented or limited to the only HA gene. In this study we provide for the first time the characterization of the complete genome of the HPAI H5N1 virus identified in Lebanon in April 2016 to determine the possible origin of the virus and to monitor the probable emergence of mutations, which may impact public health.

## MATERIALS AND METHODS

Seven tracheal swabs were collected from the first chicken layer farm that had been affected by H5N1 in Lebanon, from dead birds or animals showing depression and severe respiratory signs. The lungs were congested and pathomorphological examination revealed petechiae in peri- and subepicardial locations. The Animal Resources Directorate of the Ministry of Agriculture in Beirut (Lebanon) submitted the samples to the World Organisation for Animal Health Reference Laboratory and the Food and Agriculture Organization of the United Nations (OIE/FAO) Reference Center for Animal Influenza at the Istituto Zooprofilattico Sperimentale delle Venezie (Italy), for confirmatory diagnosis and genetic analyses.

To genetically characterize swabs, testing positive for H5N1 by molecular investigation, viral RNA was purified with the use of the Nucleospin RNA II kit (Machery-Nagel, Duren, Germany). Reverse transcriptase-PCR (RT-PCR) was performed with the use of the One Step RT-PCR kit (Qiagen, Hilden, Germany) and a partial sequence (250 nucleotides) of the HA gene was obtained with a 16-capillary ABI Prism 3130xl Genetic Analyzer (Applied Biosystems, Foster City, CA), to analyze the cleavage site motif. Sequence data were assembled and edited with SeqScape software v2.5 (Applied Biosystems).

The complete genome of one sample (A/chicken/Lebanon/16vir-2783-20/2016) was processed by using the Illumina MiSeq (San Diego,

<sup>D</sup>Corresponding author. E-mail: bzecchin@izsvenezie.it

Ali El Romeh and Bianca Zecchin contributed equally to this article.

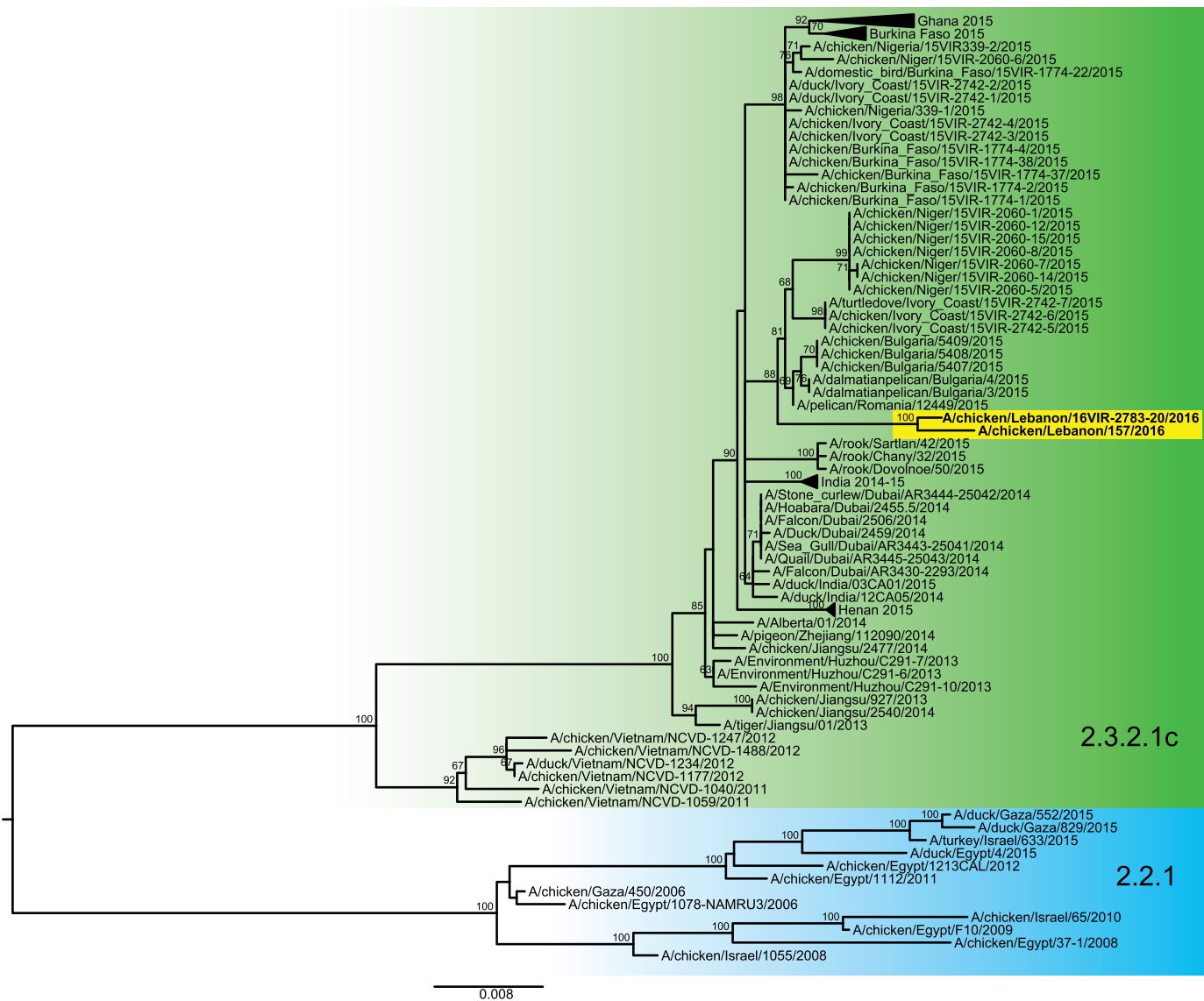


Fig. 1. Maximum-likelihood phylogenetic tree of the hemagglutinin gene segment of highly pathogenic avian influenza (H5N1) viruses belonging to clade 2.3.2.1c. The samples from Lebanon are highlighted in yellow. Clades 2.3.2.1c and 2.2.1 are colored in green and light blue, respectively. Bootstrap values higher than 60% are indicated above the nodes.

CA, USA) technology, as described by Fusaro *et al.* 2016 (2). It was then phylogenetically analyzed with the use of a general time reversible model of nucleotide substitution with a gamma distribution of among-site rate variation (with four rate categories) and a subtree pruning and regrafting branch-swapping search procedure with PhyML 3.1 (3). A nonparametric bootstrap analysis was performed with 100 replicates.

Consensus sequences of the eight gene segments of *A/chicken/Lebanon/16vir-2783-20/2016* were submitted to GISAID under the accession numbers EPI774540 (HA gene), EPI888808 (PB2 gene), EPI888812 (PB1 gene), EPI888815 (PA gene), EPI888816 (NP gene), EPI888817 (NA gene), EPI888818 (MP gene), and EPI888820 (NS gene).

## RESULTS AND DISCUSSION

All swabs collected from the primary outbreak in Lebanon tested positive to the H5N1 subtype and possessed a multibasic cleavage site typical of HPAI viruses (PQRERRRKR\*G). The 250 nucleotide sequences of these seven viruses showed a similarity of 100%;

therefore, a single sample was selected for the complete genome sequencing. The maximum likelihood (ML) tree of the HA gene (Fig. 1) showed that the virus *A/chicken/Lebanon/16vir-2783-20/2016* belonged to clade 2.3.2.1c (6) and, as expected, showed the highest identity (99.5%) with the HA gene of the virus identified from the second poultry farm affected by the H5N1 incursion in Lebanon (*A/chicken/Lebanon/157/2016*, sequence available in GenBank under Accession No. KX644141). Both the Lebanese viruses group with the H5N1 viruses collected in 2015 from wild birds and poultry in Europe (Bulgaria and Romania) and West Africa (Niger and Ivory Coast), showing a similarity of 98.2%–98.7%.

The phylogenetic trees obtained for the other gene segments (Supplemental Figs. S1\_MP gene, S2\_NA gene, S3\_NP gene, S4\_NS gene, S5\_PA gene, S6\_PB1 gene, S7\_PB2 gene) confirm the topology of the HA gene, displaying a genetic clustering between the Lebanese virus and the West African isolates. Unfortunately, only the HA gene of the European viruses was available in the public

databases at the time of writing, which has prevented us from making any assessment on their genetic relationship with the internal genes of the Lebanese samples.

Analysis of the amino acid sequences revealed a mutation in the HA protein (Thr188Ile, H5 numbering), which has been proved to increase the H5 specificity for  $\alpha$ 2,6-linked sialic acid receptors (12). This mutation occurs in both the Lebanese isolates (A/chicken/Lebanon/16vir-2783-20/2016 and A/chicken/Lebanon/157/2016) and can be observed only in 2.9% of the viruses belonging to clade 2.3.2.1c. In addition, the NA protein of the two Lebanese viruses possesses a mutation at position 129 (Val129Ala, N1 numbering), which reduces the susceptibility to the neuraminidase inhibitor zanamivir (4), never observed before in clade 2.3.2.1c.

Before 2016, incursion of the HPAI H5N1 goose/Guangdong lineage virus had never been reported in Lebanon. Our study is an attempt to describe the origin and the molecular properties of the HPAI H5N1 outbreak in Lebanon. We showed that the virus belongs to clade 2.3.2.1c and clusters with viruses from Europe and West Africa. In 2014 a virus with this gene constellation had begun to spread from Asia to Europe, the Middle East, and West Africa (7). The identification of a virus clade with an enhanced ability to cause disease in the ferret model (5), as well as the detection of mutations in the Lebanese sample possibly related to host specificity shift highlight the urgent need to find a solution to monitor the evolution and spread of the viruses circulating in the areas bordering Lebanon closely. Unfortunately, the social, political and economic instability of this geographical area, as well as the lack of sequences of the complete genome of H5N1 viruses from nearby countries that have recently been experiencing HPAI H5N1 outbreaks, makes it very difficult to identify the origin of this strain and to implement an adequate surveillance plan. According to the OIE report, the source of the outbreak could be ascribed to the illegal movement of animals (11). This claim may suggest that the circulation of the virus in bordering countries such as Syria is very likely, meaning that local authorities in the whole area are expected to maintain elevated levels of alert and control over the neighboring regions. This is even more necessary after the emergence of the new HPAI H5N8 strain, which has been spreading in Europe, Asia, the Middle East, and Africa since June 2016 (9).

Supplemental data associated with this article can be found at <http://dx.doi.org/10.1637/AVIANDISEASESJOURNAL-11544-113016-CaseR.s1>

## REFERENCES

1. Almada Press. Avian influenza: Iraq (WA) poultry, HPAI H5N1, spread. ProMed (2016 Jul 14). Available from: <http://www.promedmail.org>, archive no. 20160714.4345054.
2. Fusaro, A., L. Tassoni, A. Milani, J. Hughes, A. Salviato, P. R. Murcia P. Massi, G. Zamperin, L. Bonfanti, S. Marangon, G. Cattoli, and I. Monne. Unexpected interfarm transmission dynamics during a highly pathogenic avian influenza epidemic. *J. Virol.* 90(14):6401–6411. 2016.

3. Guindon S, J.-F. Dufayard, V. Lefort, M. Anisimova, W. Hordijk, and O. Gascuel. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst. Biol.* 2010;59:307–321.

4. Naughtin, M., J. C. Dyason, S. Mardy S. Sorn, M. von Itzstein, and P. Buchy. Neuraminidase inhibitor sensitivity and receptor-binding specificity of Cambodian clade 1 highly pathogenic H5N1 influenza virus. *Antimicrob. Agents Chemother.* 55(5):2004–2010. 2011.

5. Pearce, M. B., C. Pappas, K. M. Gustin, C. Todd Davis, M. J. Pantin-Jackwood, D. E. Swayne, T. R. Maines, J. A. Belser, and T. M. Tumpey. Enhanced virulence of clade 2.3.2.1 highly pathogenic avian influenza A H5N1 viruses in ferrets. *Virology* 502:114–122. 2017.

6. Smith, G. J., and R. O. Donis. World Health Organization/World Organisation for Animal Health/Food and Agriculture Organization (WHO/OIE/FAO) H5 Evolution Working Group. Nomenclature updates resulting from the evolution of avian influenza A(H5) virus clades 2.1.3.2a, 2.2.1, and 2.3.4 during 2013–2014. *Influenza Other Respir. Viruses* 9(5):271–276. 2015.

7. Tassoni, L., A. Fusaro, A. Milani, P. Lemey, J. A. Awuni, V. B. Sedor, O. Dogbey, A. N. Commey, C. Meseke, T. Joannis, G. L. Minoungou, L. Ouattara, A. M. Haido, D. Cisse-Aman, E. Couacy-Hymann, G. Dauphin, G. Cattoli, and I. Monne. Genetically different highly pathogenic avian influenza A(H5N1) viruses in West Africa, 2015. *Emerging Infect. Dis.* 22(12):2132–2136. 2016.

8. [WHO] World Health Organization. Antigenic and genetic characteristics of zoonotic influenza viruses and development of candidate vaccine viruses for pandemic preparedness, WHO. 2015 Sep [cited 2016 Jul 28]. Available from: [http://www.who.int/influenza/vaccines/virus/201509\\_zoonotic\\_vaccinevirusupdate.pdf?ua=1](http://www.who.int/influenza/vaccines/virus/201509_zoonotic_vaccinevirusupdate.pdf?ua=1)

9. WHO. Assessment of risk associated with influenza A(H5N8) virus, WHO [cited 2016 Nov 17]. Available from: [http://www.who.int/influenza/human\\_animal\\_interface/avian\\_influenza/riskassessment\\_AH5N8\\_201611/en/](http://www.who.int/influenza/human_animal_interface/avian_influenza/riskassessment_AH5N8_201611/en/)

10. World Organization for Animal Health. Update on highly pathogenic avian influenza in animals, 2015—Turkey, Palestinian Auton. Territories, Israel [cited 2016 Jul 28]. Available from: <http://www.oie.int/animal-health-in-the-world/update-on-avian-influenza/2015/>

11. World Organization for Animal Health. Update on highly pathogenic avian influenza in animals, 2016—Lebanon. Follow-up report no.1 (final report). 2016 Jun 14 [cited 2016 Jul 28]. Available from: [http://www.oie.int/wahis\\_2/public%5C..%5Ctemp%5Creports/en\\_fup\\_000020278\\_20160614\\_131721.pdf](http://www.oie.int/wahis_2/public%5C..%5Ctemp%5Creports/en_fup_000020278_20160614_131721.pdf)

12. Yang, Z.Y., C. J. Wei, W. P. Kong, L. Wu, L. Xu, D. F. Smith, and G. J. Nabel. Immunization by avian H5 influenza hemagglutinin mutants with altered receptor binding specificity. *Science* 317(5839):825–828. 2017.

## ACKNOWLEDGMENTS

We acknowledge the originating and submitting laboratories of the sequences from GISAID's EpiFlu Database on which this research is based in part (online Technical Appendix).

This study was made possible through the technical support provided by the United Nations Food and Agriculture Organization (Letter of Agreement with FAO 321149). We thank Annalisa Salviato, Alessia Schivo, and Francesca Ellero for their excellent technical assistance.

Submitted: 02/06/2017

Accepted: 15/09/2017

Published: 29/09/2017

## Seroprevalence of Schmallenberg virus and other Simbu group viruses among the Lebanese sheep

Alain Abi-Rizk<sup>1,\*</sup>, Tony Kanaan<sup>1</sup> and Jeanne El Hage<sup>1,2,3</sup>

<sup>1</sup>Faculty of Agricultural and Food Sciences, Holy Spirit University of Kaslik (USEK), P.O.Box 446, Jounieh, Lebanon

<sup>2</sup>Lebanese Agricultural Research Institute, Fanar, Bekaa P.O. Box 287, Lebanon

<sup>3</sup>Ecole Pratique des Hautes Etudes, 75014 Paris, France

### Abstract

In order to evaluate for the first time, the serological prevalence of Schmallenberg virus (SBV) and other Simbu group viruses in Lebanon, sheep originating from 15 Lebanese regions were sampled in September 2016. A total number of 750 serum samples from Awassi sheep were tested by ELISA for viral nucleoprotein antibodies. From the sampled animals, 122 animals were seropositive to SBV/Simbu group viruses. The seropositive sheep were mainly located in South Lebanon. At herd-level, a seroprevalence of 53.33% was recorded in the Seven Lebanese governorates. The animal-level seroprevalence was 16.26% and both animal and herd-level seroprevalences were negative in Mount-Lebanon. Despite that there was some serological evidence showed the presence of some Simbu group viruses in the Middle East, no study was done in Lebanon. In this study, we report for the first time the prevalence of SBV and other Simbu group viruses in Lebanon.

**Keywords:** Lebanon, SBV, Schmallenberg virus, Sheep, Simbu group viruses.

### Introduction

Schmallenberg virus (SBV) was identified in November 2011 by Friedrich Loeffler Institute (FLI, Island of Riems, Germany) following a metagenomic analysis of a pool of blood samples coming from a farm of the town of Schmallenberg (the Rhineland of North-Westphalia, Germany) (Gibbens, 2012; Hoffmann *et al.*, 2012). The percentages of nucleotide homology presented in the new genetic sequences made it possible to classify this new virus in the family of *Bunyaviridae*, genus *Orthobunyavirus*, serogroup Simbu (Tarlinton *et al.*, 2012).

The viruses group normally noncontagious, are transmitted by hematophagous arthropods, in particular of the mosquitos and culicoides (Hoffman *et al.*, 2012; Tarlinton *et al.*, 2012).

Between the month of November 2011 and mid-March 2012, the virus was highlighted in sheep, goats and cattle in Germany, the Netherlands, Belgium, the United Kingdom, France, Italy, Luxembourg and Spain (Lievaart-Peterson *et al.*, 2012, 2015). Thus constituting the first occurrence of indigenous circulation of *Orthobunyavirus* of the serogroup Simbu in Western Europe.

However, other *Orthobunyavirus* were identified in Europe (case of the Batai virus in Germany), Africa, Asia, the Middle East, and Australia (cases of Akabane virus), either sporadically by the analysis of pools of mosquitos (Jöst *et al.*, 2011; Horne and

Vanlandingham, 2014), or due to an endemic presence (case of the Tahyna virus) (Bennett *et al.*, 2011).

In fact, back in 1980 an outbreak of Akabane virus was confirmed in cattle in the Turkish Province of Aydin. Thereafter, many Middle East countries were screened for the presence of neutralizing antibodies to Akabane virus. The results showed that the virus was present in the south Turkish coast, Cyprus, the Orontes river valley in Syria and the lower Jordan River valley.

Interestingly, the fact that this virus failed to persist in southern Turkey for more than two years indicates that the Middle East may be open to epidemic rather than endemic infection.

The presence of neutralizing antibodies in the eastern Turkish Provinces of Gaziantep and Diyarbakir suggests that this might be the route whereby Akabane virus and probably other Simbu group viruses could invade the Middle East region (Taylor and Mellor, 1994).

Regarding the infection by SBV, its associated disease appears in the adult cattle by a fall of the dairy production, fever, diarrhea being able to be severe and sometimes abortions. A congenital attack is also described in lambs, calves and kids, is characterized by malformations of the arthrogryposes / hydranencéphalitis type (Lievaart-Peterson *et al.*, 2015).

Previous work conducted by the authors on brain samples collected from malformed ovine fetuses

\*Corresponding Author: Alain Abi-Rizk. Faculty of Agricultural and Food Sciences, Holy Spirit University of Kaslik (USEK), P.O.Box: 446, Jounieh, Lebanon. Tel.: +961 9 600 876. Email: [alainabirizk@usek.edu.lb](mailto:alainabirizk@usek.edu.lb)

(submitted for post mortem examination in December 2015), detected SBV by RT-q PCR in 2 out of the 7 tested animals (protocol generously provided by Dr. Bernd Hoffmann from the Friedrich-Loeffler-Institut). This study therefore aimed to evaluate for the first time, the seroprevalence of SBV and other Simbu group viruses in Lebanon.

### Materials and Methods

#### Country location and background

Lebanon, located at the eastern shores of the Mediterranean, with an area of 10452 Km<sup>2</sup> most of it mountainous. The length is almost three times its width and it have the Mount-Lebanon and the anti-Lebanon that are parallel to each other, they are separated by the Bekaa plain. Narrows from north to south. Lebanon has 454 Km of land boundaries and a coast line of 225 Km (Ministry of Agriculture, 2009).

#### Blood sampling

In September 2016, a total of 750 sheep above 12 month of age belonging to 15 large herds (> 500 sheep per herd) were sampled. Farms were chosen randomly throughout the seven Lebanese governorates: Lebanon South (4 farms), Mount Lebanon (3 farms) Lebanon North (1 farm) Nabatiyeh (2 farms), Bekaa (2 farms), Baalbeck-Hermel (2 farms), Akkar (1 farm) (Table 1).

**Table 1.** Number of sampled sheep per governorate and region.

Governorate	Herd / Region	Number of samples
Lebanon South	Saida	50
	Tyr	50
	Koura	50
	Jezzine	50
Mount- Lebanon	Baabda	50
	Keserwan	50
	Chouf	50
Nabatiyeh	Bent jbeil	50
	Marjeoun	50
Akkar	Akkar	50
Bekaa	Bekaa	50
	Zahle	50
Lebanon North	Zgharta	50
Baalbeck-Hermel	Baalbak	50
	Hermel	50

A number of 50 blood sample was collected from each herd, representing at least 10% of individuals in each herd. Blood samples (25ml per sample) were collected from the external jugular vein and the serum was separated into 2ml crayo-vial and preserved at -20°C until analysis.

#### ELISA test

In accordance with the manufacturer's instructions, all collected sera were evaluated for anti-SBV antibody using a commercially available ELISA kit (IDEXX Schmallenberg Ab Test) (IDEXX Switzerland AG Laboratories). This kit detects SBV and cross-react with other Simbu serogroups.

Briefly, all collected sera and controls were diluted 1:10 and distributed in the wells of the microtiter plate (100 µl per well). After 60 minutes of incubation in darkness at 22–24°C, all the wells were washed 3 times with 300 µl of wash solution and added 100 µl per well of the conjugate. After 60 minutes of incubation in darkness at 22–24°C, the wells were washed once again 3 times with 300 µl wash solution and added with 100 µl of TMB substrate. After 10 minutes of incubation at 22–24°C, the color reaction was stopped by adding into each well 100 µl of the Stop solution.

The plates were read with a photometer at a wavelength of 450 nm, and the optical density (OD) of the samples was analyzed in relation to the negative and the positive controls with the formula:  $S/P\% = (\text{sample OD} - \text{negative control OD}) / (\text{positive control OD} - \text{negative control OD})$ . The sample is considered negative if S/P% is <30%, suspect if S/P% is between 30% and 40%, and positive if S/P% is >40%.

The true prevalence of serologically positive animals was estimated by adjusting the apparent prevalence to the sensitivity and specificity of the test as previously described (Rogan and Gladen, 1978)

### Results

#### Schmallenberg virus or other Simbu group viruses prevalence in Lebanon

The results of the anti-SBV/Simbu group viruses antibody detection tests (Table 2) showed that 122 of the 750 tested sheep had anti-SBV/Simbu group viruses' antibodies, representing an individual prevalence of 16.26%.

**Table 2:** Prevalence of SBV and other Simbu viruses in Lebanon.

Item	Total	SBV + (±STDV)	Prevalence %
Sheep	750	122 (±3.53)	16.26
Herds	15	8 (±4.24)	53.33

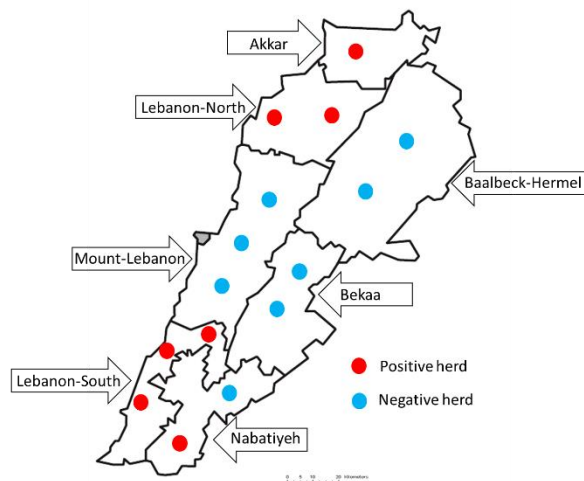
The sheep that were tested positive were from the 8 of the 15 herds tested for SBV Simbu group viruses. The infection rate at herd level was therefore 53.33 %.

According to the results shown in Table 2, SBV/Simbu group viruses infection varied substantially between the different Lebanese governorates, with a maximum recorded in Lebanon South and Nabatiyeh, where 30% of the analyzed animal were positive, and no seropositivity was recorded in Bekaa , Baalbeck-Hermel and Mount-Lebanon (Table 3).

**Table 3:** Prevalence of SBV in the different Lebanese governorates.

Governorates	Number of animals	SBV +	Prevalence %
Nabatieh	100	30 ( $\pm 1.4$ )	30
Lebanon South	150	48 ( $\pm 2.1$ )	32
Lebanon North	100	32 ( $\pm 2.1$ )	32
Mount Lebanon	150	0 ( $\pm 1$ )	0
Bekaa	100	1 ( $\pm 0.5$ )	1
Akkar	50	11 ( $\pm 0$ )	22
Baalbak Hermel	100	0 ( $\pm 0$ )	0

A diverse geographical variation of the seropositivity was observed in the different Lebanese regions. In the southern part of the country, 80% of the tested herds contained SBV-seropositive sheep where in Mount-Lebanon (were the lowest population of sheep is located) the percentage of infected herds was 0% (Figure 1).



**Fig. 1.** Prevalence of SBV/Simbu group Viruses herds in the different Lebanese governorates.

### Discussion

This is the first study that investigates the epidemiology of SBV and other Simbu group viruses' infection in Lebanon.

Herd seroprevalence in this study (53.3%) was alarming and somehow expected after the several outbreaks and sever economic losses from SBV detected from 2011 till now in many European and Mediterranean countries like Germany, France, Belgium, Italy, Spain, Greece, turkey and others.

Interestingly, our results were lower to what was previously shown in Belgium for example (above 80%

of the herds) (Méroc *et al.*, 2013) and higher than Turkey (39.8%) (Azkur *et al.*, 2013).

This study present for the first time, a clear evidence that SBV and Simbu group viruses are already spread in Lebanon and that most of the Lebanese flocks had been in contact with SBV or other Simbu group viruses somehow.

The absence of severe symptoms and any vaccination program could mean that most of the Lebanese sheep have acquired a natural protective immunity against these viruses (Sailleau *et al.*, 2013). But till today, it remains unknown how long this natural immunity could last, mainly that the Lebanese , sheep and goats are kept out on pasture day and night with only a shelter against extreme weather conditions and are not protected against parasites, mosquitos, culicoides and viruses.

For that, even in highly immunized flocks, because of all newborn animals, part of the sheep population will continuously remain susceptible.

This study should be taken into consideration as part of comprehensive SBV surveillance strategy in the country, mainly because no study is done for the detection of SBV or other Simbu group viruses and thus no prevention action are made to control the spreading of these virus.

In Lebanon, around 55% of the total population of sheep is located in the Eastern part at the Bekaa Valley. In this region, the majority of herds are large, which increase the stocking density and eventually increase the likelihood of transmission knowing that no animal in this region is yet seropositive. The absence of seropositivity in the Bekaa could be related to the absence of the vectors in this region which is characterized by rainy winters and extremely dry, and hot summers. More detailed epidemiological studies are needed to elucidate other factors, such as climate variations, presence of culicoides and rearing system that could contribute to high SBV seroprevalence and a continuous surveillance must be done in order to avoid any undesirable outbreaks.

For the moment, insufficient data are present in order to establish the origin and the time of SBV and or other Simbu viruses' infection of the sheep in Lebanon. One hypothesis is that transmission could have coincided with the importation of new animals. Yet, that there is not enough evidence to implicate the animal importation as the main source of Simbu viruses. The presence SBV already confirmed by RT-q PCR and Simbu virus-specific antibodies in most of the Lebanese herds suggests that these viruses have already spread to neighboring countries such as Syria, Palestine and Cyprus.

On the other hand, SBV coming from Greece could have been already dispersed in the area before these importations.

Even if the ELISA method used in this work does not provide a definite diagnosis for the detection of SBV, it is a recommended technique when a large number of animals needs to be tested.

Further molecular detection and phylogenetic analysis of SBV and other Simbu viruses in all the Lebanese ruminants and/or culicoides midges could reveal the source of the infections in Lebanon.

#### **Conflict of interest**

The authors declare that they have no competing interests.

#### **References**

- Azkur, A.K., Albayrak, H., Risvanli, A., Pestil, Z., Ozan, E., Yilmaz, O., Tonbak, S., Cavunt, A., Kadi, H., Macun, H.C., Acar, D., Ozenc, E., Alparslan, S. and Bulut, H. 2013. Antibodies to Schmallenberg virus in domestic livestock in Turkey. *Trop. Anim. Health. Prod.* 45(8), 1825-1828.
- Bennett, R.S., Gresko, A.K., Murphy, B.R. and Whitehead, S.S. 2011. Tahyna virus genetics, infectivity, and immunogenicity in mice and monkeys. *Virology* 438, 135. doi: 10.1016/j.virol.2011.08.018
- Gibbens, N. 2012. Schmallenberg virus: a novel viral disease in northern Europe. *Vet. Rec.* 170(2), 58.
- Hoffmann, B., Scheuch, M., Höper, D., Jungblut, R., Holsteg, M., Schirrmeier, H., Eschbaumer, M., Goller, K.V., Wernike, K., Fischer, M., Breithaupt, A., Mettenleiter, T.C. and Beer, M. 2012. Novel orthobunyavirus in cattle, Europe, 2011. *Emerg. Infect. Dis.* 18(3), 469-472.
- Horne, K.M. and Vanlandingham, D.L. 2014. Bunyavirus-Vector Interactions. *Viruses*, 6(11), 4373-4397.
- Jöst, H., Bialonski, A., Maus, D., Sambri, V., Eiden, M., Groschup, M.H., Günther, S., Becker, N. and Schmidt-Chanasit, J. 2011. Isolation of Usutu virus in Germany. *Am. J. Trop. Med. Hyg.* 85(3), 551-553.
- Lievaart-Peterson, K., Luttikholt, S. J. M., Van den Brom, R., & Vellema, P. 2012. Schmallenberg virus infection in small ruminants—First review of the situation and prospects in Northern Europe. *Small Ruminant Res.* 106(2-3), 71-76.
- Lievaart-Peterson, K., Luttikholt, S., Peperkamp, K., Van den Brom, R. and Vellema, P. 2015. Schmallenberg disease in sheep or goats: Past, present and future. *Vet. Microbiol.* 181, 147-153.
- Méroc, E., De Regge, N., Riocreux, F., Caij, A.B., van den Berg, T. and van der Stede, Y. 2013. Distribution of Schmallenberg Virus and Seroprevalence in Belgian Sheep and Goats. *Transbound. Emerg. Dis.* 61(5), 425-431.
- Ministry of agriculture. 2009. The Agriculture in Lebanon 2008 – 2009. The National project for agriculture development – FAO. pp: 64.
- Rogan, W.J., Gladen, B., 1978. Estimating prevalence from the result of a screening test. *Am. J. Epidemiol.* 107(1), 71-76.
- Sailleau, C., Boogaerts, C., Meyrueix, A., Laloy, E., Bréard, E., Viarouge, C., Desprat, A., Vitour, D., Doceul, V., Boucher, C., Zientara, S., Nicolier, A. and Grandjean D. 2013. Schmallenberg virus infection in dogs, France, 2012. *Emerg. Infect. Dis.* 19(11), 1896-1898.
- Tarlinton, R., Daly, J., Dunham, S., & Kydd, J. 2012. The challenge of Schmallenberg virus emergence in Europe. *Vet. J.* 194(1), 10-18.
- Taylor, W.P. and Mellor, P.S. 1994. The distribution of Akabane virus in the Middle East. *Epidemiol. Infect.* 113(1), 175-185.

## Résumé

Le mouton était l'un des premiers animaux à être domestiqués dans le croissant fertile. Ces événements de domestication, probablement initiés au début du Néolithique, ont génétiquement construit les races contemporaines du Moyen-Orient mais aussi du monde entier. L'élevage de moutons, principalement mouton de la race Awassi, représente une activité économique essentielle du Liban ; cependant, jusqu'à présent, il n'existe que très peu de données génétiques sur cette race. L'objectif principal de mon projet de thèse était de caractériser génétiquement la race Awassi du Liban.

La première approche utilisée est celle des rétrovirus endogènes (enJSRV) qui sont polymorphiques par insertion dans les génomes du mouton domestique. Nos résultats montrent i) une prédominance du type R2 confirmant que le mouton Awassi du Liban est une race moderne et ii) une présence du rétrotype R4, une caractéristique commune des populations de moutons du bassin méditerranéen.

Et comme approche complémentaire, deux marqueurs mitochondriaux ont été utilisés, le cytochrome b et la D-Loop, pour étudier l'origine maternelle de cette race et sa relation phylogénétique au sein de la famille *Ovis aries*. Des moutons d'Awassi analysés, les cinq HPGs maternels connus à ce jour ont été trouvés avec l'analyse de la super-séquence, l'alignement Cyt-b\_D-Loop, dont l'HPG D ; un HPG extrêmement rare et limité jusqu'à présent aux moutons à queue grasse tel que l'Awassi. Dans l'ensemble, les résultats obtenus au cours de cette étude fournissent une caractérisation génétique complète ainsi que quelques idées sur la structure phylogéographique des populations de moutons de la race Awassi au Liban.

## Mots Clés

Domestication, mouton, Awassi, Rétrovirus endogènes, ADN mitochondrial.

## Abstract

Sheep was one of the first animals to be domesticated in the Fertile Crescent. These domestication events, probably initiated in the early Neolithic, have genetically built the contemporary races of the Middle East but also of the whole world. Sheep farming, mainly sheep of Awassi breed, represents an essential economic activity of Lebanon; however, so far, only very few genetic data exist on this breed. Hence, the main objective of my PhD project was to genetically characterize the Awassi breed of Lebanon.

The first approach used is the endogenous retrovirus (enJSRV) that are insertionally polymorphic within the genomes of domestic sheep. Our results show i) a predominance of the R2 retrotype confirming that the Awassi sheep of Lebanon is a modern breed, ii) the presence of R4 retrotype a common feature of the sheep populations present within the Mediterranean area.

As a complementary approach, two mitochondrial markers were used, the cytochrome b and D-Loop, to investigate the maternal origin of this breed and its phylogenetic relationship within the *Ovis aries* family. From the Awassi sheep analyzed, the five maternal haplogroups (HPGs) known today were found with the super-sequence analysis, Cyt-b\_D-Loop alignment, of which HPG D, an extremely rare HPG, limited till now to fat-tailed sheep such as Awassi. Overall, the results obtained during this study provide a comprehensive genetic characterization as well as some insights into the phylogeographic structure of the sheep populations of the Awassi breed in Lebanon.

## Keywords

Domestication, Sheep, endogenous retroviruses, Mitochondrial DNA.