

Université PARIS EST-CRETEIL

Université PARIS SUD



Ecole nationale Vétérinaire d'Alfort

MASTER 2<sup>eme</sup> ANNEE

Santé publique Paris Sud-Saclay et Santé UPEC

Dominante

**SURVEILLANCE EPIDEMIOLOGIQUE DES MALADIES  
HUMAINES ET ANIMALES**

---

# RAPPORT DE STAGE

L'épidémiologie moléculaire de la tuberculose bovine – Rôle de la faune sauvage  
et de bovins dans la transmission en Dordogne et Haute-Vienne

Présenté par

Gabriela MODENESI

Réalisé sous la direction de : Laetitia CANINI et Lorraine MICHELET

Organisme et pays : Unité Epidémiologie, Laboratoire de Santé Animale, Anses - France

Période du stage : 28 janvier 2019 au 26 juillet 2019

Date de soutenance : 26 juin 2019

Année universitaire 2018-2019

## Remerciements

A Laetitia Canini et Lorraine Michelet

Pour tous les enseignements, la disponibilité et la patience

A Barbara Dufour

Pour le soutien et les encouragements

A Maud Marsot

Pour tout l'aide

A l'équipe de l'unité EPI

Pour le chaleureux accueil et leur disponibilité

A l'équipe du UZB

Pour l'accueil et les manipulations qui ont généré les données

A Valentina

Pour être ma famille en France

A Cynthia

Pour l'écoute, les rires et pour faire mes journées plus légères

A Marcelo

Pour le soutien et la compréhension admirables

## Table des matières

Index des figures .....	3
Index des tableaux.....	4
Résumé court .....	5
Résumé long .....	6
Synthèse bibliographique.....	9
1 La tuberculose bovine .....	9
1.1 Pathologie et épidémiologie.....	9
1.2 La surveillance de la tuberculose bovine en France.....	9
2 Séquençage du génome .....	10
3 Phylogénétique et épidémiologie moléculaire .....	11
3.1 Principe.....	11
4 Inférence phylogénétique .....	14
5 Epidémiologie moléculaire .....	15
TRAVAIL PERSONNEL.....	16
1 Contexte .....	16
2 Matériels et Méthodes.....	17
2.1 Sélection des souches.....	17
2.2 Séquençage complet des souches.....	17
2.3 Modèle évolutif .....	18
2.4 Inférence phylogénétique .....	19
2.5 Distribution spatiale des clades.....	20
3 Résultats .....	20
3.1 Zones d'étude et souches sélectionnées.....	20
3.2 Analyse des données génomiques .....	22
3.3 Modèle sélectionné et analyse phylogénétique .....	23
3.4 Analyse phylogénétique .....	25
3.5 Distribution spatiale des clades.....	28
4 Discussion .....	34
5 Conclusion .....	37
6 Références .....	38
7 Annexes .....	43
Annexe 1 – Mesures de police sanitaire selon statu du troupeau de bovine (Arrêté du 15/09/2003 – Chapitre 5).....	43
Annexe 2 – Modèles de substitution de nucléotides .....	44
Annexe 3 – Modèles évolutifs testés avec BEAST .....	46

Annexe 4 – Arbre phylogénétique divisé par clade et probabilité du trait des nœuds hypothétiques .....	47
Annexe 5 – Liste de souches par clade.....	48

## Index des figures

Figure 1. Un arbre non enraciné (a) et les arbres enracinés correspondants (b) selon l'emplacement la racine.....	12
Figure 2 Modèle de population constante et (b) modèle de population de croissance exponentielle. Le temps entre les événements de coalescence est inversement proportionnel à la taille de la population. ....	14
Figure 3. Carte du nombre de souches disponibles par commune dans la zone sélectionnée en Dordogne/Haute Vienne. Les deux départements sont représentés ainsi que l'ensemble de ses communes. Les polygones gris représentent les communes qui ne sont pas incluses dans la zone sélectionnée. Le nombre de souches disponible correspond uniquement aux souches retenues après les analyses de qualité. Les communes de Saint-Saud-Lacoussière et Villars sont identifiées par une étoile bleue et une étoile verte, respectivement.....	21
Figure 4. Arbre de maximum parcimonie non enraciné construit avec Bionumerics. Chaque nœud représente une séquence. La taille des nœuds est proportionnelle au nombre de souches partageant la même séquence génomique. Les couleurs désignent l'espèce hôte. L'arbre représente les 300 souches incluses ainsi que la 10 <sup>ème</sup> souche exclue (étoile noire).....	22
Figure 5. Extrait de la matrice de SNP générée par Bionumerics. Chaque ligne représente une souche. Chaque colonne qui suit représente un site du génome où il y a au moins une mutation pour au moins une souche. Les couleurs désignent les 4 nucléotides (A, T, C, G). Les SNP sont représentés par les nucléotides divergents.....	23
Figure 6. Arbre phylogénétique circulaire de maximum crédibilité de clade (« Maximum credibility clade tree »). La racine (R) représente l'ancêtre commun le plus récent et est placé au centre de l'arbre. Les losanges représentent les nœuds hypothétiques et les cercles représentent les souches séquencées. Les couleurs désignent l'espèce hôte. Le zoom met en évidence la transition de la mycobactérie de blaireaux vers un bovin. Les grands cercles précisent l'échelle temporelle. ....	27
Figure 7. Arbre phylogénétique consensus circulaire divisé par clade. Chaque couleur et chaque numéro représentent un clade. Les valeurs de probabilité postérieures sont représentées à l'intérieur de l'arbre. ....	28
Figure 8. Distribution spatiale des souches de chaque clade. Les formes représentent les différentes espèces et sont placées sur les communes où les souches ont été collectées. Les chiffres à l'intérieur des formes indiquent le nombre de souches provenant d'une même commune. Chaque clade est représenté par une couleur différente. ....	31
Figure 9. Distribution spatiale des souches de chaque clade (suite). Les formes représentent les différentes espèces et sont placées sur les communes où les souches ont été collectées. Les chiffres à l'intérieur des formes indiquent le nombre de souches provenant d'une même commune. Chaque clade est représenté par une couleur différente. ....	32

## Index des tableaux

Tableau 1. Interprétation du BF .....	19
Tableau 2. Espèce hôte des souches séquencées et année de l'isolement. ....	21
Tableau 3. Les modèles testés sur BEAST2, les détails concernant la convergence et l'ESS .....	24
Tableau 4. Le log BF calculé pour chaque paire de modèles. Pour chaque comparaison, le modèle favorisé est mis en gras et est utilisé pour le calcul suivant. +++ indique un niveau de preuve décisif	25
Tableau 5. Les paramètres d'intérêt estimés par BEAST.....	26
Tableau 6. Distribution des souches par clade selon l'espèce et selon l'année d'isolement. ....	29
Tableau 7. Les distances intra-clades en km: minimum, moyenne et maximum. ....	33
Tableau 8. Statistique de la moyenne de la distance intra-clade.....	33
Tableau 9. Résultats du test de Mantel pour comparer les distances géographiques et génétiques intra-clades. ....	34

## Résumé court

La détection des foyers de tuberculose bovine, une maladie entraînant des impacts économiques importants, augmentent régulièrement dans le Sud-Ouest de la France, les cas étant détectés chez les bovins et chez la faune sauvage. Pour explorer le rôle de la faune sauvage dans la transmission de la tuberculose bovine, nous avons utilisé des données de séquençage complet du génome de 300 souches de *Mycobacterium bovis* isolées dans le cadre de la surveillance en Dordogne et Haute-Vienne de bovins, blaireaux, sangliers, renards, cerfs et un chevreuil de 2001 à 2017. L'inférence Bayésienne a été utilisée pour estimer les relations phylogénétiques entre les souches et estimer le taux de mutation et de transmission inter-espèce. La distribution spatiale des clades obtenus a été explorée. Le taux de mutation estimé (0,28 mutations/génome/an ; IC<sub>95%</sub> [0,24 ; 0,34]) est cohérent avec ceux trouvés dans des autres études. Nous avons prédit que la plupart des transmissions a lieu entre bovins et que la plupart des transmissions inter-espèces a lieu des bovins vers la faune sauvage, les sauts inter-espèces de la faune sauvage vers les bovins n'étant pas fréquents.

## Résumé long

La tuberculose bovine est une maladie zoonotique à progression lente dont l'agent pathogène, *Mycobacterium bovis*, peut infecter plusieurs mammifères domestiques et sauvages (Corner, 2006), bien que les bovins en soient le principal réservoir. La transmission a principalement lieu par voie respiratoire.

La France est officiellement indemne depuis 2001, malgré la persistance ou réémergence de la tuberculose bovine dans plusieurs zones. Depuis 2014, le nombre de foyers a augmenté, notamment en Côte d'Or, dans les Ardennes, et dans deux zones du Sud-Ouest (Pyrénées-Atlantiques/Landes et Dordogne/Charentes/Haute-Vienne), l'infection étant identifiée chez les bovins et chez la faune sauvage (sangliers, blaireaux, renards, chevreuil et cerfs). La tuberculose bovine en France est surveillée dans les élevages et dans la faune sauvage (sylvatub), de façon événementielle et programmée, avec un rythme basé sur le risque local.

Plusieurs méthodes de typage moléculaire sont intégrées dans la surveillance pour différencier des souches de *M. bovis*, comme le spoligotypage (Kamerbeek et al., 1997) et le typage MLVA (Roring et al., 2002) qui analysent des régions définies du génome et permettent l'identification de l'origine d'un grand nombre de foyers (Hauer et al., 2015). Cependant, dans les zones à forte prévalence les souches isolées appartiennent à un nombre limité de génotypes. Le séquençage complet du génome (WGS) suivi de l'identification des modifications ponctuelles du génome (SNP), maintenant moins cher et moins complexe, permet une résolution plus fine.

L'épidémiologie moléculaire associe les données génomiques et épidémiologiques pour évaluer les relations phylogénétiques entre les individus échantillonnés d'une population. L'inférence phylogénétique demande souvent des modèles évolutifs, généralement composés de modèles de substitution de nucléotides et de variation du taux d'évolution par site du génome. Ce modèle de base peut être complété de modèles de variation du taux d'évolution par branche de l'arbre (horloge moléculaire) et de modèles de variation de la structure de la population au cours du temps, développés plus récemment, et qui permettent la génération d'arbres phylogénétiques avec l'intégration d'une dimension temporelle (Drummond et Rambaut, 2007).

En France, où la tuberculose bovine fait face à une expansion des zones d'infection à *M. bovis* dans le Sud-Ouest, avec des foyers détectés chez les bovins et dans la faune sauvage, l'épidémiologie moléculaire se présente comme un outil pour éclairer la dynamique des contacts infectieux entre les différentes espèces. Des études ont été conduites en Irlande du Nord et en Nouvelle Zélande et ont permis d'explorer la transmission inter-espèce de la maladie (Biek et al., 2012; Crispell et al., 2017; Orloski et al., 2018; Trewby et al., 2016).

Cette étude a pour objectif l'exploration du rôle de la faune sauvage dans la transmission de la tuberculose bovine, à partir de l'inférence phylogénétique, et s'est focalisée sur les souches isolées dans une zone sélectionnée en Dordogne et Haute-Vienne.

La zone sélectionnée comprend 78 communes et il s'agit d'une zone isolée, dont 95% de la limite est partagée avec des communes non infectées. Dans cette zone, 310 souches de *M. bovis* provenant de 244 bovins, 41 blaireaux, 17 sangliers, 5 renards, 2 cerfs et un chevreuil ont été collectées entre 2001 et 2017. Ces souches ont été cultivées au LNR pour en extraire l'ADN à séquencer. La qualité du séquençage a été contrôlée et les séquences ont été alignées avec une séquence de *M. bovis* isolée d'un bovin en 2014 en France en Côte-d'Or. Des SNP ont été identifiés et utilisés pour l'inférence

phylogénétique, réalisée par inférence bayésienne avec un modèle évolutif sélectionné parmi plusieurs modèles testés.

L'inférence phylogénétique a permis d'obtenir les distributions *a posteriori* des paramètres du modèle et un arbre phylogénétique consensus, ainsi que la prédiction des espèces hôtes des ancêtres hypothétiques. L'arbre a été divisé par clades et leur distribution spatiale a été explorée.

Suite aux analyses de qualité, 300 souches ont pu être gardées et 366 SNP ont été identifiés. Le modèle sélectionné est composé du modèle de substitution HKY, du modèle d'horloge moléculaire stricte et du modèle de population avec un taux de croissance exponentielle.

L'inférence phylogénétique a généré un arbre d'une hauteur de 45 ans, ce qui suggère que l'ancêtre le plus récent commun aux souches circulait en 1972, et provenait d'un bovin (probabilité de 0,81). Le taux de substitution estimé est de 0,29 substitutions/site/an. Les taux de transmission inter-espèces ont été estimés pour chaque paire d'espèces de manière symétrique (mêmes valeurs pour les deux directions de transmission dans chaque paire). La plupart des valeurs sont similaires, variant entre 0,07 et 0,10 transitions/individu/an. Le taux de transitions est moins élevé pour les transmissions entre bovin et sanglier (0,04 transitions/individu/an) et plus élevé pour celles entre sangliers et renards (0,17 transitions/individu/an).

La phylogénie générée prédit que *M. bovis* circule la plupart du temps entre bovins. Après la transmission de *M. bovis* de bovins vers la faune sauvage, *M. bovis* circule majoritairement au sein de la faune sauvage, principalement entre individus de la même espèce. Une seule transmission de la faune sauvage (blaireau) vers des bovins a été prédite. Avant cet événement, *M. bovis* avait circulé et évolué pendant 6 ans chez des blaireaux.

Toutes les souches provenant de blaireaux ont un bovin comme ancêtre hypothétique, ainsi que toutes les souches provenant de sangliers et celle provenant du chevreuil. Quatre des 5 souches isolées chez des renards ont un sanglier comme ancêtre hypothétique, la cinquième proviendrait d'un bovin. Parmi les 2 souches provenant de cerfs, une a un bovin comme ancêtre hypothétique et l'autre un blaireau. Il est possible d'observer dans l'arbre que certaines souches sont temporellement éloignées de leur ancêtre le plus récent. Cette situation arrive plus souvent dans la faune sauvage que chez les bovins. Ce saut temporel indique probablement que des souches intermédiaires n'ont pas été détectées et échantillonnées.

Onze clades ont été identifiées, contenant de 8 à 48 souches. La plupart des clades est bien définie géographiquement et les distances moyennes intra-clade varient de  $1,53 \pm 3,93$  km à  $16,65 \pm 9,61$  km. La distance intra-clade moyenne est de  $11,53 \pm 5,32$  km.

Pour l'ensemble des souches de l'étude, une corrélation positive et significative ( $p=0.0001$ ) a été identifiée entre distance génétique et distance géographique, par le test de Mantel.

Pour la première fois des données de séquençage complet du génome ont été utilisées pour étudier l'épidémiologie moléculaire de la tuberculose bovine en Dordogne et Haute-Vienne. Le nombre élevé de souches disponibles, la diversité des espèces hôtes testées et la grande profondeur temporelle constituent un jeu de données informatif pour l'inférence phylogénétique.

Le taux de mutation de *M. bovis* estimé dans cette étude est de 0,28 mutations/génome/an ( $IC_{95\%}$  [0,24 ; 0,34]). Cette valeur est cohérente avec celles trouvées dans les autres études. Trewby *et al.* (2015) ont estimé un taux de 0,2 mutations/génome/an ( $IC_{95\%}$  [0,1 ; 0,3]), alors que Crispell *et al.* (2017) ont estimé 0,53 mutations/génome/an ( $IC_{95\%}$  [0,22 ; 0,94]). Néanmoins, l'étude actuelle a permis de d'estimer un taux de substitution plus précis, probablement grâce à la taille plus grande de



l'échantillon et à la profondeur temporelle moins grande. Ce taux de mutation de *M. bovis* a également été estimé pour une autre région en France, la Côte-d'Or à 0,29 mutations/génome/an. Il s'agit d'une étude qui n'est pas encore publiée.

Les estimations de taux de transmissions entre espèces doivent être traités avec du recul. En effet, le biais d'échantillonnage, plus important pour la faune sauvage, bien que non négligeable pour les bovins, peut surestimer certains taux (ceux qui impliquent les bovins) et sous-estimer les autres. Dans le modèle utilisé les transmissions ont été considérées symétriques pour chaque paire d'espèces. La direction de transmission entre deux espèces peut avoir des taux divergents et cette différence peut varier au cours du temps. Des modèles supplémentaires avec des taux asymétriques de transmission entre espèces doivent être développés afin de déterminer si une transmission asymétrique serait à privilégier. Les modèles asymétriques demandent l'estimation de plus de paramètres, ce qui peut poser des problèmes d'identifiabilité des paramètres.

Une corrélation positive significative entre les distances géographiques et les distances génétiques est attendue, étant donné que les espèces animales concernées, principalement les bovins et les blaireaux, sont des animaux territoriaux. Si cette corrélation est mise en évidence pour l'ensemble des 300 souches de l'étude, on ne retrouve cette corrélation que pour les clades 4 des 11 clades. Pour les autres clades, l'absence de corrélation pourrait s'expliquer par la petite taille des clades ou l'existence de pâtures rattachées à une exploitation donnée mais situées dans une autre commune. Pour explorer la présence des souches chez des bovins géographiquement éloignées des autres d'un même clade, il serait également intéressant d'analyser les mouvements de ces bovins.

Le cas des communes comme Villars et Saint-Saud-Lacoussière avec nombre important et croissant de APDI, suggère l'importance d'évaluer une possible lacune des mesures de dépistage et de biosécurité implantées.

En conclusion, *M. bovis* circule davantage chez les bovins et les sauts de la faune sauvage vers les bovins ne sont pas fréquents. En considérant notre jeu de données, la faune sauvage ne semble pas avoir un rôle important dans la transmission et le maintien de l'infection chez les bovins.

Les prochaines étapes incluent la comparaison de ces résultats avec des autres régions en France, une analyse temporelle pour explorer comment la maladie progresse au cours du temps et pour comparer les distances génétiques avec les distances temporelles, l'analyse des territoires occupés par la faune sauvage et ses mouvements et la modélisation de chaînes de transmission.

# Synthèse bibliographique

## 1 La tuberculose bovine

### 1.1 Pathologie et épidémiologie

La tuberculose bovine est une maladie zoonotique à progression lente dont l'agent pathogène, *Mycobacterium bovis*, peut infecter plusieurs mammifères domestiques et sauvages (Corner, 2006). Les bovins en sont le principal réservoir. *M. bovis* est une mycobactérie à croissance lente membre du complexe *Mycobacterium tuberculosis* et partage 99,95% de son génome avec l'agent pathogène de la tuberculose humaine (Garnier et al., 2003).

La transmission de la tuberculose bovine se produit principalement par contact direct, suite à l'inhalation d'aérosols. L'ingestion ou le léchage de matières contaminées contribuent plus rarement à la transmission (Neill et al., 2001). Etant donné que la contamination a principalement lieu par voie respiratoire, les lésions se trouvent le plus souvent dans les nœuds lymphatiques de la cage thoracique. Les lésions sont souvent petites et des signes cliniques associés sont frustes dans la plupart des cas (McIlroy et al., 1986).

Les facteurs de risque d'infection par *M. bovis* pour les bovins incluent les mouvements de bovins (Bessell et al., 2012; Johnston et al., 2011), les élevages de grande taille et l'exploitation extensive (Martínez-López et al., 2014), le voisinage au pâturage (Johnston et al., 2011; White et al., 2013) et la présence de la faune sauvage infectée (Bessell et al., 2012; Martínez-López et al., 2014). Dans les zones à forte prévalence de tuberculose bovine en France, les facteurs de risques trouvés sont les mêmes (Marsot et al., 2016).

### 1.2 La surveillance de la tuberculose bovine en France

La tuberculose bovine est inscrite sur la liste des maladies réglementées à déclaration obligatoire de l'Organisation mondiale de la santé animale (OIE) et est classée danger sanitaire de première catégorie en France (JORF n°0187 du 13 août 2013 page 13832 texte n° 112). Selon le Code sanitaire pour les animaux terrestres de l'OIE (OIE, 2018), pour qu'un pays soit qualifié indemne de tuberculose bovine, un programme de surveillance reposant sur des inspections *ante mortem* et *post mortem* dans les abattoirs et sur le dépistage régulier des troupeaux doit être mis en œuvre depuis trois ans, permettant de démontrer, au cours de cette même période, l'absence d'infection par le complexe *M. tuberculosis* chez au moins 99,8 % des troupeaux, représentant au moins 99,9 % des bovins du pays. Afin de conserver le statut indemne, le taux d'infection des troupeaux doit rester au-dessous de 0,1%.

La surveillance et la lutte contre la tuberculose bovine en France sont motivées par son caractère zoonotique et ses impacts économiques liés à la baisse de production des cheptels ainsi qu'aux entraves au commerce des animaux et de leurs produits aux niveaux européen et international. Il est donc important de maintenir un statut indemne. La France est officiellement indemne depuis 2001, malgré la persistance ou réémergence de la tuberculose bovine dans plusieurs zones. Depuis 2014, le nombre de foyers a augmenté d'une cinquantaine à une centaine par an. Cette augmentation est plus importante dans certaines zones : en Côte d'Or, dans les Ardennes, et dans deux zones du Sud-Ouest (Pyrénées-Atlantiques/Landes et Dordogne/Charentes/Haute-Vienne). L'objectif à court terme de la surveillance est la détection des foyers et le suivi du niveau de prévalence. L'objectif à long terme est l'éradication de la maladie (Pandolfi et al., 2018).

Les foyers détectés en France ne concernent pas uniquement les bovins. L'infection a également été identifiée chez certaines espèces de la faune sauvage comme les sangliers, les blaireaux, les renards et les cerfs. La tuberculose bovine a été détectée chez des cerfs élaphe pour la première fois en 2001, en forêt de Brotonne en Normandie (Zanella et al., 2008). La présence de la faune sauvage infectée

rend le contrôle de la maladie plus difficile. En 2011, un dispositif national de surveillance de la tuberculose bovine dans la faune sauvage (Sylvatub) a été mis en place, pour la détection et le suivi du taux d'infection apparent et pour mieux connaître les interactions entre faune sauvage et faune domestique (Note de service DGAL/SDSPA N2011-8214).

Le système de surveillance de la tuberculose bovine dans les élevages français et Sylvatub sont étroitement liés. La surveillance reposant sur les élevages prévoit un rythme de dépistage qui diffère entre les départements en fonction de la prévalence chez les bovins et dans la faune sauvage. Les campagnes de dépistage sont réalisées généralement d'octobre à avril par intradermo-tuberculation (IDT) simple (IDS) ou comparative (IDC). Le test est fait par l'injection de tuberculine (tuberculine bovine normale ou forte pour l'IDS ou tuberculine normale et tuberculine aviaire pour l'IDC) suivi de l'inspection du lieu d'application 72 heures après. L'occurrence d'une réaction par hypersensibilité retardée suggère que l'animal a déjà été exposé à une mycobactérie. L'interprétation correcte du résultat dépend du suivi précis du protocole de test. Etant donné que la sensibilité (Se) et la spécificité (Sp) des tests (IDS : Se [80%-91%]<sub>95%</sub> et Sp [75%-99,9%]<sub>95%</sub> et IDC : Se [55%-93%]<sub>95%</sub> et Sp [89%-100%]<sub>95%</sub> (Fediaevsky et al., 2012) peuvent être assez basses, une combinaison des deux et un recontrôle par dosage de l'interféron gamma suite à un résultat non-négatif peuvent être adoptés (Note de service DGAL/SDSPA/2016-1001).

En ce que concerne les inspections *post mortem* des bovins à l'abattoir, dans le cas d'une suspicion, l'organe ou le nœud lymphatique lésé doivent être prélevés et envoyés à un laboratoire agréé pour la réalisation de la PCR, l'histologie et la culture (Note de service DGAL/SDSPA/N2013-8202). La PCR, basée sur la détection de l'ADN possède une sensibilité plus élevée que la culture bactérienne (PCR : 87,7% [82,5-92,5 %]<sub>95%</sub> ; culture : 78,1 % [72,9-82,8%]<sub>95%</sub>) et toutes les 2 ont une spécificité élevée (PCR : 97,0 % [94,3-99,0%]<sub>95%</sub> ; culture : 99,1 % [97,1-100,0%]<sub>95%</sub>). La PCR a l'avantage de confirmer l'infection plus rapidement que la culture (Courcoul et al., 2014).

Suite à l'identification d'un troupeau infecté, caractérisé par la présence de signes cliniques associés à une réaction positive par IDT ou l'isolement et identification de *Mycobacterium sp.* ou une réaction d'IDT comparative positive associée à l'observation de lésions histologiques ou un résultat de PCR positive, un Arrêté Préfectoral portant Déclaration d'Infection (APDI) est mis en place. L'APDI entraîne des limitations de mouvements vers et en provenance de l'élevage infecté et un abattage total ou partiel des animaux. L'annexe 1 rappelle l'ensemble des mesures de police sanitaire appliquées selon statut du troupeau de bovins.

Le système de surveillance Sylvatub repose sur des composantes événementielles et programmées et dépend du risque local. Pour la composante événementielle, des lésions évocatrices de tuberculose sont recherchées lors de l'examen de carcasses par les chasseurs et dans le cadre du réseau Sagir, qui analyse les causes de mortalité des animaux sauvages trouvés morts ou mourants. La composante programmée repose sur la surveillance des blaireaux piégés et des cerfs et sangliers tués à la chasse. Les cas suspects sont confirmés par PCR. Le niveau de surveillance est déterminé par la présence de foyers de tuberculose bovine chez les bovins et chez la faune sauvage ou la proximité géographique avec une zone infectée considérée à haut risque (Réveillaud et al., 2018; Rivière et al., 2013).

## 2 Séquençage du génome

Plusieurs méthodes de typage moléculaire ont été utilisées pour différencier des souches de *M. bovis*. Le spoligotypage est basée sur la présence ou l'absence d'oligonucléotides « espaceurs » dans une région conservée de l'ADN (Kamerbeek et al., 1997). Il s'agit d'une méthode simple et rapide, peu

onéreuse et qui permet le typage à partir d'échantillons facilement disponibles (Roring et al., 2000). Le spoligotypage est une méthode couramment utilisée sur le terrain.

Le typage MLVA (Multi-Locus Variable number tandem repeat Analysis) détecte le nombre de répétitions de plusieurs marqueurs de type MIRU-VNTR (Mycobacterial Interspersed Repetitive Unit – Variable Number Tandem Repeat) identifiés dans des régions particulières du génome de la mycobactérie (Roring et al., 2002). Plusieurs loci ont été identifiés mais leur capacité de discrimination est variable selon les groupes clonaux.

En France, l'analyse de 4020 souches de *M. bovis* isolées dans des foyers bovins et chez des animaux sauvages entre 1978 et 2013, a permis l'identification de 176 spoligotypes, dont 3 sont prédominants (SB0120, SB0134 et SB0121), correspondent à 43% des souches. La même étude a identifié 706 génotypes différents après le génotypage combiné par spoligotypage et MLVA, réalisé sur 2332 souches collectées entre 1978 et 2013, ce qui montre la capacité de différenciation plus fine en combinant les deux méthodes (Hauer et al., 2015).

Le spoligotypage et le VNTR permettent l'identification de l'origine d'un grand nombre de foyers et de possibles transmissions inter-espèces (Hauer et al., 2015). Cependant, dans les zones de forte prévalence les souches isolées appartiennent à un nombre limité de génotypes et des techniques moléculaires à très fine résolution sont nécessaires pour retracer les chaînes de transmission (Boschioli et al., 2015). Plus récemment, l'évolution des techniques de séquençage, maintenant moins chères et moins complexes, a rendu possible le séquençage de l'ensemble du génome (whole genome sequencing ou WGS) de pathogènes. Le génome de *M. bovis* a été entièrement séquencé pour la première fois en 2003 à partir d'une souche isolée en Grande Bretagne en 1997 (Garnier et al., 2003). Le WGS permet l'analyse de la séquence complète de chaque souche et l'identification de modifications ponctuelles (polymorphismes d'un seul nucléotide ou SNP pour « single-nucleotide polymorphism »). Cette méthode permet donc une discrimination plus fine qu'avec les méthodes précédemment utilisées (spoligotypage et VNTR).

### 3 Phylogénétique et épidémiologie moléculaire

#### 3.1 Principe

##### 3.1.1 Les arbres phylogénétiques

Un arbre phylogénétique est une représentation schématique des relations évolutives entre les objets étudiés (espèces, individus, gènes...). Ces arbres sont construits par inférence phylogénétique à partir de l'analyse de séquences moléculaires alignées. Les séquences peuvent être de l'ADN, de l'ARN, des protéines ou des codons.

Dans un arbre phylogénétique, les objets étudiés sont représentés par les nœuds terminaux (ou feuilles). Les nœuds internes représentent des ancêtres hypothétiques partagés par les deux lignées qui suivent. Les branches définissent les relations entre les nœuds. La longueur des branches représente, le plus souvent, la distance génétique entre 2 nœuds. Elle peut également représenter l'intervalle de temps après lequel deux lignées divergent ou être arbitraire. La racine d'un arbre phylogénétique représente l'ancêtre le plus récent commun à tous les nœuds (Baum et Offner, 2008). Les arbres phylogénétiques non enracinés illustrent les distances génétiques entre les nœuds terminaux sans faire d'hypothèse sur les ancêtres communs. Ces arbres n'intègrent pas de dimension temporelle, donc ils ne fournissent pas d'information sur la direction de l'évolution, et peuvent générer des différents arbres enracinés (Figure 1).

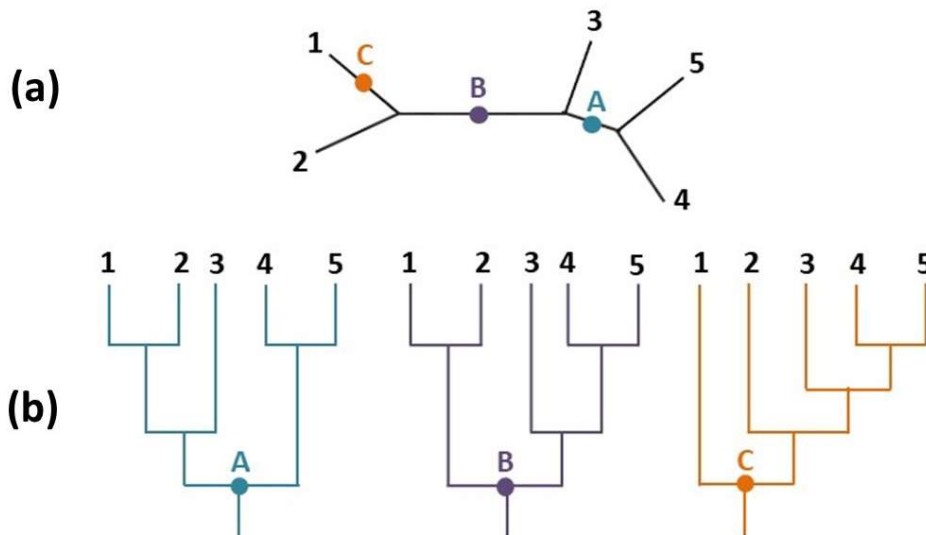


Figure 1. Un arbre non enraciné (a) et les arbres enracinés correspondants (b) selon l'emplacement la racine

Pour inférer les arbres phylogénétiques, on utilise souvent des modèles évolutifs, généralement composés de modèles de substitution de nucléotides et de variation du taux d'évolution par site du génome. Ce modèle de base peut être complété de modèles de variation du taux d'évolution par branche de l'arbre (horloge moléculaire) et de modèles de variation de la structure de la population au cours du temps, développés plus récemment, et qui permettent l'intégration d'une dimension temporelle pour générer des arbres enracinés (Drummond et Rambaut, 2007).

### 3.1.2 Les modèles de substitution et de variation du taux d'évolution par site

Les modèles de substitution décrivent les taux de modifications ponctuelles (ou substitutions) de nucléotides, ainsi que la fréquence des bases dans les séquences étudiées. Le modèle de substitution le plus simple, proposé par Jukes et Cantor (JC69) (Jukes et Cantor, 1969), suppose le même taux de substitution entre n'importe quelle paire de nucléotides et la même fréquence pour les 4 bases. Plusieurs modèles ont ensuite été développés avec des taux de substitution différents (K80) (Kimura, 1980), des fréquences de bases différentes (F81) (Felsenstein, 1981) ou les deux (Hasegawa-Kishino-Yano ou HKY85, General time reversible ou GTR) (Hasegawa et al., 1985; Tavaré, 1986). Les différents modèles de substitution sont présentés en détail dans l'annexe 2.

Le taux de substitution peut varier entre les différentes régions (ou site) du génome, ce qui est expliqué par des contraintes de sélection différentes. Par exemple, en raison de pressions sélectives différentes, les gènes codant des protéines hautement exprimées ou avec des fonctions critiques ont tendance à évoluer plus lentement (Echave et al., 2016). La distribution gamma ( $\Gamma$ ) continue ou discrète est largement utilisée pour représenter cette variabilité, en raison de sa capacité à modéliser divers degrés de variation. Une proportion de sites invariables du génome (I) peut également être ajoutée au modèle, pour prendre en compte des sites où des mutations seraient délétères (Yang, 1996).

Les modèles de substitution sont importants pour tenir compte des effets de substitutions multiples sur un même site et qui ne sont pas observable. Si les substitutions multiples sur un même site et la variation du taux d'évolution par site ne sont pas pris en compte, l'estimation de la longueur des branches peut être biaisée, ainsi que les relations phylogénétiques et l'échelle temporelle (Jia et al., 2014).

### 3.2.2 Horloge moléculaire et les modèles de variation du taux d'évolution par branche

L'hypothèse de l'horloge moléculaire suppose que les séquences d'ADN et/ou de protéines évoluent à un taux constant au cours du temps et entre les lignées. Cette supposition implique que la différence génétique entre n'importe quel couple d'espèces ou d'organismes est proportionnelle au temps qui s'est passé depuis qu'il existait un ancêtre commun. Afin d'inférer les temps de divergence entre les nœuds d'un arbre phylogénétique, il conviendrait de supposer un taux d'évolution constant entre les branches, en accord avec l'hypothèse de l'horloge moléculaire. Néanmoins, le taux de substitution peut varier significativement au cours du temps, entre les branches ou les lignées, sous l'influence de contraintes environnementales par exemple. Ces limitations ont conduit au développement de modèles d'horloge moléculaire relâchée (Rutschmann, 2006).

Les modèles qui décrivent une horloge moléculaire relâchée (« relaxed-clock ») permettent que le taux de substitution varie entre les branches d'un même arbre phylogénétique, sans nécessairement attribuer un taux unique pour chaque branche. En fonction des hypothèses concernant la variation du taux de mutation, plusieurs modèles d'horloge relâchée peuvent être utilisés. Ainsi, suivant que l'on considère que les taux de mutations varient au niveau des nœuds internes ou le long des branches, on utilisera une horloge relâchée exponentielle ou lognormale, respectivement. (Aris-Brosou et Yang, 2002; Drummond et al., 2006; Thorne et al., 1998).

### 3.3.3 Les modèles de population

Les modèles de génétique des populations ou modèles de population sont employés pour représenter les relations généalogiques (la topologie de l'arbre phylogénétique et la hauteur des branches) entre un ensemble de séquences génomiques échantillonnées simultanément ou à différentes périodes. La plupart des modèles adaptés à ces échantillons est basée sur la théorie de la coalescence.

La théorie de la coalescence décrit les relations entre l'histoire démographique d'une population et l'ancestralité partagée par des individus tirés au hasard dans cette population et représentées par un arbre phylogénétique (Drummond et al., 2003). Selon la théorie de la coalescence (Kingman, 1982), la généalogie d'un échantillon de  $n$  séquences génomiques tirées d'une population implique  $n-1$  événements coalescents qui remontent dans le temps, et chaque événement réduit de 1 le nombre de lignées. A la fin du processus stochastique, la seule lignée restant après le dernier événement coalescent est l'ancêtre le plus récent commun à toutes les séquences.

Le modèle coalescent fournit une distribution de probabilité de temps entre les événements de coalescence et requiert un modèle démographique qui décrit la variation de la taille de la population effective et le taux de croissance de la population au cours du temps. Par exemple le modèle de population constante (Figure 2a) suppose une taille de population constante, le modèle de croissance exponentielle (Figure 2b) suppose un taux de croissance constant. Le modèle de croissance logistique suppose un taux de croissance qui diminue dans le temps et le modèle coalescent Bayésien skyline estime la variation de la taille de la population effective (Drummond et Rambaut, 2007; Drummond et al., 2006).

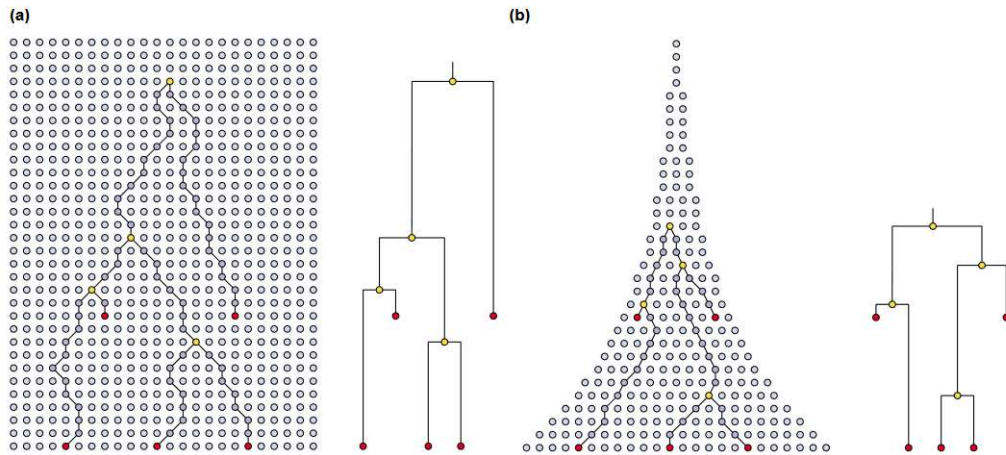


Figure 2 Modèle de population constante et (b) modèle de population de croissance exponentielle. Le temps entre les événements de coalescence est inversement proportionnel à la taille de la population.

### 3.3.4 Traits

Les modèles évolutifs peuvent aussi comprendre des traits, qui représentent des états relatifs aux individus échantillonnés, par exemple la localisation spatiale, l'espèce hôte, la présence de gènes de résistance aux antibiotiques, compartiments cellulaires, etc. Dans ce cas, en plus d'estimer l'arbre phylogénétique il faut estimer l'état des nœuds internes, à partir de l'état des nœuds terminaux et selon le modèle précisé. Le modèle de transition pour les traits peut être symétrique, supposant que les taux de transition entre les états sont réversibles, ou asymétrique (Baele et al., 2016). Dans le cas de la tuberculose bovine, un modèle symétrique supposerait que le taux de transmission de bovin à blaireau et le taux de transmission de blaireaux à bovins sont identiques.

## 4 Inférence phylogénétique

L'inférence phylogénétique peut être faite avec des méthodes basées sur la distance ou sur les caractères. Avec les méthodes basées sur la distance, la distance évolutive est inférée à partir d'une matrice de dissimilarités, obtenue après la comparaison de chaque paire de séquences et l'observation de la proportion de sites différents (SNP). Les méthodes basées sur les caractères comparent simultanément toutes les séquences alignées, en considérant un site à la fois, pour inférer l'arbre phylogénétique. Ces méthodes comprennent le maximum parcimonie, le maximum de vraisemblance et l'inférence Bayésienne (Yang et Rannala, 2012).

Les méthodes basées sur la distance demandent un calcul moins complexe, car ils n'impliquent pas la comparaison de nombreux arbres. Cependant, ces méthodes sont peu efficaces pour des séquences très divergentes et sont sensibles aux lacunes dans l'alignement. Pour ces raisons, ces méthodes sont utiles pour analyser de grands ensembles de données qui comprennent des séquences peu divergentes. (Yang et Rannala, 2012).

Les méthodes basées sur la distance, le maximum de vraisemblance et l'inférence bayésienne demandent un modèle de substitution pour décrire les données, alors que la méthode de maximum parcimonie n'a pas de modèle explicite. Le manque d'hypothèses explicites fait du maximum de parcimonie une méthode qui ne prend pas en compte des effets de substitutions multiples pour un même site. En comparaison de la méthode du maximum de vraisemblance, l'inférence bayésienne est plus flexible lorsqu'il s'agit de modèles complexes et de grands ensembles de données.

#### 4.1 L'approche bayésienne

L'approche bayésienne est actuellement largement utilisée pour estimer la phylogénie des pathogènes à partir de séquences génomiques. Ces méthodes permettent l'intégration des données de séquençage et de connaissances *a priori* sur la distribution des paramètres dans le modèle, avec pour but d'estimer une distribution *a posteriori* des paramètres. Les connaissances *a priori* décrivent la connaissance que l'on a des valeurs du paramètre avant d'avoir observé les données et peuvent être basées sur la littérature, l'avis des experts ou des études précédentes. Un *a priori* peut être très informatif, peu informatif ou non-informatif, ce qui va moduler son influence sur la distribution *a posteriori*. La spécification d'un *priori* trop informatif ou mal ajusté peut biaiser les résultats en diminuant l'importance des données. L'estimation bayésienne repose sur des simulations, des processus itératifs d'échantillonnage à partir de distributions, par exemple l'algorithme stochastique Monte Carlo Markov chaîne (MCMC). Pendant l'échantillonnage par MCMC, les distributions *a posteriori* sont explorées aléatoirement un nombre déterminé de fois et chaque étape de cette exploration fournit un échantillon.

MrBayes (Ronquist et al., 2012), BAli-Phy (Suchard et Redelings, 2006) et BEAST (Drummond et al., 2012) sont des exemples de logiciels pour faire de l'inférence phylogénétique bayésienne. Plus spécifiquement, BEAST permet de développer des modèles phylogénétiques et d'en estimer les paramètres et l'arbre phylogénétique. BEAST se distingue d'autres outils par la possibilité de modéliser des arbres phylogénétiques avec un modèle d'horloge moléculaire et un modèle de population, ce qui génère des arbres phylogénétiques avec une dimension temporelle (Drummond et Rambaut, 2007). BEAST est composé de plusieurs packages pour chaque étape de l'inférence phylogénétique. BEAUti est une interface graphique pour définir les modèles évolutifs. BEAST estime les paramètres du modèle. Tracer permet l'affichage des paramètres estimés et l'évaluation de la convergence. Tree Annotator et Fig Tree sont employés pour l'analyse et la manipulation des arbres phylogénétiques.

#### 4.2 Sélection de modèle

Une fois les différents modèles évolutifs ajustés, il est nécessaire de sélectionner le modèle le mieux ajusté. Seuls les modèles qui convergent correctement et avec une taille d'échantillon effective (« effective sample size » - ESS) suffisant sont considérés pour la sélection. L'ESS d'un paramètre échantillonné par une MCMC est le nombre de tirages effectivement indépendants réalisés par la chaîne. Si l'ESS est trop petite, l'estimation de la distribution *a posteriori* du paramètre n'est pas suffisamment informative (Lenth, 2001).

Plusieurs critères sont disponibles pour comparer deux modèles, suivant la méthode d'estimation utilisée, par exemple le test de ratio de vraisemblance (LRT), le critère d'information d'Akaike (AIC), le critère d'information bayésien (BIC) et le facteur de Bayes (Bayes Factor, BF). Le BF ( $BF = VM1/VM2$ ) ou ( $\log BF = \log VM1 - \log VM2$ ) est largement utilisé dans le cadre de l'inférence bayésienne et correspond au ratio des vraisemblances marginales (VM), une tâche qui demande des calculs intensifs mais qui s'affranchit des poids *a priori* des deux modèles testés (Maturana et al., 2019). BEAST fournit un algorithme implémenté pour estimer la vraisemblance marginale, ce qui permet de calculer le BF.

## 5 Epidémiologie moléculaire

Dans le domaine des maladies infectieuses, les études d'épidémiologie moléculaire fournissent des enseignements sur les causes, la pathogénie, la circulation, la transmission, la prévention et la thérapie (Eyboosh et al., 2017). Les séquences génétiques peuvent être utilisées pour identifier des groupes



(ou clades) d'individus infectés par des pathogènes proches génétiquement. Cette information peut être associée aux données épidémiologiques, et ainsi aider à la traçabilité des sources d'infection.

Des études ont été conduites sur l'épidémiologie moléculaire de *M. bovis* en utilisant des données de séquençage total du génome et ont permis de mieux caractériser les liens entre bovins et faune sauvage. En Irlande du Nord, le séquençage de 30 souches isolées d'un foyer a montré une proximité génotypique entre les souches isolées des bovins et des blaireaux, la persistance de certaines souches dans un élevage et l'existence d'une corrélation entre proximité génotypique et distribution spatiale. Toutefois, les résultats suggèrent que la technique est plus adaptée à une échelle inter-cheptel qu'à une échelle intra-cheptel, à cause de l'évolution lente de la mycobactérie (Biek et al., 2012). Cette étude a été étendue pour inclure plusieurs souches isolées au cours de plusieurs années. Encore une fois une corrélation entre proximité génotypique et distribution spatiale a été montrée à une échelle inter-cheptel. Néanmoins, selon les auteurs, il demeure incertain si le séquençage total du génome de *M. bovis* peut fournir une résolution suffisante pour caractériser le rôle de la faune sauvage sous-échantillonnée dans la transmission de la maladie et la direction de la transmission, surtout dû à l'évolution lente du pathogène (Trewby et al., 2016).

Une étude menée en Nouvelle Zélande, avec le séquençage de 168 souches de *M. bovis* a suggéré que la tuberculose bovine circulait entre les bovins et la faune sauvage et que la faune sauvage avait un rôle potentiel de réservoir de l'infection, malgré l'impossibilité de déterminer la direction de la transmission (Crispell et al., 2017). Dans le même pays, il a été montré que le séquençage complet du génome de *M. bovis* permet de mieux identifier la source des nouvelles infections en discriminant des sources potentielles (Price-Carter et al., 2018).

Aux Etats-Unis, après le séquençage complet du génome de 1248 souches de *M. bovis*, une vaste diversité entre les souches isolées de bovins et celles isolées de cervidés domestique a été constatée. Les résultats suggèrent que la transmission se passe surtout entre les cheptels de bovins (Orloski et al., 2018).

Enfin, en ce qui concerne la tuberculose bovine, avec sa longue période d'incubation, la traçabilité des contacts infectieux est difficile et l'épidémiologie moléculaire pourrait permettre d'estimer et reconstruire les chaînes de transmission, ce que a été fait pour d'autres maladies (Campbell et al., 2019; Mollentze et al., 2014; Morelli et al., 2012).

## TRAVAIL PERSONNEL

### 1 Contexte

La tuberculose bovine fait face à une expansion régulière des zones d'infection à *M. bovis* dans le Sud-Ouest de la France, notamment dans les départements de la Dordogne, de la Haute-Vienne et de la Charente. Les foyers sont détectés chez les bovins, mais aussi chez la faune sauvage. La question du rôle de la faune sauvage dans le maintien de l'infection et l'augmentation du nombre de cas chez les bovins se pose dès lors.

L'inférence phylogénétique à partir du séquençage complet du génome de *M. bovis* commence à être utilisée en France en intégrant des données épidémiologiques. L'épidémiologie moléculaire se

présente donc comme un outil potentiel pour éclairer la dynamique des contacts infectieux entre bovins et entre bovins et faune sauvage.

Cette étude a comme objectif l'exploration du rôle de la faune sauvage dans la transmission de la tuberculose bovine, à partir de l'inférence phylogénétique, et s'est focalisée sur les séquences isolées d'une zone sélectionnée en Dordogne et en Haute-Vienne.

## 2 Matériels et Méthodes

### 2.1 Sélection des souches

Les souches de *M. bovis* analysées dans cette étude avaient été choisies dans la souchothèque du Laboratoire National de Référence (LNR) de la tuberculose bovine. Elles proviennent de foyers de tuberculose bovine déclarés en Dordogne et en Haute Vienne entre 2001 et 2017, dans le cadre de la surveillance de la maladie en France. Les souches avaient été isolées chez des bovins et chez la faune sauvage (blaireau, sanglier, chevreuil, cerf et renard). L'étude fait partie d'un projet qui s'est focalisé initialement sur 3 régions à forte prévalence : Ardennes, Pyrénées Atlantiques/Landes et Dordogne/Haute-Vienne. Toutes les souches avaient été typées précédemment par spoligotypage et profil VNTR.

La totalité des souches prélevées sur la faune sauvage et un maximum de 3 souches par élevage et par APDI (Arrêté Préfectoral de Déclaration d'Infection) avaient été incluses dans cette étude. Le nombre maximal des souches pouvant être analysées pour le projet qui englobe cette étude étaient de 576. Deux-cent-soixante souches de la zone précédemment étudiée en Pyrénées Atlantiques/Landes (Bouchez-Zacria et al., 2018) et des Ardennes ont été incluses dans le projet final. En conséquence, l'inclusion de 318 souches était possible pour la région Dordogne/Haute Vienne. La sélection de la zone d'étude pour cette région a été réalisée en suivant l'algorithme stochastique suivant: 1. Sélection de la zone initiale ; 2. à chaque itération, ajout d'une commune limitrophe à la zone initiale choisie au hasard et définissant la zone présélectionnée ; 3. évaluation de la zone présélectionnée ; 4. Arrêt lorsque le nombre maximal d'échantillons est atteint. La zone initiale correspond à Villars et aux communes limitrophes, en raison du nombre important d'APDI (n=8 entre 2004 et 2017, concernant 21 bovins dans 6 exploitations dont 2 exploitations avec 2 APDI) et des cas observés dans la faune sauvage (2 sangliers, 3 blaireaux, 4 renards). Pour qu'une commune soit incluse, au moins une souche devait être disponible et la commune ne devait pas être traversée par une autoroute (ici A20 et A89), étant donné que les animaux ne la traversent pas. Mille zones d'études ont été simulées. Les critères de sélection pour la zone d'étude finale étaient une délimitation maximale (définie comme le ratio maximal :  $\frac{\text{longueur des frontières avec des communes indemnes}}{\text{longueur totale des frontières}}$ ), l'inclusion de la commune de Saint-Saud-Lacoussière (14 APDI dont 8 en 2017, concernant 55 bovins dans 13 exploitations).

### 2.2 Séquençage complet des souches

Les souches sélectionnées ont été cultivées au LNR pour en extraire l'ADN. Elles ont été séquencées à l'Institut du cerveau et de la moelle épinière (ICM) avec la méthode Illumina, basée sur la réplication et fragmentation du génome et l'amplification de ces fragments ("reads", appelés ici fragments de lecture). La qualité du séquençage a été contrôlée avec FASTQC, en utilisant les scores de qualité par

base, déterminés pour chaque souche lors du séquençage. Le score de qualité par base est calculé à partir de la probabilité d'erreur moyenne de positionnement pour chaque base de chaque fragment de lecture ( $p$ ). Un score de qualité ( $Q$ ) égal à 30 ( $Q = -10 * \log_{10} p$ ) a été fixé comme minimum acceptable, ce qui correspond en moyenne au risque d'erreur de positionnement de base sur un site donné du fragment de lecture de 1/1000.

Le logiciel Bionumerics version 7.6 (Applied Maths) a été employé pour l'alignement et l'assemblage (« mapping ») des fragments de lectures et l'obtention du génome consensus de chaque souche. L'assemblage a été fait en utilisant la séquence de *M. bovis* Mb3601, isolée d'un bovin en 2014 en France en Côte-d'Or, de spoligotype BCG. Son utilisation comme souche de référence n'a pas encore été publiée dans la littérature.

Pour identifier les SNP, des variations d'une seule paire des bases dans une séquence, espacées de plus de 12 paires de bases et avec une couverture minimale de 5 fragments ont été recherchées. Ces critères stricts ont été choisis afin d'obtenir les SNP les plus informatifs. Enfin les SNP avec un nucléotide indéterminé ou ambigu ont été éliminés.

Les souches avec moins de 200 000 fragments alignés pendant l'assemblage n'ont pas été incluses pour la suite des analyses, leur alignement étant considéré de faible qualité. Etant donné les critères stricts appliqués pour la recherche de SNP, les génomes obtenus par des alignements de basse qualité ont un impact important sur l'identification des polymorphismes et entraînent la perte de plusieurs SNP informatifs.

Un premier arbre non enraciné (sans modèle évolutif et sans échelle temporelle) a été construit pour avoir un aperçu de la distance génétique et des relations phylogénétiques entre les souches sélectionnées. Cet arbre a été généré avec Bionumerics par maximum parcimonie.

### 2.3 Modèle évolutif

Pour construire les arbres phylogénétiques, plusieurs modèles évolutifs ont été testés en utilisant BEAUti pour l'annotation de chaque modèle.

Les ajustements de modèles ont été réalisées avec BEAST2, en utilisant 500 000 000 chaînes MCMC. Des échantillons ont été enregistrés toutes les 50 000 chaînes MCMC. Le même nombre d'enregistrements a été adopté pour tous les arbres phylogénétiques générés.

La convergence a été évaluée par inspection visuelle des graphiques pour chaque paramètre en recherchant un aspect de « chenille poilue » (« hairy caterpillar ») observé lorsqu'aucune tendance à des valeurs extrêmes n'est observée et que les MCMC deviennent stationnaires. Les ESS ont également été vérifiés en utilisant un seuil fixé à 200, qui est le seuil communément accepté. Ces analyses diagnostiques sont réalisées avec Tracer.

Les modèles ont été élaborés d'abord en considérant les différents modèles de substitutions de nucléotides disponibles, JC69 (Jukes et Cantor, 1969), HKY85 (Hasegawa et al., 1985) et GTR (Tavaré, 1986), associés avec une horloge moléculaire stricte et une population constante. Par la suite, les modèles d'horloge moléculaire relâchée exponentielle et lognormale ont été évalués, en combinaison avec le modèle de substitution retenu et en gardant le modèle de population constante. Une fois l'horloge moléculaire la mieux ajustée sélectionnée, les modèles de population exponentielle et lognormale ont été testés, en gardant les modèles de substitution et d'horloge retenus.

Pour sélectionner le modèle le mieux ajusté nous avons utilisé le facteur de Bayes (BF), en comparant une paire de modèles à la fois. La vraisemblance marginale (VM) des modèles nécessaire au calcul du BF a été calculée par « Nested Sampling » (Maturana et al., 2019) ce qui fournit aussi une estimation de l'écart type de la VM. Cette méthode basée sur l'algorithme stochastique MCMC demande la définition du nombre de points actifs, qui détermine le nombre de particules gardées à chaque itération de la chaîne. Ce paramètre affecte la précision de l'estimation. La vraisemblance marginale a été estimée pour tous les modèles ayant convergé, avec 2 points actifs. Ensuite le calcul a été fait avec 10 points actifs pour les modèles dont l'écart type ne permettait pas une discrimination de la VM.

L'interprétation du BF a été faite selon le Tableau 1 et en considérant l'écart type estimé.

Tableau 1. Interprétation du BF

LOG(BF)	INTERPRETATION
0 - 0,5	Niveau de preuve faible
0,5 - 1,3	Niveau de preuve substantiel
1,3 - 2,2	Niveau de preuve fort
> 2,2	Niveau de preuve décisif

Tous les modèles testés sont présentés en détail dans l'annexe 3.

## 2.4 Inférence phylogénétique

L'inférence phylogénétique permet d'obtenir les distributions *a posteriori* des paramètres du modèle ainsi que le registre des arbres phylogénétiques générés pendant les chaînes MCMC. Les paramètres estimés sont compilés avec Tracer.

La **hauteur de l'arbre** (« Tree Height ») est estimée en années et indique quand circulait l'ancêtre le plus récent commun à toutes les souches. Le **taux d'horloge** (« Clock Rate ») correspond au taux de substitution de nucléotides par site du génome par an. Le paramètre **Kappa** estime le ratio entre les taux des transitions (substitution entre pyrimidines – T et C) et de transversions (substitution entre purines – A et G) du génome (Hasegawa et al., 1985). Le paramètre **alpha** correspond à la forme de la distribution gamma représentant la variation de taux de substitution par site du génome.

Les **taux relatifs de transitions** (« Relative Geo Rates ») représentent les taux de transmission entre une paire d'espèces. Les taux absolus peuvent être obtenus en les multipliant par le **taux d'horloge du trait** (« traitClockRate ») qui indique le taux moyenne de transition entre espèces (transmissions par individu par an).

La taille de la population au temps  $t$  (en ans) est déterminée par  $N(t) = N_e \exp(-gt)$ , où  $N_e$  est la taille de population effective et  $g$  le **taux de croissance exponentielle** ( $g$ ) (« exponential growth rate »). La **taille de population exponentielle** ( $P$ ) représente la taille actuelle de la population et fournit la taille de population effective quand elle est divisée par le temps de génération en années ( $N_e = P * \text{temps de génération}$ ).

Tree Annotator a été utilisé pour obtenir l'arbre consensus, en trouvant l'arbre avec le produit des probabilités postérieures de tous ses nœuds le plus élevé (Arbre phylogénétique de maximum crédibilité de clade ou « Maximum clade credibility tree »), à partir des arbres enregistrés. FigTree (version 1.4.4) permet la visualisation de l'arbre consensus et des traits estimés pour les nœuds ancêtres, ainsi que les probabilités des états de ces nœuds.

## 2.5 Distribution spatiale des clades

### 2.5.1 Identification des clades

Un clade est un groupe monophylétique qui inclut un ancêtre commun et tous ces descendants. A partir de l'analyse visuelle de l'arbre phylogénétique retenu, des clades ont été déterminés. BEAST fournit des probabilités postérieures de clades qui correspondent à la probabilité que ces clades soient vrais sachant les données utilisées. Ceux qui avaient une probabilité supérieure à 0,60 ont été gardés.

### 2.5.2 Distribution intra-clade

La distribution spatiale des clades a été explorée avec le logiciel R 3.5.2.

La matrice des distances euclidiennes entre chaque isolat d'un même clade a été calculée. Le centroïde de la commune où la souche a été isolée a été utilisé pour le calcul des distances. La distance moyenne intra-clade est la moyenne de cette matrice.

La matrice des distances génétiques a été calculée pour chaque clade et pour l'ensemble du jeu de donnée en utilisant le modèle de substitution Felsenstein (F84) (McGuire et al., 1999) disponible dans le package *ape* et similaire au modèle HKY85.

La corrélation entre la matrice de distances géographiques et la matrice de distances génétiques a été évaluée avec un test de Mantel (avec 10 000 permutations) (package *ecodist*).

## 3 Résultats

### 3.1 Zones d'étude et souches sélectionnées

La zone sélectionnée en Dordogne/Haute Vienne comprend 78 communes (Figure 3). Il s'agit d'une zone isolée, dont 95% de la limite est partagée avec des communes non infectées. Dans cette zone, 310 souches de *M. bovis* provenant de 244 bovins, 41 blaireaux, 17 sangliers, 5 renards, 2 cerfs et un chevreuil ont été collectées entre 2001 et 2017 (Table 1). Les souches provenant de bovins ont été isolées toutes les années sauf en 2002. Les souches provenant de blaireaux et de sangliers ont été isolées toutes les années entre 2010 et 2017, sauf en 2011 pour les sangliers. En ce que concerne les souches provenant des autres espèces, elles ont été isolées en 2015 et 2017 de renards, en 2010 et 2017 de cerfs et en 2011 du chevreuil (Tableau 2).

Les bovins proviennent de 124 exploitations, dont 66 contenaient plus d'un animal infecté. La totalité des 310 souches isolées de ces échantillons sont du spoligotype BCG. Quatre profils VNTR complets ont été identifiés : 5 3 5 3 9 3 5 6, 5 3 5 3 9 4 5 6 et 5 3 5 3 9 4 5 (5+6), 5 3 5 3 9 (2+4) 5 6 ainsi qu'un profil incomplet : - 3 5 3 - - 5 6.

Nombre de souches disponibles:

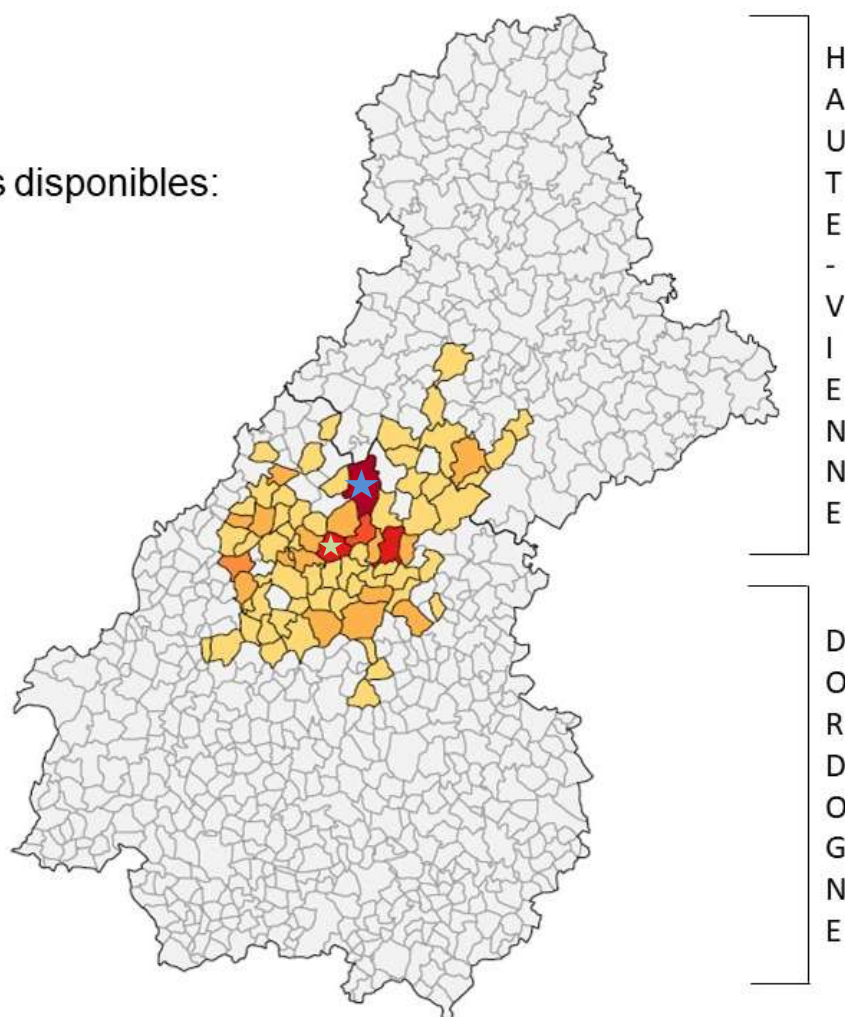


Figure 3. Carte du nombre de souches disponibles par commune dans la zone sélectionnée en Dordogne/Haute Vienne. Les deux départements sont représentés ainsi que l'ensemble de ses communes. Les polygones gris représentent les communes qui ne sont pas incluses dans la zone sélectionnée. Le nombre de souches disponible correspond uniquement aux souches retenues après les analyses de qualité. Les communes de Saint-Saud-Lacoussière et Villars sont identifiées par une étoile bleue et une étoile verte, respectivement.

Tableau 2. Espèce hôte des souches séquencées et année de l'isolement.

ESPECE	ANNEE DE COLLECTE																	TOTAL
	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	
Bovin	3	-	5	2	12	23	13	9	7	4	8	19	26	17	13	22	61	244
Blaireau	-	-	-	-	-	-	-	-	-	1	1	10	2	6	9	7	5	41
Sanglier	-	-	-	-	-	-	-	-	-	3	-	1	4	3	1	2	3	17
Renard	-	-	-	-	-	-	-	-	-	-	-	-	-	-	3	-	2	5
Cerf	-	-	-	-	-	-	-	-	-	1	-	-	-	-	-	-	1	2
Chevreuril	-	-	-	-	-	-	-	-	-	-	1	-	-	-	-	-	-	1
<b>TOTAL</b>	<b>3</b>	<b>0</b>	<b>5</b>	<b>2</b>	<b>12</b>	<b>23</b>	<b>13</b>	<b>9</b>	<b>7</b>	<b>9</b>	<b>10</b>	<b>30</b>	<b>32</b>	<b>26</b>	<b>26</b>	<b>31</b>	<b>72</b>	<b>310</b>

### 3.2 Analyse des données génomiques

Le séquençage a produit des fragments de lecture de 150 paires de bases et de 250 paires de bases. La totalité des données de séquençage avait un score de qualité moyen par base supérieur au seuil fixé de Q30.

Neuf des 310 souches avaient moins de 200 000 fragments de lectures utilisés dans l'alignement final, et ont été éliminées de l'étude. Une dixième souche a été éliminée suite à la construction de l'arbre de maximum de parcimonie non enraciné (Figure 4). L'arbre a montré qu'une souche était très éloignée de toutes les autres, malgré les critères de qualité vérifiés. Il a été constaté que cette souche n'était pas du spoligotype BCG comme les autres, mais du spoligotype F041, et ne provenait pas de la zone d'étude. Au total, 10 souches ont été écartées de l'analyse. Toutes les 10 souches provenaient des bovins, isolées en 2004 (1), 2012 (1), 2014 (1), 2015 (1), 2016 (3) et 2017 (3).

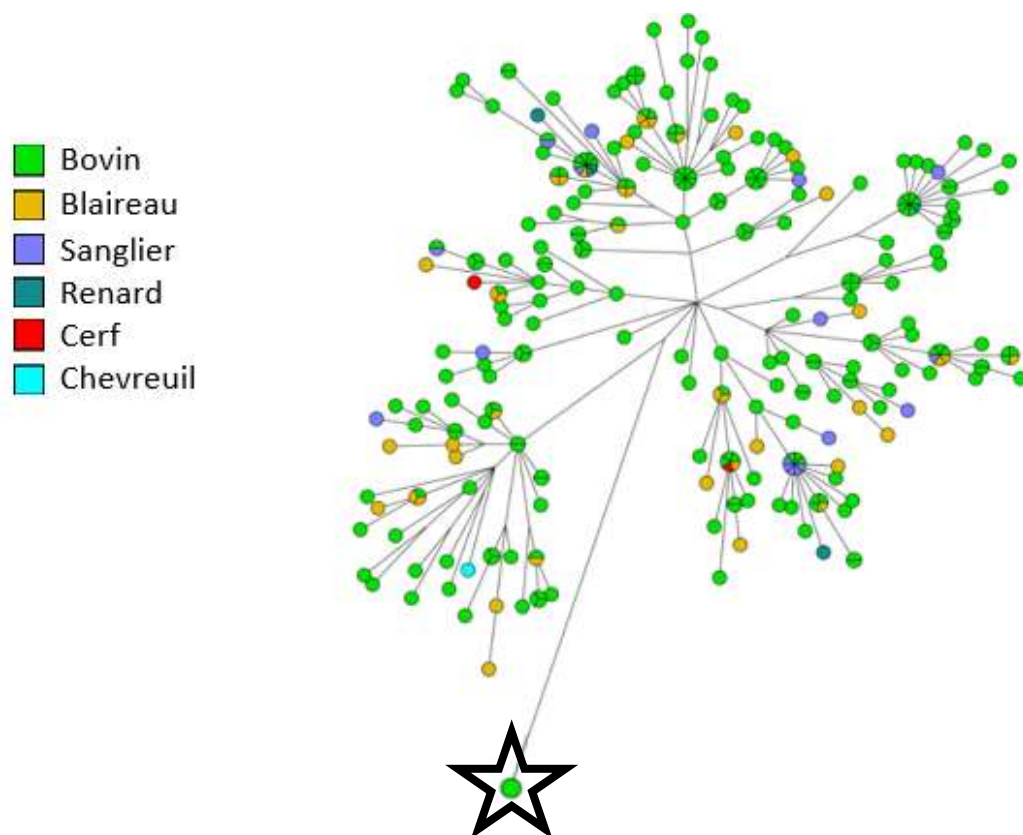


Figure 4. Arbre de maximum parcimonie non enraciné construit avec Bionumerics. Chaque nœud représente une séquence. La taille des nœuds est proportionnelle au nombre de souches partageant la même séquence génomique. Les couleurs désignent l'espèce hôte. L'arbre représente les 300 souches incluses ainsi que la 10<sup>ème</sup> souche exclue (étoile noire)

La recherche de SNP sur les 300 séquences retenues a permis d'identifier un total de 366 SNP. Parmi eux, 154 étaient partagés par  $\geq 2$  souches et 212 étaient SNP ont été identifiés pour une seule souche. La Figure 5 représente un extrait de la matrice de SNP générée par Bionumerics.

SNP matrix																							
	2662707	2665931	2667478	2670212	2676053	2682220	2689772	2720541	2725692	2729038	2776259	2839553	2859861	2937511	2944309	2946018	2954004	2955725	2966952	2977034	3006164	3032645	3037255
15Z006492	T	T	C	G	G	A	C	A	C	T	T	G	A	C	A	C	T	G	A	G	A	G	A
15Z006495	T	T	T	G	G	A	C	A	C	T	T	G	A	C	A	C	T	G	A	G	A	G	A
15Z006541	T	T	T	G	G	A	C	A	C	T	T	G	A	C	A	C	T	G	A	G	A	G	A
15Z006760	T	T	T	G	G	A	T	A	C	T	T	G	A	C	A	C	T	G	A	G	A	G	A
15Z008880	T	T	T	G	G	A	C	A	C	T	T	G	A	C	A	C	T	G	A	G	A	G	A
15Z010210	T	T	T	A	G	A	C	A	C	T	T	G	A	C	A	C	T	G	A	G	A	G	A
15Z010962	T	T	T	G	G	A	C	A	C	T	T	G	A	C	A	C	T	G	A	G	A	G	A
16Z000595	T	T	T	A	G	A	C	A	C	T	T	G	A	C	A	C	T	G	A	G	A	G	A
16Z000964	T	T	T	G	G	A	C	A	C	T	T	G	A	C	A	C	T	A	A	G	A	G	A
16Z000965	T	T	T	G	G	A	C	A	C	T	T	G	A	C	A	C	T	G	A	G	A	G	A
16Z001097	T	T	T	A	G	A	C	A	C	T	T	G	A	C	A	C	T	G	A	G	A	G	A
16Z002887	C	T	T	G	G	A	C	A	C	C	T	G	A	C	A	C	T	G	A	G	A	G	A
16Z004107	T	T	T	G	G	A	C	A	C	T	T	G	A	C	A	C	T	G	A	G	A	G	A
16Z004271	T	T	T	G	G	A	C	A	C	T	T	G	A	C	A	C	T	G	A	G	A	G	A
16Z004272	T	T	T	G	G	A	C	A	C	T	T	G	A	C	A	C	T	G	A	G	C	G	A
16Z004274	T	T	C	G	G	A	C	A	C	T	T	G	A	C	A	C	T	G	A	G	A	G	A
16Z004560	T	T	T	G	G	A	C	A	C	T	T	G	A	C	A	C	T	G	A	G	A	G	A
16Z004566	T	T	C	G	G	A	C	A	C	T	T	G	A	C	A	C	T	G	A	G	A	G	A
16Z004569	T	T	T	G	G	A	C	A	C	T	T	G	A	C	A	C	T	G	A	G	A	G	A
16Z004699	C	T	T	G	G	A	C	A	C	C	T	G	A	C	A	C	T	G	A	G	A	G	A
16Z005204	T	T	T	G	G	A	C	A	C	T	T	G	A	C	A	C	T	G	A	G	A	G	C
16Z005206	T	T	T	G	G	A	C	A	C	T	T	G	A	C	A	C	T	G	A	G	A	G	A
16Z005207	T	T	T	G	G	A	T	A	C	T	T	G	A	C	A	C	T	G	A	G	A	G	A
16Z005209	T	T	T	G	G	A	C	A	C	T	T	G	A	C	A	C	T	G	A	G	A	G	A
16Z005210	T	T	T	G	G	A	C	A	C	T	T	G	A	C	A	C	T	G	A	G	A	G	A
16Z006257	T	T	T	G	G	A	C	A	C	T	T	G	A	C	A	C	T	G	A	G	A	G	A
16Z006707	T	T	T	G	G	A	C	A	C	T	T	G	A	C	A	C	T	G	A	G	A	G	A
16Z007121	T	T	T	G	G	A	C	A	C	T	T	G	A	C	A	C	T	G	A	G	A	G	A
16Z007476	T	T	C	G	G	A	C	A	C	T	T	G	A	C	A	C	T	G	A	G	A	G	A

Figure 5. Extrait de la matrice de SNP générée par Bionumerics. Chaque ligne représente une souche. Chaque colonne qui suit représente un site du génome où il y a au moins une mutation pour au moins une souche. Les couleurs désignent les 4 nucléotides (A, T, C, G). Les SNP sont représentés par les nucléotides divergents.

### 3.3 Modèle sélectionné et analyse phylogénétique

Les modèles testés sont représentés dans le Tableau 3 ainsi que des détails concernant la convergence. Seuls les modèles ayant convergés et avec des ESS >200 ont été gardés pour le calcul du BF. Les vraisemblables marginales estimées sont représentés dans le Tableau 4.



Tableau 3. Les modèles testés sur BEAST2, les détails concernant la convergence et l'ESS

Test	Modèles			Convergence	ESS > 200
	Substitution	Horloge	Population		
1	JC	Stricte	Constante	oui	oui
2	HKY	Stricte	Constante	oui	oui
3	GTR	Stricte	Constante	non	non
4	HKY	Exponentielle	Constante	oui	oui
5	HKY	Log Normale	Constante	oui	oui
6	HKY	Stricte	Exponentielle	oui	oui
7	HKY	Stricte	Skyline	oui	oui

Le modèle GTR (avec horloge moléculaire stricte et modèle de population constante) n'a pas convergé et a été éliminé de l'analyse. La comparaison entre les modèles de substitution de nucléotides JC69 et HKY85, tous les deux avec une horloge moléculaire strict et un modèle de population constante, a montré que le HKY85 est le mieux ajusté ( $\log(\text{BF}) = 186,8$ ). Selon le Tableau 1, le résultat indique un niveau de preuve décisif en faveur de ce modèle.

La comparaison des modèles d'horloge moléculaire (stricte, exponentielle et lognormale) avec un modèle de substitution HKY et un modèle de population constante, permet de conclure que l'horloge moléculaire stricte est la mieux ajustée. Enfin, la comparaison des modèles de population constante, exponentielle et skyline (avec un modèle de substitution HKY et un modèle d'horloge moléculaire stricte) en utilisant le log de la vraisemblance marginale estimée avec 2 points actifs, n'a pas permis une discrimination suffisant entre eux, à cause de l'écart type important (les  $\log(\text{BF})$  étaient inférieures à la somme des écarts types des  $\log(\text{VM})$ ). Le calcul du BF en utilisant la vraisemblance marginale estimée avec 10 points actifs permet de conclure que le modèle de population exponentielle est le mieux ajusté.

Au final, le modèle sélectionné est le modèle HKY/Horloge moléculaire stricte/Population exponentielle. Seuls les résultats issus de ce modèle seront présentés par la suite. Les résultats de calculs du BF sont représentés dans le Tableau 4.

Tableau 4. Le log BF calculé pour chaque paire de modèles. Pour chaque comparaison, le modèle favorisé est mis en gras et est utilisé pour le calcul suivant. +++ indique un niveau de preuve décisif

MODELE 1	MODELE 2	Log VM 1	Log VM 2	SD	Log BF	Niveau de preuve
<b>2 POINTS ACTIFS</b>						
JC Horloge Stricte Pop Constante	<b>HKY</b> <b>Horloge Stricte</b> <b>Pop Constante</b>	-5155	-4968	26	-187	+++
<b>HKY</b> <b>Horloge Stricte</b> <b>Pop Constante</b>	HKY Horloge Exponentielle Pop Constante	-4968	-5100	26	132	+++
<b>HKY</b> <b>Horloge Stricte</b> <b>Pop Constante</b>	HKY LogNormale Pop Constante	-4968	-5029	26	61	+++
HKY Horloge Stricte Pop Constante	<b>HKY</b> <b>Horloge Stricte</b> <b>Pop Exponentielle</b>	-4968	-4929	26	-39	+++
HKY Horloge Stricte Pop Exponentielle	<b>HKY</b> <b>Horloge Stricte</b> <b>Pop Skyline</b>	-4929	-4895	26	-34	+++
<b>10 POINTS ACTIFS</b>						
HKY Horloge Stricte Pop Constante	<b>HKY</b> <b>Horloge Stricte</b> <b>Pop Exponentielle</b>	-4907	-4885	11	-22	+++
<b>HKY</b> <b>Horloge Stricte</b> <b>Pop Exponentielle</b>	HKY Horloge Stricte Pop Skyline	-4885	-5364	11	479	+++

### 3.4 Analyse phylogénétique

La hauteur moyenne de l'arbre phylogénétique est de 45 ans, ce qui suggère que l'ancêtre le plus récent commun aux souches circulait en 1972, et provenait d'un bovin (probabilité de 0,81). Le nombre de substitutions par site du génome par an est d'en moyenne  $7.85 \cdot 10^{-4} \pm 7 \cdot 10^{-5}$ . Etant donné que 366 SNP (ou sites) ont été identifiés, le taux estimé de substitution de 0,29 substitutions/site/an.

Les taux absolus de transition d'état, équivalant aux taux de transmission inter-espèces, ont été estimés pour chaque paire d'espèces (Tableau 5). Les valeurs sont les mêmes pour les taux de transition entre deux espèces (c'est-à-dire que le taux de transition bovin → blaireau est le même que le taux de transition blaireau → bovin), le modèle prédit donc des transitions symétriques. La plupart des valeurs sont similaires, variant de 0,07 à 0,10 transitions/individu/an. Le taux de transitions est moins élevé pour les transmissions entre bovins et sangliers (0,04 transitions/individu/an) et plus élevé pour celles entre sangliers et renards (0,17 transitions/individu/an) (Tableau 5).

Tableau 5. Les paramètres d'intérêt estimés par BEAST.

PARAMETRE		UNITE	MOYENNE	ECART TYPE	IC <sub>95%</sub>	ESS	
Hauteur de l'arbre		ans	45,15	5,85	[34,69 ; 57,05]	5222	
Taux d'horloge		substitutions/site/an	0,00078	0,000070	[0,00065 ; 0,00092]	2632	
Taux de substitution		substitutions/génome/an (taux d'horloge*nb SNP)	0,29	-	[0,24 ; 0,34]	-	
Taux absolus de transition entre les états	Blaireau	Bovin	Transition/individu/an  (taux de transition relatif * taux d'horloge du trait)	0,09	-	-	-
	Blaireau	Cerf		0,07	-	-	-
	Blaireau	Chevreuril		0,08	-	-	-
	Blaireau	Renard		0,08	-	-	-
	Blaireau	Sanglier		0,09	-	-	-
	Bovin	Cerf		0,08	-	-	-
	Bovin	Chevreuril		0,08	-	-	-
	Bovin	Renard		0,08	-	-	-
	Bovin	Sanglier		0,04	-	-	-
	Cerf	Chevreuril		0,10	-	-	-
	Cerf	Renard		0,09	-	-	-
	Cerf	Sanglier		0,10	-	-	-
	Chevreuril	Renard		0,09	-	-	-
	Chevreuril	Sanglier		0,09	-	-	-
Renard	Sanglier	0,17	-	-	-		

L'arbre phylogénétique consensus circulaire et contenant l'espèce hôte estimée pour les ancêtres hypothétiques est représenté dans la Figure 6.

- Blaireau
- Bovin
- Cerf
- Chevreuil
- Renard
- Sanglier

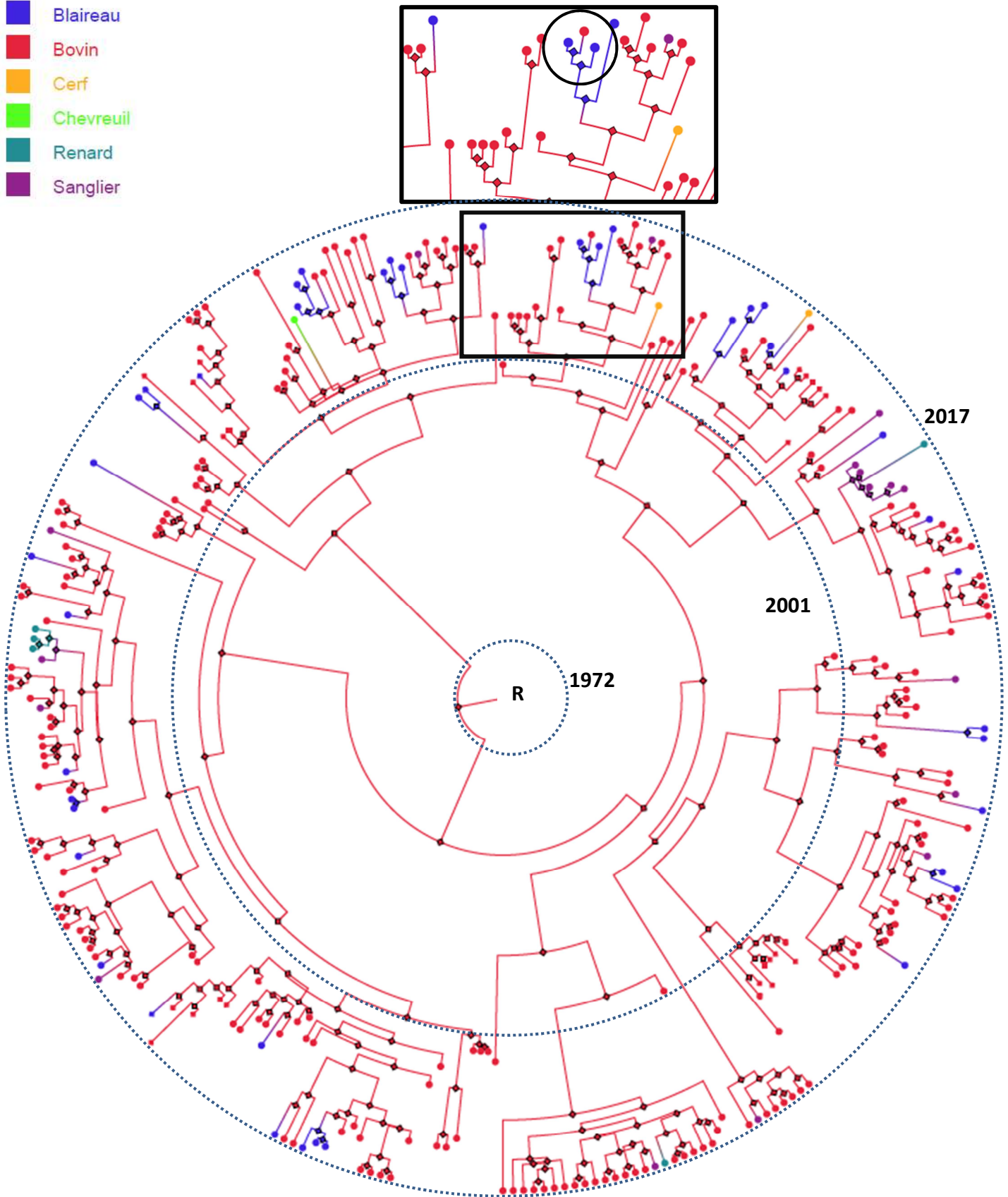


Figure 6. Arbre phylogénétique circulaire de maximum crédibilité de clade (« Maximum credibility clade tree »). La racine (R) représente l'ancêtre commun le plus récent et est placé au centre de l'arbre. Les losanges représentent les nœuds hypothétiques et les cercles représentent les souches séquencées. Les couleurs désignent l'espèce hôte. Le zoom met en évidence la transition de la mycobactérie de blaireaux vers un bovin. Les grands cercles précisent l'échelle temporelle.

La phylogénie générée prédit que l'ancêtre commun le plus récent, représenté par la racine de l'arbre, ainsi que la plupart des nœuds hypothétiques, proviennent de bovins. Après la transmission de *M. bovis* de bovins vers la faune sauvage, *M. bovis* circule majoritairement au sein de la faune sauvage, principalement entre individus de la même espèce. Une seule transmission de la faune sauvage (blaireau) vers des bovins a été prédite. Avant cet événement, *M. bovis* a circulé et évolué pendant 5 années chez des blaireaux.

Toutes les souches provenant de blaireaux ont un bovin comme espèce hôte de l'ancêtre hypothétique, ainsi que toutes les souches provenant de sangliers et celle provenant du chevreuil. Quatre des 5 souches isolées chez des renards ont un sanglier comme espèce hôte de l'ancêtre hypothétique, la cinquième provenant d'un bovin. Parmi les 2 souches provenant de cerfs, une a un bovin comme espèce hôte de l'ancêtre hypothétique et l'autre a un blaireau.

Il est possible d'observer dans l'arbre que certaines souches sont temporellement éloignées de leur ancêtre le plus récent. Cette situation arrive plus souvent dans la faune sauvage que chez les bovins. Ce saut temporel indique probablement que des souches intermédiaires n'ont pas été détectées et échantillonnées.

### 3.5 Distribution spatiale des clades

#### 3.5.1 Identification des clades

A partir de l'arbre consensus, 11 clades ont été identifiés (Figure 7 et Tableau 6). L'annexe 4 contient l'arbre phylogénétique linéaire divisé par clades, avec les probabilités de l'espèce hôte estimés pour les nœuds ancêtres. L'annexe 5 contient la liste des souches de chaque clade, ainsi que l'année et la commune d'isolement.

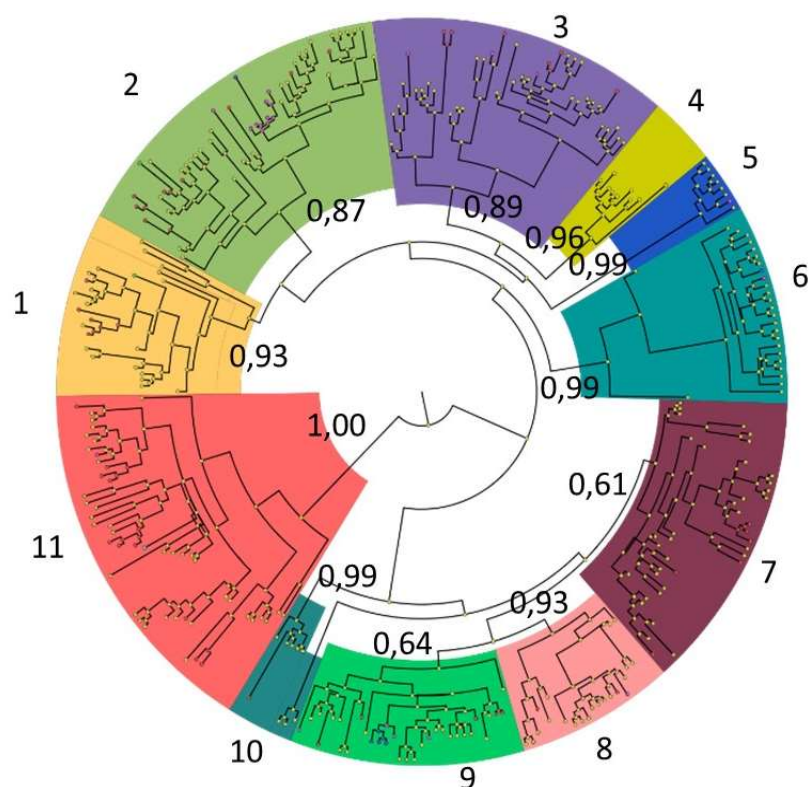


Figure 7. Arbre phylogénétique consensus circulaire divisé par clade. Chaque couleur et chaque numéro représentent un clade. Les valeurs de probabilité postérieures sont représentées à l'intérieur de l'arbre.

Tableau 6. Distribution des souches par clade selon l'espèce et selon l'année d'isolement.

ESPECE	CLADE										
	1	2	3	4	5	6	7	8	9	10	11
Bovin	19	29	31	9	7	24	33	18	20	8	36
Blaireau	3	8	6	-	-	-	6	2	5	1	10
Sanglier	1	6	3	-	1	1	-	1	3	-	1
Renard	-	1	-	-	-	1	-	-	3	-	-
Cerf	1	1	-	-	-	-	-	-	-	-	-
Chevreuil	-	-	-	-	-	-	-	-	-	-	1

ANNE	CLADE										
	1	2	3	4	5	6	7	8	9	10	11
2001	1	-	1	-	-	1	-	-	-	-	-
2002	-	-	-	-	-	-	-	-	-	-	-
2003	-	-	-	-	-	-	3	-	-	-	2
2004	-	-	-	-	-	1	-	-	-	-	-
2005	-	1	2	4	-	-	3	-	-	-	2
2006	4	2	3	1	-	-	1	-	-	5	7
2007	4	2	-	3	-	-	3	-	-	-	1
2008	-	-	3	-	-	-	3	2	-	-	1
2009	-	-	3	-	-	-	2	-	1	-	1
2010	2	4	-	-	-	-	3	-	-	-	-
2011	1	2	5	1	-	-	-	-	-	-	1
2012	-	11	5	-	-	-	2	1	5	-	5
2013	1	5	7	-	-	-	5	1	5	-	8
2014	3	3	1	-	-	-	2	2	8	-	6
2015	4	4	1	-	-	1	1	2	5	2	5
2016	3	3	4	-	-	1	2	7	1	2	5
2017	1	8	5	-	8	22	9	6	6	-	4
<b>TOTAL</b>	<b>24</b>	<b>45</b>	<b>40</b>	<b>9</b>	<b>8</b>	<b>26</b>	<b>39</b>	<b>21</b>	<b>31</b>	<b>9</b>	<b>48</b>

Le nombre de souches par clade varie de 8 à 48, avec 3 clades (4, 5 et 10) contenant moins de 10 souches, 4 clades (1, 6, 8 et 9) contenant entre 21 et 31 souches et les 4 restantes (2, 3, 7, 11) 39 souches ou plus (Tableau 6, Annexes 4 et 5).

Le **clade 4** contient des souches provenant uniquement de bovins (9). Huit de ces souches ont été collectées entre 2005 et 2007 et une en 2011, étant la souche la plus génétiquement isolée, avec donc une absence de détection de souches provenant de ce clade pendant 3 ans. Toutes les souches du **clade 5** ont été isolées en 2017, et proviennent de bovins (7) et d'un sanglier (1).

Le **clade 6** contient 2 souches isolées en 2001 et 2004 (bovins) et les 24 souches restantes entre 2015 et 2017 (22 bovins, 1 sanglier et 1 renard). La souche de 2004 est la souche la plus isolée génétiquement du restant du clade.

Pour les **clades 8 et 9**, la plupart des souches est isolées à partir de 2012 et proviennent de bovins et de la faune sauvage. Le **clade 10** contient 9 souches dont 5 ont été isolées en 2006 (bovins) puis 4 entre 2015 et 2016 (bovins et 1 blaireau), avec donc une absence de détection pendant 8 ans. L'arbre illustré dans l'annexe 4 montre que le blaireau de 2015 est plus proche génétiquement des 5 souches de bovins de 2006 que des 3 souches de bovins de 2015/16.

Les **clades 1, 2, 3, 7 et 11** ont un grand nombre de souches et les plus grandes profondeurs temporelle avec des souches provenant d'au moins 4 espèces (sauf le clade 7). Le **clade 1** contient de souches de bovins (19), de blaireaux (3), d'un sanglier et d'un cerf, isolées en 2001, 2006, 2007, 2010, 2011 et entre 2013 et 2017. Le **clade 2** contient de souches de bovins (29), de blaireaux (8), de sangliers (6) d'un renard et d'un cerf, isolées entre 2005 et 2007 puis entre 2010 et 2017. Le **clade 3** contient de souches de bovins (31), de blaireaux (6) et de sangliers (3), isolées en 2001, 2005, 2006, 2008 et 2009 puis entre 2011 et 2017. Le **clade 7** contient de souches de bovins (33) et de blaireaux (6), isolées en 2003, entre 2005 et 2010, puis entre 2012 et 2017. Enfin, le **clade 11** contient de souches de bovins (36), de blaireaux (10), d'un sanglier et d'un chevreuil, isolées en 2003, entre 2005 et 2009 puis entre 2011 et 2017. Tous ces clades ont des sauts temporels avec des années où aucune souche n'était isolée.

### *3.5.2 Distances intra-clade*

La plupart des clades est bien définie géographiquement (Figure 8 et Figure 9, Tableau 7 et Tableau 8). La distance intra-clade moyenne est de  $11,53 \pm 5,32$  km. Les distances moyennes intra-clade varient de  $1,53 \pm 3,93$  km pour le clade 6 (dont 21 souches ont été isolées dans une seule commune) à  $16,65 \pm 9,61$  km pour le clade 11 (celle avec le plus grand nombre total de souches, n=48).

La distance maximale intra-clade peut être utilisée pour caractériser la couverture géographique. Le clade 1 (n=24) a la plus grande distance maximale intra-clade, 64,50 km, suivie du clade 11 (57,14 km, n=48) et du clade 4 (53,41 km, n=9). Un grand nombre de souches peut impliquer une large couverture géographique, comme observé pour le clade 11. Mais la présence de souches éloignées, comme dans le clade 4 contribue aussi pour une distance maximale élevée.

○ Bovin □ Blaireau △ Sanglier ◻ Renard ◡ Cerf ◻ Chevreuril

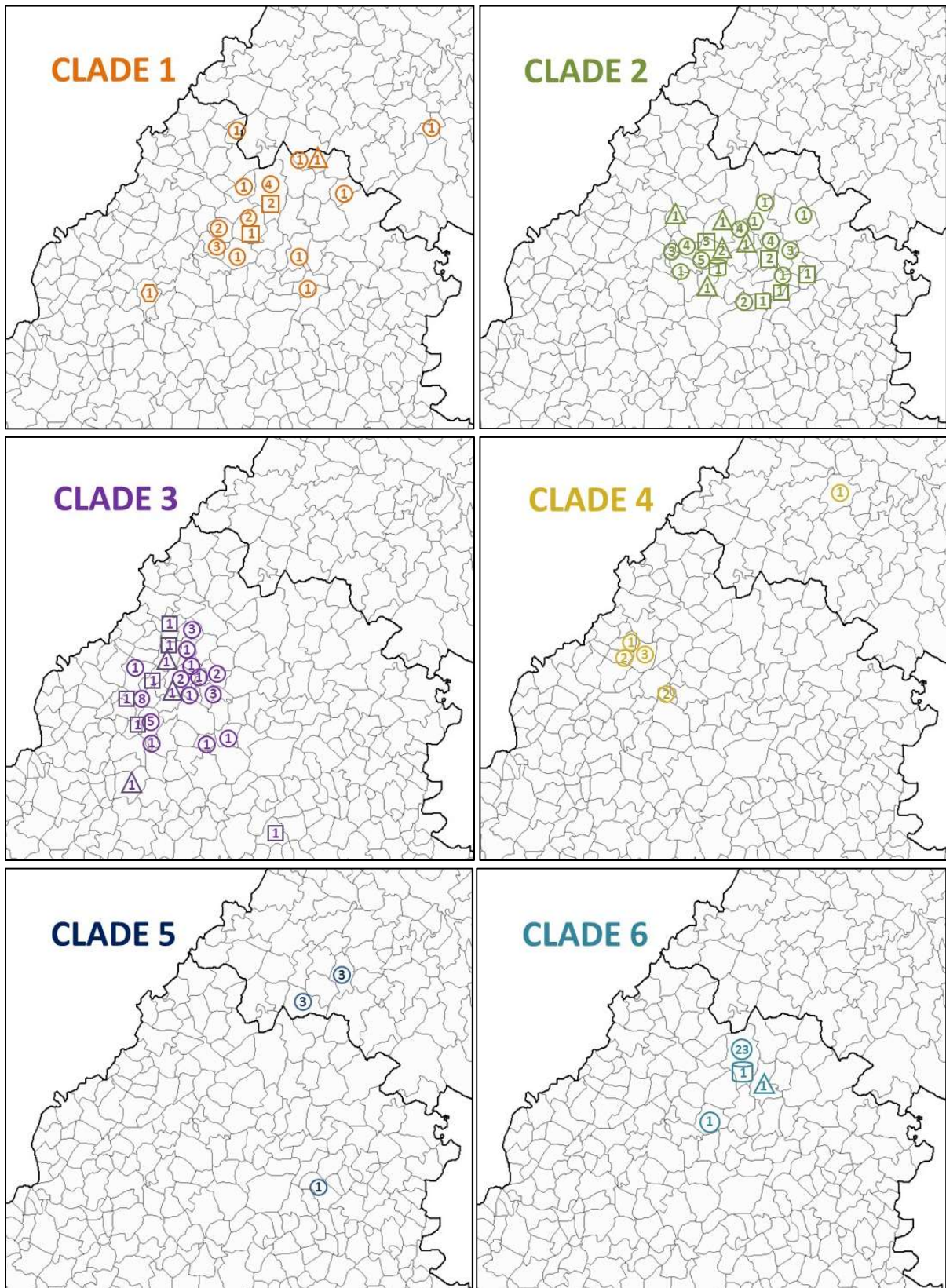


Figure 8. Distribution spatiale des souches de chaque clade. Les formes représentent les différentes espèces et sont placées sur les communes où les souches ont été collectées. Les chiffres à l'intérieur des formes indiquent le nombre de souches provenant d'une même commune. Chaque clade est représenté par une couleur différente.



○ Bovin □ Blaireau △ Sanglier ◻ Renard ◊ Cerf ▭ Chevreuil

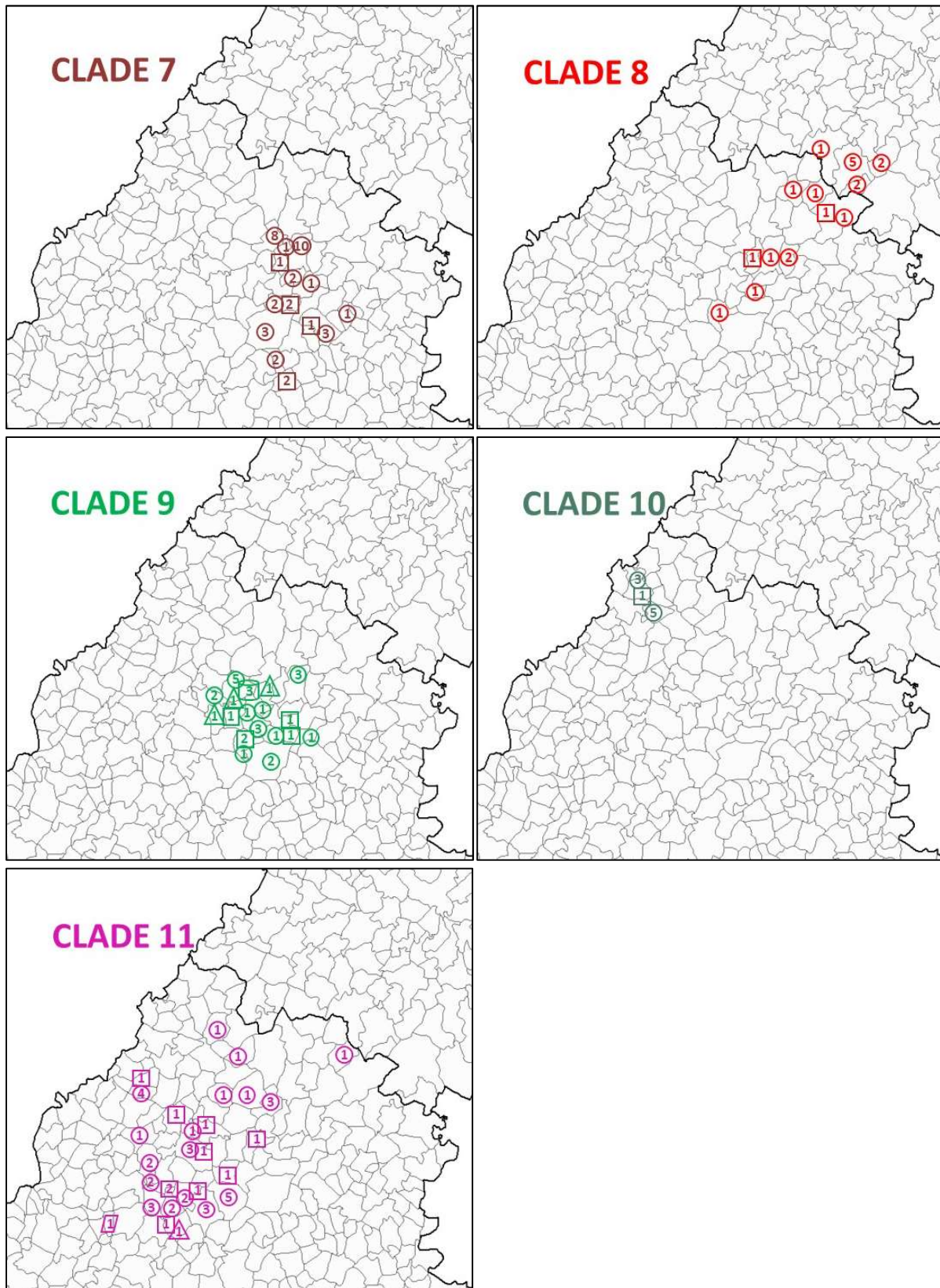


Figure 9. Distribution spatiale des souches de chaque clade (suite). Les formes représentent les différentes espèces et sont placées sur les communes où les souches ont été collectées. Les chiffres à l'intérieur des formes indiquent le nombre de souches provenant d'une même commune. Chaque clade est représenté par une couleur différente.

Tableau 7. Les distances intra-clades en km: minimum, moyenne et maximum.

CLADE	MIN	MOYENNE	ECART TYPE	MAX	Nombre de souches
1	0	15,92	11,03	64,50	24
2	0	10,18	6,06	28,76	45
3	0	12,22	8,16	47,17	40
4	0	16,60	19,26	53,41	9
5	0	15,54	16,51	42,96	8
6	0	1,53	3,93	14,02	26
7	0	10,67	7,34	27,77	39
8	0	15,57	10,77	45,24	21
9	0	8,66	4,55	20,27	31
10	0	3,27	2,96	5,88	9
11	0	16,65	9,61	57,14	48

Tableau 8. Statistique de la moyenne de la distance intra-clade.

DISTANCE INTRA-CLADE (KM)	MINIMUM	MOYENNE	ECART TYPE	MAXIMUM
	1,53	11,53	5,32	16,65

Pour l'ensemble des souches de l'étude, une corrélation positive et significative ( $p=0.0001$ ) a été identifiée entre distance génétique et distance géographique.

Le test de Mantel (Tableau 9) a montré une corrélation positive significative entre distance géographique et distance génétique pour les clades 1 ( $p=0,0049$ ), 5 ( $p=0,05$ ), 7 ( $p=0,03$ ) et 11 ( $p=0,04$ ). Pour les clades 4 et 10 moins de 10 souches ont été disponibles, par conséquent un manque de puissance pourrait expliquer que la corrélation ne soit pas significative. Dans le cas du clade 6, 23/26 des souches sont isolées dans la même commune (Saint-Saud-Lacoussiere), donc la majorité des distances entre souches sont égales à 0.

Tableau 9. Résultats du test de Mantel pour comparer les distances géographiques et génétiques intra-clades.

Clade	Nombre de souches	Nombre de communes	Coefficient de corrélation de Mantel [95% CI]		P
1	24	13	0,22	[0,15 – 0,28]	<b>0,0049</b>
2	45	16	0,066	[-0,0086 – 0,14]	0,15
3	40	19	0,015	[-0,032 – 0,063]	0,39
4	9	5	0,29	[0,14 – 0,42]	0,08
5	8	3	0,036	[0,020 – 0,050]	<b>0,05</b>
6	26	3	-0,26	[-0,30 – 0,029]	0,99
7	39	11	0,19	[0,11 – 0,25]	<b>0,03</b>
8	21	12	0,086	[-0,033 – 0,20]	0,19
9	31	13	-0,28	[-0,34 - -0,22]	1
10	9	2	-0,31	[-0,60 – 0,016]	0,98
11	48	23	0,12	[0,066 – 0,18]	<b>0,04</b>
GLOBAL	300	74	0,049	[0,033 – 0,062]	<b>0,0001</b>

#### 4 Discussion

J’ai utilisé des données de séquençage complet du génome pour étudier pour la première fois l’épidémiologie moléculaire de la tuberculose bovine en Dordogne et Haute-Vienne. Le nombre élevé de souches disponibles (n=300), la diversité des espèces hôtes testées (bovins, blaireaux, sangliers, renards, cerfs, chevreuils) et la grande profondeur temporelle (2001-2017) constituent un jeu de données informatif pour l’inférence phylogénétique.

Les souches ont été isolées dans le cadre de la surveillance de la tuberculose bovine en France. Les souches provenant des bovins ont été collectées suite aux campagnes de dépistage avec des résultats non négatifs aux tests réalisés ou les inspections *ante mortem* et *post mortem* dans les abattoirs. Les souches provenant de la faune sauvage ont été collectées dans le cadre de Sylvatub.

Les caractéristiques intrinsèques des tests de dépistage (sensibilité et spécificité imparfaites), les conditions d’application et l’interprétation des tests sont des facteurs qui jouent un rôle important dans la qualité de la surveillance et peuvent causer une sous-détection de cas. En ce qui concerne la faune sauvage, la sous-détection de cas est inhérente à l’état sauvage des espèces. Il est aussi important de considérer le fait que le séquençage complet du génome de *M. bovis* n’a pas été obtenu à partir de l’ADN pur, et donc de possibles contaminations rendant les échantillons inutilisables. Pour ces raisons nous attendons un biais d’échantillonnage dans notre étude.

Deux autres études phylogénétiques de *M. bovis* ont été conduites en utilisant la même méthode d’inférence phylogénétique avec le même outil (BEAST): Crispell *et al.* (2017) ont analysé 168 souches de *M. bovis* provenant de bovins (n=104) et de la faune sauvage (opossum, porc, furet et cerf, n= 64) et isolées entre 1980 et 2015 en Nouvelle Zélande ; Trewby *et al.* (2015) ont analysé 144 souches de *M. bovis* de bovins (n=139) et de blaireaux (n=6), isolées entre 1996 et 2011 en Irlande du Nord. Par rapport à l’étude actuelle, les deux études avaient des échantillons plus petits, mais une plus grande profondeur temporelle.

Le modèle évolutif sélectionné comprend un modèle de substitution HKY85, une horloge moléculaire stricte et un modèle de population exponentielle. Le modèle de substitution de nucléotides sélectionné, HKY85, permet d’estimer des fréquences de bases différentes et considère des taux de transition et transversion différents. Ce modèle a également été utilisé par Crispell *et al.* (2017), mais

pas par Trewby *et al.* (2016) qui ont utilisé un modèle de substitution plus simple (JC69). Le modèle de substitution HKY85 semble le mieux adapté à *M. bovis*. En effet, il a été démontré qu'un de taux de bases GC ~65% est caractéristique de ce pathogène (Garnier *et al.*, 2003).

Le modèle d'horloge moléculaire sélectionné, une horloge stricte, suppose un taux d'évolution constant le long des branches. Le même modèle a été sélectionné par Trewby *et al.* (2016). Crispell *et al.* (2017) ont utilisé un modèle d'horloge moléculaire relâchée exponentielle. Une hypothèse qui pourrait expliquer cette différence est la profondeur temporelle plus importante du jeu de données de Crispell *et al.* (2017) (35 ans, contre 15 ans pour l'étude de Trewby *et al.* (2016) et 16 ans pour notre étude).

Le modèle de population sélectionné, un modèle de population avec un taux de croissance exponentielle (taille de population qui augmente à un taux constant), diffère des modèles sélectionnés par les deux autres études. Crispell *et al.* (2017) a utilisé un modèle de population skygrid et Trewby *et al.* (2016) un modèle de population skyline. Ce sont des modèles qui estiment tous les paramètres des variations de taille de la population. Ces différences peuvent être dues à des différentes tailles d'échantillon.

Le taux de mutation de *M. bovis* estimé dans cette étude est de 0,28 mutations/génome/an (IC<sub>95%</sub> [0,24 ; 0,34]). Cette valeur est cohérente avec celles trouvées dans les autres études. Trewby *et al.* (2015) ont estimé un taux de 0,2 mutations/génome/an (IC<sub>95%</sub> [0,1 ; 0,3]), alors que Crispell *et al.* (2017) ont estimé 0,53 mutations/génome/an (IC<sub>95%</sub> [0,22 ; 0,94]). Néanmoins, l'étude actuelle a permis de d'estimer un taux de substitution plus précis, probablement grâce à la taille plus grande de l'échantillon et à la profondeur temporelle moins grande. Ce taux de mutation de *M. bovis* a également été estimé pour une autre région en France, la Côte-d'Or à 0,29 mutations/génome/an. Il s'agit d'une étude qui n'est pas encore publiée.

Les estimations de taux de transmission entre espèces sont similaires pour la plupart des paires d'espèces, variant entre 0,07 et 0,10 transitions/individu/an. Le taux estimé est moins élevé pour les transmissions entre bovin et sanglier (0,04 transitions/individu/an) et plus élevé pour celles entre sangliers et renards (0,17 transitions/individu/an). Ces valeurs doivent être traités avec du recul. En effet, le biais d'échantillonnage, plus important pour la faune sauvage, bien que non négligeable pour les bovins, peut surestimer certains taux (ceux qui impliquent les bovins) et sous-estimer les autres.

Nous avons prédit que la plupart des transmissions a lieu entre bovins et que la plupart des transmissions inter-espèces a lieu des bovins vers la faune sauvage, comme illustré dans l'arbre phylogénétique. Suite à une transmission de bovins vers la faune sauvage, il est prédit que *M. bovis* circule ensuite majoritairement au sein de la faune sauvage et principalement au sein de la même espèce. La seule transmission prédite de la faune sauvage vers des bovins a un blaireau comme source estimé. Au sein de cette lignée *M. bovis* a circulé et évolué quelques années chez des blaireaux avant d'être transmise à nouveau à un bovin. L'étude de Crispell *et al.* (2017) en Nouvelle Zélande a suggéré que la tuberculose bovine circulait entre les bovins et la faune sauvage, malgré l'impossibilité de déterminer la direction de la transmission.

Parmi les 5 souches provenant de renards, il est prédit que 4 proviendraient d'un sanglier comme ancêtre hypothétique. Il s'agit d'un résultat surprenant, étant donné que les renards partagent des terriers avec des blaireaux (Mori *et al.*, 2015). Une hypothèse qui pourrait expliquer ce résultat est que des déchets de chasse provenant de sangliers infectés pourraient être laissés par les chasseurs et être consommés par des renards.

Dans le modèle utilisé les transmissions ont été considérées symétriques pour chaque paire d'espèces. La direction de transmission entre deux espèces peut avoir de taux divergents et cette différence peut varier au cours du temps. Des modèles supplémentaires avec des taux de transmission entre espèces asymétrique doivent être développés afin de déterminer si une transmission asymétrique serait à privilégier. Les modèles asymétriques demandent l'estimation de plus de paramètres, ce qui peut poser des problèmes d'identifiabilité des paramètres. Un nombre plus important de souches isolées

de la faune sauvage serait idéal pour ces analyses, ces paramètres étant influencés par l'échantillonnage (Kao et al., 2016). Pourtant, l'obtention de nouveaux échantillons n'est pas une possibilité.

Dans l'arbre consensus, certaines souches de la faune sauvage sont éloignées temporellement de leur ancêtre le plus récent. Ceci est également observé pour les bovins, mais moins souvent. Bien que la tuberculose bovine soit une maladie chronique d'évolution lente et donc que les dates d'échantillonnage et d'infection soient différentes et qu'il ne semble pas probable que ces souches aient évolué plus lentement que les autres, il est possible que d'autres souches aient circulé dans la faune sauvage entre temps, mais n'aient pas été détectées et échantillonnées. De plus, le réseau Sylvatub est actif depuis 2009 et la faune sauvage n'a pas été échantillonnée sur la période 2001-2009. Il est donc important de considérer que le biais d'échantillonnage est réel et exerce un impact important sur la reconstruction de la phylogénie, par une sous-estimation du rôle de la faune sauvage dans les transmissions inter-espèce.

Une nouvelle analyse sur les souches collectées en 2009-2017 serait intéressante pour comparer les estimations des taux de transmissions entre espèces à celles obtenues pour les souches collectées en 2001-2017.

La distribution temporelle des souches dans les clades 1, 2, 3, 4, 6, 7, 8, 9 et 11 suggère que des cas n'ont pas été détectés par la surveillance, illustrant aussi le sous-échantillonnage.

Une corrélation positive significative entre les distances géographiques et les distances génétiques est attendue, étant donné que les espèces animales concernées, principalement les bovins et les blaireaux, sont des animaux territoriaux. Une corrélation positive significative entre distance géographique et distance génétique a également été trouvée dans deux autres études sur l'épidémiologie moléculaire de *M. bovis* en Irlande du Nord (Biek et al., 2012; Trewby et al., 2016). Si cette corrélation est mise en évidence pour l'ensemble des 300 souches de l'étude, on ne retrouve cette corrélation que pour les clades 1, 5, 7 et 11. Pour les autres clades, l'absence de corrélation pourrait s'expliquer par l'existence de pâtures rattachées à une exploitation donnée mais situées dans une autre commune. Pour explorer la présence des souches géographiquement éloignées des autres d'un même clade, il serait intéressant d'analyser les mouvements des bovins.

Les communes de Villars et de Saint-Saud-Lacoussière méritent une attention spéciale en raison du nombre important d'APDI. A Villars (8 APDI entre 2004 et 2017, concernant 21 bovins dans 6 exploitations dont 2 exploitations avec 2 APDI) des souches appartenant à 4 clades (1, 2, 6 et 9), provenant de bovins, blaireaux, sangliers et renards ont été isolées majoritairement entre 2013 et 2017. A Saint-Saud-Lacoussière (14 APDI dont 8 en 2017, concernant 55 bovins dans 13 exploitations) des souches appartenant à 2 clades (1 et 6), provenant de bovins, blaireaux et un renard, ont été isolées dans la plupart de cas entre en 2017. Une étude épidémiologique dans ces communes peut être d'intérêt pour essayer de trouver les causes du grand nombre de foyers. Par exemple, il est possible qu'il y ait une lacune dans des mesures de dépistage et de biosécurité qui pourrait être identifiée et réglée.

## 5 Conclusion

Cette étude a pour objectif d'explorer le rôle de la faune sauvage dans la transmission de la tuberculose bovine, à partir de l'inférence phylogénétique. Nous avons analysé 300 souches de *M. bovis* isolées entre 2001 et 2017, provenant de bovins, blaireaux, sangliers, renards, cerfs et un chevreuil, d'une zone sélectionnée en Dordogne et Haute-Vienne.

Il a été constaté que *M. bovis* circule davantage chez les bovins et que les sauts inter-espèces de la faune sauvage vers les bovins ne sont pas fréquents. En considérant notre jeu de données, la faune sauvage ne semble pas avoir un rôle important dans la transmission et le maintien de l'infection chez les bovins.

Des événements de transmissions de sangliers vers des renards ont été mis en évidence causant un certain étonnement. Les relations entre ces espèces peuvent être étudiées pour mieux expliquer le lien entre elles.

Les prochaines étapes incluent la comparaison de ces résultats avec des autres régions en France, une analyse temporelle pour explorer comment la maladie avance au cours du temps et pour comparer les distances génétiques avec les distances temporelles, l'analyse des territoires occupés par la faune sauvage et ses mouvements et la modélisation de chaînes de transmission.

## 6 Références

- Aris-Brosou, S., and Yang, Z. (2002). Effects of Models of Rate Evolution on Estimation of Divergence Dates with Special Reference to the Metazoan 18S Ribosomal RNA Phylogeny. *Syst. Biol.* *51*, 703–714.
- Baele, G., Suchard, M.A., Rambaut, A., and Lemey, P. (2016). Emerging Concepts of Data Integration in Pathogen Phylodynamics. *Syst. Biol.* *syw054*.
- Baum, D.A., and Offner, S. (2008). Phylogenies & Tree-Thinking. *Am. Biol. Teach.* *70*, 222–229.
- Bessell, P.R., Orton, R., White, P.C.L., Hutchings, M.R., and Kao, R.R. (2012). Risk factors for bovine Tuberculosis at the national level in Great Britain. *BMC Vet. Res.* *8*, 51.
- Biek, R., O’Hare, A., Wright, D., Mallon, T., McCormick, C., Orton, R.J., McDowell, S., Trewby, H., Skuce, R.A., and Kao, R.R. (2012). Whole Genome Sequencing Reveals Local Transmission Patterns of *Mycobacterium bovis* in Sympatric Cattle and Badger Populations. *PLoS Pathog.* *8*, e1003008.
- Boschiroli, M.L., Hénault, S., Palisson, A., Karoui, C., Biet, F., and Zanella, G. (2015). Tuberculose bovine en France : cartographie des souches de *Mycobacterium bovis* entre 2000-2013. *Bull. Épidémiologique* *7*.
- Bouchez-Zacria, M., Courcoul, A., and Durand, B. (2018). The Distribution of Bovine Tuberculosis in Cattle Farms Is Linked to Cattle Trade and Badger-Mediated Contact Networks in South-Western France, 2007–2015. *Front. Vet. Sci.* *5*.
- Campbell, F., Cori, A., Ferguson, N., and Jombart, T. (2019). Bayesian inference of transmission chains using timing of symptoms, pathogen genomes and contact data. *PLOS Comput. Biol.* *15*, e1006930.
- Corner, L.A.L. (2006). The role of wild animal populations in the epidemiology of tuberculosis in domestic animals: How to assess the risk. *Vet. Microbiol.* *112*, 303–312.
- Courcoul, A., Moyen, J.-L., Brugère, L., Faye, S., Hénault, S., Gares, H., and Boschiroli, M.-L. (2014). Estimation of Sensitivity and Specificity of Bacteriology, Histopathology and PCR for the Confirmatory Diagnosis of Bovine Tuberculosis Using Latent Class Analysis. *PLoS ONE* *9*, e90334.
- Crispell, J., Zadoks, R.N., Harris, S.R., Paterson, B., Collins, D.M., de-Lisle, G.W., Livingstone, P., Neill, M.A., Biek, R., Lycett, S.J., et al. (2017). Using whole genome sequencing to investigate transmission in a multi-host system: bovine tuberculosis in New Zealand. *BMC Genomics* *18*.
- Drummond, A.J., and Rambaut, A. (2007). BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol. Biol.* *7*, 214.
- Drummond, A.J., Pybus, O.G., Rambaut, A., Forsberg, R., and Rodrigo, A.G. (2003). Measurably evolving populations. *Trends Ecol. Evol.* *18*, 481–488.
- Drummond, A.J., Ho, S.Y.W., Phillips, M.J., and Rambaut, A. (2006). Relaxed Phylogenetics and Dating with Confidence. *PLoS Biol.* *4*, e88.
- Drummond, A.J., Suchard, M.A., Xie, D., and Rambaut, A. (2012). Bayesian Phylogenetics with BEAUti and the BEAST 1.7. *Mol. Biol. Evol.* *29*, 1969–1973.

- Echave, J., Spielman, S.J., and Wilke, C.O. (2016). Causes of evolutionary rate variation among protein sites. *Nat. Rev. Genet.* *17*, 109–121.
- Eybpoosh, S., Haghdoost, A.A., Mostafavi, E., Bahrapour, A., Azadmanesh, K., and Zolala, F. (2017). Molecular epidemiology of infectious diseases. *Electron. Physician* *9*, 5149–5158.
- Fediaevsky, A., Courcoul, A., Boschioli, M.L., and Reveillaud, E. (2012). Tuberculose bovine en France en 2012 : des signaux favorables mais une situation toujours complexe dans certaines zones. *Bull. Épidémiologique Santé Anim. Aliment.* *7*.
- Felsenstein, J. (1981). Evolutionary trees from DNA sequences: A maximum likelihood approach. *J. Mol. Evol.* *17*, 368–376.
- Garnier, T., Eiglmeier, K., Camus, J.-C., Medina, N., Mansoor, H., Pryor, M., Duthoy, S., Grondin, S., Lacroix, C., Monsempe, C., et al. (2003). The complete genome sequence of *Mycobacterium bovis*. *Proc. Natl. Acad. Sci.* *100*, 7877–7882.
- Hasegawa, M., Kishino, H., and Yano, T. (1985). Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *J. Mol. Evol.* *22*, 160–174.
- Hauer, A., De Cruz, K., Cochard, T., Godreuil, S., Karoui, C., Henault, S., Bulach, T., Bañuls, A.-L., Biet, F., and Boschioli, M.L. (2015). Genetic Evolution of *Mycobacterium bovis* Causing Tuberculosis in Livestock and Wildlife in France since 1978. *PLOS ONE* *10*, e0117103.
- Jia, F., Lo, N., and Ho, S.Y.W. (2014). The Impact of Modelling Rate Heterogeneity among Sites on Phylogenetic Estimates of Intraspecific Evolutionary Rates and Timescales. *PLOS ONE* *9*, e95722.
- Johnston, W.T., Vial, F., Gettinby, G., Bourne, F.J., Clifton-Hadley, R.S., Cox, D.R., Crea, P., Donnelly, C.A., McInerney, J.P., Mitchell, A.P., et al. (2011). Herd-level risk factors of bovine tuberculosis in England and Wales after the 2001 foot-and-mouth disease epidemic. *Int. J. Infect. Dis.* *15*, e833–e840.
- Jukes, T.H., and Cantor, C.R. (1969). “Evolution of protein molecules”. In Munro, H.N. (ed.). *Mammalian protein metabolism*. (Academic Press).
- Kamerbeek, J., Schouls, L., Kolk, A., Agterveld, M.V., Soolingen, D.V., Kuijper, S., Bunschoten, A., Molhuizen, H., Shaw, R., and Goyal, M. (1997). Simultaneous Detection and Strain Differentiation of *Mycobacterium tuberculosis* for Diagnosis and Epidemiology. *35*, 8.
- Kao, R.R., Price-Carter, M., and Robbe-Austerman, S. (2016). Use of genomics to track bovine tuberculosis transmission: -EN- -FR- L’ utilisation de la génomique pour retracer la transmission de la tuberculose bovine -ES- Uso de la genómica para rastrear la transmisión de la tuberculosis bovina. *Rev. Sci. Tech. OIE* *35*, 241–268.
- Kimura, M. (1980). A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol.* *16*, 111–120.
- Kingman, J.F.C. (1982). The coalescent. *Stoch. Process. Their Appl.* *13*, 235–248.
- Lenth, R.V. (2001). Some Practical Guidelines for Effective Sample Size Determination. *Am. Stat.* *55*, 187–193.



- Marsot, M., Béral, M., Scoizec, A., Mathevon, Y., Durand, B., and Courcoul, A. (2016). Herd-level risk factors for bovine tuberculosis in French cattle herds. *Prev. Vet. Med.* *131*, 31–40.
- Martínez-López, B., Barasona, J.A., Gortázar, C., Rodríguez-Prieto, V., Sánchez-Vizcaíno, J.M., and Vicente, J. (2014). Farm-level risk factors for the occurrence, new infection or persistence of tuberculosis in cattle herds from South-Central Spain. *Prev. Vet. Med.* *116*, 268–278.
- Maturana, P., Brewer, B.J., Klaere, S., and Bouckaert, R. (2019). Model selection and parameter inference in phylogenetics using Nested Sampling. *Syst. Biol.* *68*, 219–233.
- McGuire, G., Prentice, M.J., and Wright, F. (1999). Improved Error Bounds for Genetic Distances from Dna Sequences. *Biometrics* *55*, 1064–1070.
- McIlroy, S., Neill, S., and McCracken, R. (1986). Pulmonary lesions and Mycobacterium bovis excretion from the respiratory tract of tuberculin reacting cattle. *Vet. Rec.* *118*, 718.
- Mollentze, N., Nel, L.H., Townsend, S., le Roux, K., Hampson, K., Haydon, D.T., and Soubeyrand, S. (2014). A Bayesian approach for inferring the dynamics of partially observed endemic infectious diseases from space-time-genetic data. *Proc. R. Soc. B Biol. Sci.* *281*, 20133251–20133251.
- Morelli, M.J., Thébaud, G., Chadœuf, J., King, D.P., Haydon, D.T., and Soubeyrand, S. (2012). A Bayesian Inference Framework to Reconstruct Transmission Trees Using Epidemiological and Genetic Data. *PLoS Comput. Biol.* *8*, e1002768.
- Mori, E., Menchetti, M., and Balestrieri, A. (2015). Interspecific den sharing: a study on European badger setts using camera traps. *Acta Ethologica* *18*, 121–126.
- Neill, S.D., Bryson, D.G., and Pollock, J.M. (2001). Pathogenesis of tuberculosis in cattle. *Tuberculosis* *81*, 79–86.
- OIE (2018). Code sanitaire pour les animaux terrestres. 562.
- Orloski, K., Robbe-Austerman, S., Stuber, T., Hench, B., and Schoenbaum, M. (2018). Whole Genome Sequencing of Mycobacterium bovis Isolated From Livestock in the United States, 1989–2018. *Front. Vet. Sci.* *5*.
- Pandolfi, F., Cavalerie, L., Chevalier, F., Jabert, P., Calavas, D., Benet, J., Keck, N., Girard, S., Lefebvre, A.-L., Martin, F., et al. (2018). SURVEILLANCE DE LA TUBERCULOSE BOVINE EN2018 (POINT AU 4 AVRIL).
- Price-Carter, M., Brauning, R., de Lisle, G.W., Livingstone, P., Neill, M., Sinclair, J., Paterson, B., Atkinson, G., Knowles, G., Crews, K., et al. (2018). Whole Genome Sequencing for Determining the Source of Mycobacterium bovis Infections in Livestock Herds and Wildlife in New Zealand. *Front. Vet. Sci.* *5*.
- Réveillaud, É., Desvaux, S., Boschioli, M.-L., Hars, J., Faure, É., Fediaevsky, A., Cavalerie, L., Chevalier, F., Jabert, P., Poliak, S., et al. (2018). Infection of Wildlife by Mycobacterium bovis in France Assessment Through a National Surveillance System, Sylvatub. *Front. Vet. Sci.* *5*.
- Rivière, J., Faure, E., Hendrikx, P., and Fediaevsky, A. (2013). Sylvatub : bilan d’une première année de surveillance de la tuberculose bovine dans la faune sauvage en France. *Bull. Épidémiologique Santé Anim. Aliment.* *6*.

Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D.L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M.A., and Huelsenbeck, J.P. (2012). MrBayes 3.2: Efficient Bayesian Phylogenetic Inference and Model Choice Across a Large Model Space. *Syst. Biol.* *61*, 539–542.

Roring, S., Hughes, M.S., Skuce, R.A., and Neill, S.D. (2000). Simultaneous detection and strain differentiation of *Mycobacterium bovis* directly from bovine tissue specimens by spoligotyping. *Vet. Microbiol.* *74*, 227–236.

Roring, S., Scott, A., Brittain, D., Walker, I., Hewinson, G., Neill, S., and Skuce, R. (2002). Development of Variable-Number Tandem Repeat Typing of *Mycobacterium bovis*: Comparison of Results with Those Obtained by Using Existing Exact Tandem Repeats and Spoligotyping. *J. Clin. Microbiol.* *40*, 2126–2133.

Rutschmann, F. (2006). Molecular dating of phylogenetic trees: A brief review of current methods that estimate divergence times. *Divers. Distrib.* *12*, 35–48.

Suchard, M.A., and Redelings, B.D. (2006). BAli-Phy: simultaneous Bayesian inference of alignment and phylogeny. *Bioinformatics* *22*, 2047–2048.

Tavaré, S. (1986). Some Probabilistic and Statistical Problems in the Analysis of DNA Sequences. *Lect. Math. Life Sci.* *17*, 57–86.

Thorne, J.L., Kishino, H., and Painter, I.S. (1998). Estimating the rate of evolution of the rate of molecular evolution. *Mol. Biol. Evol.* *15*, 1647–1657.

Trewby, H., Wright, D., Breadon, E.L., Lycett, S.J., Mallon, T.R., McCormick, C., Johnson, P., Orton, R.J., Allen, A.R., Galbraith, J., et al. (2016). Use of bacterial whole-genome sequencing to investigate local persistence and spread in bovine tuberculosis. *Epidemics* *14*, 26–35.

White, Paul.W., Martin, S.W., De Jong, M.C.M., O’Keeffe, J.J., More, S.J., and Frankena, K. (2013). The importance of ‘neighbourhood’ in the persistence of bovine tuberculosis in Irish cattle herds. *Prev. Vet. Med.* *110*, 346–355.

Yang, Z. (1996). Among-site rate variation and its impact on phylogenetic analyses. *Trends Ecol. Evol.* *11*, 367–372.

Yang, Z., and Rannala, B. (2012). Molecular phylogenetics: principles and practice. *Nat. Rev. Genet.* *13*, 303–314.

Zanella, G., Duvauchelle, A., Hars, J., Moutou, F., Boschioli, M.L., and B., D. (2008). Patterns of lesions of bovine tuberculosis in wild red deer and wild boar. *Vet. Rec.* *163*, 43–47.

### **Textes réglementaires**

Arrêté du 29 juillet 2013 relatif à la définition des dangers sanitaires de première et deuxième catégorie pour les espèces animales.

Note de service DGAL/SDSPA/N2011-8214 du 20 septembre 2011 : Surveillance épidémiologique de la tuberculose dans la faune sauvage en France : réseau Sylvatub.

Note de service DGAL/SDSPA/N2013-8202 du 04 décembre 2013 : Tuberculose bovine : diagnostic de laboratoire post-mortem

Note de service DGAL/SDSPA/2016-1001 du 22 décembre 2016 : Note de service relative aux modalités techniques de gestion des suspicions de tuberculose bovine (mise à jour pour la campagne de prophylaxie 2016-2017)

OIE (2018). Code sanitaire pour les animaux terrestres. 562.

## 7 Annexes

### Annexe 1 – Mesures de police sanitaire selon statu du troupeau de bovine (Arrêté du 15/09/2003 – Chapitre 5)

#### I. Troupeaux susceptible :

- lien épidémiologique à risque établi avec un animal infecté

##### Mesures de police sanitaire : Arrêté préfectoral de surveillance

1. Qualification suspendue ;
2. Investigations et possible abattage diagnostique.

#### II. Troupeaux suspects d'être infectés de tuberculose:

- constatation de lésions *post mortem*
- PCR positif sur un animal issu d'un troupeau indemne
- réactions non négatives par intradermo-tuberculation (IDT) ou au test de l'interféron gamma ou à la sérologie

##### Mesures de police sanitaire : Arrêté préfectoral de surveillance

1. Qualification suspendue ;
2. Mesures 1, 2, 6, 7 des troupeaux infectés (listées au-dessous) ;
3. Investigations épidémiologiques et analytiques, contrôles documentaires, contrôles par IDT et par test de dosage de l'interféron gamma de tout ou partie des animaux et contrôles des pratiques d'élevage utiles à la détermination du statut sanitaire du troupeau ;
4. Mesures de gestion du lait et des produits laitiers.

#### III. Troupeaux infectés de tuberculose :

- signes cliniques associés à une réaction positive par IDT
- isolement et identification de *Mycobacterium* sp.
- réaction d'IDT comparative positive associée à l'observation de lésions histologiques
- analyse PCR positive

##### Mesures de police sanitaire : Arrêté préfectoral portant déclaration d'infection (APDI)

1. Visite, recensement et contrôle de l'identification des bovinés et d'autres espèces sensibles;
2. Isolement et séquestration de tous les animaux du troupeau reconnu infecté jusqu'à leur abattage;
3. Investigations cliniques, allergiques, analytiques et épidémiologiques à l'égard des animaux détenus sur l'exploitation;
4. Isolement et séquestration des animaux d'autres espèces sensibles reconnus tuberculeux;
5. Marquage ou repérage et abattage de tous les animaux reconnus infectés;
6. Interdiction de laisser entrer des animaux sensibles provenant d'autres troupeaux;
7. Interdiction de laisser sortir de l'exploitation des animaux sensibles;
8. Enquête épidémiologique approfondie;
9. Mesures de gestion du lait et des produits laitiers;
10. Mesures de nettoyage de désinfection pouvant être assorties d'une période de vide sanitaire.

## Annexe 2 – Modèles de substitution de nucléotides

Le modèle de substitution le plus simple, le Jukes et Cantor (Jukes and Cantor, 1969), suppose le même taux de substitution entre n'importe quelle paire de nucléotides et les mêmes fréquences (0,25) pour les 4 bases (A, T, G, C).

Le modèle Kimura (K80) (Kimura, 1980) suppose des différents taux de substitution pour les transitions (substitution entre pyrimidines – T et C) et les transversions (substitution entre purines – A et G).

Le modèle Felsenstein's 1981 (F81) (Felsenstein, 1981) est une variation du JC69, en supposant que les fréquences de bases peuvent être différents de 0,25.

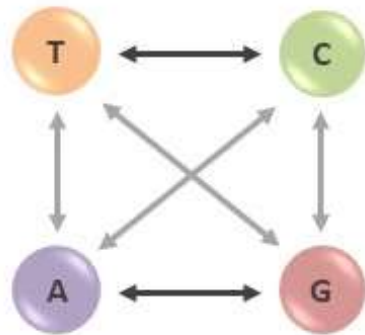
Le modèle HKY85 (Hasegawa et al., 1985) suppose une fréquence de bases variable et des taux de substitution différents pour les transitions et les transversions.

Le modèle Tamura-Nei TN93 (Tavaré, 1986) diffère du HKY85 en ce que concerne le taux de transitions, qui est variable.

Le modèle GTR (« General Time Reversible ») est le plus complexe. Selon ce modèle, la fréquence de bases est variable et chaque paire de nucléotides présente un taux de substitution différent.

Les modèles utilisés dans cette étude (JC69, HKY85 et GTR), sont illustrés dans la figure au-dessous.

### JC69

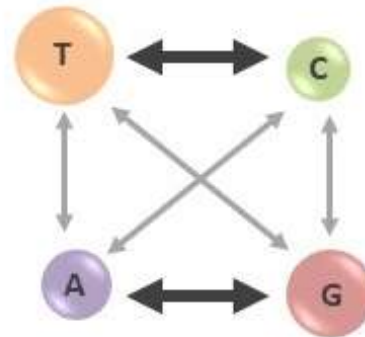


$$\pi_A = \pi_G = \pi_C = \pi_T = 0,25$$

$$Q = \begin{pmatrix} * & \alpha & \alpha & \alpha \\ \alpha & * & \alpha & \alpha \\ \alpha & \alpha & * & \alpha \\ \alpha & \alpha & \alpha & * \end{pmatrix}$$

$\alpha$  = taux de n'importe quelle substitution

### HKY85

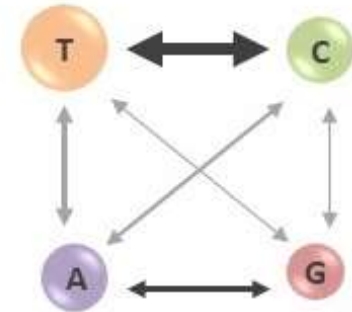


$$\pi_A \neq \pi_G \neq \pi_C \neq \pi_T$$

$$Q = \begin{pmatrix} * & \pi_C\beta & \pi_G\beta\kappa & \pi_T\beta \\ \pi_A\beta & * & \pi_G\beta & \pi_T\beta\kappa \\ \pi_A\beta\kappa & \pi_C\beta & * & \pi_T\beta \\ \pi_A\beta & \pi_C\beta\kappa & \pi_G\beta & * \end{pmatrix}$$

$\beta$  = taux de transversions  
 $\beta\kappa$  = taux de transitions

### GTR



$$\pi_A \neq \pi_G \neq \pi_C \neq \pi_T$$

$$Q = \begin{pmatrix} * & \pi_C a_{\mu} & \pi_G b_{\mu} & \pi_T c_{\mu} \\ \pi_A a_{\mu} & * & \pi_G d_{\mu} & \pi_T e_{\mu} \\ \pi_A b_{\mu} & \pi_C d_{\mu} & * & \pi_T f_{\mu} \\ \pi_A c_{\mu} & \pi_C e_{\mu} & \pi_G f_{\mu} & * \end{pmatrix}$$

↔ Transitions (entre pyrimidines ou entre purines)

↔ Transversions (entre pyrimidines et purines)

Annexe 3 – Modèles évolutifs testés avec BEAST

TEST	1	2	3	4	5	6	7	8	9	10
Modèle Substitution	HKY	HKY	HKY	HKY	HKY	JC	HKY	HKY	GTR	HKY
Cocher Estimate	NON	NON	NON	NON	NON	NON	OUI	OUI	NON	NON
Gamma category count	4	4	4	4	4	4	4	4	4	4
Proportion Invariant	0	0	0	0	0	0	0	0	0	0
Frequence Bases	Empirical	Empirical	Empirical	Empirical	Empirical	-	Empirical	Empirical	Empirical	Empirical
Horloge Moléculaire	Stricte	Exponentielle	Log Normal	Stricte	Stricte	Stricte	Stricte	Exponentielle	Stricte	Exponentielle
Cocher estimate	OUI	OUI	OUI	OUI	OUI	OUI	OUI	OUI	OUI	OUI
Modèle Population	Constante	Constante	Constante	Skyline	Exponentielle	Constante	Constante	Constante	Constante	Yule
Modèle TRAIT	Symetric	Symetric	Symetric	Symetric	Symetric	Symetric	Symetric	Symetric	Symetric	Symetric
Cocher Estimate Substitution Rate TRAIT	NON	NON	NON	NON	NON	NON	OUI	OUI	NON	NON
Horloge Moléculaire	Stricte	Stricte	Stricte	Stricte	Stricte	Stricte	Stricte	Stricte	Stricte	Stricte
MCMC	500 000 000	500 000 000	500 000 000	500 000 000	500 000 000	500 000 000	500 000 000	500 000 000	500 000 000	500 000 000
Log every	50 000	50 000	50 000	50 000	50 000	50 000	50 000	50 000	50 000	50 000
Screen Log	1 000	1 000	1 000	1 000	1 000	1 000	1 000	1 000	1 000	1 000
Tree Log	50 000	50 000	50 000	50 000	50 000	50 000	50 000	50 000	50 000	50 000
Tree Log TRAIT	50 000	50 000	50 000	50 000	50 000	50 000	50 000	50 000	50 000	50 000
Temps d'analyse (h)	80	88	75	90	81	95	118	71	81	80
Commentaires										Hauter de l'arbre aberrant
Convergence	oui	oui	oui	oui	oui	oui	non	non	non	oui

Annexe 4 – Arbre phylogénétique divisé par clade et probabilité du trait des nœuds hypothétiques



## Annexe 5 – Liste de souches par clade

ID	Espèce	Année	Clade	INSEE	Commune
15Z006495	Bovin	2015	1	24101	Champs Romain
07-01114-00	Bovin	2007	1	24134	Corgnac-sur-l'Isle
15Z002145	Bovin	2015	1	24180	Firbeix
07-00879-00	Bovin	2007	1	24271	Milhac-de-Nontron
07-01013-00	Bovin	2007	1	24271	Milhac-de-Nontron
06-01579-00	Bovin	2006	1	24346	Quinsac
06-01580-00	Bovin	2006	1	24346	Quinsac
06-01780-00	Bovin	2006	1	24346	Quinsac
16Z004107	Bovin	2016	1	24381	Saint-Barthélemy-de-Bussière
13Z004785	Bovin	2013	1	24410	Saint-Front-la-Rivière
10Z001989	Cerf	2010	1	24319	Paussac-et-Saint-Vivien
14Z000013	Bovin	2014	1	24410	Saint-Front-la-Rivière
01-03584-00	Bovin	2001	1	24489	Saint-Priest-les-Fougères
07-00780-00	Bovin	2007	1	24498	Saint-Saud-Lacoussière
14Z006798	Blaireau	2014	1	24498	Saint-Saud-Lacoussière
14Z007047	Blaireau	2014	1	24498	Saint-Saud-Lacoussière
15Z001806	Bovin	2015	1	24498	Saint-Saud-Lacoussière
15Z004066	Bovin	2015	1	24498	Saint-Saud-Lacoussière
17Z003238	Bovin	2017	1	24498	Saint-Saud-Lacoussière
11Z001983	Bovin	2011	1	24551	Thiviers
16Z000965	Sanglier	2016	1	24180	Firbeix
06-01691-00	Bovin	2006	1	24582	Villars
16Z008443	Blaireau	2016	1	24271	Milhac-de-Nontron
10Z000099	Bovin	2010	1	87077	Janailhac
15Z004345	Bovin	2015	2	24096	Champagnac de Belair
06-01573-00	Bovin	2006	2	24134	Corgnac-sur-l'Isle
06-01692-00	Bovin	2006	2	24304	Nantheuil
12Z004494	Bovin	2012	2	24304	Nantheuil
12Z004928	Bovin	2012	2	24304	Nantheuil
12Z006994	Bovin	2012	2	24308	Négrondes
10Z008544	Sanglier	2010	2	24582	Villars
10Z008546	Sanglier	2010	2	24107	La Chapelle-Faucher
10Z009059	Sanglier	2010	2	24271	Milhac-de-Nontron
12Z007220	Bovin	2012	2	24308	Négrondes
11Z002157	Blaireau	2011	2	24551	Thiviers
11Z002149	Bovin	2011	2	24346	Quinsac
12Z003430	Sanglier	2012	2	24528	Sceau-Saint-Angel
17Z000058	Bovin	2017	2	24346	Quinsac
17Z000059	Bovin	2017	2	24346	Quinsac
17Z000060	Bovin	2017	2	24346	Quinsac
16Z004271	Bovin	2016	2	24428	Saint-Jory-de-Chalais
12Z004995	Blaireau	2012	2	24551	Thiviers
12Z005163	Blaireau	2012	2	24308	Négrondes

07-01115-00	Bovin	2007	2	24453	Saint-Martin-de-Fressengeas
07-01914-00	Bovin	2007	2	24453	Saint-Martin-de-Fressengeas
17Z000898	Bovin	2017	2	24474	Saint-Pancrace
17Z002558	Bovin	2017	2	24474	Saint-Pancrace
17Z003109	Bovin	2017	2	24474	Saint-Pancrace
05-01691-00	Bovin	2005	2	24481	Saint-Paul-la-Roche
12Z004505	Bovin	2012	2	24496	Saint-Romain-et-Saint-Clément
13Z005637	Sanglier	2013	2	24582	Villars
14Z003302	Blaireau	2014	2	24582	Villars
14Z003304	Blaireau	2014	2	24582	Villars
14Z005576	Blaireau	2014	2	24505	Saint-Sulpice-d'Excideuil
13Z002705	Bovin	2013	2	24496	Saint-Romain-et-Saint-Clément
15Z005835	Blaireau	2015	2	24134	Cognac-sur-l'Isle
15Z006760	Blaireau	2015	2	24582	Villars
15Z010962	Sanglier	2015	2	24453	Saint-Martin-de-Fressengeas
10Z002121	Bovin	2010	2	24551	Thiviers
12Z003372	Bovin	2012	2	24551	Thiviers
12Z004930	Bovin	2012	2	24551	Thiviers
12Z006182	Bovin	2012	2	24551	Thiviers
13Z001903	Bovin	2013	2	24582	Villars
13Z002215	Bovin	2013	2	24582	Villars
17Z000353	Cerf	2017	2	24453	Saint-Martin-de-Fressengeas
13Z003644	Bovin	2013	2	24582	Villars
16Z005207	Bovin	2016	2	24582	Villars
16Z008197	Bovin	2016	2	24582	Villars
17Z008465	Renard	2017	2	24582	Villars
12Z004741	Bovin	2012	3	24079	Cantillac
11Z002528	Bovin	2011	3	24096	Champagnac de Belair
11Z003570	Bovin	2011	3	24096	Champagnac de Belair
11Z003894	Bovin	2011	3	24096	Champagnac de Belair
13Z003770	Bovin	2013	3	24144	Creyssac
13Z001990	Bovin	2013	3	24170	Eyvirat
05-02882-00	Bovin	2005	3	24111	La Chapelle-Montmoreau
12Z003438	Bovin	2012	3	24235	Léguillac-de-Cercles
12Z003519	Bovin	2012	3	24235	Léguillac-de-Cercles
12Z004424	Bovin	2012	3	24235	Léguillac-de-Cercles
12Z004996	Bovin	2012	3	24235	Léguillac-de-Cercles
13Z001898	Bovin	2013	3	24235	Léguillac-de-Cercles
17Z000349	Bovin	2017	3	24235	Léguillac-de-Cercles
17Z000350	Bovin	2017	3	24235	Léguillac-de-Cercles
17Z000351	Bovin	2017	3	24235	Léguillac-de-Cercles
05-03537-00	Bovin	2005	3	24319	Paussac-et-Saint-Vivien
06-00415-00	Bovin	2006	3	24319	Paussac-et-Saint-Vivien
06-00416-00	Bovin	2006	3	24319	Paussac-et-Saint-Vivien
13Z002708	Bovin	2013	3	24319	Paussac-et-Saint-Vivien

13Z002709	Bovin	2013	3	24319	Paussac-et-Saint-Vivien
11Z003734	Bovin	2011	3	24346	Quinsac
16Z002887	Bovin	2016	3	24346	Quinsac
09-00132-00	Bovin	2009	3	24391	Saint-Crépin-de-Richemont
09-00516-00	Bovin	2009	3	24391	Saint-Crépin-de-Richemont
06-01469-00	Bovin	2006	3	24411	Saint-Front-sur-Nizonne
08-00650-00	Bovin	2008	3	24451	Saint-Martial-de-Valette
08-00986-00	Bovin	2008	3	24451	Saint-Martial-de-Valette
13Z005634	Sanglier	2013	3	24099	Champeaux-et-la-Chapelle-Pommier
13Z010142	Sanglier	2013	3	24391	Saint-Crépin-de-Richemont
14Z009345	Sanglier	2014	3	24553	Tocane-Saint-Apre
15Z004344	Blaireau	2015	3	24319	Paussac-et-Saint-Vivien
08-00989-00	Bovin	2008	3	24451	Saint-Martial-de-Valette
16Z004569	Blaireau	2016	3	24099	Champeaux-et-la-Chapelle-Pommier
16Z004699	Blaireau	2016	3	24235	Léguillac-de-Cercles
16Z007121	Blaireau	2016	3	24248	Lussas-et-Nontronneau
01-07278-00	Bovin	2001	3	24474	Saint-Pancrace
09-01046-00	Bovin	2009	3	24530	Sencenac-Puy-de-Fourches
11Z002529	Bovin	2011	3	24579	Vieux-Mareuil
17Z006451	Blaireau	2017	3	24283	Monsec
17Z008476	Blaireau	2017	3	24166	Eyliac
05-03037-00	Bovin	2005	4	24099	Champeaux-et-la-Chapelle-Pommier
05-03263-00	Bovin	2005	4	24099	Champeaux-et-la-Chapelle-Pommier
05-03264-00	Bovin	2005	4	24099	Champeaux-et-la-Chapelle-Pommier
05-04174-00	Bovin	2005	4	24503	Saint-Sulpice-de-Mareuil
06-00856-00	Bovin	2006	4	24503	Saint-Sulpice-de-Mareuil
07-00027-00	Bovin	2007	4	24079	Cantillac
07-00306-00	Bovin	2007	4	24079	Cantillac
07-00596-00	Bovin	2007	4	87191	Sereilhac
11Z006170	Bovin	2011	4	24221	Rudeau Ladosse
17Z007283	Bovin	2017	5	24137	Coulaures
17Z002865	Bovin	2017	5	87060	Dournazac
17Z004428	Bovin	2017	5	87060	Dournazac
17Z005394	Bovin	2017	5	87060	Dournazac
17Z006454	Bovin	2017	5	87112	Pageas
17Z007631	Bovin	2017	5	87112	Pageas
17Z008463	Bovin	2017	5	87112	Pageas
17Z010552	Sanglier	2017	5	87	
01-09608-00	Bovin	2001	6	24498	Saint-Saud-Lacoussière
04-01062-00	Bovin	2004	6	24582	Villars
15Z004068	Bovin	2015	6	24498	Saint-Saud-Lacoussière
16Z005204	Bovin	2016	6	24498	Saint-Saud-Lacoussière
17Z000064	Sanglier	2017	6	24428	Saint-Jory-de-Chalais
17Z001046	Bovin	2017	6	24498	Saint-Saud-Lacoussière
17Z001576	Bovin	2017	6	24498	Saint-Saud-Lacoussière

17Z001577	Bovin	2017	6	24498	Saint-Saud-Lacoussière
17Z002196	Bovin	2017	6	24498	Saint-Saud-Lacoussière
17Z002197	Bovin	2017	6	24498	Saint-Saud-Lacoussière
17Z002553	Bovin	2017	6	24498	Saint-Saud-Lacoussière
17Z002744	Bovin	2017	6	24498	Saint-Saud-Lacoussière
17Z003110	Bovin	2017	6	24498	Saint-Saud-Lacoussière
17Z003236	Bovin	2017	6	24498	Saint-Saud-Lacoussière
17Z003245	Bovin	2017	6	24498	Saint-Saud-Lacoussière
17Z003570	Bovin	2017	6	24498	Saint-Saud-Lacoussière
17Z004077	Bovin	2017	6	24498	Saint-Saud-Lacoussière
17Z004080	Bovin	2017	6	24498	Saint-Saud-Lacoussière
17Z004088	Bovin	2017	6	24498	Saint-Saud-Lacoussière
17Z004275	Bovin	2017	6	24498	Saint-Saud-Lacoussière
17Z005047	Bovin	2017	6	24498	Saint-Saud-Lacoussière
17Z005495	Bovin	2017	6	24498	Saint-Saud-Lacoussière
17Z005496	Bovin	2017	6	24498	Saint-Saud-Lacoussière
17Z005671	Bovin	2017	6	24498	Saint-Saud-Lacoussière
17Z005672	Renard	2017	6	24498	Saint-Saud-Lacoussière
17Z006770	Bovin	2017	6	24498	Saint-Saud-Lacoussière
17Z001581	Bovin	2017	7	24134	Corgnac-sur-l'Isle
16Z005209	Blaireau	2016	7	24137	Coulaures
17Z002146	Bovin	2017	7	24137	Coulaures
17Z002198	Bovin	2017	7	24137	Coulaures
17Z002740	Bovin	2017	7	24137	Coulaures
09-01041-00	Bovin	2009	7	24171	Eyzerac
09-01353-00	Bovin	2009	7	24171	Eyzerac
16Z006257	Blaireau	2016	7	24103	Le Change
17Z008467	Blaireau	2017	7	24103	Le Change
08-01232-00	Bovin	2008	7	24308	Négrondes
08-01460-00	Bovin	2008	7	24308	Négrondes
15Z006541	Blaireau	2015	7	24308	Négrondes
17Z008478	Blaireau	2017	7	24308	Négrondes
17Z005395	Bovin	2017	7	24448	Saint-Martial-d'Albarède
07-01117-00	Bovin	2007	7	24453	Saint-Martin-de-Fressengeas
07-01384-00	Bovin	2007	7	24453	Saint-Martin-de-Fressengeas
07-01910-00	Bovin	2007	7	24453	Saint-Martin-de-Fressengeas
10Z005043	Bovin	2010	7	24453	Saint-Martin-de-Fressengeas
10Z008545	Bovin	2010	7	24453	Saint-Martin-de-Fressengeas
12Z004751	Bovin	2012	7	24453	Saint-Martin-de-Fressengeas
13Z003478	Bovin	2013	7	24453	Saint-Martin-de-Fressengeas
13Z005787	Bovin	2013	7	24453	Saint-Martin-de-Fressengeas
06-01095-00	Bovin	2006	7	24496	Saint-Romain-et-Saint-Clément
10Z009733	Blaireau	2010	7	24496	Saint-Romain-et-Saint-Clément
14Z005351	Bovin	2014	7	24521	Sarliac-sur-l'Isle
14Z005364	Bovin	2014	7	24521	Sarliac-sur-l'Isle

13Z003643	Bovin	2013	7	24540	Sorges et Ligueux en Périgord
13Z004129	Bovin	2013	7	24540	Sorges et Ligueux en Périgord
17Z002194	Bovin	2017	7	24540	Sorges et Ligueux en Périgord
03-02908-00	Bovin	2003	7	24551	Thiviers
03-04656-00	Bovin	2003	7	24551	Thiviers
03-04658-00	Bovin	2003	7	24551	Thiviers
05-02879-00	Bovin	2005	7	24551	Thiviers
05-02880-00	Bovin	2005	7	24551	Thiviers
05-02950-00	Bovin	2005	7	24551	Thiviers
08-00803-00	Bovin	2008	7	24551	Thiviers
12Z003442	Bovin	2012	7	24551	Thiviers
13Z010139	Bovin	2013	7	24551	Thiviers
17Z006498	Bovin	2017	7	24551	Thiviers
14Z004188	Bovin	2014	8	24218	Jumilhac-le-Grand
16Z004272	Blaireau	2016	8	24218	Jumilhac-le-Grand
15Z002980	Bovin	2015	8	24133	La Coquille
16Z006707	Bovin	2016	8	87096	La Meyze
16Z007479	Bovin	2016	8	87096	La Meyze
14Z008376	Bovin	2014	8	87082	Ladignac le Long
15Z006243	Bovin	2015	8	87082	Ladignac le Long
17Z000914	Bovin	2017	8	87082	Ladignac le Long
17Z003235	Bovin	2017	8	87082	Ladignac le Long
17Z004423	Bovin	2017	8	87082	Ladignac le Long
16Z005206	Bovin	2016	8	87031	Le Chalard
16Z007478	Bovin	2016	8	87031	Le Chalard
17Z002151	Bovin	2017	8	24304	Nantheuil
17Z004128	Bovin	2017	8	24304	Nantheuil
08-01321-00	Bovin	2008	8	24408	Saint-Front-d'Alemps
16Z005210	Bovin	2016	8	87027	Saint-Nicolas-Courbefy
08-02561-00	Bovin	2008	8	24489	Saint-Priest-les-Fougères
13Z001275	Blaireau	2013	8	24496	Saint-Romain-et-Saint-Clément
12Z004749	Bovin	2012	8	24551	Thiviers
16Z004560	Bovin	2016	8	24567	Vaunac
17Z010330	Sanglier	2017	8	87	
17Z004425	Bovin	2017	9	24096	Champagnac de Belair
17Z005388	Bovin	2017	9	24096	Champagnac de Belair
14Z008498	Sanglier	2014	9	24129	Condat-sur-Trincou
14Z000006	Sanglier	2014	9	24107	La Chapelle-Faucher
15Z005833	Bovin	2015	9	24107	La Chapelle-Faucher
17Z006771	Blaireau	2017	9	24107	La Chapelle-Faucher
14Z006581	Bovin	2014	9	24238	Lempzours
14Z007947	Bovin	2014	9	24238	Lempzours
14Z008062	Bovin	2014	9	24238	Lempzours
17Z001578	Bovin	2017	9	24540	Ligueux
17Z003568	Bovin	2017	9	24540	Ligueux

12Z004999	Blaireau	2012	9	24308	Négrondes
12Z007219	Bovin	2012	9	24308	Négrondes
12Z003373	Blaireau	2012	9	24408	Saint-Front-d'Alemps
12Z004182	Blaireau	2012	9	24408	Saint-Front-d'Alemps
13Z002277	Bovin	2013	9	24408	Saint-Front-d'Alemps
16Z000964	Sanglier	2016	9	24425	Saint-Jean-de-Côle
15Z002493	Bovin	2015	9	24429	Saint-Jory-las-Bloux
09-01094-00	Bovin	2009	9	24485	Saint-Pierre-de-Côle
13Z010140	Bovin	2013	9	24551	Thiviers
14Z000010	Bovin	2014	9	24551	Thiviers
14Z000016	Bovin	2014	9	24551	Thiviers
12Z005339	Blaireau	2012	9	24567	Vaunac
13Z004051	Bovin	2013	9	24582	Villars
13Z004667	Bovin	2013	9	24582	Villars
13Z005631	Bovin	2013	9	24582	Villars
14Z005338	Bovin	2014	9	24582	Villars
15Z005356	Renard	2015	9	24582	Villars
15Z006272	Renard	2015	9	24582	Villars
15Z008880	Renard	2015	9	24582	Villars
17Z004525	Bovin	2017	9	24582	Villars
06-01345-00	Bovin	2006	10	24458	Saint-Martin-le-Pin
06-01578-00	Bovin	2006	10	24458	Saint-Martin-le-Pin
06-01779-00	Bovin	2006	10	24458	Saint-Martin-le-Pin
06-02226-00	Bovin	2006	10	24458	Saint-Martin-le-Pin
06-02227-00	Bovin	2006	10	24458	Saint-Martin-le-Pin
15Z004377	Blaireau	2015	10	24548	Teyjat
15Z010210	Bovin	2015	10	24548	Teyjat
16Z000595	Bovin	2016	10	24548	Teyjat
16Z001097	Bovin	2016	10	24548	Teyjat
05-03418-00	Bovin	2005	11	24002	Agonac
05-03419-00	Bovin	2005	11	24002	Agonac
06-01486-00	Bovin	2006	11	24002	Agonac
06-01487-00	Bovin	2006	11	24002	Agonac
06-02101-00	Bovin	2006	11	24002	Agonac
15Z004335	Bovin	2015	11	24016	Augignac
12Z007111	Blaireau	2012	11	24042	Biras
13Z002208	Bovin	2013	11	24042	Biras
14Z004539	Bovin	2014	11	24042	Biras
12Z000029	Blaireau	2012	11	24055	Bourdeilles
15Z004939	Blaireau	2015	11	24055	Bourdeilles
12Z002503	Blaireau	2012	11	24064	Brantôme
14Z005365	Bovin	2014	11	24069	Bussac
14Z006089	Blaireau	2014	11	24069	Bussac
14Z006436	Bovin	2014	11	24069	Bussac
06-02579-00	Bovin	2006	11	24079	Cantillac

15Z005359	Blaireau	2015	11	24079	Cantillac
16Z004566	Bovin	2016	11	24101	Champs Romain
13Z001729	Bovin	2013	11	24115	Chateau l'Eveque
13Z001730	Bovin	2013	11	24115	Chateau l'Eveque
14Z004540	Bovin	2014	11	24115	Chateau l'Eveque
13Z002699	Bovin	2013	11	24144	Creyssac
13Z004012	Bovin	2013	11	24144	Creyssac
11Z001496	Chevreuil	2011	11	24154	Douchapt
13Z000269	Blaireau	2013	11	24170	Eyvirat
13Z010141	Sanglier	2013	11	24108	La Chapelle-Gonaguet
07-01014-00	Bovin	2007	11	24235	Léguillac-de-Cercles
03-01690-00	Bovin	2003	11	24243	Lisle
03-04294-00	Bovin	2003	11	24243	Lisle
17Z001177	Bovin	2017	11	24243	Lisle
13Z003769	Bovin	2013	11	24271	Milhac-de-Nontron
08-00609-00	Bovin	2008	11	24319	Paussac-et-Saint-Vivien
12Z004742	Bovin	2012	11	24319	Paussac-et-Saint-Vivien
15Z004334	Blaireau	2015	11	24391	Saint-Crépin-de-Richemont
14Z000015	Bovin	2014	11	24410	Saint-Front-la-Rivière
06-01468-00	Bovin	2006	11	24064	Saint-Julien-de-Bourdeilles
06-01582-00	Bovin	2006	11	24064	Saint-Julien-de-Bourdeilles
06-01843-00	Bovin	2006	11	24064	Saint-Julien-de-Bourdeilles
17Z005849	Bovin	2017	11	24453	Saint-Martin-de-Fressengeas
17Z007633	Bovin	2017	11	24453	Saint-Martin-de-Fressengeas
17Z007634	Bovin	2017	11	24453	Saint-Martin-de-Fressengeas
15Z006492	Blaireau	2015	11	24485	Saint-Pierre-de-Côle
09-01683-00	Bovin	2009	11	24489	Saint-Priest-les-Fougères
12Z005660	Blaireau	2012	11	24503	Saint-Sulpice-de-Mareuil
16Z004274	Bovin	2016	11	24503	Saint-Sulpice-de-Mareuil
16Z007476	Bovin	2016	11	24503	Saint-Sulpice-de-Mareuil
16Z007477	Bovin	2016	11	24503	Saint-Sulpice-de-Mareuil
16Z007681	Bovin	2016	11	24503	Saint-Sulpice-de-Mareuil