MASTER THESIS

Anja Jensen Hervik

BI309F Master thesis in Marine Ecology

Next Generation Sequencing Analyses on Norwegian isolates of the tick *Ixodes ricinus*.

Faculty of Biosciences and Aquaculture

Autumn 2016

NORD University

www.nord.no

Preface

This thesis is a part of the degree “Master of science in Marine Ecology” at the Faculty of Biosciences and Aquaculture, Nord University, Bodø, Norway.

First of all, I would like to thank my supervisors Jarle Tryti Nordeide and Steinar Johansen for making this project a possibility, and for their creative thinking, problem solving and wonderful guidance throughout this project. Even though things fell through sometimes, you still believed in the project and me, for that I am eternally grateful.

I am very grateful to Tor Erik Jørgensen for guiding me through the many days of lab work and providing more than useful insights into the world of bioinformatics.
Employees at the Faculty of Biosciences and Aquaculture also deserve a thank you for always providing helpful input and a smile when needed.

I would like to thank to the project “Flått i Nord” and Dag Hvidsten for giving us tips for a sampling location in Brønnøysund, and Snorre Stuen for providing us with a sampling location in Sandnes and aid in the sampling process.

To my friends, you deserve a basket of thanks for your support and motivational input, you are my pillars. A special thank you to mom, dad and my brothers for providing motivational- and overall loving support throughout this time, you always believed in me.

Bodø, August 2016

Anja Jensen Hervik

Table of contents

Abstract

Ticks are hematophagous ectoparasites capable of carrying and transmitting a wide range of diseases, between a variety of species including *Homo sapiens*. Some of the most known disease causing agents carried by ticks are *Borrelia burgdorferi*, *Anaplasma phagocytophilum* and the tick-bone encephalitis complex of viruses (TBE). In addition, *Ixodes ricinus* harbours a recently discovered and possibly infectious endosymbiont, *Midochloria mitochondrii*. A potential further expansion in the distribution of several tick species in Norway is worrying scientists. Currently there is no knowledge regarding the genetic structure of *I. ricinus* in Norway and limited genetic/genomic research have been performed on ticks in this country. Investigating the genetic structure could aid in understanding some of the dispersal patterns of both ticks and tick-borne diseases.

In this study, IonTorrent PGM next generation sequencing technology was used to sequence fragments of the mitochondrial DNA of geographic isolates of the tick species *I. ricinus*. The ticks originated from four locations in Norway, covering a span of 1238 km; Sandnes, Sogn, Levanger and Brønnøysund. In addition, total DNA of one tick specimen was sequenced using the IonTorrent PGM platform. Here, 8 million reads of an average length of 280 bp were generated, corresponding to 2.2 billion nucleotides. These high quality sequences were applied in two different analyses. First, the complete mitochondrial genome of *I. ricinus* was mined assembled and analysed, and further applied in a phylogenetic analysis of *Ixodes* species. Second, the total DNA sequencing data set was used to investigate the presence of bacterial genome sequences corresponding to *B. burgdorferi*, *A. phagocytophilum and M. mitochondrii*.

The resulting phylogenetic analysis was rather inconclusive, one clade contained 3/5 of samples from one geographical location, and all samples originating from Sogn placed within another, clade (although spread within subclades) which could indicate a stronger genetic structure in Brønnøysund and Sogn. However, this does not necessarily indicate genetic structure due to geographic isolation, it could also be due to differences in transport of ticks by birds. Nonetheless, this result presents a need for further studies.

Phylogenetic analysis of *Ixodes* species, including the complete *I. ricinus* mitochondrial genome mined in this study, supported previous reports conducted on the phylogeny of *Ixodes*, in which a division between Australasian and non-Australasian ticks was also apparent. Interestingly, genome sequences from *B. burgdorferi* (or a closely related species), along with the mitochondrial endosymbiont *M. mitochondrii*, were readily detected from the total DNA data set. This work is to our knowledge the first study assessing the tick *I. ricinus* genome by next generation sequencing technologies.

Next generation sequencing analysis will contribute to a better understanding of the genetic structure of ticks in Norway. Furthermore, the information is highly important in trying to understand the dispersal potential and epidemiological questions of ticks and tick-borne diseases.

1 Introduction

1.1 Ticks in general

Ticks are arachnids (class Arachnida) that live as obligate, hematophagous ectoparasites, which are distributed worldwide (Lu et al., 2013, Cakic et al., 2014, Sonenshine and Roe, 2014). These parasites feed on every class of vertebrates, except from fish (Nava et al., 2009). An ectoparasitic animal resides on the exterior of its host, from which it feeds and exploits (Shao and Barker, 2007), as the tick does when it attaches to the skin of an animal and feeds on its blood. In addition to being parasites, ticks are vectors capable of transmitting a large variety of pathogens; these include bacteria (e.g. rickettsia and spirochetes), viruses and protozoa. Many of these pathogens can cause serious diseases and all are thought to be zoonoses (Sonenshine and Roe, 2014), which means that they are transmittable from other animals to humans. Ticks are thus of great concern, not only because they affect humans, but also husbandry animals and livestock (Jongejan and Uilenberg, 2004). In addition to disease transmission, large tick infestations can in the worst-case scenario cause anaemia and infections in heavily parasitized animals (Sonenshine and Roe, 2014).

Ticks have been present for a long time, and it is believed that ticks and their pathogens stretch as far back as the ancient Egyptians and Greeks (de la Fuente et al., 2015). Based on fossil records, ticks are thought to originate around the pre-middle cretaceous (Nava et al., 2009). Some of the oldest tick fossils are larva, corresponding to *Carios jerseyi*, dated to 90-94 Mya (Klompen and Grimaldi, 2001), and a hard tick larva of the cretaceous fossil genus *Cormpalpatum*, most likely from 100 Mya (Poinar and Brown, 2003, Nava et al., 2009). In the present day, there are many different species of ticks, each with different habitat/host preferences. A list of valid tick names was created by (Guglielmone et al., 2010) where 896 species of ticks divided into three families were recognized.

1.2 Classification, phylogeny and tick species in Norway

Ticks belong to the phylum Arthropoda, the class Arachnida and the lineage chelicerates (Casati et al., 2008, Nava et al., 2009) they are further categorized in the subclass Acari, and in the suborder Ixodida in the order Parasitiformes (Black and Roehrdanz, 1998). Ixodida contains

three families; Argasidae (soft ticks), Ixodidae (hard ticks) and Nuttalliellidae (Guglielmone et al., 2010). The family Ixodidae (the hard ticks), contains 12 genera. It is divided into two groups; the prostriata (containing only the genus *Ixodes*) and the metastriata (which contains the remaining 11 genera) (Sonenshine and Roe, 2014). This family of ticks is arguably the most important, both in regards to veterinary and medical importance (Barker and Murrell, 2004, Liu et al., 2013). *Ixodes* is the largest genus of hard ticks, containing 243 species (Guglielmone et al., 2010). Within this genus, there is a complex of closely related tick species called the *Ixodes ricinus* species complex, which was defined by Keirans et al. (1999) (cited in Xu et al., 2003). This contains 14 species that are distributed all over the world, and the best known are *I. scapularis*, *I. persulcatus*, *I. ricinus* and *I. pacificus* (Leger et al., 2013). The two species that are most significant in Palearctic and Oriental regions are *I. persulcatus* and *I. ricinus* (Xu et al., 2003, Jongejan and Uilenberg, 2004).

The phylogeny of the three tick families remained unresolved as of 2008, however this is mostly due to the fact that the Nuttalliellidae (containing one species) had not been collected and investigated for years (Barker and Murrell, 2008). Whether or not *Ixodes* is a paraphyletic or monophyletic group is still being discussed (Barker and Murrell, 2008), but one general working hypothesis surrounding the genus, is that there is a division between Australasian *Ixodes* and other *Ixodes*, that the genus *Ixodes* is monophyletic, and that the most closely related genus of ticks to the *Ixodes* genus is *Bothriocroton* (Barker and Murrel, 2004, 2008). Australasian Ixodes are called so due to the presumption that they are endemic to, or that they have evolved in Australasia (Barker and Murrel, 2004). The division between the two *Ixodes* types is due to the fact that Australasian *Ixodes* species (e.g. *I. uriae* and *I. holocyclus*) contain two control regions (Shao et al., 2005, Barker and Murrell, 2008), whilst other *Ixodes* (e.g. *I. ricinus*, *I. hexagonus* and *I. pavlovskyi*) possess the mitochondrial genome composition of the chelicerates, which is found in the living fossil *Limulus polyphemus* (Staton et al., 1997, Black and Roehrdanz, 1998).

Both Australasian ticks and non-Australasian ticks have been observed in Norway (Table 1), eleven species of ticks in total. These ticks have different preferences regarding their hosts and not all are endemic with stable populations. *I. ricinus* is the tick that is most known to feed on humans, which is due to its non-nidiculuos nature (see section below).

Table 1. Tick species observed in Norway (Mehl, 1983, Mehl and Braathen, 1987). The summary is based on observations from 1983 and present day may deviate from this.

Species	"Main" hosts	Presence in Norway
I. ricinus	Varies in host use	Very common
I. uriae	Seabirds	Very common
I. hexagonus	Carnivores/hedgehogs	Rare
I. trianguliceps	Rodents	Very common
Argas vespertilionis	Bats	Very common
I. arboricola	Birds	Rare
Rhipicephalus sanguineus	Dogs	Very rare
I. caledonicus	Birds	Rare
I. lividus	Birds	Rare
I. frontalis	Birds	Rare
Hyalomma marginatum	Birds/mammals	Rare

Figure 1. a) Photograph of an adult female tick of the species *Ixodes ricinus*. (By H. Krisp – personal work, CC BY 3.0, https://commons.wikimedia.org/w/index.php?curid=16316859) b) Photograph of a male specimen of the species *I. ricinus* (www.influentialpoints.com).

1.3 Ixodes ricinus

Ixodes ricinus is a hard tick with a flat body shape (Figure 1) and a protective dorsal plate that covers the entire back of males and partly for females (Mehl and Braathen, 1987). It has common

names such as the sheep tick, castor bean tick and “skogflått”. This tick is a non-nidiculuos tick, which means that it resides in the open environment, as opposed to nidiculuos ticks that live and acquire hosts in nests or caves. Thus *I. ricinus* is highly susceptible to abiotic factors when moulting, searching for prey or questing for a meal (Randolph, 2014). According to Gray (1991), *I. ricinus* is said to require a relative humidity of approximately 80% or above to be able to survive and reproduce, which means that this species favours areas where there is moderate to high rainfall with a decent vegetation cover (cited in Gray et al., 2009). However, this species is also a survivalist, and can survive in unsuitable environments. It has a high plasticity in regards to ecology and is distributed across the western Palaearctic. It favours deciduous woodland and mixed forests, where it can survive in different ranges of environmental and temperature conditions (Fournier et al., 2000, Porretta et al., 2013).

The sheep tick is one of the hard ticks that employ a three-stage life cycle with three active life stages (not including the egg stage, which is inactive): larvae, nymphs and adults. One blood meal, from a vertebrate host, is required for each of these life stages (Gilbert et al., 2014). *I. ricinus* has an indiscriminate feeding behaviour, meaning that it is not a host-specific species (Jongejan and Uilenberg, 2004). For instance, in Norway it has been found to feed on 50 different species of mammals, reptiles and birds (Mehl, 1983, Mehl and Braathen 1987). However, this species has often shown to be feeding on birds and mammals, including humans (Casati et al., 2008). *I. ricinus’* choice of hosts and affinity to humans makes its role as a vector highly important, especially in relation to human health (Jongejan and Uilenberg, 2004). As the tick also feed on a variety of birds, it is often transported by migratory species both from within Norway and from other countries (Mehl et al., 1984).

1.4 Tick-borne pathogens and a peculiar endosymbiont

In Europe, *I. ricinus* is the main vector for several disease causing agents of both medical and economical importance (Dinnis et al., 2014, Gilbert et al., 2014), the dominant tick in the temperate areas (Rumer et al., 2011) and is said to be responsible for 90-95% of tick bites that occur (Süss et al., 2008). These agents include the *Borrelia burgdorferi* sensu lato complex of bacteria, Anaplasma, Rickettsia and Babesia species, and the tick-borne encephalitis (TBE) complex of viruses (Gilbert et al., 2014, Tomkins et al., 2014). Lyme disease, anaplasmosis,

babesiosis and ehrlichisosis are examples of diseases transmitted by this tick that are zonooses (Dantas-Torres et al., 2012). Lyme borreliosis is caused by the spirochete *B. burgdorferi* sensu lato. This disease has a wide range of manifestations in humans, some of which include skin changes and neurological symptoms (Ogden et al., 2014). TBE (tick-borne encephalitis), which is a neuroinfection, is regarded as one of the most dangerous of its type in Europe and Asia (Gritsun et al., 2003). The incidence of tick-borne diseases (TBD) has shown to increase throughout the years (Randolph, 2004), and even though some of this increase can be attributed to increased awareness, the increase is still a reality. In addition to known pathogens, *I. ricinus* has more recently been shown to harbour a mitochondrial endosymbiont named *Midichloria mitochondrii* (Sassera et al., 2006). There is evidence that this bacterium is transmitted by tick saliva, but little is known about whether or not it is infectious/pathogenic (Mariconti et al., 2012).

1.5 *Ixodes ricinus* in Norway

Ticks are especially interesting to scientists due to their medical importance (Paulauskas et al., 2006). In recent years, there has been a focus on range expansion by *I. ricinus*, as it is of great concern how vectors, and thus vector-borne pathogens, can respond to upcoming climate change (Gilbert et al., 2014). Due to the high vagility of its hosts, it has been hard to determine the exact distribution of established populations of *I. ricinus* in Norway. The approximate distribution of the tick is from the coastal areas in Østfold (south) and up to Nordland (north) (Kjelland et al., 2010). According to a report from 1983, the tick was present in northern areas of Norway up to 66°N, which was also reported by Tambs-Lyche in 1943 (cited in Mehl, 1983). Seven decades later, Jore et al. (2011) claimed that *I. ricinus* has expanded its range to 69°N, a distance of 400 kilometres. More recent investigations have suggested that the tick is still scarce north of Brønnøy (approximately 66°N), and if there are established populations in these areas the number of ticks are most likely small with a patchy distribution (Hvidsten et al., 2014). In these regions, ticks are believed to originate from migratory birds (Hvidsten et al., 2014, Hvidsten et al., 2015).

As ticks (and their pathogens) are transported by different species of birds, different individuals can originate from a number of locations, both inland and foreign countries. However, the most common migratory routes are from southwest to northeast during spring time and the complete

opposite direction in autumn (Fonstad et al., 2007, Kjelland et al., 2010). In a study performed in 2010, the most commonly infested bird species were ground feeding and corresponding to the species *Turdus* spp., *Anthus trivialis*, *Fringilla coelebs*, *Sylvia* spp., *Prunella modularis* and *Erithacus rubecula* (Kjelland et al., 2010). A projection of possible future changes in climate and the distribution of the castor bean tick suggest that *I. ricinus* distribution in Norway, Sweden and Finland; will encompass the entire countries, up to 70º North (with an exception of the mountainous regions) (Jaenson and Lindgren, 2011). However, the authors claim that even though climate might permit the tick to expand, other factors might still limit its distribution (Jaenson and Lindgren, 2011). Jore et al. (2014) presented a model and according to this, significant predictors of tick expansion include abundance of hosts (red deer and farm animals), bush encroachment/ecotones and several abiotic factors such as snow cover and spring precipitation.

Pathogens carried by ticks depend on two factors; the distribution of the tick and its use of hosts. Predicting how a potential expansion of *I. ricinus* affects disease transmission will depend on knowledge of plasticity in host use (Leger et al., 2013). In addition to investigate disease transmission and all things related, population genetics and genomics are other important topics. Investigating the relation and structure within and among populations of species, such as *I. ricinus*, is highly important, as it will help scientists to understand the epidemiology and evolutionary dynamics of both the vector and disease in question (Casati et al., 2008). For instance, mitogenomic studies on chelicerates are conducted both to investigate their evolutionary biology, and to obtain sequences to use in the identification of species (Montagna et al., 2012).

1.6 Sequencing and mitochondrial genomes

In the early 2000's it was stated that molecular methods were increasingly being applied to acarology (e.g. ticks and mites) (Navajas and Fenton, 2000, Cruickshank, 2002). The improvements in next generation sequencing (NGS) in regards to speed, cost and accuracy, in addition to evolving tools in bioinformatics allow deeper studies in regards to non-model organisms (Helyar et al., 2011), such as ticks. The early knowledge of arachnid mitochondrial genomes retrieved from Sanger sequencing and traditional PCR has yielded detailed and useful

information, and NGS technologies have the possibility to increase the current knowledge surrounding sequence variation, diversity and transcriptional mechanisms. In addition, whole mitochondrial genomes can now be rapidly sequenced and assembled (Brewer et al., 2014).

Mitochondrial genomes have properties that make them suitable for investigating evolutionary questions at taxonomic levels such as species, genera or families. Some of these include maternal inheritance and a faster evolutionary rate than that of the nuclear genome (Brown, 1979, Shao and Barker, 2007). There is a higher proportion of coding sequences in the mitochondrial genome than the nuclear genome, in addition, due to their small size and the fact that mitochondria are abundant in cells, they are easier to sequence and annotate (Shao and Barker, 2007). There are currently 41 complete tick mitochondrial genomes available from the NCBI Organelle Genome Database, but only one version is available for *I. ricinus*, which originated from Italy. The entire mitochondrial genome of *I. ricinus* was reported by Montagna et al. (2012) and is available from this database, the structure of which is shown in Figure 2. Studies that have suggested that mitochondrial sequences are useful to establish closely related species of ticks and thus mitochondrial sequences should be useful in distinguishing within species, for instance the population level (Caporale et al., 1995, Xu et al., 2003, Casati et al., 2008).

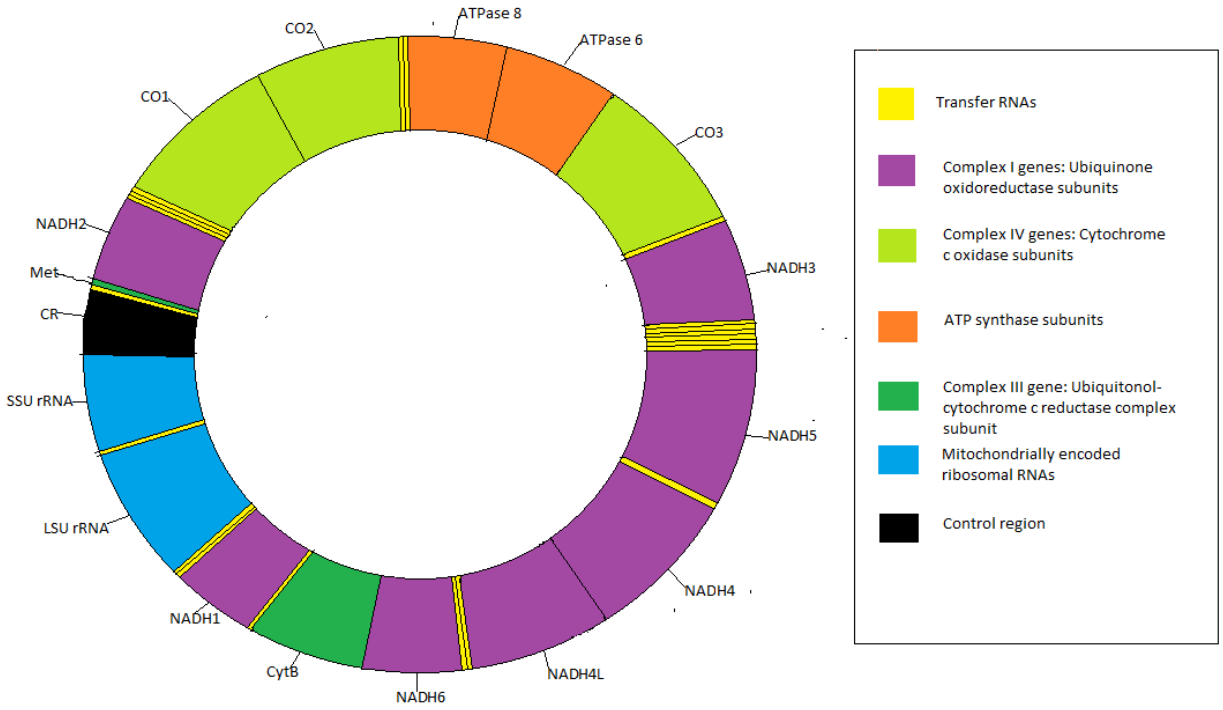


Figure 2. The organization of the mitochondrial genome of *Ixodes ricinus* (Modified from Montagna et al. (2012)).

To assess the genetic difference between individuals or populations is often fundamental within biology. The sequence depth that is now available through high-throughput sequencing (NGS) allows for comparative analyses of genomic sequences and enables detection of polymorphic loci within genomes, such as single nucleotide polymorphisms (SNPs) (Meyer and Hill, 2014). SNPs are changes in a single base pair of a sequence (Uricaru et al., 2015) and SNP's are popular in population genetics due to their abundance, in addition to their ease of scoring (Brookes, 1999). SNPs have many applications and according to Freeland et al. (2011) they are highly suitable for inferring evolutionary relationships. Phylogeny is a method of inferring such relationships and is defined by Xiong (2006) as a "*study of evolutionary relationships between organisms by using treelike diagrams as representations*" or by Freeland et al. (2011) as "*a branching diagram, usually depicted as a tree, that shows the evolutionary relationships among different genetic lineages (e.g. species or genera)*". The general gest in this type of analysis, which is depicted in a tree-like diagram, is that members of the same group (or clade) are more closely related to each other than they are to other such groups in the diagram (Brinkman and Leipe, 2001). This type of analysis can thus help resolve relationships between for instance geographically separated individuals of the same species.

1.7 Previous studies

There have been a couple of studies related to the genetic structure of populations of *I. ricinus* and some using mitochondrial markers as their resource. (Here I use the word "population" based on each authors' own definition of population in relation to their experiment.) Dinnis et al. (2014) discovered a genetic difference between two populations (Latvia and Britain) of *I. ricinus* using multiloci sequence typing of mitochondrial genes, whilst Casati et al. (2008) found no evidence of population structure when investigating five mitochondrial gene fragments from ticks originating from several countries (Switzerland, Italy, Austria, Denmark, Sweden and Finland). Delaye et al. (1997) found little genetic variation between populations within the same country using allozymic data. Noureddine et al. (2011) found no structure within populations from 14 different countries when using a combination of nuclear and mitochondrial markers.

However, none of these studies used phylogenetic analysis of SNPs within the mitochondrial genome to assess genetic variation between and within populations of the same country. Currently, there is little information available on the genetic structure of *I. ricinus*, especially in Norway, where *I. ricinus* is widespread. Thus there is a need for more studies with a focus on genetics and/or genomics to obtain more knowledge.

1.8 Aim and design of study

The aim of this study was to assess genetic variation within and between four different *Ixodes ricinus* populations in Norway: Sandnes in Rogaland, Vik in Sogn og Fjordane, Levanger in Nord-Trøndelag and Brønnøysund in Nordland. The distance from Sandnes to Brønnøysund is roughly 1238 kilometres, which represents the longest distance span between the sampling locations. The experimental approach was to sequence specific fragments of the mitochondrial genome and apply detected SNPs to create phylogenies to compare the different populations. In addition, total DNA of one individual tick from Brønnøysund was high-throughput sequenced at random. The mitochondrial genome of this tick was mined out from the sequence pool and assembled, and used as a reference for the population data, in addition this sequence was also used in a phylogenetic analysis of *Ixodes* species, based on available mitochondrial genomes. As there is currently little information available on the population genetics and genomics of this species in Norway, one could say that it is an important subject to investigate, as it would help in mapping the potential movements and gene flow of ticks in Norway.

Three hypotheses were formed based on background information surrounding the genetic structure of ticks. The first was that there would be genetic uniformity between the geographic isolates in this study due to a constant maintenance of gene flow created by birds. The second was that there would be a genetic difference between the populations due to varying degrees of migratory birds in the different locations. Third, the population in Brønnøysund is at the northern limit of the species *I. ricinus* distribution and hence one could hypothesise that ticks from this northern locality should group together in one clade due to different selection pressure at its northern limit compared to further south. In addition, phylogenetic analyses of *Ixodes* species was carried out to test previously suggested phylogenies, especially the suggested division between Australasian- and non-Australasian *Ixodes*. Moreover, total DNA of one *I. ricinus*

specimen was sequenced to examine if the data was sufficient to mine the mitochondrial genome of *I. ricinus* and to identify the presence (or non-presence) of bacterial genome sequences.

2 Materials and methods

2.1 Sampling

Ticks of the species *Ixodes ricinus* were collected in four different locations in Norway (Figure 3). The locations covered the south, middle and north of the country (Table 1). The sampling location in Levanger was chosen based on input from citizens, whereas the remaining sampling locations were chosen based on known populations of ticks.

The sampling location in Levanger was chosen based on a survey on Facebook. A status update was made where the public was asked to share the status and notify where ticks had been observed. There were several locations mentioned, but the chosen one was the location most frequently mentioned.

Ticks were sampled by the method called flagging and/or dragging (Figure 4). This method involves dragging a piece of cloth attached to a rod, across the vegetation. The cloth was checked approximately every two metres and attached ticks were removed by tweezers. Appropriate security gear was worn by the sampler to avoid tick bites. Ticks were placed in sealed tubes along with moist paper and/or a piece of vegetation to maintain moisture. The tubes were aired out once a day to keep the containers as habitable as possible, as transport to Bodø from the sampling location did not happen immediately. Ticks were supposed to be used in another experiment, which was why they were kept alive. The first experiment fell through, and the ticks were placed on 96% ethanol and stored in a -20°C freezer.

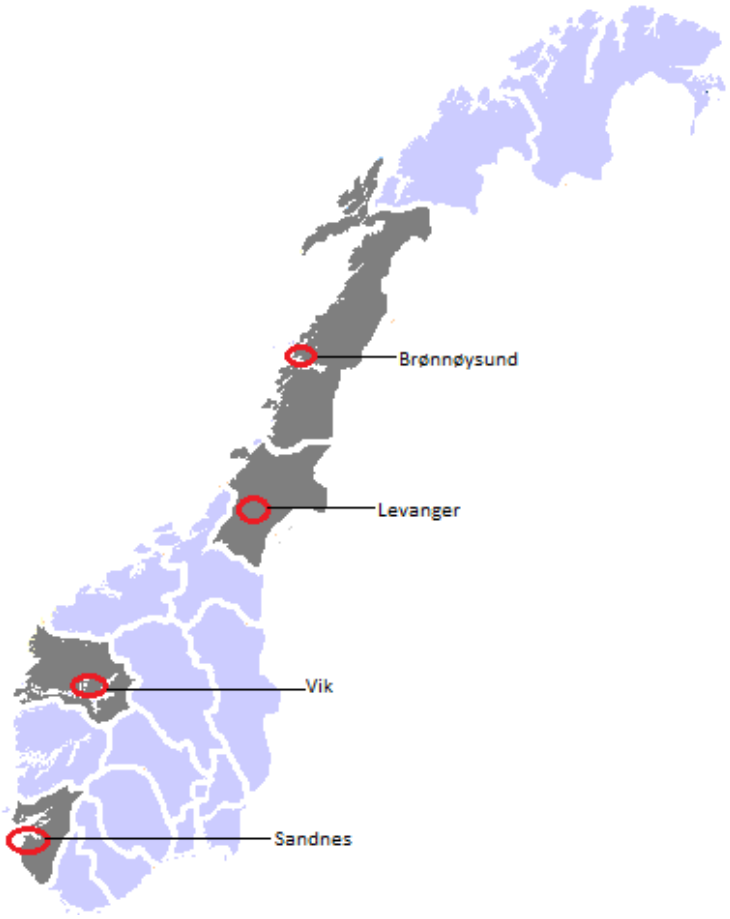


Figure 3. A map of Norway showing the sampling counties (in grey) and the sampling location (red circles).

Table 1. A table showing the sampling location, coordinates and date of sampling.

Location	Coordinates	Date of sampling
Levanger (Åsenfjord), Nord Trøndelag	63°34'28.3"N 10°53'46.6"E	04.08.15
Vik, Sogn	61°5'33.6 "N 6°32'21.6"E	05.08.15
Torghatten, Brønnøysund	65°23'53.8"N 12°5'26.6"E	14.07.15
Sandnes (Kjosavik), Rogaland	58°51'47.5"N 5°54'53.5"E	14.07.15

Figure 4. Photograph of the flagging/dragging sampling method being performed. Duct tape was worn around the arms and feet to protect the sampler. The instrument shown was a common bath towel attached to a wooden stick.

2.2 Nucleic acid isolation and quantification

MasterPure™ Complete DNA Purification Kit (Illumina™) was used to extract DNA from sampled ticks. Whole ticks were transferred to an Eppendorf tube, air-dried and rinsed in Nuclease-Free water. The ticks were divided in two parts by a scalpel and mixed with a solution containing 300µl Tissue and Cell Lysis Solution and 2µl Proteinase K(Illumina™). The samples were incubated on a heat block at 65°C for one hour and turned every fifteen minutes. Purification and precipitation of total DNA then proceeded according to the MasterPure ™ (Illumina™) protocol for tissue samples. Nuclease-free water was added to the final extracted product; 10µl for nymphs, 30 µl for adult ticks. The samples containing the extracted DNA was stored in a -20°C freezer. DNA was extracted from a total of 40 individuals, then from each sampling location. The amount of nymphs vs. adults varied within each population. A list of life stages for each sample is available in the appendix.

Qubit™ dsDNA BR Assay Kit (Invitrogen™) was used to measure the amount of DNA in each sample. Two assay tubes for standard and one assay tube for each sample was prepared. A working solution was prepared by diluting 1µl of Qubit® Reagent in 199µl of Qubit Dilution Buffer for each sample tube. The concentration was measured by a Qubit® Flurometer.

2.3 Polymerase Chain Reaction and Gel electrophoresis

Eight regions of the mitochondrial genome (Table 2; Figure 5) for each sample were amplified by Polymerase Chain Reaction (PCR) following the protocol for Q5™ High-Fidelity DNA Polymerase, New England BioLabs™. A 25 µl reaction was set up for each primer pair, which contained: 1µl DNA, 1.25µl of forward primer, 1.25µl of reverse primer, 15.75µl of Nuclease-free water, 5 µl 5X Q5 Reaction Buffer, 0.5µl 10 mM dNTP, 0.25µl Q5 High-Fidelity DNA polymerase. The PCR was performed on a thermocycler (Thermofisher ™ Applied Biosciences™) were the initial denaturation was set to 98°C for 30 seconds, 35 cycles of 98°C and 10 second denaturation, 30 seconds of primer-pair specific annealing time (Table 4) and 72°C for 2.50 minutes of elongation. The final extension was set at 72° C for 2 minutes. Deviations from these settings included an elongation time of 2 minutes for shorter amplicons (see Table 4).

Table 2. A table showing the final primer pairs used in the experiment. IRM refers to "*Ixodes ricinus* mitochondrial (genome)" and numbers following refers to the base pair position in the mitochondrial genome.

Primer pairs	
1	IRM_13370 + IRM_143
2	IRM_143 + IRM_3067
3	IRM_2959 + IRM_4195
4	IRM_2959 + IRM_5000
5	IRM_6650 + IRM_7168
8	IRM_10308 + IRM_11305
9	IRM_11092 + IRM_11856
10	IRM_11489 + IRM_12153
Additional primer pairs	
12	IRM_143 + IRM_1632
13	IRM_1243 + IRM_3067

Primer pair 2 was hard to amplify in a couple of individuals. Two alternate primer pairs, called 12 and 13, were made, which together corresponded to the length of primer pair two (IRM_143 + IRM_1632 and IRM_1243 + IRM_3067). The additional and original primer sequences are shown in table 3.

Table 3. Direction and sequences of the primers used in this experiment. IRM refers to "*Ixodes ricinus* mitochondrial (genome)" and numbers following refers to the base pair position in the mitochondrial genome.

Primer name	Direction	Sequence
IRM_13370	Forward	5' TGTTACGACTTATCTCACCT 3'
IRM_143	Reverse	5' TTCTAAAGATACCCAAAGTGG 3'
IRM_143	Forward	5' CCACTTTGGGTATCTTTAGAA 3'
IRM_3067	Reverse	5' CATAAGATCAATATCATTGATGTCC 3'
IRM_2959	Forward	5' TGAACAATTATCCCAGCAATTAC 3'
IRM_4195	Reverse	5' TATTAGCAGATAACCGAACTG 3'
IRM_5000	Reverse	5' CCAATGATTACATGGATTCCATGA 3'
IRM_6650	Forward	5' k0TACCTAAATATCGAATATCCTGAA 3'
IRM_7186	Reverse	5' ATCGAGTAGGAGATGTGACAG 3'
IRM_10308	Forward	5' AGTTGATAATAATACACTCAC 3'
IRM_11305	Reverse	5' CGTAGTCCATTTGATTTAACTG 3'
IRM_11092	Forward	5' TATCGATAACGCACTAAAGTAC 3'
IRM_11856	Reverse	5' CTTTGTGTTTTAGTAAGAGTT 3'
IRM_11489	Forward	5' TAACTTGAGCTACACCTCGATA 3'
IRM_12153	Reverse	5' TACATGATCTGAGTTCAGACC 3'
Additional primers	**Direction**	**Sequence**
IRM_1243	Forward	5' TTCCAATGTCTTTATGGTTAGTAG 3'
IRM_1632	Reverse	5' ATTGCTATGTCAACGGAAGC 3'

Table 4. Thermocycling conditions for the different primer pairs. Forward and Reverse primers are referred to as their starting position in the reference mitochondrial genome (NC_018369.2).

Primer pair number	Forward	Reverse	Annealing temperature (Celsius)	Elongation time (minutes)
1	13370	143	55°C	2.50
2	143	3067	55°C	2.50
3	2959	4195	55°C	2.50
4	2959	5000	55°C	2.50
5	6650	7168	54°C	2.50
8	10308	11305	53°C	2.00
9	11092	11856	53°C	2.00
10	11489	12153	56°C	2.50

12	143	1632	52°C	2.00
13	1243	3067	52°C	2.00

A few primer pairs for certain individuals had to be run at different annealing temperatures than the determined settings. These were random primer pairs within the different populations and were run at 52°C and 53°C.

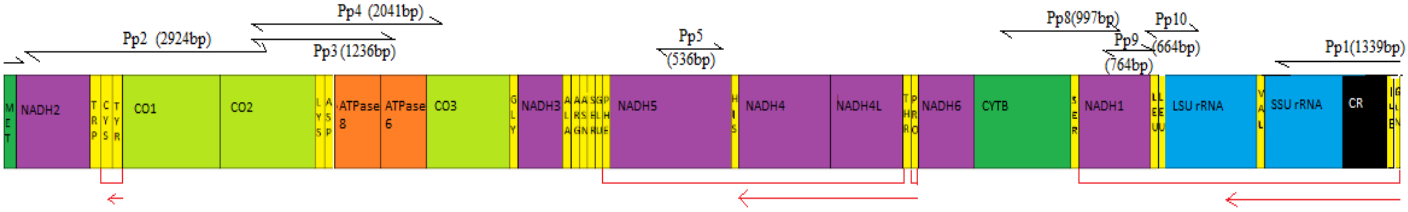


Figure 5. Schematic representation of mitochondrial genome of *Ixodes ricinus* with the approximate location and length of the amplicons shown as black arrows at the top of the figure. The red lines and arrows indicate genes coded for on the non-dominant coding strand. Modified from Montagna et al. (2012).

Verification of PCR amplicons was performed by gel electrophoresis (see appendix, Figure A). 5µl of PCR product was combined with 1µl 6x purple Gel Loading Dye (New England Biolabs ™) and run on agarose gel with 5µl SYBR® Safe (Invitrogen ™). 2-Log DNA ladder (New England BioLabs® Inc.) was run on the agarose gel along with the PCR product to determine the size of the amplicons. The samples were run at 92 Volts for approximately two hours, depending on size of the gel. Gel Logic 200 Imaging System (Kodak™) was used to visualize the results.

2.4 Purification, concentration and quantification of PCR products

PCR products were purified with the PureLink™ Purification Kit (Invitrogen™). 80µl of PureLink™ Binding buffer was added to the PCR product and placed in a PureLink™ Spin Column. The procedure followed the user manual from Invitrogen™. The resulting purified PCR products were quantified by a Qubit® Flurometer. The purified sample was eluted in 50 µl of Nuclease-free water.

The DNA samples were concentrated with the DNA Clean & Concentrator™ -5 (Zymo Research©). About 40 µl of each sample was mixed with 5X the sample volume of DNA

Binding Buffer and transferred to Zymo-Spin™ Column within a Collection Tube, where the samples were pooled together. The purification process followed the User Manual. 20 µl of 1x Low-TE buffer (10 mM Tris-HCl and 0.1 mM EDTA, pH 8) was added to the final product.

2.5 Ion Torrent Total DNA sequencing

Fragmentation of total genomic DNA was performed using the Ion Xpress™ Plus Fragment Library Kit and User Guide (Thermofisher Scientific Inc). The fragmentation was performed with the Covaris™ System S2 Sonicator, using the standard program for 400 bp. This sample, called B9AM from now on, followed the same protocols as the rest, except that it was not concentrated after PCR purification nor amplified before sequencing. Library preparation was performed with Ion Xpress ™ Plus Fragment Library kit and proceeded according to the manual. E-Gel® electrophoresis system was used for size selection of the 400-bp fragmented samples. Agilent D1000ScreenTape System protocol and reagents was used to determine the quality of the fragmented DNA samples. 1 µl of sample was mixed with 3 µl of D1000 Sample Buffer and loaded in the 2200 Tape Station instrument. Emulsion PCR was performed according to the Ion PGM™ OT2Kit User Guide with Ion PGM™ Template OT2 400 Kit (Life Technologies ™). Sequencing of total DNA was performed according to the Ion PGM™ Hi-Q™ Sequencing User Guide. The samples were sequenced on 318 v.2 chips and chip loading followed the user guide above.

2.6 Ion Torrent Sequencing of geographic isolates

Fragmentation of mitochondrial genome fragments was performed using the Ion Xpress™ Plus Fragment Library Kit and User Guide (Thermofisher Scientific Inc). The fragmentation was performed with the Covaris™ System S2 Sonicator, using the standard program for 200 bp + 60 seconds. Library preparation was performed with Ion Xpress ™ Plus Fragment Library kit and proceeded according to the manual. E-Gel® electrophoresis system was used for size selection of the 200-bp fragmented samples. Due to the low yield of the libraries, amplification by PCR had to be performed. Agilent D1000 High Sensitivity ScreenTape System was used to measure the quantity and quality of the amplified libraries. The barcoded libraries were pooled and emulsion PCR was performed according to the Ion PGM™ OT2Kit User Guide with Ion PGM™

Template OT2 200 Kit (Life Technologies ™). Sequencing of the mitochondrial genome fragments was performed according to the Ion PGM™ Hi-Q™ Sequencing User Guide. The samples were sequenced on 318 v.2 chips.

2.7 Bioinformatics

Design of primers from the template mitochondrial genome (Montagna et al., 2012) was performed in BioEdit Sequence Alignment Editor (Hall, 1999). Analysis of next generation sequencing data was performed in CLC Genomics Workbench 9.0.1. A quality control report was created for all sequences (both the population data and the single tick sequencing), followed by quality and adapter trimming. The two data sets were also mapped to the reference mitochondrial genome of *I. ricinus*. The consensus from this mapping was extracted and used as a reference for mapping the population sequence data. A consensus sequence was extracted from each mapping of the population data and saved as a FASTA file. The consensus sequences were aligned manually in BioEdit. Homopolymer errors and the validity of SNPs were checked against the original population data mappings and corrected. The nucleotide sequences were trimmed and used in phylogenetic analyses.

2.7.1 Detection of SNPs, amino acid changes and phylogenetic analysis of geographic isolate data

When analysing the data for SNPs, B9AM was used as a reference against the alignment. The protein coding genes were cut out and analysed in UGENE, from which amino acid changes were investigated. The sequences were translated according to the invertebrate mitochondrial genetic code, which differs slightly from the standard genetic code (Table 5).

Table 5. The differences in the invertebrate mitochondrial genetic code from the standard genetic code. Information from NCBI (http://www.ncbi.nlm.nih.gov/Taxonomy/Utils/wprintgc.cgi#SG5)

	Code 5	**Standard**
AGA	Ser S	Arg R
AGG	Ser S	Arg R
AUA	Met M	Ile I
UGA	Trp W	Ter *

In the phylogenetic analysis of the population data, the reference sequence from NCBI was used as an outgroup, which was trimmed down to be equal to the amplicons represented by the population data. All trimmed nucleotide data obtained from sequencing was used in the phylogenetic analysis. A models test was performed and based on this, Maximum likelihood (ML), maximum parsimony (MP) and neighbour joining (NJ) trees were made from the alignment in MEGA 7 (http://www.megasoftware.net/).

2.7.2 Detection of SNPs and phylogenetic analysis of total DNA sequencing data

The entire sequence was aligned against the reference sequence from NCBI (NC_018369.2) and SNPs were detected. The consensus sequence of *Ixodes ricinus* was then aligned to *I. hexagons* (NC_002010.1), *I. uriae* (NC_006078.1), *I. holocylcus* (NC_005293.1), *I. persulcatus* (NC_004370.1), *I. pavlovskyi* (NC_023831.1) and *Bothriocroton concolor* (NC_017756.1). The latter served as an outgroup. The alignment was trimmed down to only contain protein coding segments, and only the coding regions present in all species were retained and used in the analysis. A models test was run and based on this a maximum likelihood tree was constructed in MEGA 7 (http://www.megasoftware.net/). Maximum parsimony and Neighbour joining analyses were also performed to provide statistical support for the ML tree.

2.7.3 Mining for bacterial genome sequences in total DNA sequencing data

Mining for potential bacterial genome sequences within the total DNA sequencing data was performed in CLC Genomics workbench (http://www.clcbio.com/). The complete total DNA data set was mapped against the complete genome sequences of the *Borrelia burgdorferi* complete genome (AE000783.1), the *Anaplasma phagocytophylum* complete genome (NC_007797.1) and the intramitochondrial endosymbiont *Midichloria mitochondrii* complete genome (NC_015722.1). The sequences were obtained through NCBI.

3 Results

3.1 Mining the complete mitochondrial genome of *Ixodes ricinus* from total DNA sequencing data

Several different methods for accessing the genetic material of the tick prior to DNA extraction were tried out. The first method followed the mentioned protocol and used the entire tick as a whole (i.e. not divided, crushed), which gave non-optimal results. The second method was to submerge the sample in liquid nitrogen and crush it with a small glass mantle. This method also proved to be non-optimal due to the small size and flat shape of nymphs used in the experiment. The third method involved slicing the tick in half by a scalpel, which was the method used in this experiment and explained in the materials and methods section, gave relatively good results, both for adult ticks and nymphs.

3.1.1 Total DNA sequencing by IonTorrent PGM

The tick B9AM was sequenced to obtain a complete mitochondrial genome to use as a reference in this experiment. The total DNA of this tick was sequenced by the IonTorrent Personal Genome Machine and the mitochondrial genome was identified by mapping the resulting reads to a reference (NC_018369.2) in CLC.

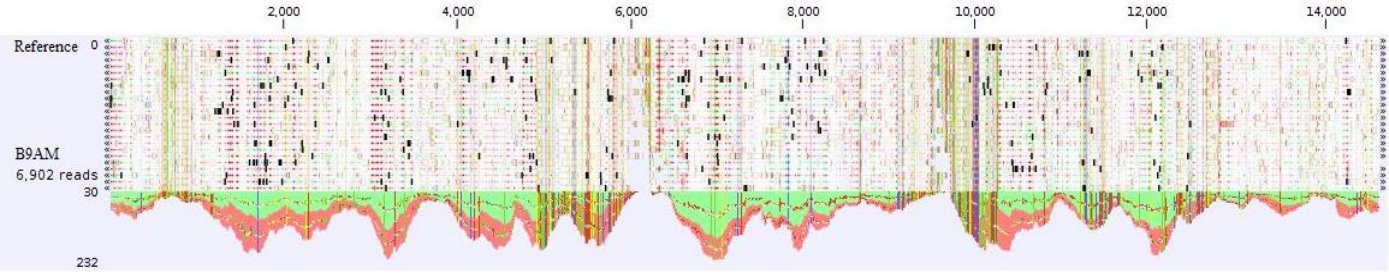


Figure 6. An illustration of the coverage of the mitochondrial genome (B9AM) mapped to the reference sequence (Accession number: NC_018369.2). The length of the mitochondrial genome is shown in the top of the figure and the coverage is visible in the coloured sections at the bottom. How many of the reads

of B9AM that matched the mitochondrial genome is shown in the left of the figure below the name of the tick.

The coverage of the two sequencing runs of B9AM is illustrated in Figure 6, where the data from the runs were mapped against the reference mitochondrial genome. B9AM stands for Brønnøysund tick number 9 Adult Male, which is the location and a number assigned early in the laboratory analysis, and the life stage and gender. The figure shows varying coverage, with two regions (at positions 6000 and 9700) having the smallest amount of coverage. With these reads it was possible to create a consensus sequence encompassing the entire mitochondrial genome of *I. ricinus*.

Table 6. Summary of the total DNA sequencing results. Total sequences before and after trimming are represented by numbers. GC-content is listed in percentage and the average PHRED-score which is a quality measurement.

	Run 1	Run 2
Total reads	4677829	3796354
Total reads after trim	4425128	3535097
Total nucleotides	1.255.032.845	1.036.554.514
Total nucleotides after trim	1.223.700.018	1.008.973.128
GC-content	45 %	45 %
Average PHRED	30	30

A quality report was performed on the results from the sequencing runs before and after trimming, this was to obtain a general overview of the quality resulting from IonTorrent PGM sequencing. The results of the report are shown in Table 6, and in total about 8 million trimmed reads with an average length of 280 nucleotides were generated. Results from both runs are generally similar, however, run 2 resulted in slightly less generated sequences and thus nucleotides. The GC-content and PHRED score were similar between both runs.

3.1.2 Sequence comparison of the two complete *I. ricinus* mtDNA genomes

In addition to the mapping results presented in Figure 5 showing the coverage of B9AM against the reference, Table 7 shows the similarity between the two mitochondrial genomes. This comparison of B9AM and the reference sequence was made in order to analyse the similarity between the two sequences, both length and nucleotide content.

Table 7. Comparison of the B9AM consensus sequence against the reference sequence, showing sequence similarity, total mitochondrial genome length, GC content, AT content, and individual nucleotide content.

	B9AM	Reference
SIMILARITY	96 %	
LENGTH (bp)	14567	14566
G+C CONTENT	20.71 %	21.34 %
A+T CONTENT	79.29 %	78.66 %
A	5631	5594
T	5919	5864
C	1876	1961
G	1141	1147

The two sequences were found to be 96% similar and B9AM has one nucleotide more in length. The reference sequence has a slightly higher GC content and a lower AT content than B9AM. This difference is due to a higher number of C nucleotides and a lower number of T nucleotides in the reference sequence, as shown in the last part of Table 6.

A graphical representation of the circular view of the mitochondrial genome is shown in Figure 7, which was created to obtain a colour-coded overview of the gene content in the mitochondrial genome of *I. ricinus*. Genes coded for by different strands are presented in this figure in such a way that genes located on the light strand are on the inside of the circle and genes on the heavy strand are on the outside. When comparing to the reference sequence, data confirmed that 37 genes are present in the *I. ricinus* mitochondrial genome; 22 tRNAs, 2 rRNAs and 13 protein coding genes. In addition, *I. ricinus* possesses one major control region. Thirteen of the 37 genes are coded for by the light strand.

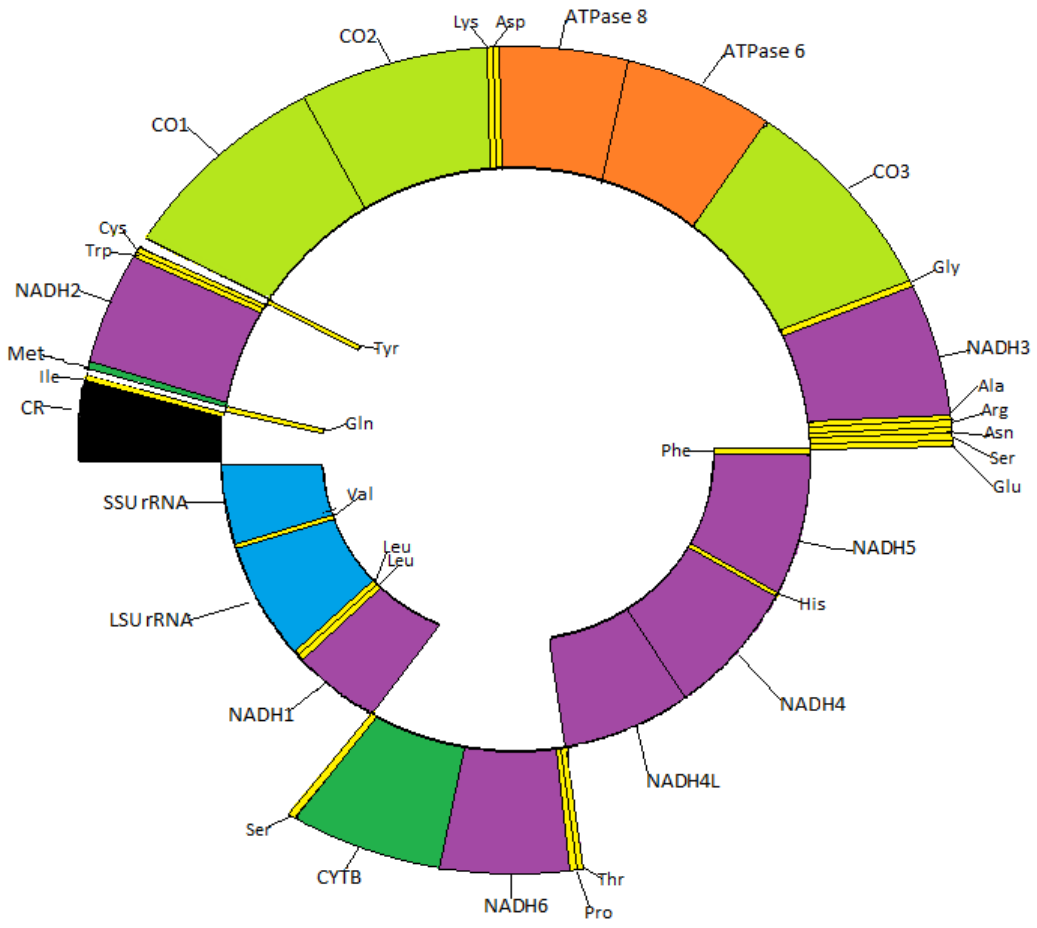


Figure 7. Detailed graphical representation of the mitochondrial genome of *Ixodes ricinus*. The different groups of genes are divided in to colour sections: Dark green represents the start of the mitochondrial genome numbering system (the transfer RNA Met). Yellow represents transfer RNAs, which are represented in the figure by their amino acid abbreviations. Purple represents mitochondrial complex 1 genes, ubiquinone oxidoreductase subunits (NADH1, 2, 3, 4, 4L, 5 and 6). Lime green represents mitochondrial complex IV, cytochrome c oxidase subunits (CO1, 2 and 3). Orange represents mitochondrial complex V, ATP synthase subunits (ATPase6 and 8). Green represents mitochondrial complex III, ubiquitonol-cytochrome c reductase complex subunits (Cytochrome B; CYTB). Blue represents mitochondrially encoded ribosomal RNAs (large sub unit and small sub unit, LSU, SSU). Black represents the control region.

In addition to the graphical representation of the mitochondrial genome, a table was created to show the exact location for each gene component in the sequence. This was done by locating the genes using the reference sequence and annotating B9AM in BioEdit.

The results of the annotation of the consensus sequence are represented in Table 8. The length of each gene, start and stop location (bp) and which strand the genes are located on is present in this table. The location and length coincide with the reference genome in NCBI and showed that the entire mitochondrial genome was obtained by the two sequencing runs of B9AM.

Table 8. A table showing the length, start and stop nucleotide positions and coding strand of the content of the *Ixodes ricinus* mitochondrial genome. Start and stop refers to the exact base pair in which the gene starts and stops and length shows the length of the gene in the amount of base pairs. Heavy strand means that the gene is coded for by the dominant strand, i.e. the strand coding for the majority of genes, whilst the light strand codes for the minority of genes.

Gene	Start	Stop	Length	Strand
tRNA-Met	1	64	64	Heavy
NADH2	65	1035	970	Heavy
tRNA-Trp	1036	1101	65	Heavy
tRNA-Cys	1095	1154	59	Light
tRNA-Tyr	1163	1247	84	Light
CO1	1221	2757	1536	Heavy
CO2	2766	3441	675	Heavy
tRNA-Lys	3442	3510	68	Heavy
tRNA-Asp	3510	3575	65	Heavy
ATP8	3576	3731	155	Heavy
ATP6	3725	4390	665	Heavy
CO3	4397	5174	777	Heavy
tRNA-Gly	5175	5238	63	Heavy
NADH3	5247	5573	326	Heavy
tRNA-Ala	5574	5634	60	Heavy
tRNA-Arg	5634	5697	63	Heavy
tRNA-Asn	5697	5765	68	Heavy
tRNA-Ser	5765	5820	55	Heavy
tRNA-Glu	5821	5901	80	Heavy
tRNA-Phe	5886	5949	63	Light
NADH5	5950	7618	1168	Light

tRNA-His	7619	7679	60	Light
NADH4	7682	8997	1315	Light
NADH4L	8991	9266	275	Light
tRNA-Thr	9269	9330	61	Heavy
tRNA-Pro	9329	9397	68	Light
ND6	9398	9826	428	Heavy
CYTB	9830	10910	1080	Heavy
tRNA-Ser	10911	10977	66	Heavy
NADH1	10994	11912	918	Light
tRNA-Leu	11913	11978	65	Light
tRNA-Leu	11190	11250	60	Light
LSU rRNA	11985	13279	1294	Light
tRNA-Val	13282	13345	63	Light
SSU rRNA	13282	14062	780	Light
Control region	14063	14416	353	Light
tRNA-Ile	14417	14485	68	Heavy
tRNA-Gln	14492	14563	71	Light
Total Length of coding regions			13741	
Total Length of mitochondrial genome			14567	

3.1.3 Single nucleotide polymorphisms in B9AM

Detection of SNPs in B9AM was done by comparing the annotated sequence to the reference sequence and investigating changes at the single nucleotide level. Information for each SNP was noted, such as where it was located (gene and nucleotide position) and which nucleotide change the SNP represented. SNP values were normalized due to the differing lengths of each gene, which was done by: SNP count / (gene length/100).

Table 9. Table showing the single nucleotide polymorphism (SNP) count for each region, type of SNPs, total SNP count and percentage of each SNP type. Each gene is represented by their abbreviated name.

Gene	SNP count	C-T	G-A	A-T	G-T	A-C	C-G
tRNA-Met							
NADH2	52	19	8	22		2	1
tRNA-Trp							
tRNA-Cys	3			3			
tRNA-Tyr							
CO1	6	3	3				
CO2	4	3		1			
tRNA-Lys							
tRNA-Asp							
ATP8							
ATP6	6	2	2	1	1		
CO3	29	14	2	11		2	
tRNA-Gly	1					1	
NADH3	45	23	14	8			
tRNA-Ala	3	2		1			
tRNA-Arg	2	2					
tRNA-Asn	3	2		1			
tRNA-Ser	1		1				
tRNA-Glu	3	1		2			
tRNA-Phe	1			1			
NADH5	47	18	7	17	1	4	
tRNA-His							
NADH4	10	5	2	1		1	1
NADH4L	24	12	3	8		1	
tRNA-Thr	3		1	2			
tRNA-Pro	4	1		2		1	
NADH6	62	25	2	22	6	7	
CYTB	52	29	1	14	3	5	
tRNA-Ser							
NADH1	37	16	3	15	1	2	
tRNA-Leu							
tRNA-Leu							
LSU rRNA	67	19	3	29	5	11	
tRNA-Val							
SSU rRNA	8		3	4		1	
CR	4	2		1		1	
tRNA-Ile	1		1				
tRNA-Gln							
IGR	2	1		1			
Total SNP	480	199	56	167	17	39	2
Total percentage		**41.46 %**	**11.67 %**	**34.79 %**	**3.54 %**	**8.12 %**	**0.42 %**

Table 10. Normalized single nucleotide polymorphism (SNP) counts represented as SNPs per 100 nucleotides.

Gene (Protein coding)	SNP per 100 NT	Gene (non-protein coding)	SNP per 100 NT
NADH2	5.36	**tRNA-cys**	5.08
CO1	0.38	**tRNA-gly**	1.59
CO2	0.59	**tRNA-Ala**	5
ATP6	0.9	**tRNA-Arg**	3.17
CO3	3.73	**tRNA-Asn**	4.41
NADH3	13.8	**tRNA-Ser**	1.82
NADH5	4.02	**tRNA-Glu**	3.75
NADH4	0.74	**tRNA-Phe**	1.59
NADH4L	8.73	**tRNA-Thr**	4.92
NADH6	14.48	**tRNA-Pro**	5.88
CYTB	4.81	**LSU**	5.18
NADH1	4.03	**SSU**	1.03
		CR	1.13
		tRNA-Ile	1.47

Table 9 shows a general overview of the SNP count for each region, in total and the percentage of each type of SNP, whilst Table 10 gives the normalized SNP count. As can be observed in Table 9, the coding region containing the largest amount of SNPs per nucleotide was NADH6, with a count of 14.48 SNP/100nt, closely followed by NADH3 with 13.8 SNP/100nt. Of the coding regions CO1 only contained 0.38 SNP/100nt, which was the lowest amount across the entire dataset. Of the non-protein coding regions tRNA-Pro with 5.88 SNP/100nt and LSU with 5.18 SNP/100nt contained the highest amount, whilst the smallest amount was represented by 1.03 SNP/100nt in SSU.

Regions containing no SNPs at all were tRNA-Met, tRNA-Trp, tRNA-Tyr, tRNA-Lys, tRNA-Asp, ATP8, tRNA-His, tRNA-Ser, tRNA-Leu, tRNA-Val, tRNA-Gln. Three SNP's were found in intergenic regions, in between tRNA-Cys and CO1, tRNA-Gly and NADH3 and SSU rRNA and tRNA-Ile. Changes from C to T represented the majority, being 41.46% of the SNPs. An entire overview of each SNP can be seen in the Appendix (Table A).

3.2 Assessing mitochondrial DNA sequence variation among geographical isolates of *Ixodes ricinus*

Ticks used in the experiment were named after their location and numbered after the order in which their DNA was extracted. The ticks represented in this experiment are not in direct numerical order due to varying results in DNA extraction. Samples abbreviated with B originate from Brønnøysund (B3, B4, B5, B9AM, B20 and B21), So originate from Sogn (So9, So12, So13, So14 and So15), L originates from Levanger (L25, L26, L27) and S originate from Sandnes (S5, S6, S8, S9 and S10). B9AM is the only sample with an additional feature to its name, gender and life stage, however, life stages (and gender) of the remaining samples are listed in table B in the appendix.

Primers obtained from Montagna et al. (2012) were tried out first in an attempt to obtain amplicons from the mitochondrial genome of *Ixodes ricinus*. However, only a few primers gave decent results. The combination of a new set of primers designed from the mitochondrial genome itself was used as the final set of primers, resulting in 8 amplicons. The main intention was to recover the entire mitochondrial genome. However, the eight successful amplicons covered about 8640 base pairs after trimming and homopolymer repair, corresponding to about 60% of the mitochondrial genome.

3.2.1 Amplicon sequencing by IonTorrent PGM

To illustrate the coverage of the population data, the tick So13 was chosen as a representative (Figure 7). This was achieved by mapping the trimmed reads of So13 to the consensus sequence of B9AM in CLC. What can be observed is how much of the mitochondrial genome the sequences from the amplicon sequencing covers which is approximately 8677 base pairs of the 14566 base pairs that is the mitochondrial genome.

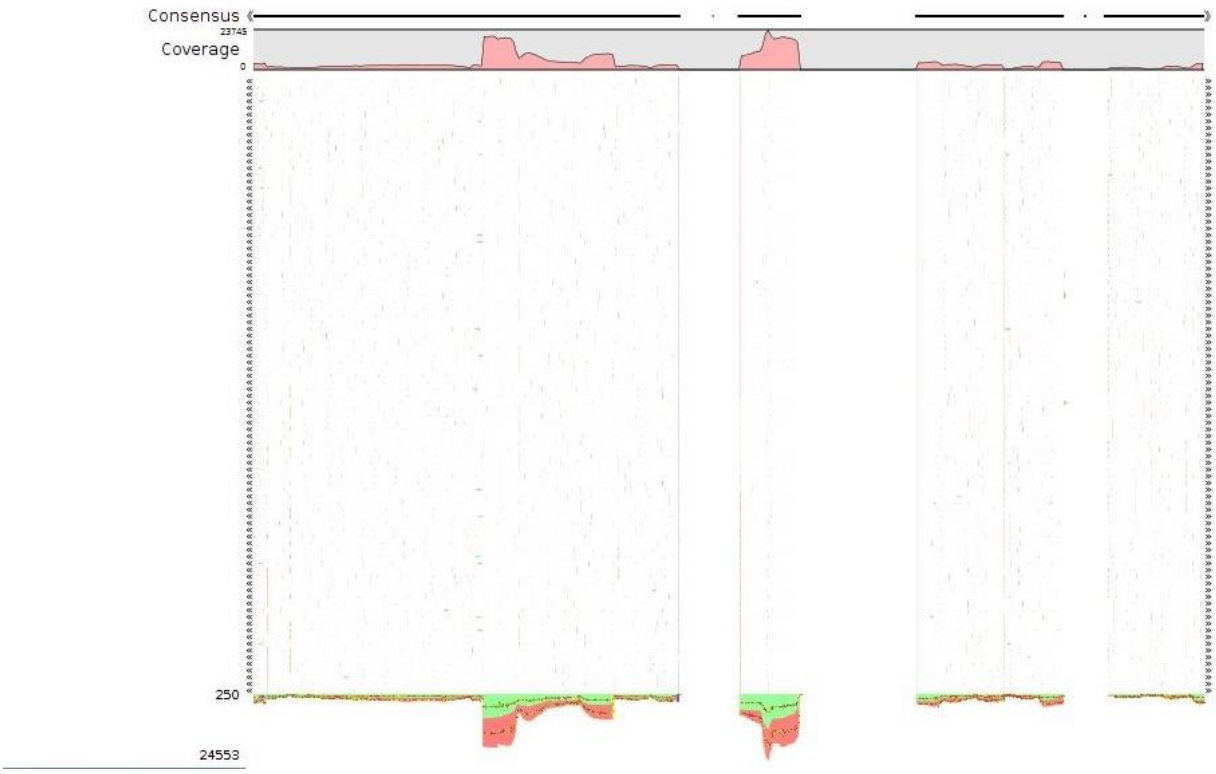


Figure 8. A representative example of the coverage of the population sequencing data using the consensus sequence of B9AM as a reference. Tick So13 was chosen for this figure. The numbers at the top, following the length of the mapping represents the location in the sequence at each point. The number located under the consensus name represents the coverage, i.e. how many reads that are included, and the sequence coverage at the reference sequence positions.

As shown in Figure 8, the coverage of the different amplicons varied somewhat within the sample, however the average coverage in each individual was more than sufficient to create a reliable partial mitochondrial genome consensus sequence from each sample. The coverage of each sample also varied amongst the tick samples, ranging from approximately 120 000 to 19 000, but most samples had a coverage between 20-30 000. S5 deviated most from the data due to a missing section of approximately 300 bp in CO3.

A table was created to show a detailed view of the result of the amplicon sequencing (Table 11), which aims to give an exact view of each gene that was obtained by the amplicon sequencing. The sequence regions were compared to B9AM and the nucleotide locations where each sequence started and stopped in relation to this sequence are noted.

Table 11. Genes represented after amplicon sequencing of the population data. The length and if a gene is complete or partial. The nucleotide positions are based on a comparison with B9AM, thus the numbers correspond to this sequence.

Gene	Present	Start to stop
tRNA-Met	Complete	1-64
NADH2	Complete	65-1035
tRNA-Trp	Complete	1036-1101
tRNA-Cys	Complete	1095-1154
tRNA-Tyr	Complete	1163-1247
CO1	Complete	1221-2757
CO2	Complete	2766-3441
tRNA-lys	Complete	3442-3510
tRNA-Asp	Complete	3510-3575
ATP8	Complete	3576-3731
ATP6	Complete	3725-4390
CO3	Partial	4397-5022
tRNA-Gly	No	
NADH3	No	
tRNA-Ala	No	
tRNA-Arg	No	
tRNA-Asn	No	
tRNA-Ser	No	
tRNA-Glu	No	
tRNA-Phe	No	
NADH5	Partial	6660-7202
tRNA-His	No	
NADH4	No	
NADH4L	No	
tRNA-Thr	No	
tRNA-Pro	No	
NADH6	No	
CYTB	Partial	10312-10910
tRNA-Ser	Complete	10911-10977
NADH1	Complete	10994-11912
tRNA-Leu	Complete	11913-11978
tRNA-Leu	Complete	11190-11250
LSU rRNA	Partial	11985-12175
tRNA-Val	No	
SSU rRNA	Partial	13393-14062

Control region	Complete	14063-14416
tRNA-Ile	Complete	14417-14485
tRNA-Gln	Complete	14492-14563

As can be observed in Table 11, sixteen genes were complete, 5 genes were partial and 16 were not included in the amplicons. Of the 16 complete genes, 6 were protein coding genes, whilst the rest was represented by tRNA genes. Of the partial sequences, 3 were protein coding. CO3 was only missing 152 nucleotides to be completed, but no primer pairs proved successful in this region.

3.2.2 Sequence variation features among geographic isolates

A table of the general SNP features was created to obtain an overview of the entire alignment of the population data. SNPs were investigated by comparing the nucleotide sequences with B9AM in BioEdit, which was a procedure performed for all SNP data in this section. A differentiation between singleton- and parsimony sites, and protein coding and non-coding regions was made. The percentage and count of each nucleotide change was also calculated. A detailed view of each SNP is available in the appendix (Table C, D).

Table 12. General SNP features for the entire alignment of population data. General SNP features show the amount of singleton sites vs. parsimony sites, SNPs in protein coding regions and non-protein coding regions and a count and percentage view of the nucleotide changes within the SNPs.

General SNP features	**Count**	**Percentage**
Singleton sites	54	
Parsimony sites	62	
SNPs in Protein coding regions	359	
SNPs in Non-protein coding regions	121	
Nucleotide changes		
C/T	163	33.9 %
A/T	119	24.8 %

A/G	129	26.9 %
C/A	38	7.9 %
G/T	30	6.3 %
G/C	1	0.2 %
TOTAL SNP	480	

A total of 480 SNPs were found in the population data alignment (Table 12), where 54 of these were singleton sites and the remaining 426 SNPs were located in 62 parsimony sites. Most SNPs were located in protein coding regions (359), whilst the remaining 121 were located in tRNA and rRNA genes, as well as in IGRs and the control region. The general count and percentage of nucleotide changes showed that C/T represented 33.9% of the changes, A/T 24.8%, A/G 26.9%, C/A 7.9%, G/T 6.3% and G/C having the smallest percentage of 0.2%.

Table 13. An overview of the SNP counts in each individual in protein coding regions and non-protein coding regions. The total SNP count for each tick and for each location group is also included.

Tick	SNPs in coding regions	SNPs in non-coding regions	Tick total	Location group total
B3	14	33	47	
B4	15	30	45	
B5	6	19	25	**178**
B20	1	13	14	
B21	15	32	47	
S5	16	40	56	
S6	1	16	17	
S8	8	20	28	**142**
S9	1	8	9	
S10	8	24	32	
L25	15	32	47	
L26	8	25	33	**113**
L27	9	24	33	

So9	1	6	7	
So12		11	11	
So13	2	8	10	**47**
So14		6	6	
So15	1	12	13	
TOTAL	**121**	**359**	**480**	

Table 13 presents the SNP count in the protein coding regions, non-protein coding regions and the total for each sample from all location groups. In addition, total SNPs were calculated for each geographic location in order to compare SNPs between location groups. S5 shows the highest count both in coding and non-coding areas and So14 contained the least amount of SNPs. The group with the smallest number of detected SNPs is Sogn with 47 SNPs, whilst Brønnøysund contained the highest count of 178. The Levanger group, containing only three individuals still contained a large amount of SNPs. A complete overview of each SNP can be found in the appendix (Tables B and C).

Table 14. Overview of total alignment SNP counts in each gene, including the count of which change the SNP represented.

	C/T	A/T	G/T	A/G	A/C	G/C	Total gene SNP count
NADH2	17	1		6			**24**
CO1	18	17	7	27		1	**70**
CO2	18	4		5	1		**28**
ATP6	10	8		25			**43**
CO3	40	18		5	5		**68**
NADH5	18	12	6	1	1		**38**
CYTB	17	1	4	1	5		**28**
NADH1	5	23		18	14		**60**
CR	19	8	4	14	11		**56**
IGR					1		**1**
tRNA-ASP/-LYS				5			**5**
LSU rRNA			8	6			**14**
SSU rRNA	1	27		16	1		**45**
Total change count	**163**	**119**	**29**	**129**	**39**	**1**	**480**

In addition to the sample overview, a summary of SNPs and their changes in each gene was made. CO1 was the region containing the most SNPs with a count of 70, followed by CO3 (68) and NADH1 (60) as shown in table 13. NADH2 and CO2 contained the least SNPs of all protein coding genes, having only 24 and 28 SNPs respectively. C/T and A/G were the most common changes accounting for 163 and 129 of the total SNP count of 480, followed by A/T accounting for 119. The smallest change count was G/C, which was represented by one change, followed by C/T and A/C with 29 and 39.

Table 15. An overview of the amino acid changes caused by SNPs polymorphisms in the geographic isolate data set. New refers to the new codon originating from the change of nucleotide and the amino acid caused by the change in the codon. Original refers to the codon in the reference sequence (NC_018369.2) and the amino acid represented by this codon.

TICK	NEW	ORIGINAL	GENE	NEW	ORIGINAL
L26	AGC	ACC	CO1	A	P
So9	CCA	TCA	CO1	L	S
So14	TTA	TCA	CO1	P	S
S6	ATT	GTT	CO1	V	I
S6	ACA	ATA	CO3	T	M
L27	AGA	GGA	NADH5	S	G
B20	TAT	AAT	NADH5	Y	N
B3	TAT	AAT	NADH5	Y	N
B5	GAT	AAT	NADH5	D	N
So9	TAT	AAT	NADH5	Y	N
S10	TAT	AAT	NADH5	Y	N
S5	TAT	AAT	NADH5	Y	N
S5	ATA	AGA	NADH5	M	S
B5	ATA	CTA	NADH5	M	L
L26	ATA	CTA	NADH5	M	L
L27	ATA	CTA	NADH5	M	L
S10	ATA	CTA	NADH5	M	L
S8	ATA	CTA	NADH5	M	L
S5, S6, S8, S10, So15, So12, B3, B4, B5, B20, B21, L25, L26, L27	GTT	ATT	NADH5	V	I
S5	ATT	ATA	CYTB	I	M
S5	ATA	CTA	CYTB	M	L
B21, B3, B4, L25	GTG	ATG	NADH1	V	M

A table showing the amino acid changes in the alignment of population data was created to see if any of the 480 SNPs caused amino acid changes in their respective sequences. This was achieved by comparing the data set to the reference sequence. The correct reading frame was found by using the reference from NCBI and the sequences were translated and analysed in UGENE. Few amino acid changes were observed in the data set (Table 15). The most common change was from CTA to ATA, which equals a change from L to M.

3.3 Phylogenetic analyses of geographic isolates of *Ixodes ricinus* and closely related *Ixodes* species

3.3.1 Geographic isolates

A phylogenetic analysis was performed to investigate genetic similarity between the samples from the different locations. The analysis was based on an alignment of the entire data set and the outgroup sequence, which was aligned in BioEdit. A maximum likelihood tree constructed from the phylogenetic analysis of the population data is shown in Figure 9. Maximum parsimony and neighbour joining analyses were also performed on the same dataset, which are presented in Figure B and C in the appendix. Complete mitochondrial DNA sequences were not obtained for any *I. ricinus* sample from the different locations. The entire nucleotide consensus of each individual was used in the phylogenetic analysis, which contained 8638 of the 14566 base pairs of the *I. ricinus* mitochondrial genome.

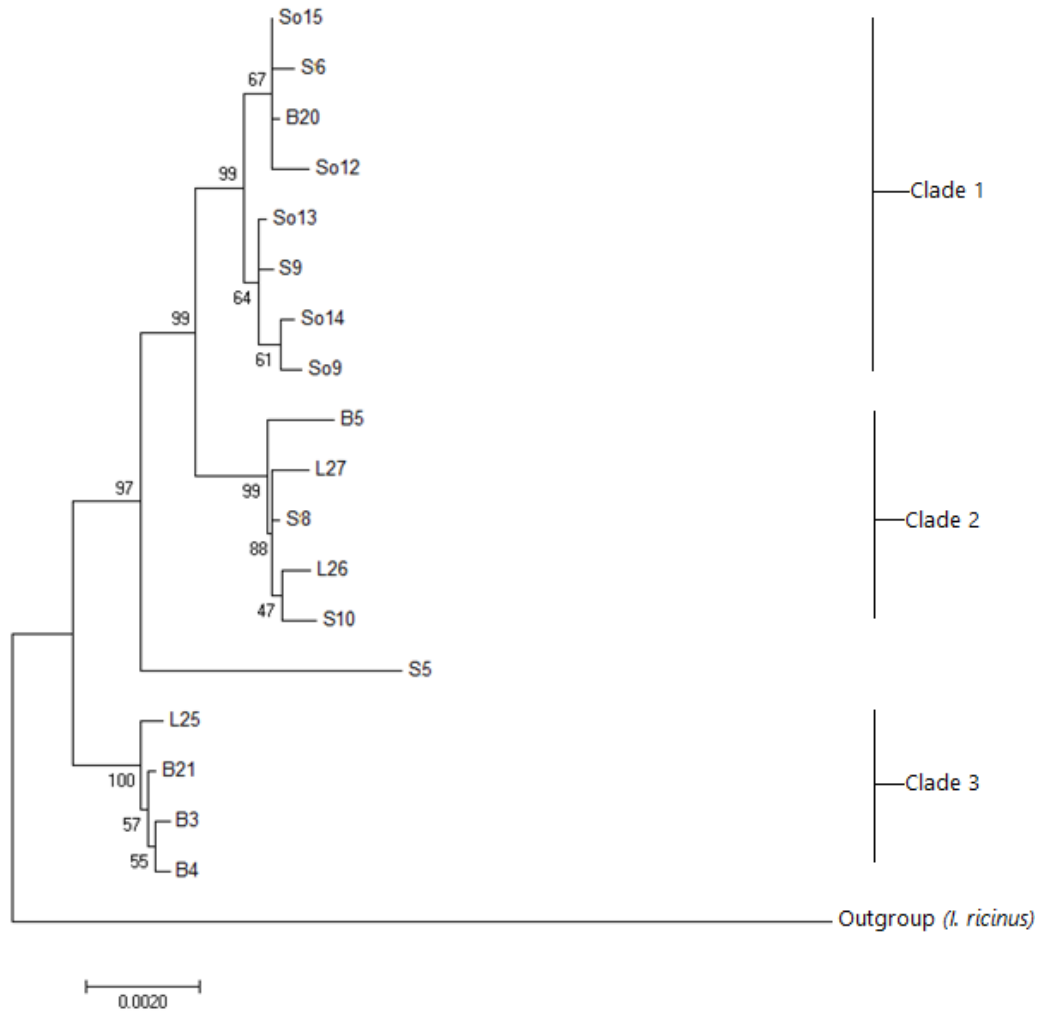


Figure 9. Maximum likelihood tree based on partial mitochondrial genome sequences. The numbers next to the nodes on the tree represents bootstrap percentages for 2000 replications of the maximum likelihood tree model. The outgroup is a distant *Ixodes ricinus* mitochondrial DNA (Accession number: NC_018369.2) sequence trimmed down to match the population data. The name of each sample is represented by the location and a number.

Three clades, supported by high bootstrap values, were identified by the maximum likelihood model (Figure 9). Interestingly, each clade appeared to harbour 2 or 3 locations, suggesting no distinct clustering between sampling locations. Thus no clear evidence of a structure between the geographic isolates was detected, however there are indications of weak structuring within Sogn and Brønnøysund. Ticks from different groups cluster together, indicating variation within populations. S5 was not placed in any cluster, while the rest of Sandnes ticks were placed amongst other populations. B3, B4, B21 and L25 placed in a clade that is more isolated than the rest (clade 3), whilst the other clades seem to be more closely related. Clade 3 contained three of the five samples from the Brønnøysund location. The clade with the highest content of individuals from the same location in the ML-tree was the uppermost clade, containing all samples from Sogn in two sub-clades (clade 1). Two trees, Neighbour joining(NJ) and Maximum parsimony(MP), were created to provide additional support to the ML-tree (Appendix figures B and C). The NJ-tree supported the phylogenetic resolution provided in results. The MP-tree, however, did not root and placed the outgroup in a clade with the geographic isolate samples and presented clade 3 slightly different.

3.3.2 *Ixodes* species.

The maximum likelihood (ML) tree was constructed from protein coding regions of the mitochondrial DNA of all species used in this phylogenetic reconstruction A total of 7 sequences, from 6 *Ixodes* species and one outgroup, was included in the alignment. A models test was performed to find the best suited substitution model and the ML tree was constructed in MEGA 7 based on the test results.

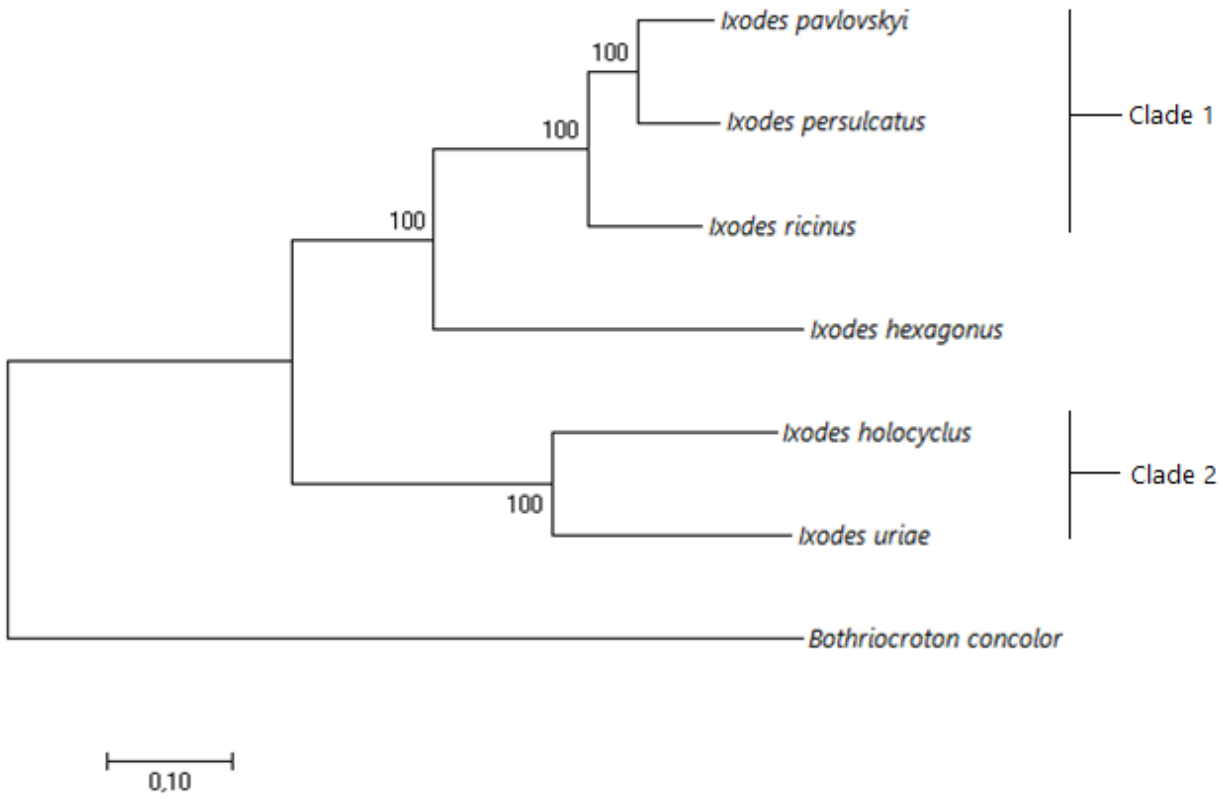


Figure 10. Maximum likelihood tree based on protein coding genes. Numbers to the right of the nodes represent bootstrap values for 2000 replicates.

The statistical support for the ML tree provide full support for all nodes and clades in Figure 10. The outgroup *B. concolor* distanced itself well away from the *Ixodes* species and thus worked well as an outgroup. *I. holocylcus* and *I. uriae* created clade 2 and are seemingly less closely related to the remaining four *Ixodes* species than they are to each other. *I. hexagonus* placed itself alone in between the clade 2 and the three remaining species. *I. pavlovsyi* and *I. persulcatus* created a subclade, within clade 1, in which *I. ricinus* also placed within. The NJ and MP trees (see Appendix, Figure D and E) fully supported the phylogenetic resolution in the maximum likelihood tree.

3.4 Mining for pathogenic/bacterial sequences in the Ixodes ricinus total DNA sequencing data

By sequencing the total DNA of B9AM (8 million reads, 2,2 billion nucleotides; Table 6) it was possible to search for the presence of pathogenic sequences within the sequencing reads. This was done by mapping reads against the reference sequence of the pathogen in question in CLC. If the bacteria are present in the tick sample, one would expect hits when mapping onto the sequences. The reads were mapped to the *Borrelia burgdorferi* complete genome (AE000783.1), the *Anaplasma phagocytophylum* complete genome (NC_007797.1) and the intramitochondrial endosymbiont *Midichloria mitochondrii* complete genome (NC_015722.1). Figures were created by screenshotting from CLC and were made to show the overall coverage of the mappings. Mapping to the first 80.000-100.000 nucleotides are presented in this study, whilst the remaining sections of each mapping are presented in the Appendix.

3.4.1 *Borrelia Burgdorferi*

The mapping of total DNA sequencing data against the first 100.000 nucleotides of the *B. burgdorferi* complete genome was chosen as a representative for the overall mapping. The complete mapping is available in Figure F in the Appendix.

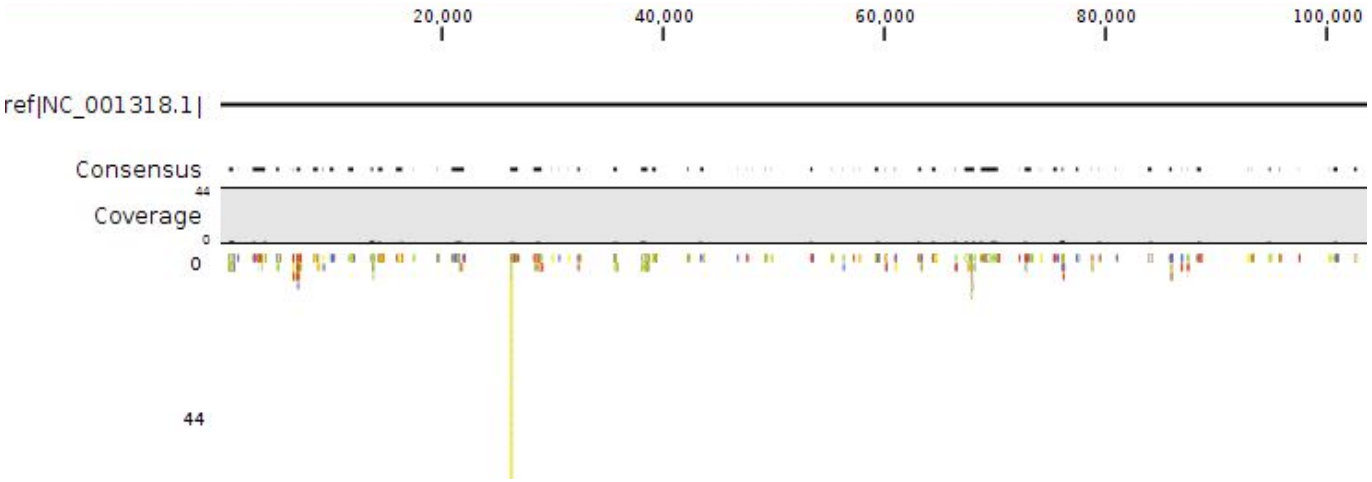


Figure 11. Graphical representation of the mapping of all B9Am reads (trimmed) to the first 100.000 nucleotides of the *Borrelia burgdorferi* genome. The values at the top, following the figure, represent the sequence length at each point. In each row the uppermost number to the right represents the coverage of

the mapping, whilst the bottom value represents how many reads that are mapped in this exact segment of mapping.

Reads from the total DNA sequencing mapped throughout the entire genome of *B. burgdorferi*, the mapping was rather fragmented, but with only two Ion Torrent PGM sequencing runs this appeared expected. A couple of regions showed a higher coverage of mapped reads, otherwise the reads count remained relatively uniform, with an average coverage of 44 reads. With the results in Figure 11 (and Figure F in Appendix), it is possible to say that there is a positive hit for a *Borrelia* bacterium, such as *B. burgdorferi* or a closely related species in this tick sample.

3.4.2 *Anaplasma phagocytophilum*

A figure of the mapping against the first 80.000 nucleotides of the complete *Anaplasma phagocytophilum* genome was chosen as a representative for the entire genome mapping. The complete mapping is available in Figure G in the Appendix.

Figure 12. Representative figure of the mapping of total DNA sequencing runs against the complete genome of *Anaplasma phagocytophilum*. The values at the top, following the figure, represent the sequence length at each point. In each row the uppermost number to the right represents the coverage of the mapping, whilst the bottom value represents how many reads that are mapped in this exact segment of mapping.

The results of the mapping onto the *A. phagocytophilum* genome, Figure 12 (and Figure G in Appendix) showed a minimal amount of coverage across the genome, which probably is not highly significant. Hits across the genome are observed, but these are few and highly non-continuous, indicating no presence of *A. phagocytophilum* in this sample.

3.4.3 *Midichloria mitochondrii*

A figure of the mapping against the first 120.000 nucleotides of the complete genome sequence of *M. mitochondrii* was chosen as a representative for the entire genome mapping. The complete mapping is available in Figure H in the Appendix.

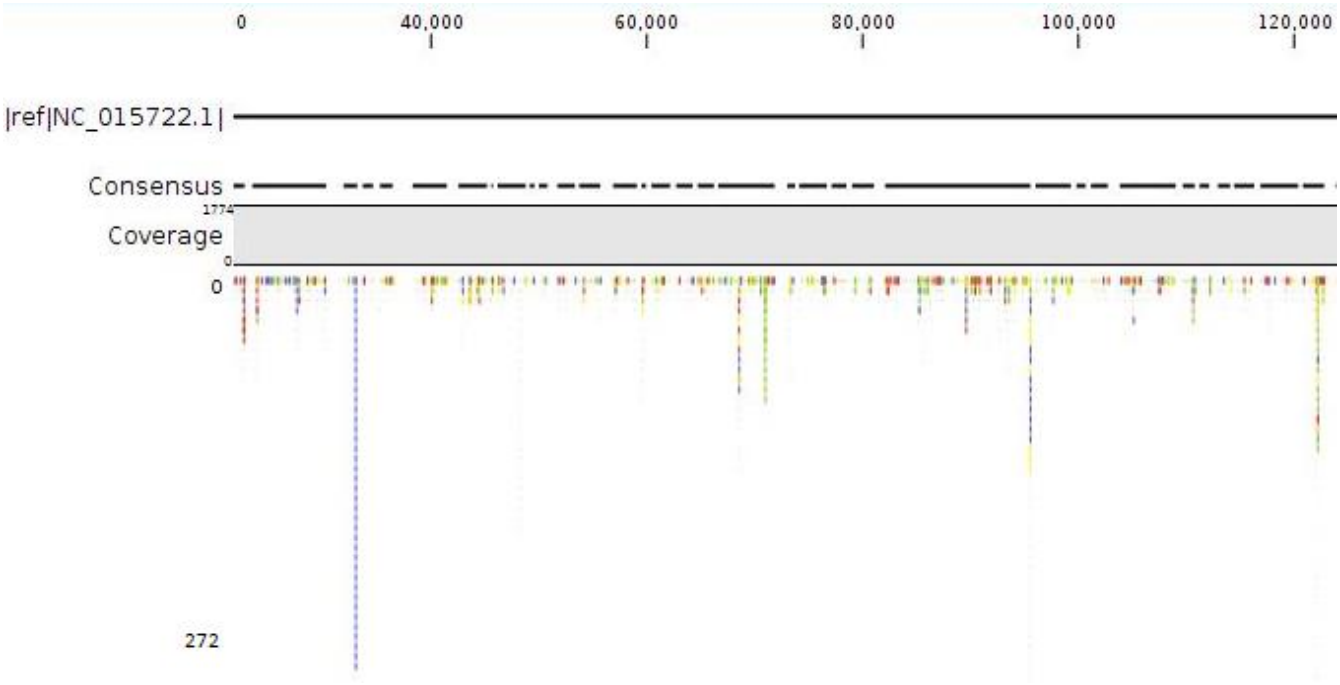


Figure 13. Mapping of the total DNA sequencing runs against the first 120.000 nucleotides of the genome of *Midichloria mitochondrii*. The values at the top, following the figure, represent the sequence length at each point. In each row the uppermost number to the right represents the coverage of the mapping, whilst the bottom value represents how many reads that are mapped in this exact segment of mapping.

Reads from the total DNA sequencing runs of B9AM mapped all throughout the genome of *M. mitochondrii* and the overall coverage of 1774 was relatively high based on the amount of reads recovered from B9AM (Table 6). Based on the results in Figure 13 (and Figure H in the Appendix) is was possible to confirm the presence of *M. midochlorii* in the sequenced tick.

4 Discussion

This study used the IonTorrent next generation sequencing technology to gain sequence information and variability in the mitochondrial genome in geographic isolates of the tick species *Ixodes ricinus*, and to sequence the total DNA of one specimen of *I. ricinus*. The results revealed weak indications of a small difference between groups of *I. ricinus* ticks originating from the four locations in Norway, located up to 1238 km apart, but the results were rather inconclusive. The study showed that ticks from Brønnøysund (3/5 samples) and Sogn (all) placed within own clades, however, these clades where not exclusive, and ticks from other locations were also present. Whether or not there is a genetic structuring between these locations remains to be demonstrated.

Sequencing of total DNA proved to be a valuable method for obtaining large amounts of information. Here, the entire mitochondrial genome sequence of *I. ricinus* was successfully mined out from the total DNA pool and subsequently assembled and presented. The results of the phylogenetic analysis of *Ixodes* species showed that *I. ricinus* is more closely related to *I. hexagonus*, *I. pavlovskyi* and *I. persulcatus* than to *I. uriae* and *I. holocylcus*. The total DNA sequencing contained genomic sequences from pathogenic and endosymbiotic bacterial candidates, from which *B. burgdorferi (*or a close relative) and *M. mitochondrii* were identified.

4.1 Sequence variation and phylogenetic analysis of geographical isolates

A decent amount of nucleotide variation was detected in the geographic isolate dataset (Tables 13 and 14), but relatively few amino acid changes were detected (Table 15). This feature is a supporting argument of using nucleotide sequence information (and not amino acids) in the phylogenetic analysis. A pattern was already visible when inspecting Tables C and D in appendix. Here it was apparent that non-geographically linked "groups" of ticks shared several polymorphic sites and as supported by the results in Figure 10 there was no clear evidence of a genetic structure among isolates. However, two groups showed signs of a relatively weak structure.

Active movement and dispersal in ticks is highly limited. Nymphal *I. ricinus* were shown to walk up to 9.65 m during night-time, but only during optimal conditions (Perret et al., 2003). This means that dispersal is mainly due to host movements, which could contribute to genetic structure (Noureddine et al., 2011). An infected hosts movement determines both rates and patterns of gene flow and can create homogeneity among connected populations. The hosts most responsible are mammals and birds, which are highly mobile (Casati et al., 2008). When groups of species are isolated, either geographically or any other reason, signs of a genetic structure would be apparent. The results obtained in this study are not in total agreement, but still relatively similar to that found in Eurasian ticks (Delaye et al., 1997, Casati et al., 2008, Noureddine et al., 2011), where little-to-no significant structure was found between ticks originating from different locations. At the time being, the results of this study does not confirm nor reject the general belief that *I. ricinus* in Europe is panmictic (Dinnis et al., 2014).

The mixture of low and no apparent genetic structure found within geographic isolates (Figure 10) could very much be due to migratory animals, and due to the large span in distance within sampling locations one could suspect migratory birds. Roe deer for instance, which are known to carry ticks (Kiffner et al., 2011), has shown to have an average migration distance of 3.8 km (males) and 12.4 km (females) in southern Norway (Mysterud, 1999), a distance probably not adequate enough to account for the low genetic structure observed here. Birds (especially passerine birds) are frequently infected by Ixodid ticks (Olsén et al., 1995, Kjelland et al., 2010), and there is also evidence that birds function as a reservoir for *Borrelia* spp. whilst simultaneously distributing the vector (Comstedt et al., 2006). A quick search on

Artsdatabanken.no, without further investigation, gives an indication that the most common tick-transporting birds (Kjelland et al., 2010) have been observed at all sampling locations in this study. It is easy to hypothesize that birds might be the main culprit in this scenario of the observed low genetic structure, however, it is more likely that they work in concert with cervids and other animals. In addition, XU and co-workers have hypothesised that species in the *Ixodes ricinus* complex are the most recently evolved group within the genus *Ixodes* (Xu et al., 2003), which could be an explanation of homogeneity within *I. ricinus* isolates (Casati et al., 2008). If this is the case, *I. ricinus* in Europe would not have had time to establish a distinct geographic structure.

As there were signs of some genetic structure among Sogn and Brønnøysund, one could hypothesise that the Norwegian *I. ricinus* actually do maintain some genetic structuring within geographic locations, and that the two locations are more isolated (in some way) compared to the other two. At least in the case of Brønnøysund, one might speculate that Brønnøy, being one of the northernmost stable tick "populations" known (Hvidsten et al., 2014, Hvidsten et al., 2015), has different abiotic factors which may create a difference in the selection pressure compared to the other locations. Sogn on the other hand is not geographically unique compared to Levanger and Sogn. As the common tick-transporting birds are present in all locations, a lack of these could not explain the result obtained in this study. However, one could suggest that there is a possibility of a lower influx of ticks dropped off and/or transported to by birds in these distinct locations as opposed to the others, but the amount of ticks transported by birds is not known in these locations. As the results were relatively inconclusive, it is not possible to conclude with anything without further studies involving a larger amount of samples and including geographic isolates from other locations.

4.2 Phylogenetic analyses on *Ixodes* species

The phylogenetic analysis of *Ixodes* species in this experiment supported the conclusions from previously reported phylogenetic arrangements. The phylogenetic relationships of *Ixodes* species has been investigated and reported by several researchers e.g. (Xu et al., 2003, Shao et al., 2005, Montagna et al., 2012). Most of these included sequences of other ticks in addition to using different molecular markers. Montagna used the same species (including others) and mitochondrial DNA as a marker, and thus is the most comparable. The results in the present

experiment (Figure 10) were similar to those reported in (Montagna et al., 2012) which provided further support for the relationship among these particular species. Species placed within the *Ixodes ricinus* species complex; *I. ricinus, I. persulcatus* and *I. pavlovskyi* all grouped within the same clade. Ticks within the prostriata (genus *Ixodes*), such as ticks within the *Ixodes ricinus* complex, are basal members within the family Ixodidae. These ticks possess the basic mitochondrial gene order of chelicerates that was reported for the living fossil *Limulus polyphemus* (Staton et al., 1997, Black and Roehrdanz, 1998), which is believed to have the same arrangement as the hypothetical ancestor of Arthropods (Lavrov et al., 2000, Shao et al., 2005). *I. uriae* and *I. holocyclus* are called Australasian ticks and has been found to contain a duplicated control region (Shao et al., 2005). However, no gene rearrangements have occurred between the *Ixodes* species used in this analysis. The Australasian species were placed within their own clade, as they do in previous reports. *I. hexagonus* is not located in any of the clades, which makes sense due to the fact that it is not in any of the groups mentioned above. A phylogenetic analysis based on amino acid sequence variation could also have been performed to obtain additional data. However, since time was an issue this analysis was not performed. It would also have been more interesting to include the mitochondrial genome of more Ixodes species, however, only those used in this experiment is available at the moment.

4.3 Mining for genomic sequences from pathogenic-bacterial candidates

Based on disease data (through MSIS), Hvidsten et al. (2014) reported that cases of Lyme borreliosis in Brønnøy account for 29% of reported cases in northern Norway. Based on sampling of ticks from different locations in Brønnøy, Hvidsten et al. (2015) found that overall percentage of Borrelia content in adult ticks was 46%. Of adult *I. ricinus*, females present the highest disease risk, as a blood meal is required to lay eggs. Males do not have the same need (Apanaskevich and Oliver, 2014) and are thus less likely to transmit disease. However, they are still capable of carrying the bacteria in question (Gern et al., 1999) and other pathogens, in addition to being able to feed and transmit tick-borne diseases.

The male tick sample screened for *B.burgdorferi* in this experiment originated from Brønnøy. The presence of a bacterium equal or similar to *B.burgdorferi* was present in this sample (Figure

11), however, more data and investigation is needed to identify the bacterium of high confidence and to determine if it is a pathogen or only a part of the ticks normal bacterial flora.

The *Anaplasma phagocytophilum* prevalence in Norway was investigated by Henningsson et al. (2015) and a prevalence of 3.0% was found in northern Norway which is notably lower than that of *B. burgdor*feri. *A. phagocytophilum* was not detected in the tick sample from Brønnøysund (Figure 12) and as the percentage prevalence seems generally low overall in Norway (e.g. 2.1% in southern Norway), this observation was to be expected.

The presence of *Midichloria mitochondrii* (or a closely related bacterium) was detected in the total DNA sequencing analysis (Figure 13). This bacterium (order *Rickettsiales*) has been found to reside within the mitochondria of *I. ricinus* (Beninati et al., 2004, Sassera et al., 2006). This symbiotic bacterium has a prevalence of 100% in female *I. ricinus,* especially in ovarian cells (Sassera et al., 2006). Presence in males has been reported only by evidence from PCR and the observed prevalence was estimated to be 44% (Lo et al., 2006). The results from this study prove that one can detect such bacteria by sequencing total DNA, however, more detailed investigations are needed to obtain further information.

4.4 Sample storage and DNA extraction

In this experiment, DNA extraction and PCR amplification was slightly problematic. Issues with tick DNA extraction have been mentioned in previous literature, where it has been noted that extracted tick DNA seem to be susceptible to degradation (Hubbard et al., 1995, Hill and Gutierrez, 2003, Halos et al., 2004, Ammazzalorso et al., 2015). The hard exoskeleton of ticks can also create problems during extraction (Halos et al., 2004). There have been questioned whether or not there are potential Taq-polymerase inhibitors present (Halos et al., 2004), but as Q5® High-Fidelity DNA polymerase was used in this experiment, the influence of this inhibitor is not relevant. If optimized kit and methods for high DNA yield that was determined for *I. scapularis* (Ammazzalorso et al., 2015) would have been used, maybe DNA extraction would have performed better. Unfortunately, these kits were not available.

4.5 Next generation sequencing

In regards to next generation sequencing technologies, Sanger sequencing is referred to as first generation technology. More recent methods are considered next generation, where one of the major advantages is the production of large amounts of data at a very low price (Metzker, 2010), in addition to the fact that they perform massively parallel high-throughput analyses (Liu et al., 2013). Using next generation sequencing platforms in analyses of geographic isolates, as in this experiment, one can study large datasets at a low cost and short handling time and increase the amount of comparable studies, thus increase the current knowledge. The IonTorrent PGM platform is a recently introduced platform that uses semiconductor technology to detect protons that are released when nucleotides are incorporated during the synthesis of DNA (Rothberg et al., 2011, Quail et al., 2012,). This platform was the first among the NGS platforms to not employ fluorescence and camera scanning, which has resulted in a smaller instrument size, higher speed, and as mentioned, lower cost (Liu et al., 2013). Each sequencing platform comes with its own bias however, and the most common error with Ion PGM sequencing is an inaccurate determination of homopolymers, causing insertion/deletion errors (Bragg et al., 2013). However, all sequencing platforms have their own biases that we have to be aware of and take that in to account when analysing sequences.

The method used sequencing of the total DNA of a specimen by IonTorrent PGM and further mining the mitochondrial genome of *I. ricinus* from the sequencing results were successful. Four sequencing chips in total were used to obtain the information used in this experiment for both geographic isolate data and total DNA sequencing. The latter reads were also useful in several different analyses in this experiment, which means that this particular method can aid in large experiments with many targets, even with relatively limited resources.

4.6 The mitochondrial genome of *Ixodes ricinus*

The two sequencing runs added up to approximately 8 million reads, corresponding to 2,2 billion nucleotides. The resulting coverage after reference mapping (Figure 6) was more than sufficient to create a consensus sequence. The successful mining of the mitochondrial genome (Tables 7 and 8) from the two sequencing reads indicated that the method used in this study to obtain it

was sufficient to obtain a complete and confident mitogenome sequence. There were no apparent differences from the already sequenced mitogenome reported by Montagna et al. (2012). Xu et al. (2003) stated that it is widely known that tick mitochondrial genomes are A-T rich, containing about 70% A-T. As apparent in Table 7 this is true and even higher than 70%, since both complete mitochondrial genomes (reference and consensus) have a percentage of A-T surrounding 79% (Table 6).
A number of SNPs were detected in the consensus sequence when comparing to the reference (Table 9,10), which was not a surprising observation as the reference sequence was mined from a tick originating from Italy.

4.7 Mitochondrial DNA as a marker

The ideal molecular marker should be present as a single copy or as multiple homogenous copies, be easy to align, the substitution rate should be at an optimal level (i.e. not too high or not too low), universal primers should be available and the marker should be widely used for comparisons with other studies (Cruickshank, 2002, Patwardhan et al., 2014).
One single marker cannot fulfil all these conditions, and one has to choose the optimal marker for its purpose (Cruickshank, 2002). In both the phylogenetic approaches presented in this study, selected parts of the mitochondrial genome were used as markers. Here all the protein coding regions were assessed in species phylogeny and all obtained nucleotide sequences were included in the geographic isolate phylogeny (Table 11).

Animal mitochondrial DNA possess characteristics that makes it useful as a marker, for instance it is present as multiple copies of the same sequences, which makes it easy to amplify (Cruickshank, 2002), The go-to mitochondrial markers for ticks have been SSU rRNA, LSU rRNA and CO1, while the remaining genes had as of 2004 been relatively ignored (Barker and Murrell, 2004). CO1 is said to evolve more slowly in relation to other genes of the mitochondrial genome (Russo et al., 1996, Patwardhan et al., 2014) and CO1 combined with CO2 have been widely used in phylogenetic reconstructions (Patwardhan et al., 2014). However, the latter authors also stated that recent works have implied that there is much potential in other regions of the mitochondrial genome of ticks for inference of phylogeny and evolutionary studies. As is

apparent in Table 10, some of the genes containing the highest variation were NADH3 and NADH6, which seem rather unexplored in tick phylogenetic assessments.

Nuclear genes are another option for phylogenetic analysis, and nuclear ribosomal genes are thought to be powerful in phylogenetic analyses of mites and ticks, however in particular for distantly related species (Cruickshank, 2002). Nuclear genomes are in no doubt larger, and thus contain more informative sites.

Shao and Barker stated almost 10 years ago that most investigations of parasitic arthropods focusing on population- and evolutionary studies have used sequences of single mitochondrial genes, whilst combining the sequences of several genes has not received its deserved attention (Shao and Barker, 2007). Combinations of with different rates of evolution and substitution rate (as was done in this study) has become more common and apparently provides a higher phylogenetic resolution.

4.8 Conclusion

Three hypotheses about the genetic structure of *I. ricinus* in Norway were proposed for this study and of these, none could be tentatively rejected. This study, based on mitogenome sequences, gave rather inconclusive results, where two groups of geographic isolates showed weak genetic structuring (where Sogn was most distinct), whilst the remaining groups did not. Still, this result suggests that ticks have some method of dispersal between the study sites, most likely birds since the distance between geographic locations in this study is large. While more data is needed to further support this result, this project provides a foundation for future studies, and hypothesis one and two could not be rejected. Hypothesis number three stating that the samples originating from Brønnøysund would form its own clade in phylogenetic analysis might perhaps not be completely rejected yet, since three of the five individuals ended up together in one clade, but more studies are needed to challenge this possibility. The *Ixodes* species phylogeny supported previous research, showing a separation between Australasian and non-Australasian ticks and total DNA sequencing proved to be a useful tool, with successful mining of the mitochondrial genome and screening for potentially pathogenic bacteria.

4.9 Future perspectives

Further research should focus on sequencing the entire genome of *I. ricinus* based on next generation sequencing technologies. The genome is not yet fully sequenced and a complete record of its genomic content could help further future research. A more extensive study on the relationships between geographical isolates of *I. ricinus* in Norway, using a larger amount of ticks and using a larger span of locations would provide a deeper insight into the genetic structure of *I. ricinus* in Norway. A more complete report of pathogen prevalence in ticks across Norway is needed, as this kind of information still remains unreported in several regions, for instance Trøndelag. By coupling such investigations with the increasing knowledge of birds transporting ticks, in addition to analysing the last blood meal of collected ticks, one would obtain useful information surrounding dispersal potential and epidemiology, which could further be used in reaching the ultimate goal, decreasing the incidence of tick-borne diseases.

References

AMMAZZALORSO, A. D., ZOLNIK, C. P., DANIELS, T. J. & KOLOKOTRONIS, S.-O. 2015. To beat or not to beat a tick: comparison of DNA extraction methods for ticks (*Ixodes scapularis*). *PeerJ,* 3**,** e1147.

APANASKEVICH, D. A. & OLIVER, J., J.H. 2014. Life cycles and natural history of ticks. *In:* SONENSHINE, D. E. & ROE, M. R. (eds.) *Biology of Ticks.* New York: Oxford University Press, 1, p. 59-73.

BARKER, S. C. & MURRELL, A. 2004. Systematics and evolution of ticks with a list of valid genus and species names. *Parasitology,* 129**,** S15-S36.

BARKER, S. C. & MURRELL, A. 2008. Systematics and evolution of ticks with a list of valid genus and species names. *In:* BOWMAN, A. S. & NUTTALL, P. A. (eds.) *Ticks: Biology, Disease and Control.* New York: Cambridge University Press, p. 1-39.

BENINATI, T., LO, N., SACCHI, L., GENCHI, C., NODA, H. & BANDI, C. 2004. A novel alpha-proteobacterium resides in the mitochondria of ovarian cells of the tick *Ixodes ricinus*. *Applied and Environmental Microbiology,* 70**,** 2596-2602.

BLACK, W. C. T. & ROEHRDANZ, R. L. 1998. Mitochondrial gene order is not conserved in arthropods: prostriate and metastriate tick mitochondrial genomes. *Molecular Biology and Evolution,* 15**,** 1772-85.

BRAGG, L. M., STONE, G., BUTLER, M. K., HUGENHOLTZ, P. & TYSON, G. W. 2013. Shining a light on dark sequencing: characterising errors in Ion Torrent PGM data. *PLoS Computational Biology,* 9**,** e1003031.

BREWER, M. S., COTORAS, D. D., CROUCHER, P. J. P. & GILLESPIE, R. G. 2014. New sequencing technologies, the development of genomics tools, and their applications in evolutionary arachnology. *Journal of Arachnology,* 42**,** 1-15.

BRINKMAN, F. S. L. & LEIPE, D. D. 2001. Phylogenetic Analysis. *In:* BAXEVANIS, A. D. & OUELLETTE, B. F. F. (eds.) *Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins.* 2nd ed.: John Wiley & Sons, Inc, p. 323-357.

BROOKES, A. J. 1999. The essence of SNPs. *Gene,* 234**,** 177-186.

BROWN, W. M., GEORGE, M. & WILSON, A. C. 1979. Rapid evolution of animal mitochondrial DNA. *Proceedings of the National Academy of Sciences,* 76**,** 1967-1971.

CAKIC, S., MOJSILOVIC, M., MIHALJICA, D., MILUTINOVIC, M., PETROVIC, A. & TOMANOVIC, S. 2014. Molecular characterization of CO1 gene of *Ixodes ricinus* (linnaeus, 1758) from Serbia. *Archives of Biological Sciences,* 66**,** 683-690.

CAPORALE, D. A., RICH, S. M., SPIELMAN, A., TELFORD, S. R. & KOCHER, T. D. 1995. Discriminating between *Ixodes* ticks by means of mitochondrial DNA sequences. *Molecular Phylogenetics and Evolution,* 4**,** 361-365.

CASATI, S., BEMASCONIB, M. V., GERN, L. & PIFFARETTI, J. C. 2008. Assessment of intraspecific mtDNA variability of European *Ixodes ricinus* sensu stricto (Acari : Ixodidae). *Infection Genetics and Evolution,* 8**,** 152-158.

COMSTEDT, P., BERGSTROM, S., OLSEN, B., GARPMO, U., MARJAVAARA, L., MEJLON, H., BARBOUR, A. G. & BUNIKIS, J. 2006. Migratory passerine birds as reservoirs of Lyme borreliosis in Europe. *Emerging Infectious Diseases,* 12**,** 1087-95.

CRUICKSHANK, R. H. 2002. Molecular markers for the phylogenetics of mites and ticks. *Systematic and Applied Acarology,* 7**,** 3-14.

DANTAS-TORRES, F., CHOMEL, B. B. & OTRANTO, D. 2012. Ticks and tick-borne diseases: a one health perspective. *Trends in Parasitology,* 28, 437-446.

DE LA FUENTE, J., ESTRADA-PENA, A., CABEZAS-CRUZ, A. & BREY, R. 2015. Flying ticks: anciently evolved associations that constitute a risk of infectious disease spread. *Parasites & Vectors,* 8, 538

DELAYE, C., BÉATI, L., AESCHLIMANN, A., RENAUD, F. & DE MEEÛS, T. 1997. Population genetic structure of *Ixodes ricinus* in Switzerland from allozymic data: No evidence of divergence between nearby sites. *International Journal for Parasitology,* 27**,** 769-773.

DINNIS, R. E., SEELIG, F., BORMANE, A., DONAGHY, M., VOLLMER, S. A., FEIL, E. J., KURTENBACH, K. & MARGOS, G. 2014. Multilocus sequence typing using mitochondrial genes (mtMLST) reveals geographic population structure of *Ixodes ricinus* ticks. *Ticks and Tick-Borne Diseases,* 5**,** 152-160.

FONSTAD, T., GENSBØL, B. & GÜNTHER, M. 2007. *Aschehougs Store Fuglebok,* Oslo, Norway, H. Aschehoug & Co, p. 330-502.

FOURNIER, P. E., GRUNNENBERGER, F., JAULHAC, B., GASTINGER, G. & RAOULT, D. 2000. Evidence of *Rickettsia helvetica* infection in humans, eastern France. *Emerging Infectious Diseases,* 6**,** 389-392.

FREELAND, J. R., KIRK, H. & PETERSEN, S. D. 2011. *Molecular Ecology,* West Sussex, UK, Wiley-Blackwell, p. 71, 379.

GERN, L., HU, C. M., KOCIANOVA, E., VYROSTEKOVA, V. & REHACEK, J. 1999. Genetic diversity of *Borrelia burgdorferi* sensu lato isolates obtained from Ixodes ricinus ticks collected in Slovakia. *European Journal of Epidemiology,* 15**,** 665-669.

GILBERT, L., AUNGIER, J. & TOMKINS, J. L. 2014. Climate of origin affects tick (*Ixodes ricinus*) host- seeking behavior in response to temperature: implications for resilience to climate change? *Ecology and Evolution,* 4**,** 1186-1198.

GRAY, J. S., DAUTEL, H., ESTRADA-PENA, A., KAHL, O. & LINDGREN, E. 2009. Effects of climate change on ticks and tick-borne diseases in europe. *Interdisciplinary Perspectives on Infectious Diseases,* 2009**,** 593232-593232.

GRITSUN, T. S., LASHKEVICH, V. A. & GOULD, E. A. 2003. Tick-borne encephalitis. *Antiviral Research,* 57**,** 129-146.

GUGLIELMONE, A. A., ROBBINS, R. G., APANASKEVICH, D. A., PETNEY, T. N., ESTRADA-PENA, A., HORAK, I. G., SHAO, R. & BARKER, S. C. 2010. The Argasidae, Ixodidae and Nuttalliellidae (Acari: Ixodida) of the world: a list of valid species names. *Zootaxa***,** 1-28.

HALL, T. A. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series* 41**,** 95-98.

HALOS, L., JAMAL, T., VIAL, L., MAILLARD, R., SUAU, A., LE MENACH, A., BOULOUIS, H.-J. & VAYSSIER-TAUSSAT, M. 2004. Determination of an efficient and reliable method for DNA extraction from ticks. *Veterinary Research,* 35**,** 709-713.

HELYAR, S. J., HEMMER-HANSEN, J., BEKKEVOLD, D., TAYLOR, M. I., OGDEN, R., LIMBORG, M. T., CARIANI, A., MAES, G. E., DIOPERE, E., CARVALHO, G. R. & NIELSEN, E. E. 2011. Application of SNPs for population genetics of nonmodel organisms: new opportunities and challenges. *Molecular Ecology Resources,* 11**,** 123-136.

HENNINGSSON, A. J., HVIDSTEN, D., KRISTIANSEN, B.-E., MATUSSEK, A., STUEN, S. & JENKINS, A. 2015. Detection of Anaplasma phagocytophilum in *Ixodes ricinus* ticks from Norway using a realtime PCR assay targeting the Anaplasma citrate synthase gene gltA. *BMC Microbiology,* 15**,** 153.

HILL, C. & GUTIERREZ, J. 2003. A method for extraction and analysis of high quality genomic DNA from ixodid ticks. *Medical and Veterinary Entomology,* 17**,** 224-227.

HUBBARD, M., CANN, K. & WRIGHT, D. 1995. Validation and rapid extraction of nucleic acids from alcohol-preserved ticks. *Experimental and Applied Acarology,* 19**,** 473-478.

HVIDSTEN, D., STORDAL, F., LAGER, M., ROGNERUD, B., KRISTIANSEN, B. E., MATUSSEK, A., GRAY, J. & STUEN, S. 2015. Borrelia burgdorferi sensu lato-infected *Ixodes ricinus* collected from vegetation near the Arctic Circle. *Ticks and Tick-borne Diseases,* 6**,** 768-73.

HVIDSTEN, D., STUEN, S., JENKINS, A., DIENUS, O., OLSEN, R. S., KRISTIANSEN, B. E., MEHL, R. & MATUSSEK, A. 2014. *Ixodes ricinus* and Borrelia prevalence at the Arctic Circle in Norway. *Ticks and Tick-borne Diseases,* 5**,** 107-12.

JAENSON, T. G. T. & LINDGREN, E. 2011. The range of *Ixodes ricinus* and the risk of contracting Lyme borreliosis will increase northwards when the vegetation period becomes longer. *Ticks and Tick-borne Diseases,* 2**,** 44-49.

JONGEJAN, F. & UILENBERG, G. 2004. The global importance of ticks. *Parasitology,* 129**,** S3-S14.

JORE, S., VANWAMBEKE, S. O., VILJUGREIN, H., ISAKSEN, K., KRISTOFFERSEN, A. B., WOLDEHIWET, Z., JOHANSEN, B., BRUN, E., BRUN-HANSEN, H. & WESTERMANN, S. 2014. Climate and environmental change drives *Ixodes ricinus* geographical expansion at the northern range margin. *Parasites & Vectors,* 7**,** 1.

JORE, S., VILJUGREIN, H., HOFSHAGEN, M., BRUN-HANSEN, H., KRISTOFFERSEN, A. B., NYGÅRD, K., BRUN, E., OTTESEN, P., SÆVIK, B. K. & YTREHUS, B. 2011. Multi-source analysis reveals latitudinal and altitudinal shifts in range of *Ixodes ricinus* at its northern distribution limit. *Parasites & Vectors,* 4**,** 1.

KIFFNER, C., LÖDIGE, C., ALINGS, M., VOR, T. & RÜHE, F. 2011. Attachment site selection of ticks on roe deer, *Capreolus capreolus*. *Experimental and Applied Acarology,* 53**,** 79-94.

KJELLAND, V., STUEN, S., SKARPAAS, T. & SLETTAN, A. 2010. *Borrelia burgdorferi* sensu lato in *Ixodes ricinus* ticks collected from migratory birds in southern Norway. *Acta Veterinaria Scandinavia,* 52**,** 59.

KLOMPEN, H. & GRIMALDI, D. 2001. First mesozoic record of a parasitiform mite: a larval argasid tick in Cretaceous amber (Acari : Ixodida : Argasidae). *Annals of the Entomological Society of America,* 94**,** 10-15.

LAVROV, D. V., BOORE, J. L. & BROWN, W. M. 2000. The complete mitochondrial DNA sequence of the horseshoe crab *Limulus polyphemus*. *Molecular Biology and Evolution,* 17**,** 813-824.

LEGER, E., VOURC'H, G., VIAL, L., CHEVILLON, C. & MCCOY, K. D. 2013. Changing distributions of ticks: causes and consequences. *Experimental and Applied Acarology,* 59**,** 219-244.

LIU, G.-H., CHEN, F., CHEN, Y.-Z., SONG, H.-Q., LIN, R.-Q., ZHOU, D.-H. & ZHU, X.-Q. 2013. Complete mitochondrial genome sequence data provides genetic evidence that the brown dog tick *Rhipicephalus sanguineus* (Acari: Ixodidae) represents a species complex. *International Journal of Biological Sciences,* 9**,** 361-369.

LO, N., BENINATI, T., SASSERA, D., BOUMAN, E., SANTAGATI, S., GERN, L., SAMBRI, V., MASUZAWA, T., GRAY, J. & JAENSON, T. 2006. Widespread distribution and high

prevalence of an alpha-proteobacterial symbiont in the tick *Ixodes ricinus*. *Environmental Microbiology,* 8**,** 1280-1287.

LU, X., LIN, X.-D., WANG, J.-B., QIN, X.-C., TIAN, J.-H., GUO, W.-P., FAN, F.-N., SHAO, R., XU, J. & ZHANG, Y.-Z. 2013. Molecular survey of hard ticks in endemic areas of tick-borne diseases in China. *Ticks and Tick-Borne Diseases,* 4**,** 288-296.

MARICONTI, M., EPIS, S., GAIBANI, P., DALLA VALLE, C., SASSERA, D., TOMAO, P., FABBI, M., CASTELLI, F., MARONE, P., SAMBRI, V., BAZZOCCHI, C. & BANDI, C. 2012. Humans parasitized by the hard tick *Ixodes ricinus* are seropositive to *Midichloria mitochondrii*: is *Midichloria* a novel pathogen, or just a marker of tick bite? *Pathogens and Global Health,* 106**,** 391-396.

MEHL, R. 1983. The distributions and host relations of Norwegian ticks (Acari, Ixodides). *Fauna Norvegica,* Ser.B**,** 46-51.

MEHL, R., BRAATHEN L.R. 1987. Skogflåtten *Ixodes ricinus*. *Tidsskrift for den Norske Lægeforening,* 107**,** 1642-1644.

MEHL, R., MICHAELSEN, J., LID G. 1984. Ticks (Acari, Ixodides) on migratory birds in Norway. *Fauna Norvegica,* Ser B**,** 46-58.

METZKER, M. L. 2010. Sequencing technologies—the next generation. *Nature Reviews Genetics,* 11**,** 31-46.

MEYER, J. M. & HILL, C. A. 2013. Tick genetics, genomics, and transformation. *In:* SONENSHINE, D. E. & ROE, M. R. (eds.) *Biology of Ticks.* New York: Oxford University Press, 2**,** p. 61-87.

MONTAGNA, M., SASSERA, D., GRIGGIO, F., EPIS, S., BANDI, C. & GISSI, C. 2012. Tick-box for 3 '-end formation of mitochondrial transcripts in Ixodida, basal chelicerates and *Drosophila*. *PLoS ONE,* 7, e47538

MYSTERUD, A. 1999. Seasonal migration pattern and home range of roe deer (*Capreolus capreolus*) in an altitudinal gradient in southern Norway. *Journal of Zoology,* 247**,** 479-486.

NAVA, S., GUGLIELMONE, A. A. & MANGOLD, A. J. 2009. An overview of systematics and evolution of ticks. *Frontiers in Bioscience-Landmark,* 14**,** 2857-2877.

NAVAJAS, M. & FENTON, B. 2000. The application of molecular markers in the study of diversity in acarology: A review. *Experimental and Applied Acarology,* 24**,** 751-774.

NOUREDDINE, R., CHAUVIN, A. & PLANTARD, O. 2011. Lack of genetic structure among Eurasian populations of the tick *Ixodes ricinus* contrasts with marked divergence from north-African populations. *International Journal for Parasitology,* 41**,** 183-192.

OGDEN, N., H., ARTSOB, H., MARGOS, G. & TSAO, J. 2014. Non-rickettsial tick-borne bacteria and the diseases they Cause. *In:* SONENSHINE, D. E. & ROE, M. R. (eds.) *Biology of Ticks.* New York: Oxford University Press, 2, p. 279-312.

OLSÉN, B., JAENSON, T. & BERGSTRÖM, S. 1995. Prevalence of *Borrelia burgdorferi* sensu lato-infected ticks on migrating birds. *Applied and Environmental Microbiology,* 61**,** 3082-3087.

PATWARDHAN, A., RAY, S. & ROY, A. 2014. Molecular markers in phylogenetic studies-A review. *Journal of Phylogenetics & Evolutionary Biology,* 2, 131.

PAULAUSKAS, A., RADZIJEVSKAJA, J., ROSEF, O., TURCINAVICIENE, J., AMBRASIENE, D. & MAKAREVICIUTE, D. 2006. Genetic variation of ticks (*Ixodes ricinus* L.) in the Lithuanian and Norwegian populations. *Experimental and Applied Acarology,* 40**,** 259-270.

PERRET, J.-L., GUERIN, P. M., DIEHL, P. A., VLIMANT, M. & GERN, L. 2003. Darkness induces mobility, and saturation deficit limits questing duration, in the tick *Ixodes ricinus*. *Journal of Experimental Biology,* 206**,** 1809-1815.

POINAR, G. & BROWN, A. E. 2003. A new genus of hard ticks in Cretaceous Burmese amber (Acari : Ixodida : Ixodidae). *Systematic Parasitology,* 54**,** 199-205.

PORRETTA, D., MASTRANTONIO, V., MONA, S., EPIS, S., MONTAGNA, M., SASSERA, D., BANDI, C. & URBANELLI, S. 2013. The integration of multiple independent data reveals an unusual response to Pleistocene climatic changes in the hard tick *Ixodes ricinus*. *Molecular Ecology,* 22**,** 1666-1682.

QUAIL, M. A., SMITH, M., COUPLAND, P., OTTO, T. D., HARRIS, S. R., CONNOR, T. R., BERTONI, A., SWERDLOW, H. P. & GU, Y. 2012. A tale of three next generation sequencing platforms: comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers. *BMC Genomics,* 13**,** 1-13.

RANDOLPH, S. E. 2004. Evidence that climate change has caused 'emergence' of tick-borne diseases in Europe? *International Journal of Medical Microbiology Supplements,* 293**,** 5-15.

RANDOLPH, S. E. 2014. Ecology of non-nidicolous ticks. *In:* SONENSHINE, D. E. & ROE, M. R. (eds.) *Biology of Ticks.* New York: Oxford University Press, 2**,** p. 3-60.

ROTHBERG, J. M., HINZ, W., REARICK, T. M., SCHULTZ, J., MILESKI, W., DAVEY, M., LEAMON, J. H., JOHNSON, K., MILGREW, M. J., EDWARDS, M., HOON, J., SIMONS, J. F., MARRAN, D., MYERS, J. W., DAVIDSON, J. F., BRANTING, A., NOBILE, J. R., PUC, B. P., LIGHT, D., CLARK, T. A., HUBER, M., BRANCIFORTE, J. T., STONER, I. B., CAWLEY, S. E., LYONS, M., FU, Y., HOMER, N., SEDOVA, M., MIAO, X., REED, B., SABINA, J., FEIERSTEIN, E., SCHORN, M., ALANJARY, M., DIMALANTA, E., DRESSMAN, D., KASINSKAS, R., SOKOLSKY, T., FIDANZA, J. A., NAMSARAEV, E., MCKERNAN, K. J.,

WILLIAMS, A., ROTH, G. T. & BUSTILLO, J. 2011. An integrated semiconductor device enabling non-optical genome sequencing. *Nature,* 475**,** 348-352.

RUMER, L., SHESHUKOVA, O., DAUTEL, H., MANTKE, O. D. & NIEDRIG, M. 2011. Differentiation of medically important Euro-Asian tick species *Ixodes ricinus, Ixodes persulcatus, Ixodes hexagonus,* and *Dermacentor reticulatus* by polymerase chain reaction. *Vector-Borne and Zoonotic Diseases,* 11**,** 899-905.

RUSSO, C. A., TAKEZAKI, N. & NEI, M. 1996. Efficiencies of different genes and different tree-building methods in recovering a known vertebrate phylogeny. *Molecular Biology and Evolution,* 13**,** 525-536.

SASSERA, D., BENINATI, T., BANDI, C., BOUMAN, E. A., SACCHI, L., FABBI, M. & LO, N. 2006. 'Candidatus *Midichloria mitochondrii*', an endosymbiont of the tick *Ixodes ricinus* with a unique intramitochondrial lifestyle. *International Journal of Systematic and Evolutionary Microbiology,* 56**,** 2535-2540.

SHAO, R. & BARKER, S. 2007. Mitochondrial genomes of parasitic arthropods: implications for studies of population genetics and evolution. *Parasitology,* 134**,** 153-167.

SHAO, R., BARKER, S. C., MITANI, H., AOKI, Y. & FUKUNAGA, M. 2005. Evolution of duplicate control regions in the mitochondrial genomes of metazoa: a case study with Australasian *Ixodes* ticks. *Molecular Biology and Evolution,* 22**,** 620-629.

SONENSHINE, D. E. & ROE, M. R. 2014. Ticks, people and animals. *In:* SONENSHINE, D. E. & ROE, M. R. (eds.) *Biology of ticks.* Oxford University Press, 1, p. 3-16

SÜSS, J., KLAUS, C., GERSTENGARBE, F. W. & WERNER, P. C. 2008. What makes ticks tick? Climate change, ticks, and tick-borne diseases. *Journal of Travel Medicine,* 15**,** 39-45.

STATON, J. L., DAEHLER, L. L. & BROWN, W. M. 1997. Mitochondrial gene arrangement of the horseshoe crab *Limulus polyphemus* L.: conservation of major features among arthropod classes. *Molecular Biology and Evolution,* 14**,** 867-874.

TOMKINS, J. L., AUNGIER, J., HAZEL, W. & GILBERT, L. 2014. Towards an evolutionary understanding of questing behaviour in the tick *Ixodes ricinus*. *Plos One,* 9, e110028

URICARU, R., RIZK, G., LACROIX, V., QUILLERY, E., PLANTARD, O., CHIKHI, R., LEMAITRE, C. & PETERLONGO, P. 2015. Reference-free detection of isolated SNPs. *Nucleic Acids Research,* 43**,** e11.

XIONG, J. 2006. *Essential Bioinformatics,* New York, Cambridge University Press, p. 326.

XU, G., FANG, Q. Q., KEIRANS, J. E. & DURDEN, L. A. 2003. Molecular phylogenetic analyses indicate that the *Ixodes ricinus* complex is a paraphyletic group. *Journal of Parasitology,* 89**,** 452-457.

Appendix

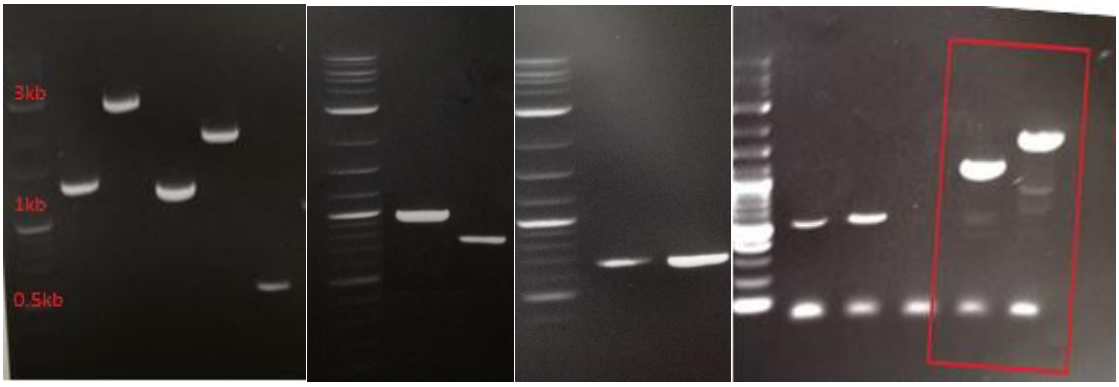


Figure A. Snapshot of the gel electrophoresis run verifying amplification of primer pairs; in order 1,2,3,4,5,8,9,10. The reference points in the 2-log ladder is shown to the right in red lettering. Two bands of primer pair 10 are showed due to the low quality of band 1. The latter image represents the additional primer pairs, 12+13, in numerical order (marked with a red square).

Table A. Table representing each SNP found in B9AM, the original nucleotide, base pair position and location in the genome.

SNP	Original (reference)	Base pair position	Base pair position reference	Gene
C	T	343	343	NADH2
G	A	523	523	NADH2
T	C	542	542	NADH2
A	T	583	583	NADH2
A	G	593	593	NADH2
A	T	603	603	NADH2
T	C	604	604	NADH2
A	T	608	608	NADH2
A	G	609	609	NADH2

T	C	611	611	NADH2
T	C	613	613	NADH2
A	G	622	622	NADH2
T	A	628	628	NADH2
T	C	630	630	NADH2
T	C	640	640	NADH2
T	A	649	649	NADH2
T	A	656	656	NADH2
A	T	678	678	NADH2
T	C	688	688	NADH2
T	C	701	701	NADH2
T	A	703	703	NADH2
T	C	706	706	NADH2
T	A	709	709	NADH2
T	C	712	712	NADH2
A	C	733	733	NADH2
T	A	742	742	NADH2
T	A	745	745	NADH2
C	T	760	760	NADH2
T	A	766	766	NADH2
A	T	775	775	NADH2
A	T	781	781	NADH2
T	C	811	811	NADH2
A	T	814	814	NADH2
G	C	815	815	NADH2
A	T	817	817	NADH2
T	A	826	826	NADH2
C	T	838	838	NADH2
T	C	839	839	NADH2
A	G	841	841	NADH2
T	A	848	848	NADH2
T	C	851	851	NADH2

A	G	859	859	NADH2
A	C	865	965	NADH2
T	C	872	972	NADH2
A	T	874	874	NADH2
A	G	875	875	NADH2
T	A	877	877	NADH2
T	C	886	886	NADH2
C	T	912	912	NADH2
A	T	925	925	NADH2
A	T	967	967	NADH2
G	A	1000	1000	NADH2
A	T	1139	1138	tRNA-Cys
T	A	1140	1139	tRNA-Cys
A	T	1141	1140	tRNA-Cys
T	A	1161	1160	/NA
C	T	1692	1691	CO1
G	A	2037	2036	CO1
G	A	2271	2270	CO1
A	G	2301	2300	CO1
T	C	2622	2621	CO1
T	C	2715	2714	CO1
T	C	2807	2808	CO2
T	C	3163	3162	CO2
A	T	3293	3294	CO2
T	C	3407	3408	CO2
A	G	3997	3998	ATPase6
G	T	4009	4010	ATPase6
C	T	4042	4023	ATPase6
C	T	4162	4163	ATPase6
A	G	4201	4202	ATPase6
A	T	4266	4267	ATPase6
T	C	4552	4553	CO3

G	A	4936	4937	CO3
A	T	4942	4943	CO3
T	A	4949	4950	CO3
C	T	4960	4961	CO3
A	T	4963	4964	CO3
T	C	4966	4967	CO3
T	A	4972	4973	CO3
A	C	4975	4976	CO3
T	C	4984	4985	CO3
C	T	4999	5000	CO3
T	C	5008	5009	CO3
T	A	5014	5015	CO3
T	C	5017	5018	CO3
C	T	5029	5030	CO3
T	A	5045	5046	CO3
T	A	5047	5048	CO3
T	C	5048	5049	CO3
A	T	5053	5054	CO3
A	T	5071	5072	CO3
A	C	5078	5079	CO3
A	T	5079	5080	CO3
T	C	5083	5084	CO3
T	C	5086	5087	CO3
G	A	5092	5093	CO3
T	C	5101	5102	CO3
C	T	5122	5123	CO3
A	T	5134	5135	CO3
T	C	5143	5144	CO3
C	A	5189	5190	Trna-Gly
T	C	5244	5245	/NA
C	T	5248	5249	NADH3
G	A	5250	5251	NADH3

A	T	5252	5253	NADH3
A	T	5261	5262	NADH3
T	C	5265	5266	NADH3
A	G	5271	5272	NADH3
T	C	5272	5273	NADH3
A	T	5282	5283	NADH3
T	C	5285	5286	NADH3
G	A	5294	5295	NADH3
T	C	5299	5300	NADH3
T	C	5303	5304	NADH3
G	A	5310	5311	NADH3
G	A	5313	5314	NADH3
C	T	5318	5319	NADH3
A	G	5321	5322	NADH3
C	T	5336	5337	NADH3
T	C	5345	5346	NADH3
C	T	5351	5352	NADH3
C	T	5354	5355	NADH3
T	C	5363	5364	NADH3
A	T	5366	5367	NADH3
C	T	5381	5382	NADH3
T	C	5384	5385	NADH3
T	C	5410	5411	NADH3
G	A	5436	5437	NADH3
T	C	5438	5439	NADH3
T	C	5441	5442	NADH3
T	C	5444	5445	NADH3
C	T	5471	5472	NADH3
A	G	5481	5482	NADH3
T	A	5486	5487	NADH3
T	C	5487	5488	NADH3
T	C	5489	5490	NADH3

G	A	5490	5491	NADH3
A	T	5492	5493	NADH3
C	T	5501	5502	NADH3
A	G	5502	5503	NADH3
A	G	5522	5523	NADH3
A	G	5525	5526	NADH3
T	C	5529	5530	NADH3
A	T	5531	5532	NADH3
G	A	5540	5541	NADH3
T	A	5558	5559	NADH3
G	A	5573	5574	NADH3
T	C	5603	5604	tRNA-Ala
A	T	5624	5625	tRNA-Ala
T	C	5649	5650	tRNA-Ala
T	C	5676	5677	tRNA-Arg
C	T	5678	5679	tRNA-Arg
T	C	5744	5745	tRNA-Asn
A	T	5751	5752	tRNA-Asn
T	C	5763	5764	tRNA-Asn
A	G	5796	5797	tRNA-Ser
A	T	5836	5837	tRNA-Glu
T	A	5862	5863	tRNA-Glu
T	C	5871	5872	tRNA-Glu
T	A	5896	5897	tRNA-Phe
T	A	5898	5899	tRNA-Phe
T	A	5953	5954	NADH5
A	T	5956	5957	NADH5
C	T	5957	5958	NADH5
A	C	5976	5977	NADH5
T	C	5992	5993	NADH5
A	G	5999	6000	NADH5
A	C	6025	6026	NADH5

T	G	6027	6022	NADH5
A	T	6033	6028	NADH5
A	T	6037	6032	NADH5
A	G	6049	6044	NADH5
A	T	6055	6050	NADH5
T	C	6062	6057	NADH5
C	T	6065	6060	NADH5
C	A	6071	6066	NADH5
T	C	6074	6069	NADH5
T	C	6092	6087	NADH5
T	C	6104	6099	NADH5
G	A	6109	6104	NADH5
T	C	6122	6117	NADH5
A	T	6129	6124	NADH5
T	A	6145	6140	NADH5
A	T	6155	6150	NADH5
A	T	6157	6152	NADH5
T	C	6167	6162	NADH5
A	G	6182	6177	NADH5
A	T	6186	6181	NADH5
T	C	6195	6190	NADH5
T	C	6200	6195	NADH5
T	C	6205	6200	NADH5
T	A	6207	6202	NADH5
T	C	6212	6207	NADH5
T	C	6214	6209	NADH5
A	T	6221	6216	NADH5
T	C	6224	6219	NADH5
A	G	6235	6230	NADH5
T	A	6242	6237	NADH5
A	C	6254	6249	NADH5
T	A	6259	6254	NADH5

A	T	6332	6327	NADH5
A	T	6344	6339	NADH5
C	T	6458	6453	NADH5
T	C	6940	6935	NADH5
T	A	6992	6987	NADH5
G	A	7084	7079	NADH5
C	T	7271	7266	NADH5
A	G	7313	7308	NADH5
A	G	7733	7728	NADH4
C	T	7841	7836	NADH4
C	T	8096	8091	NADH4
C	T	8647	8642	NADH4
T	A	8651	8646	NADH4
G	A	8663	8658	NADH4
A	C	8665	8660	NADH4
C	T	8666	8661	NADH4
T	C	8974	8970	NADH4
C	G	9073	9069	NADH4
T	C	9090	9086	NADH4L
C	T	9092	9088	NADH4L
G	A	9094	9090	NADH4L
T	A	9102	9098	NADH4L
T	C	9105	9101	NADH4L
A	T	9122	9118	NADH4L
A	T	9132	9128	NADH4L
A	G	9133	9129	NADH4L
C	T	9134	9130	NADH4L
T	A	9147	9143	NADH4L
A	G	9149	9145	NADH4L
C	T	9156	9152	NADH4L
T	A	9158	9154	NADH4L
T	A	9159	9155	NADH4L

C	T	9189	9185	NADH4L
C	T	9198	9194	NADH4L
T	C	9224	9220	NADH4L
T	A	9228	9224	NADH4L
A	T	9231	9227	NADH4L
T	C	9243	9239	NADH4L
T	C	9251	9247	NADH4L
T	C	9252	9248	NADH4L
A	C	9254	9250	NADH4L
T	C	9255	9251	NADH4L
T	A	9284	9280	tRNA-Thr
G	A	9309	9305	tRNA-Thr
A	T	9319	9315	tRNA-Thr
A	T	9341	9337	tRNA-Pro
A	T	9373	9369	tRNA-Pro
A	C	9375	9371	tRNA-Pro
T	C	9378	9374	tRNA-Pro
T	A	9421	9417	NADH6
T	C	9426	9422	NADH6
C	A	9460	9456	NADH6
A	C	9461	9457	NADH6
A	T	9466	9462	NADH6
T	G	9486	9482	NADH6
T	C	9497	9493	NADH6
T	C	9529	9525	NADH6
T	C	9530	9526	NADH6
T	C	9532	9528	NADH6
C	T	9533	9529	NADH6
C	A	9535	9531	NADH6
A	G	9547	9543	NADH6
T	A	9550	9546	NADH6
C	T	9554	9550	NADH6

G	T	9559	9555	NADH6
T	C	9572	9568	NADH6
T	A	9592	9588	NADH6
T	C	9601	9597	NADH6
T	A	9614	9610	NADH6
C	T	9615	9611	NADH6
A	T	9619	9615	NADH6
T	C	9631	9627	NADH6
A	T	9634	9630	NADH6
T	A	9638	9634	NADH6
A	T	9640	9636	NADH6
T	C	9649	9645	NADH6
A	T	9650	9646	NADH6
T	C	9653	9649	NADH6
A	T	9655	9651	NADH6
T	A	9661	9657	NADH6
C	T	9668	9664	NADH6
C	T	9669	9665	NADH6
T	C	9673	9669	NADH6
T	G	9680	9676	NADH6
C	A	9687	9683	NADH6
C	T	9691	9687	NADH6
T	C	9694	9690	NADH6
T	C	9695	9691	NADH6
A	T	9697	9693	NADH6
T	A	9698	9694	NADH6
T	G	9699	9695	NADH6
A	C	9708	9704	NADH6
T	C	9712	9708	NADH6
A	T	9721	9717	NADH6
A	T	9727	9723	NADH6
G	A	9730	9726	NADH6

C	T	9733	9729	NADH6
A	C	9737	9733	NADH6
T	C	9745	9741	NADH6
C	T	9757	9753	NADH6
T	A	9764	9760	NADH6
G	T	9765	9761	NADH6
T	A	9767	9762	NADH6
T	A	9787	9783	NADH6
T	C	9790	9786	NADH6
A	T	9794	9790	NADH6
T	A	9796	9792	NADH6
T	G	9799	9795	NADH6
C	A	9811	9807	NADH6
T	C	9814	9810	NADH6
A	T	9817	9813	NADH6
T	A	9836	9832	CYTB
T	A	9845	9841	CYTB
C	T	9847	9843	CYTB
T	A	9850	9846	CYTB
T	C	9856	9852	CYTB
T	A	9860	9856	CYTB
T	C	9868	9864	CYTB
C	T	9871	9867	CYTB
A	T	9874	9870	CYTB
T	G	9877	9873	CYTB
A	C	9889	9885	CYTB
T	C	9892	9888	CYTB
A	C	9911	9907	CYTB
T	C	9913	9909	CYTB
T	C	9914	9910	CYTB
C	T	9925	9921	CYTB
G	T	9931	9927	CYTB

T	A	9949	9945	CYTB
A	T	9961	9957	CYTB
A	T	9973	9969	CYTB
T	C	9979	9975	CYTB
A	T	9983	9979	CYTB
C	T	9997	9993	CYTB
T	A	10013	10009	CYTB
C	T	10014	10010	CYTB
A	T	10028	10024	CYTB
A	T	10037	10033	CYTB
T	C	10042	10038	CYTB
C	T	10051	10047	CYTB
C	T	10054	10050	CYTB
T	C	10066	10062	CYTB
C	T	10091	10087	CYTB
T	G	10093	10089	CYTB
C	T	10096	10092	CYTB
T	C	10102	10098	CYTB
T	C	10120	10116	CYTB
T	G	10123	10119	CYTB
T	C	10132	10128	CYTB
C	T	10165	10161	CYTB
T	C	10166	10162	CYTB
T	C	10186	10182	CYTB
T	A	10201	10197	CYTB
T	C	10202	10198	CYTB
A	C	10204	10200	CYTB
A	C	10207	10203	CYTB
C	T	10222	10218	CYTB
T	A	10240	10236	CYTB
A	C	10267	10263	CYTB
T	C	10270	10266	CYTB

T	A	10271	10267	CYTB
T	C	10276	10272	CYTB
A	C	10279	10275	CYTB
T	C	10238	10234	CYTB
T	C	10717	10713	CYTB
A	G	10837	10833	CYTB
T	A	11247	11243	NADH1
G	T	11259	11255	NADH1
T	A	11280	11276	NADH1
C	T	11295	11291	NADH1
A	T	11319	11315	NADH1
G	A	11322	11318	NADH1
C	A	11325	11321	NADH1
T	A	11363	11359	NADH1
T	A	11369	11365	NADH1
A	T	11374	11370	NADH1
T	A	11382	11378	NADH1
T	A	11391	11387	NADH1
T	C	11402	11398	NADH1
T	C	11428	11424	NADH1
C	A	11436	11432	NADH1
T	C	11466	11462	NADH1
A	T	11468	11464	NADH1
G	A	11469	11465	NADH1
A	T	11478	11474	NADH1
C	T	11481	11477	NADH1
A	T	11502	11498	NADH1
G	A	11511	11507	NADH1
T	C	11514	11510	NADH1
T	C	11523	11519	NADH1
A	T	11526	11522	NADH1
T	C	11532	11528	NADH1

T	C	11547	11543	NADH1
C	T	11550	11546	NADH1
T	C	11553	11549	NADH1
A	T	11580	11576	NADH1
C	T	11592	11588	NADH1
A	T	11601	11598	NADH1
T	C	11670	11667	NADH1
C	T	11673	11670	NADH1
A	T	11676	11673	NADH1
T	C	11688	11685	NADH1
T	C	11912	11909	NADH1
G	T	12051	12051	LSU rRNA
T	A	12184	12185	LSU rRNA
T	C	12203	12204	LSU rRNA
T	C	12205	12206	LSU rRNA
T	C	12207	12208	LSU rRNA
A	G	12220	12221	LSU rRNA
C	T	12325	12326	LSU rRNA
T	A	12344	12345	LSU rRNA
C	T	12346	12347	LSU rRNA
T	A	12347	12348	LSU rRNA
T	A	12349	12350	LSU rRNA
T	A	12350	12351	LSU rRNA
A	C	12351	12352	LSU rRNA
A	C	12352	12353	LSU rRNA
T	A	12364	12366	LSU rRNA
A	G	12380	12382	LSU rRNA
T	C	12388	12390	LSU rRNA
A	T	12394	12396	LSU rRNA
T	A	12400	12402	LSU rRNA
T	A	12402	12404	LSU rRNA
A	T	12403	12404	LSU rRNA

G	T	12408	12410	LSU rRNA
T	A	12409	12411	LSU rRNA
T	A	12410	12412	LSU rRNA
A	T	12412	12414	LSU rRNA
T	A	12414	12416	LSU rRNA
T	A	12421	12423	LSU rRNA
T	C	12459	12461	LSU rRNA
A	C	12477	12478	LSU rRNA
C	T	12490	12492	LSU rRNA
A	C	12495	12497	LSU rRNA
C	T	12498	12500	LSU rRNA
C	T	12511	12513	LSU rRNA
T	C	12512	12514	LSU rRNA
C	T	12584	12586	LSU rRNA
A	T	12598	12600	LSU rRNA
T	C	12623	12623	LSU rRNA
C	T	12633	12633	LSU rRNA
C	A	12635	12635	LSU rRNA
A	G	12638	12638	LSU rRNA
T	A	12669	12669	LSU rRNA
T	A	12670	12670	LSU rRNA
A	C	12685	12685	LSU rRNA
C	T	12686	12686	LSU rRNA
A	T	12688	12688	LSU rRNA
A	T	12692	12692	LSU rRNA
A	C	12700	12700	LSU rRNA
C	A	12701	12701	LSU rRNA
G	T	12705	12706	LSU rRNA
A	C	12707	12708	LSU rRNA
A	C	12708	12709	LSU rRNA
G	T	12709	12710	LSU rRNA
A	T	12711	12712	LSU rRNA

T	A	12714	12714	LSU rRNA
T	A	12715	12715	LSU rRNA
T	G	12717	12717	LSU rRNA
A	T	12722	12722	LSU rRNA
T	A	12725	12725	LSU rRNA
T	A	12726	12726	LSU rRNA
A	C	12727	12727	LSU rRNA
T	C	12729	12729	LSU rRNA
A	T	12774	12773	LSU rRNA
A	C	12817	12816	LSU rRNA
T	C	12875	12875	LSU rRNA
C	T	13052	13052	LSU rRNA
A	T	13136	13136	LSU rRNA
A	T	13475	13474	SSU rRNA
A	G	13485	13484	SSU Rrna
T	A	13721	13720	SSU rRNA
A	G	13810	13809	SSU rRNA
G	A	13855	13854	SSU rRNA
A	T	13893	13892	SSU rRNA
A	C	13960	13959	SSU rRNA
T	A	13968	13967	SSU rRNA
C	T	14324	14323	CR
A	C	14371	14370	CR
T	C	14387	14386	CR
A	T	14412	14411	CR
G	A	14436	14435	tRNA-Ile

Table B. Life stage and gender of each tick used in the geographic isolate analysis. Nymphal ticks show no specific sex characteristics, so these gender sections are left blank.

Tick	Life stage	Gender

B3	Adult	Male
B4	Adult	Male
B5	Adult	Male
B20	Adult	Male
B21	Adult	Male
L25	Nymph	
L26	Nymph	
L27	Nymph	
S5	Nymph	
S6	Nymph	
S8	Adult	Female
S9	Nymph	
S10	Nymph	
So9	Adult	Female
So12	Adult	Male
So13	Adult	Male
So14	Nymph	
So15	Nymph	

Table C. Table showing each SNP, both singleton and parsimony sites in the population data, including base pair position in according to the complete *Ixodes ricinus* mitochondrial genome.

SNP	Original	Tick(s)	BP-position	Gene
T	C	L25	89	NADH2
C	T	B21, B3, B4, L25	268	NADH2
T	C	B21, B3, B4, L25, S5	343	NADH2
C	T	B21, B3, B4, L25, S5	542	NADH2
A	T	S5	745	NADH2
T	C	S5	838	NADH2
C	T	S5	896	NADH2
A	G	B5, L26, L27, S10, S5, S8	1000	NADH2
G	C	L26	1346	CO1
A	G	S5	1372	CO1
A	G	S5	1438	CO1

G	T	B5, L26, L27, S10, S5, S8	1480	CO1
T	C	S5	1534	CO1
T	C	B21, B3, B4, L25	1690	CO1
A	G	B21, S5	1696	CO1
C	T	So9	1733	CO1
T	C	So14	1734	CO1
A	G	S6	1817	CO1
T	A	S5	1972	CO1
C	T	S5	1996	CO1
T	A	B5, L26, L27, S10, S5, S8	2017	CO1
G	A	So12	2023	CO1
A	G	L27	2030	CO1
T	A	B5, L26, L27, S10, S5, S8	2122	CO1
C	T	B3	2125	CO1
A	G	B21, B3, B4, L25, S5	2269	CO1
T	G	B5	2269	CO1
G	A	B21, B3, B4, B5, L25, L26, L27, S10, S5, S8	2299	CO1
T	A	B21, B3, B4, L25	2380	CO1
G	A	B5, L26, L27, S10, S8	2476	CO1
C	T	B21, B3, B4, L25	2620	CO1
C	T	B21, B3, B4, L25, S5	2713	CO1
C	T	B21, B3, B4, L25	2807	CO2
C	T	B20, B21, B3, B4, B5, L25, L26, L27, So12, So15, S10, S5, S6, S8	3161	CO2
A	C	S5	3284	CO2
T	A	B21, B3, B4, L25	3293	CO2
A	G	B5, L26, L27, S10, S8	3365	CO2
G	A	B21, B3, B4, B5, L25, L26, L27, S10, S5, S8	3997	ATPase6
T	C	S5	4080	ATPase6
T	C	B21, B3, B4, L25	4162	ATPase6
G	A	B20, B21, B3, B4, B5, L25, L26, L27, So12, So13, So14, So15, So9, S5, S6	4201	ATPase6

T	A	B20, B21, B3, B4, L25, So12, So15, S6	4266	ATPase6
C	T	S9	4308	ATPase6
C	T	B21, B3, B4, L25	4365	ATPase6
A	G	B21, B3, B4, L25	4449	CO3
C	T	S6	4157	CO3
A	C	B5, L26, L27, S10, S8	4611	CO3
T	C	B21, B3, B4	4618	CO3
A	G	So12	4638	CO3
C	T	ALL	5011	CO3
A	T	ALL	5017	CO3
C	T	ALL	5020	CO3
T	C	S5	6808	NADH5
C	T	B20, B21, B3, B4, L25, L26, L27, So12, So15, S10, S5, S6, S8	6940	NADH5
T	C	S5	6983	NADH5
T	G	B5, L26, L27, S10, S8	7084	NADH5
A	C	S5	7125	NADH5
C	T	B5	7195	NADH5
A	T	B20, B3, So9, S10, S5	7195	NADH5
A	T	L26	7196	NADH5
C	T	S10	7196	NADH5
C	T	B5	7198	NADH5
A	T	L26, S10, S6	7198	NADH5
G	T	L27	7199	NADH5
A	T	S10	7199	NADH5
T	A	L27	7201	NADH5
G	A	S6	7201	NADH5
T	A	L26	7202	NADH5
C	T	B5, L26, L27, S10, S8	10358	CYTB
T	A	S5	10369	CYTB
T	C	L25	103670	CYTB
T	C	B21, B3, B4, L25	10376	CYTB
A	C	S5	10376	CYTB

A	G	S5	10417	CYTB
C	T	S5	10535	CYTB
A	C	B20, So12, So15, S6	10600	CYTB
C	T	B21, B3, B4, L25, S5	10717	CYTB
C	T	S5	10840	CYTB
G	T	B21, B3, B4, L25	11196	CYTB
C	T	S5	11235	NADH1
T	A	B20, B21, B3, B4, L25, L26, L27, So13, So15, S10, S6, S8, S9	11319	NADH1
A	G	B20, B21, B3, B4, L25, L26, L27, So13, So15, S10, S6, S5, S8, S9	11322	NADH1
A	C	B20, B21, B3, B4, L25, L26, L27, So13, So15, S10, S6, S5, S8, S9	11325	NADH1
T	A	B21, B3, L25, L26, L27, So13, So14, S10, S8, S9	11502	NADH1
G	A	B20, So12, So15, S6	11610	NADH1
C	T	B21, B3, B4, L25	11912	NADH1

Table D. Detailed overview of each SNP in non-protein coding regions in the geographic isolate alignment. Base pair position refers to the position in the complete mitochondrial genome.

SNP	Original	TICK(S)	BP-position	Location
C	A	S5	1156	IGR
A	G	B5, L26, L27, S10, S8	3510	tRNA-Lys/Asp
T	G	B21, B3, B3, L25, L26, L27, S10, S8	12051	LSU rRNA
A	G	S5	12051	LSU rRNA
G	A	B21, B3, B4, L25, S5	12079	LSU rRNA
T	A	B21, B3, B4, B5, L25, L26, L27, S10, S5, S8	13475	SSU rRNA
G	A	B21, B3, B4, L25, S5	13485	SSU rRNA
A	C	S5	13523	SSU rRNA
G	A	S5	13562	SSU rRNA
C	T	B4	13627	SSU rRNA
A	T	L25	13781	SSU rRNA
T	A	S5	13810	SSU rRNA
A	G	B21, B3, B4, B5, L25, L26, L27, S10, S5, S8	13855	SSU rRNA
T	A	B20, B21, L25, L26, L27, So13, So15, So9, S10, S5, S6, S8, S9	13960	SSU rRNA
A	T	L27	13980	SSU rRNA
A	T	So13	13992	SSU rRNA

A	G	B21, B3, B4, L25	14131	CR
T	G	B21, B3, B4, L25	14180	CR
C	T	B21, B3, B4, L25	14205	CR
G	A	B5, L26, L27, S10, S8	14227	CR
G	A	S5	14315	CR
G	A	S5	14320	CR
A	C	S5	14324	CR
T	C	B21, B3, B4, L25	14324	CR
C	T	S5	14346	CR
G	A	B21, B3, B4	14352	CR
C	A	B21, B3, B4, B5, L25, L26, L27, S10, S5, S8	14371	CR
C	T	B21, B3, B4, B5, L25, L26, L27, S10, S5, S8	14387	CR
T	A	B21, B3, B4, L25	14400	CR
T	A	B21, B3, B4, L25	14412	CR

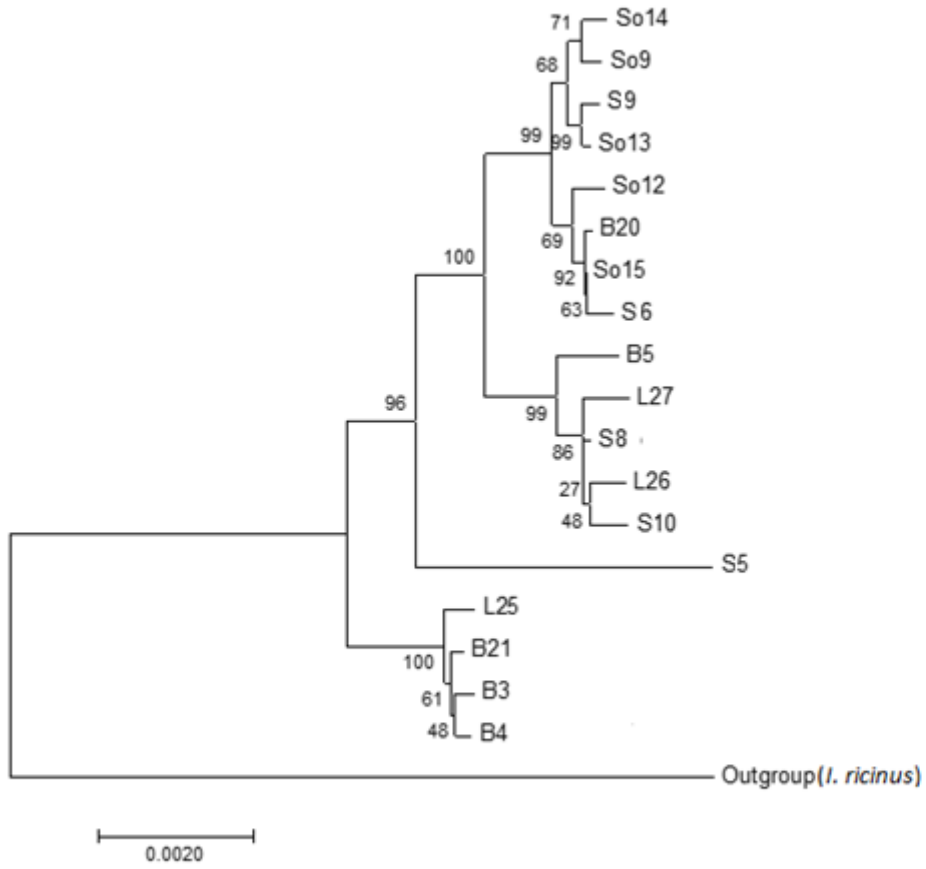


Figure B. Neighbour joining (NJ) tree created from the alignment of geographic isolates. Numbers represent bootstrap values and the number of bootstrap replications was set to 2000.

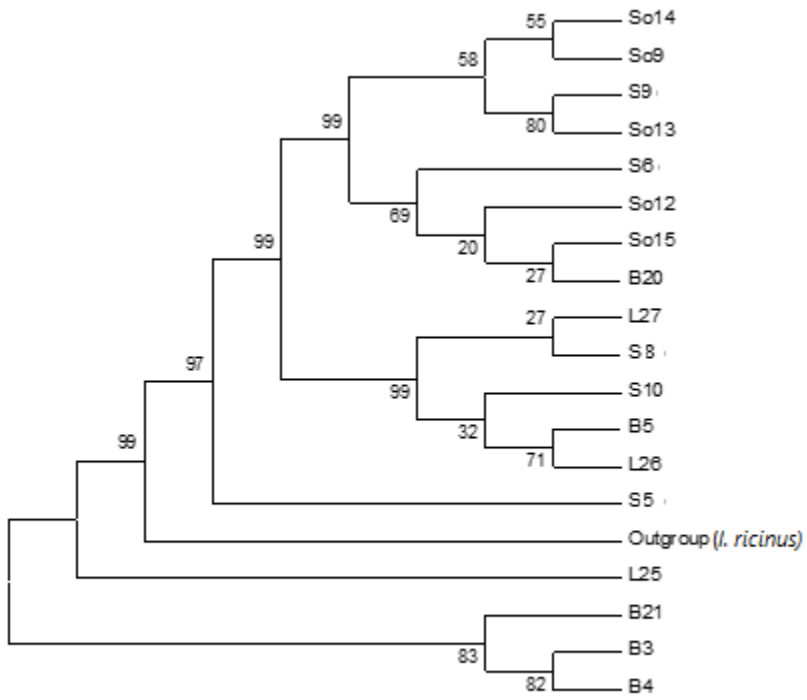


Figure C. Maximum parsimony (MP) tree based on the geographic isolate data set. Bootstrap values were set to 2000 replicates.

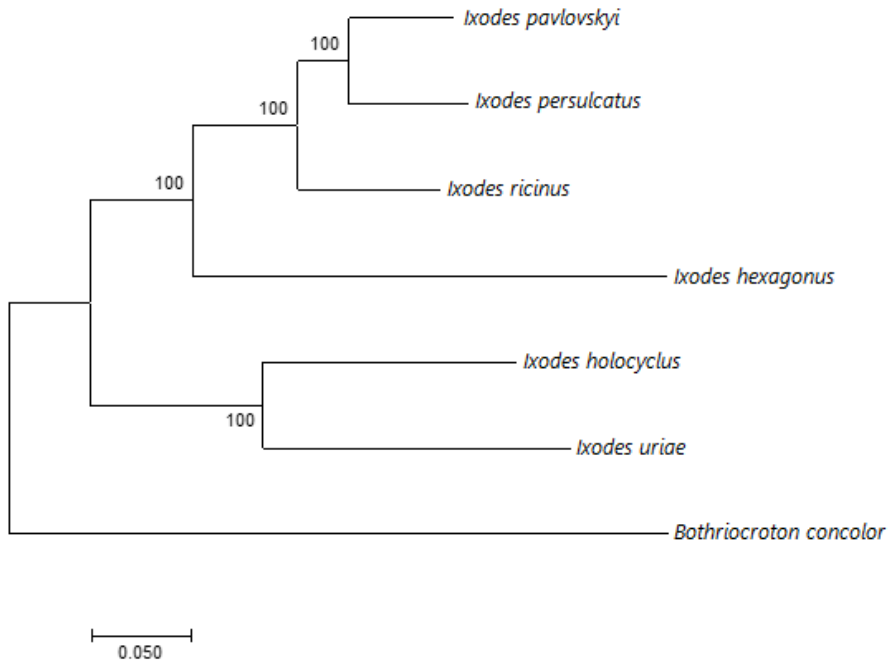


Figure D. Neighbour joining (NJ) tree constructed for the *Ixodes* species alignment. Bootstrap values were set to 2000 replicates.

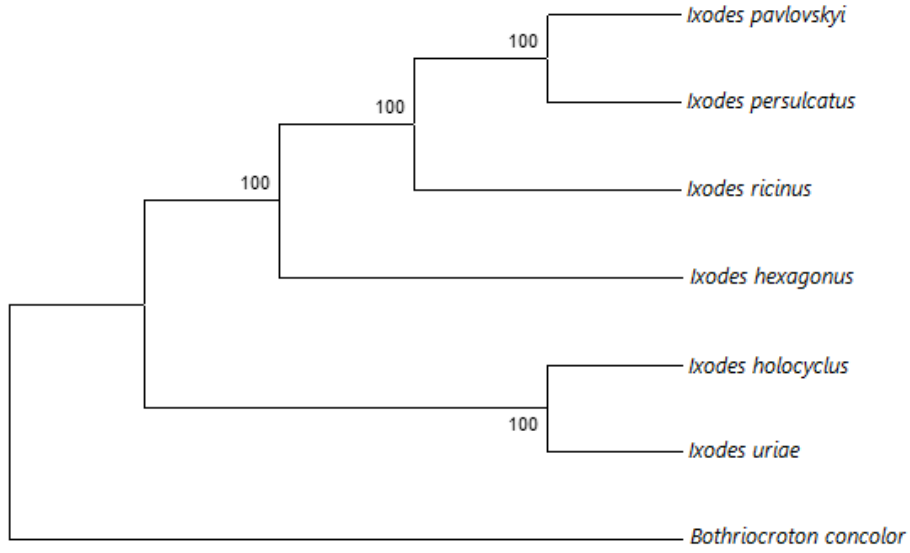


Figure E. Maximum parsimony constructed for the protein coding *Ixodes* species alignment. Bootstrap values were set to 2000 replicates.

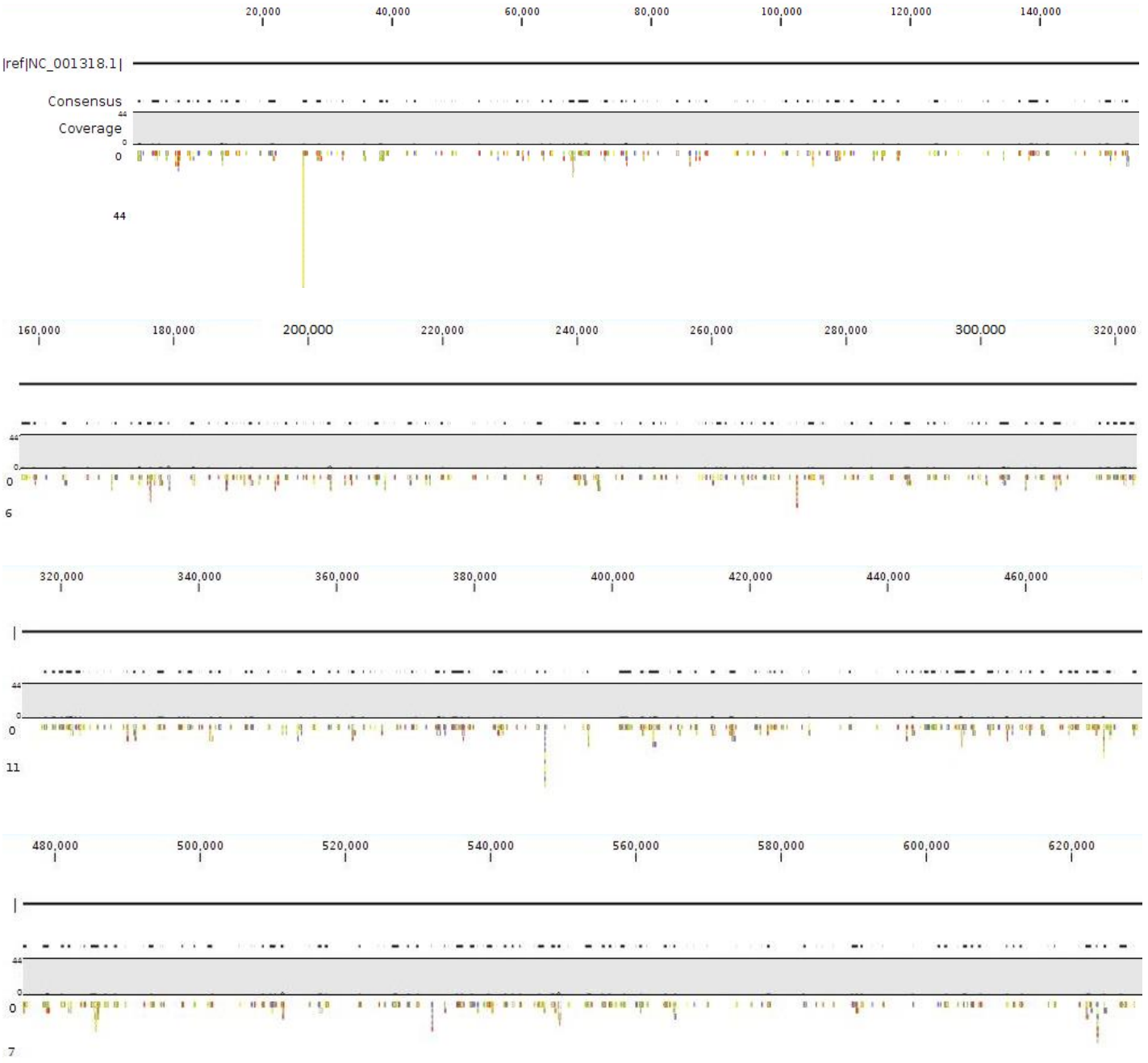

20,000
40,000
60,000
80,000
100,000
120,000
140,000
|ref|NC_001318.1|
Consensus
Coverage
44
0
0
44
160,000
180,000
200,000
220,000
240,000
260,000
280,000
300,000
320,000
44
0
0
6
320,000
340,000
360,000
380,000
400,000
420,000
440,000
460,000
44
0
0
11
480,000
500,000
520,000
540,000
560,000
580,000
600,000
620,000
44
0
0
7

Figure F. Graphical representation of the mapping of all B9AM reads against the complete genome of *Borrelia burgdorferi*. The values at the top, following the figure, represent the sequence length at each point. In each row the uppermost number to the right represents the coverage of the mapping, whilst the bottom value represents how many reads that are mapped in this exact segment of mapping.


360,000
380,000
400,000
420,000
440,000
460,000
480,000
480,000
500,000
520,000
540,000
560,000
580,000
600,000
600,000
620,000
640,000
660,000
680,000
700,000
720,000
720,000
740,000
760,000
780,000
800,000
820,000
840,000
840,000
860,000
880,000
900,000
920,000
940,000
960,000
960,000
980,000
1,000,000
1,020,000
1,040,000
1,060,000

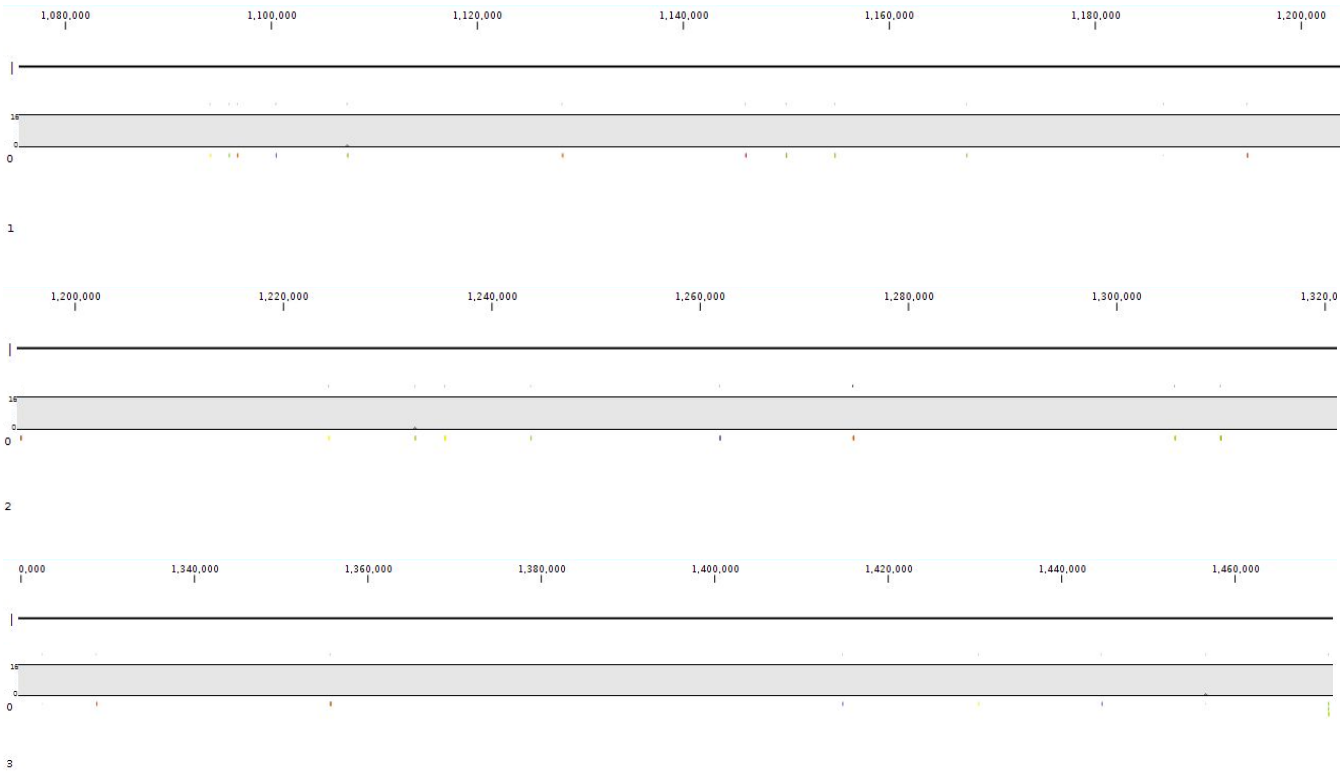
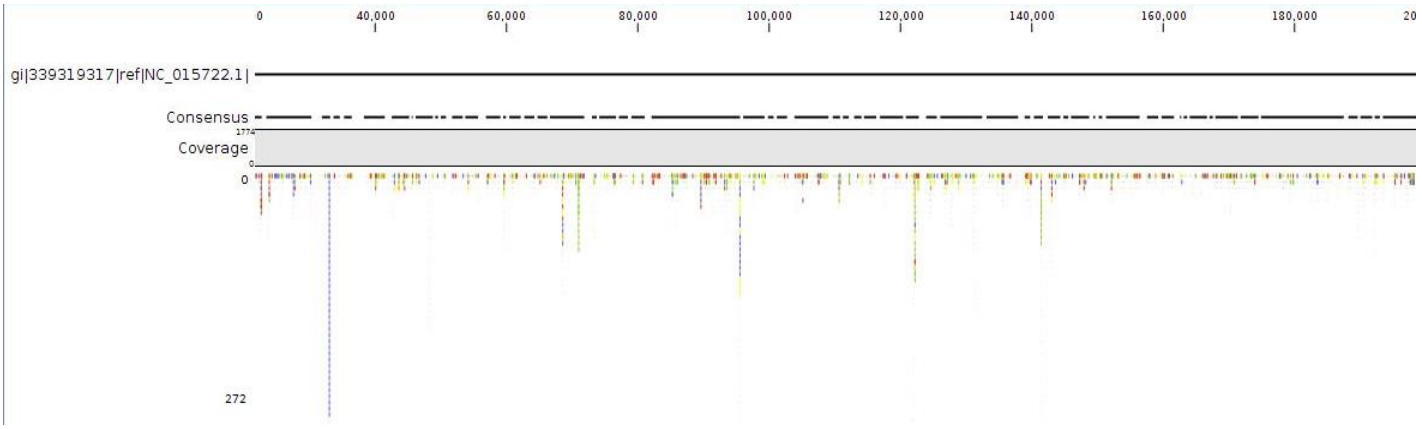
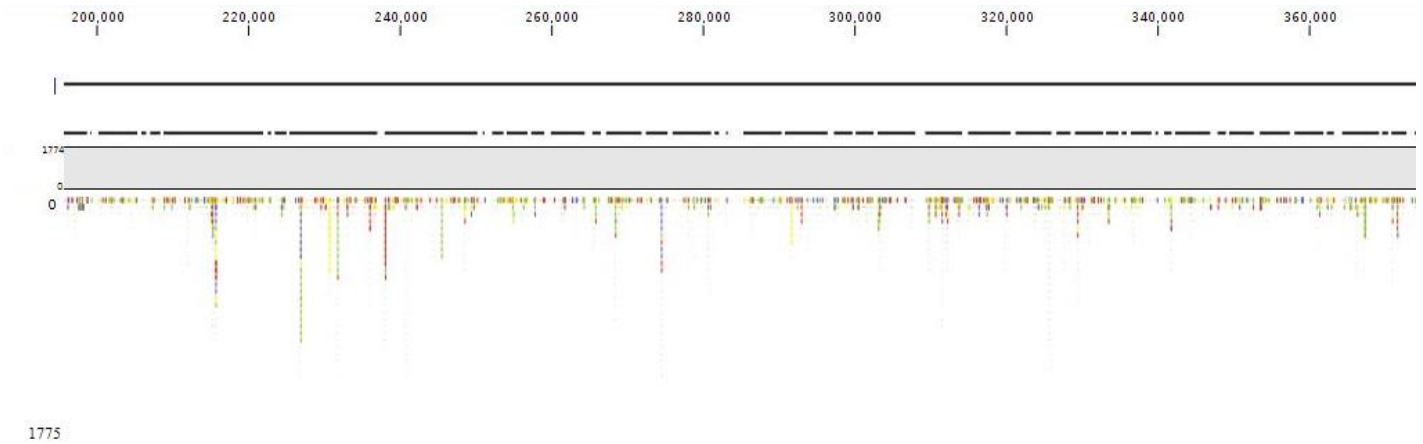
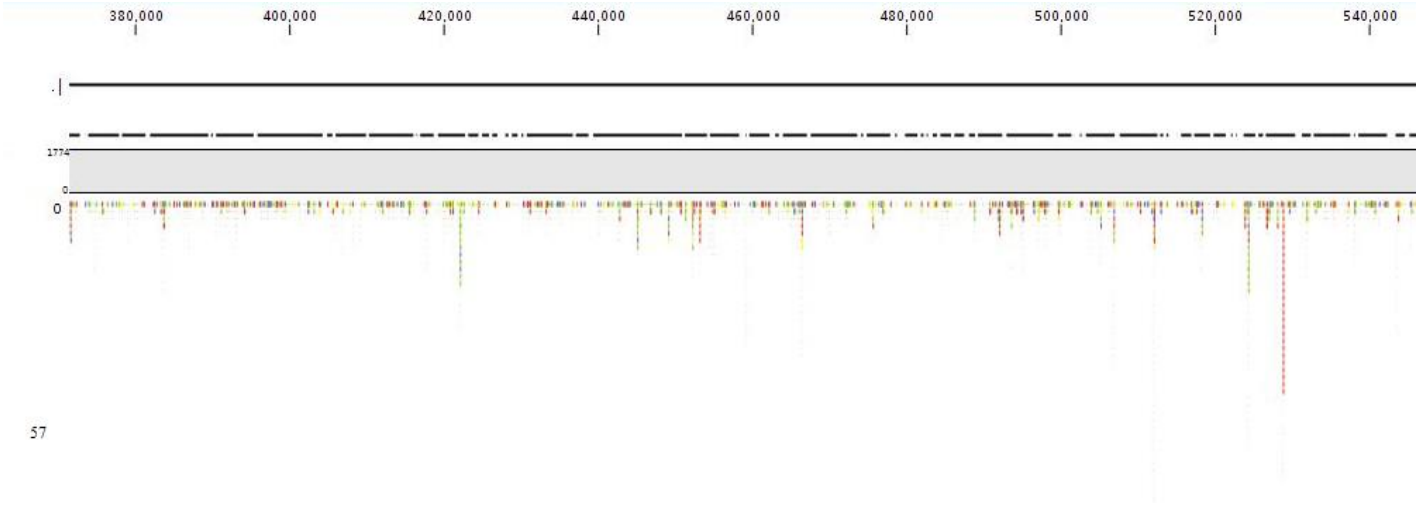


Figure G. Mapping of total DNA sequencing runs against the complete genome of *Anaplasma phagocytophilum*. The values at the top, following the figure, represent the sequence length at each point. In each row the uppermost number to the right represents the coverage of the mapping, whilst the bottom value represents how many reads that are mapped in this exact segment of mapping.

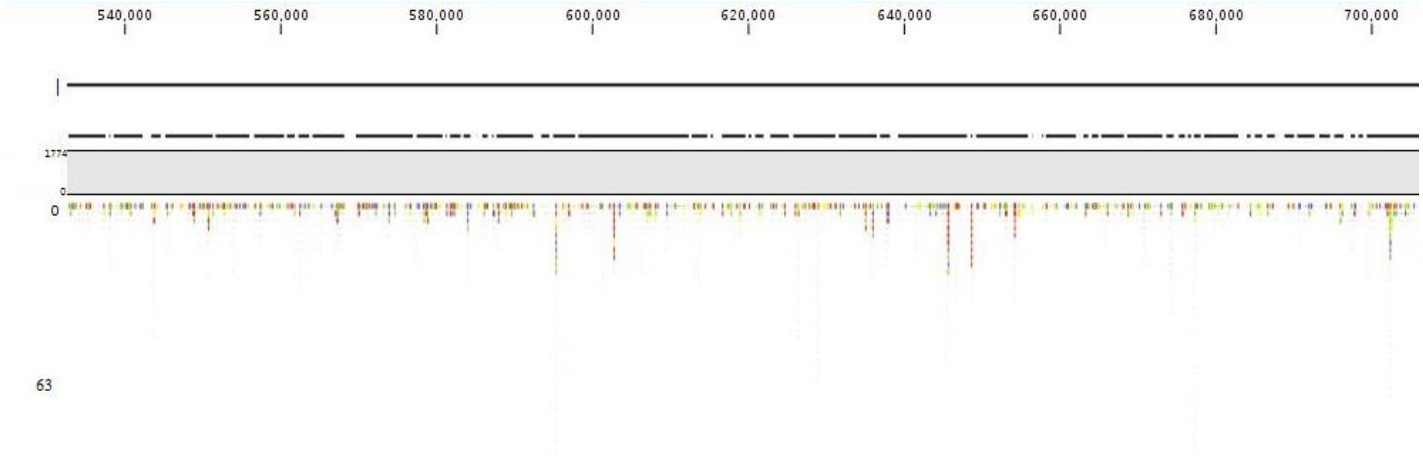



200,000
220,000
240,000
260,000
280,000
300,000
320,000
340,000
360,000
1774
0
0

1775


380,000
400,000
420,000
440,000
460,000
480,000
500,000
520,000
540,000
1774
0
0

57


540,000
560,000
580,000
600,000
620,000
640,000
660,000
680,000
700,000
1774
0
0

63

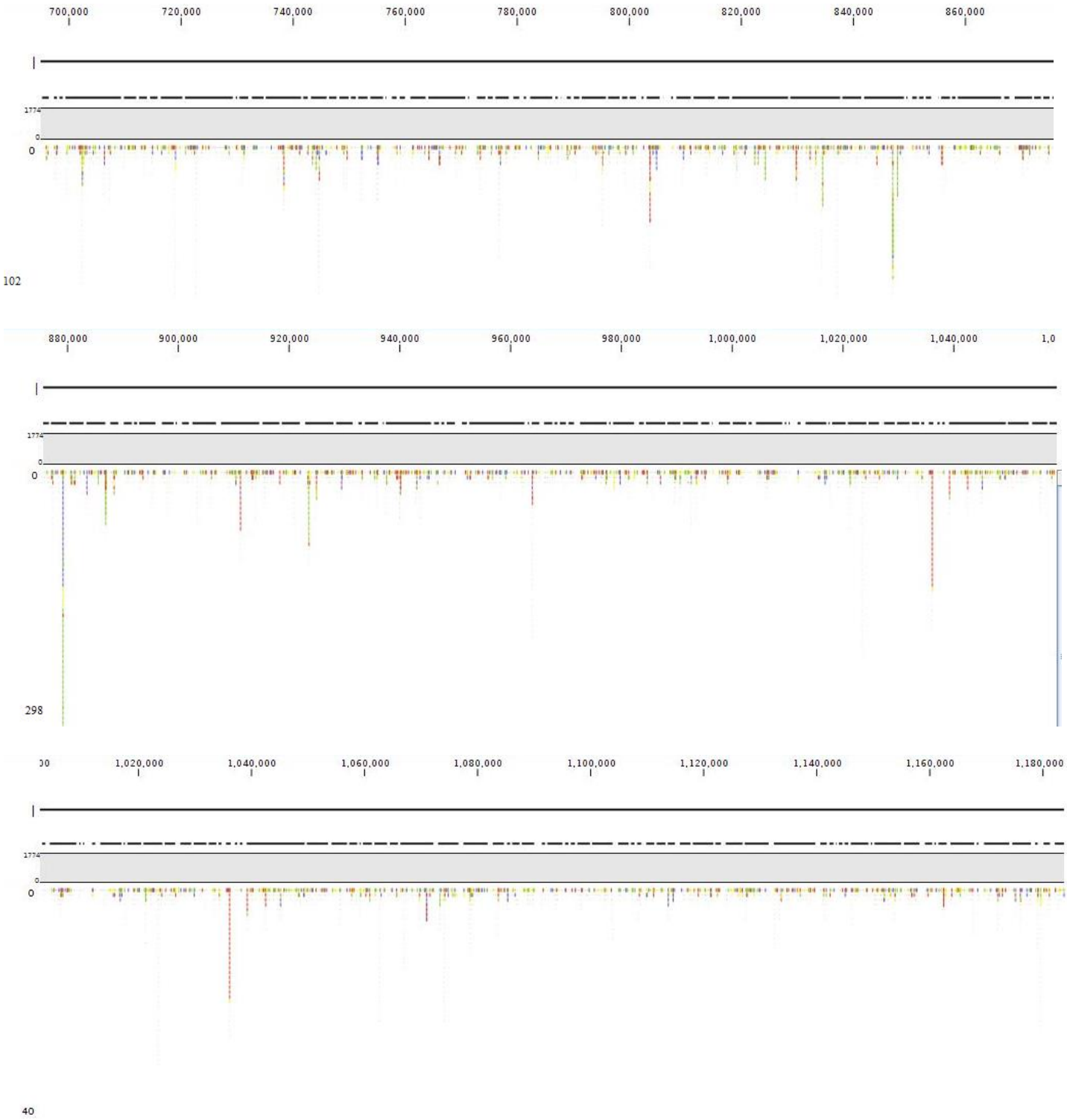


Figure H. Mapping of the total DNA sequencing runs against the complete genome of *Midichloria mitochondrii*. The values at the top, following the figure, represent the sequence length at each point. In each row the uppermost number to the right represents the coverage of the mapping, whilst the bottom value represents how many reads that are mapped in this exact segment of mapping.