

Faculté des bioingénieurs

# Analyse génétique de la vipère péliade (*Vipera berus*) pour une réintroduction durable

Auteur : Nerea Tapia Nsang

Promoteur(s) : Pr. François Chaumont, Dr. Marie-Christine Flamand

Lecteur(s) : Pr. Pierre Bertin, Mr Mathieu Bufkens

Année académique 2021-2022

Mémoire de fin d'études présenté en vue de l'obtention du diplôme de  
Bioingénieur : Chimie et bioindustries



# Table des matières

Remerciements .....	1
1. Introduction .....	2
1.1. La vipère péliade .....	2
1.1.1 Taxonomie .....	2
1.1.2 Mode de vie .....	2
1.1.3 Aire de répartition et statut législatif .....	4
1.1.4 Causes de déclin.....	5
1.1.5 Réintroduction .....	7
1.2. L'outil moléculaire : les microsatellites.....	8
1.2.1. Principe général .....	8
1.2.2. Avantages et désavantages .....	10
1.3. Concepts de génétique des populations.....	10
1.3.1. Equilibre Hardy-Weinberg .....	10
1.3.2. Dérive génétique.....	12
1.3.3. Les statistiques « F » .....	12
1.3.4. Déséquilibre de liaison.....	14
1.3.5. Allèle nul .....	15
1.3.6. Goulot d'étranglement .....	16
1.4. Contexte, objectifs et plan .....	17
2. Matériel et méthode .....	19
2.1. Matériel.....	19
2.1.1. Echantillonnage.....	19
2.1.2. Base de données .....	20
2.2. Méthode.....	21
2.2.1. Extraction d'ADN.....	21
2.2.2. Amplification.....	22
2.2.3. Electrophorèse .....	25
2.2.4. Logiciel de lecture .....	25
2.3. Logiciels utilisés.....	26
2.3.1. GENETIX .....	26
2.3.2. FSTAT.....	26

2.3.3.	MEGA .....	27
2.3.4.	Bottleneck .....	27
2.3.5.	STRUCTURE .....	28
2.3.6.	CERVUS.....	29
3.	Résultats & discussion .....	30
3.1.	Mises au point technique.....	30
3.1.1.	Conservation des écouvillons et extraction.....	30
3.1.2.	Mise au point de l'amplification : Primers et multiplex .....	32
3.1.2.1.	Formation des multiplex .....	32
3.1.2.2.	Amorces récalcitrantes.....	34
a)	Cas de VB_Bp9.....	34
b)	Cas de VB_A8 et de VB_B10.....	35
c)	Cas de VB_B18.....	36
3.1.3.	Calibration.....	37
3.2.	Analyses statistiques .....	39
3.2.1.	Diversité génétique .....	39
3.2.2.	Assignation parentale .....	43
3.2.3.	Richesse allélique des individus de l'élevage.....	44
3.2.4.	Analyse Factorielle de Correspondances (AFC) .....	46
3.2.5.	Distances génétiques et dendrogrammes .....	48
3.2.6.	Goulot d'étranglement .....	53
3.2.7.	Structure .....	55
3.2.7.1.	Vérification des hypothèses .....	55
a)	Equilibre d'Hardy-Weinberg.....	55
b)	Déséquilibre de liaison .....	57
c)	Isolement par la distance .....	58
3.2.7.2.	Résultats .....	59
4.	Conclusion et perspectives.....	63
	Bibliographie .....	65

## Remerciements

Je tiens tout d'abord à remercier mes promoteurs, Marie-Christine Flamand et François Chaumont, de m'avoir fait confiance et accueillie au sein de leur laboratoire, d'avoir répondu à mes différentes questions avec enthousiasme, et d'avoir fait de mon travail de fin d'études une expérience enrichissante. Je les remercie également pour leur disponibilité, les conseils qu'ils m'ont octroyés et leur suivi.

Je tiens également à remercier Adrian Neveux pour le temps et l'aide qu'il m'a accordés. Comme on dit, l'équipe est petite, mais le soutien est grand.

Mes remerciements vont aussi à toutes les personnes externes à l'Université qui ont grandement contribué à la réalisation de ce mémoire. Je fais ici référence à Sylvain Ursenbacher, Éric Graitson, Mathieu Bufkens et Christoffel Bonte, qui par leur contribution en termes de connaissances partagées, de temps consacré, mais aussi de matériel et échantillons fournis se sont révélés indispensables à la réussite de ce travail. Ces remerciements vont également à ceux qui les ont aidés dans ces démarches, je pense notamment au Dr Goulven Rigaux et à Vanessa Laus qui ont été d'une grande aide lors du prélèvement d'ADN dans l'élevage de Mathieu Bufkens.

Je souhaiterais ensuite remercier mes lecteurs, Pierre Bertin et Mathieu Bufkens, pour le temps consacré à la lecture et la compréhension de mon travail afin de me permettre d'affiner ma réflexion.

Un grand merci également à mon groupe d'amies ainsi qu'à ma maman, pour le soutien sans failles dont elles ont fait preuve pendant la rédaction de ce mémoire.

Je terminerai par remercier tous ceux que j'aurais pu oublier de mentionner et qui ont contribué de près ou de loin à la réalisation de ce travail.

# 1. Introduction

## 1.1. La vipère péliade

Cette partie vise à vous présenter l'espèce qui fait l'objet de ce mémoire, mais également les différents aspects qui vous permettront de mieux comprendre les enjeux liés aux objectifs de ce dernier.

### 1.1.1 Taxonomie

Le terme vipère péliade fait référence à l'espèce *Vipera berus* appartenant à la classe des Reptilia, à la sous-classe des Lepidosauria, à l'ordre des Squamata, au sous-ordre des Serpentes, à la famille des Viperidae, à la sous-famille des Viperinae et au genre *Vipera*. *Vipera berus* est comprise dans les *Vipera pelias* qui tout comme *Vipera ammodytes* et *Vipera latastei*, est issu de *Viperas s. s.* (Garrigues, Dauga, Ferquel, Choumet, & Failloux, 2005). Il existe trois sous-espèces de vipères péliades nommées *V. b. bosniensis*, *V. b. sachalinensis* et *V. b. berus*, établies du fait des différences morphologiques (dessins dorsaux, nombres d'écailles ventrales et subcaudales) les caractérisant (Nilson, Andrén, & Szyndlar, 1994). Cependant, si on se fie au test du cytochrome b, seulement deux d'entre elles (*V. b. bosniensis* et *V. b. berus*) sont génétiquement différentes (Joger, et al., 1997).

### 1.1.2 Mode de vie

Ce petit serpent de coloration variable avec une bande dorsale longitudinale en zigzag continu caractéristique et dont la taille varie de 45 à 70 cm en moyenne, avec des femelles plus grosses et plus grandes que les mâles, occupe des milieux très variés. En effet, on peut le retrouver dans des marais, des landes, des dunes, en lisière de forêts et jusqu'à 3000 mètres d'altitude (Paquay & Graitson, 2007). Les vipères mâles sont en âge de se reproduire durant leur troisième année de vie, tandis que les femelles le sont durant leur quatrième année et ont un cycle bi-triennal, signifiant qu'elles ne s'accouplent que tous les deux ou trois ans (Paquay & Graitson, 2007) et (Ursenbacher, 2005).

Les mâles sortent d'hibernation en premier, aux alentours de mars-avril, en moyenne deux semaines avant les femelles et les juvéniles. Une fois leur hibernation terminée, ceux-ci entreprennent une migration dans le but de rejoindre des zones suffisamment exposées au

soleil pour leur permettre de réaliser des bains de soleil ou "*basking*" (Paquay & Graitson, 2007). Ces bains de soleil leur permettent de sortir de leur léthargie post-hibernation en les réchauffant jusqu'à ce qu'ils atteignent leur température corporelle optimale de 24-33°C chez les adultes et 24-28°C chez les juvéniles. Une fois la température optimale atteinte les animaux alternent entre zones ombragées et zones exposées (Santune, et al., 2012).

Une fois le *basking* terminé, les vipères mâles vont muer avant de partir en recherche d'une partenaire d'accouplement. Les mâles les plus lourds ont plus de chance de réussite dans le jeu de la reproduction que les autres. Les femelles quant à elles ne muent qu'après l'accouplement et peuvent s'accoupler avec différents mâles (Ursenbacher, 2005).

L'accouplement est généralement suivi de la période d'alimentation pour les mâles et les femelles non gestantes, car pendant l'accouplement et dans leur année de procréation les vipères matures ne se nourrissent pas ; il en va de même pendant la mue (Saint-Girons, 1979). La vipère péliade a généralement 3 périodes d'alimentation : début juin, juillet et fin août. Cette affirmation fait exception des femelles gestantes qui, le plus souvent, ne s'alimentent qu'après avoir mis bas, en septembre, passant le plus clair de leur temps au soleil afin de permettre le développement embryonnaire de leurs petits (Saint-Girons, 1979) et (Paquay & Graitson, 2007).

La fin de l'accouplement marque aussi le début de la période où les mâles sont moins observables, pratiquant une sorte d'errance, et ce, jusqu'à l'hibernation en octobre. Les femelles gestantes, au contraire, seront plus percevables à cette période compte tenu de leur nécessité de s'exposer au soleil pour le développement de leur portée. Leur temps gestatoire est variable et se termine de la mi-août (lors des été ensoleillés) à la mi-septembre. Chaque femelle a en moyenne 6 à 10 vipéreaux en fonction de sa taille (Paquay & Graitson, 2007).

Etant une espèce ovovivipare, une fois sortis du cloaque, les vipéreaux sont livrés à eux-mêmes. Après leur naissance ils vont, tout comme les adultes à cette période de l'année, migrer vers un lieu d'insolation proche de leur lieu d'hibernation ou "*hibercula*" avant de procéder à cette dernière de la mi-octobre à la fin février (Vacher & Geniez, 2010), (Santune, et al., 2012) et (Paquay & Graitson, 2007). La Figure 1 ci-après synthétise le cycle saisonnier de la Vipère péliade.

	Jan	Fev	Mar	Avr	Mai	Juin	Juill	Août	Sept	Oct	Nov	Dec
Hibernation												
Migration et basking des mâles												
Migration et basking des femelles												
Mue des mâles												
Mue des femelles												
Recherche de partenaires et accouplement												
Recherche de nourriture	Mâles											
	Femelles											
Naissance												
Migration vers l'habitat d'hibernation												

Figure 1 : Cycle saisonnier de la vipère péliade (Santune, et al., 2012)

### 1.1.3 Aire de répartition et statut législatif

La vipère péliade est le serpent terrestre dont l’aire de répartition est la plus étendue sur la planète. Comme vous pouvez en partie le voir sur la Figure 2, celle-ci s’étend de la Grande-Bretagne à la côte pacifique russe en passant par les Balkans et la Scandinavie (Vacher & Geniez, 2010).

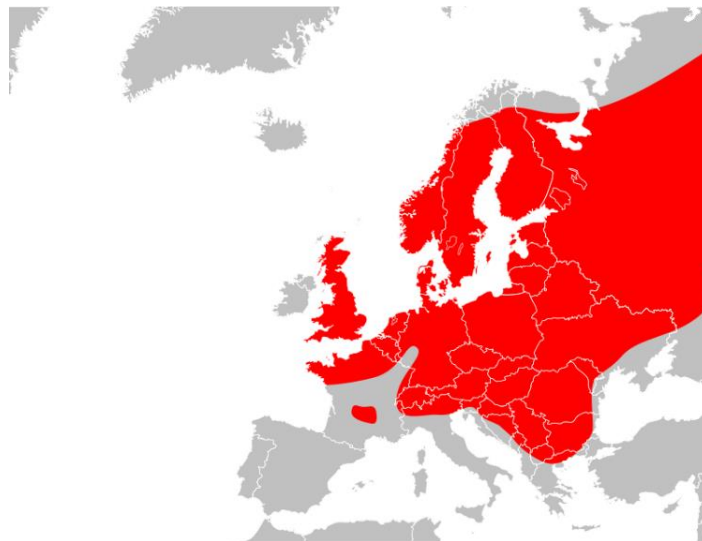


Figure 2 : Carte représentant l'aire de répartition de vipera berus

Malgré sa grande aire de répartition, elle est actuellement en danger de disparition dans l’ouest de l’Europe (Ursenbacher, 2005). Elle est cependant classée “taxon à préoccupation mineure” dans la liste rouge mondiale de l’Union internationale pour la conservation de la nature (UICN) ainsi que dans la liste rouge des reptiles d’Europe. L’UICN précise cependant

que les populations de vipère péliade sont en déclin (Santune, et al., 2012). De fait, les statuts diffèrent en fonction du pays considéré. En France, par exemple, elle est protégée par l'arrêté ministériel du 19 novembre 2007, mais considérée comme "taxon à préoccupation mineure", tandis qu'en Suisse elle est classée en tant qu'"espèce menacée" à "espèce très menacée" en fonction des régions considérées (Santune, et al., 2012) (Ursenbacher, 2005). La Belgique quant à elle possède deux listes différentes en fonction de la région considérée : Wallonie ou Flandre. Dans la première, où sa répartition est relictuelle, elle est classée comme "espèce en danger" ; tandis que dans la deuxième, où seules deux populations isolées y subsistent, elle est considérée comme "espèce en voie d'extinction" (Santune, et al., 2012).

#### 1.1.4 Causes de déclin

Les vipères péliades sont vivipares avec une fécondité et une vigueur basse et une haute philopatrie. En effet, la plupart des femelles se reproduisent pour la première fois à quatre ans puis tous les deux-trois ans et produisent en moyenne 1.3 portées durant leur vie reproductive (Julian & Hodges, 2019). Ces caractéristiques les rendent historiquement vulnérables aux extinctions locales et sont amplifiées du fait que les serpents sont parmi les vertébrés les moins appréciés par l'Homme (Ball, et al., 2020).

Ces facteurs, en association avec la consanguinité, augmentent nettement la probabilité d'extinction. D'après (Ball, et al., 2020) la perte de diversité génétique et la dépression de consanguinité représentent la principale menace génétique pouvant mener à une extinction. Ils ajoutent que la taille des populations est un facteur important dans la perte de diversité génétique. En effet, la consanguinité panmictique désigne une probabilité élevée de s'accoupler avec un apparenté lors d'un choix aléatoire de partenaire et elle est directement liée à la diminution du nombre d'individus dans une population (De Fazio, 2010). Elle s'oppose à la consanguinité systémique qui elle fait référence à un choix non aléatoire du partenaire et qui est indépendante de la taille de la population. Ce facteur de taille de population a d'ailleurs conduit au concept de taille minimale viable de population (Ball, et al., 2020). Cette association consanguinité/taille de la population n'est pourtant pas une condition suffisante à l'établissement d'un risque d'extinction.

En effet l'étude de la population de l'île de Hallands Väderö en Suède (Julian & Hodges, 2019) a permis de démontrer que, du fait du système de reproduction de l'espèce, la consanguinité

est un risque d'extinction uniquement si la population concernée a une base génétique très restreinte. En effet, une population de 200 individus a réussi à maintenir une haute diversité génétique malgré un isolement de 6 000 ans. Ceci implique qu'une fois que la diversité génétique d'une population souffrant de perte d'hétérozygotie est boostée, celle-ci devrait voir sa population augmenter et se stabiliser durablement. Ceci a d'ailleurs été confirmé via l'expérience réalisée à Smygehuk (Madsen, Ujvari, & Olsson, 2004), autre population isolée de Suède, qui elle a connu, dans les années 80, un déclin lié à la perte de diversité de sa population. Après introduction d'une vingtaine de mâles, la population s'est vu croître comme on peut l'observer sur la Figure 3.

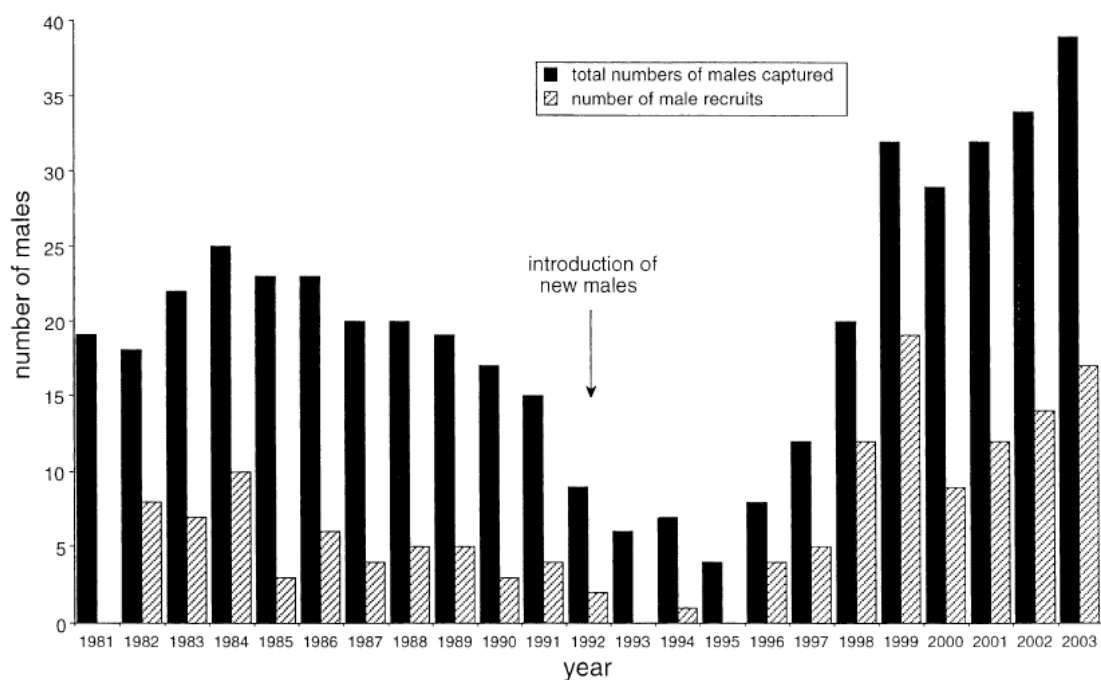


Figure 3 : Evolution du nombre de mâles dans la population de Smygehuk avant et après l'introduction de nouveaux mâles (Madsen, Ujvari, & Olsson, 2004)

Les analyses génétiques du succès de Smygehuk et les séquençages qui s'en sont suivis ont permis de définir le boost génétique comme un gain d'hétérozygotie, notamment dans les gènes responsables de réponses innées aux maladies (Julian & Hodges, 2019), gènes primordiaux à la survie. Cette définition n'est cependant pas optimale puisqu'elle revient à dire qu'un boost génétique est une hausse du degré de polymorphisme des gènes d'intérêt d'une population, étant donné que le degré de polymorphisme est défini comme étant le pourcentage de gènes polymorphes (Serre, 2006). Le problème avec cette définition est lié à l'indicateur de la diversité génétique globale d'une population qu'il sous-tend (le degré de

polymorphisme). En effet, cet indicateur ne prend en compte ni le nombre d'allèles pour un gène, ni la valeur des fréquences alléliques, mais uniquement le fait qu'un gène soit ou non polymorphe (Serre, 2006). Sachant qu'un gène est considéré polymorphe à partir du moment où au moins deux formes alléliques de fréquences supérieures ou égales à 1% sont détectées. Serre (Serre, 2006) conseille plutôt l'utilisation du taux moyen d'hétérozygotie d'une population comme indicateur. Ce dernier est défini comme la moyenne des taux d'hétérozygotie pour un grand nombre de gènes étudiés.

D'après une étude menée en zoo sur des espèces non menacées (Asmodé, 1990), la mortalité des jeunes à la naissance et dans les premiers mois de vie avoisine 60% dans les populations consanguines contre 20% dans les populations normales. D'après (Asmodé, 1990), plus une espèce a une population abondante sans séparation dans l'aire de répartition dans la nature plus elle est sensible à la consanguinité en captivité.

Cependant, (Ball, et al., 2020) s'accordent pour dire qu'il est difficile de savoir comment les résultats d'études génétiques devraient influencer la gestion de la conservation de l'espèce, puisqu'on ignore dans quelle mesure la perte de diversité contribue au déclin et à l'extinction dans la nature. L'isolement ne conduit pas forcément à la consanguinité (Julian & Hodges, 2019). En effet, il se peut qu'il n'y ait aucune perte d'hétérozygotie significative malgré qu'il existe un lien étroit entre les individus d'une population. Dans ce cas-ci d'autres facteurs peuvent être responsables du déclin des petites populations de vipères péliades. Nous pouvons citer des facteurs comme la diminution des milieux favorables à l'espèce faisant suite à une intensification de l'utilisation des espaces naturels, le changement climatique (qui affecte, entre autres, la durée de gestation des femelles et donc la probabilité d'avoir un vipéreau en bonne santé) ou encore le braconnage et la hausse du nombre de prédateurs (sangliers et faisans) (Paquay & Graitson, 2007).

#### 1.1.5 Réintroduction

La réintroduction est définie par (Asmodé, 1990) comme étant l'établissement d'une population animale dans une zone d'habitat naturel à l'intérieur de son ancienne répartition historique. Asmodé ajoute qu'il y a généralement quatre phases à cette réintroduction. Celles-ci sont les suivantes : tout d'abord l'étude de faisabilité, qui est une étude montrant le potentiel de succès de la réintroduction, elle est suivie de la phase préparatoire qui a pour

but de choisir la zone de relâcher et les sources, nombre et sexe des animaux relâchés. Les deux dernières phases étant le relâché et le suivi.

Malgré le respect de ces étapes, la réintroduction peut s'avérer infructueuse. En effet, comme l'indique (Lecomte, 1990), les transferts d'animaux d'un milieu à un autre résultent souvent en un échec du fait de leur impossibilité à s'adapter aux nouvelles conditions rencontrées et on obtient souvent de meilleurs résultats en introduisant des groupes sociaux constitués plutôt que des individus isolés. Bien qu'il ajoute que ces difficultés et observations sont sans doute plus vraies pour les mammifères que pour les autres vertébrés pour lesquels on manque généralement d'analyses sérieuses.

Pour finir, d'après (Asmodé, 1990), pour qu'un projet de réintroduction maximise ses chances d'être fructueux il faut que les fondateurs de la population soient les plus nombreux possibles, qu'ils ne soient pas apparentés et que le nombre d'individus dans la population augmente rapidement via la reproduction pour minimiser la perte de diversité génétique.

## 1.2. L'outil moléculaire : les microsatellites

Cette section a pour but de vous permettre de vous familiariser avec l'outil moléculaire à l'origine des résultats

### 1.2.1. Principe général

Le présent travail a nécessité l'utilisation de microsatellites. Un microsatellite ou séquence microsatellite est une séquence d'ADN formée par une répétition en tandem de motifs composés de 1 à 5 nucléotides qui est alors dit di-, tri-, tétra- ou pentanucléotidique. Ces séquences sont dispersées le long du génome, dans les parties non-codantes. Etant donné que plus de 95% de l'ADN total est intergénique ou intronique dans la majorité des espèces domestiques et que la sélection naturelle y tolère une plus grande variabilité, ces régions sont les principales pourvoyeuses de polymorphisme. Ceci fait des microsatellites des marqueurs à haut potentiel d'informations et de variations.

Leur localisation dans le génome est relativement conservée entre espèces phylogénétiquement proches. Leurs longueurs, quant à elles, varient selon l'espèce, mais aussi d'un individu à l'autre et d'un allèle à l'autre chez un même individu, voire d'une cellule à l'autre du fait d'erreurs au cours de la réplication de l'ADN. Ces loci mutent par insertion ou

délétion d'une ou plusieurs répétitions et le taux de mutations augmente généralement avec l'augmentation des séquences répétées (Wierdl, Dominska, & Petes, 1997). La Figure 4 tirée de (Srivastava, Avvaru, Sowpati, & Mishra, 2019) donne un aperçu des longueurs moyennes de microsatellites chez les reptiles ainsi que pour d'autres sous-groupes d'organismes eucaryotes.

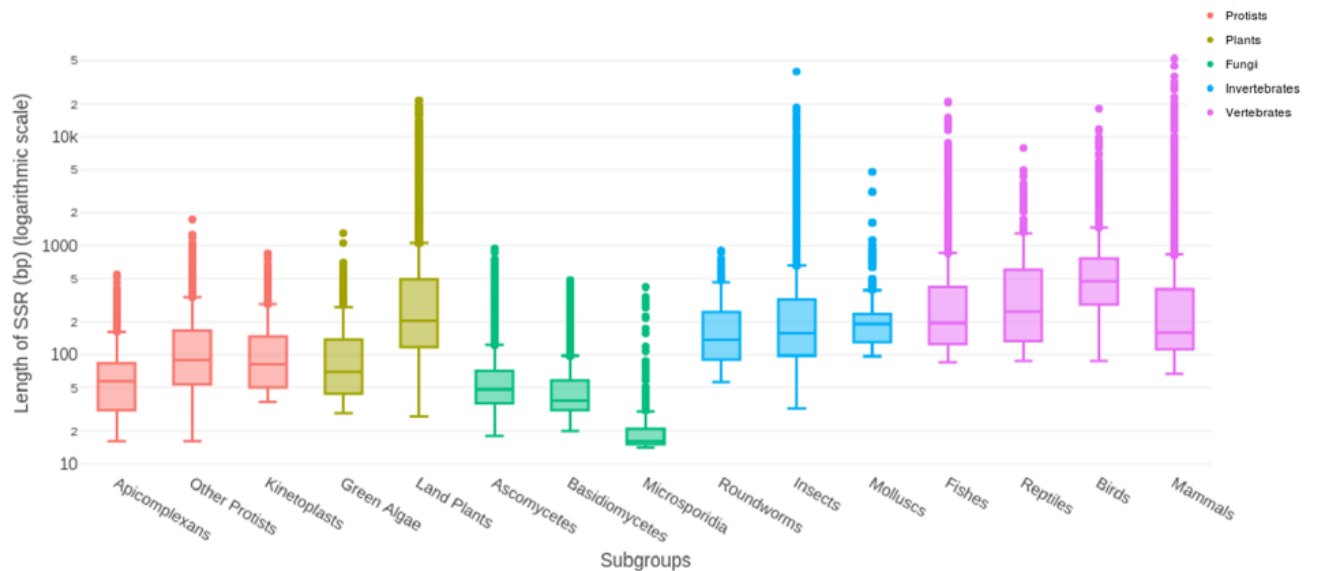


Figure 4 : Représentation graphique des longueurs moyennes des microsatellites en fonction des règnes et des phylums (Srivastava, et al., 2019)

Deux mécanismes ont été proposés pour expliquer l'instabilité des microsatellites : le glissement de l'ADN polymérase avec une recombinaison inégale et le crossing-over entre séquences répétées mal alignées.

La première explication implique la dissociation des brins d'ADN en réplication, suivie d'une réassociation mal alignée, résultant en unités de répétitions non-appairées et donc à la formation de boucles. Si la réplication de l'ADN s'achève sans que ces boucles n'aient été résolues, le nombre de répétitions sur le brin nouvellement synthétisé sera soit plus important soit plus faible que le nombre original, en fonction du brin sur lequel la boucle s'est formée. Si la boucle est sur le brin d'amorce, le nombre de répétitions augmente ; tandis que si elle est sur le brin modèle les répétitions diminuent (Streisinger, et al., 1966).

La seconde explication est plus simple puisqu'il s'agit simplement d'un échange d'informations génétiques entre deux chromatides sœurs ou chromosomes homologues

aussi appelés crossing-over. Si le crossover est réciproque, on obtient à la fois un microsatellite plus court que l'original et un plus long (Smith, 1973).

### 1.2.2. Avantages et désavantages

Les microsatellites ont, comme expliqué ci-avant, un haut taux de polymorphisme dû aux mutations aléatoires qu'ils subissent. Ces mutations sont de l'ordre de  $10^6$  à  $10^7$  fois plus fréquentes que les mutations ponctuelles classiques. Cet aspect n'est cependant pas le seul atout de ces loci. En effet, ils sont également codominants. En découle qu'il est ainsi possible de détecter les deux allèles en présence. Mais surtout ils sont répartis de façon homogène dans le génome. De plus, la technique dans laquelle ils sont employés est facilement répétable et reproductible d'un laboratoire à un autre.

La technique utilisée est basée sur l'amplification de l'ADN par réaction en chaîne par polymérase, ou PCR, avec des amorces spécifiques à des séquences en amont des séquences microsatellites d'intérêt. Ceci implique que le nombre de marqueurs utilisables est restreint à ceux dont l'emplacement des séquences est connu ; impliquant qu'on sait si un microsatellite suit cette séquence et quel motif est répété.

## 1.3. Concepts de génétique des populations

Cette section décrit de façon succincte les différents concepts fondateurs des analyses statistiques qui ont été menées dans le cadre de ce mémoire.

### 1.3.1. Equilibre Hardy-Weinberg

Le modèle communément appelé « équilibre d'Hardy-Weinberg » a été conçu par Hardy, un mathématicien britannique, et Weinberg, un médecin allemand, de façon indépendante. Ce modèle est un concept central à de nombreux modèles de diversité et de différenciation génétique et sert de base à une immense partie des études de génétique des populations (De Meeûs, 2012). Les hypothèses de ce modèle sont les suivantes :

- La taille de la population est infinie
- Cette population ne subit pas de mutations
- Cette population ne subit pas de migrations
- La population d'intérêt ne connaît pas de pression de sélection (neutralité)
- La reproduction y est sexuée avec rencontre au hasard des gamètes (panmixie)

- Il n'y a pas de chevauchement de générations

Ce modèle considère un locus à deux allèles A et B de fréquences  $p_1$  et  $p_2$  dans une telle population nommée N. La part de la population de génotypes AA est noté  $N_{AA}$ , celle de génotypes AB,  $N_{AB}$  et celle de génotypes BB,  $N_{BB}$ . On a donc  $N = N_{AA} + N_{AB} + N_{BB}$ . Les fréquences  $p_1$  et  $p_2$  dans la population sont donc :

$$p_1 = \frac{N_{AA} + \frac{N_{AB}}{2}}{N} \quad \text{Et} \quad p_2 = \frac{N_{BB} + \frac{N_{AB}}{2}}{N} \quad (1)$$

Puisque les gamètes se rencontrent au hasard, nous pouvons construire le Tableau 1 pour décrire leur rencontre.

*Tableau 1 : Tableau de rencontre de 2 allèles A et B d'un locus donné. Leurs fréquences  $p_1$  et  $p_2$  et celles résultantes de la rencontre sont reprises entre parenthèses.*

Femelle/Mâle	A ( $p_1$ )	B ( $p_2$ )
A ( $p_1$ )	AA ( $p_1^2$ )	AB ( $p_1p_2$ )
B ( $p_2$ )	AB( $p_1p_2$ )	BB ( $p_2^2$ )

Nous attendons donc, dans les zygotes ainsi formés, les proportions de génotypes :  $p_1^2$ ,  $2p_1p_2$  et  $p_2^2$  pour AA, AB et BB respectivement (\*). Et puisqu'il n'y a ni sélection, ni mutation, ni migration et que la population est infinie, ces proportions resteront les mêmes chez les adultes de la génération suivante. En nous rappelant que la somme  $p_1 + p_2$  est nécessairement égale à 1, nous démontrons facilement en appliquant (\*) dans l'équation (1) que les nouvelles fréquences alléliques  $p_1'$  et  $p_2'$  seront :

$$p_1' = \frac{p_1^2 + \frac{2p_1p_2}{2}}{p_1^2 + 2p_1p_2 + p_2^2} = \frac{p_1^2 + p_1p_2}{(p_1 + p_2)^2} = \frac{p_1(p_1 + p_2)}{(1)^2} = \frac{p_1(1)}{(1)^2} = p_1 \quad (2)$$

De même  $p_2' = p_2$

C'est cette immobilité du modèle qui fait que l'on appelle l'équilibre de Hardy-Weinberg (De Meeûs, 2012).

Cet équilibre peut être rompu dans les cas où les reproductions sont non-aléatoires et/ou consanguines, lorsqu'il y a stratification de la société, que la population est de taille finie ou prône à la migration et aux mutations. Dans ces conditions, la fréquence des gènes subit des

fluctuations plus ou moins importantes et la population est dite en déséquilibre de Hardy-Weinberg (Serre, 2006).

### 1.3.2. Dérive génétique

Il a été précisé que l'équilibre d'Hardy Weinberg pouvait être altéré, entre autres, par migration. La migration peut notamment engendrer l'altération par dérive génétique. Cette dernière est définie comme une modification notable dans les fréquences des gènes d'une sous-population après sa formation et son isolement d'une population plus importante. La dérive génétique peut donc favoriser l'enracinement d'un allèle, qui n'est pas nécessairement favorable, par augmentation de sa fréquence dans la population (Serre, 2006). Cette dérive sera d'autant plus forte que la sous-population est petite.

La perte de variabilité par dérive génétique est inversement proportionnelle à la taille effective ( $N_e$ ) de la population (Guyomard, 1989).  $N_e$  se définit comme la taille idéale de la population pour qu'elle ait des comportements similaires à la population réelle (Crow & Denniston, 1988). Dans une population naturelle, ce terme  $N_e$  dépend du nombre total de géniteurs, du sex-ratio, de la variance de la taille des familles, du degré de chevauchement entre générations et du taux de migration. La dérive dépend donc de l'effectif participant à la reproduction (Guyomard, 1989).

Dans (Crow & Denniston, 1988), les auteurs font référence à une formule, reprise ci-après, reliant  $N_e$  à l'hétérozygotie de la population aux générations  $t$  et  $t-1$ .

$$H_t = H_{t-1} \left(1 - \frac{1}{2N_e}\right) \quad (3)$$

Cette formule nous permet de lier la variation de l'hétérozygotie d'une population sur une période donnée à la variation de la taille effective, et donc indirectement à la dérive génétique.

### 1.3.3. Les statistiques « F »

Les statistiques « F » comprennent les indices de fixation ou de différenciation plus communément dénommées  $F_{ST}$  de Weir et Cockerham (Weir & Cockerham, 1984) ainsi que les coefficients de consanguinité ou  $F_{IS}$  de Wright (Wright, 1951) et sont un ensemble

d'indicateurs ayant pour but de mesurer l'excès ou le manque d'hétérozygotie observée par rapport au nombre attendu sous l'équilibre d'Hardy-Weinberg.

Le premier indicateur, le  $F_{ST}$ , compare l'hétérogénéité des fréquences alléliques entre subdivisions d'une population. Il a pour but de quantifier les divergences génétiques à partir de fréquences alléliques, en retournant des valeurs comprises entre 0 et 1. Une valeur de 0, signifie que les deux populations comparées sont génétiquement identiques (pas de structuration) puisque la diversité génétique est nulle ou minimale. Tandis qu'une valeur de 1 signifie qu'elles sont totalement différentes (toutes les populations sont fixées pour l'un ou l'autre des allèles) et donc, qu'elles ont une diversité génétique maximale (De Meeûs, 2012). On présuppose qu'il n'y a eu aucune mutation dans les populations. L'expression du  $F_{ST}$  est reprise ci-après.

$$F_{ST} = \frac{H_{exp,Tot} - H_{exp,Sp}}{H_{exp,Tot}} \quad (4)$$

Avec  $H_{exp,Tot}$  l'hétérozygotie attendue en cas de panmixie globale (tous les individus, de toutes les sous-populations se croisent au hasard), et  $H_{exp,Sp}$  l'hétérozygotie moyenne attendue en cas de panmixie locale (tous les individus d'une sous-population se croisent au hasard).

Le second, le  $F_{IS}$ , différencie les individus à l'intérieur des sous-populations. Il a pour but de quantifier l'écart entre l'hétérozygotie moyenne attendue sous l'équilibre d'Hardy-Weinberg et l'hétérozygotie moyenne des individus en renvoyant des valeurs comprises entre -1 et 1. Une valeur de 1 indique une déficience en hétérozygotes, tandis qu'une valeur négative implique un excès d'hétérozygotes, les deux signifiant un écart de la population par rapport à l'équilibre d'Hardy-Weinberg. Une valeur nulle quant à elle indique un parfait équilibre d'Hardy-Weinberg (De Meeûs, 2012). Il donne donc une indication du déficit relatif en hétérozygotes dans les populations. Il est à noter que la valeur du  $F_{IS}$  peut être influencée par l'effet Wahlund et les allèles nuls.

$$F_{IS} = \frac{H_{exp,Sp} - H_{obs,ind}}{H_{exp,Sp}} \quad (5)$$

Avec  $H_{obs,ind}$  l'hétérozygotie moyenne observée chez un individu.

Un troisième indicateur, le  $F_{IT}$ , existe. Il différencie les individus par rapport à la population entière et a pour but de mesurer l'homozygotie des individus de la population totale. Tout

comme le  $F_{IS}$ , il renvoie des valeurs qui varient de -1 à 1 (De Meeûs, 2012). Cet indicateur ne sera cependant pas utilisé dans le cadre de ce mémoire. Son expression est la suivante :

$$F_{IT} = \frac{H_{exp,Tot} - H_{obs,ind}}{H_{exp,Tot}} \quad (6)$$

Ces trois indicateurs sont liés par la relation suivante :

$$F_{ST} = 1 - \frac{(1 - F_{IT})}{(1 - F_{IS})} \quad (7)$$

#### 1.3.4. Déséquilibre de liaison

On parle de déséquilibre de liaison, également appelé association gamétique ou association allélique, lorsque la fréquence des gamètes porteurs de deux gènes A et B est différente du produit des fréquences de A et de B (Feingold, 1991).

Si on considère que les loci auxquels appartiennent ces gènes sont di-alléliques - A et a pour le premier et B et b pour le second - leurs fréquences géniques respectives sont alors les suivantes :  $P_A$ ,  $P_a$ ,  $P_B$  et  $P_b$ . En observant que  $P_{AB}$  est la fréquence des gamètes porteurs des gènes A et B, on peut définir le déséquilibre de liaison D comme étant :

$$D = P_{AB} - P_A * P_B \quad (8)$$

Si D est nul cela revient à dire que  $P_{AB} = P_A * P_B$ . Autrement dit, il n'y a pas de déséquilibre de liaison. Tout résultat différent de 0 implique un déséquilibre (Feingold, 1991).

Si déséquilibre de liaison il y a, sa valeur évoluera avec les générations dépendamment du taux de recombinaison et tendra vers l'équilibre, comme le montre la Figure 5. Ceci s'explique par le fait que le déséquilibre de liaison après t générations ( $D_t$ ) est lié au déséquilibre initial ( $D_0$ ) via cette formule.

$$D_t = (1 - \theta)^t D_0 \quad (9)$$

Où  $\theta$  représente la fréquence des recombinaisons entre les deux loci considérés.

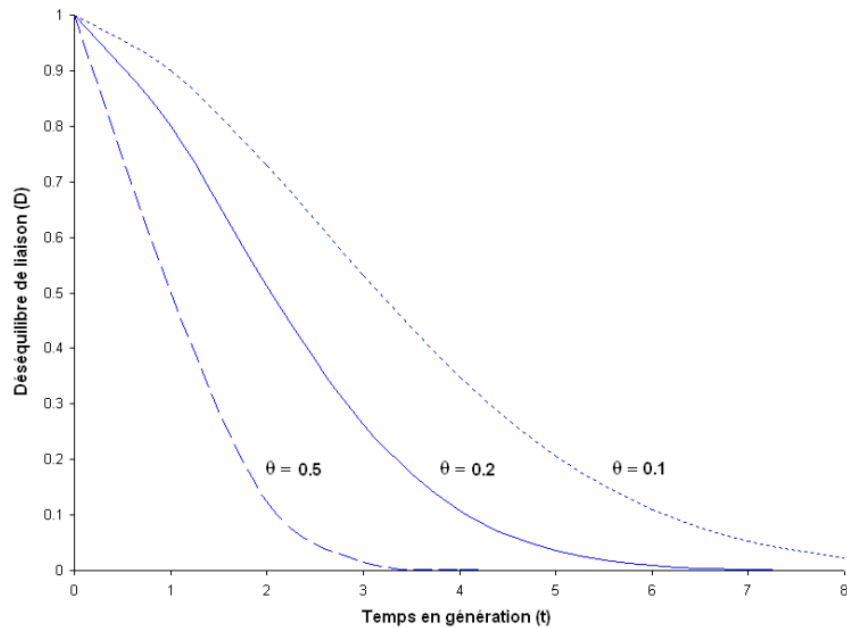


Figure 5 : Evolution du déséquilibre de liaison ( $D$ ) en fonction du taux de recombinaison ( $\vartheta$ ), au cours des générations ( $t$ ) (Jomphe, 2006)

Les causes du déséquilibre de liaison sont réparties en deux groupes : celles liées à l'évolution démographique (fusion entre deux populations à fréquences géniques différentes) et celles liées à la sélection naturelle (mutation et/ou recombinaison) (Feingold, 1991) & (Jomphe, 2006).

### 1.3.5. Allèle nul

Les allèles nuls sont des allèles présents au sein du génome, mais qui ne sont pas détectés. Dans notre cas, un allèle nul microsatellite est un allèle au niveau d'un locus microsatellite qui ne parvient pas à s'amplifier jusqu'aux niveaux détectés par la réaction en chaîne par polymérase (PCR) (Dakin & Avise, 2004).

Il existe différentes causes potentielles aux allèles nuls. Parmi elles il y a le mauvais appariement des amorces dû à une divergence de séquence nucléotidique (pouvant impliquer des mutations ponctuelles par exemple) dans l'une ou les deux amorces (Dakin & Avise, 2004). Cette implication est d'autant plus notable que d'après (Kwok, et al., 1990), les mutations à l'extrémité 3' du site d'amorçage sont particulièrement préjudiciables aux amplifications PCR.

Une autre cause possible est liée à l'amplification différentielle d'allèles de tailles variables. En effet, du fait de la nature compétitive des PCR, les allèles de courte longueur s'amplifient souvent plus efficacement que les grands. Ainsi, le plus petit des deux allèles a plus de chances d'être détecté chez un individu hétérozygote. Les allèles nuls causés par une amplification différentielle sont parfois appelés "nuls partiels" car ils peuvent souvent être rendus visibles en chargeant davantage d'échantillons ou en ajustant le contraste (Dakin & Avise, 2004).

Une troisième source possible d'allèles nuls est l'échec de la PCR dû à une qualité inégale de la matrice d'ADN ou à une faible quantité de matrice. Ces problèmes sont insidieux, car, dans certains cas, seul un ou quelques loci (ou allèles) ne parviennent pas à s'amplifier, alors que d'autres s'amplifient assez facilement à partir de la même préparation d'ADN. Lorsque la matrice d'ADN à un locus est pauvre dans certains exemplaires, mais pas dans d'autres, les échantillons pauvres peuvent apparaître "homozygotes" plutôt qu'hétérozygotes pour l'allèle nul (Dakin & Avise, 2004).

La présence d'un ou plusieurs allèles nuls biaise les analyses provoquant la détection de plus d'homozygotes pour certains loci qu'il y en a en réalité. Ils peuvent donc être la cause d'une déviation fictive d'un déséquilibre d'Hardy-Weinberg (Pemberton, et al., 1995).

#### 1.3.6. Goulot d'étranglement

Le goulot d'étranglement, aussi appelé goulet d'étranglement, se réfère à un épisode de réduction sévère de la population suivi d'une extinction ou d'un accroissement de celle-ci en fonction de la situation (Cornuet & Luikart, 1996). Ces événements sont généralement dus à des changements environnementaux tels que la sécheresse ou encore le manque de proies par exemple.

Ces goulots d'étranglement sont détectables du fait que lorsqu'une population subit une réduction de sa taille effective, elle développe généralement un excès d'hétérozygotie sur des loci sélectivement neutres. Ceci revient à dire que l'hétérozygotie calculée à partir d'un échantillon de gènes est plus importante que l'hétérozygotie attendue si la population était à l'équilibre de dérive de mutation. Cet excès d'hétérozygotie ne persiste qu'un certain nombre de générations jusqu'à ce qu'un nouvel équilibre soit établi (Cornuet & Luikart, 1996).

Deux tests statistiques nécessitant des mesures du nombre d'allèles et de l'hétérozygotie à plusieurs loci d'un échantillon de population permettent de détecter un excès d'hétérozygotie. Le premier test détermine si la proportion de loci présentant un excès d'hétérozygotie est significativement plus importante que celle attendue à l'équilibre. Le second test établit si la moyenne des différences standardisées entre les hétérozygoties observées et attendues est significativement différente de zéro (Cornuet & Luikart, 1996).

Tout comme le dit (Cornuet & Luikart, 1996) , il est tout de même important de noter que, si on ne connaît pas la taille historique des populations étudiées, la conclusion qu'elles ont ou non subi un grave goulot d'étranglement et une perte de variation génétique sont déductives.

#### 1.4. Contexte, objectifs et plan

Comme vous aurez pu le constater dans la partie 1.1 la vipère péliade est une espèce endémique belge en voie d'extinction. Natagora, une association sans but lucratif (ou asbl) ayant comme objectif de protéger la nature et notamment d'enrayer la dégradation de la biodiversité, a donc, par le biais d'Éric Graitson qui y est un chargé de projets, mis en place un plan d'action visant à la conserver. Ce plan d'action comprend notamment sa réintroduction dans de grands espaces adaptés et classés réserves naturelles.

Pour permettre une réintroduction sur ces sites, un élevage, tenu par Matthieu Bufkens, constitué de 5 enclos a été mis en place. Deux grands enclos contiennent des individus ayant des origines génétiques différentes, mais connues. Il a été porté attention à ce qu'il n'y ait pas de mélanges et aucun frères et sœur dans un même enclos. Deux des trois autres enclos de taille plus petite contiennent des serpenteaux à des phases différentes de maturité tandis que le cinquième contient des femelles gravides.

Sachant que l'une des problématiques liées à l'extinction des péliades est la forte consanguinité, il nous a ainsi été demandé d'optimiser la diversité génétique des individus destinés à la réintroduction. Pour se faire, des échantillons d'ADN de chaque individu de l'élevage nous ont été transmis ainsi que ceux de quelques individus néerlandophones de la région de Kalmthout et ceux d'une population des Hautes-Fagnes.

Nos objectifs étaient donc de déterminer les parents des juvéniles de l'élevage pour pouvoir déterminer s'il y a eu un boost génétique entre les deux générations. Il était aussi question d'essayer de déterminer les origines des populations des Hautes Fagnes et de Kalmthout. En déterminant leurs origines, qui étaient alors uniquement hypothétiques, une introduction de certains des individus de ces populations dans l'élevage pourrait être envisagée si leur origine est wallonne et qu'il y a une nécessité et/ou une plus-value à le faire.

Pour pouvoir déterminer l'origine de ces deux populations, nous avons reçu des résultats obtenus par une mémorante, Fanny Demez, en 2012 dans le laboratoire de Sylvain Ursenbacher qui est chargé de recherches à l'université de Bâle (Suisse) dans le département des sciences environnementales. Son domaine d'étude s'étend de la phylogéographie et la taxonomie des vipères européennes à la conservation des populations ; en passant par les effets génétiques dans les petites populations ou par les impacts des espèces invasives sur la biodiversité et les fonctions des écosystèmes. Il a notamment fait son doctorat sur *Vipera berus*.

Pour pouvoir utiliser les résultats de Fanny Demez, un calibrage entre les deux laboratoires était nécessaire et Sylvain Ursenbacher nous a pour cela envoyé des échantillons de références.

## 2. Matériel et méthode

### 2.1. Matériel

#### 2.1.1. Echantillonnage

Les analyses génétiques ont été effectuées à partir de prélèvements salivaires de vipères péliades récoltés via écouvillons. Ces prélèvements ont été réalisés dans 3 lieux différents : l'élevage de Mathieu Bufkens, les Hautes Fagnes et Kalmthout. Deux de ces 3 lieux sont représentés sur la carte à la Figure 6.



Figure 6 : Représentation des lieux naturels desquels proviennent les échantillons reçus

Les écouvillons utilisés et les dates de prélèvement diffèrent en fonction de leur provenance. En effet, les 29 échantillons de l'élevage ont été prélevés avec des écouvillons *E-SWAB liquid amies* en septembre 2021 par le propriétaire de l'élevage, Monsieur Mathieu Bufkens, aidé du Dr Goulven Rigaux et de Madame Vanessa Laus. Les 30 échantillons provenant des Hautes Fagnes ont été prélevés sur une étendue de 5 kilomètres entre mai 2018 et avril 2021 par Monsieur Eric Graitson, avec des écouvillons *Sarstedt*. Les Hautes Fagnes sont réparties en trois régions (Eupen-Ternell, Mont Rigi-Baraque Michel et SPA-Berinzene), mais les échantillons reçus proviennent tous de la région d'Eupen-Ternell. Le Tableau 2 reprend le nombre d'individus échantillonnés par région et leur localisation « exacte ». Enfin les 12 échantillons provenant de Kalmthout ont été prélevés entre novembre 2019 et septembre

2021. Il est à noter qu'une trentaine d'échantillons supplémentaires en provenance de Kalmthout auraient dû être ajoutés à l'étude, mais ces derniers ont malheureusement été perdus lors de leur envoi par la poste.

Tableau 2 : Tableau de répartition des échantillons des Hautes Fagnes en fonction de leur région.

Région	Localisation	Nombre d'individus
Eupen-Ternell	Brackvenn	4
	Brackvenn Nord	10
	Brackvenn Centre	1
	Algemeenes Venn	2
	Steinley callebotis	1
	Hautes Fagnes	1
	Aachener Kreuz	9
	Fagnes de l'est (Eupen)	2

En plus de ces prélèvements salivaires, des extractions d'ADN, au nombre de vingt, réalisées par Sylvain Ursenbacher dans son laboratoire en Suisse ont été reçues. Ces échantillons provenaient notamment de deux régions françaises, le Cantal et la Lozère, mais aussi de certains villages à la frontière franco-belge tels que Vierves, Tailles Claude et Roche Madoux.

### 2.1.2. Base de données

Les 20 échantillons déjà extraits envoyés par Sylvain Ursenbacher étaient accompagnés des résultats que ce dernier avait obtenus dans son laboratoire pour ces échantillons. A également été reçu, de Monsieur Graitson, un tableau contenant les résultats obtenus par Fanny Demez lors de son mémoire intitulé « Caractérisation génétique et conservation des populations de vipères péliades (*Vipera berus*) en Wallonie, et comparaison avec des populations suisses et françaises. » (Demez, 2012-2013) réalisé dans le laboratoire de Sylvain Ursenbacher ; ainsi que la liste des provenances des différents échantillons qu'elle a analysés. Ces données devaient nous permettre de réaliser une calibration entre le laboratoire de Sylvain Ursenbacher et le nôtre et d'ainsi pouvoir intégrer les résultats de Fanny Demez à notre étude.

## 2.2. Méthode

### 2.2.1. Extraction d'ADN

Pour les échantillons sous forme d'écouvillon, nous avons extrait l'ADN selon le même protocole que celui utilisé par Fanny Demez lors de son mémoire. Nous avons donc utilisé le kit QIAGEN DNeasy® Blood & Tissue en suivant essentiellement le protocole proposé pour la purification de l'ADN total à partir de tissus animaux.

La première étape de cette extraction consiste à déposer dans des tubes microcentrifuges de 1.5 ml, ou tubes Eppendorf, la partie en coton des écouvillons, puis d'y ajouter 200 µl de tampon ATL, qui est un tampon de Lyse, de vortexer 15 s avant d'ajouter 20 µl de protéinase K (20mg/ml), de vortexer de nouveau pendant 15 s et d'incuber le tout à 56°C « overnight ». Cette étape sert à solubiliser l'ADN.

Une fois ceci fait, 200 µl de tampon AL (tampon de lyse) ont été ajoutés au mélange avant de vortexer pendant 15 s et d'ajouter 200 µl d'éthanol 99%. L'ajout d'éthanol est suivi d'un vortexage de 30 s après quoi, le tout (écouvillon compris) est transvasé dans des colonnes "DNeasy Mini Spin Column" placées dans des tubes collecteurs de 2 ml. Après une centrifugation de 1 min à 8 500 rpm dans une centrifugeuse Eppendorf, les cotons sont retirés et les tubes collecteurs de 2 ml changés avant d'ajouter 500 µl de tampon de lavage AW1 et de recentrifuger pendant 1 min à 8 500 rpm. Cette centrifugation est suivie d'un nouveau remplacement des tubes collecteurs et d'un ajout de 500 µl de tampon de lavage AW2 avant une nouvelle centrifugation de 3 min, cette fois à 14 000 rpm. Après quoi, les colonnes ont été transférées dans des tubes Eppendorf 1,5 ml et centrifugés pendant 1 min à 14 000 rpm. Pour veiller à ce que la membrane de la colonne soit bien sèche, cette étape a été répétée jusqu'à ce qu'on ne récupère plus de liquide en fin de centrifugation, en veillant à bien remplacer le tube Eppendorf à chaque fois. Les colonnes ont ensuite été placées dans de nouveaux tubes Eppendorf puis 50 µl de tampon d'élution AE ont été directement déposés sur les membranes. Une incubation de 15 min à température ambiante s'en est suivie, finalisée par une dernière centrifugation de 1 min à 8 500 rpm. Une fois l'ADN extrait, il est dosé et conservé à -20°C.

Nous avons également réalisé une extraction sur le morceau de queue d'un animal mort. Pour ce dernier le protocole diffère puisqu'ici c'est celui établi par Madame Flamand dans

(Flamand, 2001-2002) qui est appliqué. Il consiste à introduire une infime fraction (approximativement 3 x 2 mm) du matériel biologique dans un tube Eppendorf 1,5ml dans lequel sont ajoutés 400 µl de tampon CTAB 2x et 10 µl de protéinase K (20mg/ml) pour lyser l'échantillon et ainsi solubiliser l'ADN. Le tampon CTAB 2x est composé de 2% (p/v) de cetyltriméthylammonium bromide, de 1% (p/v) de polyvinylpyrrolidone, de 1,4 M de NaCl, 20 mM d'EDTA, de 100 mM de Tri-HCl à pH 8 et enfin de 0,2% (v/v) de β-mercaptoéthanol. Le tout est incubé à 56°C « overnight ». Au produit de lyse sont ajoutés 800 µl de chloroforme/alcool isoamylique (CIA) (24:1), avant de vortexer puis de centrifuger le tout à 14 000 rpm pendant 5 min afin de séparer les membranes cellulaires, les protéines dénaturées et les polysaccharides de la phase aqueuse. Cette dernière est transférée dans un nouveau tube Eppendorf contenant 800 µl de CIA avant de vortexer le tout et de nouveau centrifuger à 14 000 rpm pendant 5 min. Cette étape de nettoyage est suivie d'une étape de précipitation de l'ADN qui commence avec l'ajout de 375 µl d'isopropanol froid suivi de 5-6 inversements avant de centrifuger à 14 000 rpm pendant 15 min. Le surnageant ainsi obtenu est retiré avant de procéder au lavage du culot avec 200 µl d'éthanol 70% (v/v) froid. S'en suivent une centrifugation de 5 min à 14 000 rpm, une aspiration du liquide et un séchage de 5-10 min du culot au speedvac avant une resuspension de ce dernier dans plus ou moins 50 µl d'eau stérile en fonction de la taille du culot. On laisse l'extrait ainsi obtenu reposer 2 à 3 h à température ambiante avant de le conserver au congélateur.

Ces extractions ont toutes été suivies d'un dosage. Pour cela, une solution A composée de 200 µl de tampon Qubit™ dsDNA BR et de 1 µl de réactif Qubit™ dsDNA BR pour chaque échantillon à doser a été réalisée. Ensuite, 198 µl de cette solution A sont prélevés avant d'être mélangés à 2 µl de l'ADN extrait dans un « tube de réaction » de 0.5 ml. Ce tube est ensuite analysé à l'aide d'un fluoromètre Qubit®2.0.

### 2.2.2. Amplification

Comme expliqué au point 1.2, les résultats obtenus se basent sur un outil moléculaire que sont les séquences microsatellites. Différents marqueurs microsatellites spécifiques à des séquences d'ADN sont impliqués dans l'utilisation de cet outil. Afin d'obtenir des quantités conséquentes de microsatellites d'intérêt, une amplification a été réalisée.

Les échantillons ont été amplifiés via des réactions en chaîne par polymérase ou PCR avant analyse. Ces dernières ont été réalisées dans des plaques de 96 puits de 200 µl à l'aide de l'appareil ABI Applied Biosystems 9902 Veriti PCR Thermal cycler.

Les amorces utilisées nous ont toutes été envoyées par Sylvain Ursenbacher. Certaines ont spécialement été développées pour *V. berus* et commencent par le sigle VB tandis que d'autres ont été développées pour *V. aspic* et possèdent le sigle Va. Du fait de la proche phylogénie de ces deux espèces, les deux types d'amorces donnent des réponses chez *V. berus*. Le protocole et les répartitions des amorces dans les différents multiplex nous ont également été envoyés par Sylvain Ursenbacher, mais ont toutefois subi quelques modifications comme vous pourrez le constater au point 3.1.

Tableau 3 : Caractéristiques des amorces utilisées d'après (Ursenbacher, et al., 2009) , (Carlsson, et al., 2002) et (Geser, et al., 2013)

Amorces	Séquences	Motif répété	Longueurs attendues des allèles	Couleurs	Température théorique d'hybridation [°c]
VB A11	F: 5'-GGTCAGTAACATCAGTCTGC-3' R: 5'-AGTCTTCCCTTACTTTGGCG-3'	CA	136-150	Rouge	58
VB B'10	F: 5'GCAAATTATTTTTGGGGTAGG-3' R: 5'-GAGATAGAGAACAAGTAGAGG-3'	GA	256-290	Vert	58
VB b'9	F: 5'-GATGGCTTTGAGTGGCACAG-3' R: 5'-CACAGAAATAGGCTGACTCC-3'	GA	260-349	Bleu	58
VB D17	F: 5'-TTTCTGCCCCATTTTACGAC-3' R: 5'-TGTAAGATGTTCCGAGTAGC-3'	AAG	135-268	Jaune	56
VB A8	F: 5'-ATTTACCATGCCTCCAGAA-3' R: 5'-GGTACACTCATTGTGATGAAC-3'	CA	186-240	Bleu	55
VB b'2	F: 5'-CTGAACAGAACAGGAGGAG-3' R: 5'-GGAAAACGAAGCAGTCAGGC-3'	GA	183-380	Rouge	58
VB b10	F: 5'-CGTGAGGTGTGTAATGAAG-3' R: 5'-CTATTTGAATCCCACAGTG-3'	GA	144-263	Jaune	55
VB B18	F: 5'-TGCTGGTGGGAATCACAATG-3' R: 5'-CACAAAAGCTGCCTGCAAAG-3'	GA	74-162	Bleu	55

VB D'10	F: 5'-GTCCTCCTTATCATCTATCC-3' R: 5'-CCTGGGTGCTCTCTCAG-3'	AAG	209-348	Jaune	58
VB 3	F: 5'-CAAGAAATGGAGATGAGC-3' R: 5'-GAAACCTATGAGCCAGTA-3'	AC	80-160	Jaune	52
VB 64	F: 5'-AGGCTCTGCTAAATGACC-3' R: 5'-GATCCCCTGAATTGATTA-3'	TG	220-260	Bleu	56
Va8	F: 5'-CCCTGATCTCCCTTGTTAATGC-3' R: 5'-CTGGACAGCCACTTGTCTG-3'	CAAT	185-205	Rouge	56
Va29	F: 5'-CACATTGATGATCCACAC-3' R: 5'-GGGAGTTTGGTCAGTTGACAG-3'	AGT	152-165	Jaune	52
Va51	F: 5'-GTAAGCATCTTCGCCAGCC-3' R: 5'-ACCAGTATCTCTGCCACGG-3'	GT	237-270	Vert	57
Va25	F: 5'-GACCAGTTTCCAAACCAGGG-3' R: 5'-CTGTGCCTCTCTATCCTTG-3'	CA	233-256	Jaune	54
Va20	F: 5'-GGATCTTCTCTAAGACAGAGC-3' R: 5'-GGAGTGCTTGTAACCTTCCC-3'	TG	191-204	Vert	54
Va69	F: 5'-GGTATAGGGTTTACTGCC-3' R: 5'-CGTTCTGAGGAAGATAGC-3'	CAA	115-159	Rouge	50
Va35	F: 5'-GCTTCATCATTGGTTGAGC-3' R: 5'-GCTGTTGGAATGTACCGCC-3'	CAA	174-196	Rouge	57
Va26	F: 5'-CACTAACCTAATGGCACAC-3' R: 5'-CTGAACACTCAAATTGACTGGC-3'	AAT	198-239	Jaune	52

Ces amorces sont en réalité des couples d'amorces dont l'une des deux est couplée à un fluorochrome, permettant ainsi l'identification du microsatellite obtenu sur l'électrophorégramme après électrophorèse. Les amplifications ont été faites dans des volumes finaux de 5 µl contenant 0.5 µl d'amorces de concentration variable en fonction du multiplex utilisé (Tableau 6), 2.5 µl de master mix et 2 µl d'extrait d'ADN. La solution master mix provient du kit Qiagen® Multiplex PCR.

Toutes les PCR ont débuté par 1 min de dénaturation à 95°C avant de poursuivre par 48 cycles divisés en 3 étapes de 45 secondes chacune que sont : la dénaturation de 30 s à 94°C, l'hybridation des amorces d'1 min 30 s à 56°C et l'élongation d'1 min à 72°C.

### 2.2.3. Electrophorèse

Les produits d'amplification sont séparés par électrophorèse sur un séquenceur ABI Prism<sup>®</sup>3130 xl Genetic Analyser produit par Applied Biosystems. Ce dernier permet la migration des microsatellites à travers des capillaires de 30 cm de longueur, chargés d'un polymère POP-7™. Il analyse 16 échantillons à la fois avec une précision au nucléotide près sur des microsatellites pouvant atteindre 500 paires de bases.

Avant leur introduction dans le séquenceur, utilisé ici comme analyseur de fragments, les produits d'amplification sont dénaturés. Pour se faire, on introduit 10 µl d'un marqueur de poids moléculaire nommé 600Liz (50 µl de 600 liz dans 1 ml de formamide) dans les puits d'une nouvelle plaque de 96 puits de 200 µl. A ces 10 µl on ajoute 1,5 µl des produits d'amplification avant de mettre la plaque dans la machine Veriti PCR Thermal cycler. Cette plaque est chauffée à 95°C pendant 5 min puis refroidie 5 min à 4°C avant d'être introduite directement dans le séquenceur.

Dans le séquenceur ou analyseur de fragments, les microsatellites migrent dans les capillaires en fonction de leur de taille, les fragments les plus petits migrant plus rapidement. A leur sortie des capillaires, ils sont détectés grâce aux fluorochromes auxquels ils sont associés. En effet, un laser va venir détecter les zones d'excitation des fluorochromes et émettre à une longueur donnée permettant de déterminer la taille du pic (autrement dit, celle du microsatellite qu'il représente) en comparant à la longueur obtenue pour un marqueur de poids moléculaire.

### 2.2.4. Logiciel de lecture

Les longueurs sont directement calculées sur *Genemapper* à partir du poids moléculaire et ce via les résultats obtenus de l'analyseur de fragments. *Genemapper* est un logiciel qui détecte les pics des réponses spectrales correspondant à des allèles. Pour chaque pic au-dessus de la limite de détection et dans la gamme de longueurs attribuées au microsatellite considéré il retourne une valeur. Cette dernière correspond à la longueur, en paires de bases, de l'allèle

détecté et constitue ce qu'on appelle un électrophorégramme. Cet électrophorégramme est analysé par l'opérateur qui une fois les résultats validés les introduits dans une matrice bilan.

## 2.3. Logiciels utilisés

### 2.3.1. GENETIX

Le Logiciel GENETIX version 4.04 permet de mesurer les hétérozygoties observée ( $H_o$ ), attendue sous l'équilibre d'Hardy-Weinberg ( $H_E$ ) et non biaisée ( $H_{nb}$ ). La dernière étant une correction de  $H_E$  pour le biais d'échantillonnage.

Le logiciel permet aussi la réalisation d'analyses factorielles de correspondances ou AFC. Pour ce faire, le logiciel positionne les individus dans un hyperespace qui a autant de dimensions qu'il y a d'allèles différents dans l'ensemble des données (Belkhir, Borsa, Chikhi, Raufaste, & Bonhomme, 2004). Par le nombre élevé de dimensions que cela représente, cet hyperespace n'est pas représentable graphiquement. Les individus sont cependant représentés selon les axes orthogonaux qui maximisent la somme des carrés de la distance entre chaque point et le centre de l'hyperespace (Belkhir, et al., 2004). Autrement dit, le logiciel représente les individus selon les deux axes (dans le cas d'une AFC 2D) qui expliquent la plus grande variation entre les individus. De ce fait, cette méthode n'utilise, ou du moins ne représente, qu'une partie de l'information génétique disponible.

### 2.3.2. FSTAT

Les distances obtenues sur GENETIX et converties en format FSTAT ont été utilisées sur le logiciel FSTAT 2.9.3. En effet, FSTAT est un logiciel qui estime et teste les diversités génétiques et les statistiques de différenciation à partir de marqueurs génétiques codominants. Il calcule les estimateurs de diversité génétique de Nei (Nei, 1987) et de Weir & Cockerham (Weir & Cockerham, 1984) aussi appelées statistiques F, et les teste en utilisant des méthodes de randomisation.

Il permet également de déterminer s'il y a présence ou non de déséquilibre de liaison dans les données qu'on lui introduit.

### 2.3.3. MEGA

Les distances génétiques entre les populations obtenues via FSTAT peuvent être utilisées sur MEGA 10.2.5 (pour Molecular Evolutionary Genetics Analysis) afin de les représenter via des dendrogrammes. De nombreuses méthodes existent pour construire un arbre à partir d'une matrice de distances. La méthode qui a été retenue ici est celle des Neighbour-Joining. Cette méthode tient compte du biais des différences de vitesse d'évolution entre les différentes branches de l'arbre phylogénétique à reconstruire et fournit un arbre non enraciné et non ultramétrique (Saitou & Nei, 1987). Elle consiste à regrouper les unités les plus proches tout en minimisant la longueur totale de l'arbre. Ce dernier ne reflète donc pas la similarité globale entre les différentes espèces, mais bien leurs relations de parenté (Saitou & Nei, 1987).

### 2.3.4. Bottleneck

Le programme Bottleneck 1.2.02 calcule, pour chaque échantillon de population et pour chaque locus, la distribution de l'hétérozygotie attendue à partir du nombre observé d'allèles  $k$ , étant donné la taille  $n$  de l'échantillon, sous l'hypothèse d'un équilibre mutation-dérive. Cette distribution est obtenue en simulant le processus de coalescence de  $n$  gènes sous trois modèles de mutation possibles : le modèle d'allèle infini ou « Infinite allele model (IAM) », le modèle de mutation par étapes ou « stepwise mutation model (SMM) » et le modèle à deux phases ou « two phase model (TPM) » (Piry, Luikar, & Cornuet). En fonction des résultats obtenus, il est alors possible de déterminer si les populations analysées ont connu un goulot d'étranglement ou « bottleneck » récent (dans les derniers  $2N_e-4N_e$  avec  $N_e$  taille effective de la population) ou non (Piry, Luikar, & Cornuet).

Le premier modèle, IAM, est un modèle qui tente de déterminer la quantité de loci homozygotes dans une population diploïde finie. Ce modèle présuppose qu'il y a suffisamment d'allèles pour que toute mutation conduise à un nouvel allèle, mais aussi que les mutations entraînent des résultats allant de neutres à délétères (Kimura & Crow, 1964).

Le second modèle, SMM, permet d'étudier la distribution d'équilibre des fréquences alléliques dans une population finie où les allèles neutres sont produits de manière progressive. Le SMM se distingue du modèle de Kimura-Crow, par le fait que lorsque la taille de la population augmente jusqu'à l'infini, la taille effective de la population ( $N_e$ ) et le taux de mutation restent fixes, le nombre moyen d'allèles différents dans la population atteint alors

rapidement un pic et se stabilise, cette valeur étant alors presque identique au nombre effectif d'allèles (Kimura & Ohta, 1978).

Tout comme le modèle SMM, le modèle TPM suppose qu'un microsatellite mute à un taux constant sans biais de contraction (perte de répétitions) ou d'expansion (gain de répétitions) quelle que soit leur longueur. La différence entre les deux est que SMM ne considère que des mutations liées à la perte ou au gain d'une seule unité de répétition tandis que TPM accepte différentes variations (Sainudiin, et al., 2004). En effet, ce dernier aborde le problème en autorisant des mutations de 1 unité de répétition (dites à une phase) avec une probabilité  $p$  et des mutations  $\geq 1$  unité(s) (dites à deux phases) avec une probabilité  $1 - p$ , tandis que la distribution des longueurs des mutations à plusieurs unités est géométrique. Dans le cas à une phase, l'allèle descendant a donc une chance égale d'être plus grand ou plus petit d'une unité de répétition que son ancêtre (Di Rienzo, et al., 1993).

### 2.3.5. STRUCTURE

Le logiciel STRUCTURE 2.3.1 fonctionne en considérant successivement plusieurs modèles avec un nombre de sous-populations ( $K$ ) différent. Les individus sont assignés aux sous-populations sur base de leur génotype tout en estimant simultanément les fréquences alléliques au sein de ces sous-populations. Son but étant de tenter d'assigner les individus à leur population d'origine.

Pour cela il se base sur une série d'hypothèses qui sont les suivantes (Pritchard, et al., 2000) :

- Les populations suivent l'équilibre d'Hardy-Weinberg
- Les marqueurs génétiques utilisés pour mettre en évidence les allèles ne doivent pas être liés ; quand la localisation des loci est inconnue, on considère par défaut que cette hypothèse est vraie.
- La distribution des allèles à chaque locus est indépendante
- Les populations sont bien distinctes, il n'y a pas d'isolement par la distance (dans un cas idéal).

Deux cas sont possibles lors de l'analyse via STRUCTURE. Le premier étant de considérer qu'il n'y a pas d'échanges entre les populations. Dans ce cas, chaque individu est supposé

n'appartenir qu'à une seule des K populations. Le second cas considère au contraire un possible échange entre populations.

### 2.3.6. CERVUS

Cervus 3.0.7 est un programme conçu pour l'analyse de parenté à grande échelle, qui a été utilisé pour réaliser les assignations parentales des juvéniles de l'élevage.

Il génère les fréquences d'allèles nécessaires aux modules de simulation et d'analyse de la filiation, et calcule diverses statistiques récapitulatives afin de pouvoir évaluer la pertinence des loci pour l'analyse de la filiation (Marshall, 2001). Pour se faire, il lit un fichier texte contenant les génotypes à un ou plusieurs loci, compte le nombre d'occurrences de chaque allèle à chaque locus et, à l'aide de ces données, calcule les diverses statistiques sur chaque locus. Pour chaque juvénile un score  $\Delta$  représentant la différence de scores LOD obtenus par les deux parents candidats les plus probables est calculé, où la LOD ou limite de détection représente la probabilité de filiation du vrai parent. Après quoi, Cervus calcule, pour chaque juvénile testé, les scores LOD pour chacun des parents candidats et le score  $\Delta$  correspondant (Marshall, 2001).

En comparant la distribution des scores  $\Delta$  pour les descendants pour lesquels le parent candidat le plus probable était le vrai parent avec la distribution des scores  $\Delta$  pour les descendants pour lesquels le parent candidat le plus probable était un individu non apparenté, on peut trouver un score critique suffisamment important pour distinguer les vrais parents des parents candidats non apparentés à un certain niveau de confiance prédéterminé.

## 3. Résultats & discussion

### 3.1. Mises au point technique

#### 3.1.1. Conservation des écouvillons et extraction

Les premières extractions, réalisées sur les échantillons de l'élevage, ayant donné des résultats peu habituels pour le laboratoire en termes de concentration (de l'ordre de 2-3 ng/ $\mu$ l contre 15-20 ng/ $\mu$ l habituellement), de nouveaux échantillons ont été demandés au responsable de l'élevage, Monsieur Bufkens. A l'époque, les seuls individus qu'il pouvait rééchantillonner étaient les juvéniles. En effet, ce second échantillonnage a eu lieu pendant la période d'hibernation de l'espèce, ce qui rendait les individus inaccessibles pour un prélèvement. Les juvéniles étant eux sous hibernation artificielle, la démarche de rééchantillonnage était plus abordable.

Les concentrations obtenues avec ces nouveaux échantillons ont été multipliées par 11, en moyenne, par rapport à celles obtenues initialement. Ceci nous a poussés à transférer au congélateur, tous les écouvillons initialement conservés au réfrigérateur, en pensant que ces différences dans les quantités d'ADN observées étaient dues à l'attente dans le réfrigérateur et qu'il était donc préférable de les conserver au congélateur avant analyse. Cette hypothèse ne s'est cependant pas totalement confirmée et n'est peut-être pas applicable aux autres écouvillons car d'autres facteurs pourraient expliquer cette observation.

Le premier étant que lors du second prélèvement, les juvéniles étaient en hibernation programmée et donc moins mobiles, ce qui peut avoir facilité le prélèvement et donc la quantité d'ADN présente sur les écouvillons. Le second facteur pouvant être que les écouvillons n'aient pas été correctement utilisés. En effet, la brochure *Eswab liquid amies* de Copan nous dit que « ce système innovant élue plus de 90 % de l'échantillon du patient dans le milieu liquide » or ce milieu liquide a été retiré des écouvillons que nous avons reçus. Ces deux observations ne permettent donc pas d'étendre nos conclusions aux écouvillons *Sarstedt*. Le troisième facteur plausible, est que notre technique d'extraction se soit améliorée entre les premiers prélèvements et les nouveaux. La dernière possibilité est que l'erreur provienne de notre application du protocole de Fanny Demez à la lettre. En effet, une étape du protocole établi par Qiagen a été négligée par Fanny (ou du moins non rapportée dans son mémoire) et pourrait également expliquer ces observations. Cette étape consiste à

vérifier que le pH de la solution soit bien inférieur à 7 après ajout du tampon AL pour maximiser l'adhérence des molécules d'ADN à la membrane de la colonne. Ce dernier facteur ayant été mis à jour après la fin de toutes les extractions, il n'a pas été possible d'en vérifier l'impact.

Lors de l'assignation parentale Monsieur Bufkens a de nouveau été sollicité afin de nous renvoyer des prélèvements des individus mâles, ceux-ci ayant donné trop peu de résultats pour réaliser une assignation fiable. Ceci nous a permis de constater que notre première conclusion n'était peut-être pas la bonne et, de plus, considérer les deux premiers facteurs cités ou le dernier. En effet, lors de ce second prélèvement, les mâles étaient sortis d'hibernation et les échantillons ont été extraits dans les mêmes délais que les deuxièmes échantillons des juvéniles. Pourtant, alors qu'on observait une augmentation importante dans les quantités obtenues pour les juvéniles entre les deux extractions, ce n'est pas le cas pour les mâles, comme vous pourrez le constater dans le Tableau 4. Avec 0.83 fois moins d'ADN en moyenne (si on retire l'échantillon VPM4), nous avons obtenu de bien meilleurs résultats que précédemment, indiquant tout de même une meilleure qualité.

*Tableau 4 : Tableau des concentrations des extraits d'ADN des juvéniles et des mâles de l'élevage au premier et au second prélèvement.*

	Échantillons	Concentration [ng/μl]		
		1er prélèvement	2ème prélèvement	2ème/1er
Juvéniles	VBM8	2,47	29,6	11,98
	VBM9	2,71	37,1	13,69
	VBM10	3,07	34,9	11,37
	VBF18	3,62	39,20	10,84
	VBF19	4,09	37,70	9,22
	VBF20	3,06	28,20	9,21
Mâles	VPM1	2,63	1,34	0,51
	VPM2	3	3,90	1,30
	VPM3	1,4	1,98	1,41
	VPM4	2,54	301 <sup>1</sup>	118,50

<sup>1</sup> Cette concentration élevée est attribué au fait que c'est un échantillon de tissu que nous avons reçu dans ce cas-ci plutôt qu'un écouvillon car l'individu est décédé entre les deux prélèvements.

	VPM5	6,12	1,92	0,31
	VPM6	4,57	1,45	0,32
	VPM7	3,28	3,67	1,12

### 3.1.2. Mise au point de l'amplification : Primers et multiplex

#### 3.1.2.1. Formation des multiplex

Comme rapporté dans la partie Matériel et Méthodes, nous avons utilisé les amorces qui nous ont été envoyées par Sylvain Ursenbacher. Ce dernier utilisait ces amorces en 7 multiplex différents repris au Tableau 5. Il a donc d'abord été question de tester les amorces individuellement pour en contrôler la réponse avant de les combiner en d'autres multiplex pour en minimiser le nombre puisque 4 amorces, Va16, Va91, Va81 et Va70, étaient manquantes. Il est à noter que lorsqu'on parle par exemple de l'amorce Va 8, il s'agit en fait d'un couple d'amorces qui permettent l'amplification du marqueur microsatellite.

Tableau 5 : Tableau contenant les différents multiplex établis par Sylvain Ursenbacher

Multiplex	Primers	Longueurs attendues	Couleurs	Température Théorique [°c]	Température de travail [°c]
Mix 1	Vb A11	136-150	Rouge	58	56
Mix 1	Vb B'10	256-290	Vert	58	56
Mix 1	VB b'9	287-349	Bleu	58	56
Mix 1	VB D17	135-268	Jaune	56	56
Mix 2	VB A8	186-240	Bleu	55	56
Mix 2	VB b'2	183-380	Rouge	58	56
Mix 2	VB b10	144-263	Jaune	55	56
Mix 2	VB B18	74-162	Bleu	55	56
<del>Carlsson</del>	VB D'10	209-348	Jaune	58	55
Carlsson	VB 3	80-160	Jaune	52	55
Carlsson	VB 64	220-260	Bleu	56	55
Mix 3	Va8	185-205	Rouge	56	56
Mix 3	Va29	152-165	Jaune	52	56
Mix 3	Va51	237-270	Vert	57	56
Mix 4	Va25	233-256	Jaune	54	56

Mix 4	Va20	191-204	Vert	54	56
Mix 4	Va 16		-	-	-
Mix 5	Va69	115-159	Rouge	50	56
Mix 5	Va35	174-196	Rouge	57	56
Mix 5	Va91		-	-	-
Mix 6	Va81		-	-	-
Mix 6	Va26	198-239	Jaune	52	56
Mix 6	Va 70		-	-	-

Ces tests préliminaires nous ont permis de constater que les amorces Va29 et VB 3 ne donnaient aucun résultat et d'ainsi les supprimer des multiplex ou mix établis. Ces constatations, en plus des marqueurs manquants, ne laissaient que 17 marqueurs possiblement utilisables à répartir en différents multiplex. Les multiplex de Sylvain Ursenbacher ont été conservés pour les 8 marqueurs (ceux contenus dans les mix 1 et 2) avec lesquels les résultats de Fanny Demez ont été établis, pour ainsi minimiser les facteurs influençant les résultats obtenus. Les autres multiplex ont été créés en fonction des fluorochromes portés par les amorces, mais aussi en fonction des longueurs attendues des microsatellites, de manière à ce que deux microsatellites avec le même fluorochrome ne se chevauchent pas pour permettre la lecture de ces derniers sans biais. Les températures théoriques d'hybridation des amorces ont aussi été facteur de discussion pour la formation des multiplex, mais après avoir constaté que toutes les amorces donnaient des résultats satisfaisants à 56°C la tâche s'est avérée plus simple et nous a permis de former les mix supplémentaires 3 et 4 repris dans le Tableau 6 avec les mix 1 et 2.

Une fois les multiplex formés, il a été question de déterminer les concentrations optimales des différentes amorces dans chacun d'entre eux (celles-ci sont reprises dans le Tableau 6). Pour se faire, des analyses ont été faites sous forme de multiplex avec un extrait d'ADN et comparés aux résultats obtenus avec le même extrait en simplex. En comparant les intensités des pics obtenus, les concentrations des différentes amorces ont ensuite été réajustées. Ainsi celles qui voient l'intensité de leurs résultats diminuer en multiplex par rapport à ce qu'elles donnaient en simplex voient leur concentration dans le multiplex augmenter tandis que celles pour qui l'intensité augmente en multiplex par rapport au simplex voient leur concentration

diminuer. Ces multiplex ont tous subi 48 cycles d'amplification pour compenser les faibles concentrations d'ADN obtenues lors des extractions. Ces cycles pouvant être divisés en trois étapes que sont la dénaturation, l'hybridation et l'élongation, respectivement menés à 94°C, 56°C et 72°C, et ce pour tous les mix.

Tableau 6 : Tableau contenant les différents multiplex après réajustement

Multiplex	Primers	Motif répété	Longueurs attendues	Couleurs	Température théorique [°c]	Température de travail [°c]	Concentration dans le Mix
Mix 1	Vb A11	CA	136-150	Rouge	58	56	1 µM
Mix 1	Vb B'10	GA	256-290	Vert	58	56	1 µM
Mix 1	VB b'9	GA	287-349	Bleu	58	56	2 µM
Mix 1	VB D17	AAG	135-268	Jaune	56	56	1 µM
Mix 2	VB A8	CA	186-240	Bleu	55	56	1 µM
Mix 2	VB b'2	GA	183-380	Rouge	58	56	1 µM
Mix 2	VB b10	GA	190-263	Jaune	55	56	1 µM
Mix 2	VB B18	GA	74-162	Bleu	55	56	1 µM
Mix 3	Va8	CAAT	185-205	Rouge	56	56	0,6 µM
Mix 3	Va26	AAT	198-239	Jaune	52	56	1 µM
Mix 3	Va51	GT	237-270	Vert	57	56	0,5 µM
Mix 3	VB 64	TG	220-260	Bleu	56	56	0,9 µM
Mix 4	Va25	CA	233-256	Jaune	54	56	1 µM
Mix 4	Va20	TG	191-204	Vert	54	56	1,2 µM
Mix 4	VB D'10	AAG	209-348	Jaune	58	56	0,5 µM
Mix 4	Va69	CAA	115-159	Rouge	50	56	0,5 µM
Mix 4	Va35	CAA	174-196	Rouge	57	56	0,5 µM

### 3.1.2.2. Amorces récalcitrantes

Dans les multiplex 1 et 2 certaines amorces se sont révélées difficilement interprétables pour différentes raisons.

#### a) Cas de VB\_Bp9

Le premier cas a été celui de VB\_b'9 aussi noté VB\_Bp9. Le problème avec cette amorce est lié au nombre de pics trop élevé obtenus sur l'électrophorégramme. En réalité cette amorce

peut être lue de deux manières différentes. Nous voulons dire par là qu'elle couvre deux loci différents qu'on va interpréter comme deux marqueurs avec des longueurs attendues différentes par souci de compréhension. Le premier VB\_Bp9a présente des longueurs comprises entre 260 et 285 pdb tandis que celles de VB\_Bp9b peuvent aller de 287 à 349 pdb. Une illustration de l'électrophorégramme obtenu avec ces deux marqueurs est reprise à la Figure 7. Le marqueur VB\_Bp9b donnant que trop peu souvent des résultats lisibles, VB\_Bp9a lui a été préféré. Cependant, ce marqueur forme régulièrement plus de 2 pics alors qu'il n'est pas possible d'avoir plus de 2 allèles différents pour un même loci. Un changement des conditions de PCR (avec et sans Q, simplex ou multiplex) n'a eu aucun effet sur les résultats obtenus.

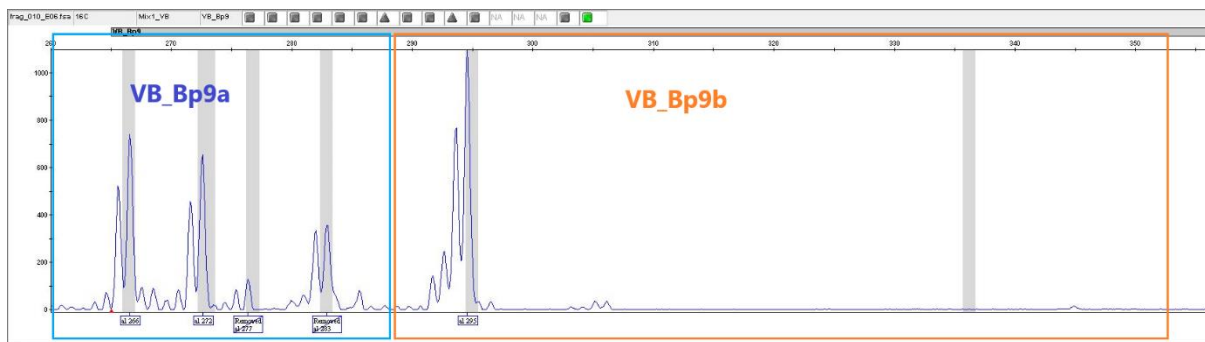


Figure 7 : Electrophorégramme de l'individu 15C pour le marqueur VB\_Bp9

C'est finalement en comparant les résultats obtenus dans notre laboratoire à ceux obtenus dans celui de Sylvain Ursenbacher pour les ADN de référence qu'une interprétation a pu être établie.

#### b) Cas de VB\_A8 et de VB\_B10

Le cas de VB\_A8 et de VB\_B10 est intéressant car il aurait été plus problématique si le problème en question n'avait pas été mis en lumière. En effet, généralement un électrophorégramme ayant l'aspect de celui représenté à la Figure 8 A est interprété comme hétérozygote. Cependant, en le comparant à l'électrophorégramme d'un hétérozygote ayant deux groupes de pics bien distincts (Figure 8 b), on se rend compte que la première interprétation est fautive. Le marqueur produit ce que Mme Flamand qualifie de "pics en crête de stégosaure". L'individu auquel appartient l'électrophorégramme Figure 8 A a donc été considéré homozygote pour ce marqueur.

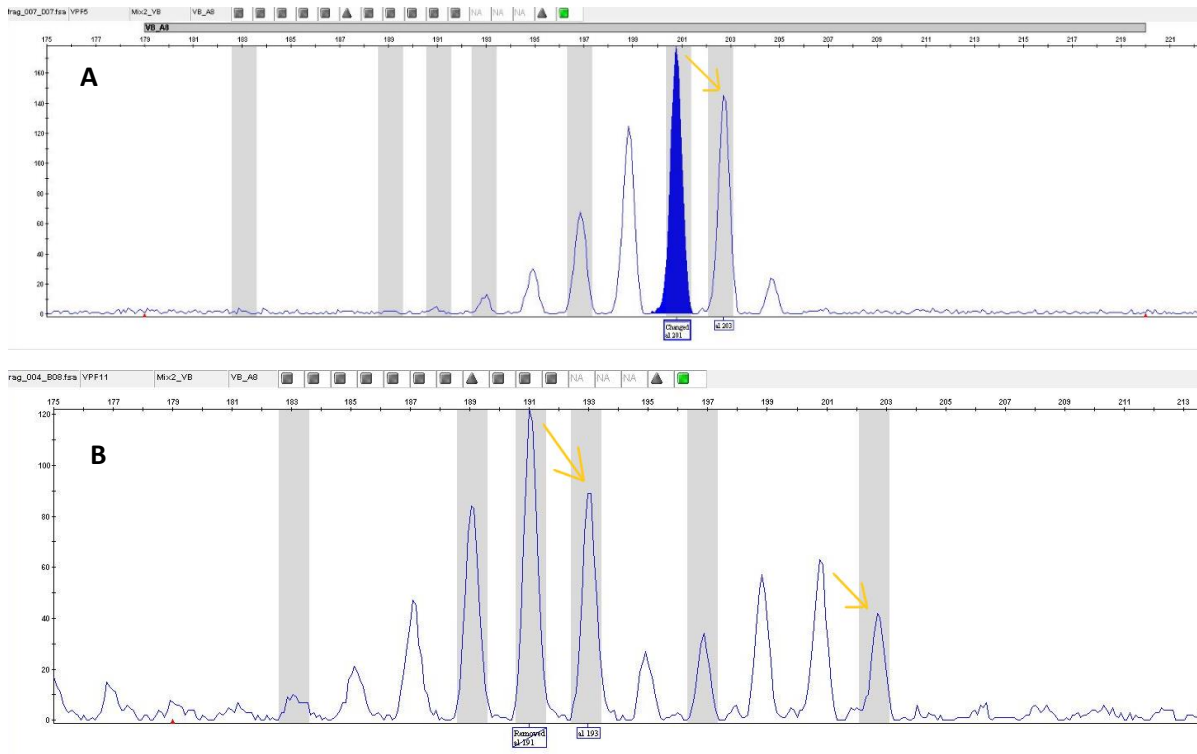


Figure 8 : Electrophorégrammes obtenus avec le marqueur VB\_A8 pour les individus (A)VPF5 et (B) VPF11

Ainsi, si le problème n'avait pas été détecté, nous aurions sûrement retrouvé un excès d'hétérozygotes. Cependant, par la même, il est possible qu'il y ait des allèles nuls. En effet, s'il y a vraiment un individu hétérozygote avec ses deux pics dans le même groupe de pic (comme dans le cas A de la Figure 8), il sera interprété homozygote car il sera impossible de le différencier d'un cas de "queue de stégosaure".

### c) Cas de VB\_B18

Le problème avec cette amorce était, tout comme pour VB\_Bp9, lié au nombre de pics obtenus sur l'électrophorégramme. En effet, un trop grand nombre de pics d'intensités relativement égales pour certains rendaient difficile l'interprétation des résultats. Tout comme pour VB\_Bp9 les changements de conditions d'amplification n'ont rien donné. Sylvain Ursenbacher ayant, de ses dires, souvent eu à retirer les résultats obtenus avec ce marqueur à la suite de résultats peu concluants au test d'Hardy-Weinberg, possiblement en raison de la lecture. C'est pourquoi une autre méthode a été entreprise : celle de l'addition non-programmée de nucléotides ou « pig-tail ».

La pig-tail consiste en l'ajout d'une séquence en 5' de l'amorce reverse du marqueur. Cette manipulation a pour but de réduire les erreurs de génotypage liées à la variabilité dans

l'addition d'adénosine (Brownstein, Carpten, & Smith, 1995). Autrement dit elle a pour but de diminuer le nombre de pics obtenus lors du génotypage avec un marqueur donné.

Après ajout de la dite pig tail, nous avons constaté un affinement des pics et une diminution des contaminations lors des PCR simplex (Figure 9 A et B). De plus, de nombreux pics parasites apparaissent en utilisant l'amorce additionnée de la pig tail en multiplex (Figure 9 C). Le choix a donc été fait de ne réaliser que des PCR simplex pour l'analyse de ce marqueur.

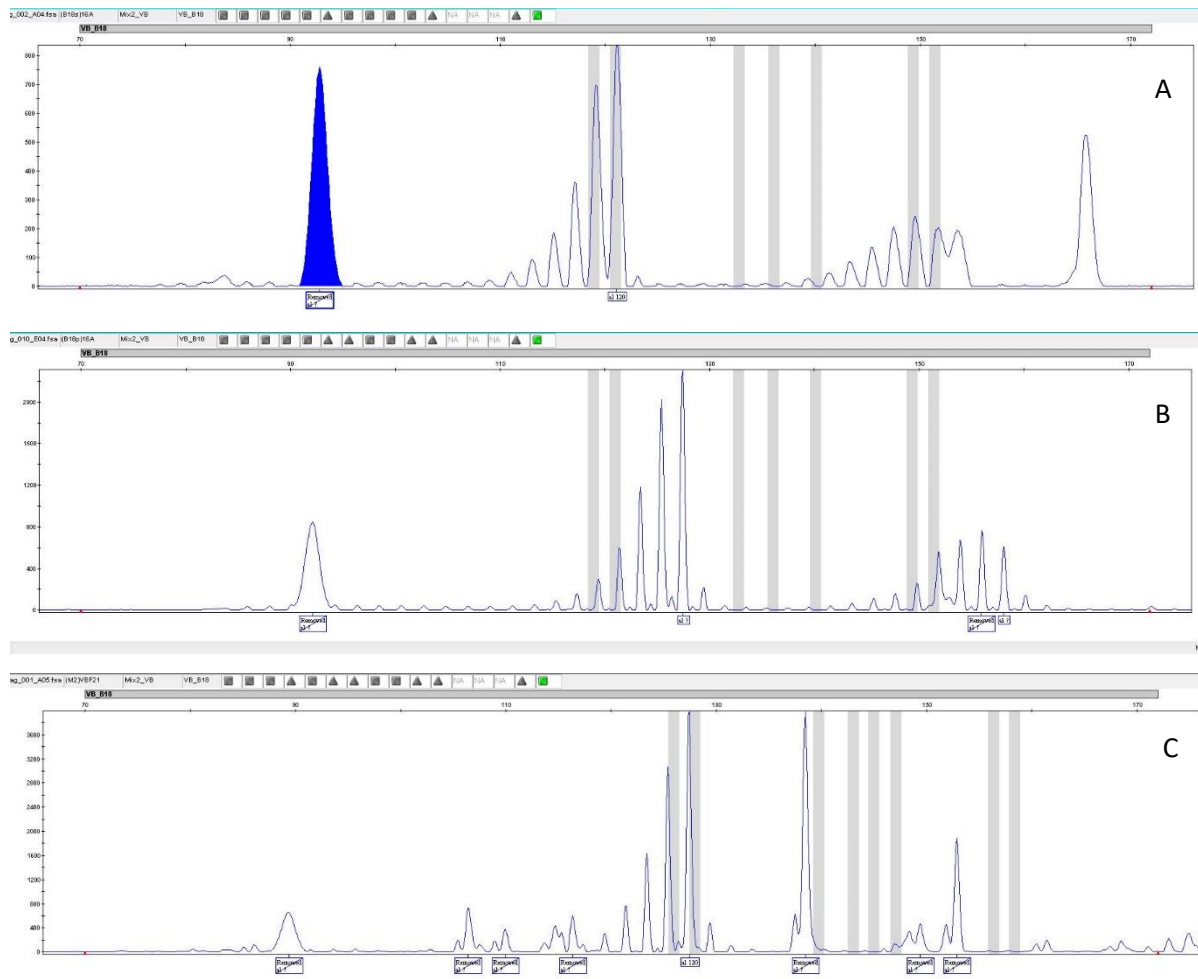


Figure 9 : Electrophorégrammes obtenus avec le marqueur VB\_B18 (A) sans pig tail ou (B) avec pig tail en PCR simplex pour l'individu 16A et (C) avec pig tail en PCR multiplex pour l'individu VBF21

### 3.1.3. Calibration

Comme précisé plus tôt, une calibration entre notre laboratoire et celui de Sylvain Ursenbacher a été nécessaire pour que nous puissions utiliser les résultats de Fanny Demez. Cette calibration a d'abord nécessité l'analyse des 20 échantillons de référence avant comparaison des résultats obtenus avec ceux de Sylvain Ursenbacher.

Cette comparaison a permis d'associer les bins (zones, en nucléotides, dans lesquelles au moins un pic a été obtenu pour un marqueur donné) obtenus dans notre laboratoire aux bins équivalents dans le laboratoire de Sylvain Ursenbacher. Ces bins permettent, après leur établissement, au logiciel Genemapper d'attribuer la longueur correspondante aux pics qui atteignent leurs zones.

Il y a cependant une limite à cette calibration. En effet, du fait du nombre restreint d'échantillons de référence en notre possession, seule une partie infime des bins a pu être couverte. Ainsi certains bins que nous avons obtenus pour d'autres individus que ceux de référence n'ont aucun équivalent chez Sylvain Ursenbacher et inversement avec certains des bins obtenus par Fanny Demez dans son mémoire.

Cette calibration nous a tout de même permis d'ajouter les résultats de 43 des 272 individus analysés par Fanny Demez lors de son mémoire. Ces individus appartiennent aux populations que l'on va nommer BE\_01, 04, 11, 17 et 19. La localisation de ces populations belges est reprise sur la carte représentée sur la Figure 10. Les 229 individus manquants n'ont pas pu être ajoutés par manque de correspondance entre les longueurs obtenues par Fanny Demez et celles de notre laboratoire ou dû à un nombre trop élevé de données manquantes pour que l'ajout soit pertinent.

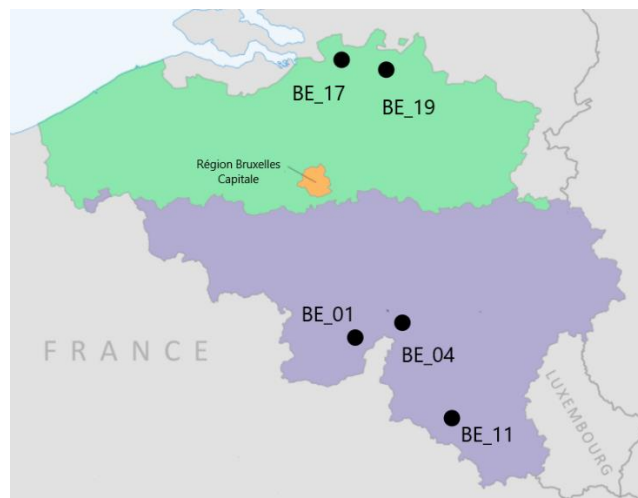


Figure 10 : Carte représentant la localisation des populations dont proviennent les individus analysés par Fanny Demez, ajoutés à notre étude.

## 3.2. Analyses statistiques

Il est à noter que les individus de référence utilisés pour la calibration ont également été conservés dans le pool de résultats pour les analyses statistiques. Ils forment trois populations nommées Cantal, Lozère et Ardennes. La dernière population contient des individus appartenant au parc naturel régional des Ardennes à la frontière belgo-française ou ayant des localisations proches de ce dernier.

### 3.2.1. Diversité génétique

La diversité génétique peut-être déterminée de différentes manières. On peut notamment citer le nombre d'allèles moyens par locus ou « allele ratio » (AR), le polymorphisme (P) ou encore le  $F_{IS}$ . Toutes ces données, en plus des hétérozygoties attendue ( $H_{exp}$ ), non biaisé ( $H_{nb}$ ) et observée ( $H_{obs}$ ), sont reprises dans le Tableau 7, le Tableau 8 et le Tableau 9. Ces derniers ont été respectivement établis avec les résultats obtenus pour les 87 individus analysés avec 17 marqueurs, pour les 28 individus des Hautes Fagnes avec 17 marqueurs et enfin pour les 87 individus analysés auxquels on a ajouté 43 des individus de Fanny Demez, le tout avec 8 marqueurs. Ces résultats ont été établis via le logiciel GENETIX version 4.04.

Il est à noter que l'AR est dépendant du nombre d'individus analysés par population. On ne peut donc comparer des AR entre eux que s'ils ont été établis pour un nombre  $n$  similaire d'individus ou si, à moindre individus analysés, on a AR aussi élevé que celui d'une population où le  $n$  est plus important. On peut ainsi dire en analysant les données obtenues dans le Tableau 7, que les populations françaises (Cantal et Lozère) semblent plus diversifiées génétiquement que celle des Ardennes. Il est aussi important de noter que l'élevage n'est pas une population à proprement parlé, puisqu'étant séparés dans différents enclos, la rencontre entre tous ces individus n'est pas équiprobable (De Meeûs, 2012).

Le second paramètre interprétable est le polymorphisme ( $p$ ). Une population est dite polymorphe pour un locus donné si la fréquence allélique de l'allèle le plus fréquent est inférieure à 0.95. La valeur de P reprise ici correspond au polymorphisme global de la population, donc sur tous les locus considérés. Un P de 0.7059 à une probabilité de 0.95, comme observé pour la population des Ardennes, peut ainsi être interprété comme le fait qu'au taux de 95%, significativement 70.59% des loci sont polymorphes. On constate ainsi

qu'au minimum, dans toutes les populations analysées, 70.59% des loci considérés sont polymorphes.

L'hétérozygotie attendue est déterminée via l'équation (10) tirée de la relation d'Hardy-Weinberg où  $p_i$  est la fréquence du  $i^{\text{ème}}$  allèle du gène (Serre, 2006).

$$H = 1 - \sum p_i^2 \quad (10)$$

Ainsi, si un locus est composé de 2 allèles équifréquents  $H_{\text{exp}} = 1 - (0.5^2 + 0.5^2) = 0.5$ . Ici ce sont bien les hétérozygoties moyennes (sur tous les locus) qui sont reprises. Une valeur de H élevée indique donc une variabilité allélique avec des fréquences correspondantes importantes.

Tableau 7 : Tableau contenant le nombre d'individus (n), les hétérozygoties attendues sous l'équilibre d'Hardy-Weinberg ( $H_{\text{exp}}$ ), sans biais ( $H_{\text{n.b}}$ ) et observée ( $H_{\text{obs}}$ ), le  $F_{\text{IS}}$  avec son degré de significativité et le polymorphisme (P) aux seuils de 95% et 99% et ce pour les 87 individus analysés avec les 17 marqueurs retenus.

	n	$H_{\text{exp}}$	$H_{\text{n.b}}$	$H_{\text{obs}}$	P(0.95)	P(0.99)	$F_{\text{IS}}$	AR
Ardennes	9	0.2802	0.3006	0.2655	0.7059	0.7059	0.1187 <sup>NS</sup>	1.94
Cantal	4	0.5195	0.6027	0.5938	0.8125	0.8125	0.0172 <sup>NS</sup>	3.25
Elevage	29	0.3705	0.3773	0.3193	0.7647	0.8824	0.1562 <sup>***</sup>	3.18
Hautes Fagnes	28	0.6842	0.6969	0.5558	1.0000	1.0000	0.2055 <sup>***</sup>	6.94
Kalmthout	12	0.5012	0.5238	0.4481	1.0000	1.0000	0.1511 <sup>***</sup>	3.71
Lozère	5	0.4880	0.5435	0.3588	0.8824	0.8824	0.3673 <sup>***</sup>	3.00

\*\* : significatif pour  $p < 0.05$ , \*\*\* : significatif pour  $p < 0.01$  et NS : non significatif

L'hétérozygotie attendue ou  $H_{\text{exp}}$  correspond donc à ce qu'on obtiendrait si la population respectait l'équilibre d'Hardy-Weinberg tandis que l'hétérozygotie observée ou  $H_{\text{obs}}$  fait référence aux vraies variabilités/fréquences constatées. Le  $H_{\text{n.b}}$  est quant à lui le  $H_{\text{exp}}$  corrigé pour le biais d'échantillonnage. Ainsi, pour 17 marqueurs, seule la population du Cantal a un  $H_{\text{obs}}$  supérieur au  $H_{\text{exp}}$ , indiquant un excès d'hétérozygotes. En prenant en compte le  $H_{\text{n.b}}$  plutôt que le  $H_{\text{exp}}$ , nous remarquons cependant que le  $H_{\text{obs}}$  lui est toujours inférieur.

Sachant qu'un  $F_{\text{IS}}$  supérieur à 0 est signe d'un déséquilibre d'Hardy-Weinberg lié à un déficit en hétérozygotes (cf 1.3.3), nous constatons que les valeurs obtenues pour l'élevage, les Hautes Fagnes, Kalmthout et la Lozère sont significativement supérieures (pour  $p = 0.99$ ) à 0. Si on se fie à (De Meeûs, 2012), la formule (11) donne une estimation approximative du taux de croisements consanguins nécessaires pour expliquer le  $F_{\text{IS}}$  obtenu.

$$b = \frac{4F_{IS}}{1+3F_{IS}} \quad (11)$$

En réalisant le calcul pour le  $F_{IS}$  le plus élevé (celui de Lozère), nous constatons qu'il faudrait 70% de croisements pour expliquer ce résultat. Ce chiffre est élevé, mais reste à relativiser puisqu'il a été obtenu pour un échantillon de population constitué de 5 individus uniquement. Il est aussi intéressant de noter qu'un  $F_{IS}$  positif peut aussi refléter la présence d'allèles nuls ou un mélange de populations dans l'échantillon considéré (effet Wahlund).

La population des Hautes Fagnes peut d'ailleurs être divisée en 3 sous-populations en fonction du lieu d'échantillonnage précis des individus. Ces 3 sous-populations sont Brackvenn, le Aachener Kreuz et autres, contenant des individus échantillonnés à Allegemeines Venn, Steinley Caillebotis et Hautes Fagnes. Les résultats obtenus pour les Hautes Fagnes ainsi fractionnées sont retrouvés dans le Tableau 8. Nous constatons que relativement au nombre d'individus analysés, la sous-population de Aachener Kreuz est plus diversifiée génétiquement que celle de Brackvenn. De plus, malgré les sous-divisions de la population des Hautes Fagnes, les valeurs de  $F_{IS}$  obtenues restent supérieures à 0, écartant l'effet Wahlund comme raison à ce résultat.

La présence d'allèles nuls pourrait aussi justifier les valeurs de  $F_{IS}$  obtenue. Le logiciel Microchecker a d'ailleurs déterminé que la sous-population de Brackvenn et la population de Kalmthout présentaient, pour plusieurs marqueurs, des signes d'allèles nuls. Ces allèles nuls pourraient donc expliquer en partie les résultats obtenus, mais pas totalement puisque, malgré une absence d'allèles nuls dans la population de Lozère, nous observons un  $F_{IS}$  conséquent.

*Tableau 8 : Tableau contenant le nombre d'individus (n), les hétérozygoties attendues sous l'équilibre d'Hardy-Weinberg ( $H_{exp}$ ), sans biais ( $H_{n.b.}$ ) et observée ( $H_{obs}$ ), le  $F_{IS}$  avec son degré de significativité et le polymorphisme (P) aux seuils de 95% et 99% et ce pour les 28 individus des Hautes Fagnes répartis en 3 sous-populations avec les 17 marqueurs retenus.*

	n	$H_{exp}$ .	$H_{n.b.}$	$H_{obs}$	P(0.95)	P(0.99)	$F_{IS}$	AR
Aachener Kreuz	9	0.6543	0.6928	0.6340	1.0000	1.0000	0.0897**	4.9412
Brackvenn	15	0.6518	0.6751	0.5209	1.0000	1.0000	0.235***	5.7647
Autres	4	0.5958	0.6889	0.5049	0.9412	0.9412	0.294***	3.6471

*\*\* : significatif pour  $p < 0.05$ , \*\*\* : significatif pour  $p < 0.01$  et NS : non significatif*

En faisant la même analyse avec 8 marqueurs plutôt que 17 et en ajoutant les populations analysées par Fanny Demez, nous constatons que la population BE\_01 montre un excès d'hétérozygotes, la valeur du  $F_{IS}$  étant négative, mais celui-ci est non significatif (Tableau 9).

Tableau 9 : Tableau contenant le nombre d'individus ( $n$ ), les hétérozygoties attendues sous l'équilibre d'Hardy-Weinberg ( $H_{exp}$ ), sans biais ( $H_{n.b.}$ ) et observée ( $H_{obs}$ ), le  $F_{IS}$  avec son degré de significativité et le polymorphisme ( $P$ ) aux seuils de 95% et 99% et ce pour les 87 individus analysés et les 43 individus de Mme Demez avec les 8 marqueurs des Mix 1 et 2.

	$n$	$H_{exp}$	$H_{n.b.}$	$H_{obs}$	$P(0.95)$	$P(0.99)$	$F_{IS}$	AR
BE_01	8	0.1973	0.2104	0.2344	0.5000	0.5000	-0.1230 <sup>NS</sup>	1.63
BE_04	9	0.1458	0.1544	0.1389	0.2500	0.2500	0.1062 <sup>NS</sup>	1.63
BE_11	6	0.1892	0.2064	0.1458	0.3750	0.3750	0.3137 <sup>**</sup>	1.50
BE_17	12	0.5461	0.5746	0.4684	1.0000	1.0000	0.1926 <sup>**</sup>	3.00
BE_19	8	0.3853	0.4337	0.4250	0.7500	0.7500	0.0005 <sup>NS</sup>	2.63
Ardennes	9	0.3856	0.4088	0.3559	1.0000	1.0000	0.1361 <sup>NS</sup>	2.25
Cantal	4	0.5352	0.6116	0.5938	0.8750	0.8750	0.0339 <sup>NS</sup>	3.63
Elevage	29	0.4783	0.4869	0.4309	0.8750	1.0000	0.117 <sup>**</sup>	3.88
Hautes Fagnes	28	0.7300	0.7436	0.6334	1.0000	1.0000	0.1505 <sup>***</sup>	8.38
Kalmthout	12	0.5750	0.6011	0.5052	1.0000	1.0000	0.167 <sup>***</sup>	4.63
Lozère	5	0.5325	0.5917	0.3250	1.0000	1.0000	0.480 <sup>***</sup>	3.13

\*\* : significatif pour  $p < 0.05$ , \*\*\* : significatif pour  $p < 0.01$  et NS : non significatif

En comparant les résultats obtenus pour 17 et 8 marqueurs, nous remarquons que les AR obtenus ont tendance à augmenter lorsque le nombre de marqueurs diminue. Cette observation nous indique qu'il y a moins de polymorphisme dans les 9 marqueurs de différence, initialement mis au point pour la vipère aspic.

Avec cette augmentation du AR, nous observons une augmentation du polymorphisme  $P$ . Cette observation est logique étant donné que AR et  $P$  sont positivement corrélés dans le sens où si le nombre d'allèles moyen par locus augmente c'est qu'il y a plus de chance que la fréquence allélique de l'allèle le plus fréquent du locus considéré diminue. Nous constatons également des variations dans les résultats obtenus pour les  $F_{IS}$ , mais, selon (De Meeûs, 2012), il serait plus correct de se fier aux  $F_{IS}$  obtenus pour 17 marqueurs car la fiabilité de ce dernier augmente avec le nombre de marqueurs utilisés.

### 3.2.2. Attribution parentale

Le logiciel Cervus nous a permis de déterminer les possibles parents des juvéniles de l'élevage. Avant cela une analyse des marqueurs a été nécessaire afin de savoir s'ils pouvaient tous être conservés. D'après le logiciel Microchecker, le seul marqueur comprenant un ou plusieurs allèles nuls dans l'élevage est le marqueur Va26. En retirant ce marqueur de l'analyse d'attribution parentale, nous obtenons les résultats repris dans le Tableau 10.

Nous garderons comme duo de parents, ceux ayant obtenu le LOD pour le trio Juvénile-Femelle-Mâle le plus élevé, LOD qui, pour rappel, représente la probabilité de filiation du vrai parent. Nous constatons que le LOD le plus élevé par trio ne contient pas forcément les deux parents ayant individuellement les LOD les plus importants, comme c'est le cas pour VBM08 et VBM09. Ainsi, si l'on choisit les femelles et les mâles avec les LOD individuels les plus importants on n'obtiendrait pas les mêmes parents pour les juvéniles que si le choix est fait en fonction du LOD obtenu par trio Juvénile-Femelle-Mâle.

Tableau 10 : Tableau des attributions parentales des juvéniles de l'élevage.

Juvéniles	cand. F	paire juv-fem			cand. M	paire juv-mâle			Trio J-F-M		
		mismatch	LOD	c		mismatch	LOD	c	mismatch	score LOD	c
VBM08	VPF07	0	2,19		VPM06	0	0,88	+	1	3,44	-
	VPF06	1	0,10		VPM06	0	0,88	+	1	3,16	
	VPF11	0	2,32	+	VPM06	0	0,88	+	1	2,98	
VBM09	VPF09	1	0,96	-	VPM03	2	-3,89		2	0,94	
	VPF09	1	0,96	-	VPM05	1	-2,98		2	-0,38	
	VPF09	1	0,96	-	VPM04	1	-0,82	-	3	-0,44	
VBM10	VPF06	2	-1,35		VPM04	1	1,11	*	2	1,72	-
VBF17	VPF14	0	1,25		VPM04	1	-0,57		3	0,09	-
	VPF12	0	0,79		VPM04	1	-0,57		3	-0,1	
VBF18	VPF01	1	2,42	+	VPM04	0	0,44	+	1	5,93	*
VBF19	VPF01	0	1,96	-	VPM04	1	-1,05	-	1	3,06	-
VBF20	VPF07	1	-0,34		VPM04	1	3,18	*	2	3,64	-
VBF21	VPF04	1	-1,97		VPM04	0	1,92	*	1	2,09	-

Niveau de confiance de (\*) 95%, (+) de 80%, (-) < 80% et ( ) indéterminé

Les résultats obtenus ne sont pas particulièrement convaincants étant donné que seul un trio (celui obtenu pour le juvénile VBF18) a été déterminé avec une certitude supérieure à 95%. Ils pourraient s'expliquer, entre autres, par la présence d'allèles nuls non détectés par Microchecker (Dakin & Avise, 2004). En effet, d'après Dakin et Avise, si un parent candidat fortement suspecté d'être le parent véritable se trouve avoir un locus homozygote à l'endroit d'une mésalliance, ceci peut être dû aux dits allèles nuls. Mais il est important de savoir que des phénomènes tels que l'effet Wahlund (synonyme de déficit d'hétérozygotes par rapport à l'équilibre d'Hardy-Weinberg due à un mélange de populations) et la consanguinité sont des facteurs biologiques qui donnent la fausse impression d'avoir des allèles nuls (Dakin & Avise, 2004).

### 3.2.3. Richesse allélique des individus de l'élevage

Après avoir déterminé, avec plus ou moins de réussite, les parents des juvéniles de l'élevage nous avons cherché à déterminer si ces derniers sont génétiquement plus diversifiés que leurs parents. Autrement dit, on cherche à savoir si la nouvelle génération a subi le boost génétique escompté. Pour cela la richesse allélique des 2 groupes de l'élevage que sont les juvéniles et les parents ont été comparés en termes de richesse allélique, comme vous pouvez le voir au Tableau 11.

La richesse allélique correspond au nombre d'allèles observés par locus donné dans un échantillon (Foulley & Ollivier, 2006). L'avantage de cette mesure par rapport au nombre d'allèles moyens par locus (AR), c'est qu'elle prend en compte la différence de taille des échantillons. Pour se faire, deux méthodes sont possibles celle de la raréfaction ou celle de l'extrapolation. La première consiste à calculer le nombre d'allèles dans une population pour un nombre  $n$  plus petit que celui dont on dispose tandis que le second fait l'inverse, soit ajouter un certain nombre d'allèles aux plus petites populations de manière à compenser la différence de  $n$  des échantillons (Foulley & Ollivier, 2006).

Ayant réalisé ce test via le logiciel FSTAT, c'est la méthode de raréfaction qui a été appliquée. Les résultats obtenus considèrent donc que les deux populations ont 8 individus contre 8 et 21 en réalité.

Tableau 11 : Tableau des richesses alléliques des Juvéniles et des adultes de l'élevage selon un modèle de raréfaction (en jaune sont repris les résultats pour lesquels les adultes ont une richesse allélique supérieure).

Marqueur	Juvéniles	Adultes
VB_A11	1.00	1.67
VB_Bp1	2.00	2.00
VB_D17	4.00	4.29
VB_Bp9	2.00	3.22
VB_A8	4.00	2.41
VB_B10	3.00	4.41
VB_B18	3.00	4.31
VB_Bp2	2.00	2.58
Va26	2.00	1.99
Va51	2.00	2.05
Va8	2.00	2.41
Vb64	1.00	1.00
Va20	1.00	2.58
Va25	2.00	2.86
Va35	1.00	1.00
Va69	1.00	1.65
VbDp10	4.00	4.04

Les résultats obtenus semblent indiquer que la diversité génétique diminue entre les deux générations. En effet, pour 12 des 17 marqueurs analysés (soit 70,6% d'entre eux), les adultes ont une diversité allélique supérieure. Cette observation reste toutefois logique au vu des résultats obtenus dans le Tableau 12. En effet, nous constatons que l'hétérozygotie observée ( $H_{obs}$ ) chez les adultes de l'élevage est faible, impliquant de nombreux homozygotes. Ceci implique que lorsqu'ils se reproduisent, il y a peu de choix dans les gènes transmis. Et puisque, comme on l'a vu au point précédent, il semble n'y avoir que 3 mâles et 6 femelles différentes impliquées dans la parenté de ces juvéniles le nombre différent d'allèles qu'ils ont reçus est limité. Pour rappel, Asmodé (Asmodé, 1990) avait dit qu'il était important que la population augmente rapidement via reproduction pour minimiser la perte de diversité génétique et son propos est très bien illustré ici (cf 1.1.5).

En plus de ceci, nous constatons dans le Tableau 12 que l'hétérozygotie observée des juvéniles est importante, plus importante que celle des adultes. Cette observation est positive puisqu'elle implique une atténuation du risque lié à la dépression de consanguinité. Celle-ci pouvant être défini comme une augmentation de caractères délétères récessifs liée à la reproduction entre individus apparentés.

*Tableau 12 : Tableau contenant le nombre d'individus (n), les hétérozygoties attendues sous l'équilibre d'Hardy-Weinberg ( $H_{exp}$ ), sans biais ( $H_{n.b.}$ ) et observée ( $H_{obs}$ ) et le polymorphisme (P) aux seuils de 95% et 99% et ce pour les 29 individus de l'élevage avec les 17 marqueurs retenus.*

	n	$H_{exp}$	$H_{n.b.}$	$H_{obs.}$	P(0.95)	P(0.99)	AR
Juvéniles	8	0,338	0,361	0,419	0,706	0,706	2,29
Adultes	21	0,358	0,368	0,279	0,882	0,882	3,24

#### 3.2.4. Analyse Factorielle de Correspondances (AFC)

L'analyse factorielle de correspondances ou AFC permet de comparer tous les individus entre eux sans tenir compte de leur population d'origine. C'est donc la première analyse qui va tenter de nous aider à déterminer l'origine des populations des Hautes Fagnes et de Kalmthout. La première AFC, reprise à la Figure 11 A, a été réalisée pour tous les individus analysés au laboratoire, et ce, avec les 17 marqueurs qui ont été conservés.

Nous constatons sur la Figure 11 A que l'AFC permet de distinguer 4 groupes différents. Un premier constitué des populations françaises, un second de la population des Ardennes, un troisième contenant l'élevage et un dernier contenant les individus des Hautes Fagnes ainsi que ceux de Kalmthout. Le groupe le plus proche de l'élevage est ici celui contenant la population des Ardennes.

Etant donné que les données tirées du mémoire de Fanny Demez ne comprennent des résultats que pour 8 des 17 marqueurs utilisés, une seconde AFC avec les 8 marqueurs en question a été réalisée afin de voir la différence entre les résultats obtenus (Figure 11 B).

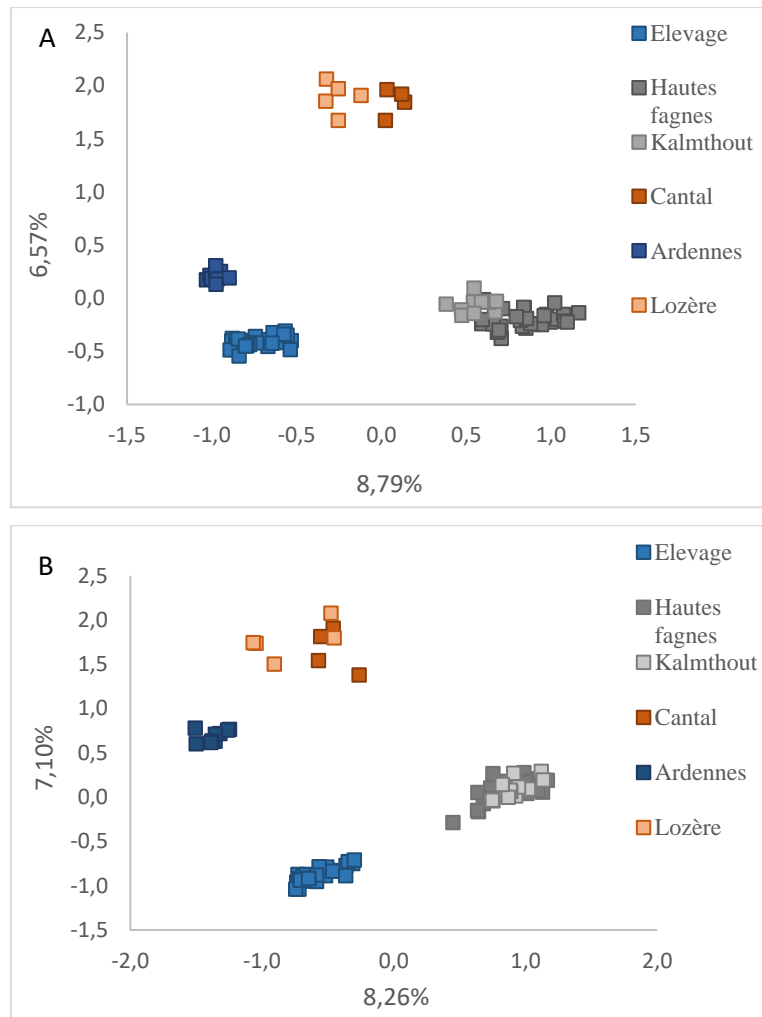


Figure 11 : Analyse factorielle de correspondances en deux dimensions, réalisée pour les (A) 83 individus analysés avec les 17 marqueurs conservés et les (B) 83 individus analysés avec 8 marqueurs.

Malgré des pourcentages d'information sensiblement équivalents, on observe que les différentes populations sont plus proches les unes des autres et que les différents groupes sont moins compacts avec les 8 marqueurs qu'avec les 17. Nous remarquons également que les groupes contenant l'élevage et la population des Ardennes se sont éloignés de l'élevage lors de la diminution du nombre de marqueurs. L'élevage semble à présent (avec 8 marqueurs) génétiquement plus proche du groupe contenant les populations des Hautes Fagnes et de Kalmthout. Même si dans les deux cas la population des Hautes Fagnes et celle de Kalmthout sont dans un même groupe, elles sont moins confondues avec 17 marqueurs qu'avec 8 marqueurs. Ceci nous permet de constater qu'avec une diminution du nombre de marqueurs, on perd en pouvoir de discrimination.

Une dernière AFC a ensuite été réalisée, cette fois-ci en incluant les résultats supplémentaires de la base de données de Fanny Demez. C'est donc bien la version avec 8 marqueurs, représentée dans Figure 11 B, à laquelle les résultats en question ont été ajoutés qui est ici considérée. Cette dernière figure nous permet de situer les individus de l'élevage près des populations 1, 4 et 11 qui sont situées à la frontière belgo-française. On constate également que les populations 17 et 19, qui sont des populations flamandes, se situent à proximité des populations des Hautes Fagnes et de Kalmthout et atteignent presque le groupe des individus de l'élevage (les différentes localisations des populations sont reprises à la Figure 10).

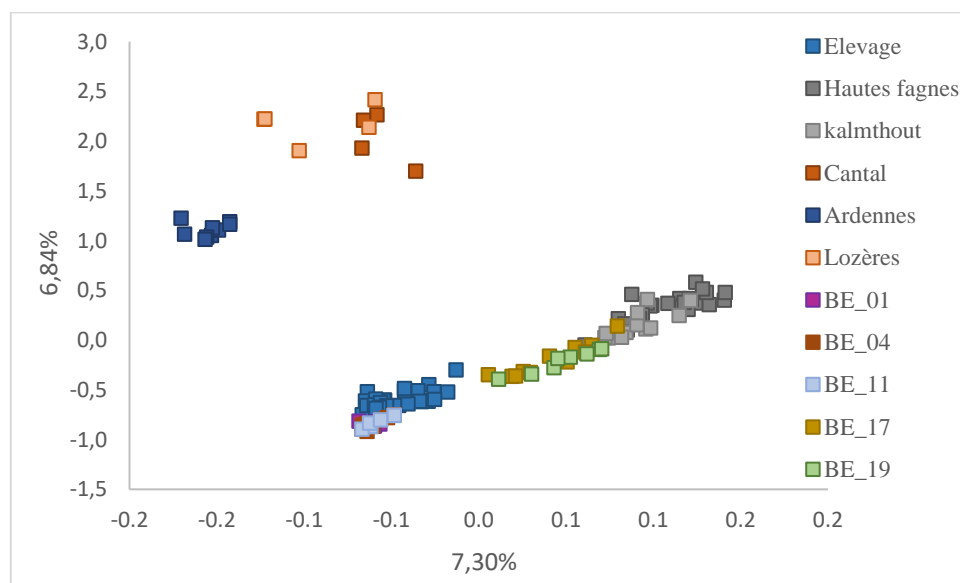


Figure 12 : Analyse factorielle de correspondances en deux dimensions, réalisée pour les 83 individus analysés et les 43 individus de Fanny Demez, et ce, avec 8 marqueurs.

### 3.2.5. Distances génétiques et dendrogrammes

La variabilité inter-populations a été mesurée via l'indicateur  $F_{ST}$  de (Weir & Cockerham, 1984). Cette variabilité a été calculée à deux reprises. La première mesure a été réalisée avec les résultats obtenus pour les 87 individus analysés par nos soins avec 17 marqueurs (Tableau 13), tandis que la seconde s'est faite sur bases des résultats obtenus pour les 87 individus avec 8 marqueurs, auxquels ont été ajoutés les résultats de 43 individus supplémentaires obtenus par Fanny Demez (Tableau 15). Sur base des matrices obtenues, deux dendrogrammes ont été établis avec la méthode du *neighbor-joining*.

Tableau 13 :  $F_{st}$  par paire de populations pour les 87 individus analysés avec 17 marqueurs. Tous les résultats obtenus sont significatifs pour  $p < 0.01$ .

$F_{ST}$	Elevage	Hautes Fagnes	Kalmthout	Cantal	Ardennes	Lozère
Elevage	0.000	0.332	0.419	0.444	0.308	0.393
Hautes fagnes	0.332	0.000	0.157	0.195	0.348	0.219
Kalmthout	0.419	0.157	0.000	0.297	0.467	0.292
Cantal	0.444	0.195	0.297	0.000	0.474	0.163
Ardennes	0.308	0.348	0.467	0.474	0.000	0.387
Lozère	0.393	0.219	0.292	0.163	0.387	0.000

En se basant sur la formule (4) retrouvée au point 1.3.3, une valeur de  $F_{ST}$  nulle indique que les hétérozygoties attendues en cas de panmixies globale ou locale sont égales. Ce qui revient à dire que les deux populations considérées lors du calcul du  $F_{ST}$  peuvent se croiser au hasard, autrement dit, elles appartiennent à une même population globale faisant d'elles des sous-populations de cette dernière. Et inversement, quand le  $F_{ST}$  est significativement différent de 0, cela signifie que les deux populations considérées lors du calcul sont bien deux populations distinctes.

Nous voyons que les indices de fixation par paire sont tous significativement différents de 0 ; il y a donc structuration (ou différenciation génétique) dans toutes les populations considérées. Le  $F_{st}$  le plus faible est obtenu pour les populations des Hautes Fagnes et de Kalmthout avec un score de 0.1575, ce sont donc les 2 populations les plus proches génétiquement. Cette constatation va dans le sens de ce qui avait été observé dans L'AFC à la Figure 11.

Le Tableau 14 contenant les  $F_{ST}$  obtenues pour les 3 sous-populations constituant la population des Hautes Fagnes et pour la population de Kalmthout nous donne un comparatif chiffré pour les distances obtenues au Tableau 13. Ainsi, on peut voir le score  $F_{ST}$  obtenu pour des groupes appartenant réellement à une même population et les comparer aux scores précédents. Il est d'ailleurs à noter que d'après (Hartl & Clark, 1997) tout  $F_{ST}$  supérieure à 0.25 indique une "très grande différenciation". Ce tableau permet également de voir de quelle sous-population des Hautes Fagnes la population de Kalmthout est la plus proche, en l'occurrence de celle constituée des individus provenant d'« autres ».

Tableau 14 :  $F_{ST}$  par paire de populations pour les 28 individus des Hautes Fagnes et les 12 de Kalmthout avec 17 marqueurs.

	Aachener Kreuz	Brackvenn	Autres	Kalmthout
Aachener Kreuz	0.000	0.029	0.003	0.188
Brackvenn	0.029	0.0000	0.017	0.173
Autres	0.003	0.017	0.000	0.166
Kalmthout	0.188	0.173	0.166	0.000

Noir= significatif pour  $p < 0.01$ , Vert= significatif pour  $p < 0.05$  et rouge = non significatif

Lorsqu'on passe de 17 à 8 marqueurs et qu'on ajoute les populations analysées par Fanny Demez, on constate que les résultats du Tableau 15 confirment les observations précédentes. Nous remarquons également que la seconde population la plus proche de celle de Kalmthout est la population flamande BE\_17, confirmant la constatation faite au point 3.2.4 selon laquelle cette population semblait appartenir au même groupe que les populations de Kalmthout et des Hautes Fagnes. De plus, l'influence du changement du nombre de marqueurs observé à la figure par le fait que la valeur de  $F_{ST}$  obtenue entre les populations Cantal et Lozère reste sensiblement la même entre les deux mesures, mais voit sa significativité diminuer (significativité passe de noir à vert entre les tableaux 13 et 15).

Tableau 15 :  $F_{ST}$  par paire de populations pour les 87 individus analysés par nos soins et les 43 individus de Fanny Demez, le tout avec 8 marqueurs.

$F_{ST}$	Elevage	HFagnes	Kalmthout	Cantal	Ardennes	Lozère	BE_01	BE_04	BE_11	BE_17	BE_19
Elevage	0.000	0.319	0.421	0.421	0.397	0.408	0.227	0.264	0.176	0.360	0.444
Hautes Fagnes	0.319	0.000	0.175	0.210	0.352	0.266	0.414	0.433	0.389	0.228	0.258
Kalmthout	0.421	0.174	0.000	0.311	0.472	0.348	0.531	0.566	0.509	0.197	0.371
Cantal	0.421	0.210	0.311	0.000	0.412	0.161	0.611	0.646	0.581	0.349	0.439
Ardennes	0.397	0.352	0.472	0.412	0.000	0.368	0.594	0.634	0.575	0.464	0.537
Lozère	0.408	0.266	0.348	0.161	0.368	0.000	0.592	0.621	0.564	0.353	0.446
BE_01	0.227	0.414	0.531	0.611	0.594	0.592	0.000	0.395	0.260	0.508	0.657
BE_04	0.264	0.433	0.566	0.646	0.634	0.621	0.395	0.000	0.067	0.543	0.697
BE_11	0.176	0.389	0.509	0.581	0.575	0.564	0.260	0.067	0.000	0.498	0.651
BE_17	0.360	0.228	0.197	0.349	0.464	0.353	0.508	0.543	0.498	0.000	0.279
BE_19	0.447	0.258	0.371	0.439	0.537	0.446	0.657	0.697	0.651	0.279	0.000

Noir= significatif pour  $p < 0.01$ , Vert= significatif que pour  $p < 0.05$  et rouge = non significatif

Le  $F_{ST}$  nous permet de déterminer s'il y a différenciation génétique entre les populations sur base de l'écart à l'équilibre d'Hardy-Weinberg, mais il ne nous indique pas la distance génétique à laquelle cette différenciation correspond. C'est pourquoi il est intéressant de travailler avec la distance de Nei (1978) notée  $D_{Nei}$ . Celle-ci, comprise entre 0 et l'infini, suppose que les différences génétiques sont causées par une mutation et une dérive génétique et nous permet de déterminer la distance génétique entre deux échantillons. Ainsi lorsque les valeurs sont relativement basses (en moyenne  $<0.3$ )<sup>2</sup>, nous pouvons considérer que les deux populations responsables du résultat obtenu appartiennent au même groupe génétique. Au plus la valeur de  $D_{Nei}$  augmente, au plus la période à laquelle les deux populations ont divergé est loin dans le temps car la distance génétique ( $D_{Nei}$ ) de Nei augmente proportionnellement au temps de divergence.

Tableau 16 : Distances génétiques de Nei pour les 87 individus analysés avec 17 marqueurs.

$D_{Nei}$	Elevage	HFagnes	Kalmthout	Cantal	Ardennes	Lozère
Elevage	0,000	0,810	0,803	0,889	0,282	0,601
HFagnes	0,810	0,000	0,373	0,694	1,045	0,863
Kalmthout	0,803	0,373	0,000	0,744	1,024	0,661
Cantal	0,889	0,694	0,744	0,000	0,825	0,396
Ardennes	0,282	1,045	1,024	0,825	0,000	0,568
Lozère	0,601	0,863	0,661	0,396	0,568	0,000

Les distances de Nei obtenues dans le Tableau 16 nous permettent de constater que les populations ayant divergé le plus récemment sont celles des Ardennes et de l'élevage suivi de celles des Hautes Fagnes et de Kalmthout. La seconde observation converge vers ce qui a été observé au niveau des  $F_{ST}$  contrairement à la première. En effet, dans le Tableau 13, les individus de l'élevage étaient dits génétiquement plus proches de la population des Hautes Fagnes, tandis qu'ici ils sont intégrés au même groupe génétique que celui des Ardennes. De même, les populations du Cantal et des Hautes Fagnes, entre autres, sont annoncées plus distantes qu'attendu au regard du score  $F_{ST}$  qui les lie.

En comparant les tableaux 16 et 17, nous remarquons que la diminution du nombre de marqueurs a une influence plus marquée sur ce test que sur celui du  $F_{ST}$ . En effet, nous voyons

<sup>2</sup> Cette valeur est variable en fonction de l'espèce étudiée (Gajardo, Da Conceicao, Weber, & Beardmore, 1995), on donne ici une valeur moyenne de différents travaux.

une augmentation globale des distances lorsqu'on passe à 8 marqueurs. Ceci est notamment observé pour l'élevage et la population des Ardennes, associés au même groupe génétique avec 17 marqueurs et fortement divergeant avec 8 marqueurs. L'élevage est à présent (dans le Tableau 17) associé au groupe génétique comprenant les populations BE\_01, 04 et 11. Ceci reflète le mouvement des individus des Ardennes observé sur les AFC lors du changement de marqueurs à la Figure 11 ainsi que les constatations faites à la Figure 12.

*Tableau 17 : Distances génétiques de Nei pour les 87 individus analysés par nos soins et les 43 individus de Mme Demez, le tout avec 8 marqueurs.*

D <sub>Nei</sub>	Elevage	HFagnes	Kalmthout	Cantal	Ardennes	Lozère	BE_01	BE_04	BE_11	BE_17	BE_19
Elevage	0,000	1,338	1,826	1,630	0,834	1,409	0,208	0,244	0,156	0,979	1,287
HFagnes	1,338	0,000	0,621	1,027	2,028	1,788	1,965	2,050	1,685	0,967	0,928
Kalmthout	1,826	0,621	0,000	1,281	2,964	1,849	1,951	2,243	1,925	0,459	1,111
Cantal	1,630	1,027	1,281	0,000	1,080	0,441	1,997	2,162	1,896	1,510	1,708
Ardennes	0,834	2,028	2,964	1,080	0,000	0,841	1,109	1,173	1,072	2,088	1,933
Lozère	1,409	1,788	1,849	0,441	0,841	0,000	2,024	2,059	2,027	1,585	2,174
BE_01	0,208	1,965	1,951	1,997	1,109	2,024	0,000	0,165	0,099	1,225	1,802
BE_04	0,244	2,050	2,243	2,162	1,173	2,059	0,165	0,000	0,025	1,350	1,815
BE_11	0,156	1,685	1,925	1,896	1,072	2,027	0,099	0,025	0,000	1,317	1,790
BE_17	0,979	0,967	0,459	1,510	2,088	1,585	1,225	1,350	1,317	0,000	0,533
BE_19	1,287	0,928	1,111	1,708	1,933	2,174	1,802	1,815	1,790	0,533	0,000

Si nous tenons uniquement compte des résultats obtenus au Tableau 17, nous aurions tendance à conclure que la population de Kalmthout est génétiquement plus proche de la population flamande BE\_17 qui, elle-même, serait à mi-distance entre la seconde population flamande (BE\_19) et Kalmthout. Les distances semblent indiquer qu'il y a d'abord eu une divergence entre la population principale BE\_17 et les individus qui ont ensuite formé la population des Hautes Fagnes. S'en est suivie une divergence entre BE\_17 et BE\_19, suivie peu de temps après par la divergence entre BE\_17 et la population de Kalmthout. Lors de sa divergence de BE\_17, la population de Kalmthout semble alors s'être génétiquement rapprochée de celle des Hautes Fagnes. Reste que les résultats ne semblent pas totalement fiables, du fait de leur forte variation avec le nombre de marqueurs utilisés, et que la détermination, à partir des distances génétiques, du temps écoulé depuis qu'un ancêtre commun a été partagé fait débat (Abatzopoulos, et al., 2002).

Les données du Tableau 13 et du Tableau 15 ont été reprises pour réaliser les dendrogrammes repris à la Figure 13. Ces arbres de *Neighbor-joining* permettent, via la longueur des traits, d'avoir une idée de la dérive génétique d'une population.

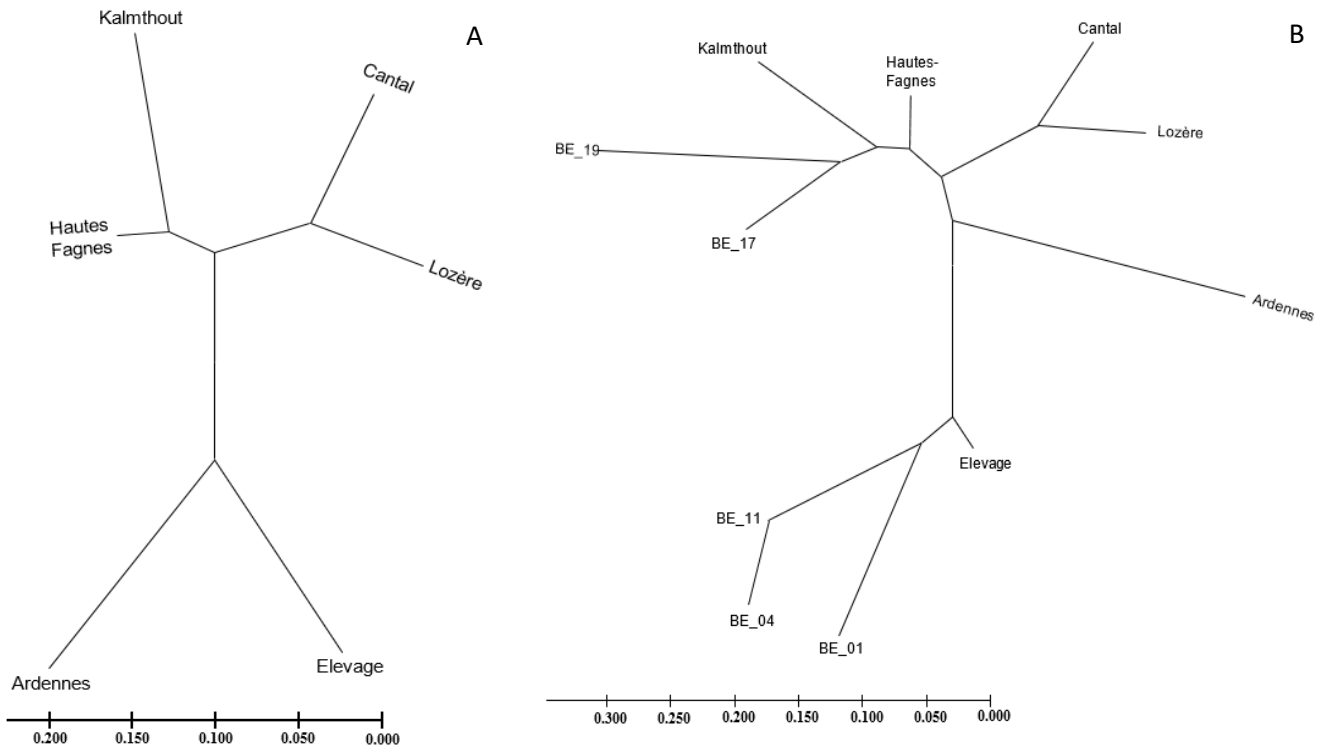


Figure 13 : Dendrogrammes construits via la méthode de neighbor-joining sur base des matrices  $F_{ST}$  reprises aux (A) tableau 10 et (B) tableau 11. Les échelles représentent les distances  $F_{ST}$ .

Ces dendrogrammes nous permettent notamment de constater que la seconde population flamande (BE\_19) est plus loin génétiquement parlant de celle de Kalmthout et de l'autre population flamande étant donné la dérive génétique que semble subir sa population. Nous remarquons également que la population des Ardennes est celle qui semble subir la plus forte dérive génétique. Les dendrogrammes permettent également de confirmer les résultats obtenus avec les distances de Nei. Les  $F_{ST}$  réajustés via *Neighbor-joining* donnent des résultats sensiblement similaires aux  $D_{Nei}$  et, par extension, aux AFC.

### 3.2.6. Goulot d'étranglement

Le test du goulot d'étranglement n'a été réalisé que sur les populations des Hautes Fagnes et de Kalmthout, étant les deux populations dont l'origine est questionnée, mais aussi parce que ce sont les deux populations avec les AR les plus élevés (Tableau 7) quand on sait que le goulot d'étranglement est défini comme une diversité génétique élevée pour un effectif réduit.

Pour rappel, 3 modèles distincts sont applicables pour déterminer la présence d'un goulot d'étranglement. Tout d'abord il y a le modèle d'allèle infinie (IAM) qui essaye de déterminer la quantité de loci homozygotes dans une population diploïde finie en considérant que toute mutation conduit à un nouvel allèle ; ces mutations entraînent des résultats allant de neutres à délétères. Il y a ensuite le modèle de mutation par étapes (SMM) et celui à deux phases (TPM) qui supposent que dans une population augmentant à l'infini, un microsatellite mute à un taux constant sans biais de contraction (perte de répétitions) ou d'expansion (gain de répétitions), quelle que soit sa longueur. Le premier des deux ne considère que des mutations liées à la perte ou au gain d'une seule unité de répétition, tandis que le second accepte des mutations avec unités de répétitions variables (cf 2.3.4).

D'après (De Meeûs, 2012), si un goulot d'étranglement a réellement eu lieu, elle sera détectera très fortement avec l'hypothèse IAM, moyennement avec le TPM et faiblement avec le SMM (Cornuet et Luikart, 1996), alors qu'en cas d'absence de goulot d'étranglement, mais en population structurée en petites sous-populations, on pourra détecter faussement une signature de goulot d'étranglement avec IAM, mais exceptionnellement (voir jamais) avec TPM et jamais avec SMM (De Meeûs *et al.*, 2010). Donc, en cas de tests très significatifs pour les trois procédures, nous pouvons être assez confiants.

Le test de Wilcoxon étant considéré comme le plus puissant et le plus robuste pour tester un goulot d'étranglement (De Meeûs, 2012), ce dernier est donc réalisé via le logiciel Bottleneck. Les 2 populations d'intérêt (Hautes Fagnes et Kalmthout) sont lancées sous les 3 hypothèses (TPM avec les paramètres par défaut). Les goulots d'étranglement récents étant décrits comme une perte d'allèles qui se fait plus vite que la baisse de la diversité génétique entraînant une diversité observée plus élevée que celle attendue, les valeurs de sortie d'intérêt sont dès lors les "one tail for H excess" obtenus pour chaque hypothèse et reprises dans le Tableau 18.

*Tableau 18 : P-valeur résultantes des tests de signature de goulot d'étranglement pour un test wilcoxon unilatéral.*

Population	IAM	TPM	SMM
Hautes Fagnes	0.00643	0.18891	0.12175
Kalmthout	0.88766	0.59121	0.86834

Nous constatons que seul le test IAM réalisé dans la population des Hautes Fagnes a une significativité suffisante pour le goulot d'étranglement. Et puisqu'il est facile de faussement conclure à un goulot d'étranglement quand seul IAM est significatif, nous allons considérer qu'aucune des deux populations n'a subi de goulot d'étranglement dans la fourchette détectable de  $[0.05 N_e^3 ; 5 N_e]$  générations (De Meeûs, 2012).

### 3.2.7. Structure

Le logiciel STRUCTURE 2.1 se base sur différentes hypothèses pour tenter de déterminer les sous-populations desquelles proviennent les différents individus. Ces hypothèses sont les suivantes :

- Les populations prédéfinies suivent presque parfaitement l'équilibre d'Hardy-Weinberg
- L'absence de déséquilibre de liaisons
- L'indépendance de la distribution des allèles à chaque locus
- L'absence d'isolement par la distance.

Ces hypothèses ont donc d'abord été testées avant de procéder aux calculs via structure.

#### 3.2.7.1. Vérification des hypothèses

##### a) Equilibre d'Hardy-Weinberg

Rappelons que si une population de taille infinie, sans mutation ni migration et dans laquelle il y a absence de sélection, se reproduit de manière sexuée avec rencontre au hasard des gamètes, elle devrait être à l'équilibre d'Hardy-Weinberg (cf 1.3.1).

Cet équilibre d'Hardy-Weinberg a été vérifié avec l'« exact test » via le logiciel GENEPOP, et ce pour les 11 populations (6 analysées au laboratoire et les 5 analysées par Fanny Demez) analysées avec 8 marqueurs différents (Tableau 19), mais aussi pour les 6 analysées avec 17 marqueurs (Tableau 20).

---

<sup>3</sup>  $N_e$  représente l'effectif efficace qui s'est mis en place après le goulot d'étranglement considéré (De Meeûs, 2012)

Tableau 19 : P-valeurs obtenues lors de l'“exact test” réalisé sur les 11 populations analysées avec 8 marqueurs.

	BE_01	BE_04	BE_11	BE_17	BE_19	Ardennes	Cantal	Elevage	HFagnes	Kalmthout	Lozère
VB_A11	-	-	-	0,7071	0,5375	-	1	1	0.0000	1	0,619
VB_A8	1	-	-	0,4811	1	0,5788	0,0835	0,185	0.0005	0,0003	0,0095
VB_B10	0,0667	-	0,0303	0,0316	-	-	0,7714	0,0319	0.0137	0,0154	0,3651
VB_B18	-	-	-	0,2727	-	0,0317	-	0,0416	0.0022	0,5699	-
VB_Bp10	1	0,9473	1	0,3584	0,1795	0,9105	1	0,8065	0.2781	0,0946	0,3333
VB_Bp2	-	-	-	0,0143	0,5641	0,9473	1	0,4685	0.3326	0,7702	0,0159
VB_Bp9	-	-	-	0,6037	-	0,0745	0,1489	0,0955	0.9999	0,1149	0,2952
VB_D17	0,5375	0,0186	0,0794	0,7947	1	1	1	0,0036	0.1791	0,5142	0,1111

Les données reprises en jaunes sont significatives pour des taux de 95% avec correcteur de Bonferroni.

Pour les résultats obtenus avec les 8 marqueurs (Tableau 19), nous remarquons que les populations analysées par Fanny Demez ne fournissent souvent pas assez d'informations pour que le test puisse donner un résultat. Ce manque d'informations n'est pas lié au nombre d'individus dans les populations considérées, puisque le Cantal et la Lozère ne sont respectivement composés que de 4 et 5 individus et donnent de meilleures réponses. Nous voyons également que seuls 3 populations et 4 marqueurs présentent des déséquilibres significatifs de l'équilibre de d'Hardy-Weinberg avec correction de Bonferroni aux taux de 95%.

Tableau 20 : P-valeurs obtenues lors de l'“exact test” réalisé sur les 7 populations analysées avec 17 marqueurs.

	Ardennes	Cantal	Elevage	HFagnes	Kalmthout	KR	Lozère
Va20	-	0,4286	0,043	0.1448	0,7071	0,3848	0,619
Va25	1	1	0,3116	0.0156	1	0,1765	0,0476
Va26	1	-	0,1241	0.0059	0,9139	0,8109	-
Va35	-	-	-	0.0014	0,2548	1	-
Va51	-	1	0,036	0.0017	1	0,3313	0,1429
Va69	-	0,6571	1	0.6744	0,5338	1	0,0857
Va8	0,5788	0,7714	0,4026	0.1919	0,5606	0,5294	0,7714
VB_A11	-	1	1	0.0000	1	0,2997	0,619
VB_A8	0,5788	0,0797	0,185	0.0009	0,0003	0,2337	0,0095

VB_B10	-	0,7714	0,0334	0.0358	0,0154	0,0915	0,3651
VB_B18	0,0317	-	0,0418	0.0047	0,5655	0,3407	-
VB_Bp10	0,9105	1	0,8065	0.2786	0,0951	0,3023	0,3333
VB_Bp2	0,9473	1	0,4685	0.2956	0,7785	0,3992	0,0159
VB_Bp9	0,0745	0,1434	0,0955	0.9998	0,1149	1	0,2952
VB_D17	1	1	0,0036	0.1668	0,5142	1	0,1111
VB_Dp10	0,0588	0,1413	0,2292	0.0000	0,0001	0,0186	1
Vb64	-	-	-	0.0037	1	0,0144	-

*Les données reprises en jaunes sont significatives pour des taux de 95% avec correcteur de Bonferroni.*

En utilisant plus de marqueurs, certains des résultats changent. En effet, la population des Hautes Fagnes était en déséquilibre de Hardy-Weinberg pour le marqueur VB\_B18 avec 8 marqueurs, mais ne l'est plus avec 17 marqueurs.

Si aucun marqueur ne semble particulièrement poser problème en termes d'équilibre d'Hardy-Weinberg, on ne peut pas en dire de même des populations. En effet, la population des Hautes Fagnes est en déséquilibre d'Hardy-Weinberg pour 3 de ses 8 marqueurs dans le premier cas, soit 37.5% d'entre eux, et pour 6 de ses 17 dans le deuxième, soit 35,3% d'entre eux. Ce non-respect de l'équilibre d'Hardy-Weinberg peut être signe d'une sous-division de la population. Sachant que cette population est divisible en 3 sous-populations en fonction du lieu d'échantillonnage des individus, il est décidé de globalement considérer les population comme étant à l'équilibre de Hardy-Weinberg.

#### b) Déséquilibre de liaison

Le déséquilibre de liaison est défini comme une association préférentielle ou non aléatoire de 2 loci (cf 1.3.4). Celui-ci entraîne une redondance des résultats pouvant conduire à une erreur de décision (De Meeûs, 2012). Un test de déséquilibre par paire de loci a donc été réalisé via GENEPOP afin de s'assurer que les loci utilisés étaient exempts de ce phénomène.

Les résultats obtenus sur l'ensemble de la population étudiée avec 8 marqueurs (soit nos 6 populations et les 5 de Fanny Demez), donnent 9 valeurs significatives ( $p < 0.05$ ) sur les 28 paires de marqueurs possibles. Après correction de Bonferroni une seule combinaison (VB\_Bp10 et VB\_B18) reste significative ( $p < 0.000162$ ). Lorsqu'on fait de même pour les 6 populations étudiées avec 17 marqueurs, nous obtenons 34 valeurs significatives ( $p < 0.05$ ) sur

les 136 paires de marqueurs obtenues. Encore une fois, après correction de Bonferroni seule la combinaison VB\_Bp10 et VB\_B18 reste significative ( $p < 0.000061$ ).

Le test de déséquilibre de liaison par paires de loci présente un problème non-négligeable lié à la répétition des tests. En effet, pour déterminer si les marqueurs sont liés entre eux, les tests ont été répétés. Or, les tests répétés ne sont pas indépendants, et peuvent donc conduire à l'inflation de la puissance du test (De Meeûs, 2012). Cette inflation peut entraîner la significativité d'un résultat qui ne l'aurait peut-être pas été si les tests avaient été indépendants. Un test exact binomial a donc été réalisé via Excel pour vérifier que les résultats obtenus ci-avant représentent bien un déséquilibre significatif. On y entre le nombre de succès (les tests non-significatifs dans notre cas, donc 27 et 135), le nombre de tests (28 et 136), la probabilité de succès (ici 1- p-valeurs de Bonferroni soit 0.999838 et 0.999939) et l'alternative "Vraie" qui représente la probabilité qu'il y ait au plus le nombre de succès renseigné. Nous obtenons respectivement pour les tests avec 8 et 17 marqueurs les p-valeurs de 0.00453 et 0.00826. Ceci nous confirme que nos observations précédentes sont bien significativement différentes de l'attendu et donc qu'il y a effectivement un déséquilibre de liaison pour la combinaison VB\_Bp10 et VB\_B18. Les données obtenues avec le marqueur VB\_B18 ont donc été supprimées pour les analyses réalisées via Structure pour éviter une redondance des résultats et respecter l'hypothèse d'absence de déséquilibre de liaisons.

### c) Isolement par la distance

L'isolement par la distance correspond à une augmentation de la différenciation avec l'éloignement géographique des individus (De Meeûs, 2012). Celui-ci a été évalué manuellement via Excel. Pour se faire, la distance génétique corrigée ( $F_{ST}/(1-F_{ST})$ ) a été calculée et exprimée en fonction du logarithme népérien de la distance en mètres. Ceci a été réalisé pour les 87 individus analysés dans notre laboratoire ajoutés aux 43 individus de Fanny Demez, le tout analysé avec 8 marqueurs (Figure 14 A), mais aussi pour nos 87 individus seuls analysés avec 17 marqueurs (Figure 14 B).

La création d'une droite de régression a permis d'apprécier la répartition des différentes paires de sous-populations. Nous constatons d'ailleurs que celle-ci ne peut expliquer que 5% de la distribution des points dans le cas où 8 marqueurs sont utilisés, comme nous l'indique le  $R^2$ , contre 18% dans celui avec 17 marqueurs. Ces résultats nous laissent penser qu'il n'y a

pas d'isolement par la distance ou, autrement dit, que la différence entre les individus de deux populations différentes n'est pas imputable à leur distance.

Le test de Mantel réalisé via GENEPOP nous donne d'ailleurs une P-valeur de 0.2699 pour les populations analysées avec 8 marqueurs contre une de 0.06502 pour celles analysées avec 17 marqueurs. Sachant que ce test a pour principe de permuter les lignes et les colonnes de deux matrices - ici celles de la distance corrigée et du logarithme de la distance- pour en tester la corrélation, cela nous permet de confirmer la non-significativité du test. Autrement dit les deux variables sont considérées indépendantes et il n'y a donc pas d'isolement par la distance des populations, quel que soit le nombre de marqueurs utilisés.

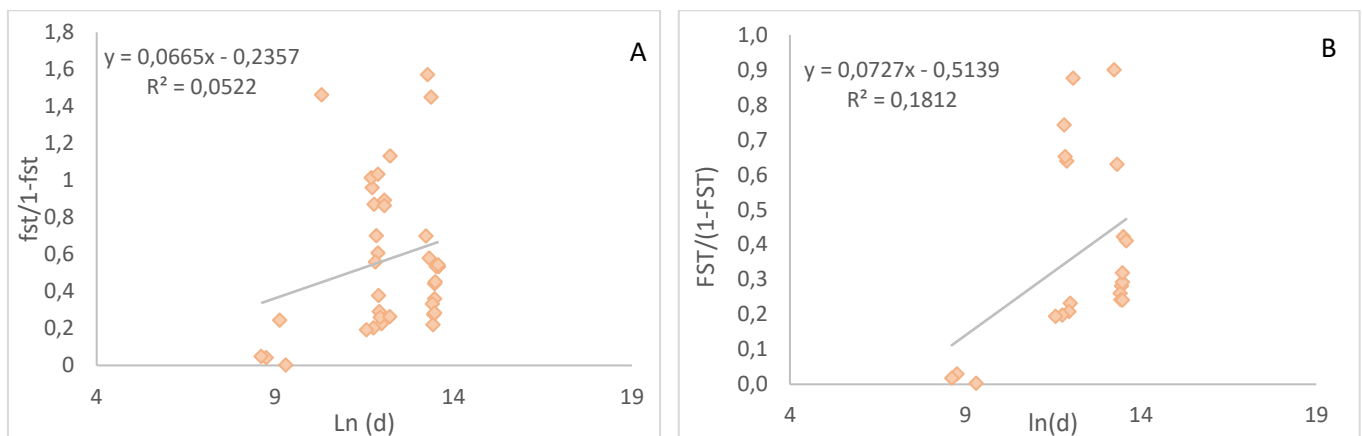


Figure 14 : Graphique représentant l'évolution de la distance génétique corrigée [ $F_{ST}/(1-F_{ST})$ ] en fonction du logarithme népérien de la distance géographique en mètre [ $\ln(d)$ ] et ce pour (A) 8 marqueurs et 11 populations et (B) 17 marqueurs et 6 populations.

### 3.2.7.2. Résultats

Les tests ont été effectués pour respectivement K allant de 1 à 9 pour les données obtenues dans notre laboratoire avec 16<sup>4</sup> marqueurs, contre K allant de 1 à 13 pour les données à 7<sup>5</sup> marqueurs combinées aux résultats de Fanny Demez. Les paramètres suivants ont été utilisés : *Burn-in* durant 200000 cycles, 500000 itérations MCMC (Monte Carlo Markov Chain) après la période de *Burn-in* et 10 répétitions pour chaque valeur de K. Un modèle de structure des populations considérant un mélange de ces dernières (admixture model) avec des fréquences corrélées a été appliqué. Ce modèle implique que les données génétiques sont mélangées entre populations. Autrement dit, les individus peuvent avoir des allèles provenant de

<sup>4</sup> Précédemment on avait 17 marqueurs mais on a retiré VBb18 suite aux résultats de déséquilibres de liaison

<sup>5</sup> Idem référence note de bas de page 4.

populations différentes. Voulant, par le biais de ce test, connaître la structuration des différentes populations il est donc plus intéressant d'appliquer ce modèle.

Le nombre de populations différentes (K) estimé par Structure pour les 87 individus analysés avec 16 marqueurs différents est de 2 ou de 4, valeurs auxquelles les pics sont obtenus sur la Figure 15 A. Ceci signifie que seuls les résultats obtenus pour K=2 et K=4 sont retenus pour établir les figures de résultats.

De même, pour nos résultats obtenus avec 8 marqueurs et associés aux résultats de Fanny Demez, les pics ont été atteints pour K=4 et K=6 (Figure 15 B). Cette augmentation du nombre estimé de populations est logique puisque le nombre de populations intégrées au test a lui aussi augmenté. De même, seuls les résultats obtenus pour K=4 et K=6 sont considérés pour la suite.

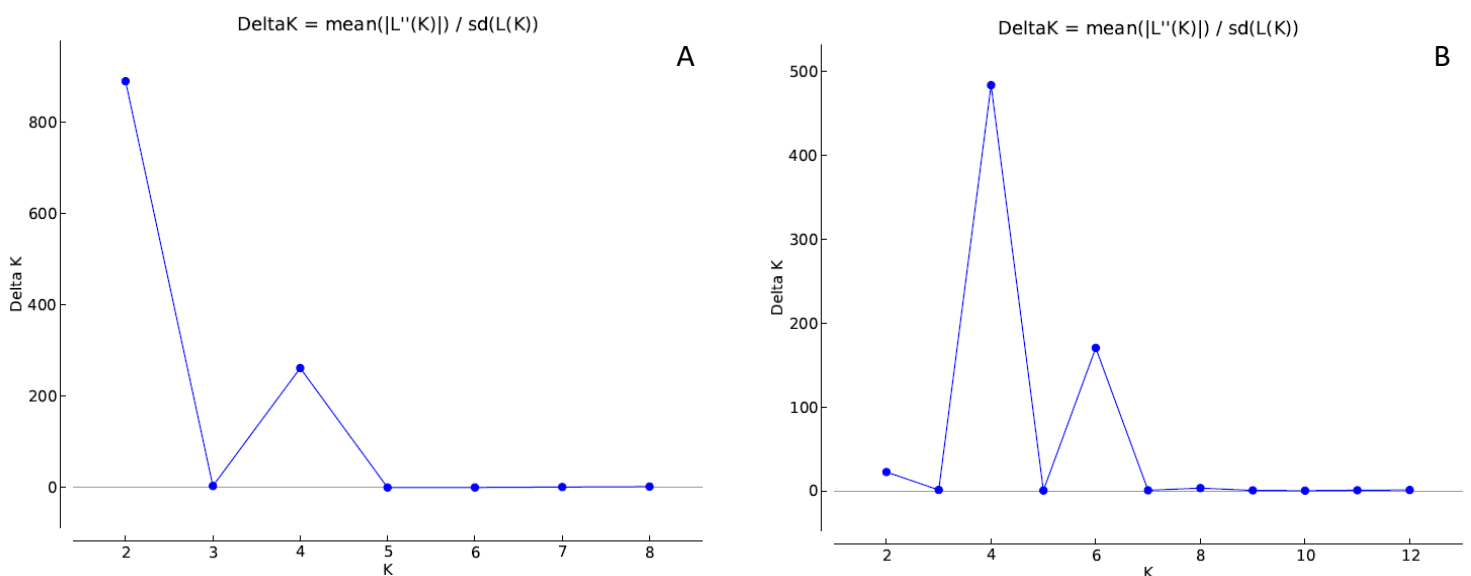


Figure 15 : Graphique représentant l'évolution du delta K en fonction du nombre de sous-populations k obtenu via le logiciel Structure pour (A) 6 populations et 16 marqueurs et pour (B) 11 populations et 7 marqueurs.

Les résultats ainsi obtenus sont repris aux figures 16 et 17. Ces figures illustrent, sur l'axe vertical, la probabilité que chaque individu appartienne à une des 2/4 (Figure 16) ou 4/6 (Figure 17) sous-populations déterminées par Structure en fonction du cas, chacune de ces populations étant représentée par une couleur différente. Les subdivisions en noir différentient les lieux de prélèvement des individus et les couleurs nous donnent les populations dans lesquelles Structure les classe.

Dans les résultats établis avec 16 marqueurs, pour  $K=2$  (Figure 16 A), les deux groupes génétiques identifiés par Structure contiennent respectivement les populations des Hautes Fagnes, de Kalmthout, du Cantal et de Lozère, puis l'élevage avec la population des Ardennes. Ceci revient à couper en 2, horizontalement, le dendrogramme obtenu à la Figure 13 A. En divisant en 4 sous-populations (Figure 16 B) plutôt qu'en 2, le premier groupe cité ci-avant se voit subdivisé en trois groupes, avec sortie des populations de Kalmthout et des Hautes Fagnes qui se retrouvent alors isolées. Ces sous-groupes semblent plus correspondre à des sous-populations prédéfinies, telles qu'échantillonnées, qu'à un apport supplémentaire d'informations.

On peut donc penser que ces deux populations semblent avoir la même origine que les deux populations françaises.

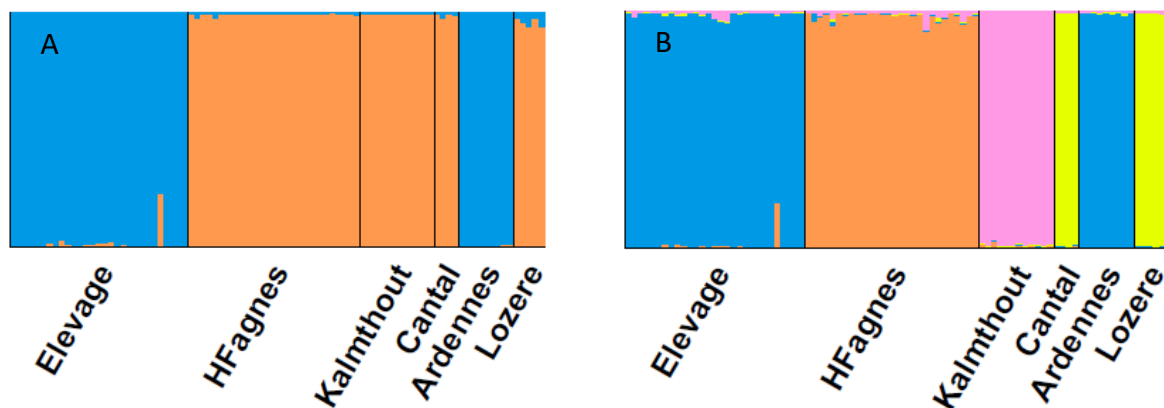


Figure 16 : Figure représentant la répartition des 6 populations analysées avec 16 marqueurs dans les sous-populations auxquelles structure les attribue pour (A)  $K=2$  et (B)  $K=4$ .

Il est intéressant de constater que lorsque les populations sont divisées en 4 sous-populations ( $K=4$ ) avec 7 marqueurs (Figure 17 A), la population des Ardennes est placée avec celles du Cantal et de la Lozère, ce qui n'était pas du tout le cas avec 16 marqueurs. Ceci confirme les observations faites précédemment, notamment lors de l'AFC (Figure 11). Au regard de la Figure 17, nous constatons également que les individus de l'élevage sont indiqués comme ayant la même origine que les populations BE\_01, 04 et 11, confirmant aussi les résultats précédents.

Les résultats obtenus à la figure équivalent à la coupure en trois endroits différents du dendrogramme repris à la Figure 13 B. De plus, plutôt que d'être associée à la population des Hautes Fagnes, comme c'était le cas à la Figure 16 A, la population de Kalmthout est associée

aux deux populations flamandes, BE\_17 et BE\_19. En passant de 4 à 6 sous-populations (Figure 17 B), une séparation de la population des Ardennes d'avec celle du Cantal et de la Lozère est observée. Ce qui concorde plus avec les résultats obtenus avec 16 marqueurs. De même, nous observons une sortie de la population de Kalmthout de son groupe précédent pour former la sixième sous-population.

En combinant les résultats obtenus avec 16 et 7 marqueurs, nous tendons à penser que les populations des Hautes Fagnes, de Kalmthout et les deux populations flamandes sont issues de populations françaises. Il serait d'ailleurs intéressant de voir si avec 16 marqueurs, les populations BE\_17 et 19 seraient associées avec celles de Kalmthout et des Hautes Fagnes.

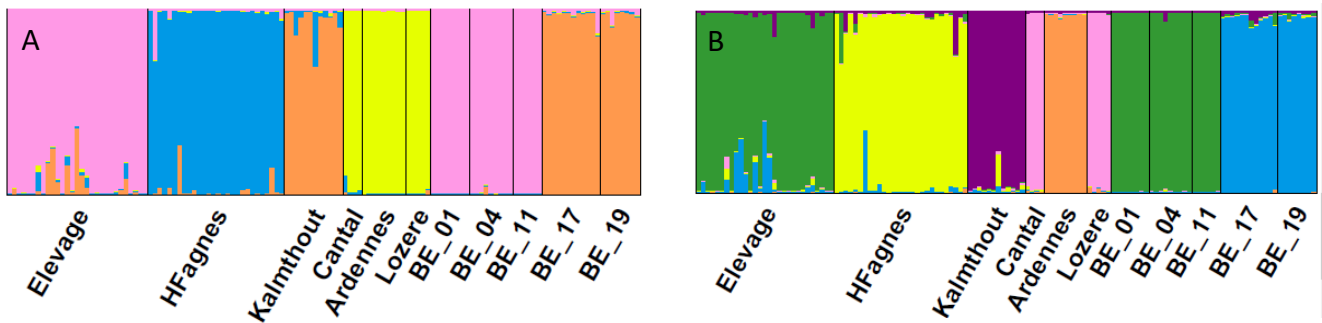


Figure 17 : Figure représentant la répartition des 11 populations analysées avec 7 marqueurs dans les sous-populations auxquelles structure les attribuees pour (A) K=4 et (B) K=6

D'après (Ursenbacher, 2005), les vipères issues de la région des Ardennes, tout comme celles de la Lozère proviennent du même clade phylogéographique, nommée clade nordique. Plus spécialement, elles proviendraient du sous-clade ouest de la clade nordique. Etant donné que ces deux populations appartiennent, d'après Structure, chacune à un groupe différent des deux groupes génétiques identifiés (Figure 16 A), cela suggère que toutes les populations considérées proviennent de ce sous-clade ouest. Il se pourrait donc qu'un groupe initial ait quitté ledit clade avant de se diviser en deux groupes distincts que nous observons via Structure. Ces deux groupes se sont, à leur tour, divisés pour former des populations dont quelques-unes sont reprises ici.

## 4. Conclusion et perspectives

En conclusion, nous remarquons d'abord que les extractions d'ADN jouent un rôle prépondérant dans les analyses génétiques étant donné que tous les résultats d'analyse en dépendent. En optimisant cette étape, le risque d'observer des allèles nuls diminue, mais pas seulement. En effet, une bonne extraction entraîne l'obtention d'un ADN de qualité et en quantité suffisante pour une interprétation facilitée des résultats. C'est donc un gain de temps puisqu'il y a alors moins de répétitions nécessaires à la confirmation des résultats, mais aussi d'argent, pour les mêmes raisons. Il serait donc intéressant de mieux comprendre les facteurs qui influencent la qualité et la quantité d'ADN lors de l'extraction.

Ensuite, la base de données établie pour *V. berus* gagnerait à être enrichie de résultats obtenus avec plus de marqueurs, notamment car certains des huit marqueurs spécialement développés pour *V. berus* sont difficilement interprétables et, donc, fortement dépendants de l'appréciation de l'opérateur entraînant un biais important sur les résultats. Les marqueurs développés pour *V. aspis* ont d'ailleurs été une source d'information importante pour la détermination des filiations dans l'élevage et ont également permis d'apprécier un gain de précision dans certaines analyses statistiques, telle que l'AFC.

Nous remarquons d'ailleurs via les différentes analyses menées sur les individus de l'élevage que celles-ci donnent des résultats encourageants, les juvéniles présentant une hétérozygotie observée en hausse par rapport à leurs parents. Le seul problème qui subsiste est lié au nombre de juvéniles voyant le jour, autrement dit, à la faible fécondité de l'espèce. En effet, il est nécessaire d'augmenter rapidement la population pour qu'un maximum de gènes soit transmis et que le pool allélique de la population ainsi créée reste important.

Même si l'origine des populations des Hautes Fagnes et de Kalmthout reste incertaine, les différentes analyses nous ont permis de constater qu'elles n'ont pas du tout les mêmes origines que les individus de l'élevage, une introduction d'individus de ces populations dans l'élevage semble donc compromise.

Les analyses factorielles de correspondances tout comme les distances génétiques de Nei et l'analyse via le logiciel structure ont permis de mettre en évidence la possible origine commune des populations de Kalmthout, des Hautes Fagnes, de deux populations flamandes

nommées BE\_17 et BE\_19, correspondant respectivement aux populations de Groot Schietveld et de Visbeekvallei, ainsi que de deux populations françaises que sont le Cantal et la Lozère. Elles ont également montré une possible appartenance des populations des Ardennes et celles à la frontière franco-belge (BE\_01, 04 et 11) à un second groupe qui comprendrait aussi les individus de l'élevage.

En se basant sur nos résultats associés à ceux de la thèse de Sylvain Ursenbacher, l'hypothèse que ces deux groupes génétiques aient été issus de la division d'une première population émanant du sous-clade ouest de la clade nordique avant divisions en différentes sous-populations a été émise. Il a également été supposé, au regard des  $D_{NEI}$  obtenus, que les populations BE\_19, Kalmthout et Hautes Fagnes soient issues de la populations BE\_17.

Différents tests seraient intéressants à réaliser pour vérifier ces hypothèses. Les séparations génétiques gardant un signe phylogénétique pouvant être, entre autres, analysé via ADN mitochondriale ; le premier test consisterait donc à travailler avec de ledit ADN mitochondrial pour voir si les populations flamandes (Kalmthout incluse) et celle des Hautes Fagnes proviennent de deux lignées différentes. Il serait également intéressant d'ajouter plus d'individus français à la base de données pour pouvoir comparer leurs populations. Refaire une analyse des différents individus analysés par Fanny Demez (si les extraits sont encore disponibles) avec 17 marqueurs serait aussi intéressant afin d'observer si les résultats obtenus via Structure restent tels qu'observés avec 8 marqueurs ou s'ils confirment nos hypothèses. Enfin, pour une significativité plus élevée des résultats obtenus, analyser un plus grand nombre d'individus différents dans chaque population serait un plus.

## Bibliographie

- Abatzopoulos, T. J., Kappas, I., Bossier, P., Sorgeloos, P., & Beardmore, J. A. (2002). Genetic characterization of *Artemia tibetiana*. *Biological journal of the linnean society*.
- Asmodé, J.-F. (1990, Octobre). Reproduction en captivité et réintroduction dans la nature, point de passage nécessaire pour les espèces menaces. *Cahiers d'Outre-mer*, pp. 464-474.
- Ball, S., Hand, N., William, F., Durrant, C., Uller, T., Claus, K., . . . Garner, T. W. (2020, avril 20). *GENETIC AND DEMOGRAPHIC VULNERABILITY OF ADDER POPULATIONS: RESULTS OF GENETIC STUDY IN MAINLAND BRITAIN*.
- Belkhir, K., Borsa, P., Chikhi, L., Raufaste, N., & Bonhomme, F. (2004, Mai 05). *Page web de Genetix*. Récupéré sur <https://kimura.univ-montp2.fr/genetix/>
- Brownstein, M., Carpten, J., & Smith, J. (1995, Décembre). Modulation of non-templated nucleotide addition by taq DNA polymerase : primer modifications that facilitate genotyping.
- Carlsson, M., Isaksson, M., Hoggren, M., & Tegelstrom, H. (2002). *Characterization of polymorphic microsatellite markers in the adder, vipera berus*. Blackwell publishing Ltd.
- Cornuet, J., & Luikart, G. (1996, Septembre 18). Description and Power Analysis of Two Tests for Detecting Recent Population Bottlenecks From Allele Frequency Data. *Genetics vol. 144*.
- Crow, J., & Denniston, C. (1988). Inbreeding and variance effective population numbers. *Evolution, 42(3)*.
- Dakin, E., & Avise, J. (2004). Microsatellite null alleles in parentage analysis. *Heredity*. USA: Nature Publishing Group.
- De Fazio, C. (2010). Rôle de la dépression de consanguinité dans le risque d'extinction des petites populations.

- De Meeûs, T. (2012). *Initiation à la génétique des populations naturelles*. Marseille: IRD éditions.
- Demez, F. (2012-2013). Caractérisation génétique et conservation des populations de vipères péliades (*Vipera berus*) en Wallonie, et comparaison avec des populations suisses et françaises.
- Di Rienzo, A., Peterson, A., Garza, J., Valdes, A., Slatkin, M., & Freimer, N. (1993, Novembre). Mutational processes of simple-sequence repeat loci in human populations.
- Feingold, J. (1991). Le déséquilibre de liaison. *médecine/sciences*.
- Flamand, M.-C. (2001-2002). Conventions de recherche Région wallonne.
- Foulley, J., & Ollivier, L. (2006). Diversité génétique et richesse allélique : concepts et applications à des races bovines.
- Gajardo, G., Da Conceicao, M., Weber, L., & Beardmore, J. A. (1995). Genetic variability and interpopulational differentiation of *Artemiu* strains from South Africa. *Hydrobiologia*.
- Garrigues, T., Dauga, C., Ferquel, E., Choumet, V., & Failloux, A.-B. (2005, janvier 10). Molecular phylogeny of *Vipera Laurenti*, 1768 and the related genera *macrovipera* (Reuss, 1927) and *Daboia* (Gray, 1842) with comments about neurotoxic *vipera aspis aspis* populations.
- Geser, S., Kaiser, L., Zwahlen, V., & Ursenbacher, S. (2013). Development of polymorphic microsatellite loci markers for the Asp viper (*Vipera aspis*) using high throughput sequencing and their use for other European vipers. Leiden.
- Guyomard, R. (1989). Gestion génétique des populations naturelles : l'exemple de la truite commune.
- Hartl, D. L., & Clark, A. G. (1997). *Principles of population genetics*. Massachusetts.
- Joger, U., Lenk, P., Baran, I., Böhme, W., Ziegler, T., Heidrich, P., & Wink, M. (1997). The phylogenetic position of *Vipera barani* and of *V. nikolskii* within the *Vipera berus* complex. *Herpetologia Bonnensis*, 185-194.

- Jomphe, V. (2006). Comparaison de la puissance de tests de déséquilibre de liaison dans les études génétiques. *Mémoire présenté à la faculté des études supérieures de l'Université de Laval*. Quebec.
- Julian, A., & Hodges, R. J. (2019). The vanishing viper: themes from a meeting to consider better conservation of *Vipera berus*. *Herpetological Bulletin* 149.
- Kimura, M., & Crow, J. (1964, Avril). *The Number of Alleles That Can Be Maintained in a Finite Population*.
- Kimura, M., & Ohta, T. (1978, Mars). Stepwise mutation model and distribution of allelic frequencies in a finite population. Mishima, Japon.
- Kwok, S., Kellogg, D., McKinney, N., Spasic, D., Goda, L., Levenson, C., & Sninsky, J. (1990, Janvier 19). Effects of primer - template mismatches on the polymerase chain reaction : human immunodeficiency type 1 model studies. *Nucleic acids research*, vol. 18, No. 4. CA, USA.
- Lecomte, J. (1990). Rev. Ecol. (Terre vie). *Réintroduction et renforcement de populations animales: Les contraintes éthologiques*.
- Madsen, T., Ujvari, B., & Olsson, M. (2004, janvier 21). *Novel genes continue to enhance population growth in adders (Vipera Berus)*. Elsevier.
- Marshall, T. (2001, Avril 24). Cervus 2.0. Royaume uni.
- Nei, M. (1987). *Molecular evolutionary genetics*. New York.
- Nilson, G., Andrén, C., & Szyndlar, Z. (1994). The systemic position of the common adder, *Vipera berus* (L.) (Reptilia, Viperidae), in North Korea and adjacent regions. *Bonner zoologische Beitrage* 45, 49-56.
- Paquay, M., & Graitson, E. (2007). La Vipère péliade, *vipera berus* (Linnaeus 1758).
- Pemberton, J., Slate, J., Brancroft, D., & Barrett, J. (1995, Avril). Nonamplifying alleles at microsatellite loci: a caution for parentage and population studies. *Molecular Ecology* 4, 249-252.

- Piry, S., Luikar, G., & Cornuet, J.-M. (s.d.). Bottleneck program. *BOTTLENECK : A program for detecting recent effective population size reductions from allele data frequencies.*
- Prestt, I. (1971). An ecological study of the viper *Vipera berus* in southern Britain.
- Pritchard, J., Stephens, M., & Donnelly, P. (2000). *Inference of Population Structure Using Multilocus Genotype Data.*
- Saint-Girons, H. (1979). Les cycles alimentaires des vipères européennes dans des conditions semi-naturelles.
- Sainudiin, R., Durrett, R., Aquadro, C., & Nielsen, R. (2004, Juin). *Microsatellite Mutation Models: Insights From a Comparison of Humans and Chimpanzees.*
- Saitou, N., & Nei, M. (1987). The Neighbor-joining Method: A New Method for reconstructing phylogenetic trees.
- Santune, V., Vanappelghem, C., Cheyrezy, T., Coquel, L., Holliday, j., & Damoy, V. (2012, février). vipère péliade (*vipera berus*). *Plan d'action régionale de la vipère péliade (Nord-pas-de-Calais).* Conservatoire d'espaces naturels du Nord et du Pas-de-Calais.
- Serre, J.-L. (2006). *Génétique des populations.*
- Smith, G. (1973). Unequal crossover and the evolution of multigene families. *Symposia on Quantitative Biology.*
- Srivastava, S., Avvaru, A., Sowpati, D., & Mishra, R. (2019). Patterns of microsatellite distribution across eukaryotic genomes.
- Streisinger, G., Okada, Y., Emrich, J., J. Newton, Tsugita, A., & al., e. (1966, janvier). Frameshift Mutations and the Genetic Code.
- Ursenbacher, S. (2005). *Phylogéographie des principales vipères européennes (vipera ammodytes, v.aspis et v. berus), structuration génétique et multipaternité chez Vipera berus.* Lausanne.
- Ursenbacher, S., Monney, J.-C., & Fumagalli, L. (2009). *Limited genetic diversity and high differentiation among the remnant adder (Vipera berus) populations in the Swiss and French Jura Mountains.*

Vacher, J.-P., & Geniez, M. (2010). *Les reptiles de France, Belgique, Luxembourg et Suisse*.

Collection Parthénope, éditions Biotope.

Weir, B., & Cockerham, C. C. (1984, Mai 07). Estimating F-statistics for the analysis of population structure.

Wierdl, M., Dominska, M., & Petes, T. (1997, Mars). *Microsatellite Instability in Yeast: Dependence on the Length of the microsatellite*.

Wright, S. (1951). *The genetical structure of populations*. Londres.





## Analyse génétique de la vipère péliade (*Vipera berus*) pour une réintroduction durable.

TAPIA NSANG Nerea

La vipère péliade ou *Vipera berus* est une espèce endémique belge en voie d'extinction. L'asbl Natagora a donc mis en place un plan d'action qui vise à la conserver, notamment en la réintroduisant dans de grands espaces adaptés, classés réserves naturelles. Pour permettre une réintroduction dans ces sites, un élevage a été constitué.

Sachant que l'une des problématiques liées à l'extinction des vipères péliades est la forte consanguinité de certaines populations, l'un des objectifs de ce mémoire a donc été de tester la diversité génétique des individus destinés à la réintroduction. Pour ce faire, des échantillons d'ADN de chaque individu de l'élevage nous ont été transmis. Il a aussi été question de tenter de déterminer l'origine de deux populations, celle de Kalmthout et celle des Hautes Fagnes. Il pourrait être envisagé, si nécessité et/ou plus-value il y a, d'introduire des individus de ces populations dans l'élevage dans le cas où ces dernières auraient une origine wallonne. Des échantillons de ces deux populations ainsi que des échantillons de référence nous ont donc été fournis pour déterminer cette ascendance.

Il a ainsi été constaté que les juvéniles de l'élevage ont une hétérozygotie plus importante que celle de leurs parents, indiquant une évolution encourageante. L'origine des populations des Hautes Fagnes et de Kalmthout n'a pas été déterminée avec certitude, mais il semblerait qu'une introduction de ces individus dans l'élevage n'est pas envisageable.

L'hypothèse d'une provenance commune des populations des Hautes Fagnes et de Kalmthout a été faite. On peut en effet, via les analyses réalisées, supposer qu'elles sont toutes deux issues d'une population flamande. Cette supposition gagnerait toutefois à être vérifiée via une analyse de l'ADN mitochondrial des différentes populations concernées.