

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**

**PHYLOGENIE ET DISTRIBUTION GEOGRAPHIQUE DE *BORRELIA GARINII*  
CHEZ DES OISEAUX COMMUNS EN FRANCE**

**Présenté par**

Margaux FRANCOIS

Réalisé sous la direction de : Maud MARSOT et Laetitia CANINI

Organisme et pays : Unité Epidémiologie, Laboratoire Santé Animale, ANSES Maisons-Alfort, FRANCE

Période du stage : 02/01/2023 – 30/06/2023

Date de soutenance : 26/06/2023

Année universitaire 2022-2023

## Remerciements :

A Laetitia Canini et Maud Marsot,

Pour leur patience, leur soutien et leur bienveillance,

A Benoît Durand, Gina Zanella, Clémence Nadal, Alex Drouin, Hélène Duault, Amaïas Avalos,

Pour la chaleur de leur accueil et leur bonne humeur,

A l'équipe pédagogique du master SEMHA,

Pour leur implication et leur disponibilité durant cette année,

Aux membres du projet OUTLYER : Amalia Rataud, Sara Moutailler, Clémence Galon, Pierre-Yves Henry, Benoit Pisanu, Laure Bournez, Alex Drouin, Nicolas Dubos, Romain Provost, Pierre Fiquet, et aux bagueurs bénévoles,

Sans qui ce travail n'aurait pu avoir lieu,

A Pascal Boireau,

Pour m'avoir permis de réaliser mon stage au Laboratoire Santé Animale,

A l'ANSES et Dim1Health,

Pour avoir financé ce projet.

## Résumé court :

La maladie de Lyme est une maladie transmise à l'homme *via* la morsure d'une tique. Elle est due à un complexe de bactéries appartenant au genre *Borrelia*, appelé *Borrelia burgdorferi sensu lato* (*Bbsl*). Certaines formes graves de la maladie sont majoritairement causées par des espèces spécifiques au sein des *Bbsl* : *Borrelia garinii* est ainsi la bactérie majoritairement responsable de neuroborrélioses chez l'homme en Europe. Cette bactérie a pour réservoir les oiseaux. Le but de notre étude était d'étudier la diversité génétique et la répartition spatiale des *Bbsl* détectées chez des larves gorgées collectées sur des oiseaux en France, en se concentrant principalement sur *B. garinii*. Pour ce faire, nous avons utilisé les séquences génétiques de 96 *Bbsl* (dont 74 *B. garinii*), correspondant à un fragment du gène *flaB* de ces bactéries. Ces *Bbsl* provenaient de larves gorgées collectées chez des oiseaux capturés dans différents sites en France dans le cadre d'un projet précédent (OUTLYER). *B. garinii* était retrouvée dans toute la France. La matrice des distances génétiques montrait une grande proximité génétique entre les séquences étudiées, les séquences provenant de *Bbsl* appartenant à une même espèce étant les plus proches. L'arbre phylogénétique construit par la méthode du maximum de parcimonie à partir des séquences des *Bbsl* et l'étude des métadonnées associées n'ont pas mis en évidence de regroupement de bactéries génétiquement proches (clades) pour le gène séquencé en fonction de l'espèce d'oiseau ou du milieu de vie de celui-ci. Une majorité de Merles noirs était cependant retrouvée au sein de certains clades, alors que les *Bbsl* présentes dans d'autres clades provenaient de larves gorgées collectées sur une plus grande diversité d'espèces aviaires. Aucune structure spatiale n'a été visualisée lors de l'étude de la répartition spatiale des *Bbsl* provenant des différents clades identifiés grâce à l'arbre phylogénétique : les *B. garinii* génétiquement proches pour le gène séquencé (*flaB*) ne semblaient pas être spatialement proches. La majorité de Merles noirs et de milieux clairsemés et landes et broussailles au sein de notre échantillon étaient des facteurs limitants la visualisation d'une structure. Il serait intéressant de reprendre le protocole mis en place lors du projet OUTLYER, mais en collectant uniquement des tiques sur les espèces aviaires dont le rôle de réservoir est connu ou suspecté pour *B. garinii*. Cela permettrait d'explorer plus en détail le lien entre diversité génétique des *Bbsl* et espèce aviaire hôte. Le gène séquencé (*flaB*) est un gène de ménage, souvent utilisé pour identifier les différentes espèces de *Bbsl*. Ainsi, ce gène est conservé au sein d'une espèce, ce qui explique la faible diversité génétique observée pour les séquences de *Bbsl* étudiées. De ce fait, il est possible que la diversité génétique observée dans cette étude pour le gène *flaB* ne reflète pas la diversité génétique existant entre les génomes des *Bbsl* étudiées. Il serait donc intéressant de reprendre cette étude en privilégiant un protocole utilisé par de nombreuses études sur *Bbsl* : ce protocole consiste à séquencer entièrement ou en partie 8 gènes de ménage, ce qui permettrait d'obtenir une plus grande variabilité génétique entre les séquences de *Bbsl* étudiées.

Mot-clés : phylogénie, arbre phylogénétique, *Bbsl*, *Borrelia garinii*, oiseaux, France, tique.

## Résumé long :

La maladie de Lyme est une maladie vectorielle due à des bactéries du genre *Borrelia* regroupées dans un complexe appelé *Borrelia burgdorferi sensu lato* (noté *Bbsl*). C'est la maladie vectorielle la plus fréquente dans l'hémisphère Nord, et elle est transmise *via* une morsure de tique. En France, le nombre de cas humains de maladie de Lyme diagnostiqués estimés par an est compris entre 25 000 et 68 530 sur la période 2009-2020, avec une tendance à l'augmentation depuis 2009. Cette maladie touche inégalement toutes les régions de France métropolitaine. Son diagnostic n'est pas toujours aisé, et des retards de prise en charge peuvent survenir et altérer la qualité de vie des malades.

Le complexe des *Bbsl* comprend une vingtaine d'espèces, toutes ne causant pas de maladie chez l'homme. Ces bactéries peuvent être assez différentes d'un point de vue éco-épidémiologique (différences au niveau des espèces réservoirs, des symptômes causés...). En Europe, les principales espèces de *Bbsl* infectant l'humain sont *B. burgdorferi sensu stricto* (ss), *B. afzelii* et *B. garinii*. Les symptômes développés par l'homme sont généralement bénins (érythème migrant), mais la maladie peut évoluer et prendre différentes formes. La fréquence et la gravité de ces formes varient avec l'espèce de *Bbsl* concernée. *B. burgdorferi* ss est ainsi plus fréquemment responsable d'arthrites, *B. afzelii* de manifestations cutanées, tandis que *B. garinii* et *B. bavariensis* sont plus fréquemment retrouvés en cas de neuroborréliose, affection pouvant nécessiter une hospitalisation. Les neuroborrélioses sont des formes particulièrement invalidantes de la maladie, car elles peuvent toucher les systèmes nerveux central et périphérique d'un individu et provoquer de nombreux symptômes, comme par exemple une paralysie faciale.

Les *Bbsl* sont transmises à l'homme *via* une tique infectée. Les tiques sont des acariens ectoparasites hématophages stricts de vertébrés, leur cycle parasitaire comprend trois stases : larve-nymphé-adulte. En Europe, *Ixodes ricinus* est l'espèce de tique majoritaire : cette tique va réaliser un unique repas sanguin lors de chaque stase, sur un hôte différent à chaque fois. *I. ricinus* peut se nourrir sur une grande variété d'hôtes, mais les larves et les nymphes de cette espèce se nourrissent principalement sur les rongeurs et les oiseaux, alors que les adultes se nourrissent principalement sur les grands mammifères (cervidés). Ainsi, une tique s'infecte la plupart du temps en réalisant un repas sanguin sur un hôte réservoir porteur de *Bbsl*.

Un réservoir peut être défini comme une population (ou un ensemble de populations épidémiologiquement connectées) dans lequel un agent pathogène est maintenu de manière persistante et depuis lequel l'infection est transmise à une population cible, différente ou non du réservoir. Ainsi, une population (ici une espèce) réservoir pour *Bbsl* est une espèce capable d'être infectée par ces bactéries *via* une morsure de tique, d'assurer leur réplication et de les transmettre ensuite à une tique naïve. Les espèces réservoirs préférentielles des *Bbsl* varient en fonction des espèces de bactéries concernées : *B. garinii* a pour réservoirs principaux les oiseaux. Certaines espèces sont connues pour leur importance en tant que réservoir de *B. garinii* : c'est le cas notamment du Merle noir (*Turdus merula*).

De plus en plus d'études soulignent l'importance de caractériser la diversité génétique des *Bbsl*, car cela permet de mieux comprendre l'histoire évolutive, la pathogénicité et l'épidémiologie de ces bactéries. La phylogénie est l'étude des liens existant entre espèces/individus/gènes apparentés (taxa). Les arbres phylogénétiques sont des représentations graphiques qui permettent le classement des différents taxa et la déduction de leur histoire évolutive. La phylogénie moléculaire est la discipline qui étudie les relations entre les taxa en se basant sur des molécules telles que l'ADN/ARN ou les protéines, elle est de plus en plus utilisée ces dernières années. La phylogénie moléculaire permet de comparer la diversité génétique et les relations entre différentes séquences génétiques de *Bbsl*, et a notamment contribué à identifier de nouvelles espèces au sein de ce complexe. De plus, la phylogénie permet de mieux appréhender l'influence des facteurs biotiques (vecteurs et hôtes potentiels) et abiotiques (distribution géographique) sur la diversité génétique des *Bbsl*. L'épidémiologie moléculaire est la discipline qui utilise des données moléculaires (ADN, ARN...) pour étudier la distribution et les déterminants des maladies chez les populations humaines et/ou animales. L'épidémiologie moléculaire peut s'appuyer sur la phylogénie, notamment lorsque la comparaison des séquences de virus ou de bactéries est mise en relation avec des données épidémiologiques (distribution spatiale de l'agent pathogène, caractéristiques des hôtes, etc). En épidémiologie moléculaire, des arbres phylogénétiques peuvent par exemple être construits pour comparer des souches virales ou bactériennes nouvellement isolées à un endroit donné à d'autres souches précédemment identifiées issues d'autres endroits. Ces arbres permettent entre autres de mettre en évidence un lien potentiel entre proximité génétique et distribution spatiale des agents pathogènes étudiés.

Plusieurs méthodes existent pour construire un arbre phylogénétique. La méthode du maximum de parcimonie utilise un algorithme qui va générer de nombreux arbres phylogénétiques, et choisir celui avec le plus petit score de parcimonie, c'est-à-dire l'arbre comportant le moins de substitutions. C'est une méthode qui est rapide à mettre en place, est qui est adaptée dans le cas de séquences faiblement divergentes. Les études phylogénétiques réalisées sur *B. garinii* tendent à montrer que cette bactérie ne présente pas de structure spatiale (hors échelle supérieure au continent) : les bactéries génétiquement proches ne sont pas spatialement proches. *B. garinii* est aussi présentée comme l'une des *Bbsl* chez qui l'on retrouve la plus grande variabilité génétique. Or, la diversité génétique des *Bbsl* peut influencer sur la pathogénicité et sur la diversité des symptômes chez l'homme, qui peuvent varier même au sein d'une espèce. Étudier la diversité génétique de *B. garinii* est donc nécessaire, car c'est l'une des *Bbsl* pour laquelle on retrouve le plus de diversité, et qu'elle est responsable de graves symptômes chez l'homme. A notre connaissance, aucune étude ne s'est concentrée sur la diversité de *Borrelia garinii* portée par les oiseaux en France.

Le but de notre étude était d'étudier la diversité génétique et la répartition spatiale des *Bbsl* détectées chez des larves gorgées collectées sur des oiseaux capturés dans différents sites en France, en se concentrant principalement sur *B. garinii*. Les séquences d'ADN des différentes *Bbsl* détectées chez ces larves gorgées étaient disponibles grâce à un projet précédent (OUTLYER). Le projet OUTLYER avait notamment pour objectif de détecter différents agents pathogènes présents chez des larves gorgées collectées sur des oiseaux en France. Chaque séquence d'ADN de *Bbsl* disponible était associée à des métadonnées telles que l'espèce et la stase de la tique positive à *Bbsl*, l'espèce et l'identifiant de l'oiseau sur lequel la tique avait été collectée, ou encore le milieu de vie de l'oiseau ainsi que les coordonnées de son lieu de capture. Au total, 1106 larves gorgées avaient été collectées sur 442 oiseaux appartenant à 36 espèces, dans 69 sites en France. 129 larves étaient positives à *Bbsl*, dont 100 à *B. garinii*. 121 *Bbsl* et leurs séquences étaient initialement disponibles, ces séquences correspondant à un fragment du gène *flaB* des *Bbsl*. Ce gène est un gène de ménage, qui est conservé au sein des différentes espèces de *Bbsl*. *FlaB* est régulièrement utilisé pour identifier les différentes espèces de *Bbsl*, comme cela a été fait lors du projet OUTLYER.

Dans le cadre de cette étude, les séquences génétiques des 121 *Bbsl* disponibles ont été nettoyées puis alignées, afin de réaliser les analyses phylogénétiques. A partir de l'alignement obtenu, une matrice des distances, qui compare la distance génétique entre les différentes paires de séquences pour l'ensemble des séquences, a été créée. Un arbre phylogénétique a été construit selon le principe du maximum de parcimonie, à partir des séquences génétiques des *Bbsl* issues des larves gorgées collectées sur les oiseaux. Les métadonnées associées aux larves gorgées ont été ajoutées à cet arbre : espèce de *Bbsl* correspondant à la séquence génétique, espèce d'oiseau dont provenait la larve gorgée, milieu de vie de l'oiseau. De plus, les larves issues d'un même oiseau ont été représentées sur l'arbre phylogénétique, lorsque l'oiseau était porteur de plus de deux larves positives à *Bbsl*. Grâce à cet arbre, des clades (groupes constitués par un ancêtre commun et l'ensemble de ses descendants) ont été identifiés, et les séquences composant ces clades ont été géolocalisées sur la carte de France.

*B. garinii* était globalement répartie dans toutes les zones de France où des larves gorgées avaient pu être collectées. 16% des oiseaux capturés sur lesquels des larves ont été collectées et analysées hébergeaient au moins une larve positive à *Bbsl*. Des larves positives à *Bbsl* ont été collectées sur 15 espèces d'oiseaux (sur les 36 échantillonnées). Le Merle noir et la Grive musicienne étaient les deux seules espèces dont le pourcentage d'individus positifs (c'est-à-dire porteurs d'au moins une larve positive à *Bbsl*) était supérieur à 50% lorsque la taille de l'échantillon était supérieure à 6 (avec des pourcentages respectifs de 67% et 60% d'individus positifs sur les 43 et 20 individus porteurs de larves gorgées). Ces deux espèces présentaient aussi un pourcentage de larves positives élevé au niveau individuel, avec 75% (15/20) des Merles noirs porteurs de 100% de larves positives à *Bbsl* contre 66% (4/6) pour la Grive musicienne (pour un nombre de larves collectées par oiseau variant de 2 à 8, les oiseaux porteurs d'une seule larve n'ayant pas été pris en compte dans ce calcul).

Les analyses phylogénétiques ont été réalisées à partir de 96 séquences génétiques de *Bbsl* (25 séquences ont été exclues des analyses lors du nettoyage des données), dont 74 de *B. garinii*. La majorité de ces séquences génétiques provenaient de larves gorgées collectées sur des Merles noirs (56/96) et sur des Grives musiciennes (15/96), et étaient majoritairement issues de milieux clairsemés (39/96) ou de milieux type « landes et broussailles » (29/96). Les séquences des *Bbsl* étudiées (fragment de *flaB*) étaient génétiquement proches : 21% des paires de séquences étudiées avaient une distance génétique nulle, et 93% une distance génétique inférieure à 5%. La distance génétique maximale observée au sein de notre échantillon était de 14%. Les séquences les plus proches génétiquement étaient issues de *Bbsl* appartenant à une même espèce, et des sous-groupes étaient

distinguables chez *B. garinii*. L'arbre phylogénétique obtenu par maximum de parcimonie comportait des multifurcations (absence de discrimination des bactéries au sein d'un groupe). Deux clades A et B ont pu être visualisés (composés respectivement de 75 et 18 séquences génétiques), ainsi que deux sous clades A1 et A2 (composés respectivement de 26 et 24 séquences génétiques). Les métadonnées associées à cet arbre ont montré que les deux clades A et B correspondaient à différentes espèces de *Bbsl*, le clade A regroupant les séquences génétiques correspondant à *B. garinii*, et le clade B formant deux sous clades, le premier regroupant majoritairement *B. valaisiana*, et le second *B. turdi*. Les séquences génétiques les plus proches sur l'arbre étaient donc issues d'une même espèce de *Bbsl*. Il n'y avait pas de regroupement strict des bactéries en fonction de l'espèce d'oiseau dont les larves étaient issues ou du milieu de vie de l'oiseau. Cependant, le grand nombre de Merles noirs et de milieux clairsemés ou landes et broussailles dans notre échantillon limitait la visualisation d'une structure. Les *B. garinii* du sous clade A1 (n = 26) étaient issues de diverses espèces d'oiseaux (8 espèces différentes, avec 6 larves collectées sur des Merles noirs, 6 larves collectées sur des Grives musciennes, et 6 larves collectées sur des Troglodytes mignons, les larves issues des autres espèces étant minoritaires). Cette diversité d'espèces aviaires allait de pair avec une diversité de milieux de vie (5 milieux différents, avec une majorité de milieux clairsemés (9), landes et broussailles (5) et forêt (5)). Les *B. garinii* du sous clade A2 (n = 24) étaient majoritairement issues de larves collectées sur des Merles noirs, et ces larves provenaient en majorité de milieux clairsemés (20 larves collectées sur des Merles noirs, et 17 larves issues de milieux clairsemés). On distinguait donc des clades où les Merles noirs étaient plutôt regroupés, et des clades avec une diversité aviaire plus importante. Globalement, les séquences génétiques des *Bbsl* issues des larves portées par un même oiseau semblaient proches entre elles. Ce n'était cependant pas toujours le cas, notamment quand les séquences étaient issues de différentes espèces de *Bbsl*. Enfin, aucune structure spatiale n'a été visualisée lors de l'étude de la répartition des *Bbsl* issues des différents clades identifiés grâce à l'arbre phylogénétique : les *B. garinii* génétiquement proches ne semblaient pas être spatialement proches, et les séquences génétiques provenant de différents clades étaient retrouvées dans des mêmes sites.

La majorité de Merles noirs présents dans notre échantillon peut s'expliquer par l'abondance de cette espèce en France (c'est l'une des espèces les plus capturées au niveau des sites de captures d'où proviennent nos données), mais aussi par le fait que le Merle noir est un réservoir majeur pour *B. garinii*. La majorité de milieux clairsemés et landes et broussailles peut s'expliquer par le fait que les sites de capture des oiseaux sont majoritairement situés dans ces milieux. Ces deux éléments nous limitent dans l'analyse du lien entre diversité génétique des *Bbsl* et espèce aviaire/milieu de vie de l'hôte. Cependant, la seule étude disponible à notre connaissance ne met pas non plus en évidence de regroupement de *B. garinii* en fonction de l'espèce d'oiseau dont les larves sont issues. Il serait intéressant de réaliser un test statistique afin de vérifier si le pourcentage de Merles noirs est significativement différent entre les différents clades observés. Un protocole similaire à celui réalisé dans le cadre du projet OUTLYER pourrait aussi être mis en place, mais en collectant uniquement des tiques chez les espèces aviaires dont le rôle de réservoir est suspecté ou prouvé pour *B. garinii*. Cela permettrait d'explorer plus en détail le lien entre diversité génétique de *B. garinii* et espèce aviaire hôte. Si l'existence de clades de *B. garinii* majoritairement issues de Merles noirs venait à être confirmée par d'autres études, cela pourrait indiquer l'existence d'une circulation préférentielle de certaines souches de *B. garinii* au sein de cette espèce, peut-être à cause de différences génétiques. Nos résultats en faveur d'une absence de structure spatiale chez *B. garinii* sont cohérents avec les autres études réalisées : hors échelle supérieure au continent, les *B. garinii* proches génétiquement ne sont pas regroupées sur un même territoire. Cependant, la présence de zones étendues sans sites pour collecter des données pour notre étude nous empêche d'avoir une vision complète concernant la répartition spatiale de ces bactéries. De plus, l'analyse de la structure spatiale de *B. garinii* a été uniquement visuelle. Il serait pertinent de compléter cette analyse visuelle par une analyse statistique de détection de cluster pour confirmer l'absence de regroupement spatial des bactéries génétiquement proches. L'arbre phylogénétique construit par maximum de parcimonie comportait de nombreuses multifurcations, qui peuvent s'expliquer par la forte proximité génétique de certains groupes de séquences, empêchant l'algorithme de discriminer les *Bbsl* présentes au sein d'un groupe. Le gène séquencé (*flaB*) est un gène de ménage, souvent utilisé pour identifier les différentes espèces de *Bbsl*. Ainsi, ce gène est conservé au sein d'une espèce, ce qui explique la faible diversité génétique retrouvée pour les séquences de *Bbsl* étudiées. De ce fait, il est possible que la diversité génétique observée dans cette étude pour le gène *flaB* ne reflète pas la diversité génétique existant entre les génomes des *Bbsl* étudiées. Il serait donc intéressant de reprendre cette étude en privilégiant cette fois le protocole défini par Margos et al (Margos et al. 2008), et repris par de nombreuses études : ce protocole consiste à réaliser un MLST (MultiLocus Sequence Typing) sur 8 gènes de ménage (séquençage de tout ou partie

d'un ensemble de gènes). Cela permettrait d'obtenir une plus grande variabilité génétique chez les *Bbs* étudiées.

Cette étude a permis de réaliser une première description de la répartition spatiale et de la diversité génétique de *B. garinii* chez les oiseaux en France, ce qui constitue une première étape vers une meilleure compréhension de l'influence de la répartition spatiale et du rôle des oiseaux dans la diversité génétique de cette bactérie. Ces résultats pourront servir de base à d'autres études et aider à une meilleure compréhension du risque associé à cet agent pathogène à l'avenir.

## Table des matières

<b>INTRODUCTION :</b> .....	<b>11</b>
<b>PARTIE 1 : SYNTHÈSE BIBLIOGRAPHIQUE</b> .....	<b>12</b>
<b>1. LA MALADIE DE LYME, UNE MALADIE VECTORIELLE COMPLEXE :</b> .....	<b>12</b>
1.1. IMPORTANCE EN SANTÉ PUBLIQUE.....	12
1.2. AGENTS PATHOGÈNES RESPONSABLES DE LA MALADIE DE LYME.....	12
1.3. <i>IXODES RICINUS</i> , PRINCIPALE TIQUE VECTRICE DES <i>BBSL</i> EN EUROPE .....	12
1.4. LES OISEAUX, PRINCIPAUX RÉSERVOIRS DE <i>BORRELIA GARINII</i> .....	14
<b>2. PHYLOGÉNIE DES <i>BBSL</i> :</b> .....	<b>15</b>
2.1. DESCRIPTION DU GÉNOME DES <i>BBSL</i> .....	15
2.2. MÉTHODES DE SÉQUENÇAGE DE <i>BBSL</i> .....	15
2.3. PHYLOGÉNIE ET ÉPIDÉMIOLOGIE MOLECULAIRE : LES PRINCIPES .....	16
2.4. ÉTAT DES LIEUX DE LA TAXONOMIE DE <i>BBSL</i> ET APPORTS DE LA PHYLOGÉNIE .....	20
2.5. LA PHYLOGÉNIE, UN OUTIL AU SERVICE D'UNE MEILLEURE COMPRÉHENSION DES <i>BBSL</i> .....	21
<b>PARTIE 2 : TRAVAIL PERSONNEL</b> .....	<b>22</b>
<b>1. MATÉRIELS ET MÉTHODES :</b> .....	<b>22</b>
1.1. COLLECTE DES DONNÉES .....	22
1.2. NETTOYAGE DES DONNÉES ET OBTENTION DES SÉQUENCES CONSENSUS POUR LES ANALYSES PHYLOGÉNÉTIQUES .....	26
1.3. ANALYSES PHYLOGÉNÉTIQUES.....	27
<b>2. RESULTATS :</b> .....	<b>28</b>
2.1. DESCRIPTION DES DONNÉES.....	28
2.2. ANALYSES DES DONNÉES : DIVERSITÉ GÉNÉTIQUE DES SÉQUENCES ET PHYLOGÉNIE .....	32
<b>3. DISCUSSION :</b> .....	<b>37</b>
<b>CONCLUSION :</b> .....	<b>40</b>
<b>REFERENCES BIBLIOGRAPHIQUES :</b> .....	<b>41</b>
<b>ANNEXES :</b> .....	<b>47</b>

## Liste des figures :

Figure 1 : Cycle de vie de la tique <i>Ixodes ricinus</i> (Lejal 2020).....	13
Figure 2 : Activité saisonnière d' <i>Ixodes ricinus</i> en Europe centrale, au nord de l'Angleterre et en Écosse (Kurtenbach et al. 2006) .....	14
Figure 3 : Exemple d'arbre phylogénétique.....	17
Figure 4 : Construction des arbres selon le principe de parcimonie .....	19
Figure 5 : Schéma de l'obtention des séquences génétiques des <i>Bbsl</i> (PCR et séquençage).....	25
Figure 6 : Schéma d'obtention d'une séquence consensus à partir des séquences F et R d'une <i>Bbsl</i> /26	
Figure 7 : Séquences non alignées (gauche) et alignement multiple des séquences (droite) .....	27
Figure 8 : Exemple d'arbre présentant une multifurcation, avec longueur de branche représentant la distance génétique entre deux nœuds .....	28
Figure 9 : Carte présentant les pourcentages de larves positives à <i>Bbsl</i> en fonction des stations en France.....	29
Figure 10 : Espèces d'oiseaux positives à <i>Bbsl</i> , et nombre d'individus porteurs de larves négatives et porteurs d'au moins une larve positive à <i>Bbsl</i> par espèce d'oiseau .....	30
Figure 11 : Nombre de de larves positives et négatives à <i>Bbsl</i> par oiseau en fonction de l'espèce ....	31
Figure 12 : Espèces de <i>Bbsl</i> correspondant aux 96 séquences génétiques (A), espèces d'oiseaux dont sont issues les 96 larves gorgées positives à <i>Bbsl</i> (B), milieux dont sont issues les 96 larves gorgées récoltées sur les 54 oiseaux capturés (C) .....	32
Figure 13 : Matrice des distances génétiques des séquences .....	33
Figure 14 : Arbre phylogénétique construit à partir des 96 séquences génétiques de <i>Bbsl</i> .....	35
Figure 15 : A) Répartition spatiale des <i>Bbsl</i> appartenant aux clades A et B, et des <i>Bbsl</i> hors clades, B) répartition spatiale des <i>B. garinii</i> appartenant aux sous clades A1 et A2 .....	36

Liste des tableaux :

Tableau 1 : Catégories utilisées pour l'habitat des oiseaux capturés, et lien avec les catégories officielles CLC.....	23
Tableau 2 : Nombre de larves récoltées en fonction de leur espèce (ensemble des larves, puis larves positives à <i>Bbsl</i> , puis larves positives à <i>B. garinii</i> ).....	28

## Introduction :

La maladie de Lyme est une maladie vectorielle due à des bactéries du genre *Borrelia* regroupées dans un complexe appelé *Borrelia burgdorferi sensu lato* (noté *Bbsl*). La maladie de Lyme est devenue ces dernières années une maladie d'intérêt médiatique et sanitaire en France. Avec un nombre de cas humains diagnostiqués estimés par an compris entre 25 000 et 68 530 sur la période 2009-2020, et une tendance à l'augmentation depuis 2009 (santepubliquefrance.fr), cette maladie, qui touche inégalement toutes les régions de France métropolitaine, est aujourd'hui un enjeu de santé publique. En effet, le diagnostic de cette maladie, qui peut être polymorphe, est parfois difficile : des erreurs ou des retards de prise en charge peuvent survenir, ce qui peut altérer la qualité de vie du patient et représenter un coût important pour le système de santé publique (Aucott et al. 2009; Cardenas-de la Garza et al. 2019). Une revue de plusieurs études s'intéressant au coût économique de la maladie de Lyme dans différents pays d'Europe a montré que ce coût pouvait varier et monter à plusieurs dizaines de millions d'euros dans certains pays, avec l'Allemagne en tête (coût annuel estimé de la maladie à plus de 30 millions d'euros, à partir de données récoltées entre 2007 et 2011) (Mac, da Silva, et Sander 2019).

Le complexe des *Bbsl* comprend un ensemble d'espèces bactériennes qui peuvent être assez différentes d'un point de vue éco-épidémiologique (différences au niveau des espèces réservoirs, des symptômes causés, etc) (Derdáková et Lencáková 2005). Les *Bbsl* sont transmises à l'homme *via* une morsure de tique, tique qui s'infecte la plupart du temps en réalisant un repas sanguin sur un hôte réservoir porteur de la bactérie. Les symptômes développés par l'homme sont généralement bénins (érythème migrant), mais la maladie peut évoluer et prendre différentes formes. La fréquence et la gravité de ces formes varient avec l'espèce de *Bbsl* concernée. Ainsi, *Borrelia garinii*, une bactérie ayant pour réservoir les oiseaux et fréquemment retrouvée chez la tique *Ixodes ricinus*, est majoritairement responsable chez l'homme (avec *B. bavariensis*), de formes neurologiques pouvant être invalidantes appelées neuroborrélioses (Coipan et al. 2016). C'est la *Bbsl* la plus fréquemment isolée chez les oiseaux en France (Rataud, Galon, et al. 2022). En 2020, d'après Santé Publique France, sur les 710 personnes hospitalisées suite à une maladie de Lyme, près de la moitié souffrait de neuroborréliose. Les deux autres bactéries majoritairement retrouvées lors d'infection chez l'homme sont *B. burgdorferi sensu stricto* (*ss*) et *B. afzelii*, toutes deux ont pour réservoirs principaux des rongeurs (Wolcott et al. 2021). *B. burgdorferi ss* est responsable d'arthrites, alors que *B. afzelii* est surtout responsable de manifestations cutanées.

De plus en plus d'études soulignent l'importance de s'intéresser à la diversité génétique des *Bbsl*, car cela permet de mieux comprendre l'histoire évolutive, la pathogénicité et l'épidémiologie de ces bactéries (Derdáková et Lencáková 2005; Seifert et al. 2015; Oppler et al. 2021). La phylogénie, *via* la construction d'arbres phylogénétiques, permet de comparer la diversité génétique et les relations entre différentes séquences génétiques de *Bbsl*, et a notamment contribué à identifier de nouvelles espèces au sein du complexe. De plus, elle permet de mieux appréhender l'influence des facteurs biotiques (vecteurs et hôtes potentiels) et abiotiques (distribution géographique) sur la diversité génétique de ces bactéries (Margos et al. 2011; Qiu et Martin 2014).

A notre connaissance, aucune étude ne s'est concentrée sur la diversité de *Borrelia garinii* portée par les oiseaux en France. Sachant que cette bactérie est retrouvée chez de nombreuses espèces d'oiseaux (Rataud, Galon, et al. 2022), il est intéressant d'étudier sa répartition spatiale en France, de chercher un lien éventuel avec la diversité génétique observée, puis de regarder si cette diversité génétique semble en lien avec des caractéristiques de l'espèce hôte ou de l'habitat.

Dans un premier temps, une recherche bibliographique a été effectuée, d'abord sur la maladie de Lyme et sur la description de l'éco-épidémiologie des *Bbsl*, en se concentrant principalement sur *B. garinii*. Ensuite, une description des principes phylogénétiques ainsi qu'une synthèse des apports de la phylogénie sur les connaissances relatives aux *Bbsl* ont été effectuées. Notre étude s'est d'abord focalisée sur la répartition géographique de *B. garinii* portée par des oiseaux en France. La diversité génétique de *B. garinii* a été étudiée grâce à une analyse phylogénétique par maximum de parcimonie, et le lien entre la phylogénie et la distribution géographique/l'espèce aviaire hôte/le milieu de vie de l'hôte a ensuite été exploré. Les autres espèces de *Bbsl* retrouvées dans cette étude ont été incluses dans les analyses, car elles peuvent aussi circuler chez les oiseaux, mais de façon minoritaire par rapport à *B. garinii*.

Cette étude permettra d'obtenir des premiers résultats concernant la phylogénie de *B. garinii* portée par les oiseaux en France, ce qui sera une première étape vers une meilleure compréhension de l'influence

de la répartition spatiale et du rôle des oiseaux dans la diversification génétique de cette bactérie. Ces résultats pourront servir de base à d'autres études et aider à une meilleure compréhension du risque associé à cet agent pathogène à l'avenir.

## Partie 1 : Synthèse bibliographique

### 1. La maladie de Lyme, une maladie vectorielle complexe :

#### 1.1. Importance en santé publique

La maladie de Lyme est la maladie vectorielle la plus fréquente dans l'hémisphère Nord (Rochlin et Toledo 2020). Elle est due à un ensemble de bactéries transmises *via* une morsure de tique, et elle peut toucher l'homme et certains animaux (chien, cheval...). L'ampleur de l'augmentation de son incidence est actuellement controversée, puisqu'elle pourrait être en partie due à une meilleure détection de la maladie et à un meilleur taux de déclaration (Sykes et Makiello 2017). Cependant, les tiques transmettant la maladie de Lyme sont en expansion à travers l'Europe (Dantas-Torres 2015). La diffusion et l'augmentation du nombre de tiques vectrices pourraient représenter un risque accru pour l'homme, bien que ce risque dépende de nombreux facteurs, et ne soit pas seulement fonction de la densité vectorielle (Li et al. 2019).

#### 1.2. Agents pathogènes responsables de la maladie de Lyme

Le genre *Borrelia*, qui comprend un ensemble de bactéries spirochètes, est divisé en deux groupes : le groupe des bactéries responsables des fièvres récurrentes et le groupe responsable de la maladie de Lyme. Les bactéries responsables de la maladie de Lyme appartiennent à un complexe appelé « *Borrelia burgdorferi sensu lato* » (*Bbsl*). Une mise à jour de la nomenclature a eu lieu récemment et préconise de garder le genre *Borrelia* seulement pour les bactéries responsables de la fièvre récurrente, et d'inclure les *Bbsl* dans un genre appelé *Borelliella* (Madison-Antenucci et al. 2020).

Le groupe des *Bbsl* comprend une vingtaine d'espèces, toutes ne causant pas de maladie chez l'homme. En Europe, les principales espèces infectant l'humain sont *B. burgdorferi sensu stricto* (*ss*), *B. afzelii* et *B. garinii* (Estrada-Peña et al. 2018; Madison-Antenucci et al. 2020). La maladie de Lyme peut prendre plusieurs formes, avec des fréquences différentes en fonction de l'espèce de bactérie concernée. *B. burgdorferi ss* est ainsi plus fréquemment responsable d'arthrites, *B. afzelii* de manifestations cutanées (notamment l'acrodermatite chronique atrophique), tandis que *B. garinii* et *B. bavariensis* sont plus fréquemment retrouvées en cas de neuroborréliose (Coipan et al. 2016; Jahfari et al. 2017). Les neuroborrélioses sont des formes particulièrement invalidantes de la maladie, car elles peuvent toucher les systèmes nerveux central et périphérique d'un individu et provoquer de nombreux symptômes. En cas de neuroborréliose, les symptômes les plus fréquents sont : la radiculonévrite (inflammation des racines des nerfs, pouvant être très douloureuse), la paralysie faciale et les maux de tête, ces symptômes pouvant être associés (Garcia-Monco et Benach 2019; Nordberg et al. 2020).

#### 1.3. *Ixodes ricinus*, principale tique vectrice des *Bbsl* en Europe

Dans l'hémisphère Nord, les tiques sont les principaux vecteurs de maladies (Jongejan et Uilenberg 2004). Les tiques sont des acariens ectoparasites hématophages stricts de vertébrés regroupant trois familles, avec des cycles parasitaires variés. Ces cycles sont caractérisés par trois stases (larve-nymphe-adulte), séparées chacune par une métamorphose, qui fait suite à un (ou plusieurs) repas sanguin sur un hôte (Mccoy et Boulanger 2015).

Le genre *Ixodes* (famille des Ixodidés ou tiques dures) comprend environ 250 espèces (Jongejan et Uilenberg 2004), parmi lesquelles *Ixodes ricinus*. C'est l'espèce de tique la plus fréquente en Europe (Jeremy Gray, Kahl, et Zintl 2021), et le vecteur principal des *Bbsl* responsables de la maladie de Lyme.

Cette tique est dite triphasique, c'est-à-dire qu'elle va réaliser des repas sanguins sur trois hôtes différents, avec entre les hôtes une phase de « vie libre » au sol où se réalise la métamorphose (Figure 1). Pour trouver un hôte, elle se met à l'affût sur la végétation et s'accroche aux animaux passant près d'elle. Les tiques ont des préférences trophiques (attirait et appétence pour un certain type d'hôte) plus ou moins marquées en fonction des espèces.

*I. ricinus* fait partie des tiques les plus « généralistes », c'est-à-dire qu'elle se nourrit sur une grande variété d'hôtes (cycle hétéroxène), homme inclus. Elle a en effet été retrouvée sur plus de 300 hôtes différents (J. Gray et al. 2002). Ainsi, *I. ricinus* peut se nourrir sur les petits et moyens mammifères (notamment les rongeurs), sur les grands mammifères comme les ongulés (et préférentiellement sur les cervidés), sur les oiseaux et moins fréquemment sur les reptiles (Hofmeester et al. 2016). Sa capacité à se nourrir sur un hôte en particulier dépend de plusieurs facteurs, notamment de sa résistance à la dessiccation. Celle-ci est plus élevée chez les adultes que chez les larves, ce qui leur permet de se mettre à l'affût en hauteur et donc d'atteindre des hôtes plus grands. Ainsi, les stases larve et nymphe sont plus fréquemment retrouvées chez les petits mammifères et les oiseaux, alors que les adultes sont plus fréquemment retrouvées chez les grands mammifères (Mccoy et Boulanger 2015).

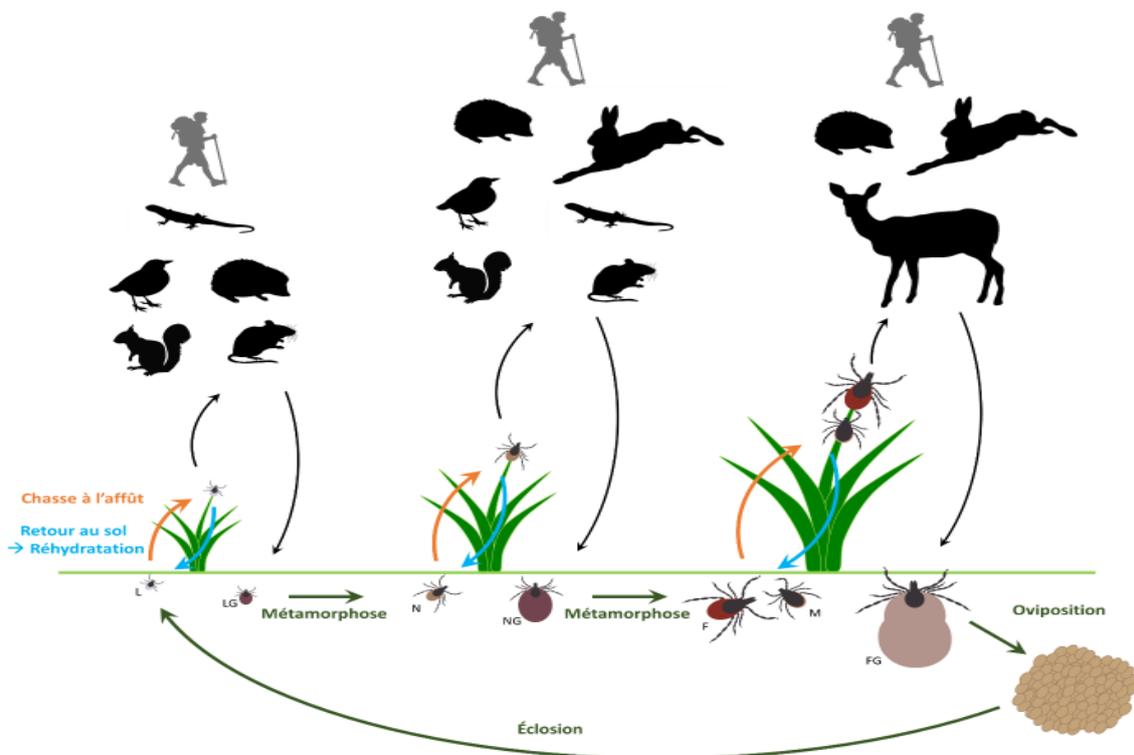


Figure 1 : Cycle de vie de la tique *Ixodes ricinus* (Lejal 2020)

Légende : L = Larve, N = Nymphe, F = Femelle, M = Mâle | LG, NG, FG = Larve, Nymphe, Femelle Gorgée

*I. ricinus* peut être porteuse de nombreux agents pathogènes et les transmettre aux hôtes sur lesquels elle se nourrit. Une tique peut acquérir un agent pathogène selon plusieurs mécanismes : soit en réalisant un repas sanguin sur un hôte infecté (infection systémique), soit de manière non systémique. L'infection peut alors se faire par co-repas, si la tique partage un même repas sanguin contaminé localement par une tique voisine infectée, ou par transmission transovarienne de la tique infectée à ses descendants (Mccoy et Boulanger 2015). Cependant, la transmission transovarienne des *Bbs1* est négligeable chez *I. ricinus* (Richter et al. 2012). L'importance de la transmission par co-repas est plus discutée (Voordouw 2015), mais elle reste marginale par rapport à la transmission systémique (Richter, Allgöwer, et Matuschka 2002).

*Ixodes ricinus* est une tique dite exophile, dont les phases de vie libre se déroulent dans des espaces ouverts ou semi-ouverts (forêts, prairies), et où l'humidité est suffisante (idéalement > 80%). Cela

conditionne aussi les types d'hôtes qu'elle est amenée à rencontrer le plus fréquemment : rongeurs, oiseaux et cerfs. L'activité des tiques est variable en fonction de l'espèce et de la stase. Chez *I. ricinus*, les nymphes et les adultes sont les plus actives au printemps, et ont une activité modérée en été et en automne, alors que le pic d'activité des larves est un peu plus tardif, entre le printemps et l'été, avec un second pic d'activité plus faible en automne (Figure 2).

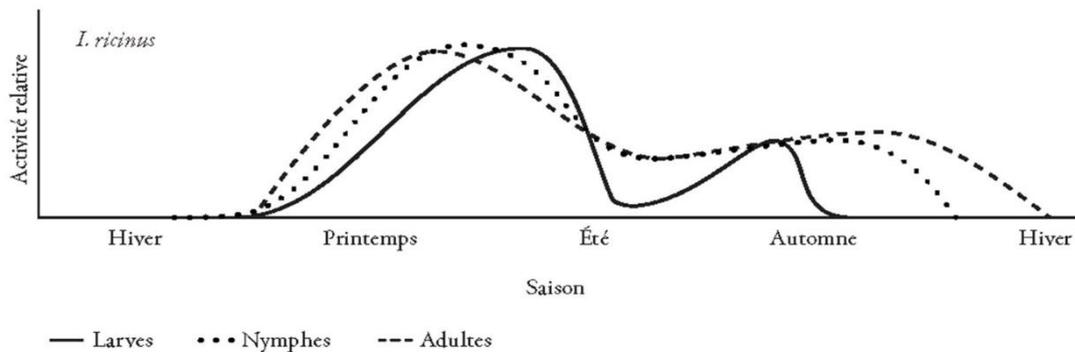


Figure 2 : Activité saisonnière d'*Ixodes ricinus* en Europe centrale, au nord de l'Angleterre et en Écosse (Kurtenbach et al. 2006)

## 1.4. Les oiseaux, principaux réservoirs de *Borrelia garinii*

### 1.4.1. Description des oiseaux étudiés dans ce mémoire

Les passereaux (ordre des Passériformes) sont les oiseaux les plus représentés en France métropolitaine, avec environ 250 espèces sur les 625 présentes sur le sol français (avibase.bsc-eoc.org). Les oiseaux d'intérêt pour ce mémoire sont principalement issus de cet ordre et vivent majoritairement dans la strate arbustive, préférant les milieux buissonnants ou les zones humides, qui sont des milieux favorables aux tiques. Bien qu'une partie de ces espèces migrent sur de longues distances, environ 44% de celles-ci sont sédentaires (ou migratrices sur de courtes distances). La période de reproduction de ces oiseaux s'étend d'avril à juillet, avec des individus qui retournent ou qui restent sur leur lieu de naissance pour s'accoupler et pondre leurs œufs (Rataud, 2022).

### 1.4.2. Rôle des différentes espèces d'oiseaux dans le portage de tiques

En fonction notamment de leur abondance respective et de leurs habitudes de vie en période de reproduction (hauteur du nid, hauteur d'alimentation, etc), certaines espèces d'oiseaux vont être plus souvent en contact avec des tiques généralistes comme *I. ricinus*, et vont avoir un taux d'infestation plus élevé. Ainsi, chez les oiseaux, on estime que 20% des individus sont porteurs de 80% des tiques (Woolhouse et al. 1997; Dieter 2016). Au sein des passériformes, la famille des *Turdidae* comprend des espèces telles que le Merle noir (*Turdus merula*) et la Grive musicienne (*Turdus philomelos*), qui figurent souvent parmi les espèces d'oiseaux les plus infestées (Klaus et al. 2016; Klitgaard et al. 2019). Leur taux d'infestation est variable selon les études mais peut dépasser 70% (Michalik et al. 2008; Dubska et al. 2009; Gryczyńska et Welc-Falęciak 2016).

### 1.4.3. Importance des espèces d'oiseaux réservoirs pour les *Bbs!*

Un réservoir peut être défini comme une population (ou un ensemble de populations épidémiologiquement connectées) dans lequel un agent pathogène est maintenu de manière persistante et depuis lequel l'infection est transmise à une population cible, différente ou non du réservoir (Haydon et al. 2002). Ainsi, une population (ici une espèce) réservoir pour les *Bbs!* est une espèce capable d'être infectée par ces bactéries *via* une morsure de tique, d'assurer leur répllication et de les transmettre ensuite à une tique naïve.

Les réservoirs principaux des *Bbsl* sont les petits mammifères (notamment les rongeurs), et les oiseaux (Stanek et al. 2012). Bien que les tiques adultes du genre *Ixodes* aient une préférence plus marquée pour les grands mammifères tels que les cervidés, ceux-ci ne sont pas des réservoirs pour les *Bbsl*, car ils ne transmettent pas ces bactéries à des tiques naïves (Telford et al. 1988). Les espèces réservoirs préférentielles des *Bbsl* varient en fonction des espèces de bactéries concernées. Ainsi, pour *B. garinii*, les hôtes réservoirs sont principalement les oiseaux (Wolcott et al. 2021).

Il a été démontré, via des expériences de laboratoire (xénodiagnostics), que les *Turdidae* jouaient aussi un rôle important dans l'acquisition des *Bbsl* par les tiques. Le Merle noir et la Mésange charbonnière (*Parus major*) sont des réservoirs pour *B. garinii* (Humair et al. 1998; Dieter et al. 2013). Enfin, certaines espèces (notamment certaines espèces de *Turdidae*) peuvent présenter des pourcentages relativement élevés de larves positives à *Bbsl*, supérieurs à 50% (Michalik et al. 2008; Dubska et al. 2009). Une étude a montré que sur 40 espèces d'oiseaux capturés en Europe Centrale, le Merle noir et la Grive musicienne portaient 95% des larves positives à *Bbsl*, dont plus de 90% étaient porteuses de *B. garinii* ou *B. valaisiana* (Taragel'ová et al. 2008). Ces espèces d'oiseaux sont des espèces importantes en santé publique puisque les larves qu'elles portent deviendront des nymphes infectées, qui seront ensuite susceptibles de transmettre la maladie à l'homme.

Les oiseaux peuvent être infestés par la tique *I. ricinus*, mais aussi par des tiques ornithophiles, dont *Ixodes arboricola* et *Ixodes frontalis*. Bien que ces trois espèces diffèrent de par leur milieu de vie et leur pic d'activité, elles partagent certains hôtes aviaires sur lesquels elles se nourrissent (Rataud 2022). Il a été montré que ces tiques ornithophiles pouvaient être porteuses de *Bbsl* (Heylen et al. 2013). Bien que des résultats contradictoires aient été obtenus quant à leur compétence vectorielle (Heylen et al. 2014), qui pourrait varier en fonction de l'espèce d'oiseau et de bactérie concernées, ces tiques pourraient participer au maintien des *Bbsl* au sein de la communauté aviaire, et *Ixodes ricinus* pourrait servir d'intermédiaire en transmettant les *Bbsl* des oiseaux aux mammifères (Heylen et al. 2017).

## 2. Phylogénie des *Bbsl* :

### 2.1. Description du génome des *Bbsl*

Le génome des *Bbsl* est un des plus petits génomes bactériens connus à ce jour (Seifert et al. 2015). Paradoxalement, celui-ci figure parmi les génomes comptant le plus de plasmides (fragment d'ADN extrachromosomique). Les séquençages complets du génome ont permis de mettre en évidence un chromosome linéaire, d'environ 900 kpb, ainsi que de nombreux plasmides (entre 7 et 21), linéaires et circulaires, de tailles diverses (Schwartz et al. 2021). La plupart des gènes de ménage, qui sont des gènes assurant les fonctions essentielles à la bactérie, se trouvent sur le chromosome. Celui-ci est globalement conservé d'une *Bbsl* à l'autre (mêmes gènes et organisation). Au contraire, les plasmides, qui portent la plupart des gènes codant les protéines de surface permettant l'interaction entre la bactérie et ses hôtes, sont beaucoup plus variables (Schwartz et al. 2021).

Certains gènes ont été particulièrement étudiés et utilisés dans le cadre des analyses phylogénétiques, c'est le cas notamment d'*ospC* (outer surface protein C), qui code pour une lipoprotéine de surface exprimée lors de la phase précoce d'infection de l'hôte vertébré. *OspC* est situé sur un plasmide, c'est un gène très variable chez les différentes souches de *Bbsl* (Schwartz et al. 2021). Un autre gène, aujourd'hui moins utilisé mais qui nous intéressera par la suite, est celui codant pour une des protéines constitutives des flagelles périsplasmiques des *Bbsl* : *flaB* (Zhang et al. 2019). Ce gène est situé sur le chromosome linéaire des *Bbsl*, et à l'inverse d'*ospC*, il est peu variable (Chao et Shih 2009).

### 2.2. Méthodes de séquençage de *Bbsl*

Les progrès techniques dans le domaine de la génétique ont permis d'étudier de plus en plus précisément le génome des bactéries (possibilité d'amplification de tout ou partie du génome, possibilité de séquencer rapidement les produits d'amplification...). Les techniques utilisées sont nombreuses et ont beaucoup évolué ces dernières années (Margos et al. 2011; Schwartz et al. 2021). Le gène de la flagelline, autrefois considéré comme un marqueur phylogénétique capable de discriminer les différentes espèces de *Borrelia*, a été utilisé lors de plusieurs études phylogénétiques (Fukunaga et al.

1996; Chao et Shih 2009; Aslam et al. 2012). D'autres gènes ou séquences ont aussi souvent été ciblés, comme *ospC* ou les espaceurs intergéniques de certaines régions (IGS) (Schwartz et al. 2021). Cependant, le développement de techniques plus précises ayant permis d'identifier de nouvelles espèces de *BbsI* a montré les limites des méthodes se concentrant sur un seul ou sur un petit nombre de gènes/séquences (Schwartz et al. 2021). Ainsi, la plupart des études phylogénétiques récentes ciblant *BbsI* utilisent le MLST (MultiLocus Sequence Typing). Le MLST consiste à cibler un ensemble de gènes et à les séquencer entièrement ou en partie, un locus étant l'emplacement d'un gène sur un chromosome. Ces gènes sont ensuite comparés à ceux de souches présentes dans une base de données (par exemple pubMLST), et un profil allélique est déterminé en fonction des divergences constatées. Le MLST cible souvent chez les *BbsI* 8 gènes de ménage, qui peuvent différer selon les études, et qui peuvent être associés au séquençage de gènes ou d'IGS supplémentaires (Margos et al. 2011; Takano et al. 2011; Rudenko et al. 2014). Le séquençage complet du génome est une méthode permettant d'avoir des arbres phylogénétiques robustes, mais les études phylogénétiques ciblant les *BbsI* et utilisant ces méthodes sont encore limitées (Becker et al. 2016; Tyler et al. 2018).

### 2.3. Phylogénie et épidémiologie moléculaire : les principes

#### 2.3.1. Définition et principe de la phylogénie

La phylogénie est l'étude des liens existant entre différents taxa, elle permet leur classement et la déduction de leur histoire évolutive. Un taxon est un groupe défini par un ensemble de caractéristiques que partagent tous les éléments de ce groupe (par exemple, différentes espèces peuvent constituer différents taxa). Parfois, le groupe peut ne comporter qu'un seul élément (par exemple lors d'études entre différentes bactéries, où chaque bactérie est un taxon). Les arbres phylogénétiques sont des représentations graphiques permettant de visualiser les liens entre les différents taxa.

La phylogénie moléculaire est la discipline qui étudie les relations entre les taxa en se basant sur des molécules telles que l'ADN/ARN ou les protéines. C'est une discipline qui a connu un grand essor ces dernières années, grâce aux progrès des techniques qui ont permis d'obtenir plus facilement et à moindre coût des séquences génétiques. Elle est aujourd'hui régulièrement utilisée pour étudier les relations entre différentes bactéries. Cependant, en fonction de la technique d'obtention des données moléculaires choisie, une connaissance relative du génome d'intérêt est nécessaire, afin de savoir quel gène (ou séquence d'ADN/ARN) utiliser pour la comparaison (Brown 2002; Schwartz et al. 2021). Un arbre phylogénétique est constitué de feuilles et de nœuds internes, reliés entre eux par des branches (Figure 3). Les feuilles sont les nœuds terminaux de l'arbre et représentent les différents taxa étudiés. Les nœuds internes correspondent à des ancêtres hypothétiques communs à plusieurs feuilles. La topologie de l'arbre correspond à la forme générale de l'arbre. La longueur des branches peut être sans signification, représenter l'importance des changements survenus entre deux nœuds, ou la quantité de temps écoulée (Baum 2008; Drummond et Bouckaert 2015).

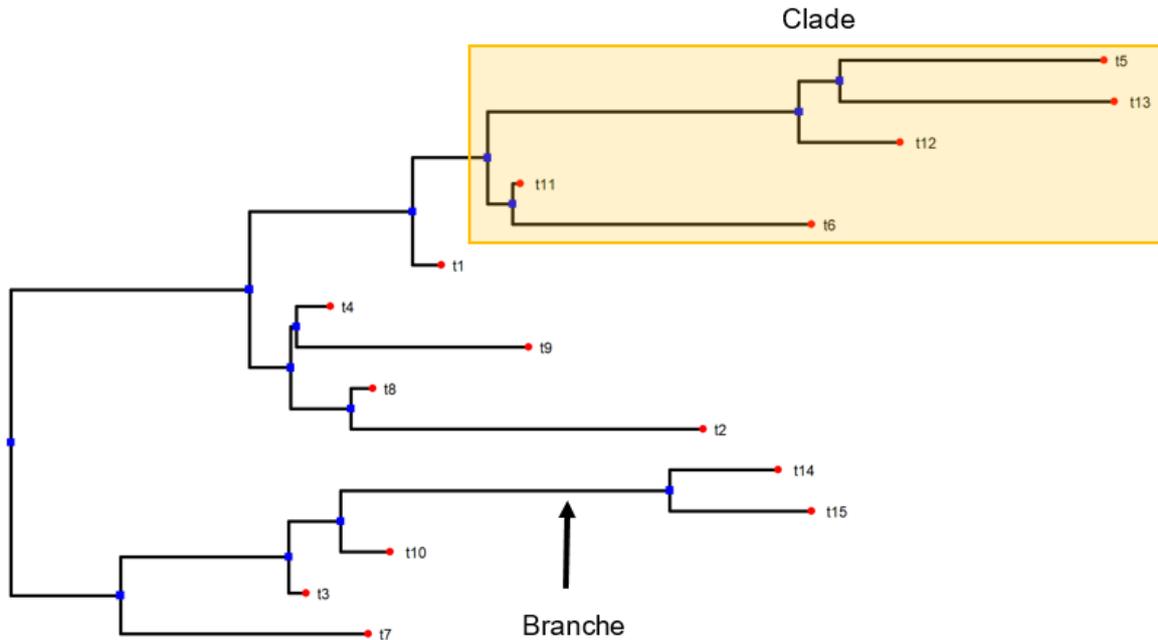


Figure 3 : Exemple d'arbre phylogénétique

*t1 à t15 sont les noms des différents taxa étudiés. Les ronds rouges sont les feuilles de l'arbre (taxa), et les carrés bleus les nœuds internes (ancêtres communs).*

Un clade est défini comme un groupe constitué par un ancêtre commun et l'ensemble de ses descendants. Les éléments présents dans un clade sont plus proches entre eux que de n'importe quels autres éléments figurant dans l'arbre.

### 2.3.2. Construction d'un arbre phylogénétique

#### *Distance génétique et modèle de substitution*

Un arbre phylogénétique est construit en comparant les séquences alignées (taxa), afin de déterminer quelles sont les séquences les plus proches pour pouvoir les rassembler et former des clades. On suppose que la distance génétique entre deux séquences représente la proximité des séquences d'un point de vue évolutif : deux séquences avec une distance faible seront supposées avoir divergé plus récemment de leur ancêtre commun que des séquences avec une distance plus grande. La méthode la plus simple pour mesurer la distance génétique entre deux séquences est de compter le nombre de sites (emplacement d'un nucléotide) pour lesquels ces séquences diffèrent (distance de Hamming) (Hamming 1950). Une proportion de sites différents entre les deux séquences est ensuite calculée grâce à ce nombre. Cependant, la comparaison à un instant donné ne permet pas de déterminer précisément la distance entre deux séquences. En effet, certains phénomènes peuvent entraîner une surestimation ou une sous-estimation de la distance génétique. Par exemple, si un nucléotide T d'une séquence mute en A puis à nouveau en T, et que cette mutation n'est pas prise en compte, la distance génétique de cette séquence avec une autre sera sous-estimée. De plus, les taux de substitutions entre les différents nucléotides peuvent varier (par exemple, la substitution A → G peut être plus fréquente que A → C). La non prise en compte de cette variation peut aussi fausser la mesure de la distance génétique. Ainsi, pour essayer de prendre en compte ces phénomènes, des modèles de substitutions de nucléotides peuvent être appliqués, afin de corriger la mesure de la distance génétique. Un modèle de substitution permet de décrire la fréquence des bases ainsi que les taux de substitutions dans les séquences étudiées. Un des modèles de substitution les plus simples est le modèle de Jukes and Cantor (1969), pour lequel tous les nucléotides ont la même fréquence, avec des taux de substitutions égaux entre les nucléotides. D'autres modèles existent, avec des fréquences et/ou des taux de substitutions variables entre les nucléotides.

### *Méthodes utilisées pour la construction d'arbres phylogénétiques*

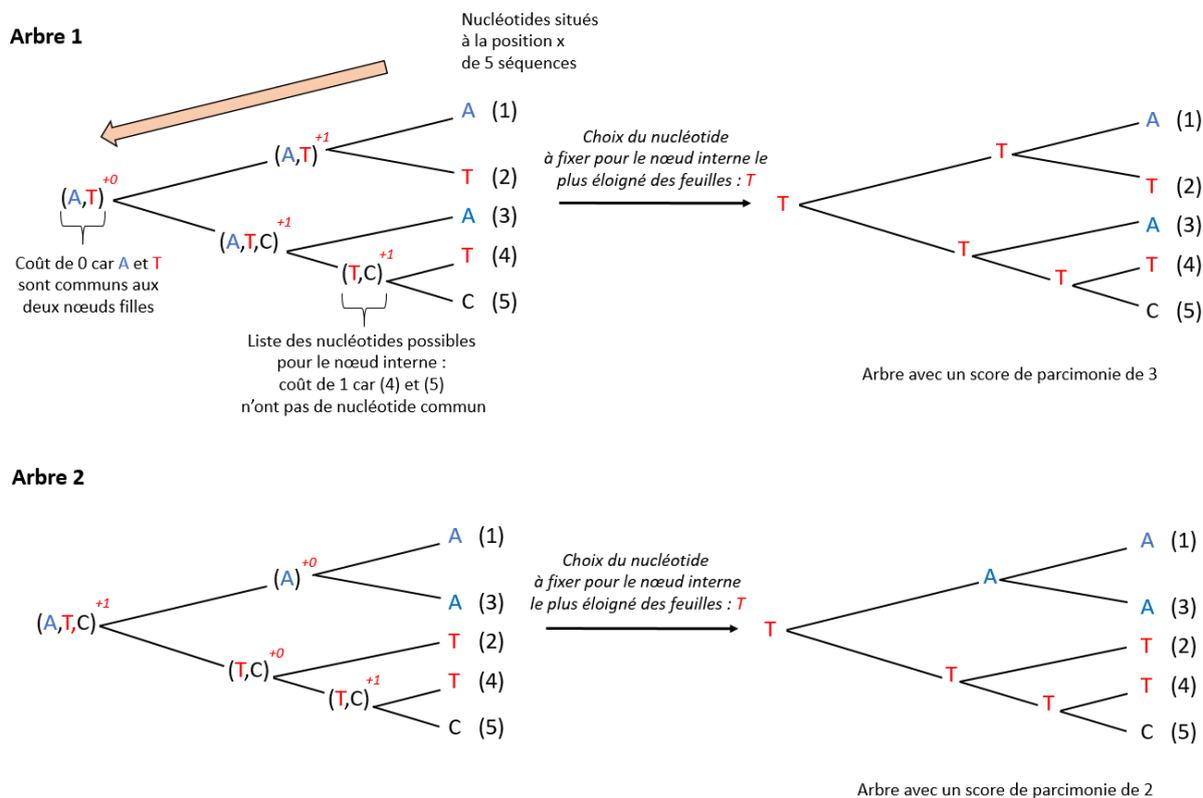
Deux types de méthodes sont utilisées pour la construction d'arbres phylogénétiques : les méthodes basées sur la distance ou les méthodes basées sur les caractères.

Les méthodes basées sur la distance reposent sur le calcul de la distance génétique entre les paires de séquences. Cette distance peut être calculée en utilisant différents modèles de substitutions. Une matrice des distances est ensuite créée, qui regroupe la distance génétique entre chaque paire de séquence. Puis un arbre phylogénétique est construit grâce à un algorithme qui va utiliser la matrice des distances pour identifier les séquences les plus proches et les regrouper (Yang et Rannala 2012). La méthode des distances la plus utilisée aujourd'hui est le Neighbor Joining. C'est un algorithme qui débute avec arbre en étoile, c'est-à-dire où chaque séquence est isolée et à égale distance des autres, et qui regroupe ensuite les séquences (ou ensemble de séquences préalablement regroupées) deux à deux. Le choix des séquences à regrouper se fait manière à minimiser la longueur des branches de l'arbre après agglomération.

Les méthodes basées sur la distance ont l'avantage d'être rapides et de ne pas nécessiter beaucoup de puissance de calcul. Cependant, elles ne sont pas adaptées dans le cas de séquences fortement divergentes, et sont sensibles aux lacunes (trous) dans l'alignement (Yang et Rannala 2012).

Les méthodes basées sur les caractères comparent en même temps l'ensemble des séquences alignées, un site à la fois, et explorent les arbres pouvant mener aux différences constatées. Un score est ainsi attribué à chaque arbre généré, et le meilleur arbre est choisi selon des critères qui varient en fonction de la méthode utilisée. Les principales méthodes utilisées sont : l'inférence bayésienne, le maximum de vraisemblance et le maximum de parcimonie. Ainsi, pour l'inférence bayésienne, le score correspond au degré de confiance *a posteriori* que l'on peut avoir dans l'arbre obtenu. Le score du maximum de vraisemblance correspond à la vraisemblance de l'arbre (probabilité), alors que le score du maximum de parcimonie représente le nombre minimum de changements nécessaires à chaque site pour obtenir l'arbre (Yang et Rannala 2012).

La méthode du maximum de parcimonie est une méthode rapide et simple à utiliser, par rapport au maximum de vraisemblance et à l'inférence bayésienne (Kapli, Yang, et Telford 2020). Cette méthode, contrairement aux deux autres, n'utilise pas de modèle de substitution explicite. La construction de l'arbre se fait des feuilles vers les nœuds internes (Figure 4).



**Figure 4 : Construction des arbres selon le principe de parcimonie**

La flèche orange représente le sens de construction de l'arbre. (1) à (5) représentent les différents taxa étudiés. Différentes couleurs sont utilisées pour représenter les nucléotides : bleu pour A, rouge pour T, noir pour C. En rouge sont notés les coûts associés à chaque nœud interne : +0 lorsqu'il y a un ou plusieurs nucléotides communs aux nœuds filles, +1 lorsqu'il n'y a aucun nucléotide en commun. Ici deux arbres sont générés, et l'arbre 2, dont le score de parcimonie est inférieur à l'arbre 1, sera gardé. Un seul site (emplacement de nucléotide) est ici représenté, le score de parcimonie de l'arbre correspond à la somme des scores de parcimonie calculés pour chaque site présent dans les séquences étudiées.

Une topologie d'arbre est générée aléatoirement, et les nœuds internes se voient attribuer une liste des nucléotides possibles pour chaque site issu des séquences filles. Une fois les listes attribuées à tous les nœuds, un arbre est produit en fixant un des nucléotides du nœud le plus éloigné de l'ensemble des feuilles, et en sélectionnant les nucléotides des autres nœuds afin de limiter les substitutions. Le nombre de substitutions de l'arbre constitue le score de parcimonie, ce nombre étant la résultante de la somme des scores de parcimonie obtenus pour chaque site des séquences étudiées (Goëffon, Richer, et Hao 2010). Plusieurs algorithmes existent pour attribuer un score à l'arbre obtenu, l'algorithme de Fitch étant l'un des plus utilisés (même fréquence de substitution pour toutes les bases) (Xia 2020a). De nombreux arbres sont ainsi générés, et leur score de parcimonie est calculé. A la fin, l'arbre avec le plus petit score est retenu (arbre le plus parcimonieux). Explorer l'ensemble des arbres possibles est très chronophage, surtout lorsque beaucoup de séquences sont incluses dans l'analyse, ou que leur longueur augmente. De ce fait, souvent les méthodes d'exploration des arbres se basent sur des algorithmes qui recherchent le meilleur arbre possible sans toutefois tous les explorer, par exemple en partant du premier arbre généré, puis en calculant un nouveau score de parcimonie après inversion de deux branches. L'obtention de l'arbre avec le meilleur score n'est donc pas toujours garantie (Goëffon, Richer, et Hao 2010).

### 2.3.3. Principes de l'épidémiologie moléculaire et lien avec la phylogénie

L'épidémiologie moléculaire est la discipline qui utilise des données moléculaires (ADN, ARN...) pour étudier la distribution et les déterminants des maladies chez les populations humaines et/ou animales (Riley et Blanton 2018). L'épidémiologie moléculaire peut s'appuyer sur la phylogénie, notamment lorsque la comparaison des souches est mise en relation avec des données épidémiologiques (distribution spatiale du pathogène, caractéristiques des hôtes, etc). En épidémiologie moléculaire, des arbres phylogénétiques sont par exemple construits pour comparer des souches virales ou bactériennes nouvellement isolées à un endroit donné à d'autres souches précédemment identifiées issues d'autres endroits. Ces arbres permettent notamment de mettre en évidence un lien potentiel entre proximité génétique et distribution spatiale. La comparaison des souches en fonction de l'espèce hôte, et l'élaboration d'hypothèses sur la chaîne de transmission inter-espèce, sont aussi des informations épidémiologiquement pertinentes pour les pathogènes touchant plusieurs espèces.

## 2.4. Etat des lieux de la taxonomie de *Bbsl* et apports de la phylogénie

### 2.4.1. Histoire de la taxonomie de *Bbsl*

La classification des espèces de *Bbsl* est en constant remaniement. *B. garinii* a été l'une des premières espèces identifiées au sein des *Bbsl*, par la méthode d'hybridation ADN-ADN (Baranton et al. 1992). Différents sérotypes (ensemble des caractéristiques antigéniques des bactéries, permettant de différencier des souches appartenant à une même espèce) de *B. garinii* pour la protéine ospA ont ensuite été découverts (Wilske et al. 1993). Or, contrairement aux autres, *B. garinii* ospA sérotype 4 avait pour réservoir des rongeurs (Huegli et al. 2002). Ultérieurement, des études phylogénétiques ont permis de mettre en évidence deux clades, avec d'un côté *B. garinii*, et de l'autre une autre espèce (*B. garinii* ospA sérotype 4), proche génétiquement de *B. garinii*, mais ayant comme réservoir préférentiel les rongeurs (Margos et al. 2009; 2013). Cette espèce a été nommée *B. bavariensis*. Depuis, plusieurs autres espèces ont été identifiées au sein des *Bbsl*.

### 2.4.2. Limites des protocoles actuels utilisés pour la taxonomie et la phylogénie

La phylogénie et les techniques d'étude du génome permettent donc de définir et de caractériser les espèces. Cependant, la notion d'espèce chez les bactéries est différente de celle proposée chez les eucaryotes. Du fait de leur grande diversité génétique, il n'est pas toujours évident de décider de l'appartenance d'une bactérie à une certaine espèce. Longtemps, la méthode principale utilisée pour différencier deux espèces était la technique d'hybridation ADN-ADN, qui s'appuie sur la mise en évidence d'une homologie inférieure à 70%. Cependant, cette méthode, bien que toujours recommandée, est longue et compliquée à mettre en œuvre, en plus de poser des problèmes d'interprétation et de reproductibilité (Postic, Garnier, et Baranton 2007; Margos et al. 2011).

Par la suite, l'identification d'espèces par MLST/MLSA (MultiLocus Sequence Analysis) a été proposée et est de plus en plus utilisée pour définir la taxonomie des bactéries (Stackebrandt et al. 2002; Glaeser et Kämpfer 2015). En 2007, une étude a ainsi identifié des souches de *Bbsl* ne pouvant être classées avec certitude par la méthode d'hybridation ADN-ADN, et mis en évidence une nouvelle espèce, *B. californiensis* (Postic, Garnier, et Baranton 2007). Pour la taxonomie de *Bbsl*, différents *loci* ont pu être utilisés, avec des variations dans les arbres phylogénétiques obtenus en fonction des gènes ciblés. Certains gènes hautement variables et portés par des plasmides, comme *ospC*, peuvent donner des arbres différents de ceux obtenus en ciblant les gènes de ménage, probablement à cause de phénomènes de recombinaison plus fréquents (Margos et al. 2009). Le choix des gènes séquencés est donc important pour avoir des arbres phylogénétiques représentatifs de la proximité et de l'histoire évolutive des bactéries analysées. Le protocole défini par Margos et al en 2008, ciblant 8 gènes de ménage, est régulièrement utilisé (Sabitova et al. 2018; Mtierová et al. 2020; Norte et al. 2020). Selon les connaissances actuelles, ces gènes sont intéressants d'un point de vue phylogénétique, car ils reflètent l'histoire évolutive des bactéries (peu de transferts horizontaux, évolution lente...) (Margos et al. 2008). Les gènes choisis déterminent aussi le seuil de différenciation de deux espèces, certains protocoles pouvant donc discriminer des bactéries plus proches génétiquement (Margos et al. 2011).

Ainsi, selon le protocole choisi, des résultats différents peuvent être obtenus, et la comparaison entre les différentes études n'est pas toujours aisée. L'émergence de nouveaux protocoles associés aux

techniques de séquençage innovantes pourraient accélérer la refonte de la taxonomie des *Bbsl* et faire apparaître de nouvelles espèces.

## 2.5. La phylogénie, un outil au service d'une meilleure compréhension des *Bbsl*

La diversité génétique des *Bbsl* est de plus en plus étudiée, et plusieurs études tentent de caractériser et d'expliquer cette diversité, en lien avec l'éco-épidémiologie de ces bactéries.

Des analyses phylogénétiques sont souvent réalisées afin de comparer des souches isolées dans une aire géographique donnée, entre elles et/ou avec d'autres souches connues. Il apparaît globalement que les *Bbsl* génétiquement proches ou identiques ne sont pas forcément situées dans une même aire géographique, même s'il peut exister des différences en fonction de l'espèce ou de l'unité géographique étudiées (pays, continent...). Une étude sur des *Bbsl* issues de plusieurs pays d'Europe a montré l'existence d'une structure spatiale chez *B. afzelii* (dont les réservoirs principaux sont les rongeurs), avec peu de profils alléliques identiques entre les pays, structure qui n'existe pas (ou faiblement) chez *B. garinii* (Vollmer et al. 2011).

Cette faible tendance à la ségrégation spatiale des *Bbsl* génétiquement proches peut s'expliquer notamment par le mélange et le déplacement importants de celles-ci, en lien avec la mobilité des vecteurs et des hôtes qu'elles colonisent. Il apparaît aussi qu'il y aurait un lien entre la mobilité d'une espèce dans l'espace, sa polyvalence et l'intensité de sa diversité génétique : les espèces plus polyvalentes (retrouvées chez de nombreuses espèces hôtes et dans différents milieux) doivent s'adapter à la diversité des milieux et des hôtes rencontrés (Oppler et al. 2021). Les espèces moins polyvalentes tendent à former des sous-populations au sein desquelles la diversité génétique est faible (du fait d'un milieu homogène), mais entre lesquelles la diversité génétique est importante (différences de milieux).

Ainsi, *B. garinii*, une bactérie présente chez de nombreux oiseaux, qui sont des hôtes très mobiles, est l'une des *Bbsl* avec le plus de variabilité génétique (Vollmer et al. 2011; Jacquot et al. 2014). Les études s'étant intéressées à la répartition géographique de *B. garinii*, en lien avec sa diversité génétique, ont montré que cette espèce avait peu tendance à s'agréger spatialement (Mtierová et al. 2020; Norte et al. 2020). Cependant, ces résultats sont à nuancer, d'une part parce que la plupart des études se sont concentrées sur des bactéries issues de tiques portées par des oiseaux migrateurs (ou à l'affût), d'autre part car il semblerait que l'existence d'une certaine structure spatiale soit possible selon l'échelle considérée, notamment lors de la comparaison de bactéries provenant de différents continents (Vollmer et al. 2013; Mtierová et al. 2020; Norte et al. 2020). De plus, *B. garinii* est aussi impliquée dans un cycle marin : elle est isolée chez *Ixodes uriae*, qui se nourrit sur de nombreuses espèces aviaires marines. Les études s'étant intéressées à la comparaison de la diversité génétique entre les bactéries issues des cycles terrestres et marins ont mis en évidence un possible mélange des souches, associé à une diversité génétique importante (Comstedt et al. 2009; Gómez-Díaz et al. 2011). La comparaison des bactéries issues de tiques portées par des oiseaux marins ne met pas en évidence de structures spatiales au sein des régions, mais met en évidence deux sous-populations distinctes entre les bassins de l'océan Pacifique et de l'océan Atlantique (Gómez-Díaz et al. 2011).

L'étude de la diversité génétique des *Bbsl* est aussi importante, car elle peut influencer sur la pathogénicité et sur la diversité des symptômes chez l'homme, qui peuvent varier même au sein d'une espèce (Hanincova et al. 2013; Cerar et al. 2016). Étudier la diversité génétique de *B. garinii* est donc nécessaire, car c'est l'une des *Bbsl* pour laquelle on retrouve le plus de diversité, et qu'elle est responsable de graves symptômes chez l'homme. Cependant, peu d'études portant sur ce sujet ont été publiées à ce jour, et aucune ne s'est focalisée sur la France. Une étude réalisée en Slovaquie a donné des éléments en faveur de l'existence de différences génétiques entre les bactéries présentes chez les tiques à l'affût et celles présentes chez les oiseaux. Elle soulève aussi l'existence d'une possible variabilité génétique de *B. garinii* en fonction des saisons (Mtierová et al. 2020). Cependant, ces résultats préliminaires nécessiteraient d'être confirmés par d'autres études. Comme mentionné précédemment, les quelques études s'étant intéressées à la diversité génétique de *B. garinii* en lien avec sa répartition géographique révèlent une structuration spatiale faible voire absente (hors études entre différents continents).

## Partie 2 : Travail personnel

Pour résumer les caractéristiques de la maladie de Lyme, il s'agit d'une maladie vectorielle causée par des bactéries appartenant à un complexe d'espèces appelé *Bbsl*. L'incidence de cette maladie a augmenté en France et dans le monde durant les dernières décennies. *B. garinii* est la *Bbsl* majoritairement responsable de neuroborréliose, affection grave pouvant nécessiter une hospitalisation. C'est l'une des *Bbsl* que l'on retrouve le plus chez l'homme en Europe. Cette bactérie a pour réservoirs principaux les oiseaux.

Ce stage s'effectue dans la continuité du projet OUTLYER, qui visait à déterminer le rôle des oiseaux en période de reproduction dans la dynamique des populations de tiques, et dans la circulation des agents pathogènes transmissibles par les tiques (notamment les *Bbsl*). Peu d'études se sont intéressées au rôle des oiseaux dans la dynamique de transmission des *Bbsl*, et la plupart de ces études se sont concentrées sur les oiseaux en période de migration (Rataud, Henry, et al. 2022). Or, outre leur rôle de potentiel disséminateurs de bactéries via les vols longue distance (Klitgaard et al. 2019; Mysterud et al. 2019), les oiseaux sont aussi à prendre en compte dans la dynamique de population locale des tiques : ils participent au nourrissage des tiques et à la circulation des pathogènes, notamment lors de leur période et sur leurs lieux de reproduction, où ils restent plusieurs mois (Marsot et al. 2012; Rataud, Galon, et al. 2022). Une densité importante d'oiseaux réservoirs pour *B. garinii* en un lieu donné participe à l'augmentation du nombre de tiques potentiellement infectées par cette bactérie. Si le nombre de tiques infectées augmente, le risque d'infection chez l'homme (risque acarologique) augmente aussi. Ce risque est d'autant plus important que l'activité maximale des tiques se superpose à la période de reproduction des oiseaux, au printemps, qui est aussi la période où les activités de plein air des hommes augmentent (Rataud, Galon, et al. 2022). Le rôle que peuvent jouer les oiseaux dans le maintien et la transmission de *B. garinii* est notamment fonction de l'espèce d'oiseau considérée et de son milieu de vie (Marsot et al. 2012).

Les séquences génétiques des *Bbsl* retrouvées chez les larves gorgées prélevées sur les oiseaux dans le cadre du projet OUTLYER permettent d'étudier la diversité génétique de ces bactéries (surtout celle de *B. garinii*, qui était la *Bbsl* la plus fréquemment retrouvée chez les oiseaux lors du projet OUTLYER) dans le cadre de ce stage. Or, l'étude de la variabilité génétique des *Bbsl* permet de mieux comprendre les déterminants et l'épidémiologie de la maladie de Lyme, notamment grâce à l'épidémiologie moléculaire et à la construction d'arbres phylogénétiques. A notre connaissance, aucune étude ne s'est concentrée sur la diversité génétique de *Borrelia garinii* portée par les oiseaux en France.

Le but de ce travail était de caractériser la diversité génétique de *B. garinii* chez les oiseaux en France. Les objectifs de ce travail étaient d'étudier la répartition géographique de *B. garinii* chez des oiseaux en France, puis d'étudier sa diversité génétique grâce à une analyse phylogénétique par maximum de parcimonie. Le lien entre la phylogénie et la distribution géographique de ces bactéries a été exploré. Les liens entre diversité génétique et espèce aviaire hôte ou milieu de vie des oiseaux ont aussi été étudiés.

Une brève description du travail préliminaire à l'analyse des séquences génétiques, réalisé par Amalia RATAUD dans le cadre d'une thèse d'Université intitulée « Eco-épidémiologie du système Oiseaux – Tiques – *Borrelia* » est fournie. Les séquences obtenues à l'issue de son travail ont été récupérées, nettoyées et mises en forme pour pouvoir réaliser les différents objectifs de l'étude.

### 1. Matériels et méthodes :

#### 1.1. Collecte des données

Les données utilisées dans ce travail correspondent à une partie des données collectées lors du projet OUTLYER.

Ces données correspondent aux caractéristiques (espèce, sexe, âge, milieu de vie, localisation...) concernant des oiseaux communs capturés en période de reproduction sur différents sites en France aux printemps 2019 et 2020, aux données issues des tiques collectées sur ces oiseaux (nombre, stase, espèce), ainsi qu'aux données issues de la recherche d'agents pathogènes dans les larves gorgées échantillonnées (genre, espèce, séquences génétiques).

Les données initiales concernant le milieu de vie des oiseaux, qui étaient nommées selon la nomenclature standard CLC (CORINE Land Cover), ont été regroupées. Certaines catégories d'habitat ont été fusionnées afin de faciliter la visualisation graphique (Tableau 1).

Tableau 1 : Catégories utilisées pour l'habitat des oiseaux capturés, et lien avec les catégories officielles CLC

Type de milieu présenté	Code et Type de milieu initial	Définition CLC
Agricole (prairie/culture)	231 - Praires	Surfaces enherbées denses de composition floristique constituée principalement de graminées, non incluses dans un assolement. Principalement pâturées, mais dont le fourrage peut être récolté mécaniquement. Y compris des zones avec haies (bocages).
	242 - Systèmes culturaux et parcellaires complexes	Mosaïque de petites parcelles de cultures annuelles diversifiées, de prairies et/ou de cultures permanentes complexes, avec éventuellement des maisons et jardins épars.
	243 - Surfaces essentiellement agricoles, interrompues par des espaces naturels importants	Surfaces essentiellement agricoles, interrompues par des zones naturelles ou semi-naturelles (y compris des zones humides, des plans d'eau ou des affleurements rocheux).
Forêt	311 - Forêts de feuillus	Formations végétales principalement constituées par des arbres, mais aussi par des buissons et arbustes en sous-étage, où dominent les espèces forestières feuillues.
	312 - Forêts de conifères	Formations végétales principalement constituées par des arbres, mais aussi par des buissons et arbustes en sous-étage, où dominent les espèces forestières de conifères.
	313 - Forêts mélangées	Formations végétales principalement constituées par des arbres, mais aussi par des buissons et arbustes en sous-étage, où ni les feuillus ni les conifères ne dominent.
Clairsemé	24 - Forêt et végétation arbustive en mutation	Végétation arbustive et herbacée avec arbres épars. Formations pouvant résulter de la dégradation de la forêt ou d'une recolonisation/régénération de la forêt.
Humide	411 - Marais intérieurs	Terres basses généralement inondées en hiver et plus ou moins saturées d'eau en toutes saisons.
	412 - Tourbières	Terrains spongieux humides dont le sol est constitué principalement de mousses et de matières végétales décomposées (principalement des sphaignes). Tourbières exploitées ou non.
Périurbain	112 - Tissu urbain discontinu	Espaces structurés par des bâtiments. Les bâtiments, la voirie et les surfaces artificiellement recouvertes coexistent avec des surfaces végétalisées et du sol nu, qui occupent de manière discontinue des surfaces non négligeables. Entre 30 et 80 % de la surface est imperméable.
Landes et broussailles	322 - Landes et broussailles	Formations végétales basses et fermées, composées principalement de buissons, d'arbustes et de plantes herbacées (bruyères, ronces, genêts, ajoncs, cytises, etc.).

### 1.1.1. Description du programme STOC-Capture

La majorité (95%) des oiseaux échantillonnés proviennent du Suivi Temporel des Oiseaux Communs nicheurs par Capture (STOC-Capture) en France. Les 5% restants ont été capturés par des bagueurs mettant en place des protocoles spécialisés ciblant des taxons aviaires plus rares. Le programme STOC-Capture est mis en place par le Centre de Recherches sur la Biologie des Populations d'Oiseaux (CRBPO) au sein du Muséum National d'Histoire Naturelle. Initié en 1989, il est basé sur la méthode de capture-marquage-recapture des oiseaux présents dans les strates buissonnantes de la végétation. Il réalise ainsi le suivi d'une communauté aviaire sur un grand nombre de sites au cours du temps (environ 250 sites en France), en maintenant un effort de capture par site constant d'une année sur l'autre, grâce à des bagueurs bénévoles qui suivent un protocole d'échantillonnage standardisé. Ce programme a pour objectif de documenter à l'échelle nationale et sur le long terme le fonctionnement démographique

des populations d'oiseaux communs et l'influence des variations climatiques et d'habitat sur leur démographie.

La capture des oiseaux nicheurs se fait au printemps, avec un minimum de trois sessions par an entre mai et début juillet, à l'aide de filets japonais. Cette technique permet de capturer les oiseaux au sol ou volant à moins de 3 mètres du sol. La méthode de capture est standardisée pour chaque station STOC (moyens matériels et humains, emplacement des filets, durée et dates de sessions). Une station STOC est une zone de capture des oiseaux préalablement définie afin de garantir un habitat homogène, et présumé stable dans le temps.

Afin de cibler le moment de la journée où les oiseaux sont le plus actifs, une session se déroule généralement de l'aube à midi. La manipulation des oiseaux se fait sur place, elle est réduite au minimum et la plus rapide possible afin de limiter leur dérangement, et les oiseaux sont ensuite relâchés. Les oiseaux capturés sont marqués ou contrôlés via une bague avec un numéro unique. L'espèce, le sexe, l'âge des oiseaux ainsi que des mesures biométriques (masse, longueur d'aile et longueur du tarse) sont relevés.

#### 1.1.2. Collecte des tiques

En plus des informations collectées sur les oiseaux, un programme de collecte de tiques au sein du réseau de bagueurs a été mis en place en 2019 et 2020. Pour cela, un kit a été fourni aux bagueurs du programme STOC-Capture afin de leur permettre de collecter les tiques présentes sur les oiseaux capturés. Ce kit a aussi été envoyé aux bagueurs ciblant les taxons aviaires plus rares, afin de couvrir la plus grande diversité d'espèces aviaires possible. Le protocole de collecte consistait au prélèvement, sur tout le corps de l'oiseau, d'un maximum suggéré de 10 tiques par individu à l'aide de pinces fines. Pour les conserver, les tiques ont été placées dans des tubes eppendorf contenant de l'éthanol à 70% ; les tiques d'un oiseau donné étant stockées dans le même tube. Ces échantillons ont ensuite été envoyés au LSA de l'Anse pour analyse.

#### 1.1.3. Identification morphologique des tiques

Une identification morphologique des tiques pour déterminer la stase et l'espèce a ensuite été réalisée par Amalia RATAUD, sous loupe binoculaire, à l'aide des clefs taxonomiques d'Estrada-Peña et de Pérez-Eid (Pérez-Eid 2007; Estrada-Peña, Mihalca, et Petney 2017).

Suite à cette identification, les larves gorgées ont été sélectionnées afin de vérifier si elles étaient ou non porteuses d'agents pathogènes, et d'identifier ceux-ci le cas échéant. Sélectionner uniquement les larves gorgées pour la détection d'agents pathogènes permet d'obtenir un indicateur indirect du statut infectieux de l'oiseau. En effet, si les transmissions non-systémiques (co-repas et transovarienne) sont négligeables ou peu fréquentes, comme fortement suspecté pour les *Bbsl* (Richter et al. 2012), une larve s'infecte *via* un hôte réservoir infecté. On considère donc qu'une larve gorgée infectée par *Bbsl* signifie que l'oiseau hôte de cette larve était infecté par *Bbsl*.

#### 1.1.4. Détection des agents pathogènes au sein des larves gorgées

##### Extraction et pré-amplification de l'ADN des larves gorgées

L'ADN des larves gorgées a été extrait grâce au kit « Nucleopsin Tissue Genomic DNA from tissue » (Machery Nagel, Düren, Allemagne). Cela s'est fait en plusieurs étapes, notamment nettoyage, broyage et centrifugation des tiques, lyse des cellules puis fixation de l'ADN. Celui-ci a ensuite été conservé à -20 °C.

Afin d'obtenir une meilleure détection de l'ADN des agents pathogènes présents dans les larves gorgées, l'ADN des pathogènes a été pré-amplifié avec le kit « PreAmp Master Mix » (Fluidigm, San Francisco, CA, USA). Ce kit permet d'augmenter le signal du matériel génétique des agents pathogènes par rapport au matériel génétique de la tique. Cette pré-amplification a nécessité l'utilisation d'amorces. Les échantillons ont ensuite été conservés à -20°C.

##### Détection des agents pathogènes présents dans les larves gorgées

Afin d'identifier les agents pathogènes présents dans les larves gorgées collectées sur les oiseaux, la technique de PCR microfluidique en temps réel (Biomark™ dynamic array, Fluidigm) a été utilisée, sur



## 1.2. Nettoyage des données et obtention des séquences consensus pour les analyses phylogénétiques

Dans l'optique de réaliser un arbre phylogénétique à partir des séquences de *Bbsl* identifiées dans les larves gorgées collectées sur les oiseaux communs lors du projet OUTLYER, nous avons utilisé les séquences génétiques d'une partie du gène *flaB* des *Bbsl*, obtenues suite à la PCR et au séquençage de l'ADN prélevé chez les larves gorgées positives à *Bbsl*. 121 *Bbsl* et leurs séquences associées étaient initialement disponibles.

Du fait de l'utilisation d'amorces Forward (F) et Reverse (R) lors du séquençage, deux séquences (une F et une R) étaient obtenues pour chaque larve gorgée positive à *Bbsl*.

A partir de ces deux séquences F et R obtenues, une séquence consensus a été déterminée, qui réunit l'information contenue dans les deux séquences (Figure 6). Au final, une séquence consensus était obtenue par larve gorgée positive à *Bbsl*.

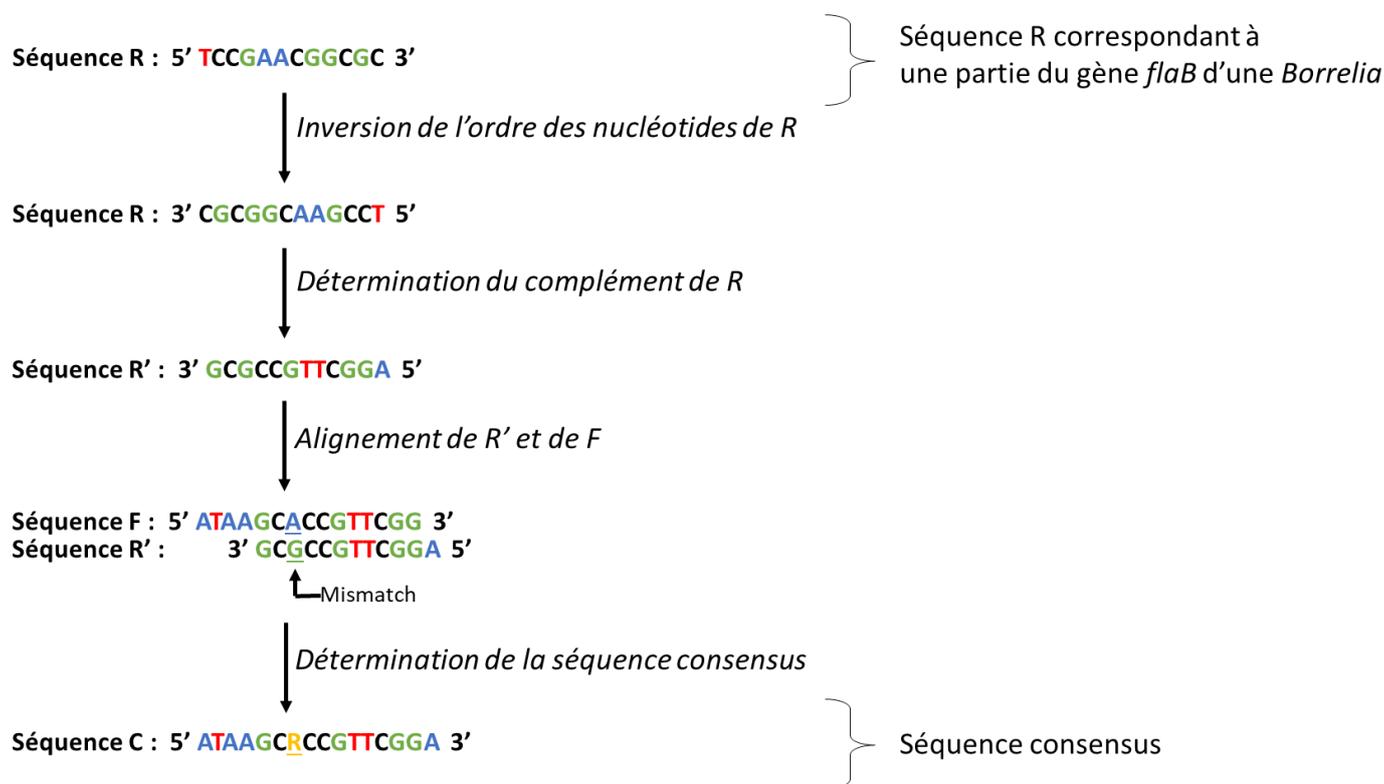


Figure 6 : Schéma d'obtention d'une séquence consensus à partir des séquences F et R d'une *Bbsl*

La lettre R en jaune correspond à un nucléotide ambigu (peut être A ou G).

En cas de divergence entre deux nucléotides lors de l'alignement (possiblement dû à une erreur lors de la synthèse des brins ou lors du séquençage), le nucléotide de la séquence consensus a pu être remplacé par un nucléotide ambigu, afin de garder le plus d'information possible sur ce site. Par exemple, un R signifie que le nucléotide de la séquence consensus est soit un A, soit un G. Sinon, il est remplacé par un N (le nucléotide peut être A, T, G, ou C).

La qualité des séquences consensus obtenues a ensuite été évaluée. Les séquences contenant moins de 300 nucléotides ont été exclues pour la suite des analyses. En effet, d'après la distribution du nombre de séquences en fonction du nombre de nucléotides, la majorité des séquences (102/121) contenaient plus de 300 nucléotides. De plus, une séquence trop courte est considérée de mauvaise qualité (amorces ne s'étant pas fixées au bon endroit, problème lors du séquençage...) et peut entraîner des erreurs lors de l'alignement de l'ensemble des séquences.

Ensuite, un alignement de l'ensemble des séquences a été réalisé, afin de rendre les séquences de chaque *Bbsl* comparables. Les séquences ne commençant et ne se terminant pas toutes au même

endroit par rapport au gène séquencé (différences de longueur), l'alignement permet de s'assurer, lors de la comparaison d'un nucléotide se trouvant à la position x des séquences, que ce nucléotide corresponde bien à celui situé à la position x du gène séquencé, et ce pour l'ensemble des séquences. Pour faire cela, l'algorithme Muscle a été utilisé (Annexe 3). Cet algorithme compare les séquences deux à deux, et crée une matrice des distances génétiques à partir de laquelle il aligne les séquences de proche en proche. A la fin, les séquences sont toutes alignées, et forment un alignement multiple (msa : multiple sequence alignment) (Figure 7).

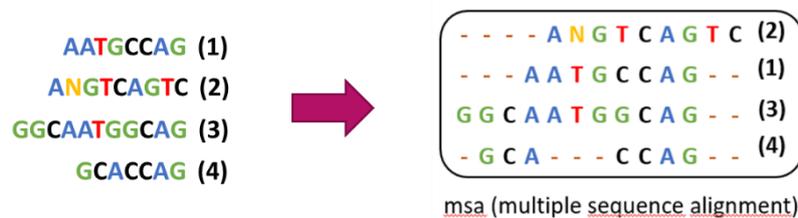


Figure 7 : Séquences non alignées (gauche) et alignement multiple des séquences (droite)

(1), (2), (3) et (4) sont les noms des 4 taxa (séquences génétiques) étudiés. La lettre N en jaune représente un nucléotide non informatif (peut être A, T, G ou C). Les tirets en marron représentent des lacunes dans l'alignement, ces lacunes correspondent soit à des insertions/délétions dans la séquence concernée, soit à des introductions artificielles (en début et/ou fin de séquence), afin que les séquences aient toutes la même longueur pour permettre la comparaison.

Une fois l'alignement réalisé, une inspection visuelle de celui-ci a été effectuée pour juger de sa qualité, conduisant à l'élimination de 6 séquences consensus des analyses. Ces séquences étaient trop différentes des autres, et ne s'alignaient pas correctement avec le reste des séquences, ce qui diminuait la qualité de l'alignement multiple.

Un nouvel alignement multiple a été effectué sur les 96 séquences restantes, et a été conservé pour la suite des analyses. Le début et la fin des séquences de l'alignement multiple comportant beaucoup de lacunes, un élagage a été réalisé (conservation d'une partie des séquences). En effet, un nombre trop important de lacunes peut modifier la topologie des arbres générés.

### 1.3. Analyses phylogénétiques

A partir des séquences alignées, une nouvelle matrice des distances génétiques des séquences a été construite. Un modèle de substitution a été choisi afin de corriger le calcul de la distance génétique entre les paires de séquences (voir paragraphe 2.3.2). Le modèle choisi (F84) est un modèle général, proche du modèle GTR (General Time Reversible). Le modèle GTR a souvent été utilisé lors de la construction d'arbres phylogénétiques à partir de séquences de *B. garinii*. Dans ces deux modèles, les fréquences des 4 nucléotides A, T, G et C sont différentes. Le modèle F84 utilise des taux de substitutions différents pour les transitions ( $A \leftrightarrow G$  ou  $C \leftrightarrow T$ ) et les transversions (toutes les autres substitutions :  $A \leftrightarrow C$ ,  $A \leftrightarrow T$ , etc), alors que le modèle GTR utilise des taux différents pour chaque type de substitutions (Xia 2020b).

Ensuite, un arbre phylogénétique a été construit grâce à la méthode du maximum de parcimonie (voir paragraphe 2.3.2). L'algorithme de Fitch a été utilisé pour attribuer un score de parcimonie aux arbres générés (même fréquence de substitution pour toutes les bases). La longueur des branches, qui représente la distance génétique entre deux nœuds ou séquences, a été calculée grâce à un algorithme (acctran). Lorsque plusieurs arbres ont le même score de parcimonie, un arbre consensus est produit, qui va présenter des multifurcations, c'est-à-dire des nœuds internes ayant plus de deux descendants (Figure 8). Ces multifurcations signifient que l'algorithme n'a pas pu déterminer quelles paires de séquences étaient les plus proches au sein du groupe.

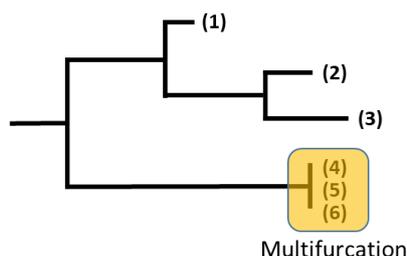


Figure 8 : Exemple d'arbre présentant une multifurcation, avec longueur de branche représentant la distance génétique entre deux nœuds

(1) à (6) sont les taxa étudiés.

Une fois l'arbre consensus généré, les données épidémiologiques d'intérêt ont été associées à cet arbre : à côté de chaque taxon de l'arbre (c'est-à-dire de chaque séquence génétique d'une *Bbsl* récoltée sur une larve gorgée provenant d'un oiseau) ont été ajoutées l'espèce de bactérie associée à la séquence, l'espèce d'oiseau dont elle provient, ainsi que son milieu de vie. Les larves d'un même oiseau ont aussi été mises en évidence pour les oiseaux portant plus de deux larves positives à *Bbsl*.

A partir de l'arbre consensus obtenu, des clades (regroupement de bactéries génétiquement proches) ont été identifiés. Les *Bbsl* issues de ces clades ont ensuite été localisées sur la carte de la France, afin de déterminer si les séquences génétiquement proches étaient aussi proches d'un point de vue spatial.

## 2. Résultats :

### 2.1. Description des données

#### 2.1.1. Description des données de terrain relatives aux oiseaux capturés, aux larves collectées et aux agents pathogènes détectés

Au total, 1040 oiseaux ont été échantillonnés au printemps en 2019 et 2020, appartenant à 56 espèces, dans 110 sites en France. 3114 tiques ont été prélevées chez ces oiseaux, parmi lesquelles 61 adultes (60 femelles, 1 mâle), 1715 nymphes, 1285 larves (dont 1106 larves gorgées) et 53 tiques dont la stase n'a pu être identifiée.

Tableau 2 : Nombre de larves récoltées en fonction de leur espèce (ensemble des larves, puis larves positives à *Bbsl*, puis larves positives à *B. garinii*)

Nombre de larves gorgées	<i>Ixodes ricinus</i>	<i>Haemaphysalis punctata</i>	<i>Haemaphysalis concinna</i>	<i>Haemaphysalis Spp *</i>	<i>Ixodes frontalis</i>	<i>Ixodes spp *</i>	Non identifiée**	Total
<b>Collectées</b>	1039	15	11	3	2	2	34	1106
<b>Positives à <i>Bbsl</i></b>	124	0	0	1	2	1	1	129
<b>Positives à <i>B. garinii</i></b>	96	0	0	1	1	1	1	100

\* Tiques pour lesquelles l'identification de l'espèce a échoué

\*\* Tiques non identifiées car trop abimées

Les 1106 larves gorgées ont été prélevées sur 442 oiseaux appartenant à 36 espèces, dans 69 sites en France. Au total, 4 espèces de tiques appartenant à 2 genres ont été identifiées (Tableau 2). Certaines larves n'ont pas pu être identifiées pour l'espèce ou le genre, car elles étaient trop abimées.

Parmi les 1106 larves analysées, 11.7% (n = 129) étaient positives pour le groupe des *Bbsl*, 5.7% (n = 63) pour le genre *Anaplasma*, 2.3% (n = 25) pour le genre *Babesia*, 1.4% (n = 15) pour le genre *Ehrlichia*, et 1% (n = 11) pour l'espèce *Borrelia miyamotoi* (espèce de *Borrelia* n'appartenant pas au complexe des *Bbsl*). Les larves positives à *Bbsl* (n = 129) ont été retrouvées dans 28 stations STOC-Capture en France (Figure 9). 78% des *Bbsl* de notre échantillon appartenait à l'espèce *B. garinii* (n = 100). Les larves positives à *B. garinii* ont été retrouvées dans 25 stations (Annexe 4).

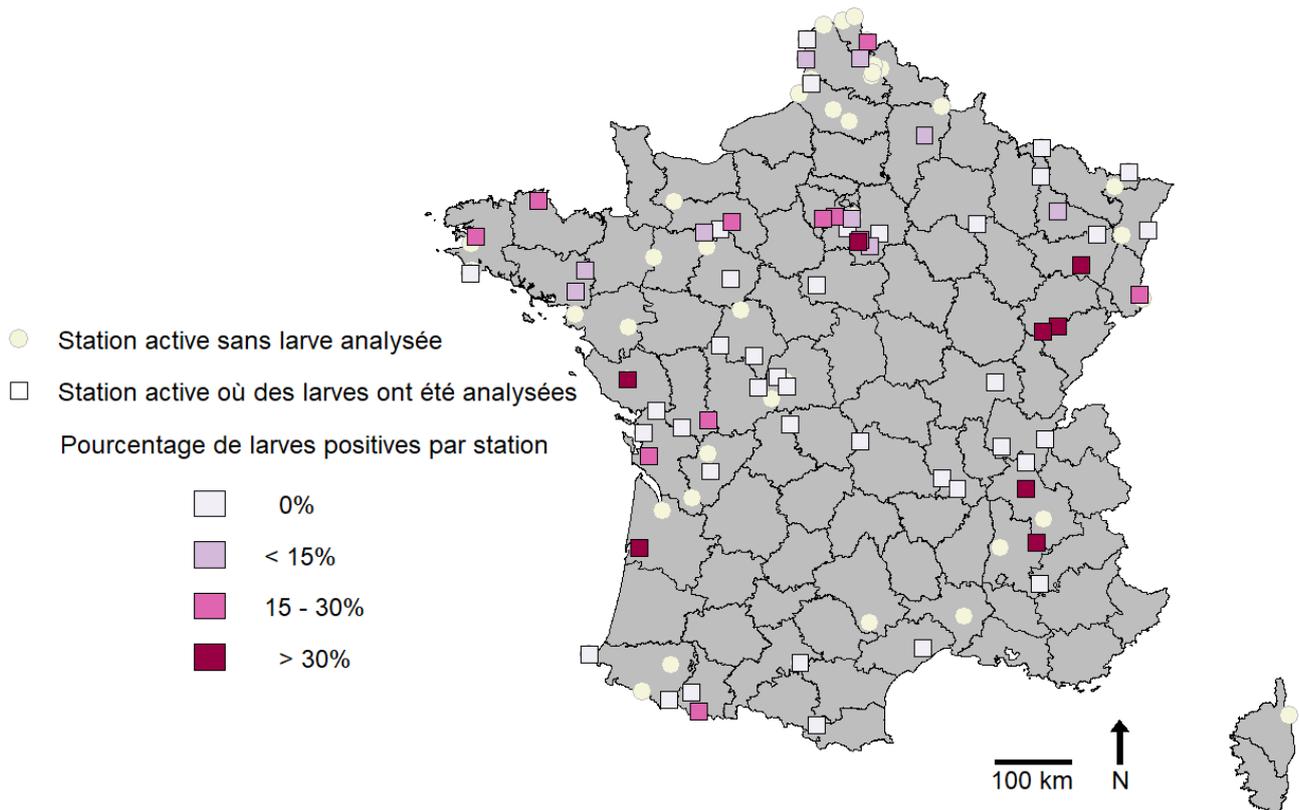


Figure 9 : Carte présentant les pourcentages de larves positives à *Bbsl* en fonction des stations en France

Les stations représentées sont les stations au sein desquelles des oiseaux ont été capturés en 2019-2020.

Parmi les 442 oiseaux sur lesquels des larves gorgées ont été collectées puis analysées pour *Bbsl*, 16% hébergeaient au moins une larve positive à *Bbsl* (n = 70), et 12% hébergeaient au moins une larve positive à *B. garinii* (n = 54). Parmi les 70 oiseaux porteurs de larves positives à *Bbsl*, 66% (n = 46) portaient plus d'une larve (Figure 11).

Les oiseaux porteurs de larves positives à *Bbsl* appartenait à 15 espèces d'oiseaux (sur les 36 espèces capturées). Les oiseaux porteurs de larves positives à *B. garinii* appartenait à 13 espèces d'oiseaux (Annexe 5). Le pourcentage d'oiseaux porteurs de larves gorgées positives à *Bbsl* était inférieure à 10% pour 6 espèces d'oiseaux (Figure 10). Pour le Merle noir et la Grive musicienne, les pourcentages d'individus positifs (porteurs d'au moins une larve positive) étaient respectivement de 67% et de 60%. Le nombre d'individus porteurs de larves gorgées était inférieur à 15 pour 6 espèces d'oiseaux.

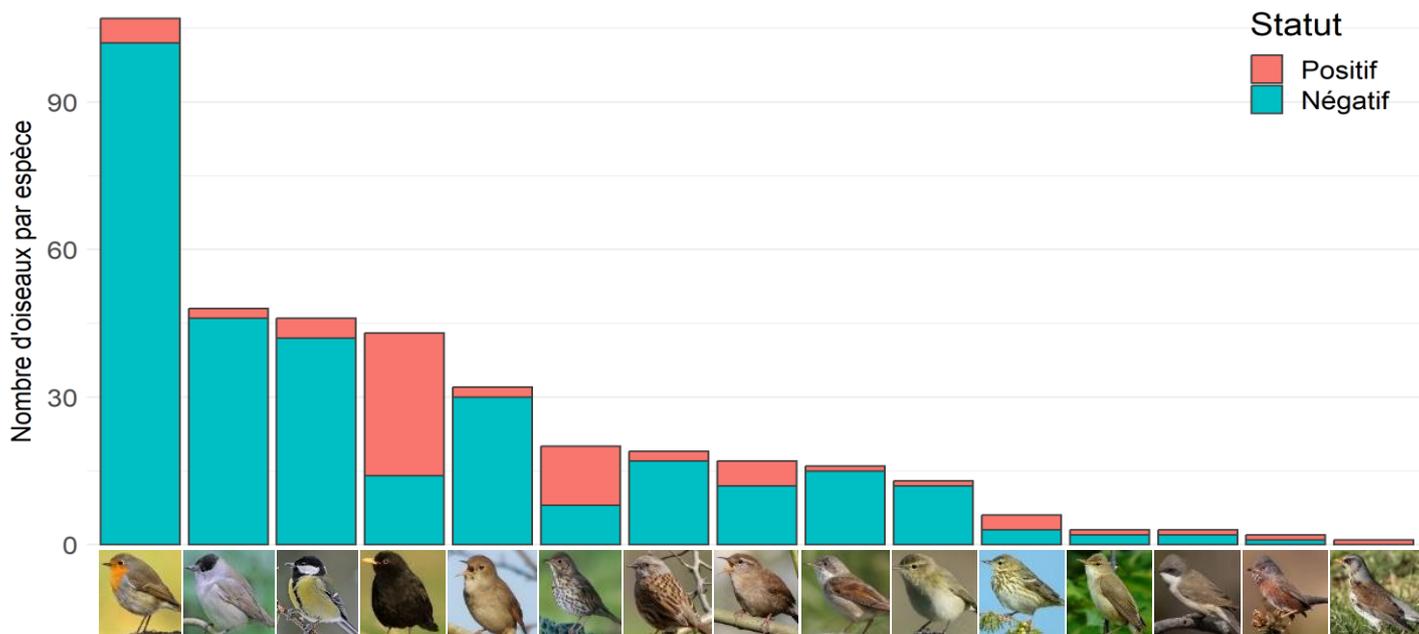


Figure 10 : Espèces d'oiseaux positives à *Bbsl*, et nombre d'individus porteurs de larves négatives et porteurs d'au moins une larve positive à *Bbsl* par espèce d'oiseau

Les photographies en bas du graphique correspondent aux espèces d'oiseaux concernées, qui sont, de gauche à droite : Rougegorge familier (*Erithacus rubecula*), Fauvette à tête noire (*Sylvia atricapilla*), Mésange charbonnière (*Parus major*), Merle noir (*Turdus merula*), Rossignol philomèle (*Luscinia megarhynchos*), Grive musicienne (*Turdus philomelos*), Accenteur mouchet (*Prunella modularis*), Troglodyte mignon (*Troglodytes troglodytes*), Fauvette grisette (*Curruca communis*), Pouillot véloce (*Phylloscopus collybita*), Pipit des arbres (*Anthus trivialis*), Rousserolle verderolle (*Acrocephalus palustris*), Fauvette babillarde (*Curruca curruca*), Fauvette pitchou (*Curruca undata*), Grive litorne (*Turdus pilaris*).

Les photographies sont toutes issues du site *oiseaux.net*.

Pour les espèces autres que le Merle noir et la Grive musicienne, le pourcentage de larves positives à *Bbsl* par oiseau tend à diminuer avec le nombre de larves collectées sur l'oiseau (Figure 11). Environ la moitié de ces oiseaux (55%) ont un pourcentage de larves positives inférieur ou égal à 50%. Pour le Merle noir et la Grive musicienne, le pourcentage de larves positives par individu est très élevé, et ce quel que soit le nombre de larves collectées sur l'oiseau : chez le Merle noir, 75% des individus positifs à *Bbsl* pour lesquels plusieurs larves ont été analysées portent 100% de larves positives à *Bbsl*, contre 66% pour la Grive musicienne. Des résultats similaires sont obtenus pour les larves positives à *B. garinii* (Annexe 6).

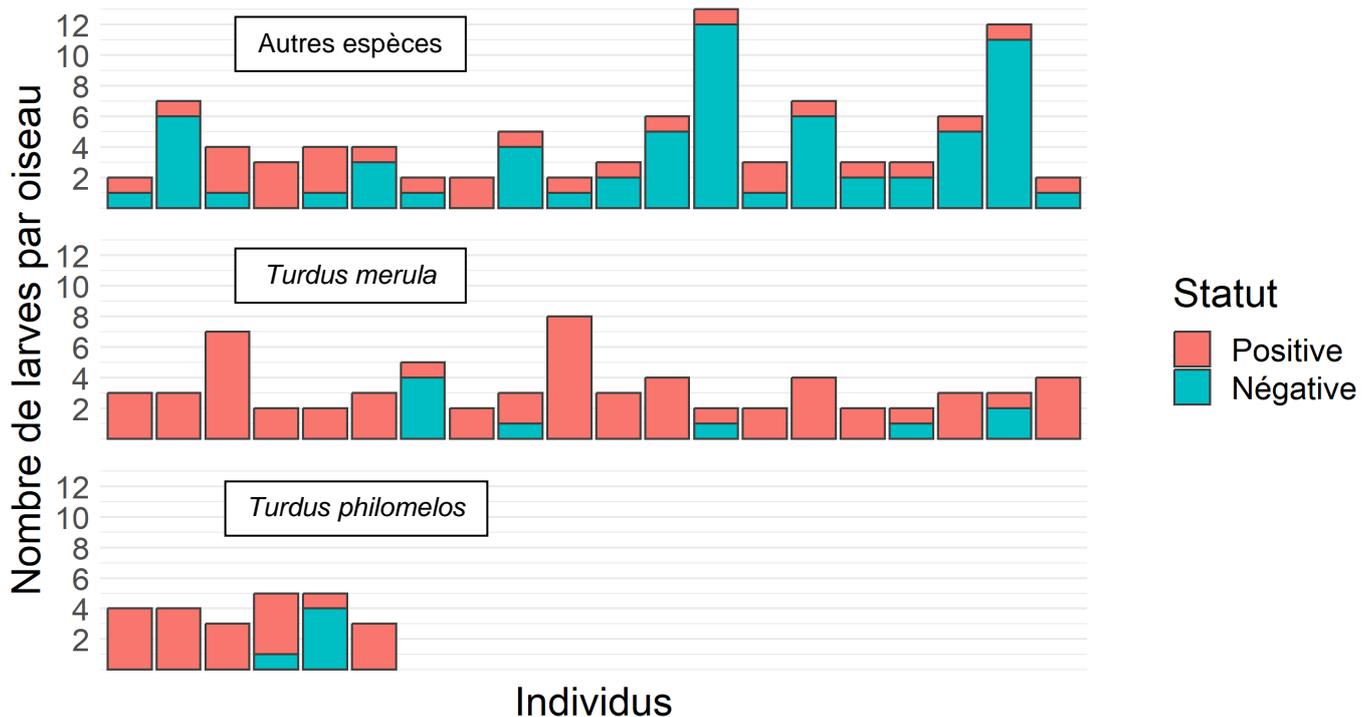


Figure 11 : Nombre de de larves positives et négatives à *Bbsl* par oiseau en fonction de l'espèce

Les oiseaux représentés ici sont les oiseaux pour lesquels au moins deux larves ont été analysées, et dont au moins une des larves était positive à *Bbsl* ( $n = 46$ ). La catégorie « Autres espèces », inclue les individus issus des espèces présentées dans la figure précédente (Figure 10), à l'exception du Merle noir et de la Grive musicienne.

### 2.1.2. Description et caractéristiques des séquences génétiques

Lors du nettoyage et de la sélection des séquences utilisées pour les analyses phylogénétiques, certaines larves gorgées n'avaient qu'une seule séquence de *Bbsl* exploitable (F ou R). Lorsqu'une seule séquence était disponible, celle-ci a été incluse telle quelle dans les analyses, après inversion de l'ordre des nucléotides pour les séquences R (ces séquences sont aussi considérées comme des séquences consensus par la suite).

121 *Bbsl* ont été séquencées, 102 séquences génétiques ont été conservées après le nettoyage, et 96 séquences ont été conservées après le premier alignement multiple et incluses dans les analyses.

La longueur des séquences issues du 1er alignement multiple était de 702 nucléotides. Seule la partie des séquences comprise entre le 67<sup>ème</sup> et le 469<sup>ème</sup> nucléotide a été gardée. L'alignement multiple final était donc constitué de séquences de 403 nucléotides.

Les 96 séquences génétiques de *Bbsl* sont issues de larves gorgées collectées sur 54 oiseaux. 24 de ces oiseaux étaient porteurs de plus d'une larve positive à *Bbsl*. Ces oiseaux appartenaient à 12 espèces regroupant 8 genres (*Turdus*, *Troglodytes*, *Erithacus*, *Anthus*, *Parus*, *Luscinia*, *Sylvia*, *Phylloscopus*). Ces genres sont tous issus de l'ordre des Passériformes (passereaux). Les oiseaux ont été capturés dans les 6 milieux définis (clairsemé, landes et broussailles, agricole, humide, forêt, périurbain).

La majorité des séquences génétiques étudiées sont des séquences correspondant aux espèces *B. garinii* (74/96) et *B. valaisiana* (12/96) (Figure 12). Les larves gorgées (et les séquences génétiques des *Bbsl* associées) proviennent majoritairement de Merles noirs (56/96), et de Grives musiciennes (15/96), et sont majoritairement issues de milieux clairsemés (39/96) ou de milieux type « landes et broussailles » (29/96) (Figure 12).

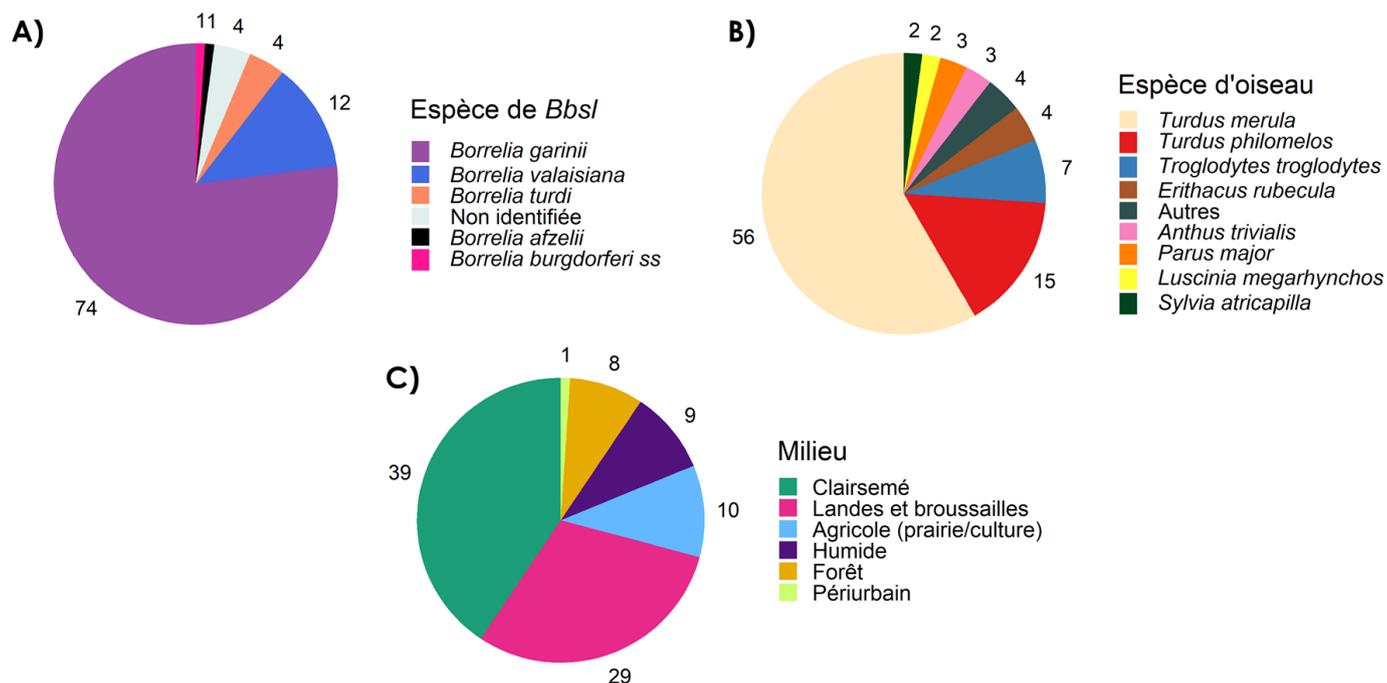


Figure 12 : Espèces de *Bbsl* correspondant aux 96 séquences génétiques (A), espèces d'oiseaux dont sont issues les 96 larves gorgées positives à *Bbsl* (B), milieux dont sont issues les 96 larves gorgées récoltées sur les 54 oiseaux capturés (C)

B) La catégorie « Autres » comprend 4 espèces pour lesquelles une seule larve gorgée positive à *Bbsl* a été trouvée : *Sylvia communis*, *Phylloscopus collybita*, *Sylvia curruca* et *Sylvia undata*.

## 2.2. Analyses des données : diversité génétique des séquences et phylogénie

Les séquences génétiques étudiées sont proches les unes des autres. En effet, la distance génétique maximale entre deux séquences est de 14% (Figure 13). 93% des paires de séquences étudiées ont une distance génétique inférieure à 5%, et 21% ont une distance génétique nulle. Plusieurs groupes se distinguent par la proximité de leurs séquences génétiques intragroupe, et par leur distance génétique plus élevée avec les séquences hors du groupe : ces groupes correspondent aux différentes espèces de *Bbsl* dont sont issues les séquences. Pour *B. garinii*, on observe aussi plusieurs sous-groupes de bactéries au sein desquelles la distance génétique est très faible ou nulle : les séquences génétiques de ces sous-groupes sont quasiment identiques. Le groupe constitué d'un mélange d'espèces de *Bbsl* contient des séquences génétiques qui présentent une grande distance génétique intra et inter-groupe (distance génétique supérieure à 5%, et allant jusqu'à 14%).

Il faut noter qu'une séquence génétique identifiée comme étant celle d'une *B. garinii* est plus proche des séquences génétiques de *B. valaisiana* que des autres *B. garinii* (inclusion dans le groupe des *B. valaisiana*). Les séquences génétiques provenant d'espèces de *Bbsl* non identifiées (4 en tout), sont toutes très proches génétiquement des *B. garinii*, suggérant leur appartenance à cette espèce.

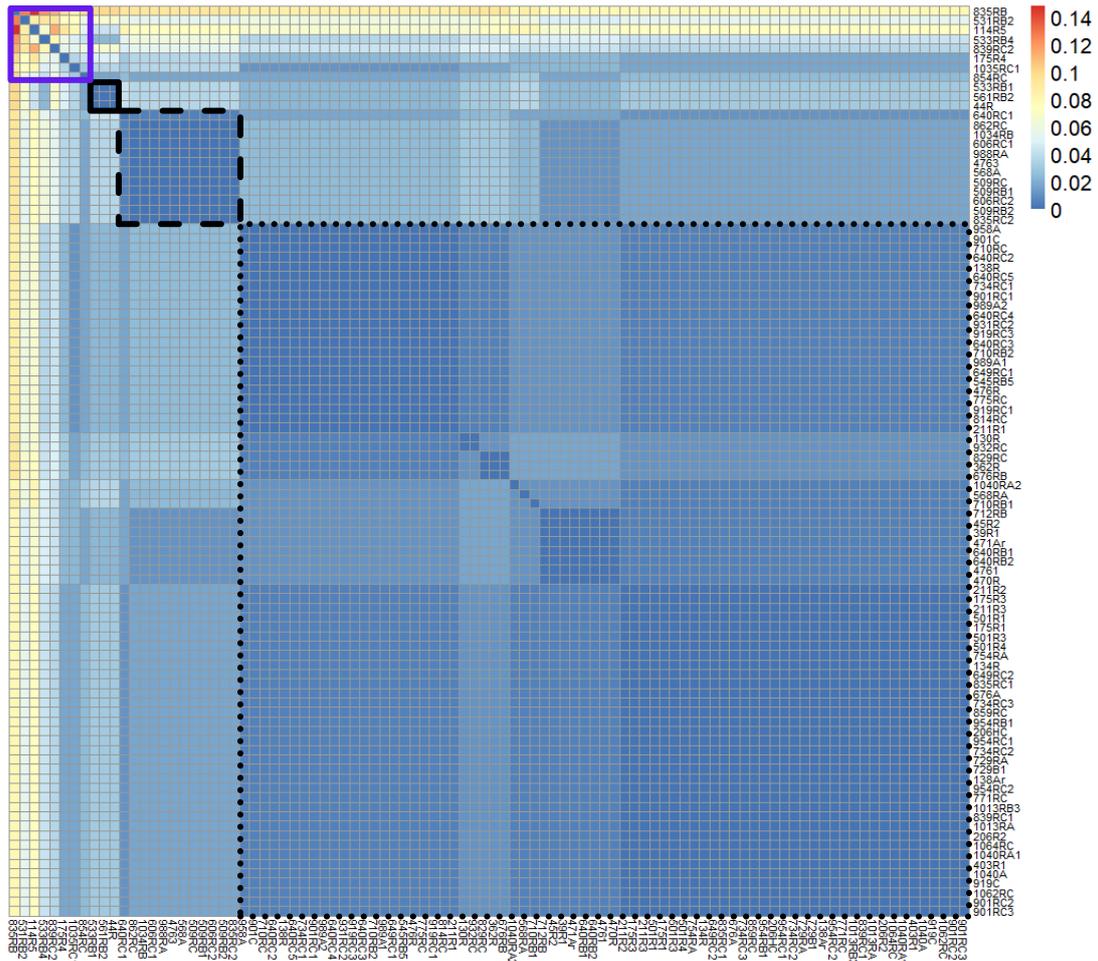


Figure 13 : Matrice des distances génétiques des séquences

Chaque case de la matrice représente la distance génétique entre deux séquences : la couleur des cases varie du bleu foncé (absence de distance génétique entre les deux séquences) au rouge foncé (distance génétique la plus élevée). Les noms des séquences génétiques comparées figurent en fin de ligne/colonne. La matrice est carrée et symétrique : sa diagonale correspond à la comparaison des séquences avec elles-mêmes. Les grands carrés noirs regroupent les distances entre les séquences génétiques issues d'une même espèce de *Bbsl*. Le carré noir en pointillés correspond à *B. garinii*, celui avec des traits discontinus correspond à *B. valaisiana*, tandis que celui en traits pleins correspond à *B. turdi*. Le grand carré violet représente un mélange d'espèces de *Bbsl*.

On remarque que de nombreuses multifurcations sont présentes sur l'arbre phylogénétique construit à partir des 96 séquences de *Bbsl* (Figure 14). En effet, l'arbre est constitué de 32 nœuds internes, alors qu'il pourrait y en avoir 95 si chaque nœud interne avait deux descendants. Deux grands clades (groupes constitués par un ancêtre commun et l'ensemble de ses descendants, au sein desquels les séquences génétiques sont les plus proches) sont visibles sur cet arbre. On définit le clade A (75 séquences) et le clade B (18 séquences), qui incluent la majorité des séquences de l'arbre. Trois séquences sont exclues de ces deux clades, car trop différentes des autres. Les 75 séquences composant le clade A sont issues de larves collectées sur des oiseaux capturés dans 23 stations, alors que les 18 séquences composant le clade B sont issues de larves collectées sur des oiseaux capturés dans 10 stations (Figure 15). De plus, on définit deux sous clades au sein du clade A : le sous clade A1 (26 séquences provenant de 17 stations) et le sous clade A2 (24 séquences provenant de 11 stations).

Les séquences composant le clade A sont toutes issues de *B. garinii*, alors que le clade B regroupe les autres espèces (sauf une *B. garinii* présente dans ce clade). Le clade B est composé de deux sous clades, un sous clade regroupant les séquences issues de *B. valaisiana* (sauf une), et un sous clade regroupant les séquences issues de *B. turdi*. Il y a donc un regroupement des séquences génétiques en fonction de l'espèce de bactérie dont elles sont issues : les séquences issues d'une même espèce

de *Bbsl* sont les plus proches génétiquement. Les séquences issues d'espèces de *Bbsl* non identifiées sont incluses dans le clade des *B. garinii* : elles proviennent probablement de cette espèce. Les séquences de *B. garinii* sont globalement plus proches entre elles que les séquences de *B. valaisiana* et *B. turdi* (moindre longueur de branche pour *B. garinii* par rapport à *B. valaisiana* et *B. turdi*).

Il n'y a pas de regroupement strict des séquences de *Bbsl* en fonction de l'espèce d'oiseau dont les larves positives sont issues, ou du milieu de vie de l'oiseau. Cependant, le grand nombre de Merles noirs présents dans notre échantillon limite la visualisation d'une structure. Le sous clade A1 présente une diversité d'espèces d'oiseaux importante, où le Merle noir est moins présent qu'ailleurs (le Merle noir n'est pas non plus présent dans le sous clade regroupant *B. turdi*). En effet, les 26 séquences du sous clade A1 sont issues de larves collectées sur 8 espèces d'oiseaux, avec 6 larves collectées sur des Merles noirs, 6 larves collectées sur des Grives musiciennes, et 6 larves collectées sur des Troglodytes mignons (les larves issues des autres espèces sont minoritaires). Cette diversité d'espèces aviaires va de pair, pour le sous clade A1, avec une diversité de milieux de vie (5 milieux différents, avec une majorité de milieux clairsemés (9), landes et broussailles (5) et forêt (5)). Les séquences du clade A2 sont majoritairement issues de larves collectées sur des Merles noirs, et ces larves proviennent en majorité de milieux clairsemés (24 séquences de *Bbsl* issues de larves collectées sur 5 espèces d'oiseaux, avec 20 larves collectées sur des Merles noirs, et 17 larves issues de milieux clairsemés). On distingue donc des clades où les Merles noirs sont plutôt regroupés, et des clades avec une diversité aviaire plus importante.

Globalement, les séquences génétiques des *Bbsl* issues des larves portées par un même oiseau sont proches entre elles (mais ce n'est pas toujours le cas, notamment quand les séquences sont issues de différentes espèces de *Bbsl*).

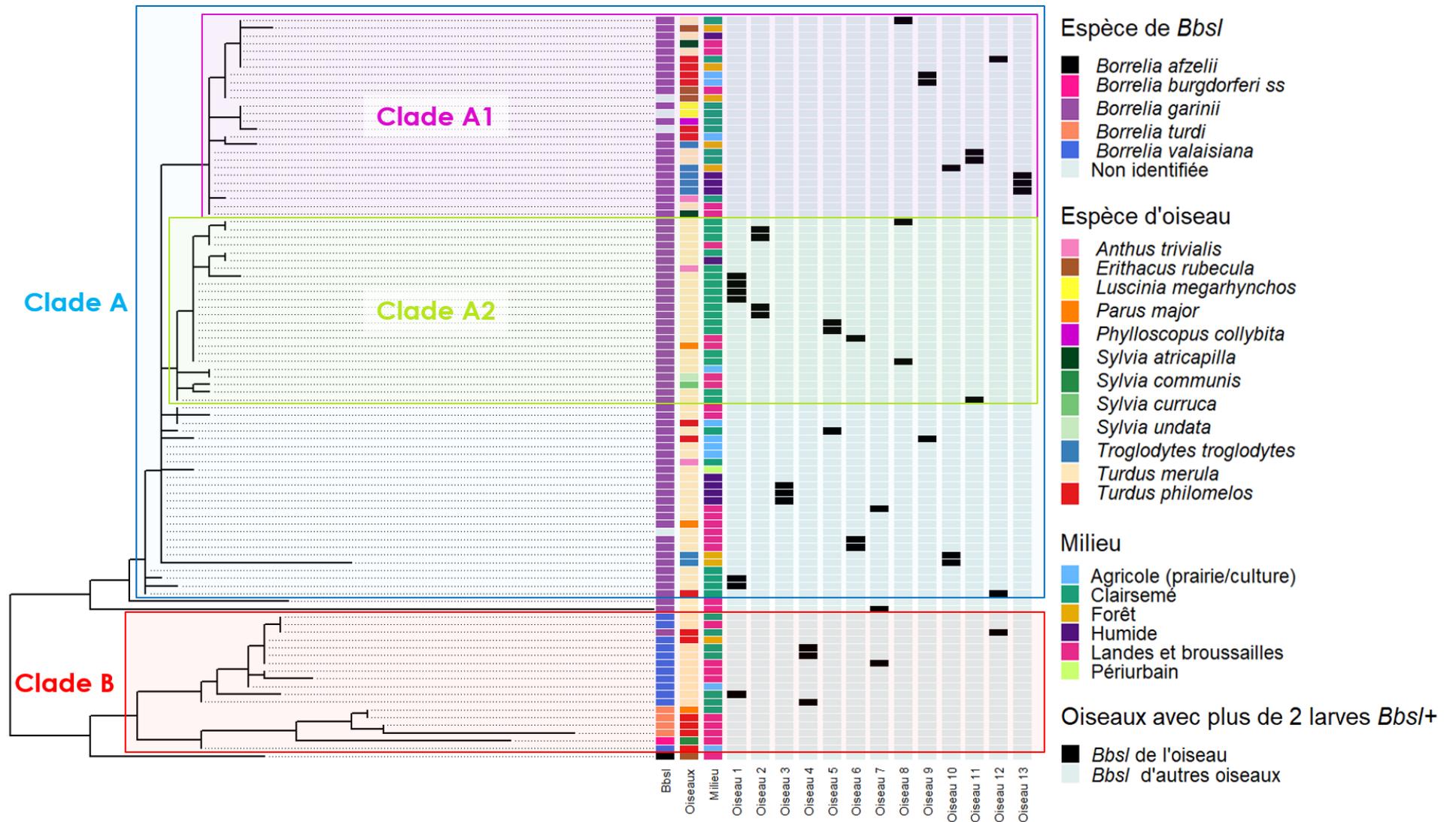


Figure 14 : Arbre phylogénétique construit à partir des 96 séquences génétiques de *Bbsl*

Arbre obtenu par la méthode du maximum de parcimonie. La longueur des branches représente la distance génétique entre les séquences.

Nous n'observons pas de regroupement spatial des séquences issues des clades A et B (Figure 15). *B. garinii* (clade A) se retrouve dans la majorité des territoires de France où des données ont pu être collectées (station STOC avec collecte de larves). Les autres espèces de *Bbsl* (clade B) se retrouvent plus au Nord de la France. Des larves positives à *B. garinii* et aux autres *Bbsl* sont présentes dans des mêmes stations : il n'y a pas de ségrégation spatiale apparente entre les clades A et B.

Les *B. garinii* appartenant aux sous clades A1 et A2 sont aussi retrouvées dans de nombreux territoires, et dans des mêmes stations sur la carte de France (Figure 15) : les *B. garinii* génétiquement proches ne semblent pas être proches d'un point de vue spatial. Visuellement, il ne semble pas y avoir de structure géographique pour *B. garinii*.

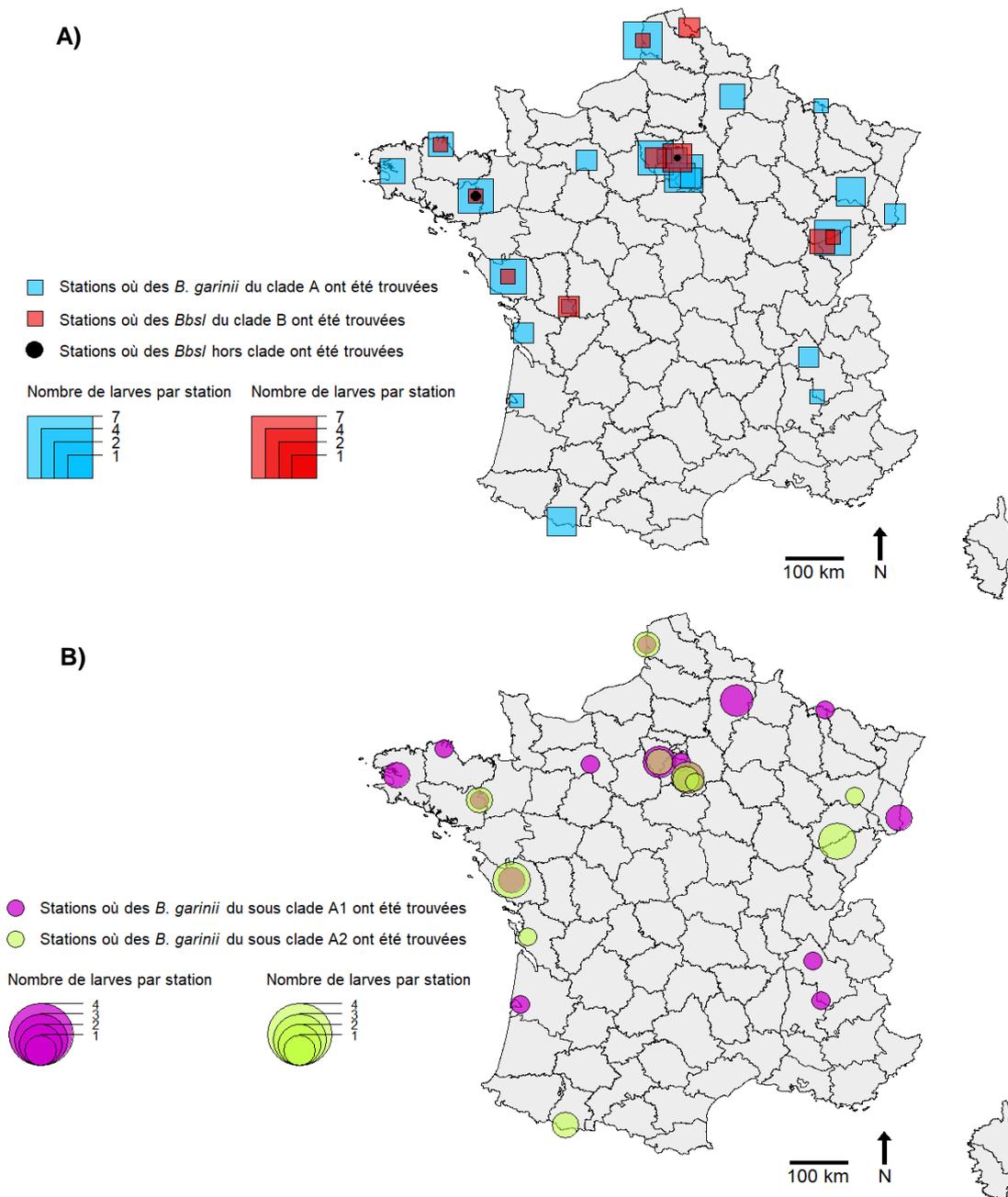


Figure 15 : A) Répartition spatiale des *Bbsl* appartenant aux clades A et B, et des *Bbsl* hors clades. B) répartition spatiale des *B. garinii* appartenant aux sous clades A1 et A2

A) Les *Bbsl* hors clade ( $n=3$ ) ont été identifiées chez des larves collectées sur des oiseaux capturés dans deux stations. Ces stations sont représentées par des points noirs de taille différente en fonction du nombre de larves trouvées (1 larve pour le plus petit point, et 2 pour le plus grand).

### 3. Discussion :

Notre étude avait pour but de caractériser la diversité génétique des *Bbsl* (et principalement de *B. garinii*) chez les oiseaux en France, de décrire leur répartition géographique et d'étudier le lien entre diversité génétique et espèce aviaire hôte. Nous avons pour cela utilisé 96 séquences génétiques de *Bbsl* (dont 74 correspondant à l'espèce *B. garinii*) détectées chez des larves gorgées collectées sur des oiseaux dans le cadre d'un projet antérieur (OUTLYER), qui avait pour but d'étudier le rôle des oiseaux dans la circulation des *Bbsl* en France. Cette étude a permis de réaliser une première description de la répartition spatiale et de la diversité génétique de *B. garinii* chez les oiseaux en France. Les séquences étudiées (fragment du gène *flaB*) étaient génétiquement très proches, et *B. garinii* était globalement retrouvée dans toute la France, là où des larves avaient pu être collectées. L'étude de l'arbre phylogénétique construit par maximum de parcimonie associé aux métadonnées d'intérêt (espèce de *Bbsl*, espèce d'oiseau, milieu de vie de l'oiseau) n'a pas mis en évidence de regroupement strict des *Bbsl* en fonction de l'espèce aviaire hôte ou de son milieu de vie. Il y avait cependant des clades au sein desquels les *Bbsl* (issues des larves gorgées) provenaient majoritairement de Merles noirs et des clades présentant une diversité aviaire plus importante. Enfin, la géolocalisation des *Bbsl* constituant les différents clades identifiés n'a pas permis de visualiser l'existence d'une structure spatiale chez *B. garinii*.

La grande majorité (94%) des larves gorgées collectées sur les oiseaux appartenaient à l'espèce *I. ricinus*. L'espèce de *Bbsl* majoritaire retrouvée chez ces larves était *B. garinii* (78%). Des larves positives à *B. garinii* (et à *Bbsl* en général) ont été retrouvées chez de nombreuses espèces d'oiseaux (13 pour *B. garinii*, 15 pour *Bbsl*). Le Merle noir et la Grive musicienne se démarquaient par un pourcentage élevé d'individus porteurs d'au moins une larve positive à *Bbsl* (respectivement 67% et 60%, pour 43 et 20 individus porteurs de larves gorgées appartenant à ces deux espèces). Pour d'autres espèces comme le Pipit des arbres ou la Fauvette pitchou, ce pourcentage était supérieur ou égal à 50%, mais le nombre d'individus porteurs de larves échantillonnés pour ces espèces était très faible ( $n \leq 6$ ) : ces pourcentages ne constituent donc probablement pas un bon indicateur du pourcentage de positivité réel des individus issus de ces espèces aviaires. De plus, le Merle noir et la Grive musicienne se démarquaient aussi par un pourcentage élevé de larves positives chez les individus positifs porteurs de plusieurs larves : pour le Merle noir, 75% (15/20) de ces individus portaient 100% de larves positives à *Bbsl*, contre 67% (4/6) pour la Grive musicienne. Ces résultats sont cohérents avec les données de la littérature : *B. garinii* est la *Bbsl* la plus fréquemment retrouvée chez les tiques récoltées sur des oiseaux pour différents pays d'Europe (Michalik et al. 2008; Dubska et al. 2009; Gryczyńska et Welc-Fałęciak 2016; Klaus et al. 2016). Dans ces études, le Merle noir et la Grive musicienne font aussi partie des espèces portant le plus de larves positives à *Bbsl*.

L'ensemble des *Bbsl* étudiées étaient génétiquement proches pour le fragment de gène séquencé (*flaB*) : 93% des paires de séquences étudiées avaient une distance génétique inférieure à 5%. La distance génétique était la plus faible pour les paires de séquences issues d'une même espèce de *Bbsl*. Quelques bactéries étaient plus proches d'autres espèces que celle qui leur avait été attribuée, ce qui pourrait s'expliquer par une mauvaise identification de l'espèce pour ces bactéries lors du séquençage. La matrice des distances montrait plusieurs sous-groupes de *B. garinii*, sous-groupes qui n'ont pas été observés chez *B. valaisiana* et *B. turdi*. Cela pourrait suggérer l'existence d'une plus grande diversité génétique au sein de *B. garinii* par rapport à d'autres *Bbsl*, comme montré dans certaines études sur des tiques à l'affût (Vollmer et al. 2011; Jacquot et al. 2014). Cependant, le faible nombre de séquences disponibles pour *B. valaisiana* et *B. turdi* nous limite dans l'interprétation de ces résultats, des sous-groupes pouvant exister chez ces espèces de *Bbsl* sans que nous ayons pu les mettre en évidence.

L'arbre phylogénétique construit par maximum de parcimonie comportait de nombreuses multifurcations, qui peuvent s'expliquer par la forte proximité génétique de certains groupes de séquences, empêchant l'algorithme de discriminer les *Bbsl* présentes au sein d'un groupe. Le gène séquencé (*flaB*) est un gène de ménage, souvent utilisé pour identifier les différentes espèces de *Bbsl*. Ainsi, ce gène est conservé au sein d'une espèce, ce qui explique la faible diversité génétique retrouvée pour les séquences de *Bbsl* étudiées. De ce fait, il est possible que la diversité génétique observée dans cette étude pour le gène *flaB* ne reflète pas la diversité génétique existant entre les génomes des *Bbsl* étudiées. Il serait donc intéressant de reprendre cette étude en privilégiant cette fois le protocole défini par Margos et al (Margos et al. 2008), et repris par de nombreuses études : ce protocole consiste à réaliser un MLST sur 8 gènes de ménage. Cela permettrait d'obtenir une plus grande variabilité génétique chez les *Bbsl* étudiées. De plus, l'inclusion de plusieurs gènes permet de limiter les biais potentiels introduits par des échanges de matériel génétique entre les *Bbsl*. Ces échanges sont

supposés rares pour les gènes de ménage, mais peuvent tout de même survenir, et fausser les analyses phylogénétiques (car deux bactéries issues d'une même lignée pourront présenter une grande distance génétique si l'une d'entre elles a par exemple échangé une partie d'un gène avec une autre bactérie).

Les *Bbsl* formant des clades sur l'arbre phylogénétique ne provenaient pas forcément de larves issues d'une même espèce d'oiseau. Une analyse poussée du lien entre diversité génétique et espèce aviaire hôte n'était pas possible du fait de la prédominance du Merle noir par rapport aux autres espèces aviaires. On a pu observer que les *Bbsl* de certains clades étaient majoritairement retrouvées chez les Merles noirs, alors que d'autres clades étaient plus diversifiés en termes d'espèces aviaires. Cependant, le nombre d'oiseaux présents dans ces analyses était limité (54), et les 56 séquences génétiques de *Bbsl* issues des Merles noirs provenaient de larves gorgées collectées sur 26 oiseaux. Ces données sont insuffisantes pour caractériser la diversité génétique présente chez les *Bbsl* circulant chez le Merle noir ou chez d'autres espèces en France.

La majorité de Merles noirs dans nos résultats peut s'expliquer d'abord par l'abondance de cette espèce en France : c'est l'une des espèces les plus capturées dans les stations STOC en 2019-2020 (130 captures), derrière le Rougegorge familier (174 individus capturés) et la Fauvette à tête noire (145 individus capturés). Ensuite, en plus de son abondance naturelle, le Merle noir est connu pour être un des réservoirs principaux de *B. garinii* (Humair et al. 1998; Taragel'ová et al. 2008). Ce n'est pas le cas du Rougegorge familier, ce qui explique que peu de larves positives aient été collectées sur cette espèce, malgré le fait que c'est l'une des principales espèces capturées et que sa charge en tique est importante. Sous réserve que ces résultats soient confirmés par des études complémentaires, la présence de clades pour lesquels les séquences génétiques de *Bbsl* sont en majorité issues de larves collectées sur des Merles noirs, et de clades pour lesquels on retrouve une plus grande diversité d'espèces aviaires pose aussi la question de la circulation des *Bbsl* au sein des différentes espèces aviaires. Certaines souches de *Bbsl* pourraient être mieux adaptées aux merles noirs qu'à d'autres espèces aviaires, par exemple en cas de différences génétiques au niveau des protéines de surface de ces bactéries, protéines qui conditionnent l'interaction de la bactérie avec l'hôte (Oppler et al. 2021). En effet, les différences génétiques entre les espèces de *Bbsl* sont le résultat de leur adaptation à leur environnement, et ces différences pourraient exister entre les lignées d'une même espèce, certaines lignées étant plus adaptées à certains hôtes. La circulation préférentielle de certaines souches de *Bbsl* chez le Merle noir pourrait aussi s'expliquer par l'importance du Merle noir dans le nourrissage des tiques (forte infestation) (Klaus et al. 2016; Klitgaard et al. 2019) et en tant que réservoir pour *B. garinii* : dans un milieu donné où le Merle noir serait l'une des espèces comptant le plus d'individus, une part importante des tiques (principalement les larves et les nymphes) infectées se nourriraient sur les Merles noirs, et ceux-ci infecteraient alors de nouvelles tiques qui seront à leur tour plus susceptibles d'infecter un Merle noir plutôt qu'un autre oiseau. Cela pourrait contribuer à maintenir une communauté de *B. garinii* génétiquement proches au sein des Merles noirs.

Les résultats obtenus lors de l'étude du lien entre diversité génétique de *B. garinii* et espèce aviaire hôte sont concordants avec la littérature : la seule étude disponible à notre connaissance ne met pas en évidence de regroupement de *B. garinii* en fonction de l'espèce d'oiseau dont les larves sont issues (Norte et al. 2020). Ces résultats sont toutefois à confirmer, le nombre d'oiseaux inclus dans l'étude de Norte et al étant faible. Cependant, des études suggèrent que les spécificités d'hôtes (vecteurs et hôtes vertébrés) jouent un rôle important dans la diversité génétique des *Bbsl*, pouvant conduire à de la spéciation. Une étude portant sur ce sujet a été réalisée en 2009 (Margos et al. 2009). Dans cette étude, une MLSA a été réalisée sur 8 gènes de ménage à partir d'une centaine de souches de *Bbsl*, provenant majoritairement de pays d'Europe. Les analyses ainsi que l'arbre phylogénétique obtenus ont permis de définir *B. bavariensis* en tant que nouvelle espèce, ainsi que de faire l'hypothèse que *B. garinii* serait une espèce plus récente que *B. bavariensis*, apparue suite à une adaptation aux oiseaux. Cependant, en 2016, une étude a suggéré que *B. bavariensis* serait l'espèce la plus récente des deux : grâce aux techniques de séquençage complet du génome, cette étude a pu inférer un arbre phylogénétique à partir d'une centaine de gènes orthologues issus de 111 souches de *Bbsl* provenant de différents pays du monde (Becker et al. 2016). Cet arbre suggère que la spéciation ayant conduit à la séparation de *B. bavariensis* et de *B. garinii* fait probablement suite à une adaptation à un nouvel hôte : *B. bavariensis* se serait différenciée de *B. garinii* suite à un changement d'hôte préférentiel en faveur des rongeurs. L'étude de Becker et al. met aussi en évidence deux populations distinctes de *B. bavariensis* : une, hétérogène, située en Asie, et l'autre, plus homogène, en Europe. Cette étude suggère que c'est l'adaptation d'une population de bactéries à un nouveau vecteur (*I. ricinus*) qui a permis à celle-ci d'envahir l'Europe à partir de l'Asie (Becker et al. 2016). Il est donc possible que, même au sein d'une

espèce de *Bbsl*, certains clades circulent préférentiellement chez différentes espèces (de vecteurs ou d'hôtes).

Il peut être pertinent de réaliser un test statistique afin de vérifier si les pourcentages de Merles noirs au sein des clades où il était majoritaire par rapport aux clades où une diversité d'espèce aviaire était présente sont significativement différents. Cela pourrait fournir un argument supplémentaire en faveur d'une circulation préférentielle de certaines souches de *Bbsl* chez le Merle noir. Cependant, au vu du faible nombre de données disponibles au sein de chaque clade, il est probable que notre étude manque de puissance statistique afin de mettre en évidence cette différence (si elle existe). Pour pouvoir explorer plus en détail l'existence d'un lien entre la diversité génétique de *B. garinii* et l'espèce aviaire hôte, il serait intéressant de reprendre le protocole de capture des oiseaux avec collecte de tiques sur une plus longue période, mais en demandant à chaque bagueur de ne collecter des tiques que chez les espèces d'oiseaux dont le rôle de réservoir est suspecté ou prouvé pour *B. garinii*, et qui sont capturables dans les stations STOC (par exemple le Merle noir, la Grive musicienne, la Mésange charbonnière, le Pipit des arbres et le Troglodyte mignon (Rataud 2022)). Si la circulation de souches de *Bbsl* plus adaptées aux Merles noirs est confirmée, il pourrait être intéressant d'étudier les différences dans le génome des *Bbsl* issues des clades regroupant une majorité de Merles noirs avec celles issues des autres espèces, afin de mettre en évidence d'éventuelles modifications des gènes régissant l'interaction hôte/bactérie.

Les *Bbsl* formant des clades sur l'arbre phylogénétique ne partageaient pas forcément le même milieu. L'analyse du lien entre diversité génétique des *Bbsl* et milieu de capture des oiseaux est aussi délicate, à cause de la prédominance des milieux arbustifs/broussailleux. Cette prédominance s'explique par le fait que les stations STOC sont majoritairement placées dans ces milieux.

Lorsque plus de deux séquences de *Bbsl* ont été isolées sur des larves provenant d'un même oiseau, leur positionnement sur l'arbre phylogénétique suggère une proximité génétique chez certaines de ces séquences. Cela pourrait être dû au fait que les *Bbsl* infectant les larves de l'oiseau sont toutes les descendantes des *Bbsl* ayant infecté l'oiseau en premier lieu. Cependant, certaines *B. garinii* détectées chez les larves issues d'un même oiseau ne partagent pas cette forte proximité génétique : ces différences génétiques pourraient notamment s'expliquer si certaines larves se sont infectées par co-repas, avec une souche de *B. garinii* provenant d'une autre tique se nourrissant sur le même oiseau.

Les *Bbsl* issues des clades identifiés à partir de l'arbre phylogénétique se répartissaient sur tout le territoire français : aucune structure spatiale pour *B. garinii* n'a pu être mise en évidence par observation visuelle, les *Bbsl* provenant de différents clades étaient retrouvées dans des mêmes stations. Ces résultats sont cohérents avec les études précédemment réalisées sur *B. garinii* : hors échelle supérieure au continent, les *B. garinii* proches génétiquement ne sont pas regroupées sur un même territoire (Vollmer et al. 2013; Mtierová et al. 2020; Norte et al. 2020). En effet, les oiseaux sont des hôtes mobiles pouvant se déplacer sur de grandes distances. Cependant, la présence de zones étendues sans station (donc sans possibilité de collecter des données) nous empêche d'avoir une vision complète de la répartition spatiale de ces bactéries. De plus, l'analyse de la structure spatiale de *B. garinii* a été uniquement visuelle. Il serait pertinent de compléter cette analyse visuelle par une analyse statistique de détection de cluster pour confirmer l'absence de regroupement spatial des bactéries génétiquement proches.

Notons que la méthode de construction d'un arbre phylogénétique par maximum de parcimonie tend à donner des arbres moins robustes et précis (dont la topologie peut varier) par rapport aux méthodes du maximum de vraisemblance et de l'inférence bayésienne (Hall 2017). Cela est notamment dû au fait que le maximum de parcimonie n'utilise pas de modèle de substitution explicite, ce qui peut mener à une sous-estimation des distances génétiques entre les séquences. De plus, la méthode du maximum de parcimonie ne prend pas en compte l'évolution parallèle des séquences (homoplasie : deux séquences seront proches car elles auront évolué de la même façon, sans toutefois appartenir à la même lignée). Elle ne prend pas non plus en compte l'évolution rétrograde des séquences (nucléotide d'une séquence qui muterait de A à G puis de nouveau en A) (Nei et Kumar 2000). Cependant, c'est une méthode rapide et qui est adaptée à notre jeu de données : en effet, cette méthode est adaptée à l'étude de gènes dont le taux de mutation est faible, ce qui est le cas pour les gènes de ménage. Au vu de la grande proximité génétique de nos séquences, l'utilisation d'autres méthodes phylogénétiques aurait probablement donné des résultats similaires à ceux obtenus.

Le séquençage de l'ADN contenu dans les larves gorgées ne permet pas de détecter d'éventuelles co-infections à *Bbsl*. Cependant, ces co-infections sont présumées rares, les larves étant généralement

non infectées par *BbsI* avant leur premier repas sanguin (transmission par co-repas et transovarienne faibles).

Enfin, l'utilisation du gène *flaB* pour l'identification des espèces de *BbsI* ne permet pas de distinguer *B. garinii* de *B. bavariensis* (deux espèces très proches génétiquement). Cependant, le nombre de *B. bavariensis* au sein de notre étude devrait être nul ou négligeable, car *B. bavariensis* a pour réservoirs principaux les rongeurs. La mise en œuvre du protocole par MLST permettrait de distinguer ces deux espèces, et l'inclusion des rongeurs dans notre étude permettrait d'étudier la diversité génétique de ces deux espèces.

## Conclusion :

Cette étude a permis de réaliser une première description de la diversité génétique et de la répartition spatiale des *B. garinii* détectées chez des larves gorgées collectées chez des oiseaux communs en France. Les séquences des bactéries échantillonnées étaient génétiquement proches, et présentes dans la majorité des zones où des données ont pu être collectées. Les séquences génétiquement proches n'étaient pas regroupées dans une même zone géographique, et des séquences issues de *BbsI* appartenant à différents clades étaient retrouvées dans des mêmes stations : aucune structure spatiale n'a été observée. L'arbre phylogénétique construit par maximum de parcimonie n'a pas mis en évidence de regroupement strict de séquences génétiques de *B. garinii* en fonction de l'espèce d'oiseau dont étaient issues les larves infectées ou de son milieu de vie. Cependant, certains clades de *BbsI* pourraient circuler majoritairement chez le Merle noir, tandis que d'autres pourraient être associés à une plus grande diversité d'espèces aviaires (ces résultats étant à confirmer par d'autres études).

Cependant, l'utilisation du fragment d'un seul gène (*flaB*) limite l'étude de la variabilité génétique et de la phylogénie de *B. garinii*. Il serait donc pertinent d'obtenir des séquences génétiques en suivant un protocole permettant d'obtenir des fragments de plusieurs gènes de ménages, ce qui pourrait permettre de mettre en évidence une variabilité génétique plus importante que celle retrouvée dans cette étude, d'améliorer la robustesse de l'arbre phylogénétique construit ainsi que de faciliter la comparaison avec d'autres études ayant utilisé ce protocole. Elargir le spectre d'hôte de notre étude et inclure les rongeurs permettrait aussi de comparer la proximité génétique de *B. garinii* et de *B. bavariensis* (sous réserve d'utiliser un protocole permettant de différencier ces deux espèces).

## Références bibliographiques :

- Aslam, Bilal, Iftikhar Hussain, Muhammad Mahmood, Sajjad Rahman, et Abu Siddique. 2012. « Flagellin gene based phylogenetic analysis of Pakistani strain of *Borrelia anserina* isolated from Argas ticks ». *African journal of microbiology research* 6 (avril): 3214-21. <https://doi.org/10.5897/AJMR11.1591>.
- Aucott, John, Candis Morrison, Beatriz Munoz, Peter C. Rowe, Alison Schwarzwald, et Sheila K. West. 2009. « Diagnostic Challenges of Early Lyme Disease: Lessons from a Community Case Series ». *BMC Infectious Diseases* 9 (juin): 79. <https://doi.org/10.1186/1471-2334-9-79>.
- Baum, David. 2008. « Reading a Phylogenetic Tree: The Meaning of Monophyletic Groups ». *Nature Education*, 2008. <http://www.nature.com/scitable/topicpage/reading-a-phylogenetic-tree-the-meaning-of-41956>.
- Becker, Noémie, Gabriele Margos, Helmut Blum, Stefan Krebs, Alexander Graf, Robert Lane, Santiago Castillo-Ramírez, Andreas Sing, et Volker Fingerle. 2016. « Recurrent evolution of host and vector association in bacteria of the *Borrelia burgdorferi* sensu lato species complex ». *BMC Genomics* 17 (septembre): 734. <https://doi.org/10.1186/s12864-016-3016-4>.
- Brown, Terence A. 2002. *Genomes*. Wiley-Liss. <https://www.ncbi.nlm.nih.gov/books/NBK21128/>.
- Cardenas-de la Garza, Jesus Alberto, Estephania De la Cruz-Valadez, Jorge Ocampo-Candiani, et Oliverio Welsh. 2019. « Clinical Spectrum of Lyme Disease ». *European Journal of Clinical Microbiology & Infectious Diseases* 38 (2): 201-8. <https://doi.org/10.1007/s10096-018-3417-1>.
- Cerar, Tjasa, Franc Strle, Dasa Stupica, Eva Ruzic-Sabljić, Gail McHugh, Allen C. Steere, et Klemen Strle. 2016. « Differences in Genotype, Clinical Features, and Inflammatory Potential of *Borrelia burgdorferi* Sensu Stricto Strains from Europe and the United States ». *Emerging Infectious Diseases* 22 (5): 818-27. <https://doi.org/10.3201/eid2205.151806>.
- Chao, Li-Lian, et Chien-Ming Shih. 2009. « Genetic Analysis of the Flagellin Gene of Lyme Disease Spirochetes (*Borrelia burgdorferi*) Isolated from Rodents in Taiwan ». *J Med Sci* 29 (octobre): 249-55.
- Coipan, E. Claudia, Setareh Jahfari, Manoj Fonville, G. Anneke Oei, Lodewijk Spanjaard, Katsuhisa Takumi, Joppe W. R. Hovius, et Hein Sprong. 2016. « Imbalanced Presence of *Borrelia burgdorferi* s.l. Multilocus Sequence Types in Clinical Manifestations of Lyme Borreliosis ». *Infection, Genetics and Evolution: Journal of Molecular Epidemiology and Evolutionary Genetics in Infectious Diseases* 42 (août): 66-76. <https://doi.org/10.1016/j.meegid.2016.04.019>.
- Comstedt, Pär, Loreta Asokliene, Ingvar Eliasson, Björn Olsen, Anders Wallensten, Jonas Bunikis, et Sven Bergström. 2009. « Complex Population Structure of Lyme Borreliosis Group Spirochete *Borrelia garinii* in Subarctic Eurasia ». *PloS One* 4 (6): e5841. <https://doi.org/10.1371/journal.pone.0005841>.
- Derdáková, Markéta, et Daniela Lencáková. 2005. « Association of Genetic Variability within the *Borrelia burgdorferi* Sensu Lato with the Ecology, Epidemiology of Lyme Borreliosis in Europe ». *Annals of Agricultural and Environmental Medicine: AAEM* 12 (2): 165-72.
- Dieter, Heylen. 2016. « 7. Ecological interactions between songbirds, ticks, and *Borrelia burgdorferi* s.l. in Europe ». In , 91-101. [https://doi.org/10.3920/978-90-8686-838-4\\_7](https://doi.org/10.3920/978-90-8686-838-4_7).
- Dieter, Heylen, Erik Matthysen, Manoj Fonville, et Hein Sprong. 2013. « Songbirds as general transmitters but selective amplifiers of *Borrelia burgdorferi* sensu lato genotypes in Ixodes ricinus ticks ». *Environmental microbiology* 16 (octobre). <https://doi.org/10.1111/1462-2920.12304>.
- Drummond, Alexei J., et Remco R. Bouckaert. 2015. *Bayesian Evolutionary Analysis with BEAST*. Cambridge University Press.
- Dubska, Lenka, Ivan Literak, Elena Kocianova, Veronika Taragelova, et Oldrich Sychra. 2009. « Differential Role of Passerine Birds in Distribution of *Borrelia* Spirochetes, Based on Data from Ticks Collected from Birds during the Postbreeding Migration Period in Central Europe ». *Applied and Environmental Microbiology* 75 (3): 596-602. <https://doi.org/10.1128/AEM.01674-08>.
- Estrada-Peña, Agustín, Sally Cutler, Aleksandar Potkonjak, Muriel Vassier-Tussaut, Wim Van Bortel, Hervé Zeller, Natalia Fernández-Ruiz, et Andrei Daniel Mihalca. 2018. « An Updated Meta-Analysis of the Distribution and Prevalence of *Borrelia burgdorferi* s.l. in Ticks in Europe ». *International Journal of Health Geographics* 17 (1): 41. <https://doi.org/10.1186/s12942-018-0163-7>.
- Estrada-Peña, Agustín, Andrei Mihalca, et Trevor Petney. 2017. *Ticks of Europe and North Africa: A Guide to Species Identification*. <https://doi.org/10.1007/978-3-319-63760-0>.

- Fukunaga, MASAHITO, KEIJI Okada, MINORU Nakao, TATSUYA Konishi, et YUKITA Sato. 1996. « Phylogenetic Analysis of *Borrelia* Species Based on Flagellin Gene Sequences and Its Application for Molecular Typing of Lyme Disease *Borreliae* ». *International Journal of Systematic and Evolutionary Microbiology* 46 (4): 898-905. <https://doi.org/10.1099/00207713-46-4-898>.
- Garcia-Monco, Juan Carlos, et Jorge L. Benach. 2019. « Lyme Neuroborreliosis: Clinical Outcomes, Controversy, Pathogenesis, and Polymicrobial Infections ». *Annals of Neurology* 85 (1): 21-31. <https://doi.org/10.1002/ana.25389>.
- Glaeser, Stefanie P., et Peter Kämpfer. 2015. « Multilocus Sequence Analysis (MLSA) in Prokaryotic Taxonomy ». *Systematic and Applied Microbiology* 38 (4): 237-45. <https://doi.org/10.1016/j.syapm.2015.03.007>.
- Goëffon, Adrien, Jean-Michel Richer, et Jin-Kao Hao. 2010. « Heuristic Methods for Phylogenetic Reconstruction with Maximum Parsimony ». In *Algorithms in Computational Molecular Biology: Techniques, Approaches and Applications*, 579-97. <https://doi.org/10.1002/9780470892107.ch26>.
- Gómez-Díaz, Elena, Thierry Boulinier, Natacha Sertour, Muriel Cornet, Elisabeth Ferquel, et Karen D. McCoy. 2011. « Genetic Structure of Marine *Borrelia* *Garinii* and Population Admixture with the Terrestrial Cycle of Lyme Borreliosis ». *Environmental Microbiology* 13 (9): 2453-67. <https://doi.org/10.1111/j.1462-2920.2011.02515.x>.
- Gray, J., O. Kahl, R. S. Lane, et G. Stanek. 2002. *Lyme Borreliosis: Biology, Epidemiology, and Control*. CABI.
- Gray, Jeremy, Olaf Kahl, et Annetta Zintl. 2021. « What Do We Still Need to Know about *Ixodes Ricinus*? » *Ticks and Tick-Borne Diseases* 12 (3): 101682. <https://doi.org/10.1016/j.ttbdis.2021.101682>.
- Gryczyńska, Alicja, et Renata Welc-Fałęciak. 2016. « Long-Term Study of the Prevalence of *Borrelia Burgdorferi* s.l. Infection in Ticks (*Ixodes Ricinus*) Feeding on Blackbirds (*Turdus Merula*) in NE Poland ». *Experimental & Applied Acarology* 70 (3): 381-94. <https://doi.org/10.1007/s10493-016-0082-x>.
- Hall, Barry G. 2017. *Phylogenetic Trees Made Easy: A How-To Manual*. Sinauer.
- Hamming, R. W. 1950. « Error Detecting and Error Correcting Codes ». *Bell System Technical Journal* 29 (2): 147-60. <https://doi.org/10.1002/j.1538-7305.1950.tb00463.x>.
- Hanincova, Klara, Priyanka Mukherjee, Nicholas H. Ogden, Gabriele Margos, Gary P. Wormser, Kurt D. Reed, Jennifer K. Meece, Mary F. Vandermause, et Ira Schwartz. 2013. « Multilocus Sequence Typing of *Borrelia Burgdorferi* Suggests Existence of Lineages with Differential Pathogenic Properties in Humans ». *PloS One* 8 (9): e73066. <https://doi.org/10.1371/journal.pone.0073066>.
- Haydon, Daniel T., Sarah Cleaveland, Louise H. Taylor, et M. Karen Laurenson. 2002. « Identifying Reservoirs of Infection: A Conceptual and Practical Challenge ». *Emerging Infectious Diseases* 8 (12): 1468-73. <https://doi.org/10.3201/eid0812.010317>.
- Heylen, Dieter, Aleksandra Krawczyk, Isabel Lopes de Carvalho, Maria Sofia Nuncio, Hein Sprong, et Ana Cláudia Norte. 2017. « Bridging of Cryptic *Borrelia* Cycles in European Songbirds ». *Environmental Microbiology* 19 (5): 1857-67. <https://doi.org/10.1111/1462-2920.13685>.
- Heylen, Dieter, Hein Sprong, Kees van Oers, Manoj Fonville, Herwig Leirs, et Erik Matthysen. 2014. « Are the Specialized Bird Ticks, *Ixodes Arboricola* and *I. Frontalis*, Competent Vectors for *Borrelia Burgdorferi* Senu Lato? » *Environmental Microbiology* 16 (4): 1081-89. <https://doi.org/10.1111/1462-2920.12332>.
- Heylen, Dieter, Ellen Tijssse, Manoj Fonville, Erik Matthysen, et Hein Sprong. 2013. « Transmission Dynamics of *Borrelia Burgdorferi* s.l. in a Bird Tick Community ». *Environmental Microbiology* 15 (2): 663-73. <https://doi.org/10.1111/1462-2920.12059>.
- Hofmeester, T. R., E. C. Coipan, S. E. van Wieren, H. H. T. Prins, W. Takken, et H. Sprong. 2016. « Few Vertebrate Species Dominate the *Borrelia Burgdorferi* s.l. Life Cycle ». *Environmental Research Letters* 11 (4): 043001. <https://doi.org/10.1088/1748-9326/11/4/043001>.
- Humair, P. F., D. Postic, R. Wallich, et L. Gern. 1998. « An Avian Reservoir (*Turdus Merula*) of the Lyme Borreliosis Spirochetes ». *Zentralblatt Fur Bakteriologie: International Journal of Medical Microbiology* 287 (4): 521-38.
- Jacquot, Maude, Mathieu Gonnet, Elisabeth Ferquel, David Abrial, Alexandre Claude, Patrick Gasqui, Valérie Choumet, et al. 2014. « Comparative Population Genomics of the *Borrelia Burgdorferi* Species Complex Reveals High Degree of Genetic Isolation among Species and Underscores Benefits and Constraints to Studying Intra-Specific Epidemiological Processes ». *PloS One* 9 (4): e94384. <https://doi.org/10.1371/journal.pone.0094384>.

- Jahfari, Setareh, Aleksandra Krawczyk, E. Claudia Coipan, Manoj Fonville, Joppe W. Hovius, Hein Sprong, et Katsuhisa Takumi. 2017. « Enzootic Origins for Clinical Manifestations of Lyme Borreliosis ». *Infection, Genetics and Evolution: Journal of Molecular Epidemiology and Evolutionary Genetics in Infectious Diseases* 49 (avril): 48-54. <https://doi.org/10.1016/j.meegid.2016.12.030>.
- Jongejan, F., et G. Uilenberg. 2004. « The Global Importance of Ticks ». *Parasitology* 129 (S1): S3-14. <https://doi.org/10.1017/S0031182004005967>.
- Kapli, Paschalia, Ziheng Yang, et Maximilian J. Telford. 2020. « Phylogenetic Tree Building in the Genomic Age ». *Nature Reviews Genetics* 21 (7): 428-44. <https://doi.org/10.1038/s41576-020-0233-0>.
- Klaus, Christine, Jörn Gethmann, Bernd Hoffmann, Ute Ziegler, Martin Heller, et Martin Beer. 2016. « Tick Infestation in Birds and Prevalence of Pathogens in Ticks Collected from Different Places in Germany ». *Parasitology Research* 115 (7): 2729-40. <https://doi.org/10.1007/s00436-016-5022-5>.
- Klitgaard, Kirstine, Jesper Højgaard, Anastasia Isbrand, Jesper J. Madsen, Kasper Thorup, et Rene Bødker. 2019. « Screening for Multiple Tick-Borne Pathogens in Ixodes Ricinus Ticks from Birds in Denmark during Spring and Autumn Migration Seasons ». *Ticks and Tick-Borne Diseases* 10 (3): 546-52. <https://doi.org/10.1016/j.ttbdis.2019.01.007>.
- Kurtenbach, Klaus, Klára Hanincová, Jean I. Tsao, Gabriele Margos, Durland Fish, et Nicholas H. Ogden. 2006. « Fundamental Processes in the Evolutionary Ecology of Lyme Borreliosis ». *Nature Reviews. Microbiology* 4 (9): 660-69. <https://doi.org/10.1038/nrmicro1475>.
- Lejal, Emilie. 2020. « La dynamique du Pathobiome des tiques, l'exemple d'Ixodes ricinus ».
- Mac, Stephen, Sara R. da Silva, et Beate Sander. 2019. « The Economic Burden of Lyme Disease and the Cost-Effectiveness of Lyme Disease Interventions: A Scoping Review ». *PloS One* 14 (1): e0210280. <https://doi.org/10.1371/journal.pone.0210280>.
- Madison-Antenucci, Susan, Laura D. Kramer, Linda L. Gebhardt, et Elizabeth Kauffman. 2020. « Emerging Tick-Borne Diseases ». *Clinical Microbiology Reviews* 33 (2): e00083-18. <https://doi.org/10.1128/CMR.00083-18>.
- Margos, Gabriele, Anne G. Gatewood, David M. Aanensen, Klára Hanincová, Darya Terekhova, Stephanie A. Vollmer, Muriel Cornet, et al. 2008. « MLST of Housekeeping Genes Captures Geographic Population Structure and Suggests a European Origin of *Borrelia burgdorferi* ». *Proceedings of the National Academy of Sciences of the United States of America* 105 (25): 8730-35. <https://doi.org/10.1073/pnas.0800323105>.
- Margos, Gabriele, Stephanie A. Vollmer, Nicholas H. Ogden, et Durland Fish. 2011. « Population Genetics, Taxonomy, Phylogeny and Evolution of *Borrelia burgdorferi* Sensu Lato ». *Infection, Genetics and Evolution: Journal of Molecular Epidemiology and Evolutionary Genetics in Infectious Diseases* 11 (7): 1545-63. <https://doi.org/10.1016/j.meegid.2011.07.022>.
- Margos, Gabriele, Stephanie Vollmer, Muriel Cornet, Martine Garnier, Volker Fingerle, Bettina Wilske, Antra Bormane, et al. 2009. « A New *Borrelia* Species Defined by Multilocus Sequence Analysis of Housekeeping Genes ». *Applied and environmental microbiology* 75 (juillet): 5410-16. <https://doi.org/10.1128/AEM.00116-09>.
- Marsot, M., P.-Y. Henry, G. Vourc'h, P. Gasqui, E. Ferquel, J. Laignel, M. Grysan, et J.-L. Chapuis. 2012. « Which Forest Bird Species Are the Main Hosts of the Tick, *Ixodes ricinus*, the Vector of *Borrelia burgdorferi* Sensu Lato, during the Breeding Season? ». *International Journal for Parasitology* 42 (8): 781-88. <https://doi.org/10.1016/j.ijpara.2012.05.010>.
- Mccoy, Karen, et Nathalie Boulanger. 2015. *Tiques et maladies à tiques : biologie, écologie évolutive, épidémiologie*.
- Michalik, Jerzy, Beata Wodecka, Maciej Skoracki, Bożena Sikora, et Joanna Stańczak. 2008. « Prevalence of avian-associated *Borrelia burgdorferi* s.l. genospecies in *Ixodes ricinus* ticks collected from Blackbirds (*Turdus merula*) and Song Thrushes (*T. philomelos*) ». *International Journal of Medical Microbiology* 298 (septembre): 129-38. <https://doi.org/10.1016/j.ijmm.2008.03.004>.
- Mtierová, Zuzana, Markéta Derdáková, Michal Chvostáč, Yuliya M. Didyk, Barbara Mangová, Veronika Rusňáková Taragelová, Diana Selyemová, Alžbeta Šujanová, et Radovan Václav. 2020. « Local Population Structure and Seasonal Variability of *Borrelia garinii* Genotypes in *Ixodes ricinus* Ticks, Slovakia ». *International Journal of Environmental Research and Public Health* 17 (10): 3607. <https://doi.org/10.3390/ijerph17103607>.
- Mysterud, Atle, Dieter J. A. Heylen, Erik Matthysen, Aïda Lopez Garcia, Solveig Jore, et Hildegunn Viljugrein. 2019. « Lyme Neuroborreliosis and Bird Populations in Northern Europe ».

- Nei, Masatoshi, et Sudhir Kumar. 2000. *Molecular Evolution and Phylogenetics*. Oxford University Press, New York.
- Nordberg, Cecilie Lerche, Jacob Bodilsen, Fredrikke Christie Knudtzen, Merete Storgaard, Christian Brandt, Lothar Wiese, Birgitte Rønde Hansen, et al. 2020. « Lyme Neuroborreliosis in Adults: A Nationwide Prospective Cohort Study ». *Ticks and Tick-Borne Diseases* 11 (4): 101411. <https://doi.org/10.1016/j.ttbdis.2020.101411>.
- Norte, Ana Cláudia, Gabriele Margos, Noémie S. Becker, Jaime Albino Ramos, Maria Sofia Nuncio, Volker Fingerle, Pedro Miguel Araújo, et al. 2020. « Host Dispersal Shapes the Population Structure of a Tick-Borne Bacterial Pathogen ». *Molecular Ecology* 29 (3): 485-501. <https://doi.org/10.1111/mec.15336>.
- Oppler, Zachary J., Kayleigh R. O’Keeffe, Karen D. McCoy, et Dustin Brisson. 2021. « Evolutionary Genetics of Borrelia ». *Current Issues in Molecular Biology* 42: 97-112. <https://doi.org/10.21775/cimb.042.097>.
- Pérez-Eid, Claudine. 2007. *Les tiques : identification, biologie, importance médicale et vétérinaire*. Lavoisier.
- Postic, Danièle, Martine Garnier, et Guy Baranton. 2007. « Multilocus Sequence Analysis of Atypical Borrelia Burgdorferi Sensu Lato Isolates – Description of Borrelia Californiensis Sp. Nov., and Genomespecies 1 and 2 ». *International Journal of Medical Microbiology* 297 (4): 263-71. <https://doi.org/10.1016/j.ijmm.2007.01.006>.
- Qiu, Wei-Gang, et Che L. Martin. 2014. « Evolutionary Genomics of Borrelia Burgdorferi Sensu Lato: Findings, Hypotheses, and the Rise of Hybrids ». *Infection, Genetics and Evolution: Journal of Molecular Epidemiology and Evolutionary Genetics in Infectious Diseases* 27 (octobre): 576-93. <https://doi.org/10.1016/j.meegid.2014.03.025>.
- Rataud, Amalia. 2022. « Eco-épidémiologie du système Oiseaux-Tiques-Borrelia ».
- Rataud, Amalia, Clemence Galon, Laure Bournez, Pierre-Yves Henry, Maud Marsot, et Sara Moutailler. 2022. « Diversity of Tick-Borne Pathogens in Tick Larvae Feeding on Breeding Birds in France ». *Pathogens* 11 (8): 946. <https://doi.org/10.3390/pathogens11080946>.
- Rataud, Amalia, Pierre-Yves Henry, Sara Moutailler, et Maud Marsot. 2022. « Research Effort on Birds’ Reservoir Host Potential for Lyme Borreliosis: A Systematic Review and Perspectives ». *Transboundary and Emerging Diseases* 69 (5): 2512-22. <https://doi.org/10.1111/tbed.14305>.
- Richter, Dania, Rainer Allgöwer, et Franz-Rainer Matuschka. 2002. « Co-Feeding Transmission and Its Contribution to the Perpetuation of the Lyme Disease Spirochete Borrelia Afzelii ». *Emerging Infectious Diseases* 8 (12): 1421-25. <https://doi.org/10.3201/eid0812.010519>.
- Richter, Dania, Alina Debski, Zdenek Hubalek, et Franz-Rainer Matuschka. 2012. « Absence of Lyme Disease Spirochetes in Larval Ixodes Ricinus Ticks ». *Vector Borne and Zoonotic Diseases (Larchmont, N. Y.)* 12 (1): 21-27. <https://doi.org/10.1089/vbz.2011.0668>.
- Riley, Lee W., et Ronald E. Blanton. 2018. « Advances in Molecular Epidemiology of Infectious Diseases: Definitions, Approaches, and Scope of the Field ». *Microbiology Spectrum* 6 (6). <https://doi.org/10.1128/microbiolspec.AME-0001-2018>.
- Rudenko, Natasha, Marina Golovchenko, Natalia Belfiore, Libor Grubhoffer, et James Oliver. 2014. « Divergence of Borrelia burgdorferi sensu lato spirochetes could be driven by the host: Diversity of Borrelia strains isolated from ticks feeding on a single bird ». *Parasites & vectors* 7 (janvier): 4. <https://doi.org/10.1186/1756-3305-7-4>.
- Sabitova, Yuliya, Nataliya Fomenko, Artem Tikunov, Oleg Stronin, Maxim Khasnatinov, Davaajav Abmed, Galina Danchinova, Irina Golovljova, et Nina Tikunova. 2018. « Multilocus Sequence Analysis of Borrelia Burgdorferi Sensu Lato Isolates from Western Siberia, Russia and Northern Mongolia ». *Infection, Genetics and Evolution* 62 (août): 160-69. <https://doi.org/10.1016/j.meegid.2018.04.015>.
- Schwartz, Ira, Gabriele Margos, Sherwood R. Casjens, Wei-Gang Qiu, et Christian H. Eggers. 2021. « Multipartite Genome of Lyme Disease Borrelia: Structure, Variation and Prophages ». *Current Issues in Molecular Biology* 42: 409-54. <https://doi.org/10.21775/cimb.042.409>.
- Seifert, Stephanie N., Camilo E. Khatchikian, Wei Zhou, et Dustin Brisson. 2015. « Evolution and Population Genomics of the Lyme Borreliosis Pathogen, Borrelia Burgdorferi ». *Trends in Genetics: TIG* 31 (4): 201-7. <https://doi.org/10.1016/j.tig.2015.02.006>.
- Stackebrandt, Erko, Wilhelm Frederiksen, George M. Garrity, Patrick A. D. Grimont, Peter Kämpfer, Martin C. J. Maiden, Xavier Nesme, et al. 2002. « Report of the Ad Hoc Committee for the Re-Evaluation of the Species Definition in Bacteriology ». *International Journal of Systematic and Evolutionary Microbiology* 52 (Pt 3): 1043-47. <https://doi.org/10.1099/00207713-52-3-1043>.

- Stanek, Gerold, Gary P. Wormser, Jeremy Gray, et Franc Strle. 2012. « Lyme Borreliosis ». *The Lancet* 379 (9814): 461-73. [https://doi.org/10.1016/S0140-6736\(11\)60103-7](https://doi.org/10.1016/S0140-6736(11)60103-7).
- Takano, Ai, Minoru Nakao, Toshiyuki Masuzawa, Nobuhiro Takada, Yasuhiro Yano, Fubito Ishiguro, Hiromi Fujita, et al. 2011. « Multilocus Sequence Typing Implicates Rodents as the Main Reservoir Host of Human-Pathogenic *Borrelia garinii* in Japan ». *Journal of Clinical Microbiology* 49 (5): 2035-39. <https://doi.org/10.1128/JCM.02544-10>.
- Taragel'ová, Veronika, Juraj Koci, Klára Hanincová, Klaus Kurtenbach, Marketa Derdakova, Nick Ogden, Ivan Literák, Elena Kocianová, et Milan Labuda. 2008. « Blackbirds and Song Thrushes Constitute a Key Reservoir of *Borrelia garinii*, the Causative Agent of Borreliosis in Central Europe ». *Applied and environmental microbiology* 74 (mars): 1289-93. <https://doi.org/10.1128/AEM.01060-07>.
- Telford, S. R., T. N. Mather, S. I. Moore, M. L. Wilson, et A. Spielman. 1988. « Incompetence of Deer as Reservoirs of the Lyme Disease Spirochete ». *The American Journal of Tropical Medicine and Hygiene* 39 (1): 105-9. <https://doi.org/10.4269/ajtmh.1988.39.105>.
- « theses.fr – Amalia Rataud, Eco-épidémiologie du système Oiseaux-Tiques-Borrelia ». s. d. Consulté le 31 janvier 2023. <https://www.theses.fr/s228218>.
- Tyler, Shaun, Shari Tyson, Antonia Dibernardo, Michael Drebot, Edward Feil, Morag Graham, Natalie Knox, et al. 2018. « Whole genome sequencing and phylogenetic analysis of strains of the agent of Lyme disease *Borrelia burgdorferi* from Canadian emergence zones ». *Scientific Reports* 8 (juillet). <https://doi.org/10.1038/s41598-018-28908-7>.
- Vollmer, Stephanie A., Antra Bormane, Ruth E. Dinnis, Frederik Seelig, Andrew D. M. Dobson, David M. Aanensen, Marianne C. James, et al. 2011. « Host Migration Impacts on the Phylogeography of Lyme Borreliosis Spirochaete Species in Europe ». *Environmental Microbiology* 13 (1): 184-92. <https://doi.org/10.1111/j.1462-2920.2010.02319.x>.
- Vollmer, Stephanie A., Edward J. Feil, Chen-Yi Chu, Stephen L. Raper, Wu-Chun Cao, Klaus Kurtenbach, et Gabriele Margos. 2013. « Spatial Spread and Demographic Expansion of Lyme Borreliosis Spirochaetes in Eurasia ». *Infection, Genetics and Evolution: Journal of Molecular Epidemiology and Evolutionary Genetics in Infectious Diseases* 14 (mars): 147-55. <https://doi.org/10.1016/j.meegid.2012.11.014>.
- Voordouw, Maarten J. 2015. « Co-Feeding Transmission in Lyme Disease Pathogens ». *Parasitology* 142 (2): 290-302. <https://doi.org/10.1017/S0031182014001486>.
- Wolcott, Katherine A., Gabriele Margos, Volker Fingerle, et Noémie S. Becker. 2021. « Host Association of *Borrelia burgdorferi* Sensu Lato: A Review ». *Ticks and Tick-Borne Diseases* 12 (5): 101766. <https://doi.org/10.1016/j.ttbdis.2021.101766>.
- Woolhouse, M. E., C. Dye, J. F. Etard, T. Smith, J. D. Charlwood, G. P. Garnett, P. Hagan, et al. 1997. « Heterogeneities in the Transmission of Infectious Agents: Implications for the Design of Control Programs ». *Proceedings of the National Academy of Sciences of the United States of America* 94 (1): 338-42. <https://doi.org/10.1073/pnas.94.1.338>.
- Xia, Xuhua. 2020a. « Maximum Parsimony Method in Phylogenetics ». In , 185-208. <https://doi.org/10.1201/9780429425875-6>.
- . 2020b. « Nucleotide Substitution Models and Evolutionary Distances ». In , 69-139. <https://doi.org/10.1201/9780429425875-3>.
- Yang, Ziheng, et Bruce Rannala. 2012. « Molecular Phylogenetics: Principles and Practice ». *Nature Reviews. Genetics* 13 (5): 303-14. <https://doi.org/10.1038/nrg3186>.
- Zhang, Kai, Zhuan Qin, Yunjie Chang, Jun Liu, Michael G. Malkowski, Saimtun Shipa, Li Li, Weigang Qiu, Jing-Ren Zhang, et Chunhao Li. 2019. « Analysis of a flagellar filament cap mutant reveals that HtrA serine protease degrades unfolded flagellin protein in the periplasm of *Borrelia burgdorferi* ». *Molecular microbiology* 111 (6): 1652-70. <https://doi.org/10.1111/mmi.14243>.

## Liste des annexes :

Annexe 1 : Espèces d'agents pathogènes et gènes ciblés par la puce microfluidique.....	47
Annexe 2 : Amorces utilisées dans les protocoles de PCR de confirmation des espèces de <i>Borrelia</i> .	48
Annexe 3a : Schéma simplifié du fonctionnement de l'algorithme Muscle, utilisé pour aligner l'ensemble des séquences et obtenir un msa (multiple sequence alignment) .....	48
Annexe 3b : Alignement des séquences grâce à l'arbre guide, suite à la création de la matrice des distances permettant de déterminer le degré de proximité des différentes séquences.....	48
Annexe 4 : Carte présentant les pourcentages de larves positives à <i>B. garinii</i> en fonction des stations en France.....	50
Annexe 5 : Espèces d'oiseaux positives à <i>B. garinii</i> , et nombre d'individus porteurs de larves négatives et porteurs d'au moins une larve positive à <i>B. garinii</i> par espèce d'oiseau .....	51
Annexe 6: Nombre de de larves positives et négatives à <i>B. garinii</i> par oiseau en fonction de l'espèce .....	52

## Annexes :

Annexe 1 : Espèces d'agents pathogènes et gènes ciblés par la puce microfluidique

<b>Espèce d'agent pathogène ciblée</b>	<b>Gène ciblé</b>
<i>Borrelia burgdorferi</i> sensu stricto	<i>rpoB</i>
<i>Borrelia garinii</i>	<i>rpoB</i>
<i>Borrelia afzelii</i>	<i>fla</i>
<i>Borrelia valaisiana</i>	<i>ospA</i>
<i>Borrelia lusitaniae</i>	<i>rpoB</i>
<i>Borrelia spielmanii</i>	<i>fla</i>
<i>Borrelia bissettii</i>	<i>rpoB</i>
<i>Borrelia miyamotoi</i>	<i>glpQ</i>
<i>Borrelia mayonii</i>	<i>fla</i>
<i>Borrelia bavariensis</i>	<i>pyrG</i>
<i>Borrelia</i> spp.	23S
<i>Anaplasma marginale</i>	<i>msp1</i>
<i>Anaplasma platys</i>	<i>groEL</i>
<i>Anaplasma phagocytophilum</i>	<i>msp2</i>
<i>Anaplasma ovis</i>	<i>msp4</i>
<i>Anaplasma centrale</i>	<i>groEL</i>
<i>Anaplasma bovis</i>	<i>groEL</i>
<i>Anaplasma</i> spp.	16S
<i>Ehrlichia</i> spp.	16S
<i>Ehrlichia canis</i>	<i>gltA</i>
<i>Neoehrlichia mikurensis</i>	<i>groEL</i>
<i>Rickettsia conorii</i>	ITS
<i>Rickettsia slovaca</i>	ITS
<i>Rickettsia massiliae</i>	ITS
<i>Rickettsia helvetica</i>	ITS
<i>Rickettsia aeschlimannii</i>	ITS
<i>Rickettsia felis</i>	<i>orfB</i>
<i>Rickettsia</i> spp.	<i>gltA</i>
<i>Bartonella henselae</i>	<i>pap31</i>
<i>Bartonella</i> spp.	<i>ssrA</i>
<i>Francisella tularensis</i>	<i>tul4</i>
<i>Francisella</i> -like endosymbionts	<i>fopA</i>
<i>Coxiella burnettii</i>	IS1111
<i>Coxiella burnettii</i>	<i>icd</i>
<i>Babesia microti</i>	CCTeta
<i>Babesia ovis</i>	18S
<i>Babesia canis</i>	3subspe_18S
<i>Babesia bovis</i>	CCTeta
<i>Babesia caballi</i>	<i>rap1</i>
<i>Babesia venatorum</i>	18S
<i>Babesia divergens</i>	<i>hsp70</i>
<i>Theileria</i> spp.	18S
<i>Hepatozoon</i> spp.	18S
<i>Ixodes ricinus</i>	ITS2
<i>Tique</i> spp.	16S
<i>Dermacentor reticulatus</i>	ITS2
<i>Babesia vulpes</i>	COI

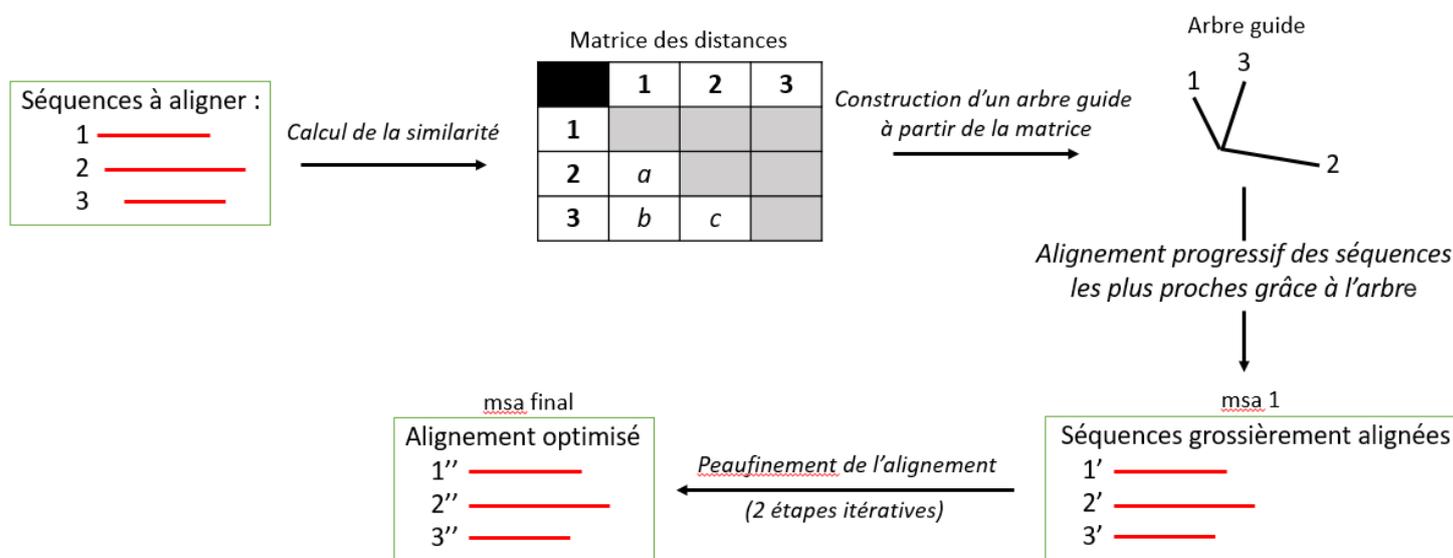
Annexe 2 : Amorces utilisées dans les protocoles de PCR de confirmation des espèces de *Borrelia*

Genre d'agent pathogène	Gène cible	Nom de l'amorce	Séquence (5'-3')	Taille de l'amplicon (pb)	T	Référence
<i>Borrelia</i> spp.	FlaB	FlaB280F	GCAGTTCARTCAGGTAACGG	645	55	[199]
		FlaRL	GCAATCATAGCCATTGCAGATTGT			
		flaB_737F	GCATCAACTGTRGTTGTAACATTAACAGG	407	59	
		FlaLL	ACATATTCAGATGCAGACAGAGGT			

Légende : F = Forward, R = Reverse, pb = paire de base, T = Température d'hybridation.

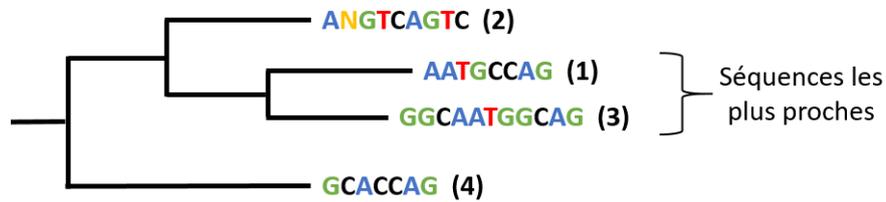
Annexe 3a : Schéma simplifié du fonctionnement de l'algorithme Muscle, utilisé pour aligner l'ensemble des séquences et obtenir un msa (multiple sequence alignment)

1, 2, et 3 sont les séquences à aligner, et a, b et c, les distances respectives entre ces séquences. L'arbre guide est créé par la méthode de l'UPGMA, qui regroupe ensemble les séquences les plus proches grâce à la matrice des distances.

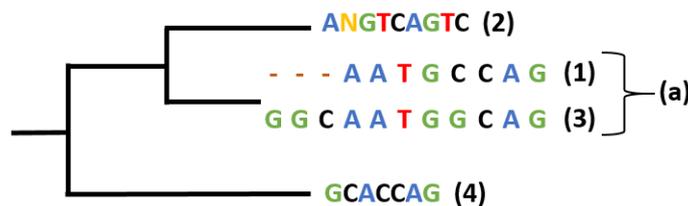


Annexe 3b : Alignement des séquences grâce à l'arbre guide, suite à la création de la matrice des distances permettant de déterminer le degré de proximité des différentes séquences

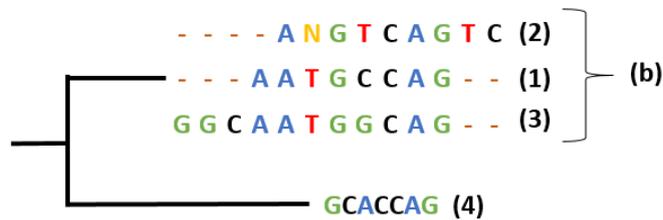
(1), (2), (3) et (4) sont les noms des 4 taxa (séquences génétiques) étudiés. (a) et (b) sont les noms donnés aux ensembles de séquences alignées, (a) et (b) constituent des nouveaux groupes qui vont pouvoir s'aligner avec les autres séquences en fonction de leur distance. La lettre N en jaune représente un nucléotide non informatif (peut être A, T, G ou C). Les tirets en marron représentent des lacunes dans l'alignement, ces lacunes correspondent soit à des insertions/délétions dans la séquence concernée, soit à des introductions artificielles (en début et/ou fin de séquence), afin que les séquences aient toutes la même longueur pour permettre la comparaison.



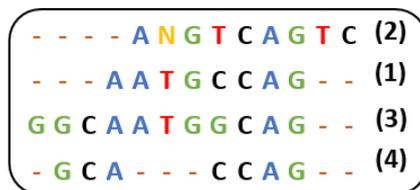
Alignement de (1) et (3)



Alignement de (2) et (a)

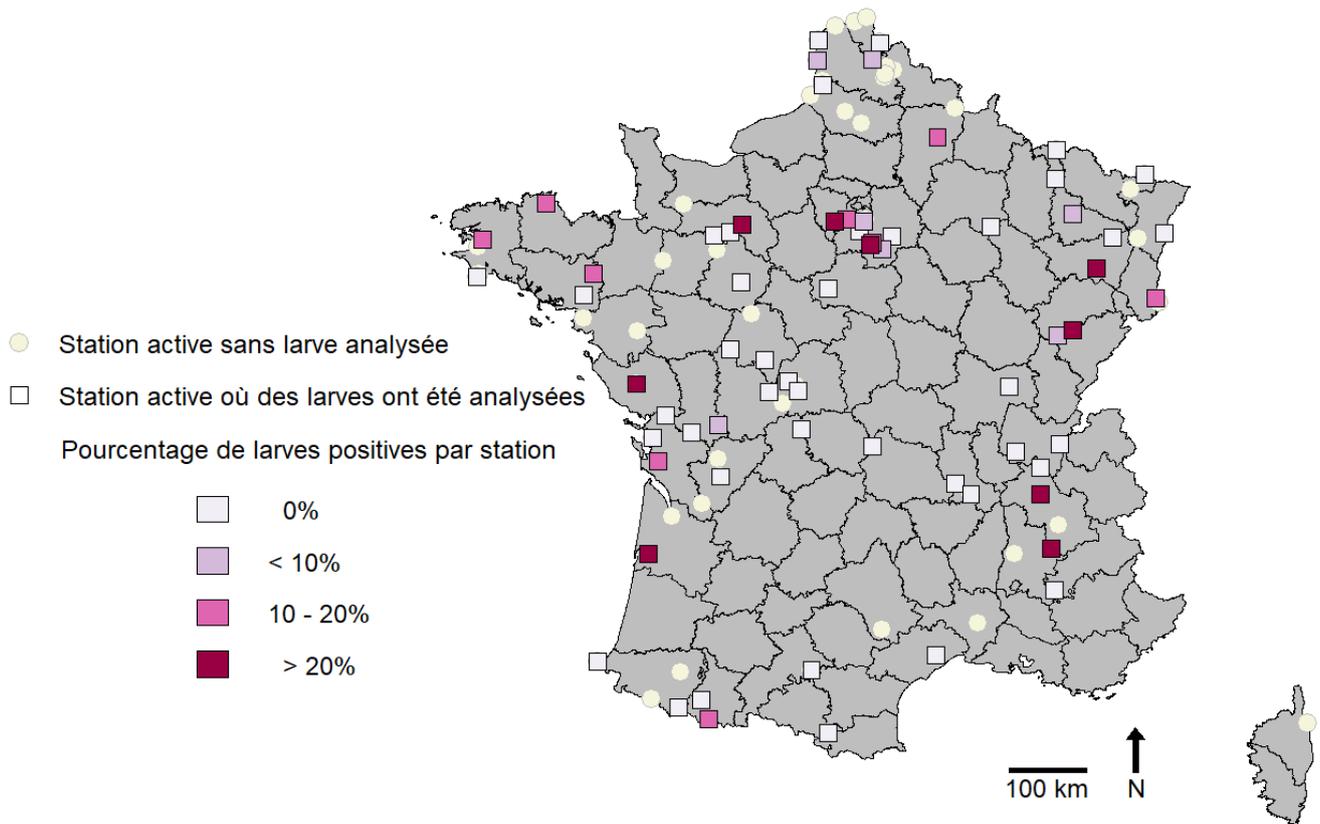


Alignement de (4) et (b)



msa (multiple sequence alignment)

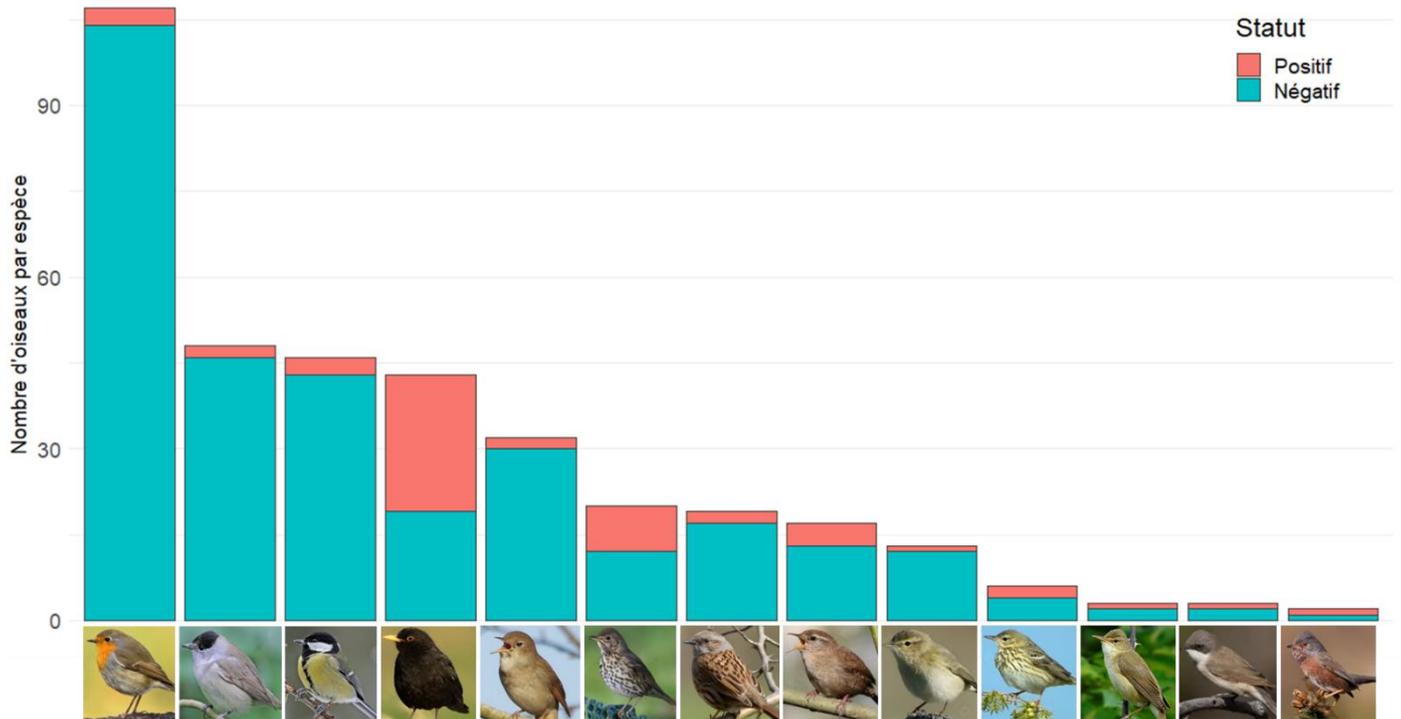
Annexe 4 : Carte présentant les pourcentages de larves positives à *B. garinii* en fonction des stations en France  
Les stations représentées sont les stations actives (ayant capturé des oiseaux) en 2019-2020.



Annexe 5 : Espèces d'oiseaux positives à *B. garinii*, et nombre d'individus porteurs de larves négatives et porteurs d'au moins une larve positive à *B. garinii* par espèce d'oiseau

Les photographies en bas du graphique correspondent aux espèces d'oiseaux concernées, qui sont, de gauche à droite : Rougegorge familier (*Erithacus rubecula*), Fauvette à tête noire (*Sylvia atricapilla*), Mésange charbonnière (*Parus major*), Merle noir (*Turdus merula*), Rossignol philomèle (*Luscinia megarhynchos*), Grive musicienne (*Turdus philomelos*), Accenteur mouchet (*Prunella modularis*), Troglodyte mignon (*Troglodytes troglodytes*), Pouillot véloce (*Phylloscopus collybita*), Pipit des arbres (*Anthus trivialis*), Rousserolle verderolle (*Acrocephalus palustris*), Fauvette babillarde (*Curruca curruca*), Fauvette pitchou (*Curruca undata*).

Les photographies sont toutes issues du site [oiseaux.net](http://oiseaux.net)



Annexe 6: Nombre de de larves positives et négatives à *B. garinii* par oiseau en fonction de l'espèce

Les oiseaux représentés ici sont les oiseaux pour lesquels au moins deux larves ont été analysées, et dont au moins une des larves était positive à *B. garinii* (n = 37). La catégorie « Autres espèces », inclue les individus issus des espèces présentées dans l'annexe précédente (Annexe 5), à l'exception du Merle noir et de la Grive musicienne.

