Inland Norway University

Faculty of Applied Ecology, Agricultural Sciences and Biotechnology

Chamara Janaka Bandara

Master's Degree Thesis

Development of a whole genome sequencing method for Tick-borne encephalitis virus in low viraemic samples

Master's Degree in Experimental Biotechnology

2022

Conducted at the

NIPH

Norwegian Institute of Public Health

Consent to lending by University College Library YES ☐NO ☐

Consent to accessibility in digital archive Brage YES ☐NO ☐

Acknowledgement

I express my deep gratitude to all of my supervisors, Prof. Åshild Kristine Andreassen, Dr. Rose Vikse and Dr. Kristian Alfsnes at Norwegian Institute of Public Health (NIPH), Oslo, Norway, for the valuable guidance, supervisions and the kind support given throughout the research activities. I also thank Prof. Robert Wilson, Department of Biotechnology, Inland Norway University of Applied Sciences, Hamar, Norway for the guidance and the support given throughout the study time.

I like to express my sincere gratitude to Prof. William Rayan Easterday, University of Oslo, Norway and Dr. Kristian Alfsnes for the development of semiquantitative real-time PCR technique for the analysis and also for the Dr. Arnulf Soleng, Senior researcher, Department of Pest control, NIPH Oslo, Norway, for the providing of tick samples collected from the Larvik in the southern coastal region in Norway for the analysis.

I especially thank Mrs. Maria J. Diekmann, NIPH, Oslo, Norway, for giving me excellent support, the initial training the introduction to the laboratory facility during the research work. I also thank Mrs. Eileen Stanzen, Master student, Northwestern University of Norway, for providing of tick RNA extracts for the further analysis and Mrs. Hilde Elshaug, NIPH, Oslo, Norway, for sharing of reagents and chemicals.

In addition, I thank all the staff members at the Norwegian Institute of Public Health for providing constant support with all the facilities, including information technology facility, purchasing of Bioinformatics software, chemicals and reagent ordering and coordination with a short notice.

I sincerely thank my family members for their help and constant support during my master study period.

This Master's thesis is one of several larger projects funded by the Norwegian Ministry of Health and Care Services (B1710), TBFVnet (Project index number: 2018-1-0659 EEA Norwegian grant for regional cooperation) and NorthTick, InterReg VB North Sea Region Programme.

Abbreviations

aa	Amino acid
Acc. no.	Accession number
ANOVA	Analysis of variance
BLAST	Basic local alignment search tool
BLASTn	Nucleotide Basic local alignment search tool
C/C-protein	Capsid/Capsid-protein
Cat. number	Catalogue number
cDNA	Complementary DNA
CDS	Coding sequence
CNS	Central nervous system
Cryo-EM	Cryo-electron microscopy
CSV	Comma-separated values
Ct	Threshold cycle
Da	Dalton
DNA	Deoxyribonucleic acid
DNase	Deoxyribonuclease
dsDNA	Double stranded DNA
E/ E-protein	Envelope/ Envelope-protein
EEA	European economic area
ELISA	Enzyme-linked immunosorbent assay
ER	Endoplasmic reticulum
Et-Br	Ethidium bromide
EU	European union
F	Forward
F/R	Forward and reversed
FBS	Fetal-bovine serum
GPS	Global Positioning System
ICTV	International Committee on Taxonomy of viruses
IDE	Integrated development environment
IgG	Immunoglobulin G
IgM	Immunoglobulin M

Kb	Kilobases
LIV	Looping ill virus
M/ M-protein	Membrane/ Membrane-protein
Mb	Megabases
MegaBLAST	Mega-Basic local alignment search tool
Method I	Semiquantitative real-time PCR developed by Prof. Åshild Kristine Andreassen
Method II	Semiquantitative real-time PCR developed by Dr. Ryan Easterday and Dr. Kristian Alfsnes
MEM	Minimal essential medium
MIR	Minimum Infection Rate
MSA	Multiple sequence alignment
NCBI	National Center for Biotechnology Information
NCR/NCRs	Noncoding region/regions
ng	Nanograms
NIPH	Norwegian Institute of Public Health
nm	Nanometre
NTC/NTCs	Non-templated control/controls
nt	Nucleotide
ONT	Oxford Nanopore Technology
ORF	Open reading frame
PCR	Polymerase chain reaction
pH	Proton ion concentration in (-log10)
poly(A)	Polyadenylated
prM/prM-protein	Pre-membrane protein
p-value	Probability value
R	Reversed
RNA	Ribonucleic aid
RNase	Ribonuclease
RNase H	Ribonuclease H
rRNA	Ribosomal ribonucleic acid
RT-LAMP	Reverse transcription-loop-mediated isothermal amplification

RT-PCR	Reverse transcription-polymerase chain reaction
RT-qPCR	Reverse transcription-Realtime polymerase chain reaction
SAM/SAMs	Sequence aligned map/maps
sDNA	Single stranded DNA
SE	Soluble envelope protein
SLA	Stem loop element-A
SLB	Stem loop element-B
SNP	Single nucleotide polymorphism
TBE	Tick-Borne encephalitis
TBEV	Tick-Borne encephalitis virus
TBEV-Bkl	Tick-Borne encephalitis-Bikalayan subtype
TBEV-Eur	Tick-Borne encephalitis-European subtype
TBEV-FE	Tick-Borne encephalitis-Far-eastern subtype
TBEV-Him	Tick-Borne encephalitis-Himalayan subtype
TBEV-Sib	Tick-Borne encephalitis-Siberian subtype
Tm	Melting temperature
UTR	Untranslated region
V	Voltage
v/v	Ratio based on volume
w/v	Ratio based on weight and the volume
WG	Whole genome
WGS	Whole genome sequencing
λ	Wavelength
μL	Microliter
μM	Micromolar

Table of Contents

ACKNOWLEDGEMENT 3

ABBREVIATIONS 4

TABLE OF CONTENTS 7

ABSTRACT 11

1. INTRODUCTION 12

1.1 TICK-BORNE ENCEPHALITIS 12

1.2 TICK-BORNE ENCEPHALITIS VIRUS 13

1.2.1 Viral genome and the RNA secondary structure 14

1.2.2 Subtypes and the quasispecies 16

1.2.3 Proteins encoded by the TBEV RNA 17

1.2.4 Replication strategy 19

1.3 VECTORS AND THE TRANSMISSION 20

1.4 DISTRIBUTION AND PREVALENCE 22

1.4.1 Distribution and prevalence in Norway 24

1.5 EPIDEMIOLOGY IN EUROPE 25

1.6 DETECTION, CONTROL AND VACCINATION 26

1.7 PREVIOUS STUDIES ON WHOLE GENOME SEQUENCING OF TBEV 28

1.8 MOLECULAR BIOLOGY TECHNIQUES EMPLOYED IN THIS STUDY 28

1.8.1 Homogenization and the total RNA extraction 28

1.8.2 Reverse transcription polymerase chain reaction (RT-PCR) 29

1.8.3 Polymerase chain reaction (PCR) 29

1.8.4 Real-time PCR for the detection of viral specific cDNA 30

1.8.5 DNA barcoding, library preparation and Oxford Nanopore sequencing 30

1.8.6 Bioinformatics and data analysis 32

1.9 AIM OF THE STUDY 34

2. MATERIAL AND METHODS 35

2.1 EXPERIMENTAL DESIGN 35

2.2 STUDY MATERIALS 38

2.3 QUANTIFICATION OF RNA USING A SPECTROPHOTOMETRIC METHOD 38

2.4 REVERSE TRANSCRIPTION AND SEMIQUANTITATIVE REAL-TIME PCR ANALYSIS OF TBEV-HOCHOSTERWITZ RNA 38
2.5 SEMIQUANTITATIVE REAL-TIME PCR WITH CDNA DERIVED THROUGH THE RANDOM HEXAMERS AND TBEV SPECIFIC PRIMERS 40
2.6 OPTIMIZATION OF RT-PCR AND PCR CONDITIONS 41
2.7 EVALUATION OF EFFECT OF TEMPERATURE OF PCR AMPLIFICATION OF TBEV-HOCHOSTERWITZ CDNA AGAINST EACH PRIMER PAIR 42
2.8 AGAROSE GEL ELECTROPHORESIS 42
2.9 PCR PRODUCT PURIFICATIONS 43
2.10 QUANTIFICATION OF DNA USING QUBIT DNA ANALYSER 43
2.11 RT-PCR ON TBEV-HOCHOSTERWITZ STRAIN RNA FOLLOWED BY PCR AMPLIFICATION AGAINST DIFFERENT OVERLAPPING PRIMER PAIR SCHEMES 44
2.12 DNA LIBRARY PREPARATION AND OXFORD NANOPORE SEQUENCING 44
2.13 REFERENCE-BASED OXFORD NANOPORE READS MAPPING AND DATA ANALYSIS . 45
2.14 *DE NOVO* SEQUENCE ASSEMBLY FOLLOWED BY REFERENCE-BASED SEQUENCE MAPPING 46
2.15 EVALUATION OF THE QUALITY OF MAPPED GENOMES 47
2.16 IDENTIFICATION OF POORLY PERFORMED PRIMER PAIRS THROUGH SEQUENCE ASSEMBLY MAP (SAM) FILES 47
2.17 DESIGNING NEW OVERLAPPING PCR PRIMERS 48
2.18 PCR AMPLIFICATION OF TBEV-HOCHOSTERWITZ CDNA AGAINST EACH NEW PRIMER PAIR 49
2.19 EVALUATION OF OXFORD NANOPORE SEQUENCED DATA DERIVED THROUGH THE COMBINATIONS OF DIFFERENT OVERLAPPING PRIMER SCHEMES 49
2.20 TOTAL RNA EXTRACTIONS FROM NYMPH TICK SAMPLES 50
2.21 REVERSE TRANSCRIPTION OF TOTAL RNA FROM TICK SAMPLES 50
2.22 REALTIME SEMIQUANTITATIVE PCR ON TICK CDNA AND THE DETECTION OF TBEV CDNA 51
2.23 PCR AMPLIFICATION OF CDNA DERIVED FROM TICK SAMPLES, PCR PRODUCT PURIFICATION AND DSDNA QUANTIFICATIONS 52
2.24 DNA LIBRARY PREPARATION, OXFORD NANOPORE SEQUENCING AND REFERENCE-BASED ASSEMBLING 52

2.25 OPTIMIZATION OF RT-PCR, PCR, DNA LIBRARY PREPARATIONS FOLLOWED BY OXFORD NANOPORE SEQUENCING AND REFERENCE-BASED ASSEMBLING OF TBEV POSITIVE TICK SAMPLES 53
2.26 IDENTIFICATION OF TBEV VIRAL SUBTYPES 54
2.27 IDENTIFICATION OF TICK SPECIES 54

3. RESULTS 56

3.1 TOTAL RNA QUANTITIES IN THE PRESERVED RNA EXTRACTS FROM THE CELL CULTURES 56
3.2 RELATIVE QUANTITIES OF TEBV-HOCHOSTERWITZ RNA 57
3.3 EFFECTS OF TYPE OF PRIMERS AND REVERSE TRANSCRIPTASE ON THE CDNA SYNTHESIS 58
3.4 EFFECTS OF PRIMER PAIR CONCENTRATIONS ON THE PCR AMPLIFICATION OF TBEV-HOCHOSTERWITZ CDNA 59
3.5 EFFECT OF TEMPERATURE ON THE PCR AMPLIFICATION OF TBEV-HOCHOSTERWITZ CDNA 61
3.6 EFFECT OF TOTAL RNA FROM TICK SAMPLES 63
3.7 QUALITY OF THE RECONSTRUCTED TBEV-HOCHOSTERWITZ GENOMES BY THE INITIAL PRIMER SCHEMES 66
3.8 PHYLOGENETIC TREE ANALYSIS FOR THE RECONSTRUCTED TBEV-HOCHOSTERWITZ SEQUENCES 69
3.9 POORLY PERFORMED PRIMERS 70
3.10 NEW OVERLAPPING PCR PRIMER PAIRS 72
3.11 EFFECT OF PCR TEMPERATURE ON NEW OVERLAPPING PRIMER PAIRS 75
3.12 DNA CONCENTRATIONS OF PCR PRODUCTS DERIVED THROUGH EACH OVERLAPPING PRIMER PAIR SCHEME 76
3.13 OPTIMUM PCR CONDITIONS FOR THE WHOLE GENOME SEQUENCING OF THE LOW TBE VIRAEMIC SAMPLES 77
3.13.1 Calculated primer pair efficiency in each primer pair scheme 77
3.13.2 Per-base sequencing depth of reconstructed TBEV-Hochosterwitz genomes by each primer pair scheme 78
3.13.3 Quality of reconstructed TBEV-Hochosterwitz genomes by overlapping primer pair schemes 79
3.14 IDENTIFIED TBEV POSITIVE TICK SAMPLES BY THE RT-QPCR METHODS 83

3.15 OXFORD NANOPORE SEQUENCE DATA ANALYSIS AND THE RECONSTRUCTED TBEV GENOMES FROM TICK SAMPLES 85
3.16 IDENTIFIED TBEV VIRAL SUBTYPES 88
3.17 SPECIES LEVEL OF IDENTIFICATION OF TBEV POSITIVE TICK SPECIES 92

4. DISCUSSION 94

4.1 THE LIMITATIONS ASSOCIATED WITH THE ALREADY EXISTING PCR METHODS AT NIPH FOR THE AMPLIFICATION OF THE WHOLE GENOME OF TBEV 94
4.2 EFFECTS OF TEMPERATURE AND THE PRIMER CONCENTRATION ON THE PCR AMPLIFICATION OF WHOLE GENOME OF TBEV 96
4.3 EFFECTS OF TOTAL RNA FROM TICK SAMPLES AND THE DILUTION FACTOR ON UNSPECIFIC AMPLIFICATIONS 97
4.4 CHALLENGES IN WHOLE GENOME MAPPING OF TBEV TO OXFORD NANOPORE SEQUENCE DATA 99
4.5 RELATIVE ACCURACY AND THE RELIABILITY OF THE SEMI QUANTITATIVE REAL-TIME PCR TECHNIQUES EMPLOYED IN THIS STUDY 101
4.6 EFFECT OF RANDOM PRIMERS AND TBEV SPECIFIC PRIMERS ON RT-PCR AND SUBSEQUENT WHOLE GENOME AMPLIFICATION OF TBEV 102
4.7 POORLY AMPLIFIED REGIONS OF TBEV-HOCHOSTERWITZ GENOME 104
4.8 CALCULATED OF PRIMER PAIR EFFICIENCIES AND THE METHOD DEVELOPED.. 105
4.9 PREVALENCE, VIRAEMIC CONTENT AND THE SEQUENCE DIVERSITY OF TBEV IN TICK SAMPLES FROM LARVIK 106
4.10 NEUROINVASIVE ABILITY AND THE VIRULENCE FACTORS OF THE IDENTIFIED TBEV STRAINS FROM TICK SAMPLES 110
4.11 VECTOR TICK SPECIES IDENTIFIED IN THE SITE OF STUDY 111
4.12 REPORTED STUDIES ON SUCCESSFUL ATTEMPTS OF WHOLE GENOME SEQUENCING OF TBEV AND A COMPARISON 112
4.13 IMPROVED PERFORMANCE, ASSOCIATED LIMITATIONS AND THE FURTHER DEVELOPMENT WITH THE PCR TECHNIQUES 114

5. CONCLUSION 119

6. REFERENCES 120

7. ANNEXURE I 136
8. ANNEXURE II 209
9. ANNEXURE III 210

Abstract

Due to the rapid increase in Tick-Borne Encephalitis (TBE) viral infections in Norway and in the Western Europe, attentions have been drawn to develop rapid screening and whole genome sequencing methods directly from the tick samples. This would facilitate the identification of Tick-Borne Encephalitis virus (TBEV) endemic area, to predict the neuroinvasive ability and the disease severity of the TBEV natural foci.

In this study, the optimization of RT-PCR, PCR and the whole genome sequencing with Oxford Nanopore technology (ONT) were carried out using tenfold dilutions of total RNA extract of Vero E6 cell culture sample incubated with TBEV-Hochosterwitz strain (NCBI Acc. No. MT311861). Total of 172 tick samples, including nymph pools of ten nymphs per pool (76) and adult males (50) and adult females (46) were collected from Larvik, Norway and separately subjected to total RNA extractions and evaluated for TBEV using two separate semiquantitative real-time PCR techniques. Whole genome amplifications and Oxford Nanopore sequencing and reference-based mapping were carried out for the TBEV positive tick RNA samples against the best optimized primer schemes. Phylogenetic tree analysis was carried out for the TBEV genome sequence fragments to identify the evolutionary closest TBEV subtype. In addition, TBEV positive tick species were identified using 18S-rRNA sequences found in Oxford Nanopore reads by phylogenetic tree analysis.

TBEV positive five nymph tick pools, five female and three male adult ticks were identified and the prevalence of TBEV at the study site found to be 0.68 ± 0.30%, 10.90% and 6.00%, respectively. Primer schemes {[(2E+2Q)+(2F+2R)] and [(G+S)+(H+T)]} were identified as suitable for the whole genome amplification of TBEV directly form the tick samples. The maximum average sequencing breadth over 20 reads and the sequencing coverage were below the 13.6% and 50.56%, respectively in the reconstructed TBEV genomes. All the TBEV genomic sequence fragments shown to have a close evolutionary relationship to the TBEV-European subtype and the TBEV positive ticks were identified as *Ixodes ricinus*.

The optimized primer schemes exhibited whole genome amplification ability over the low viraemic TBEV positive tick samples and having sequencing gaps in the reconstructed genomes. Further optimizations are necessary to improve the quality of the reconstructed TBEV genomes directly from the tick samples and which will lead to improve the accuracy in the predictions of the neuroinvasive ability of the TBEV strains.

1. Introduction

Viral infections and their impacts on humans have shaped the human evolution and the human development in many ways. Its' not only affects the health of individuals, but also affect the society through the memories and scars left throughout the past time periods (Leal & Zanotto, 2000). The frequency of new viral infections and the frequency of pandemics have increased during the last four decades (Roychoudhury *et al.,* 2020), mainly due to the humans and their interactions with the nature being intensified by the main human activities linked with deforestations, shifting of human settlements closer to the wildlife, backed with the scarcity of arable lands and also the population growth. In addition to that it is also found that the changes in human consumption patterns during the past time period has a huge effect on it (McNeely, 2021).

The majority of the viruses identified since the first official documentation in 1800s, are associated with Chordata and the Arthropoda, this has strong potential for humans to become a breading host of a virus (Harvey & Holmes, 2022). One of such diverse family of viruses is the family *Flaviviridae,* consist of three main genera, *Flavivirus, Hepacivirus* and *Pestivirus.* The genus *Flavivirus* is the largest among the others, exhibiting 53 different viruses (Chene, 2012), the majority of them are arboviruses and causing mild to deadly infections in humans and accounted for up to 400 million globally every year (Pierson & Diamond, 2020).

1.1 Tick-borne encephalitis

Tick borne encephalitis is arboviral zoonotic disease caused by the Tick-Borne Encephalitis virus (TBEV) belonging in to Genus *Flavivirus,* leading to severe acute and chronic neurological infections in humans and other vertebrates (Conze *et al.,* 2021b). The severity of the disease to the other vertebrates such as dogs, horses, monkeys, sheep and goats, known to be very rare (Michelitsch *et al.,* 2019). The viral infection caused by the TBEV infected tick bite or through the consumption of unpasteurized cheese or milk, processed through the TBEV infected goats or cows may lead to sever disease (Conze *et al.,* 2021a). The symptoms appearing in two stages, the initial unspecific symptoms appear five to six days after the infection. The first stage symptoms are fever and most likely the same as an influenza infection. The secondary stage symptoms are related to the inflammation in the central nervous system, appears after six to seven days of the appearing the first unspecific symptoms (Gerhard

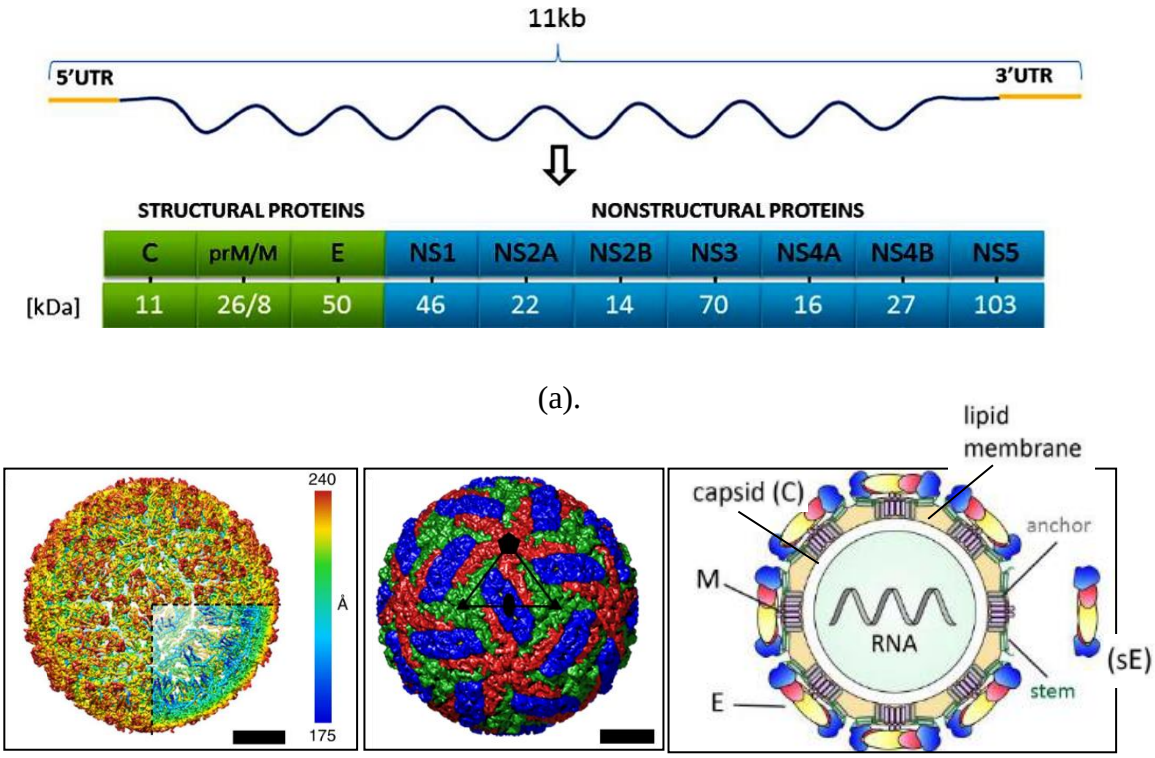
et al., 2021). It is found that most of the TBEV infections are asymptomatic and does not cause severe neurological infections and the majority of the cases are known to be unreported. This was identified through a study carried out by the participants of nature managements workers in Netherlands (Hofhuis *et al.,* 2021). The severe meningoencephalitis cases are reported for one-third of the total TBEV infections and causing long-term sequelae. It is found to have low fatality rate, which is about less than 2% to the total number of TBEV cases reported in the Europe (Jääskeläinen *et al.,* 2010). Disease symptoms and the severity can be different from person to person and it depends on the infected subtype of the TBEV (Michelitsch *et al.,* 2019). It is found to be three main subtypes of TBEV; Far-Eastern subtype is the most severe TBEV subtypes. It has been identified that the Siberian subtype of the TBEV is leading to more chronic infection than that of European TBEV subtype (Jääskeläinen *et al.,* 2010). The highest mortality rates have been reported maximum between 20 - 40% of the total TBEV cases reported in an outbreak in Russia (Gritsun *et al.,* 2003). Infections caused by the European subtype of TBEV often induce a biphasic clinical course, while the infections coursed by the Far Eastern and Siberian subtypes leading to a monophasic clinical course. Therefore it has shown the importance of identification of TBE viral subtypes in the epidemiological point of view (Hayasaka *et al.,* 2013).

1.2 Tick-borne encephalitis virus

Tick-Borne Encephalitis virus is an arbovirus belonging into the genus *Flavivirus* and one of the Tick-borne flaviviruses causing infection in mammals. It also belongs to the class of nucleocapsid enveloped viruses having an icosahedral structural geometry with pseudo-T-3-symmetry. The diameter of the virus is 50 nm and the internal structure is arranged into three distinct sections, leaving the positive sense RNA at the centre and covered by the nuclear capsid surrounded with the membrane protein and the outermost layer of envelope proteins (Figure 1.1. d) (Füzik *et al.,* 2018).

1.2.1 Viral genome and the RNA secondary structure

The TBE viral genome is about ~11 Kb long and which is a positive sense single stranded RNA with single open reading frame flanked by 5' and 3' noncoding regions (NCRs). The coding region of the TBEV genome translated by the host enzymes and produce the 3400 amino acid (aa) long polyprotein sequence.



(b). (c). (d).

Figure 1.1. Schematic representation of the TBEV genome (~ 11 Kb) and its organization, ORF region (middle) flanked by 5' and 3' UTR regions. (a) The host translation machinery translates the coding sequence (CDS) region to a 3400 aa. long polypeptide chain consists with three structural proteins and seven non-structural proteins (Gerhard et al., 2021). (b) Cryo-electron microscopic density map of TBE virion, coloured according to the distance from the centre of the virion. Lower right section shows the interior transmembrane helixes of membrane proteins and the envelope proteins and their arrangements by removing the particles of the virion, (c) Cryo-electron microscopic density map of the surface of TBE virion highlighting the envelope proteins within each icosahedral asymmetric unit (Red, Green and Blue) and the black triangle shows the borders of the one icosahedral subunit (Füzik et al., 2018). (d) Schematic representation of the TBEV viral structure, innermost the nucleocapsid filled with single stand positive sense RNA, M: Membrane protein helixes, E: Envelope protein and SE: soluble envelope protein (Gerhard et al., 2021).

The formed polyprotein is cleaved by both host and the viral proteases to form eleven different proteins, three structural proteins [capsid (C), pre-membrane/membrane (M) and envelope(E)]

and seven non-structural functional proteins (NS1, NS2A, NS2B, NS3, NS4A, NS4B and NS5) (Yun *et al.*, 2011) (Figure 1.1.).

Flavivirus RNA secondary structure plays a crucial role in the viral replication where it functions as the main driving mechanism of the RNA replication, translation and provide the protection against the RNA degradation. In TBEV and other flaviviruses, various functional elements have been identified from the 5'UTR and 3'UTR regions of the flavivirus viral RNA genomes where these regions are known to be conserved among the species of the flaviviruses (Brinton & Dispoto, 1988 & Gritsun & Gould, 2007). It has been identified the stem loop element-A (SLA, Figure 1.2.) of at the 5' end of the flaviviral RNA involved in the viral replication through the recognition as a replication initiation position for the RNA polymerase and for the RNA dependent-RNA polymerase. In addition to the RNA binding recognition, it has been identified that the SLA structure is involved in the capping of the RNA during the RNA synthesis by directing the addition of a cap at the 5' end through the host capping enzymes. The RNA sequence elements corresponding to the SLA region is preserved all across the Flaviviruses and have been identified to be playing a major role in the viral replication. The second structural element which involved in the RNA replication is the stem-loop element-B (SLB, Figure 1.2.), where it has been identified to be involved in the translation initiation by bearing the translation initiation codon in the stem portion of the stem loop.

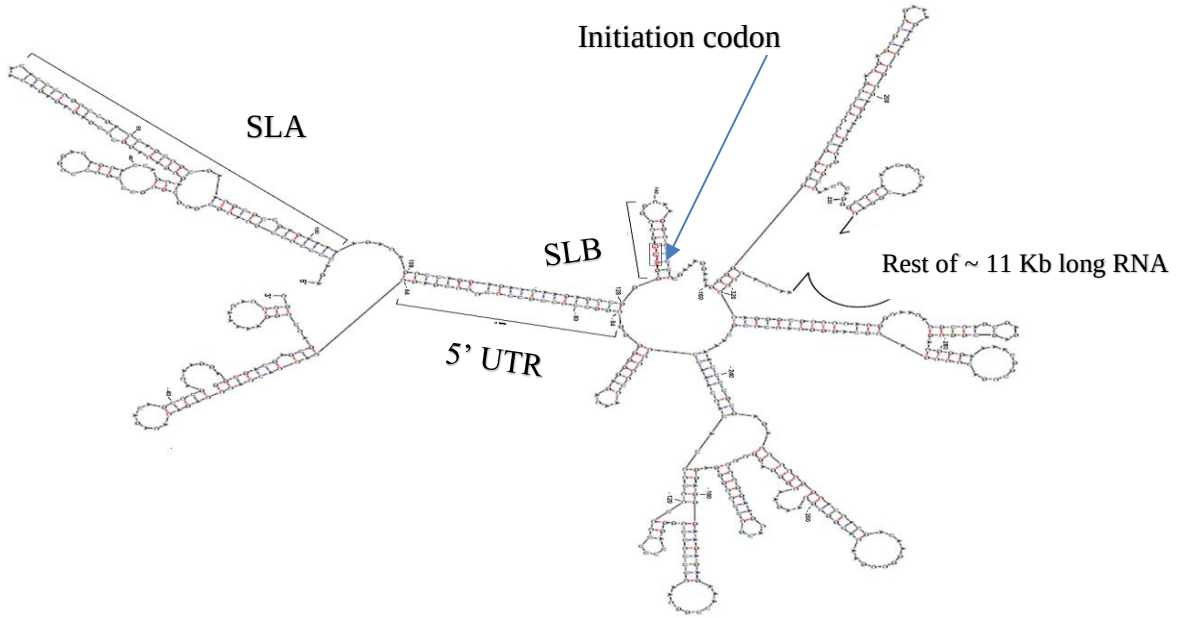


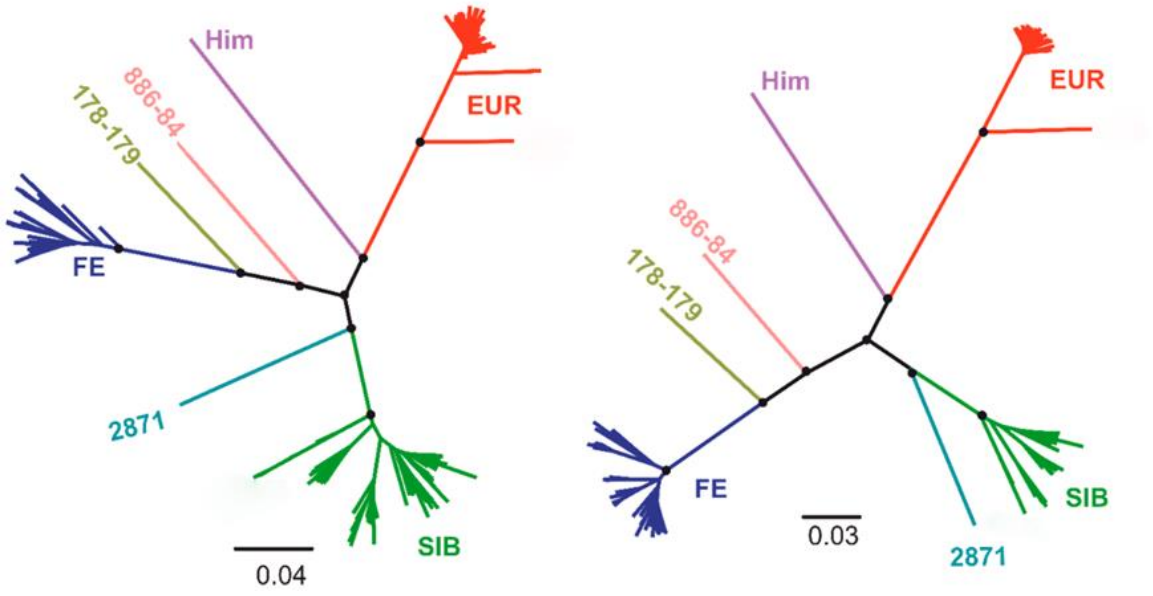
Figure 1.2. Predicted secondary structure exhibiting 5' and 3'-untranslated regions (UTR) of the ~11 Kb long positive sense TBEV genomic RNA with stem loop element-A (SLA) and stem loop element-B (SLB) (Belikov et al., 2014).

1.2.2 Subtypes and the quasispecies

According to the well-established categorization, three main subtypes of TBEV have been identified based on the phylogenetic analysis of the envelope protein region. These subtypes are European (TBEV-Eur), Far-Eastern (TBEV-FE) and the Siberian (TBEV- Sib) (Kozlova *et al.,* 2018) and antigenically and genetically related to each other. The distribution of the subtypes around the Eurasia is well established to specific regions. Western subtype is endemic to the central Europe, whereas the Far-Eastern subtype is more towards the east Asia, known to be identified in China and Japan and in the Eastern parts of Russia. The Siberian subtype is mainly identified from the regions of Eastern Russia and also in western Siberia (Yun *et al.,* 2011). Four distinct lineage from TBEV-Sib alone have been identified as Zausaev, Vasilchenko, Baltic and Obskaya in Russia (Tkachev *et al.,* 2020).

A recent study published in 2020, found that, according to the genetic distance criterion based on the TBEV envelope protein, TBEV variants should be divided into seven different subtypes namely, "TBEV-Eur, TBEV-Sib, TBEV-FE, TBEV-2871, TBEV-Him, TBEV-178-79 (TBEV-Bkl-1), and TBEV-886-84 (TBEV-Bkl-2)" (Figure 1.3.). The same study explains the complexity regarding the classification of the TBEV variants into subtypes, with the new divergent TBEV strains identified from Siberian near Baikal, Obskaya and Tibetan Highlands are considered to have deviated to be classified as subtypes. According to the International Committee on Taxonomy of Viruses (ICTV), that the phylogenetic alone is not sufficient to distinguish one specific virus from the other, but it should also need to be considered about the factors such as the association with the host, vector, disease and also geographical distribution. The same study showed that, phylogenetically, the genetic distance of the TBEV-Eur subtype is more closer to the Louping ill virus (LIV) than the other TBEV subtypes (Deviatkin *et al.,* 2020).

In addition to the highly variable subtypes found in the TBEV, it also reports the existence of the quasispecies in natural foci identified in two studies carried out in 2014 and 2017 years for the Norwegian Mandal 2009 and the Swedish Saringe 2009 TBEV variants. In the 2014 study, a depth analysis of the 3'- NRC region viral genome of the Saringe 2009, revealed that varying lengths of poly-(A) regions, confirming the existence of the quasispecies of the TBEV (Asghar *et al.,* 2014). In the subsequent deep sequencing study on the Mandal 2009 sequences was identified to have 40% single nucleotide polymorphism (SNP) among the quasispecies (Asghar *et al.,* 2017).

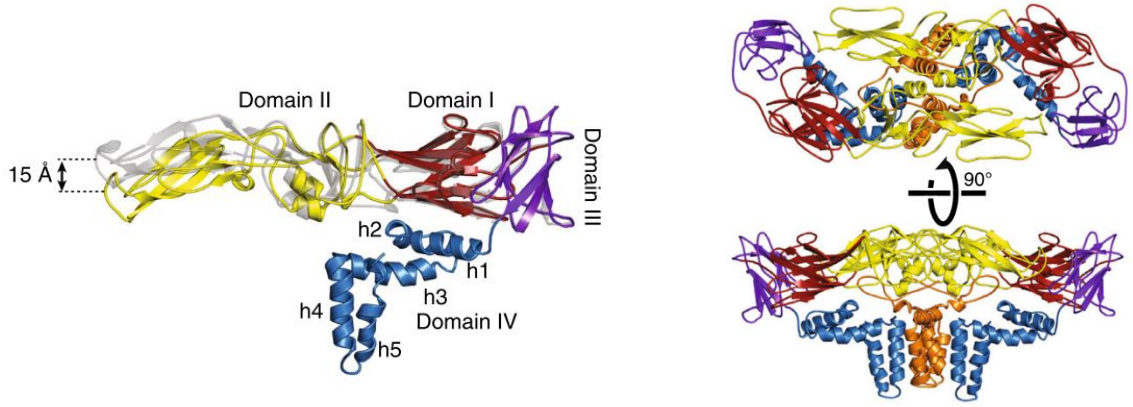


(a). (b).

Figure 1.3. Maximum likelihood trees (Unrooted) constructed by the (a) envelope gene region of the TBEV RNA, and by the (b) complete open reading frame (ORF) of TBEV RNA. Respective scalar bars represent the expected number of substitution per site (Deviatkin et al., 2020).

1.2.3 Proteins encoded by the TBEV RNA

The resulting polyprotein sequence through the translation of positive strand TBEV RNA by the host enzymes, cleaves into three structural proteins and seven non-structural proteins both by the host and the viral proteases (Yun *et al.*, 2011). Capsid protein is one of the three structural proteins and the smallest (11 KDa) of all the proteins encoded by the translation of the TBEV CDS region (Gerhard *et al.*, 2021). Membrane protein is the second protein of the structural protein with the molecular weight of 26 KDa. After the translation, it undergoes through several proteolytic cleavages resulting, premembrane protein (prM) and membrane protein (M) (Heinz *et al.*, 1994). The third structural protein is the envelope protein (50 KDa) which function as the protective layer of the virus. The envelope protein is also involved in the receptor mediated endocytosis process by binding with the host receptor molecules for the fusion of the viral and the cellular membrane to form the endosome (Figure 1.4.) (Füzik *et al.*, 2018).



(a). (b).

Figure 1.4. (a). Differently coloured four domains of the superimposed cryo-EM (coloured) and X-ray (Gray) E-protein structures. (b). Cryo-EM images of heterotetrametric structure of two envelope proteins (coloured according to each domains) and two M-proteins (shown in Orange) shown in two different orientations (Füzik *et al.*, 2018).

In addition to the three structural proteins, it translates seven non-structural proteins, mainly to regulate the viral replication machinery inside the host cell. The first non-structural protein next to the envelope protein of the polyprotein chain is the NS1, and the exact function is not yet unravelled (Lee *et al.*, 1989). NS2A (22 KDa) is the second non-structural protein, the function of it is not completely resolved yet. The NS2B (14 KDa) identified to be acting as a cofactor for the protease activity of the NS3 protein via the central hydrophilic domain of it.

The second largest non-structural protein of the TBEV is the NS3 protein (70 KDa) and it is the main regulatory protein of the viral replication by involving in the polyprotein processing through the protease activity of the N-terminal domain to generate individual proteins. It is also found that the mutational changes and their correlations to have the neuropathogenicity have been associated with the NS2B-NS3 interactive active site by showcasing the critical functions of the NS3 protein in the estimation of neuropathogenicity of TBEV variants (Růžek *et al.*, 2008). The NS4 (16 KDa) is identified to be an integral protein and involved in the bridging NS1 and the viral replication complex (Uchil & Satchidanandam, 2003). NS4B and its function is not fully understood yet, but its interactions and the localizations in the nucleus and site of replication have been identified. In addition to that it also have the ability to

interfere with the interferon signalling by blocking the activation of interferon stimulated response element promoters (Gerhard *et al.,* 2021).

The most functional and the largest protein, which is located at the c-terminal end of the polyprotein chain is the NS5 protein and function as the RNA-dependent RNA polymerase. In addition to its main function, the N-terminal domain functions in the mRNA capping process by involving tin the AdoMet-dependent methyltransferase activity by transferring the methyl group from the cofactor S-adenosyl-I-methionine onto the cap guanine and ribose moiety of the first RNA nucleotide (Gerhard *et al.,* 2021).

1.2.4 Replication strategy

It is assumed to be the viral replication initially start in the antigen presenting cells of the host after a tick bite, further it spreads to the lymph nodes and continue the incubation inside the lymph nodes. It advances towards the other organs inside the body via the lymphatic system and through the blood circulatory systems. Once it captured by the reticuloendothelial system, it starts replication further and at this stage TBEV infected patients start experiencing flu-like non-specific symptoms. After seven to ten days, secondary phase begins with the involvement of central nervous system (CNS), this leads to meningoencephalitis myelitis and in some cases even death (Gerhard *et al.,* 2021).

TBEV replication begins through the attachment of the viral particle to the cellular membrane of a host cell. The molecular attachment process of the TBEV has not been fully characterized but hypothesised to be involving receptor mediated endocytosis process through the viral envelope glycoprotein protein domain I, through a glycosylation (Fernández-Sanlés *et al.,* 2017). The receptors involved in the host cells have been not identified properly, but have been suggested to be involving more than one receptor for the triggering of endocytosis process (Gerhard *et al.,* 2021). The endocytosis process leading to form the endosomal vesicle and the lowering of the pH will trigger the conformational changes of the E-protein to form trimerized E-protein complex. Through the help of the membrane fusion ends of the trimerized envelope proteins, viral membrane and the cell membranes get fused and nucleocapsid get released into the cytoplasm (Fernández-Sanlés *et al.,* 2017). The disintegration mechanism of the nucleocapsid has not been fully understood, but the viral positive sense RNA gets released into the cytoplasm and finally reach the endoplasmic reticulum.

Initially through the translational machinery of the host cell, the positive sense viral RNA gets translated into 3400 aa. long polyprotein and by the help of protease of the host cell, polyprotein chain gets cleaved into smaller proteins (Fernández-Sanlés *et al.,* 2017). The rest of the viral proteins get formed through the proteolytic cleavage by the released viral peptides. Through the NS5 protein and also with the involvement of other replication complex associated viral proteins, viral RNA replication begins (Černý *et al.,* 2016). In this the positive sense RNA get transcribed into negative sense RNA and through the negative sense RNA as the template, starts to produce new positive sense RNA molecules (Proutski *et al.,* 1997). The translated polypeptides and the transcribed new positive sense RNA gets assembled inside the ER lumen and finally transported into the Golgi complex for the maturation. The TBEV viral particles maturation taking place via the cleavage of prM-protein to form the M-protein and finally the matured TBEV particles get released into the extracellular space via cytoplasmic vesicles and the cycle of infection begins (Fernández-Sanlés *et al.,* 2017).

1.3 Vectors and the transmission

There are 910 known species of ticks, classified into three families: Ixodidae, Nuttalliellidae and Argasidae (Beati & Klompen, 2019 & Guglielmone *et al.,* 2014). The family Ixodidae alone consist of 14 genera and all of them are hard ticks. Tick species belonging to the genus *Ixodes* in family Ixodidae are known to responsible for the transmission of TBEV, among them *Ixodes persulcatus* Schulze (“the taiga tick”) and the *I. ricinus* Linnaeus (“Castor bean tick” or the “Sheep tick”) are the main reservoirs and the vectors (Korenberg, 2000).

In addition to the species of genus *Ixodes,* three species from the genus *Dermacentor* have been identified to carry TBEV in Europe, *D. reticulatus, D. marginatus* and *D. nuttalli.* Tick species from two other genera, *Haemaphysalis puncata, H. concinna* and *Hyalomma marginatum* have been identified to be the vectors of TBEV in Europe (Nicholson *et al.,* 2018). It has been identified, two possible ways of contracting TBEV to humans, one way of transmission is through a TBEV infected tick bite, the other way of transmission is through a consumption of TBEV contaminated unpasteurized dairy products (Figure 1.5.).

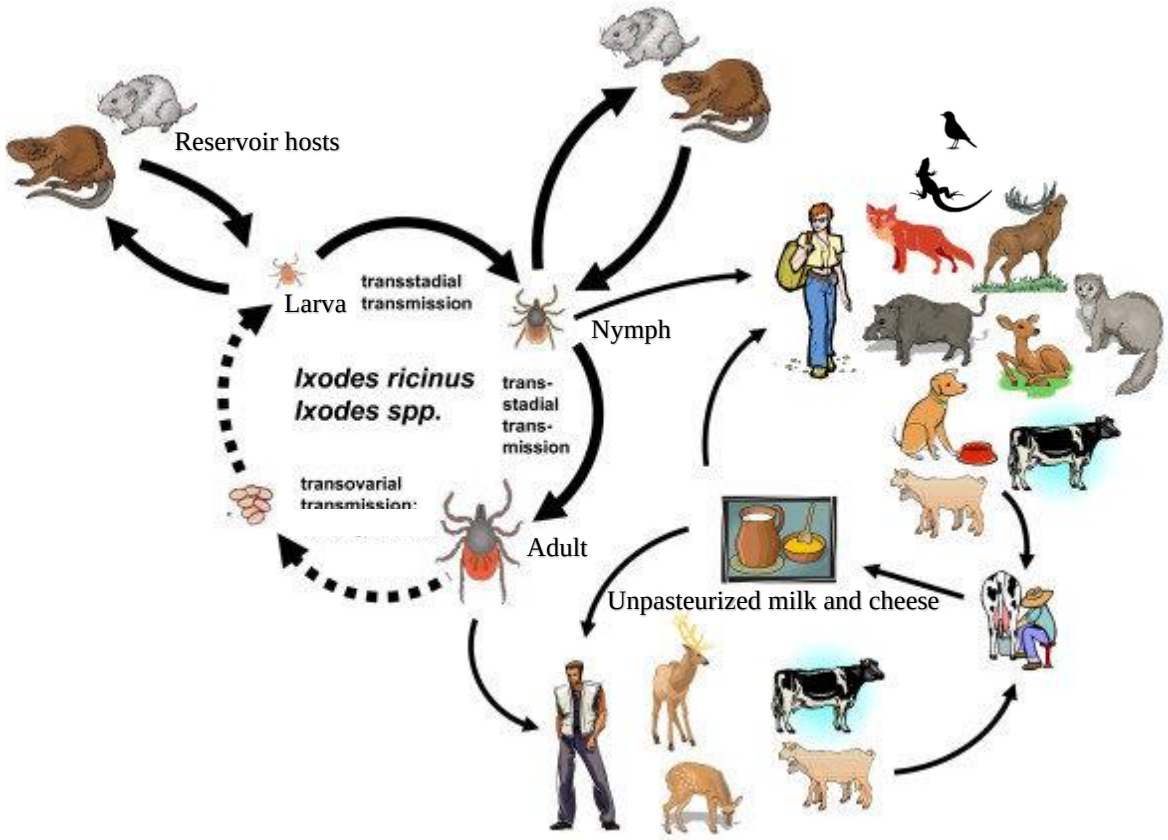


Figure 1.5. A schematic drawing of transmission cycle of TBEV between reservoir hosts, wild animals, domestic animals and also the possible transmission routes to humans (Pfeffer & Dobler, 2010).

The most common and the most frequent incidents have been reported through the TBEV infected tick bites, where the transmission through the TBEV contaminated dairy product consumptions found to be rare and limited to the Eastern Europe (Lemhöfer *et al.*, 2021). Most of the reported TBEV infections among humans are from tick bites by adult females (Guglielmone *et al.*, 2014). A recent study on TBEV infection indicates a possible transmission route from mother to infant by the breast feeding (Kerlik *et al.*, 2022).

The TBEV viral transmission between the ticks are through three main routes, transstadial, transovarial and through co-feeding ticks on the same host (Figure 1.5.) (Estrada-Peña & De La Fuente, 2014). During a blood meal a TBEV infected vector transfers the pathogen to the susceptible host. Co-feeding transmission of the TBEV occurs when the uninfected ticks are feeding closer to the TBEV infected ticks (Estrada-Peña & De La Fuente, 2014). According to both experimental and field data, previous experiments have shown that co-feeding alone may maintain a local foci (Dobler et al., 2011).

Stability of maintaining the local foci of TBEV in the wild is mainly dependent on well-established interactions between vectors and the reservoir hosts and the environment for optimal virus circulation. This is very complex and involves diverse set of species (Labuda *et al.,* 1996, Michelitsch *et al.,* 2019). The persistence of the TBEV in an endemic area is found to be caused by horizontal transmission among the wild animals from TBEV infected tick to a susceptible vertebrate host. The reservoir hosts of TBEV plays a major role in the transmission of the disease due to the long lasting viraemic period, identified to be mainly rodents, small (wildlife) carnivores and the insectivores (Conze *et al.,* 2021). Spreading of the TBEV from one region to another, mainly happens by migratory birds and large ungulates (Liebig *et al.,* 2021 and Estrada-Peña & De La Fuente, 2014).

Tick phenology resets by the winter condition each year. Due to the changes in the climate conditions the rapid distribution of TBEV foci in the natural population across Europe might have disrupted this phenological resetting. Climate changes are shown to have a negative effect on the tick natural control cycle. This might lead to higher tick population and increased cofeeding behaviour with increased maintenance of the TBEV foci (Estrada-Peña & De La Fuente, 2014).

1.4 Distribution and prevalence

Currently, TBEV have been identified in 27 European countries, Russia and some Asian countries (Figure 1.6.) (Süss, 2011). The recent studies of the global distribution of TBEV is not restricted to the TBEV endemic areas but also to TBEV non-endemic regions possibly due to adaptability to different climatic zones (Galgani *et al.,* 2017). The rapid variation in the distribution of the TBEV in non-endemic regions identified to be linked with the increased survival of *I. ricinus* and *I. persulcatus* I wild population (Estrada-Peña & De La Fuente, 2014).

The geographical distribution of TBEV is well studied and TBEV subtypes and tick vector species are correlated to specific geographical regions. The exception is the TBEV-Sib subtype, which is widely distributed in eastern Europe and Russia, but has been identified in Finland and the other Baltic countries, Ukraine, also in the Balkan peninsula. In addition to

the Europe and Russia, the TBEV-Siberian subtype has been reported in Asia, including Mongolia, Kazakhstan and Kyrgyzstan (Figure 1.6.) (Süss, 2011).

Figure 1.6. Global distribution of TBEV subtypes, borders the identified areas of TBEV-European subtypes (dotted blue line), TBEV-Siberian subtype (dotted green line) and TBEV-Far eastern subtype (dotted gray line) (Kunze *et al.*, 2022).

The TBEV-FE subtype has been reported in the far-eastern parts of the Russia, North China and in the Northern parts of Japan. In addition, there are considerable evidence of reported incidents of TBEV-FE subtype in the European parts of the former soviet union and the Urals and Siberia, where the TBEV-Sib subtype are more common (Figure 1.6.) (Kovalev *et al.*, 2010).

The distribution of the TBEV-Eur subtype found to have no clear, distinct clustering and the dispersal of the virus considered to be through rodents and mammals but also though birds over longer distances. The clustering of the TBEV-Eur subtype is randomly dispersed throughout Europe (Uzcátegui *et al.*, 2012). TBEV-Eur have been identified in most of the European countries including, Hungary, Poland, Bulgaria, Slovenia, Austria, Romania, Finland, Germany, Sweden, Norway, Switzerland and also in the Czech Republic (Daniel *et al.*, 2018, Gäumann *et al.*, 2010). The TBEV-European subtype has been reported in Korea (Figure 1.6.) (Uzcátegui *et al.*, 2012). A study carried out in Latvia and in Lithuania, identified both *I. ricinus* and *I. persulcatus* tick species and identified all three main subtypes of TBEV. In Finland both TBEV-Eur and TBEV-Sib detected (Katargina *et al.*, 2013).

The TBEV in the wild tick population is low and it has been found to be 1.3% for *I. ricinus* and for the *D. reticulatus* 1.58% in some endemic regions (Biernat *et al.,* 2014). In general, the prevalence in *I. ricinus* is found to be closer to 1% and rarely reaching between 2-5%, in ticks from migratory birds it varies from 0.5 to 14% in nymphs and 7% in larvae (Gerhard *et al.,* 2021, Gresíková & Noseck, 1966). In a study carried out in Hampshire, England in 2019, they have used deer as the sentinel species for the determination of TBEV seroprevalence by ELISA and found out 4% positive cases out of 1309 in the serum samples (Holding et al., 2019).

1.4.1 Distribution and prevalence in Norway

The distribution of TBEV in Norway has been studied since 1973 and it was first isolated from a *I. ricinus* tick in the Vestland County by Traavik and co-workers in 1973 (Traavik, 2009). Several studies have been carried out to identify the distribution of the TBEV throughout Norway using tick samples collected in field surveys, milk samples collected from farm animals and also through seroprevalence studies of humans, farm and wild animal samples. TBEV has been identified along the coastal areas of Norway to Brønnøy municipality (Vikse *et al.,* 2020; Paulsen *et al.,* 2020). The TBEV virus is endemic in the southern parts in Agder, Vestfold and Telemark and Viken County in Norway (Paulsen *et al.,* 2020) (Figure 1.7.).

The prevalence of TBEV in nymph and adult ticks have been identified ranging from 0.0 – 1.1% and 0.0 -20.6%, respectively. The seroprevalence in wild animals ranged from 41% to 1.6% in Farsund region of southern coast of Norway and Molde in Midwest coast of Norway, respectively (Ytrehus *et al.,* 2013). The overall seroprevalence in wild animals was 4.6%, 9.4% in moose samples, 1.4% in red deer, 0.7% in roedeer and 0% in reindeer (Paulsen *et al.,* 2019). The seroprevalence in humans living in the TBEV endemic area of the southern Norway was 3.1% (Thortveit *et al.,* 2020). In TBEV non-endemic area in the Viken County, towards the southern part of Norway, the seroprevalence was 0.65%, whereas in Vestland County, western Norway, the seroprevalence was 0.4% (Hjetland *et al.,* 2015a and Larsen *et al.,* 2014) . In 1979, the seroprevalence data reported to be 19.6% in Vestland County (Traavik, 2009) and which was much higher than the study carried out by Hjetland et al., 2015.

Figure 1.7. Geographical locations of identified TBEV positive samples from Ticks samples, cow milk samples and antibodies from animals in Norway (K. Paulsen *et al.*, 2020).

1.5 Epidemiology in Europe

The recent report published by the European Centre for Disease Prevention and Control (2019), showcased over 12,000 TBEV infections in humans during the four-year time period from 2012 to 2016. Only in EU/EEA countries 3246 cases were reported (2019) and the notification rate was 0.7 per 100,000 individuals (Figure 1.8.). The highest number of notification rate was from Lithuania where the 25.4 cases per 100,000 population and second highest from Czechia (7.3 : 100,000) and the third highest notifications from Estonia (6.2 : 100,000). The percentage of fatal cases in the region was 0.7%. The number of infected cases

were most frequent in the age group 45-64 years with a 1.5:1 male to female ratio (European Centre for Disease Prevention and Control, 2021).

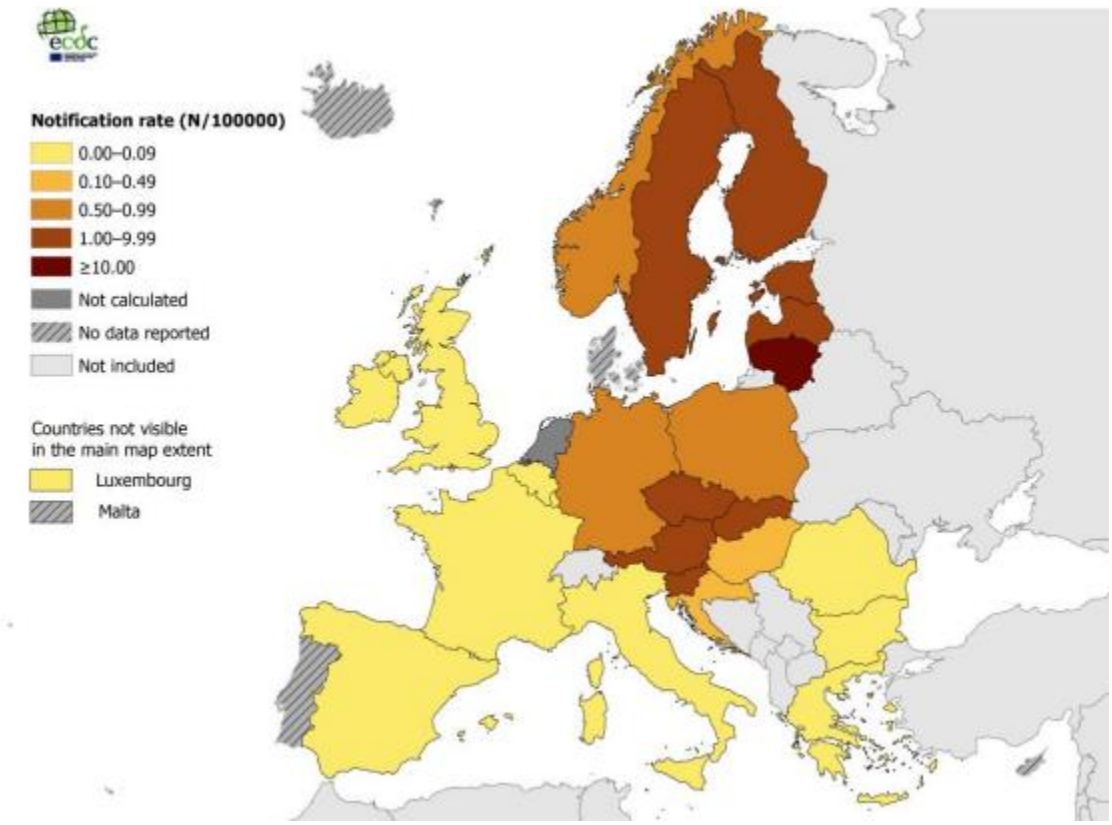


Figure 1.8. Number of confirmed TBEV cases per 100,000 population in Europe (2019) (European Centre for Disease Prevention and Control, 2021).

1.6 Detection, control and vaccination

Due to the unspecific clinical symptoms similar to influenza and other flavivirus infections, early detection of the TBEV in humans is a challenge. It requires laboratory diagnosis and clinical history of the patient, information about recent tick bites or consumption of unpasteurized dairy products. During the first phase of the disease, blood serum samples from the TBEV infected patients are used to confirm the presence of TBEV RNA by various techniques; real-time RT-PCR (Andreassen *et al.*, 2012), nested RT-PCR (Christina Schrader & Süss, 1999), RT-qPCR (Rudenko *et al.*, 2004) or RT-LAMP (Hayasaka *et al.*, 2013) techniques. The secondary phase of the disease presenting severe meningoencephalitis and related symptoms, the amount of TBEV RNA in the blood serum are undetectable through the conventional PCR techniques (Figure 1.9.). At this stage the detection of TBEV RNA is

achievable through analysis of brain samples during the post-mortem. However, in a retrospective study, TBEV RNA was detected in urine samples by RT-PCR technique 19 days after the appearance of neurologic symptoms (Veje *et al.*, 2014).

In the second phase, indirect techniques are developed to identify the TBEV infections in patients through the antigen and antibody interactions (Figure 1.9.), such as complement fixation assay, Hemagglutination inhibition tests, immune fluorescent assays, neutralization assays and ELISA techniques. The sensitivity due to the cross reactivity seems to be the main disadvantages associated with these techniques. Antibodies which are produced against other flavivirus infections may give false positive results with these assays and questioning the accuracy and the reliability of the diagnostic techniques (Varlacher *et al.*, 2015).

To the current date, no specific treatments are available against the TBEV infections. Antiviral therapy can be administered during the first phase of the infection. During the second phase of the infection, the use of antiviral therapy has no meaningful effect (Gerhard *et al.*, 2021).

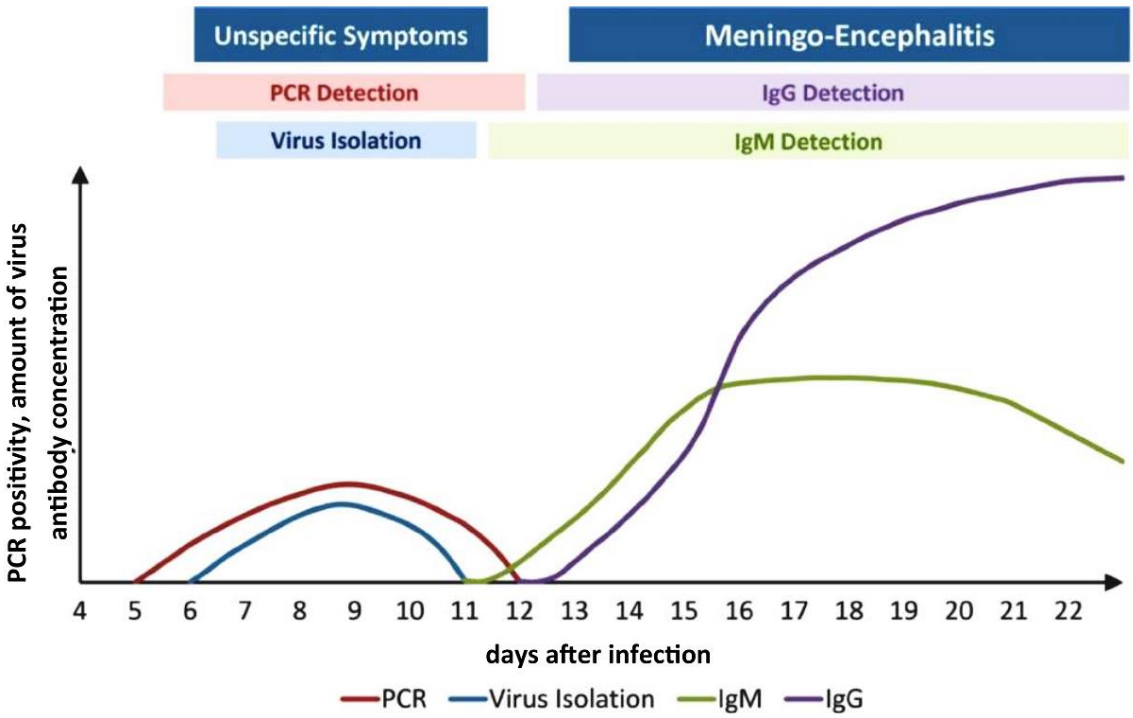


Figure 1.9. Relative proportionate changes in viral RNA concentrations, viraemia content, concentration of IgM and IgG in the blood serum samples of TBEV infected patients during two phases of the TBEV infections. Horizontal bars represent the specific detection techniques that could be employed for the identification of TBEV in patients with specific time intervals during a TBEV infection (Gerhard *et al.*, 2021).

It has been found that the control of TBEV is basically through the prevention, avoid tick bites and consumption of unpasteurized dairy products as well as vaccination. There are commercially available inactivated vaccines are available based on the TBEV subtypes commonly found in the specific country regions. To receive the full immunization against TBEV, individuals need to be vaccinated with at least three to four doses of the same vaccine. Moreover, the vaccination needs to be repeated after three to four years, depending on the manufacturer's requirements. Austria has a mass vaccination program since 1981 and the vaccine coverage is 96-99% percent of the population and as a result they have been able to prevent 4,000 cases of TBE and 15 to 30 deaths annually (Varlacher *et al.,* 2015).

1.7 Previous studies on whole genome sequencing of TBEV

There have been numerous success studies carried out to sequence the whole genome (WG) of TBEV by various approaches (Frey *et al.,* 2012, Asghar *et al.,* 2017, Kupča *et al.,* 2010 and K. M. Paulsen *et al.,* 2021). Majority of the methods have been developed based on culturing of TBEV in cell culture samples (Kupča *et al.,* 2010). Very few whole genome sequencing (WGS) studies have been carried out by direct analysis of TBE in ticks, serum or milk samples. Only one WGS analysis has been carried out directly from a sample with separate PCR reactions and finally combining the sequence fragments to generate complete genome (Gonzalez *et al.,* 2022).

1.8 Molecular Biology techniques employed in this study

1.8.1 Homogenization and the total RNA extraction

Extraction of TBEV RNA requires mechanical disruption of tissues and cells of tick samples or animal tissues. Homogenization of the samples usually carried out with tungsten carbide beads and with the addition of lysis buffer to digest the tissues and cells to release the viral particles to the extraction medium (Gäumann, Mühlemann, Strasser, & Beuret, 2010). Centrifugation of the homogenized solution facilitates removal of remaining cells, tissues and cell debris. To prepare TBEV RNA it is crucial to digest DNA by addition of DNase followed by addition of RNase inhibitors to minimize RNA degradation (Mazeaud *et al.,* 2018). The

addition of viral lysis buffer provide the inactivation of the viral particles with release of viral RNA prior to extraction of total RNA (Gäumann *et al.,* 2010, Fitzpatrick *et al.,* 2021 and Bastakoti, 2019). Various techniques for extraction of either total RNA or total nucleic acids have been employed for a large number of tick samples (Gäumann, Mühlemann, Strasser, & Beuret, 2010). Mainly affinity chromatography or partition chromatographic techniques or the combination of both, have been used to separate DNA and other cellular debris from RNA (K. M. Paulsen *et al.,* 2021 and Klaus *et al.,* 2010).

1.8.2 Reverse transcription polymerase chain reaction (RT-PCR)

Conversion of viral positive sense RNA to DNA is critically important for the WGS of viral genome because the molecular biology techniques are mainly developed for amplification of DNA. Another issue is the high degradation potential of RNA mainly due to the activity of the RNase or reactive oxygen and nitrogen species present in the sample (Li *et al.,* 2014). Freeze thaw cycles through the formation of ice crystals may cause RNA degradation especially of TBEV RNA present in low amounts (Gonzalez-Perez *et al.,* 2010).

Reverse transcription of RNA to DNA are utilized by RNA dependent DNA polymerases. Selection of the appropriate type of reverse transcriptase is based on several factors, such as maximum processivity, maximum length of the single stretched cDNA, minimum RNase H activity, maximum thermal stability and the ability to resist reverse transcriptase inhibitors (Coffin J.M., Hughes S.H., 1997). The primers employed in most of the RT-PCR of TBEV RNA studies were random hexamers for the viral RNA. The avoidance of oligo dt primers are crucial to avoid the synthesis of mRNA from the host cells and only to synthesize viral cDNA. The resulting DNA are directly use for further amplification with WGS primers (Smura *et al.,* 2019, Carpi *et al.,* 2009 and Yun *et al.,* 2011).

1.8.3 Polymerase chain reaction (PCR)

There are several well-established methods for the whole genome amplification of flaviviral genomes, majority of them have been utilizing one or more steps of PCR techniques such as nested-PCR (Puchhammer-Stöckl *et al.,* 1995), Hot-start PCR technique (Jansen *et al.,* 2021) and the touchdown PCR technique (Katargina *et al.,* 2013) rather than conventional PCR technique. These methods are used to increase the specificity and efficiency of the PCR amplifications of DNA for the library preparations and the subsequent WGS.

Selection of an appropriate DNA polymerase is critical for the quality of the resulting whole viral genome. Selection criteria are high fidelity, processivity, thermal stability, proofreading ability and low error rate. The efficiency of the PCR technique is determined by the primer's annealing temperatures and their specificity to the targeted regions of the DNA sequences. These factors will greatly reduce amplification of unspecific products and improve the overall efficiency of the PCR leading to a better DNA library for WGS. To improve the specificity of highly variable regions of the genomes degenerate primers have been employed to increase the accuracy and efficiency of the viral whole genome amplifications (Nanda *et al.,* 2008).

1.8.4 Real-time PCR for the detection of viral specific cDNA

Several methods have been developed for the detection of viral genetic materials in samples. Real-time RT-PCR methods for amplification of viral cDNA have been developed and well established (Schwaiger & Cassinotti, 2003, Achazi *et al.,* 2011 and Gäumann, Mühlemann, Strasser, Beuret, *et al.,* 2010). There are several different detection strategies developed for real-time PCR products, such as using SYBR Green as the dsDNA intercalating and fluorescent agents with specific primers (Mackay *et al.,* 2002) and Taq-Man probes with quencher and the fluorescent is annealed to the specific target sequence during the annealing step. In Taq-Man probe assays, the fluorescent tag is released due to the 5'→ 3' exonuclease activity of the polymerase and the activated fluorescent emission is detected by a laser (Andreassen *et al.,* 2012). The main advantages over the conventional PCR techniques are increased specificity and direct quantification ability. The amount of specific DNA present in the sample can be quantified by comparing the cyclic threshold (Ct) value to a standard sample. Main disadvantages are the false positive and false negative results due to the unspecific amplifications. In addition, presence of SNP in the primer binding sites leading to the low amplified PCR products. The presence of strong PCR inhibitors in the sample of interest may inhibit the PCR amplifications of the targeted region leading to false negative results (Mackay *et al.,* 2002).

1.8.5 DNA barcoding, library preparation and Oxford Nanopore sequencing

The number of samples that can be submitted for single sequencing process have been increased with the use of DNA barcoding and this is basically to reduce the cost and the

efficient use of resources and to minimize the time required for the analysis of large number of samples (Xu *et al.*, 2018). The insertion of the specific DNA barcodes is carried out with the use of fragmentation mixture consisting of transposase enzyme complexed with the extended transposase adapter with the specific DNA barcode sequence. During the DNA library preparation each individual sample reacts with fragmentation mixture to fragment the DNA and incorporate the transposase adapter containing the DNA barcodes followed by the DNA product purifications (Kono & Arakawa, 2019 & Wang *et al.*, 2021).

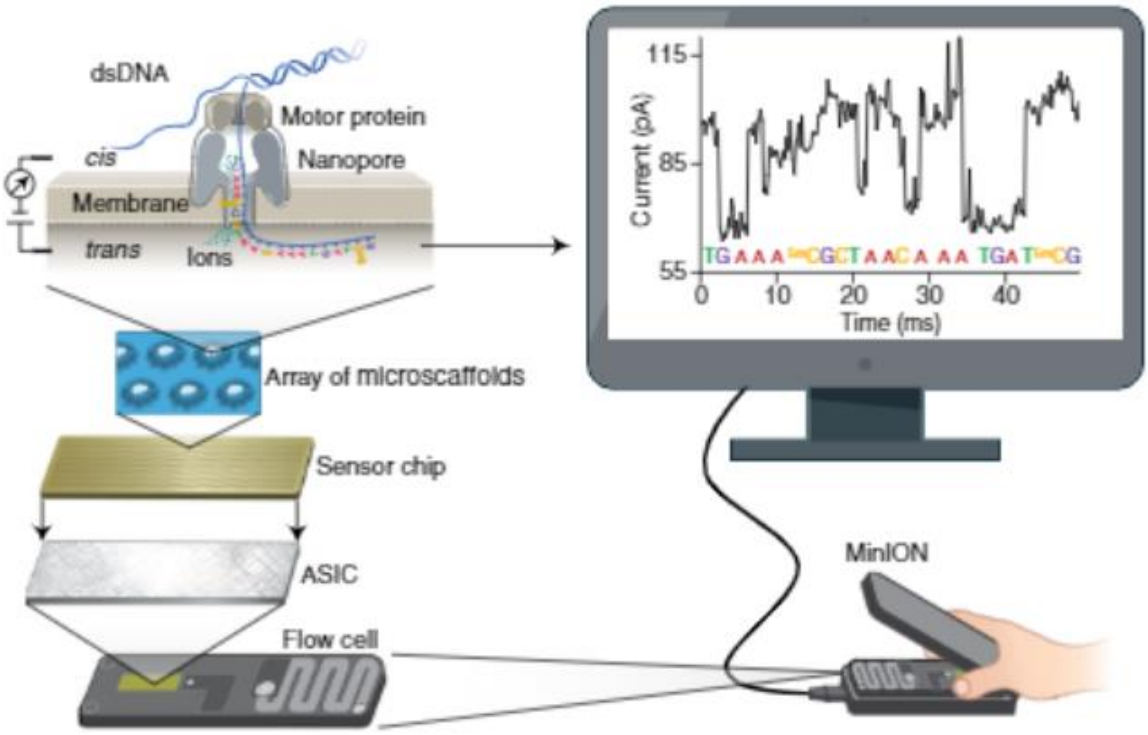


Figure 1.10. Schematic representation of the Oxford Nanopore sequencing technology, flow cell and the determination of nucleotide composition of the DNA samples using the tunnelling current fluctuations and by the base-calling (Wang *et al.*, 2021).

An added motor protein will guide the ligated adapter containing barcoded DNA to the Nanopore cell. The passing DNA sequences through the nanopore lead to the changes in the tunnelling current fluctuations measured by the device. These fluctuations converts into the corresponding nucleotide sequences using the basecalling algorithms. Finally, the nucleotide sequences are analysed by trimming the adapter sequences and followed by bioinformatic analysis (Figure 1.10.). The main advantages of the Oxford Nanopore technology are an analysis of the data in real-time and rapid diagnostic settings. In addition, long read sequencing capability with the length of 2.273 Mb and detection of DNA and RNA modifications and related epigenetic information are some of the added advantages over the other WGS

techniques currently available. The main drawbacks of the sequencing by the Oxford Nanopore Technology (ONT) are lower sequencing accuracy of repeated nucleotide sequences compared to other WGS platforms and prone to have errors that can be overcome by a better sequencing depth. The percentage of sequencing accuracy in ONT are reported to be ranging from 94%, 86% and 95% in 2D reads, 1D and $1D^2$ reads, respectively (Wang *et al.,* 2021).

1.8.6 Bioinformatics and data analysis

Oxford Nanopore sequencing of purified PCR amplified products from viral genomes, generates long reads which are required to be assessed for the quality. There are various tools developed for the determination of next generation sequencing (NGS) data quality such as FastQC, htSeqTools, Kraken, NGSQC and PRINSEQ (Pfeifer, 2017). After the quality analysis, the removal of the poor-quality sequences is required in order to improve the quality of the dataset. Further, adapter and barcoding sequences attached during the NGS, need to be removed to increase the accuracy and to reduce the complexity of the dataset. In addition, the removal of the primer sequences is crucial to avoid any bias caused by the primer sequences in the determination of the final genome construct. Removal of specific sequence regions can be performed by the trimming tools (Trimmomatic, NanoFlit & Trim Galore!), which are available for the removal of poor-quality reads, adapter, barcode sequences and primer sequences (Quick *et al.,* 2017 & Marston *et al.,* 2013).

The quality enhanced datasets can be submitted for the *de novo* assembly or for the reference-based genome mapping. The *de novo* assembly can be carried out using various tools available, such as rnaviralSPAdes, Flye and Raven. The requirement to have a reference sequence is absent and assembling is mainly carried out using the overlapping sequence regions constructed by the overlapping primer pairs during the PCR amplifications. But in the reference-based mapping, sequence reads mapped along a reference genome sequence to generate the final construct. Therefore, it is required to have a reference genome sequence to do the reference-based mapping. There are several available tools developed to do the reference-based sequence mapping such as, Minimap2, Bowtie 2 and NanoPlot. In both cases, the consensus sequence can be generated by using the sequence aligned map as specific to the requirements (El-Metwally *et al.,* 2014).

The identification and the evaluation of evolutionary relationship to a known set of viruses or tick species can be carried out by the phylogenetic tree analysis. This is basically done by the multiple sequence alignment (MSA) carried out with similar set of sequences or the genome sequences and the resulting MSA is subjected to phylogenetic tree analysis with appropriate algorithm (Kuno *et al.*, 1998 & Gorbalenya & Lauber, 2017). If the sequences and its biological origin is unknown, BLASTn analysis against a suitable nucleotide database may provide a possible answer to find out most similar sequences. In addition, the MegaBLAST technique can be used to filter the specific sequences from a large dataset in order to reduce the complexity of the analysis (Chen *et al.*, 2015).

1.9 Aim of the study

Previous studies carried out by the Norwegian Institute of Public Health (NIPH) have analysed the prevalence of TBE in Norway in tick samples collected from the different parts of Norway. Semiquantitative RT-qPCR techniques for the detection of TBEV in tick samples, serum samples from patients and animals have been developed. Furthermore, PCR methods to amplify and sequence the complete genome of TBEV from samples with high viral load by the Illumina sequencing techniques have been developed. The drawbacks of the existing methods have been identified and associated with partial PCR amplification with low viraemic TBEV samples leading to lack of information regarding the sequence.

The aim of the present study was to develop a rapid whole genome sequencing technique for the low viraemic TBEV positive samples directly obtained from the RNA extracts of ticks to generate accurate results with higher precision. Further, by the generated sequence data to identify the TBEV subtypes and to evaluate the genomic diversity of TBEV in Norway with more accuracy and with better precision if the developed methods are accurate enough. In addition, through the identified TBEV subtypes of the clinical samples, we may predict the neuropathogenicity and severity of the disease.

The aim of this study was to:

I) improve the already existing RT- real-time PCR

II) improve the whole genome sequencing method for total RNA extracts of homogenized ticks with very low viraemic content collected directly from the field

a) design primers,
b) test diffenrent PCR conditions and
c) library preparation and sequencing of the TBEV samples using Oxford Nanopore technology.

2. MATERIAL AND METHODS

2.1 Experimental design

The study was carried out to develop a PCR method which would facilitate to sequence whole TBE viral genome directly through the total RNA extracts of TBEV infected tick species collected from the fields or from serum sample extracts from the TBEV infected patients with very low TBEV viraemic content.

The study consists of three main sections, the first one is the evaluation of the limitations associated with the already existing RT-PCR and PCR methods at NIPH for the amplification of the whole genome of the TBEV-Hochosterwitz strain. To identify the limitations of the methods, identification of the lowest possible TBEV RNA concentration required to perform the already established PCR methods to amplify the TBEV whole genome was considered. To achieve that, Oxford Nanopore sequencing was carried out for 12 different DNA libraries developed through the already established RT-PCR and PCR methods using a tenfold dilution series of an RNA extract obtained from a cell culture sample incubated with TBEV-Hochosterwitz stain. Through the assembled sequenced data, the lowest possible concentration of RNA of TBEV-Hochosterwitz stain required for the whole genome amplification of the TBEV-Hochosterwitz stain was identified. In addition to that, by the analysis of assembled sequenced data, experiments were carried out to identify the poorly performing primer pairs which required a replacement. To quantify the relative concentration of TBEV RNA in tenfold diluted samples and with seven different TBEV RNA extracts obtained from different RNA extraction techniques through the TBEV-Hochosterwitz strain incubated cell culture samples, semi quantitative real-time PCRs were carried out. This was performed to facilitate the relative comparison of existing methods and their capability to amplify whole genome of TBEV in any sample below the threshold cycle (Ct) value (Figure 2.1.).

The second section of the study was mainly involved in the optimization of the already established RT-PCR and PCR techniques. Very few changes were carried out for the RT-PCR step by experimenting with the primer composition by replacing the random-hexamers with TEBV specific primers and the increase in the amount of template RNA used in the reactions. The optimizations related to the PCR conditions were mainly through the replacement of newly designed primers with the poorly performing primers, changes in the primer

concentrations, compositional changes in the primer schemes used and the changes in the PCR reaction temperatures. All these changes and their effects on the amplification of the whole TBEV genome were evaluated through the assembled Oxford Nanopore sequence data to identify the best possible primer schemes, optimum temperatures, and the optimum primer concentrations for the RT-PCR and PCR reactions (Figure 2.1.).

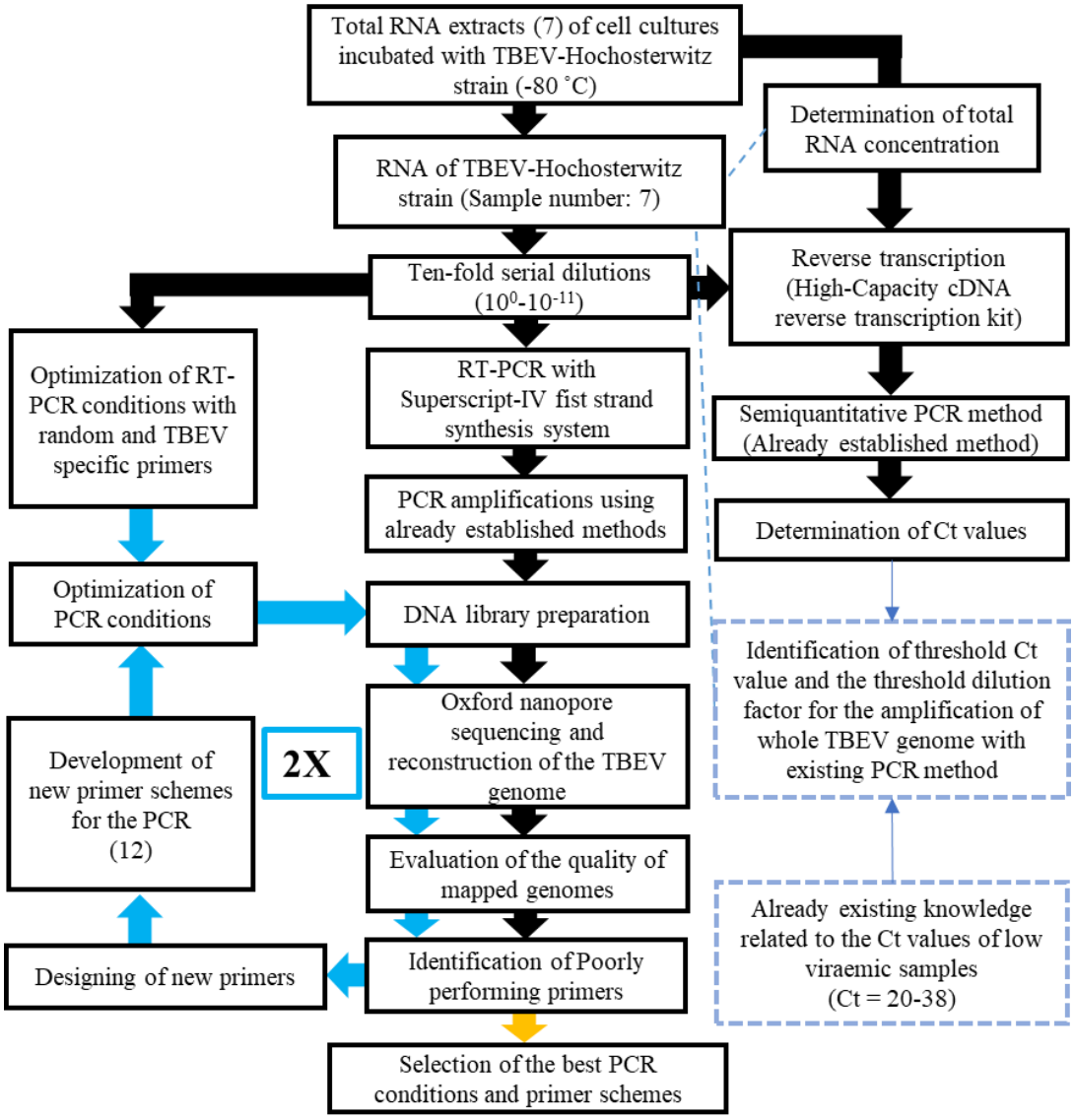


Figure 2.1. First two sections of the experimental design, the evaluation of the limitations associated with the previous RT-PCR and PCR techniques for the whole genome amplifications with low viraemic samples, optimization of the previous RT-PCR and PCR conditions to enhance the whole genome amplification of TBEV directly through the tick samples with very low viraemic content. Coloured arrows indicate the changes made into the already existing methods with the number of trials. Dotted lines indicate the information and decision flow.

The third section of the study was to evaluate amplification capability of the TBEV whole genome through the optimized RT-PCR, PCR and best primer schemes developed in the 2nd section of the study. In this section, total RNA extracts from the TBEV positive ticks collected and evaluated for the TBEV from the southern coastal region of Norway were sued for the evaluation through the RT-PCR, PCR and subsequent Oxford Nanopore sequencing. Finally, evaluation of the developed PCR method and the identification of the TBEV viral subtypes was performed (Figure 2.2.). In addition, the study was extended to identify the specific tick species which exhibited positive TBEV variants using Oxford Nanopore sequenced data derived through the third part of the study (Figure 2.2.).

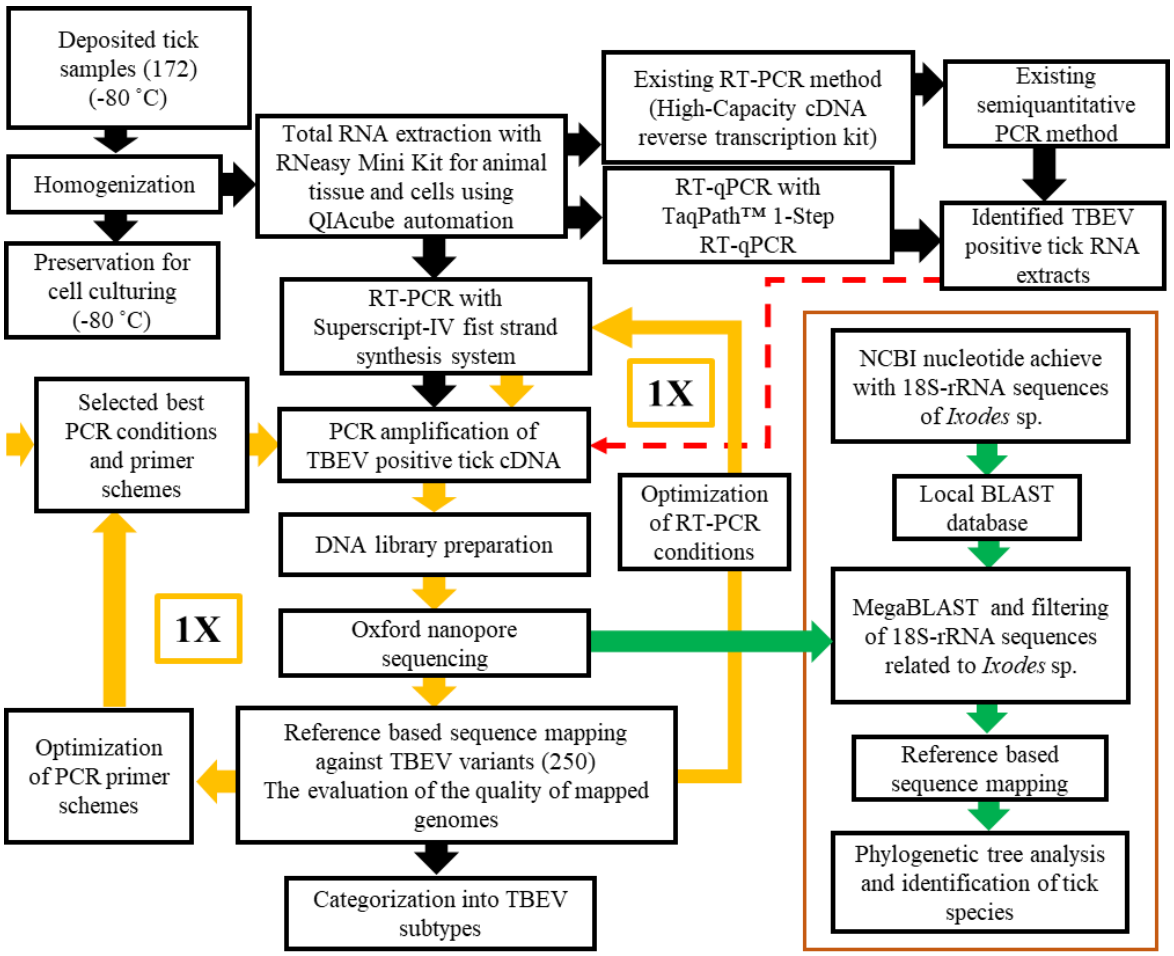


Figure 2.2. The evaluation of the optimized RT-PCR and PCR conditions and the capability to amplify whole genome of TBEV directly through the total RNA extracts from tick samples with very low viraemic content and additionally the identification of TBEV infected tick species using phylogeny.

2.2 Study materials

In 2020 there was an increase of TBEV cases near Larvik. Ticks were collected from a mixed forest in Larvik at site 1 (GPS coordinates: 58°98'N & 09°94'E), in 12th August 2021 during a field survey carried out by Dr. Arnulf Soleng, Senior researcher, Department of Pest control, NIPH. A total of 172 tick samples were collected, consisting of 76 nymph pools of ten nymphs per pool and 50 adult males and 46 adult females. All ticks were kept on ice conditions from the collection site until it was stored at -80 °C freezer until further analysis.

The seven different RNA extracts of TBEV-Hochosterwitz stain were obtained for a previous study carried out by Bastakoti, 2019, Paulsen *et al.,* 2021 and Maria J. Diekmann, 2021 (stored at -80 °C, at NIPH). Extraction was carried out using seven different RNA extraction techniques (Table 2.1.) utilizing the Vero E6 cell cultured samples of TBEV-Hochosterwitz strain. The TBEV-Hochosterwitz strain was isolated from an *Ixodes ricinus* tick collected in 1971 near Hochosterwitz, Austria (Paulsen *et al.,* 2021).

2.3 Quantification of RNA using a spectrophotometric method

RNA concentration (ng/µL) of seven different RNA extracts (Table 2.1.), were determined using a spectrophotometer (DeNovix, DS-11) at 260 nm wavelength with reference to the nuclease-free H_2O.

2.4 Reverse transcription and semiquantitative real-time PCR analysis of TBEV-Hochosterwitz RNA

Semiquantitative real-time PCR was carried out to identify the relative quantities of TBEV-Hochosterwitz strain RNA using cycle threshold (Ct) values. Seven different RNA extracts of TBEV-Hochosterwitz strain (each with 5.0 µL) (Table 2.1.), tenfold serial dilutions of TBEV-Hochosterwitz RNA from the sample number seven (3.34 to 3.34 $\times 10^{-10}$ ng/µL, 5.0 µL) and five nuclease-free H_2O samples (5.0 µL) as nontemplate controls (NTC) were separately subjected to RT-PCR using High-Capacity cDNA reverse transcription kit (Applied Biosystems™, Thermo Fisher Scientific, Cat. number: 4368813), according to the manufacturers protocol (Section A.1., Annexure I). Reverse transcription reaction was carried

out by initial primer annealing at 25 °C for 10 min., the RT-polymerization and cDNA synthesis at 37 °C for 120 min., the quenching of the reaction at 85 °C for 5 min. and finally the storing at 4 °C using the Thermal Cycler (2720), Thermo Fisher Scientific.

Table 2.1. TBEV-Hochosterwitz RNA extracts and extraction techniques

Sample number	Sample name	Total RNA extraction technique from cell culture sample
1	Hochosterwitz_Trizol_11-04-2016	Trizol™ reagent (Cat. No. 15596026, Invitrogen, CA)
2	Hochosterwitz_Qiagen_11-04-2016	RNeasy Mini Kit for animal tissue and cells (Qiagen, Inc., Valencia, CA, USA)
3	Hochosterwitz_RLT_28-04-2016	RLT lysis buffer (250 µL, added 1% β-mercaptoethanol, Cat. No. 79216, Qiagen, CA, USA)
4	Hochosterwitz_250416_26-04-2016	RNeasy Mini Kit for animal tissue and cells (Cat. No. 74106, QIAGEN, Hilden, Germany)
5	Hochosterwitz_2_AVL_28-04-2016	Lysis with AVL buffer (Cat. No. 19073) followed by extraction using RNeasy Mini Kit for animal tissue and cells (Qiagen, Inc., Valencia, CA, USA)
6	Hochosterwitz_ 17-06-2021_MAJD	RLT lysis buffer (250 µL, added 1% β-mercaptoethanol, Cat. No. 79216, Qiagen, CA, USA) and RNeasy Mini Kit for animal tissue and cells (Cat. No. 74106, QIAGEN, Hilden, Germany)
7	Hochosterwitz_2_ 26-04_2016	RNeasy Mini Kit for animal tissue and cells (Qiagen, Inc., Valencia, CA, USA)

Resulting cDNA samples from the High-Capacity cDNA reverse transcriptase reaction in the above were separately subjected to semiquantitative real-time PCR analysis using a TaqMan probe (TBE 339) assay method I (Andreassen *et al.,* 2012) . Semiquantitative real-time PCR reaction was carried out by adding the cDNA (3.0 µL) from the each of the above RT-PCR reactions separately to each reaction mixture (22.0 µL) (Composition of the reaction mixture is given in the Section A.2., Annexure I). The real-time PCR reactions were carried out on the Corbett Rotor-gene™ 6000 (Qiagen Corbett, Hilden, Germany) with the initial denaturation at 95 °C for 2 minutes and followed by 45 cycles: 95 °C for 15 seconds, 60 °C for 45 seconds for the annealing of the primers and the probes to the template, 72 °C for 30 seconds for the extension and at the end of the cycles storing at 4 °C. Fluorescent data acquired by the instrument software Rotor-Gene Q Software (Version: 2.3.1.49), at the very beginning and also at each cycle at the primer annealing and the extension stages were used to determine the cycle threshold (Ct) value of each sample by keeping the normalized fluorescence

threshold cutoff value (0.02). Finally, the corresponding Ct values received for the tenfold dilutions were compared with the previously known Ct values (Previous studies carried out at NIPH) given by the TBEV positive tick samples as a guidance for the identification of the appropriate dilution level which would represent low viremic TBEV positive tick samples.

2.5 Semiquantitative real-time PCR with cDNA derived through the random hexamers and TBEV specific primers

Seven different TBEV-Hochosterwitz strain RNA extracts (Table 2.1.), TICOVAC vaccine (5.0 µL, Pfizer Inc.), three nontemplate controls (NTCs) and one negative control (Hela cell culture RNA: provided with the SuperScript™ IV First-Strand synthesis system, Thermo Fisher Scientific, Cat. No. 18091050) were separately subjected to RT-PCR using SuperScript™ IV First-Strand synthesis system, (Thermo Fisher Scientific, Cat. number: 18091050) against five different primer compositions [Random hexamers (Thermo Fisher Scientific, Cat. No. 18091050), JK_1-F, JK_7-R, Eur_39-R, combination of all the primers with equal proportions (v/v), details of each primer are given (Table A.1., Annexure I)]. Initial primer annealing was separately carried out by heating each TBEV-Hochosterwitz strain RNA (1.0 µL) according to the manufacturer's protocol (Details of the protocol are given in the section A.4, Annexure I) in a thermal cycler [Thermal Cycler (2720), Thermo Scientific]. All the cDNA samples (50 samples, Table A.1., Annexure I) resulted from the SuperScript™ IV reverse transcription reaction in the above procedure and additionally with RNA extract of TBEV-Hochosterwitz strain RNA sample number six (3.0 µL, a negative control) (Table 2.1) were separately subjected to semiquantitative real-time PCR analysis using the TaqMan probe (TBE 339) assay method I (Andreassen *et al.*, 2012), according to the procedure mentioned (Section A.4, Annexure I). Finally, the resulting Ct values were used to compare the effect of cDNA synthesis with different primer compositions.

2.6 Optimization of RT-PCR and PCR conditions

To evaluate the total RNA from tick samples and primer concentration changes and their effects on cDNA synthesis with the TBEV RNA and subsequent PCR amplification a series of RT-PCR reactions were separately carried out against random primers (1.0 μL, 50 μM), template TBEV-Hochosterwitz strain RNA (1.0 μL) from five different serial dilutions (10^{-4} – 10^{-8}) prepared in the above procedure (Section 2.4) with /or/ without adding Tick RNA extract (1.0 μL). In addition to above RT-PCR reactions, another two RT-PCR reactions were parallelly carried out for RNA extracts from two different tick samples, one sample is the above tick RNA extract (1.0 μL) identified to be negative for the TBEV sample and the other RNA extract of tick sample (2.0 μL), have been suspected as positive for the TBEV in a previous study carried out at NIPH. For each dilution level of the RT-PCR, one NTC was carried out and all the reaction conditions were maintained identical to the previous procedure mentioned in the above (Section 2.5).

Resulting cDNA products (2.50 μL) derived through the above RT-PCR reactions were separately amplified using Q5® High-Fidelity PCR kit (New England BioLabs® Inc., Cat. No. E0555L) against six different primer pairs (JK_1-F/R, JK_3-F/R, JK_7-F/R, EUR_1-F/R, EUR_19-F/R & EUR_39-F/R, Table A.2., Annexure I) under two different concentrations each with the final the concentration of 0.015 M and 0.5 μM (Details of the PCR reaction mixture are given in the section A.5, Annexure I). The PCR reactions were carried out using a thermal cycler profile by keeping initial denaturation temperature at 98 °C for 30 seconds and 40 cycles by maintaining the temperature at 98 °C for 15 seconds required for the denaturation at the beginning of each cycle and aneling and extensions were carried out at 65 °C for 5 min and finally holding at 4 °C in a thermal cycler [Thermal Cycler (2720)].

2.7 Evaluation of effect of temperature of PCR amplification of TBEV-Hochosterwitz cDNA against each primer pair

Each primer pair (Eur primer pairs, Eur_1-F/R to Eur_39-F/R, Table A.2, Annexure I) was separately subjected to PCR amplifications using Q5® High-Fidelity PCR kit (New England BioLabs® Inc., Cat. No. E0555L) at two different primer annealing and extension temperatures (63 and 65 °C). As the template cDNA from the RT-PCR product of 4th ten-fold dilutions (10^{-4}) of the TBEV-Hochosterwitz sample prepared against the random hexamers without adding total tick RNA, in the above procedure was used (Section 2.6). For both PCR reaction conditions, each primer concentration was maintained at 0.5 µM and the rest of the PCR reaction conditions were maintained according to the procedure (Section 2.6) with a modification. The PCR conditions related to the primer extension and annealing temperatures were maintained 63 °C for a one PCR condition of each primer pair. The other PCR reactions for each primer pair, primer annealing and the primer extension temperatures were maintained at 65 °C.

2.8 Agarose gel electrophoresis

Seven precast agarose (1% w/v) gels and six precast agarose (2%) gels (1% and 2% E-Gel® EX, Invitrogen™, Cat. number: G401001) stained with SYBR® Gold II DNA were used for the electrophoresis of PCR products from the above procedures (Section 2.6). For the 2% and 1% precast agarose gels as the reference, 100-b ladder (4.0 µL, Low DNA Mass Ladder, Invitrogen™, Cat. number: 10068013) and 1 Kb ladder (4.0 µL, E-Gel 1Kb Plus DNA Ladder, Invitrogen™, Cat. Number:10488-090) were introduced, respectively to the well labelled as "M" of each gel, combining with nuclease free water (16.0 µL). PCR products (10.0 µL) derived through the PCR amplification in the above step with the expected amplification of 1000-2000 nt were separately introduced to the 1% agarose gels by combining with nuclease free water (10.0 µL) and for the 2% agarose gels each PCR product (10.0 µL) with expected amplification of 300-500 nt were introduced to each well separately by combining with PCR grade water (10.0 µL). The electrophoresis and the visualization were carried out using E-

Gel™ iBase™ and E-Gel™ Safe imager™ Combo Kit (Invitrogen™, G6465EU) for 10 min. at 45 V.

The resulting PCR products from the amplification of each primer pair in the above procedure (Section 2.7) were separately subjected to agarose gel electrophoresis using agarose gel (Et-Br, E-Gel® 48, Invitrogen, Thermo Fisher Scientific, Cat. No. G800802). Each PCR product (10 µL) was separately introduced to each well by combining with PCR grade nuclease free water (10 µL). As reference standards, two DNA size ladders [1 Kb DNA ladder (4 µL, E-Gel 1Kb Plus DNA Ladder, 0.035 µg/µL, Invitrogen™, Catalog Number:10488-090) and a 100 bp ladder (4 µL, Low DNA Mass Ladder, Invitrogen™, Catalog number: 10068013)] were introduced to 1st and 18th well positions of the gel, respectively and electroporated (10 min., 45.0 V) using E-Gel™ iBase™. Finally, electroporated gels were visualized and imaged under UV-light (λ = 280 nm) using a biomolecular imager (Image quanta 300, GE Healthcare).

2.9 PCR product purifications

DNA products derived through the PCR amplifications in the above procedure (Section 2.6) were separately subjected to PCR products purification using Agencourt® AMPure® XP beads (Product number: A63881, Beckman Coulter) with 0.6X based size cut clean-up by following the standard procedure given by the manufacturer (Section A.6., Annexure I).

2.10 Quantification of DNA using Qubit DNA analyser

The concentration of each DNA sample which derived through the purification steps of the above PCR reactions (Section 2.9) was separately subjected to dsDNA quantification using Qubit 4 Fluorometer (Invitrogen™, Thermo Fisher Scientific) by following the High sensitivity protocol. Preparation of samples were carried out using the Qubit™ 1x dsDNA High-sensitivity (HS) Assay Kit, (Invitrogen™, Thermo Fisher Scientific, Cat. number: Q33231) for each PCR product with compared to the PCR grade water.

2.11 RT-PCR on TBEV-Hochosterwitz strain RNA followed by PCR amplification against different overlapping primer pair schemes

RT-PCR products derived through four different dilutions (10^{-4}, 10^{-5}, 10^{-6} & 10^{-7}) from the above procedure (Section 2.6) were selected. PCR amplifications were carried out for each resulting cDNA (2.50 µL) from the above RT-PCR reactions against four different overlapping primer pair schemes (Sequencing experiment No. 1, Table A.3., Annexure I) using Q5® High-Fidelity PCR kit (New England BioLabs® Inc., Cat. No. E0555L). A total of four parallel NTC were also prepared and the PCRs conditions were maintained identical to the conditions mentioned in the above procedure (Section 2.6). Derived PCR products were separately subjected to PCR products purification using Agencourt® AMPure® XP beads (Product number: A63881, Beckman Coulter) and the concentration of each purified dsDNA sample was determined using Qubit 4 Fluorometer, (Invitrogen™, Thermo Fisher Scientific) instrument by following the dsDNA High sensitivity protocol according to the procedure (Section A.6, Annexure I & Section2.10), respectively.

2.12 DNA library preparation and Oxford Nanopore sequencing

The Oxford Nanopore flow cell (R9.4.1, Oxford Nanopore Technologies) was used for the sequencing with the GridION MK1 (GXB03008, Oxford Nanopore Technologies) instrument. Purified PCR products from the above procedure (Section 2.11) and four different purified PCR products at the dilution levels (1×10^{-4}) of TBEV-Hochosterwitz strain cDNA against JK_1-F/R and Eur_1-F/R primer pairs in the presence and the absence of tick RNA from the procedure (Section 2.6) with all together 12 samples each with 7.50 µL were subjected to Oxford Nanopore sequencing in GridION MK1 instrument for 52 hours by utilizing the Rapid Barcoding Sequencing kit (SQK-RBK004). Flow-cell priming and the DNA library preparation were carried out according to the manufacturer's protocols (Section A.7, Annexure I). (Details of barcode assignment for each sample are given in sequencing experiment No. 1, Table A.3., Annexure I).

2.13 Reference-based Oxford Nanopore reads mapping and data analysis

Reads derived through the Oxford Nanopore sequencing (Corresponding to 12 different barcode reads and also the unspecific set of reads, the minimum quality score for base-calling: 9) were combined according to the details given (Sequencing experiment No. 1, Table A.4., Annexure I) and subjected to trimming of primer sequences (Trim tool with primers only, Mismatches: 5 b, Minimum match length: 10 b, Geneious Prime®, version: 2022.0.1) followed by referenced based genome mapping against the TBEV-Hochosterwitz strain sequence (NCBI Acc. No. MT311861) using Minimap2 (version: 2.0.0) plugin in Geneious Prime® (version: 2022.0.1). The following parameters were maintained for each reference mapping against the TBEV reference genome with Minimap2 plugin (Data type: Oxford Nanopore, option: more sensitive, maximum secondary alignment: 1000, minimum secondary to primary alignment score ratio: 0.8, K-mer length: 5). SAM (Sequence alignment map) files generated for each reference-based mapping with Minimap2, the percentage of coverage to the reference sequence, the percentage of identical sites, percentage of pairwise identity to the mapped reads were tabulated. Consensus sequences were generated for each reference-based mapping using the SAM files by keeping the parameters: Threshold: Highest quality (60%), based on total coverage above 20 reads using the consensus sequence generator, Geneious Prime® (version: 2022.0.1). The consensus sequence derived through each reference-based mapping with each barcode and combined barcodes were separately subjected to pairwise alignment (Geneious alignment tool, Alignment type: Global alignment with free end gaps, cost matrix: 93% similarity, Gap open penalty: 12, Gap extension penalty: 3, with considering both directions) with the reference TBEV sequence (MT311861) and the pairwise identity and the gaps in the mapped genomes were determined. Finally the MSA file was subjected to phylogenetic tree construction against 250 different sequences of TBEV variants obtained from the NCBI nucleotide database by the MEGA software (Version 11.0.11, Tamura *et al.,* 2021) with following parameters; statistical method: Maximum Likelihood, test of phylogeny: Bootstrap method (replicates: 1000), substitution model: Kimura 2-parameter model).

2.14 *De novo* sequence assembly followed by reference-based sequence mapping

Reads derived through the Oxford Nanopore sequencing (Corresponding to each barcode, FASTQ file format) were separately subjected to quality analysis using NanoPlot tool (Version: 1.28.2, De Coster *et al.*, 2018) in the Galaxy online server (version: 21.09, "https://usegalaxy.org.au/", Afgan *et al.*, 2018). After the quality analysis, file format conversion was carried out using FASTQ to FASTA file conversion tool (version: 1.15, Blankenberg *et al.*, 2010) and data corresponding to each barcode was concatenated into a single FASTA file and subjected to quality trimming to remove any anchored remaining barcode sequences and primer sequences using the Porechop tool (version: 0.2.4, Wick, 2017), before continuing to the de novo assembly. The quality trimmed FATSA files were separately (based on the barcode used to tag the sequences) submitted to the de novo assembly using Trinity tool (version: 2.9.1, Grabherr *et al.*, 2011) with the minimum contigs length of 20 bases for single pair reeds and leaving the rest of the parameters as default values. Resulting assembled short reads were mapped against reference sequences (250 different TBEV genomes obtained from NCBI) using Minimap2 (version: 2.0.0) plugin in Geneious Prime® (version: 2022.0.1) with the identical set of parameters mentioned in the above procedure (Section 2.13). Sam files generated from each reference-based mapping with Minimap2, the percentage of coverage to the reference sequence, the percentage of identical sites, percentage of pairwise identity to the mapped reads were tabulated to identify the best possible reference genome sequence of TBEV. The consensus sequence derived through each reference-based mapping with each barcode and combined barcodes were separately subjected to pairwise alignment (Geneious Prime® alignment tool, Alignment type: Global alignment with free end gaps, cost matrix: 93% similarity, Gap open penalty: 12, Gap extension penalty: 3, with considering both directions) with each respective reference TBEV genome and the pairwise identity and the gaps in the mapped genomes were determined.

2.15 Evaluation of the quality of mapped genomes

The SAM (Sequence alignment map) files generated from each reference-based mapping with Minimap2, the percentage of coverage to the reference sequence TBEV-Hochosterwitz, percentage of identical sites, percentage of pairwise identity to the mapped reads were tabulated though the data generated by the Geneious Prime® (version: 2022.0.1) to identify the best possible reference genome sequence of TBEV and to evaluate each primer pair efficiency and the quality of the reconstructed genome. Pairwise alignments were carried out for the consensus sequence derived through each reference-based mapping with each barcode and through the combined barcodes using Geneious alignment tool by keeping the following parameters, Alignment type: Global alignment with free end gaps, cost matrix: 93% similarity, Gap open penalty: 12, Gap extension penalty: 3, considering aliment through both directions).

The SAM files which generated for each of the reference based mapping to the TBEV-Hochosterwitz (NCBI, Acc. No. MT311861) derived from each of the barcode were subjected to average sequencing depth, per base sequencing depth and the breadth of sequencing above 20 reads analysis with SAMtools (Danecek *et al.*, 2021, Version: 1.3.1+htslib-1.3.1) installed with an IDE (*Anaconda Software Distribution*, 2020, Version: 4.11.0) in a Linux Mint (Version: 20.3) operating system.

2.16 Identification of poorly performed primer pairs through sequence assembly map (SAM) files

The relative performance of each primer pair used in the study was determined using two different approaches. The first approach by the determination of primer pair efficiency through the determination of number of sequences generated by each primer pair using the SAM files for each concentration (Python program was written to determine the number of sequences generated under each primer pair) and finally through the linear regression analysis using the Microsoft® Excel® [Microsoft 365 MSO (Version 2205 Build 16.0.15225.20172) 64-bit], with the selection of intercept values of each equation (primer pair efficiency) and keeping the threshold cut off (< 1.00) value. The second approach was by through the determination of per-base depths of the mapped genomes using SAMtools (Threshold < 50 reads) (Danecek *et al.*, 2021, Version: 1.3.1+htslib-1.3.1). Finally, through the relative primer efficiencies, poorly

performing primers were identified and compared with the primer sequence identities (Threshold: < 95%) to the reference TBEV-Hochosterwitz (NCBI, Acc. No. MT31186).

2.17 Designing new overlapping PCR primers

To replace the low efficient primers identified in the above procedure (Section 2.16), 30 different new overlapping primers (Table A.2, Annexure I) were designed according to the procedure mentioned in a previous study (Bandara, 2021). According to the procedure, first a multiple sequence alignment (Geneious alignment tool, Alignment type: Global alignment with free end gaps, cost matrix: 93% similarity, Gap open penalty: 12, Gap extension penalty: 3, with considering both directions, Geneious Prime® 2022.0.1) was carried out for 249 different TBEV sequences obtained from the NCBI database (Table A.5., Annexure I). Consensus sequence and the per base identity values corresponding to each nucleotide of the multiple sequence alignment were exported to a comma separated (CSV) file using a Python program written. Using the Python program (Bandara, 2021), conserved score values containing output CSV file was generated by keeping the primer length at 20 nt. Based on the conserved score value, GC% and Tm [Calculation of melting temperature (Tm) was carried out using the online Tm calculator from the NEW ENGLAND BioLabs® inc., according to the requirements of Q5® High-Fidelity PCR kit (New England BioLabs® Inc., Cat. No. E0555L) (https://tmcalculator.neb.com/#!/main)] (Viewed: 11/09/2022), the values of the output CSV file data, 42 different new primers were designed to replace the poorly performing primers identified in the above procedure (Section 2.16). Finally, self-dimerization and heterodimerization abilities of each primer pairs were calculated using IDT OligoAnalyzer (https://eu.idtdna.com/calc/analyzer) (Viewed: 11/09/2022) to select the best suitable primer pairs for the whole genome amplification of TBEV.

2.18 PCR amplification of TBEV-Hochosterwitz cDNA against each new primer pair

To evaluate the amplification ability of TBEV-Hochosterwitz cDNA by the new primers (CJB primer pairs, Table A.2., Annexure I) designed to replace the identified poorly performing primers, each newly designed primer pair was separately subjected to two separate PCR amplifications using Q5® High-Fidelity PCR kit (New England BioLabs® Inc., Cat. No. E0555L) by following the identical procedure (Section 2.7).

2.19 Evaluation of Oxford Nanopore sequenced data derived through the combinations of different overlapping primer schemes

RT-PCR products of four different dilutions (10^{-4}, 10^{-5} & 10^{-6}) prepared in the above procedure (Section 2.6) were selected. PCR amplifications were carried out for the resulting cDNA (2.50 µL) against 16 different overlapping primer schemes (Sequencing experiments No. 2 & 3, Table A.3., Annexure I), using Q5® High-Fidelity PCR kit (New England BioLabs® Inc., Cat. No. E0555L) with eight NTC according to the procedure given in the above (Section 2.6) with the only additional change in the thermal cycler profile, the primer extension and annealing temperature were maintained at 63 °C. PCR products purifications and subsequent DNA quantifications were carried out according to the procedures mentioned in the above (Sections 2.9 & 2.10), respectively. The resulting purified PCR products derived through 8 different primer schemes with three different concentrations (10^{-4}, 10^{-5} & 10^{-6}), paired together (Details related to the combinations of PCR products given in the Sequencing experiment No. 2 & 3, Table A.4., Annexure I) to generate 24 different samples belonging to eight different combined primer schemes in each concentration. DNA library preparation followed by Oxford Nanopore sequencing were carried out according to the procedure given (Section 2.12). Reads derived through the Oxford Nanopore sequencing (Minimum quality score: 9) were separately subjected to reference-based genome mapping (NCBI Acc. No. MT311861) followed by the sequencing quality analysis according to the procedure given in the above (Section 2.13) to identify the best PCR conditions for the DNA library development.

2.20 Total RNA extractions from nymph tick samples

A total of 76 cryovials containing pools of ten nymph ticks, each stored at -80 °C were separately transferred into lysing matrix tubes (2 ml) consist of six steal beads (Ø = 3.8 mm) (MP biomedicals, Santa Ana, California, USA) and Minimum Essential Medium (MEM) (400 µL, Cat. No. 11095080, Gibco™, Thermo Fisher Scientific). Homogenization was carried out with the CY-24 option for 40 S using Fast-Prep-24™ (MP biomedicals, Santa Ana, California, USA) and cell debris were removed by the centrifugation (14000 rpm, 5 min) (Eppendorf Centrifuge 5430, Hamburg, Germany). From each supernatant, an aliquot (about 200 µL) was separately transferred into 76 cryovials containing fetal bovine serum (FBS) (10%, Cat. No. 11095080, Gibco™, Thermo Fisher Scientific) and stored at -80 °C for cultivations with cell culture. The remaining volume (200 µL) was lysed by RTL buffer (200 µL, QIAGEN RNeasy mini kit, Cat. No. 74106, QIAGEN, Hilden, Germany). Total RNA extraction was performed by QIAcube automated RNA extractor using the RNeasy mini kit (Cat. No. 74106, QIAGEN, Hilden, Germany) according to protocol given by the manufacturer (Section A.8, Annexure I). The resulting RNA extracts were immediately revers transcribed into cDNA. The remaining RNA extracts were stored at -80 °C for further analysis.

2.21 Reverse transcription of total RNA from tick samples

Total RNA extracts (1.0 µL) from each 76 nymph tick pools were separately subjected to cDNA synthesis using both High-Capacity cDNA reverse transcription kit (Applied Biosystems™, Thermo Fisher Scientific, Cat. number: 4368813) and SuperScript™ IV First-Strand synthesis system, (Thermo Fisher Scientific, Cat. number: 18091050) according to the procedure (Sections A.1 & A.4, Annexure I). In addition, total RNA extracts (1.0 µL) from each tick sample of three female ticks and seven male ticks received from another study (Carried out by another MSc. student attached to the NIPH, Miss. Eileen Stanzen), were also subjected to cDNA synthesis using both methods (Sections A.1 & A.4, Annexure I).

2.22 Realtime semiquantitative PCR on tick cDNA and the detection of TBEV cDNA

Two separate methods of semiquantitative real-time RT-PCR were carried out to identify the TBEV positive total RNA extracts of nymph pools (76), adult female (46) and adult male tick samples (50). Semiquantitative real-time PCR method I was separately carried out for each of the 76 different nymph tick cDNA samples (3.0 µL from each) derived from High-Capacity cDNA reverse transcriptase reactions following the identical conditions as mentioned in the above procedure (Section 2.5).

Semi-quantitative RT-qPCR method II (Ryan Easterday and Kristian Alfsnes) was separately carried out for each total RNA extracts of nymph ticks (76), total RNA extracts of female ticks (46) and total RNA extracts of male ticks (50), separately with individual reaction consists with total tick(s) RNA (1.0 µL), TaqPath™ 1-Step RT-qPCR Master Mix, CG (2.50 µL, Cat. No. A15299, Applied Biosystems®, Life Technologies™, Thermo Fisher Scientific), forward primer [TBE320F, 0.60 µL, 25 µM, 5'-GGGAGCGCAAAACTGGAA-3', Acc.no (NCBI): KF991107: 1628-1645], reverse primer [TBE373R, 0.60 µL 25 µM, 5'-TGAGGAGCCCCAAATTCAAC-3', Acc.no (NCBI): KF991107: 1681-1662], probe [0.13 µL, TaqMan probe was designed with 5'-Fluorescein amidite (6-FAM™) labeled -3'-minor groove binding protein-eclips (MGB-eclips) anchored (TBE339-probe, FAM-5'-AACGCAGAAAGAC-3'-MGB-eclips), to target the TBEV envelope protein coding region of the RNA [1647-1659 (reference to the TBEV genome, Acc. No. KF991107, NCBI)] and the PCR grade water (5.17 µL). The actual real-time PCR was carried out on the Corbett Rotor-gene™ 6000 (Qiagen Corbett, Hilden, Germany) with the initial reverse transcription related random hexamer primer annealing at 25 °C for 10 minutes, the reverse transcription and the cDNA synthesis at 50 °C for 15 minute, the inactivation of Reverse transcriptase, denaturation of DNA and the TaqPath™ enzyme activation at 95 °C for 2 minutes and followed by 40 cycles: 95 °C for 3 seconds, 60 °C for 30 seconds for the denaturation, annealing of the primers and for the extension; at the end of the cycles storing at 4 °C. Fluorescent data acquired by the instrument software Rotor-Gene Q Software (Version: 2.3.1.49), at the very beginning and also at each cycle in the primer annealing and the extension stages were used to determine the cycle threshold (Ct) value of each sample by keeping the normalized fluorescence threshold cutoff value (0.02). Through the results derived by the semiquantitative real-time PCR, the

prevalence of TBEV in nymph tick pools, female adult ticks and male adult ticks in the sample collected site near to Larvik in Norway were calculated using EPITOOLS (Sergeant, 2018).

2.23 PCR amplification of cDNA derived from tick samples, PCR product purification and dsDNA quantifications

Through the results generated by the real time semiquantitative PCR in the above procedure, TBEV positive cDNA samples derived from each of the TBEV positive RNA extracts (Section 2.21) with six NTC were separately subjected to PCR amplification against six different primer schemes [(A, 0.5 µM, 1.25 µL), (B, 0.5 µM, 1.25 µL) and Sequencing experiment No. 4, 5 & 6, Table A. 3., Annexure I] using Q5® High-Fidelity PCR kit (New England BioLabs® Inc., Cat. No. E0555L) according to the procedure mentioned (Section A.5, Annexure I). The resulting PCR products were separately subjected to PCR product purification using Agencourt® AMPure® XP beads, following the identical procedure (Section A.6 Annexure I). Purified PCR products were separately subjected to dsDNA quantification using Qubit 4 Fluorometer, (Invitrogen™, Thermo Fisher Scientific) instrument by following the High sensitivity protocol as mentioned in the procedure (Section 2.10).

2.24 DNA library preparation, Oxford Nanopore sequencing and reference-based assembling

Purified PCR products derived from each of the PCR reactions in the above procedure (Section 2.23) were selected for Oxford Nanopore sequencing based on the dsDNA concentration values and to represent each of TBEV positive PCR amplified tick samples through two combined primer schemes (Sequencing experiment No. 4, 5 & 6, Table A.4., Annexure I). The DNA library preparation, Oxford Nanopore sequencing and reference-based mapping were carried out according to the identical procedures mentioned in the above (Sections 2.12 & 2.13), respectively.

2.25 Optimization of RT-PCR, PCR, DNA library preparations followed by Oxford Nanopore sequencing and reference-based assembling of TBEV positive tick samples

All the TBEV positive RNA extract (8.0 µL) from each tick sample in the above procedure (Section 2.20), were used for the cDNA synthesis using SuperScript™ IV First-Strand synthesis system, (Thermo Fisher Scientific, Cat. number: 18091050), according to the procedure (Section 2.21) with additional minor changes. The amount of template RNA from each TBEV positive tick sample was increased from 1.0 µL to 8.0 µL as the modification to the procedure and the rest of the reaction conditions were maintained as the same as described in the procedure (Section 2.21).

Each of the cDNA (7.0 µL) derived from the above RT-PCR reactions was subjected to PCR amplification against four different primer pools with the respective primer concentrations (Sequencing experiment No. 7, Table A.3., Annexure I). PCR product purifications and DNA quantifications were carried out according to the identical procedures (Section A.6, Annexure I & Section 2.10), respectively. DNA library preparation was carried out for the purified PCR products according to the procedure (Section 2.12) mentioned in the above with a minor modification, the volume of each purified PCR product was increased from 7.50 µL to 20.0 µL for each barcode with appropriate proportionate changes in the volumes of other reagents used in the procedure and finally Oxford Nanopore sequencing was carried out by following the procedure (Section 2.12). Reference based sequence mapping was carried out for reads derived under each barcode, against 250 different nucleotide sequences of TBEV variants (Table A.5., Annexure I) obtained from the NCBI nucleotide database, according to the procedure (Section 2.13). The resulting consensus sequences derived from each mapping against TBEV variants corresponding to each barcode were sorted to identify the highest number of reads mapped (Manually curated to identify and to remove the non-TBEV related poly-(A) regions containing reads mapped to each reference genome). Top ~10 reference sequences with the heist reads mapped were separately selected corresponding to each barcode and separately mapped against the identified and selected reference TBEV sequences according to the procedure (Section 2.13).

2.26 Identification of TBEV viral subtypes

The resulting SAM files derived through each reference-based mapping in the above procedure (Section 2.25) were separately subjected to generate consensus sequences by maintaining the per-base sequencing depth above 20 reads. Multiple sequence alignments were carried out for each of the consensus sequences against 250 different sequences of TBEV variants obtained from the NCBI nucleotide database. (Table A.7., Annexure I). As the MSA tool, MAFFT alignment option of the Geneious Prime® (Geneious Prime®, 2022.0.1) was used with the following parameters: Algorithm : FFT-NS-ix1000, scoring matrix: 1PAM/k=2, Gap open penalty: 1.53, Offset value: 0.123, with considering both directions. The resulting sequence alignments were separately subjected to phylogenetic tree constructions using PhyML (version: 3.3.20180621) tool of the Geneious Prime® 2022.0.1 by maintaining the following parameters: Substitution model: K80, Tamura, Tree, Outgroup: Louping ill virus (NCBI Acc. No. NC_001809), Branch support: Bootstrap, Number of bootstraps: 1000, Substitution rate: 4 and all the parameters by the estimations. Finally, the genetically closest TBEV known subtype was identified through the smallest number of substitutions per sites. To evaluate the accuracy of the phylogenetic method used in this study, the nucleotide sequence of reconstructed TBVE-Hochosterwitz strain using combine primer scheme [(G+H)+(S+T)] at the tenfold dilution (log10: -6) was incorporated into the MSA and in the phylogenetic tree construction, [MT311861_GSHT_(-6)].

2.27 Identification of tick species

The generated unmapped sequences of each reference based mapping, MegaBLAST (Altschul *et al.,* 1990) (BLAST plugin, Geneious Prime®, 2022.0.1) sequence filtering was carried out using a local BLAST database created with the sequences obtained from the NCBI nucleotide archive by searching under "18S rRNA *Ixodes"* species (Total of 1043 individual sequences). Initially, BLAST-hits were obtained as hits and non-hits by maintaining the MegaBLAST parameters (Max E-value: $1\times e^{-300}$, word size: 20, max target steps:100). MegaBLAST analysis was carried out again on the separated BLAST-hist to generate hit table with full sequence annotations by maintaining the Max E-value: $1\times e^{-300}$ and to receive one single BLAST-hit. Resulting BLAST-hits with annotations were grouped according to the name of each Tick-

species in the description. Each grouped sequences were mapped against the corresponding reference sequences identified through BLAST-hits using Minimap2 (Geneious Prime®, 2022.0.1) with identical set of parameters mentioned in the above (Section 2.13) (Table A.16., Annexure I). Generated consensus sequences were separately subjected to multiple sequence alignment against 22 different sequences of 18S-rRNA regions of tick species obtained from the NCBI nucleotide database. (Table A.6., Annexure I). As the MSA tool, MAFFT alignment option of the Geneious Prime® (Geneious Prime®, 2022.0.1) was used with the following parameters: algorithm: FFT-NS-ix1000, scoring matrix: 1PAM/k=2, Gap open penalty: 1.53, Offset value: 0.123, with considering both directions. The resulting sequence alignment was subjected to phylogenetic tree constructions using PhyML tool in the Geneious Prime® 2022.0.1 by maintaining the following parameters: Substitution model: K80: Kimura, Branch support: Bootstrap (1000), Number of substitution rate: 4, Optimize based on substitution rate and assigned an outgroup: *Argas lahorensis* (NCBI Acc. No. AGG18SRA) 18S-rRNA gene, Finally, genetically closest tick species was identified through the smallest number of substitutions per sites. Selection of the 18S-rRNA region was decided base on the literature reference (Mangold *et al.*, 1997) and also based on the availability of 18S-rRNA gene fragments of the Oxford Nanopore sequence reads.

3. RESULTS

The TBEV-Hochosterwitz strain used in this study was obtained from the Norwegian Institute of Public Health (NIPH), Oslo and the study was a continuation of a previous similar study carried out at the institution. The first section of the results is mainly related to the identification of limitations associated with the already existing method. The second section is related to the optimization of RT-PCR, RT-qPCR, PCR and DNA library development techniques to sequence the whole genome of TBEV-Hochosterwitz stain in very low viraemic samples using Oxford Nanopore sequencing. The final part of the results associated with the evaluation of whole genome amplification and sequencing capability of the optimized PCR method was performed on low viraemic TBEV positive tick samples received from a field survey carried out by the NIPH in Larvik, Norway.

3.1 Total RNA quantities in the preserved RNA extracts from the cell cultures

Quantitative determination of total RNA from the seven Vero E6 cell cultures incubated with TBEV-Hochosterwitz strain (NCBI Acc. No. MT311861) were carried out and tabulated to compare the RNA quantity and the stability based on the RNA extraction techniques employed in previous studies (Table 3.1.). The recently extracted (17th June 2021) sample (Sample No. 6, Table 3.1.) exhibited the highest amount of total RNA (179.80 ng/µL**)** compared to all the other extracts. Among the other RNA extracts (Extractions carried out in 2016), the highest amount of RNA was preserved in the sample number five (7.22 ng/µL) and the RNA extraction was carried out by the RNeasy Mini kit for animal, tissue and cells (Qiagen inc. Valencia, CA, USA) after the lysis with AVL buffer (Cat. No. 19073). The amount of total RNA preserved in sample number one (7.20 ng/µL) also contained almost equivalent amount of total RNA present in sample number five, extracted with the Trizol™ reagent (Cat. No. 15596026, Invitrogen, CA). The lowest amount of total RNA was preserved with sample number three (2.44 ng/µL), extracted with RTL lysis buffer (200 µL, Cat. No. 79216, Qiagen, CA, USA) added 1% β-mercaptoethanol (Table 3.1.).

Table 3.1. Total RNA quantities measured at 260 nm and the Ct values determined by two similar Semiquantitative RT-qPCR analysis of TBEV-Hochosterwitz RNA

Sample number	Sample name	RNA concentration (ng/µL)	Cycle threshold	
			Ct*	Ct**
1	Hochosterwitz_Trizol_11-04-2016	7.20	13.05	
2	Hochosterwitz_Qiagen_11-04-2016	5.81	13.42	
3	Hochosterwitz_RLT_28-04-2016	2.44	12.15	
4	Hochosterwitz_250416_26-04-2016	5.29	8.83	
5	Hochosterwitz_2_AVL_28-04-2016	7.22	8.85	
6	Hochosterwitz_ 17-06-2021_MAJD	179.80	7.62	
7	Hochosterwitz_2_ 26-04_2016	6.68	9.01	
8	Sample number 7 (10^{-1})		15.85	14.41
9	Sample number 7 (10^{-2})		18.19	16.48
10	Sample number 7 (10^{-3})		22.03	20.67
11	Sample number 7 (10^{-4})		25.28	23.57
12	Sample number 7 (10^{-5})		29.89	27.46
13	Sample number 7 (10^{-6})		34.04	32.13
14	Sample number 7 (10^{-7})		40.54	37.98
15	Sample number 7 (10^{-8})		--	44.61
16	Sample number 7 (10^{-9})		--	--
17	Sample number 7 (10^{-10})		--	--
18	NTC-1		--	--
19	NTC-2		--	--
20	NTC-3		--	--
21	NTC-4		--	--
22	NTC-5		--	--

Blank (“ “): Not determined, NTC: Nontemplate control samples, (--): Below the threshold level, Ct: Semiquantitative RT-qPCR method I, Ct**: Semiquantitative RT-qPCR (method II).*

3.2 Relative quantities of TEBV-Hochosterwitz RNA

The relative quantities of TBEV-Hochosterwitz RNA in total RNA extracts and also in the tenfold dilutions were determined using the Ct values received by two similar semiquatitative real-time-PCR techniques (Table 3.1.). The highest amount of TBEV-Hochosterwitz RNA was detected in the recently extracted RNA sample number six, exhibiting the lowest Ct vlaue (Ct = 7.62). Among the samples prepared in 2016, the highest amount of RNA preserved in

two samples (Sample No. 4 & 5) extracted through the RNeasy Mini kit for animal tissue and cells (Qiagen inc. Valencia, CA, USA) and Lysis with RTL buffer followed by the extraction through the RNeasy Mini kit for animal tissue and cells (Qiagen inc. Valencia, CA, USA), respectively. The lowest amount of TBEV-Hochosterwitz RNA was exhibited in the sample number two (Ct = 13.42) extracted by the RNeasy Mini kit for animal tissue and cells and the sample number one (Ct = 13.05), and the sample number three (Ct = 12.15) also contained reasonably low amounts of TBEV-Hochosterwitz RNA extracted with Trizol™ reagent and with RTL lysis buffer (1%, β-mercaptoethanol), respectively. The sample number seven contained moderate amount of TBEV-Hochosterwitz RNA (Ct = 9.01) and the following tenfold dilutions exhibited proper linearity with the received Ct values (R^2 = 0.9975, Figure A.8., Annexure I) carried out using semiquantitative RT-qPCR method I. The Ct values of the tenfold dilutions obtained through the semiquantitative RT-qPCR method II, exhibited significantly different Ct values compared to the RT-qPCR method I, with preserved linearity (R^2 = 0.9973).

3.3 Effects of type of primers and reverse transcriptase on the cDNA synthesis

The effects of the primers and the type of reverse transcriptase used in the cDNA synthesis of TBEV-Hochosterwitz RNA were determined by the semiquantitative RT-qPCR method. The cDNA derived through the reverse transcription of seven different TBEV-Hochosterwitz RNA extracts (Table 2.1.) against random hexamers and TBEV-specific primers (Table A.7., Annexure I) using both High-Capacity cDNA reverse transcriptase and the SuperScript™ IV First-Strand synthesis system, were subjected to semiquantitative RT-qPCR analysis. The lowest Ct values were exhibited for the cDNA samples derived with random-hexamers and the High-Capacity cDNA reverse transcriptase for all the seven samples indicating proper cDNA synthesis compared to the other combinations. The cDNA samples derived through the SuperScript™ IV First-Strand synthesis system against random hexamers exhibited higher Ct values for all the seven total RNA extracts as well (Figure 3.1 & Table A.7, Annexure I). To find out any significant difference between Ct values given by the two methods, single factor analysis of variance (ANOVA) test was carried out and found no significant differences between the two methods giving p-value ($p = 0.065 > 0.05$).

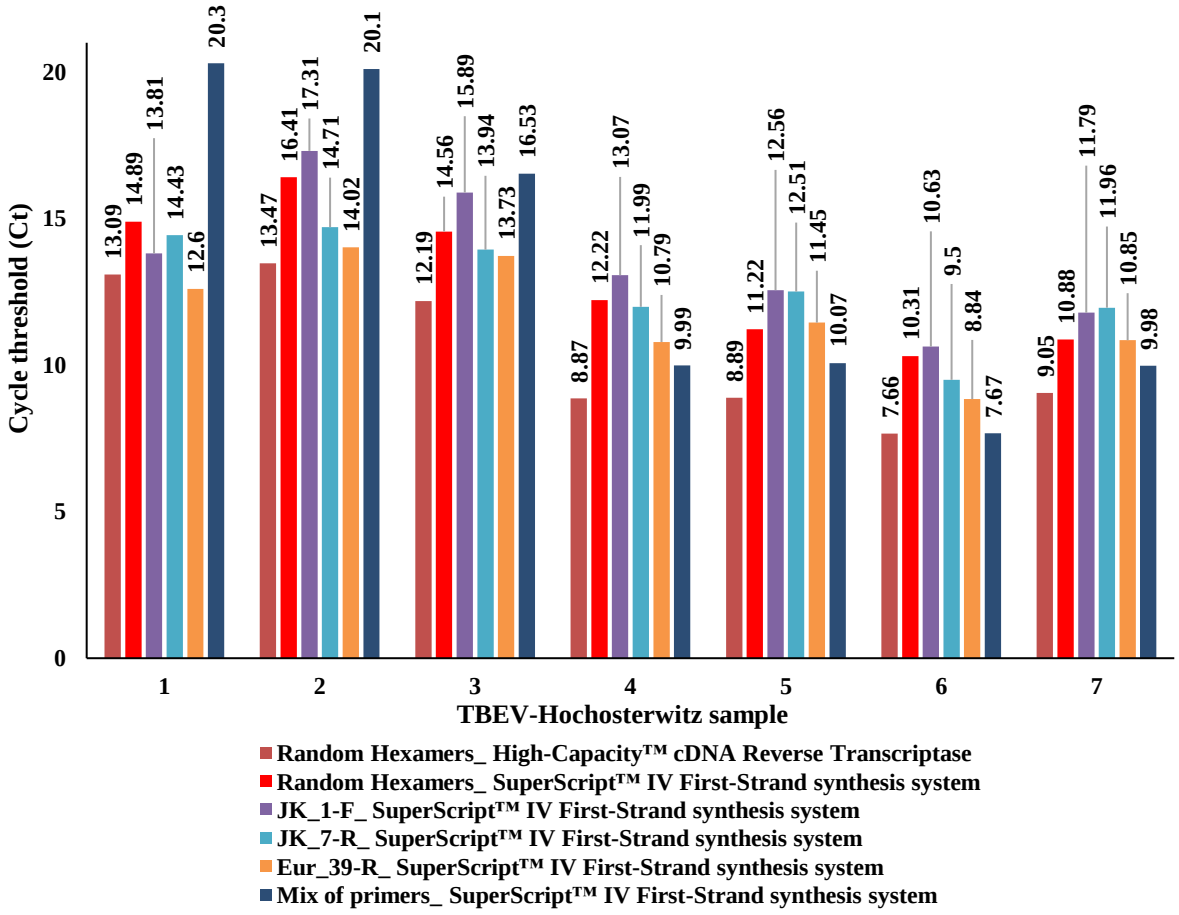


Figure 3.1. The effect of primer compositions and the type of reverse transcriptase on the cDNA synthesis of TBEV-Hochosterwitz RNA, determined by the semiquantitative real-time PCR (Table A.7., Annexure I).

3.4 Effects of primer pair concentrations on the PCR amplification of TBEV-Hochosterwitz cDNA

Effects of primer concentrations (0.15 µM & 0.5 µM) on the amplification of tenfold dilutions of TBEV-Hochosterwitz cDNA were determined using PCRs carried out against five different primer pairs (JK_1-F/R, JK_3-F/R, JK_7-R/R, Eur_19-F/R and Eur_39-F/R). The dsDNA quantities of purified PCR products were determined and compared to identify any improvement with the increase of primer concentration on the amplified products (Figure 3.2), (Table A.8., Annexure I).

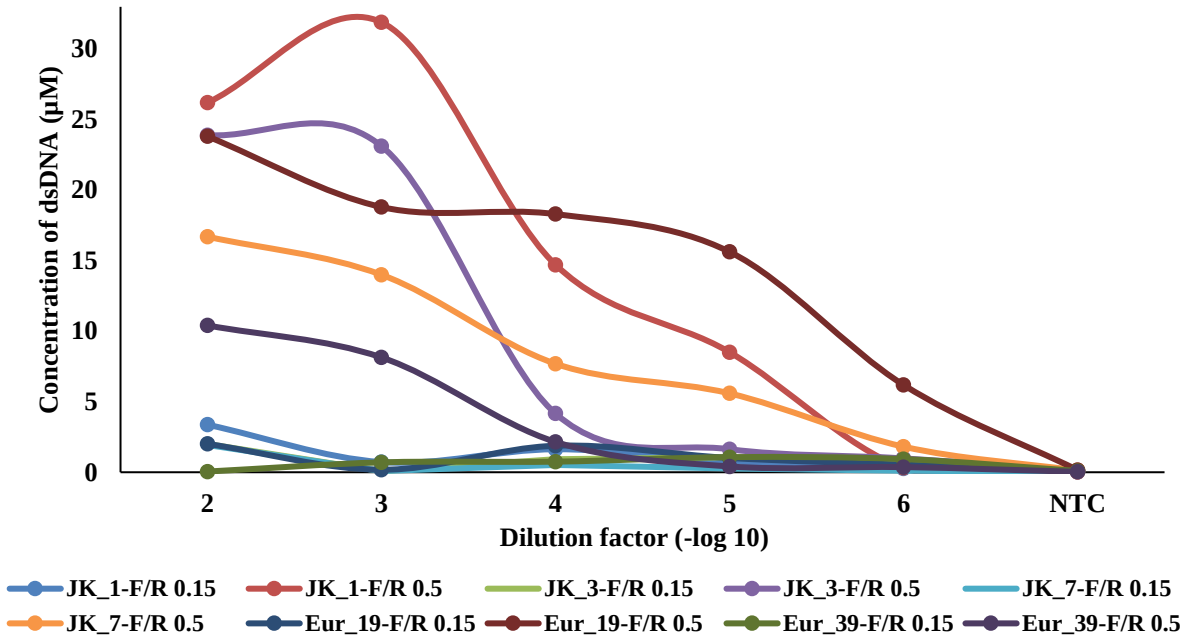


Figure 3.2. Changes in the purified dsDNA concentration in the PCR products derived through five different tenfold serial dilutions of TBEV-Hochosterwitz cDNA, amplified against five different primer pairs at two different primer concentrations (0.15 µM & 0.5 µM) (Data source: Table A.7., Annexure I).

Amplified PCR products derived through the higher primer pair concentrations (0.5 µM) exhibited higher quantity of dsDNA than the PCR products derived through the lower primer pair concentrations (0.15 µM) (Figure 3.2). The highest quantity of dsDNA exhibited in the amplified PCR products derived through the primer pair (JK_1-F/R) and after the initial two dilutions, dsDNA amount decreased rapidly and the sequence of changes were similar with the dsDNA amounts derived from the PCR reaction against JK_3-F/R primer pair at different dilutions. All the nontemplate controls (NTC) were performed without adding template TBEV-Hochosterwitz cDNA, exhibited very less quantity of dsDNA, indicating the amount of primer dimer formations were significantly low among the compared primer pairs.

3.5 Effect of temperature on the PCR amplification of TBEV-Hochosterwitz cDNA

Visualized electrophoresed agarose gels (2%) containing the PCR products from the TBEV-Hochosterwitz cDNA at the dilution level of (1×10^{-4}), amplified against each primer pairs at two different temperatures (63 & 65 °C), exhibited differences in PCR bands corresponding to the unspecific amplifications. The number of bands and the intensities of the unspecific bands in the PCR carried out at 63 °C was higher than the respective products of the PCR carried out at 65 °C (Figure 3.3.). The intensities of the expected PCR products were high in the PCR carried out at 63 °C compared to 65 °C. The expected PCR products from the primer pairs (Eur_6-F/R and the Eur_11-F/R) were absent when the PCR amplification carried out at 63 °C, but only a very small quantity of amplified products appeared in the PCR amplification carried out at 65 °C (Figure 3.3.). Similarly, PCR products derived from the primer pairs (Eur_29-F/R and Eur_32-F/R) were higher in the PCR carried out at 65 °C than at 63 °C. A Considerable amounts of unspecific amplifications were identified in the PCR products derived from the primer pairs (Eur_13-F/R, Eur_15-F/R, Eur_16-F/R, Eur_29-F/R & Eur_30-F/R) and showed at least one additional shorter band appearing closer to the main PCR product at 63 °C. But the corresponding unspecific amplified products were absent in the visualized agarose gel images when the PCR was carried out at 65 °C, except for the primer pair Eur_16-F/R and Eur_29-F/R. Very short unspecific PCR products (< 100 nt) were found among the PCR products derived through the PCR primer pair (Eur_6-F/R, Eur_12-F/R, Eur_14-F/R, Eur_21-F/R to Eur_30-F/R) carried out at both temperatures. This is assumed to be due to dimerized products of primers.

Even though the corresponding sequence positions are not in the previously published TBEV-Hochosterwitz sequence (Accession number MT311861), PCR products with appropriate size lengths were exhibited in the analysis, derived through the primer pairs (Eur_38-F/R & Eur_39-F/R). This indicates the possibility of the detection of an extended region of the TBEV-Hochosterwitz genome. The NTC derived from the RT-real-time PCR followed by PCR reactions, did not contain any amplified PCR bands by indicating the lack of any cross contaminations during the RT-PCR and PCR steps of the analysis.

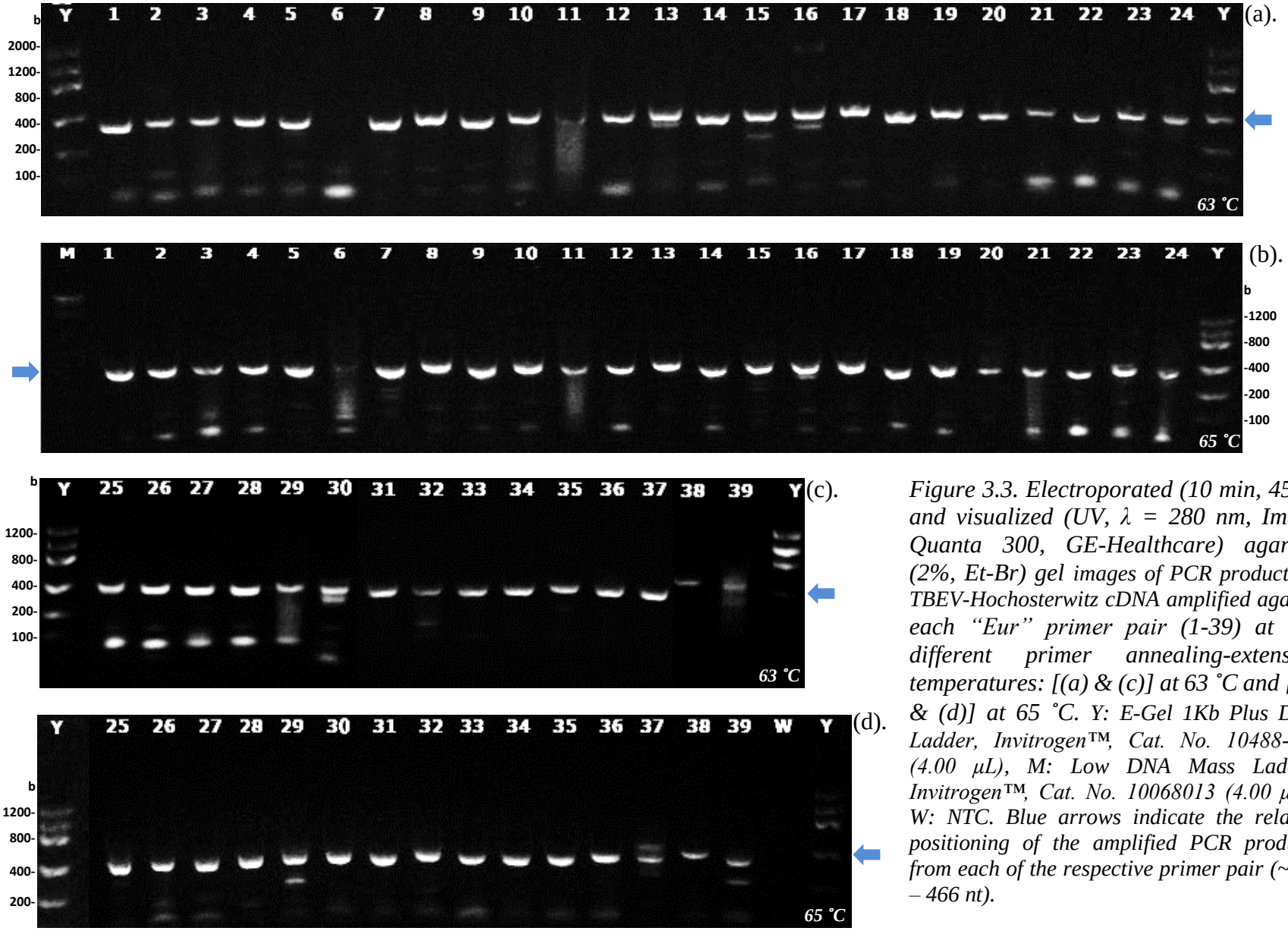


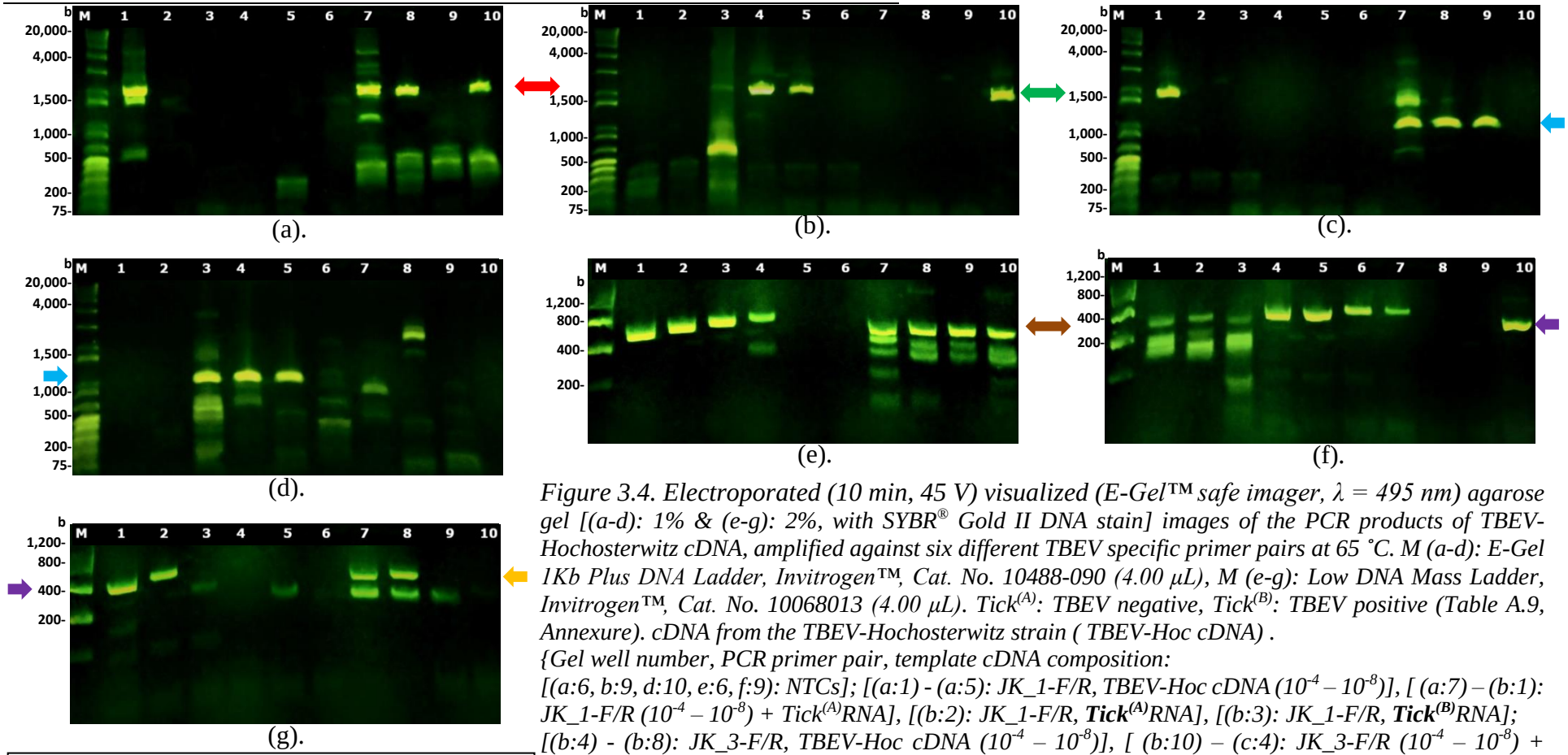
Figure 3.3. Electroporated (10 min, 45 V) and visualized (UV, λ = 280 nm, Image Quanta 300, GE-Healthcare) agarose (2%, Et-Br) gel images of PCR products of TBEV-Hochosterwitz cDNA amplified against each "Eur" primer pair (1-39) at two different primer annealing-extension temperatures: [(a) & (c)] at 63 ˚C and [(b) & (d)] at 65 ˚C. Y: E-Gel 1Kb Plus DNA Ladder, Invitrogen™, Cat. No. 10488-090 (4.00 µL), M: Low DNA Mass Ladder, Invitrogen™, Cat. No. 10068013 (4.00 µL,), W: NTC. Blue arrows indicate the relative positioning of the amplified PCR products from each of the respective primer pair (~363 – 466 nt).

3.6 Effect of total RNA from tick samples

Five tenfold serial dilutions (1×10^{-4} – 1×10^{-8}) of TBEV-Hochosterwitz RNA were reverse transcribed separately with /or/ without the addition of total RNA from a TBEV negative tick sample. In addition, the same experiment was performed with /or/ without the addition of total RNA from a TBEV positive tick sample. The resulting cDNA were amplified using PCR with six different TBEV specific primer pairs (Section 2.6). Each PCR product was subjected to agarose gel electrophoresis, PCR product purification and followed by dsDNA quantification. The agarose gels were analyzed for the interfering effect from the tick RNA and for the detection of amplified TBEV-Hochosterwitz regions against five different TBEV specific primer pairs under five different dilutions (1×10^{-4} – 1×10^{-8}). In the visualized electrophoresed agarose gels, all the "JK" primers (JK_1-F/R, JK_3-F/R & JK_39-F/R) performed well with the higher concentration against cDNA of reverse transcribed TBEV-Hochosterwitz RNA with dilutions ranging from 1×10^{-4} to 1×10^{-5} and giving corresponding appropriate size lengths around (1961-1851 nt) with compared to the reference E-Gel 1Kb Plus DNA Ladder, Invitrogen™, Cat. No. 10488-090 [Figure 3.4, (a), (b), (c) & (d)]. Corresponding amplified PCR products were absent when the template RNA concentrations fell below the dilution level (1×10^{-5}). In the visualized gel bands corresponding to the "JK" primers, interference from the tick RNA extract were absent and only the unspecific amplifications with the shorter PCR products were present. The PCR bands and their intensities in corresponding amplified products had no significant differences in the presence and absence of tick RNA, indicating uniform amplification until the dilutions up to 1×10^{-5} level. In the further diluted samples, PCR bands related to the unspecific amplifications were more dominant and did not contain any expected PCR bands. The PCR products derived through the cDNA of TBEV positive tick sample exhibited very faint PCR band against only the "JK_3-F/R" primer pair with corresponding electroporated distance to the PCR band amplified through the same primer pair from the TBEV-Hochosterwitz RNA Figure 3.4, [(a), (b), (c) & (d)]. Corresponding PCR bands derived through the PCR amplification of TBEV-Hochosterwitz cDNA at the dilutions (1×10^{-4} to 1×10^{-7}) against "Eur" primer pairs (Eur_1-F/R, Eur_19-F/R and Eur_39-F/R) were identified and in the presence and the absence of the total RNA extracts of TBEV negative tick samples which gave fragments with 361 to 419 nt in lengths compared to the reference ladder. Those PCR bands were absent at the dilution (1×10^{-8}) in both cases and only the bands related to the unspecific amplifications were exhibited in the PCR products. The corresponding

PCR bands relevant to the cDNA from the TBEV positive total RNA of tick sample exhibited in all the PCR products amplified against “Eur” primer pairs Figure 3.4 [(e), (f) & (g)]. The relative intensities of the PCR bands exhibited the same for both PCR conditions carried out in the presence and absence of cDNA derived from the total RNA extract of TBEV negative tick sample. This indicates no inhibition of the PCR product from the tick extract RNA. The corresponding wells of the NTC samples did not contain any PCR bands neither the presence or in the absence of total RNA from tick samples. This shows there were no cross contaminations during the PCR.

The remaining amount of PCR products were separately subjected to PCR product purification and finally the dsDNA was quantified using the Qubit DNA analyzer. The addition of the total RNA from the TBEV negative and TBEV positive tick samples to the cDNA synthesis of the serially diluted RNA from TBEV-Hochosterwitz and the subsequent PCR amplification have increased the dsDNA quantities. Higher amount of dsDNA was observed in the PCR products derived through the JK_1-F/R and Eur_1-F/R when compared to the other primer pairs (Table A.9, Annexure I). Concentrations of purified dsDNA in the TBEV negative tick samples were lower than the concentration of dsDNA in the PCR products derived from the TBEV positive total RNA extracts of the tick sample (Table A.9, Annexure I). The overall effect by the total RNA from the TBEV positive and negative tick samples on the RT-PCR and PCR amplifications were positive.



Coloured arrows: Indicate the positioning of PCR amplified products of "JK" & "Eur" primer pairs with the amplified lengths: (Red: JK_1-F/R, 1961 b), (Green: JK_3-F/R, 1833 b), (Blue: JK_7-F/R, 1851 b), (Brown: Eur_1-F/R, 361 b), (Purple: Eur_19-F/R, 419 b), (Orange: Eur_39-F/R, 365 b).

Figure 3.4. Electroporated (10 min, 45 V) visualized (E-Gel™ safe imager, λ = 495 nm) agarose gel [(a-d): 1% & (e-g): 2%, with SYBR® Gold II DNA stain] images of the PCR products of TBEV-Hochosterwitz cDNA, amplified against six different TBEV specific primer pairs at 65 ˚C. M (a-d): E-Gel 1Kb Plus DNA Ladder, Invitrogen™, Cat. No. 10488-090 (4.00 µL), M (e-g): Low DNA Mass Ladder, Invitrogen™, Cat. No. 10068013 (4.00 µL). Tick$^{(A)}$: TBEV negative, Tick$^{(B)}$: TBEV positive (Table A.9, Annexure). cDNA from the TBEV-Hochosterwitz strain (TBEV-Hoc cDNA) .
{Gel well number, PCR primer pair, template cDNA composition:
*[(a:6, b:9, d:10, e:6, f:9): NTCs]; [(a:1) - (a:5): JK_1-F/R, TBEV-Hoc cDNA (10^{-4} – 10^{-8})], [(a:7) – (b:1): JK_1-F/R (10^{-4} – 10^{-8}) + Tick$^{(A)}$RNA], [(b:2): JK_1-F/R, **Tick$^{(A)}$**RNA], [(b:3): JK_1-F/R, **Tick$^{(B)}$**RNA];*
*[(b:4) - (b:8): JK_3-F/R, TBEV-Hoc cDNA (10^{-4} – 10^{-8})], [(b:10) – (c:4): JK_3-F/R (10^{-4} – 10^{-8}) + Tick$^{(A)}$RNA], [(c:5): JK_3-F/R, **Tick$^{(A)}$**RNA], [(c:6): JK_3-F/R, **Tick$^{(B)}$**RNA];*
*[(c:7) - (d:1): JK_7-F/R, TBEV-Hoc cDNA (10^{-4} – 10^{-8})], [(d:3) – (d:7): JK_7-F/R (10^{-4} – 10^{-8}) + Tick$^{(A)}$RNA], [(d:8): JK_7-F/R, **Tick$^{(A)}$**RNA], [(d:9): JK_7-F/R, **Tick$^{(B)}$**RNA];*
*[(e:1) - (e:5): Eur_1-F/R, TBEV-Hoc cDNA (10^{-4} – 10^{-8})], [(e:7) – (f:1): Eur_1-F/R (10^{-4} – 10^{-8}) + Tick$^{(A)}$RNA], [(f:2): Eur_1-F/R, **Tick$^{(A)}$**RNA], [(f:3): Eur_1-F/R, **Tick$^{(B)}$**RNA];*
*[(f:4) - (f:8): Eur_19-F/R, TBEV-Hoc cDNA (10^{-4} – 10^{-8})], [(f:10) – (g:4): Eur_19-F/R (10^{-4} – 10^{-8}) + Tick$^{(A)}$RNA], [(g:5): Eur_1-F/R, **Tick$^{(A)}$**RNA], [(g:6): Eur_1-F/R, **Tick$^{(B)}$**RNA];*
[(g:7) - (g:10): Eur_39-F/R, TBEV-Hoc cDNA (10^{-4} – 10^{-7})]} (Table A.9, Annexure).

3.7 Quality of the reconstructed TBEV-Hochosterwitz genomes by the initial primer schemes

Oxford nanopore sequencing was carried out using rapid barcoding sequencing kit for purified PCR products obtained from the PCR amplification of TBEV-Hochosterwitz cDNA from the reverse transcribed products of tenfold dilutions (1×10^{-4}, 1×10^{-6} and 1×10^{-7}) against four different primer schemes [primers from the already established methods, (M), (N), (A) & (B), Table A.2, Annexure I]. The resulting data were mapped against the TBEV-Hochosterwitz strain (NCBI Acc. No. MT311861) and the generated SAM files were analyzed for the determination of the quality of the reconstructed genomes. The overview of the resulting data is tabulated (Sequencing experiment number 1, Table A.4, Annexure I).

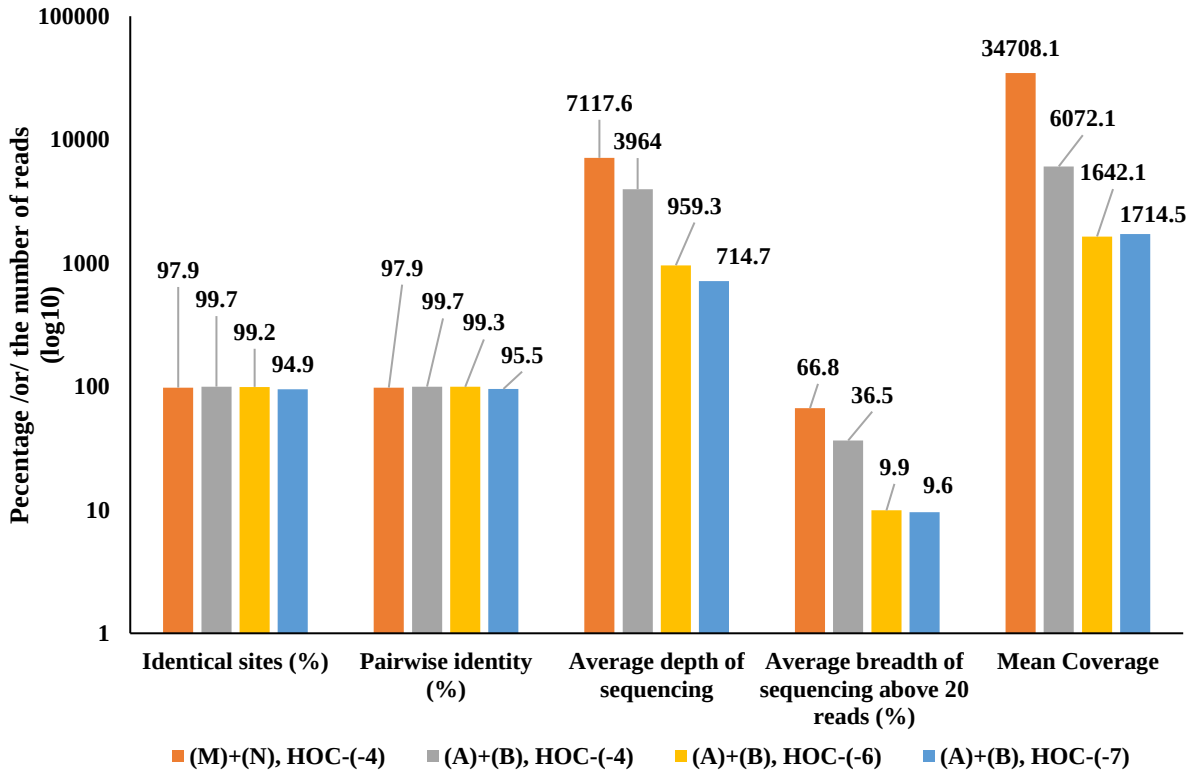


Figure 3.5. A comparison between the quality of the reconstructed TBEV-Hochosterwitz genomes through the Oxford Nanopore sequenced data generated by the amplified PCR products from three tenfold dilutions (1×10^{-4}, 1×10^{-6} and 1×10^{-7}) of reverse transcribed TBEV-Hochosterwitz RNA against four different existing primer schemes [(M), (N), (A) & (B)].

Though the data analysis of each SAM file generated for each reference-based mapping, the average depth, breadth and mean coverage of the reference-based mappings were evaluated. The sequencing data generated through the PCR amplification of reverse transcribed TBEV-Hochosterwitz RNA against combined primer scheme (M)+(N) were compared to the combined primer scheme (A)+(B) at the dilution level (1×10^{-4}). The highest level of reduction in the average depth, breadth and mean coverages of combined primer scheme (A)+(B) were found when the dilution factor increased from 1×10^{-4} to 1×10^{-6} levels and the corresponding differences were minimum between the dilution level 1×10^{-6} and the 1×10^{-7}. The highest percentage of pairwise identity and the highest percentage of identical sites exhibited in the SAM file generated from the reference-based assembly of sequencing data generated by the PCR products from the combined primer schemes (A)+(B) at the dilution level (1×10^{-4}) and the percentage of reduction was minimum between the dilutions at 1×10^{-4} - 1×10^{-6} and was maximum between the dilutions at 1×10^{-6} - 1×10^{-7} (Figure 3.5).

Pairwise sequence alignments were carried out for each of the consensus sequences generated by the reference-based mapping, by the *de-novo* assembly of the combined primer schemes under different dilution levels, identified the pairwise identities and the gaps (Figure 3.6) in each of the reconstructed TBEV-Hochosterwitz sequences. The consensus sequence generated by the reference-based mapping (TBEV-Hochosterwitz, NCBI Acc. No. MT311861) of the Oxford Nanopore sequenced data derived through the combined primer schemes (A)+(B) at the 1×10^{-4} dilution, exhibited one small gap (10630-10670 nt) in the sequence with 41 bases in length and further with the dilutions at 1×10^{-6}, number of gaps in the consensus sequence increased slightly, but at the dilution (1×10^{-7}), longer and higher number of gaps in the consensus sequences were identified. The consensus sequence derived through the combined primer schemes [(M)+(N)] at the 1×10^{-4} dilution level by the referenced-based mapping, exhibited single long stretch of 216 nt gap (10493-10708 nt). The number of gaps and the length of the gaps found to be much higher in the corresponding consensus sequences derived from the *de-novo* assembly followed by reference-based mapping with TBEV-Hochosterwitz (TBEV-Hochosterwitz, NCBI Acc. No. MT311861) (Figure 3.6).

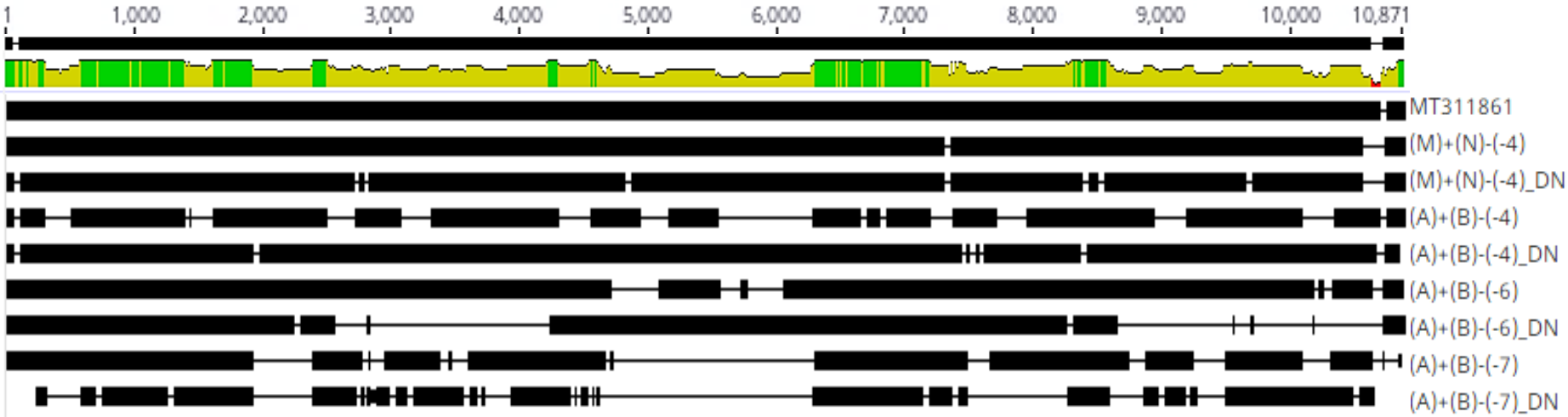


Figure 3.6. The effect on the quality of the reconstructed TBEV genomes based on the sequence coverage with the changes in overlapping primer pair composition and the template TBEV-Hochosterwitz RNA concentration. The multiple sequence alignment [MAFFT Alignment: FFT-NS-i×1000), 1PAM/k = 2, Geneious Prime® (version: 2022.0.1)] of the reconstructed consensus sequences (Threshold = 20 per base depth) derived through the Oxford nanopore sequencing of PCR products from tenfold dilutions (1×10^{-4}, 1×10^{-6} & 1×10^{-7}) of the reverse transcribed TBEV-Hochosterwitz RNA, amplified against two different combined existing primer schemes {[(M)+(N)] & [(A)+(B)]}. PI: Pairwise identity (%), DN: De novo assembly using Trinity tool (version: 2.9.1, Grabherr et al., 2011) followed by reference-based mapping with Minimap2. All the other consensus sequences (Threshold = 20 per-base depth) derived through the reference-based mapping against TBEV-Hochosterwitz strain sequence (MT311861) using Minimap2 plugin in the Geneious Prime® (version: 2022.0.1).

3.8 Phylogenetic tree analysis for the reconstructed TBEV-Hochosterwitz sequences

The consensus sequences received by the Oxford Nanopore sequencing of the already existing combined primer schemes {[(M)+(N)] & [(A)+(B)]} derived from three different tenfold dilutions (1×10^{-4}, 1×10^{-6} & 1×10^{-7}) were subjected to phylogenetic tree analysis to evaluate the accuracy of the reconstructed sequences to the reference TBEV-Hochosterwitz sequence (NCBI. Acc. No. MT311861).

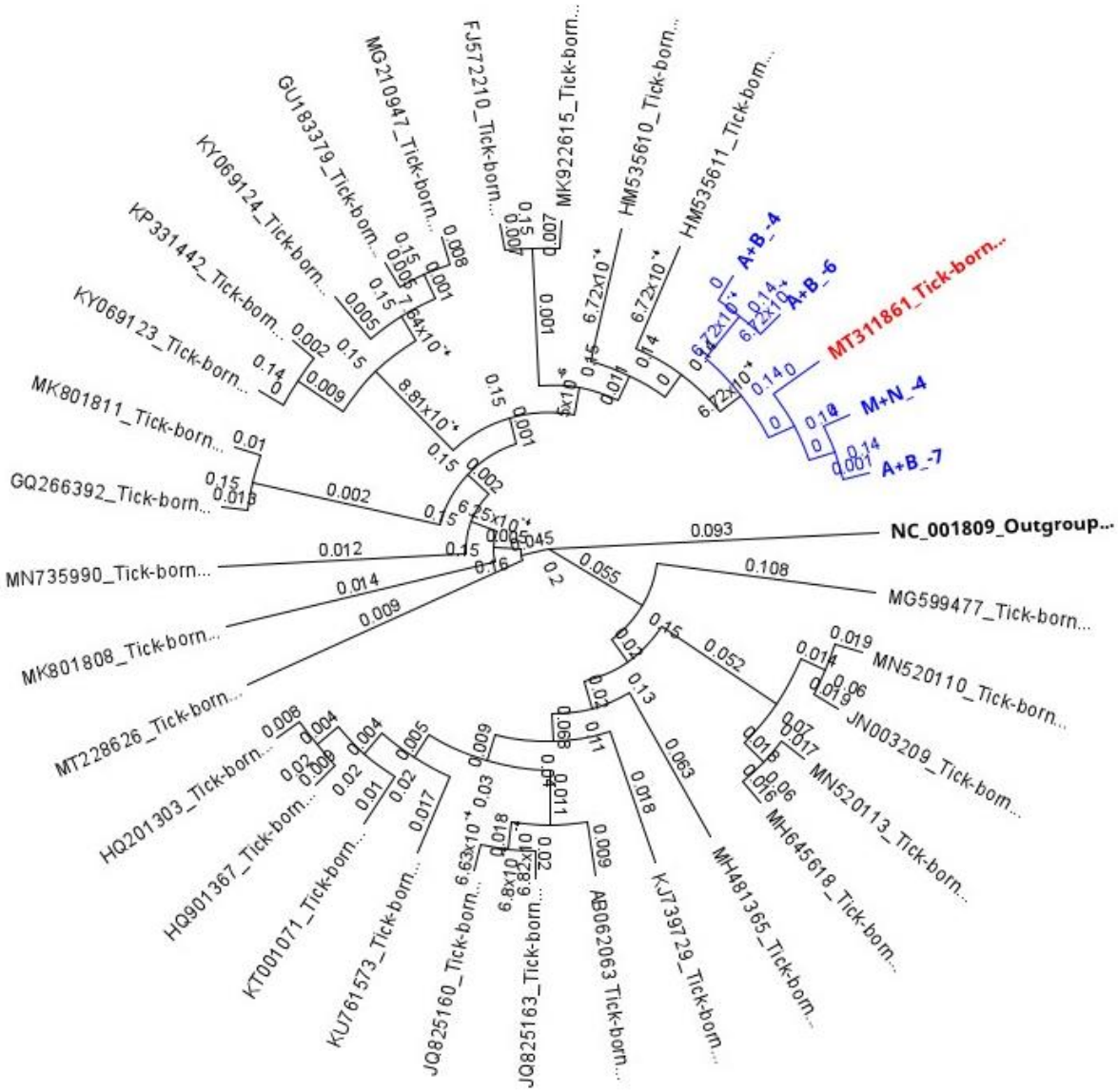


Figure 3.7 A phylogenetic tree constructed using the maximum-likelihood method, depicting the sequence similarity through the substitution per site values of the reconstructed TBEV-Hochosterwitz sequences against two different combined primer schemes {[(M)+(N)] & [(A)+(B)]} with three tenfold dilutions (log10:-4, -6, -7), against 28 different TBEV sequence variants obtained from the NCBI nucleotide database and an outgroup sequence of Louping ill virus (NC_001809).

All the reconstructed sequences from two different combined primer pools at three different concentrations were found closely related to the reference sequence of TBEV-Hochosterwitz strain (MT311861) in the constructed phylogenetic tree, compared to the other 28 variants of TBEV strains and to the outgroup, Louping ill virus strain (NC_001809) (Figure 3.7). Among the reconstructed sequences, the sequence derived through the Oxford Nanopore sequencing of the PCR products from the combined primer scheme (M)+(N) at the dilution (1×10^{-4}), exhibited nearest to the reference sequence of TBEV-Hochosterwitz strain (MT311861) with zero substitution per site. The highest deviation exhibited with the reconstructed sequence derived from the combined primer scheme [(A)+(B)] at the dilution level (1×10^{-6}) with the substitution per site (1.344×10^{-3}), but the sequence derived from the higher dilution (1×10^{-7}) exhibited only lower substitution per site (1×10^{-3}) to the reference sequence (MT311861). The outgroup considered here exhibited the highest deviation to all the reconstructed sequences giving highest substitution values compared to all the other sequences of TBEV variants (Figure 3.7). Reconstructed TBEV genome sequences and the deviations to the reference sequence (MT311861) found to be due to the error associated with the sequencing and due to the poor sequencing depth. This is very clear, the increased substitution per site values and the tenfold dilution factor associated with the TBEV-Hochosterwitz RNA used for the reconstruction of the TBEV-Hochosterwitz genomes.

3.9 Poorly performed primers

Poorly performed primers were identified based on the combination of calculated primer pair efficiency (Threshold < 1.00), sequence identities (< 95%), the sum of sequencing depth of each base (Threshold < 50 reads) and the sequence positional accuracy values (± 50 nt). The primer pair efficiencies, primer identities and positional accuracies for each primer pair are tabulated in Table 3.2. A total of 20 overlapping primer pairs were identified as poorly performed primers according to the calculated primer pair efficiency values and the primer identities to the reference sequence of TBEV-Hochosterwitz. The primer pairs (Eur_1-F/R, Eur_29-F/R, Eur_38-F/R & Eur_39-F/R) exhibited inaccurate primer positioning in the reference genome (MT311861). Out of 39 overlapping primer pairs, 12 overlapping primer pairs exhibited lower primer pair efficiency compared to the threshold level (<1.00) (Purple text, Table 3.2), six of them exhibited inaccurate positional value (Blue coloured fill, Table

3.2) and exhibited the average sequencing depth values bellow 50 reads (Green text, Table 3.2). Primer pairs which exhibited lower primer pair efficiency also exhibited lower primer

Table 3.2. Poorly performing primers identified through the calculated primer efficiency, primer pair identity (%), average per-based depths and the positional accuracy values.

Primer pair	Dilution factor (-log10)			Linear regression (Graph intercept)	PI (%) to the reference sequence#	
	4	6	7			
	Number of reads derived through each primer pair			Primer pair efficiency	Forward primer	Reverse primer
Eur_1_L/R	255	4	0	0.60		100
Eur_2_L/R	9540	221	4	60.06	95	100
Eur_3_L/R	2104	5	1	-8.58	100	100
Eur_4_L/R	2770	0	0	-15.26	90	100
Eur_5_L/R	1104	5	1	-3.08	95	100
Eur_6_L/R	280	6	2	2.47	95	100
Eur_7_L/R	1994	911	0	444.93	95	100
Eur_8_L/R	2471	43	1	8.41	95	100
Eur_9_L/R	981	741	0	365.43	95	100
Eur_10_L/R	5227	533	297	387.95	90	100
Eur_11_L/R	1486	117	0	50.37	100	100
Eur_12_L/R	3760	1103	841	956.04	100	100
Eur_13_L/R	2289	902	0	438.80	100	95
Eur_14_L/R	3938	376	252	293.75	86	95
Eur_15_L/R	2063	1	907	447.22	95	100
Eur_16_L/R	6878	2702	3619	3142.13	86	100
Eur_17_L/R	476	4	3	0.89	95	95
Eur_18_L/R	670	1	0	-3.19	95	100
Eur_19_L/R	1354	127	0	56.10	100	100
Eur_20_L/R	45	0	0	-0.25	81	95
Eur_21_L/R	30	2	0	0.84	90	100
Eur_22_L/R	1709	17	0	-0.91	91	100
Eur_23_L/R	1295	82	22	45.01	100	95
Eur_24_L/R	3575	159	31	75.53	95	100
Eur_25_L/R	2278	2514	1384	1944.59	95	95
Eur_26_L/R	7284	2305	2325	2287.67	95	100
Eur_27_L/R	1523	1286	2	636.20	100	100
Eur_28_L/R	5244	388	311	322.36	100	100
Eur_29_L/R	3750	2875	2191	2524.72		100
Eur_30_L/R	8372	3042	2841	2911.12	100	100
Eur_31_L/R	674	6	4	1.31	90	100
Eur_32_L/R	3555	219	119	150.11	95	100
Eur_33_L/R	1028	49	2	19.87	95	95
Eur_34_L/R	2974	72	100	70.15	95	100
Eur_35_L/R	487	44	28	33.48	100	100
Eur_36_L/R	932	3	6	-0.60	100	100
Eur_37_L/R	1561	360	923	637.73	95	95
Eur_38_L/R	202	0	0	-1.11	95	
Eur_39_L/R	561	13	0	3.42		

Red coloured text: Considered as poorly performing primers, purple coloured text: Selection based on the primer efficiency (Threshold < 1.00), Green coloured text: Selection based on both primer identities (Threshold < 95%) and the sequencing depth of each base (Figure A.3, Annexure I), Blue coloured fill: Selection based on the inaccurate primer positioning. #Reference sequence: TBEV-Hochosterwitz sequence (MT311861). Details of each primer are given in the Table A.2, Annexure I.

identity to the reference sequence or with the inaccurate primer positioning in the reference sequence. All the considered poorly performing primer pairs contained a lower number of reads mainly in the higher tenfold dilution levels (1×10^{-5} - 1×10^{-7}).

3.10 New overlapping PCR primer pairs

To replace the 20 different poorly performing overlapping primer pairs, 41 different new overlapping primers were designed. The identified poorly performing primers, new primers as substitutions and their physical properties are listed in the Table 3.3. The details of the expected PCR product lengths of each overlapping primer pairs are given in the Table A.12, Annexure I. Primer lengths of the newly designed primers kept nearly closer to 20 nt and some of the primers extended up to 25 nt long to increase the melting temperature nearly to 65 °C as required for the PCR reaction. One degenerate primer (CJB_4-F) was designed due to the poor conservedness of the region (1070-1091 nt) in the TBEV genome sequences of the variants analysed. Self-dimerization ability and the heterodimerisation ability were measured using IDT online OligoAnalyser and data were tabulated in the Table 3.3 and found less possibility to form any self-dimerized or heterodimerized PCR products at the reaction temperature through the Gibbs energy values associated with the dimerization. Primer sequence identity to the reference sequences (Consensus sequence of European subtype of TBEV sequences, considered 57 variants) were maintained above 95% for each of the primers and maintained the 100% identity minimum over 87 out of 250 TBEV sequence variants in the collection (Table A.5, Annexure I).

Table 3.3. Physical properties of primers designed to replace poorly performing primers in this study

Poorly performing primers	Primers designed to replace the poorly performing primers									
	Primer Name	Sequence	Length (b)	Tm* (°C)	GC (%)	Self-dimerization energy (kcal/mol)	Heterodimerization energy (kcal/mol)	Primer position#		PI##
								Start	End	
Eur_1-F	CJB_1-F	CAGCTTAGGAGAACAAGAGCTG	22	65.0	50.0	-39.26	-39.26	109	131	100
Eur_1-R	CJB_1-R	TCATCACACCATGACCCCAT	20	66.0	50.0	-37.17		596	616	100
Eur_3-F	CJB_3-F	CCTATGAGTGTGTGACCATAGA	22	63.0	45.5	-35.67	-38.19	624	646	100
Eur_3-R	CJB_3-R	GTCAAGCCACACATCCATTGA	21	65.0	47.6	-38.19		1094	1115	95
Eur_4-F	CJB_4-F	ACYATAACAGCTGAGGGGAAG	21	63-67	47.6	-38.60	-44.47	1070	1091	100
Eur_4-R	CJB_4-R	GTGGCTTTCTTTTTTGCCTCACA	23	66.0	43.5	-44.47		1337	1360	100
Eur_5-F	CJB_5-F	CAGAGTGATCGAGGCTGGGG	20	70.0	65.0	-40.40	-40.4	1262	1282	100
Eur_5-R	CJB_5-R	GTGAAGGATGCCGTCTTCCT	20	67.0	55.0	-39.01		1454	1474	100
Eur_6-F	CJB_6-F	GCAAAAAAGAAAGCCACAGGAC	22	65.0	45.5	-42.52	-42.52	1343	1365	100
Eur_15-R	CJB_15-F	CATGGCTCTCATGACACAGC	20	66.0	55.0	-36.22	-42.74	3993	4013	100
Eur_17-F	CJB_17-F	GTGATGGGACTGTGGACGCT	20	70.0	60.0	-38.54	-38.54	4558	4578	100
Eur_17-R	CJB_17-R	TGAGCAATGCTGCTGACGTA	20	67.0	50.0	-37.81		5092	5112	100
Eur_18-F	CJB_18-F	CAGGTCCATGCCTTCCCACC	20	71.0	65.0	-41.82	-44.05	4897	4917	100
Eur_20-F	CJB_20-F	TGGGAGGTGGCAATCATGGA	20	69.0	55.0	-40.81	-40.81	5456	5476	100
Eur_20-R	CJB_20-R	CATCACACTGTCCAGAGTATATGT	24	63.0	41.7	-37.37		6027	6051	95
Eur_21-F	CJB_21-F	AGAGTGAGGGATGAGAAGCC	20	66.0	55.0	-36.88	-42.39	5810	5830	100
Eur_21-R	CJB_21-R	TGCTTTCTTTTCTCTTCAGTGAGTC	25	65.0	40.0	-42.39		6183	6208	100
Eur_22-F	CJB_22-F	ACCAGAACAGGACAAGATGC	20	65.0	50.0	-36.00	-39.90	6142	6162	100
Eur_22-R	CJB_22-R	CGAACTCTTTGATGTCACGTCC	22	66.0	50.0	-39.90		6425	6447	100

Table 3.3. Physical properties of primers designed to replace poorly performing primers in this study cont..

Poorly performing primer name	Primers designed to replace the poorly performing primers									
	Primer Name	Sequence	Length (b)	Tm* (°C)	GC (%)	Self-dimerization energy (kcal/mol)	Heterodimerization energy (kcal/mol)	Primer position#		PI##
								Start	End	
Eur_34-F	CJB_34-F	GGAGGCTGCATCATGGATGT	20	68.0	55.0	-38.94	-38.94	9435	9455	100
Eur_34-R	CJB_34-R	AAGTAGCTCAGCAGCCACAT	20	66.0	50.0	-36.93		9957	9977	100
Eur_36-F	CJB_36-F	TGCCGAGACCAAGATGAACT	20	66.0	50.0	-37.89	-38.96	9858	9878	100
Eur_36-R	CJB_36-R	AGCTCCCAGTGCAGGTCATG	20	70.0	60.0	-38.96		10342	10362	100
Eur_37-R	CJB_37-R	ATGATGCAGGGGGTCTCACA	20	69.0	55.0	-38.74	-42.85	10992	11012	100
Eur_38-F	CJB_38-F	TCCCGACGTAGGGCACTCTG	20	72.0	65.0	-41.05	-41.29	10963	10983	100
Eur_38-R	CJB_38-R	AGGAGGAGGAAAAATCCTGAAG	22	64.0	45.5	-41.29		11096	11118	100
Eur_39-F	CJB_39-F	TGCATCATGATAAGGCCGAAC	21	65.0	47.6	-40.78	-40.78	11004	11025	100
Eur_39-R	CJB_39-R	CTGTGTCTGGGTGATGGTGG	20	68.0	60.0	-37.19		11168	11188	100
Eur_11-F	CJB_11-F	CTCGAGATGGCCATGTGGAG	20	68	60.0	-39.16	-39.16	2653	2672	100
Eur_11-R	CJB_11-R	GCAGCTCCCATCACTCCTGT	20	70	60.0	-38.58		3007	3026	100
Eur_14-F	CJB_14-F	CTTGTTCGCTCAATGGTGGT	20	66	50.0	-38.39	-38.39	3500	3519	100
Eur_14-R	CJB_14-R	GCTGTGTCATGAGAGCCAT	19	65	53.0	-34.27		3995	4014	100
Eur_17-F	CJB_17-NF	GGTGGAGAGGTTAGCCTGCG	20	71	65.0	-41.21	-41.82	4415	4434	100
Eur_17-R	CJB_17-NR	GGTGGGAAGGCATGGACCTG	20	71	65.0	-41.82		4898	4917	100
Eur_29-F	CJB_29-F	TGCTCAGAAGAGGAGAGACCAA	22	67	60.0	-39.41	-39.41	7802	7823	100
Eur_29-R	CJB_29-R	CAGTTGGAATCTGTGCAGTG	20	64	50.0	-35.27		8258	8277	100
Eur_33-F	CJB_33-F	AACATGATGGGCAAGAGAGA	20	63	45.0	-36.37	-36.55	9034	9053	100
Eur_33-R	CJB_33-R	ACCTTTATGTTGGTGAGGGT	20	63	45.0	-36.55		9511	9530	100
Eur_34-R	CJB_34-NR	TGCTCCCATTCCCCAATGTC	20	68	55.0	-40.80	-40.80	9745	9764	100
Eur_37-R	CJB_37-NR	CACCATGTTCGGCCTTATCATG	22	66	50.0	-42.43	-42.43	11010	11031	

** Calculation of melting temperature (Tm) was done according to the requirements of Q5® High-Fidelity PCR kit (https://tmcalculator.neb.com/#!/main) (Viewed on 14/09/2022); # Reference to the consensus sequence derived from the multiple sequence alignment of 249 different TBEV strains obtained from NCBI nucleotide archive; ## Primer identity reference to the consensus sequence of European subtype of TBEV variants (57 sequences); Blank (" "): Lack of corresponding position in the reference sequence of TBEV Hochosterwitz stain (MT311861).*

3.11 Effect of PCR temperature on new overlapping primer pairs

Primer pairs (CJB) designed to replace the poorly performing primer pairs were separately evaluated for the PCR amplification ability of TBEV-Hochosterwitz cDNA at the dilution (1×10^{-4}) level under two different annealing-primer extension temperatures (63 & 65 °C) in two separate PCRs. Amplified PCR products were subjected to agarose gel electrophoresis followed by visualization under UV radiation using gel documentation system Image, Quanta 300, GE-Healthcare) (Figure 3.8).

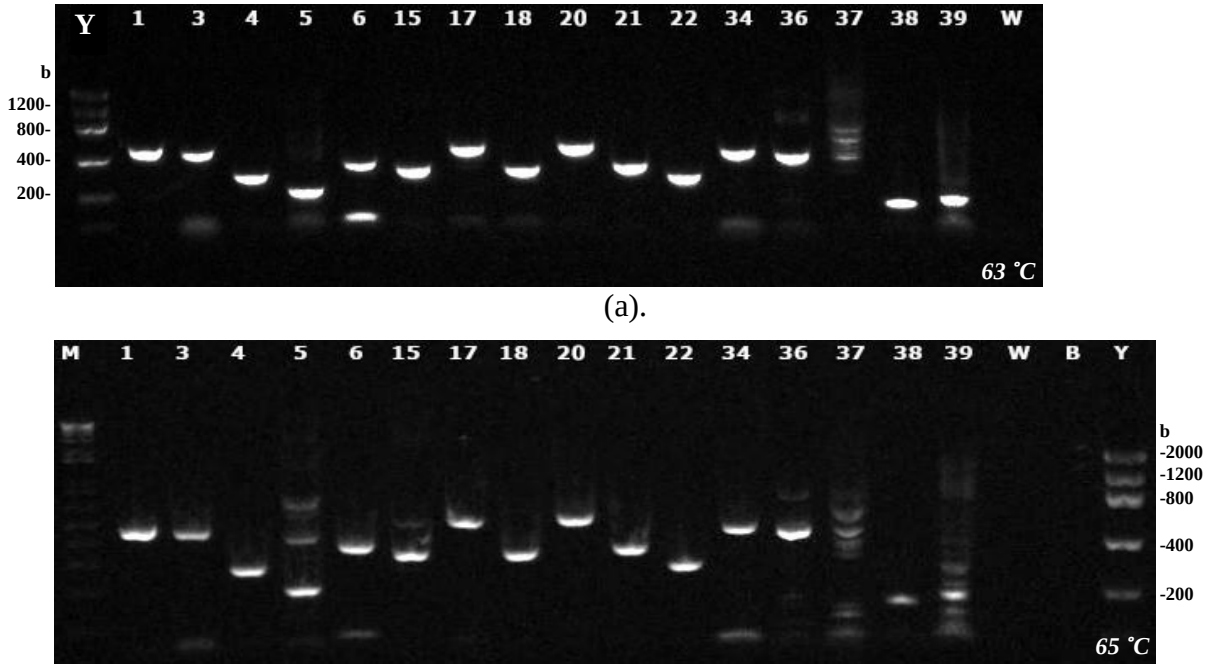


Figure 3.8. Electroporated (10 min, 45 V) and visualized (UV, λ = 280 nm, Image Quanta 300, GE-Healthcare) agarose (2%, Et-Br) gel images of the PCR products of TBEV-Hochosterwitz cDNA amplified against each "CJB" primer pairs (1-39) at two different primer annealing-extension temperatures: (a) at 63 °C and (b) at 65 °C. Y: E-Gel 1Kb Plus DNA Ladder, Invitrogen™, Cat. No. 10488-090 (4.0 µL), M: Low DNA Mass Ladder, Invitrogen™, Cat. No. 10068013 (4.0 µL,), W: NTC, B: Nuclease free water. Gel well numbers represent the corresponding primer pair numbers in the "CJB" primer scheme (Table A.2 Annexure).

Visualized agarose gel images of the PCR products of TBEV-Hochosterwitz cDNA amplified at two different temperatures (63 & 65 °C), found that the new primers (CJB) performed better at 63 °C than at 65 °C, giving comparatively intense bands of amplified products and less unspecific amplifications. The PCR carried out at 65 °C, the resulting products derived through the primer pair (CJB_5-F/R) exhibited two unspecific additional PCR amplifications lading to lower intensities of the expected PCR product with the length of 212 bp, but the PCR

amplification at 63 °C was shown to improve the intensity of the expected product with less unspecific amplifications. Similarly, CJB_38-F/R and CJB_39-F/R primer pairs exhibited proper amplified products with the sequence lengths of 155 nt & 184 b, respectively in the visualized gel carried out at 63 °C leading to intense PCR products even though the PCR primers were designed to amplify the extended region beyond the 10871 nt length of the TBEV-Hochosterwitz strain.

3.12 DNA concentrations of PCR products derived through each overlapping primer pair scheme

The concentrations of each purified PCR product of TEBV-Hochosterwitz DNA derived through the PCR amplification against nine different primer schemes were separately determined. Quantitative determination of dsDNA concentration was carried out for the purified PCR products using Qubit 4 Fluorometer (Invitrogen™, Thermo Fisher Scientific) by following the High sensitivity protocol and the corresponding values were tabulated (Table A.10, Annexure I). Through the tabulated data, a significant steep reduction in the dsDNA concentrations was found between the tenfold dilutions between the (1×10^{-5} and 1×10^{-6}). A significant difference in the dsDNA concentrations in primer scheme (A) and (B) were always exhibited, the concentration of dsDNA in the primer scheme (A) exhibited considerably low compared to the primer scheme (B) and this was consistent with the other primer schemes where the some of the primer pairs associated from primer scheme (A). In the repeated experiments associated with primer schemes (A) & (B), significant reductions in the dsDNA concentrations in the resulted PCR products were identified. This was mainly by the PCR products derived through the primer schemes (A), (B), (C+O) and (D+P). In addition, the increase of primer concentrations directly resulted in the increase of PCR products of corresponding primer schemes and the effect is more significant with the higher tenfold dilutions of the template cDNA (TBEV-Hochosterwitz cDNA). The highest amount of dsDNA concentration detected in the PCR products derived through the primer schemes (2G+2S) and (2H+2T) in all the corresponding tenfold dilutions. All the NTCs carried out by without adding TBEV-Hochosterwitz cDNA exhibited considerable low amount of dsDNA except for the samples of NTCs derived from the primer schemes (2F+2R) and (2H+2T) with the values of 3.12 & 1.23 ng/µL, respectively. This may be attributed for a contamination or the higher amount of primer dimer formations or complex interactions leading to form unspecific amplifications.

3.13 Optimum PCR conditions for the whole genome sequencing of the low TBE viraemic samples

3.13.1 Calculated primer pair efficiency in each primer pair scheme

Through the SAM files obtained through the reference-based assembly of Oxford Nanopore sequenced data derived through each PCR amplifications under three different tenfold dilutions against nine different overlapping primer schemes, best PCR conditions and best primer schemes for the PCR amplification of low viraemic samples were identified. Through the SAM files, each primer pair efficiency was determined for each primer scheme (Table A.13, Annexure I) and used to identify the poorly performing primers with the combination of different primer schemes. Through the calculated primer pair efficiency values of the sequencing experiment number 2, exhibited a higher number of better performing primer pairs compared to the corresponding data generated by the 3rd sequencing experiment. Very few numbers of primer pairs were performed poorly, mainly the primer pairs (Eur_2-F/R, Eur_24-F/R, Eur_25-F/R, CJB_1-F/R & CJB_34-F/R) in the combined primer scheme [(C+O)+(D+P)] and also seven primer pairs (Eur_21-F/R, Eur_34-F/R, Eur_36-F/R, Eur_39-F/R, CJB_3-F/R, CJB_21-F/R & CJB_34-F/R) in the combined primer scheme [(2A+O)+(2B+P)]. But the combined primer schemes of [(A+O)+(B+P)] and [(A)+(B)] did not contain any poorly performing primer pairs as the primer efficiency values generated by the calculation (Table A.13, Annexure I).

In the 3rd sequencing experiment, considerable number of primer pair contained the calculated primer efficiencies below the threshold level in each combined primer pools. During the repeated sequencing attempt in the 3rd sequencing experiment of the combine primer scheme [(C+O)+(D+P)], exhibited a higher number of poorly performed primer pairs than in the first attempt in the 2nd sequencing experiment (Table A.13, Annexure I). The least number of poorly performed primer pair were identified in the combined primer scheme [(2E+2Q)+(2F+2R)] with three primer pairs (Eur_12-F/R, Eur_16-F/R & CJB_18-F/R) in the 3rd sequencing experiment. The number of poorly performing primer pairs were significantly reduced with the increase in the concentration of the respective primer schemes (Table A.13, Annexure I).

3.13.2 Per-base sequencing depth of reconstructed TBEV-Hochosterwitz genomes by each primer pair scheme

Though the BAM files generated by each reference-based mapping against the TBEV-Hochosterwitz nucleotide sequence (NCBI Acc. No. MT311861), per-base sequencing depth values were plotted against the respective nucleotide positioning with considering each dilution level and the concentration of each primer pair scheme (Figure A.3 to Figure A.7, Annexure I). Through the plotted graphs, effects of dilution factor, the effect of concentration of each primer scheme and the regions with the minimum sequencing depths on the PCR amplification of TBEV-Hochosterwitz cDNA were identified. A significantly high reduction in per-base sequencing depths between the dilution factor 1×10^{-5} and 1×10^{-6}, identified in the graphs derived through the primer schemes of [(A)+(B)] and [(2A+O)+(2B+P)], but the difference between dilution factors 1×10^{-4} and 1×10^{-5} and also between the 1×10^{-6} and 1×10^{-7}, identified to be less (Sequencing experiment No. 1 & 2, Figure A.3-A4, Annexure I). The effect of increase in the primer scheme concentration exhibited positive effects on the per-base sequencing depths values in the graphs derived through the primer schemes [(2A+O)+(2B+P)], [(2G+2S)+(2H+2T)] & [(2E+2Q)+(2F+2R)] compared to the corresponding primer schemes with lower concentrations [(A+O)+(B+P)], [(G+S)+(H+T)] & [(2E+Q)+(2F+R)] (Figure A.4, Figure A.5 & Figure A.6, Annexure I), respectively. Through the corresponding graphs, regions where the per base sequencing depth fell below 20 were identified and found to be mainly associated with the dilution factor 1×10^{-4}, 1×10^{-5}, 1×10^{-6} and 1×10^{-7} in [(M)+(N)] and [(A)+(B)] primer schemes, but for the rest of primer pair schemes, per base sequencing depths fell below 20, found to be associated mainly with the tenfold dilutions at 1×10^{-5} and 1×10^{-6}. Several regions identified with the per-base sequencing depths fell below 20, mainly in the nucleotide positions (1379-1607, 1931-2381, 2823-2920, 3387-3541, 4612-6501, 7450-7587, 83003-1023 and 10496-10667 b) due to the poor performance of the primer pairs (Eur_6-F/R, Eur_8-F/R, Eur_10-F/R, Eur_12-F/R, Eur_17-F/R to Eur_22-F/R, Eur_27-F/R, Eur_31-F/R to Eur_34-F/R and Eur_36-F/R to Eur_39) at the tenfold dilutions of 1×10^{-6} and 1×10^{-7} of the combined primer scheme [(A)+(B)]. The combine primer scheme [(M)+(N)], had single region of per-base sequencing depth fell below 20 in (10496-10670 b) region of the TBEV-Hochosterwitz nucleotide sequence.

In addition, by the data generated through the repeated sequencing experiments of the same primer schemes against the TBEV-Hochosterwitz RNA after a certain time period with the degradation due to the freeze and thaw cycles, changes in the per-base sequencing depth of

the same primer schemes were also compared. This was carried out to understand the possible effects on the comparisons between the primer schemes which were analyzed during the latter stages of the study. In addition, through the changes in per-base depth values between the similar primer schemes carried out at two different sequencing experiments were used to evaluate the effect of random hexamers used in the reverse transcription of the TBEV-Hochosterwitz RNA on the PCR amplification. Through the repeated study with the same primer schemes, found out the per-base sequencing depths significantly fell below the previous per-base sequencing depths received for the same primer scheme and this was clearly comparable with the per-base sequencing depth graphs generated by the same primer schemes [(A)+(B)] and the [(C+O)+(D+P)] in both tenfold dilutions at 1×10^{-5} and 1×10^{-6} (Figure A.3, and Figure A.5, Annexure I). In addition, through the repeated experiments, the inconsistent random effects of the random hexamers used in the cDNA synthesis was found to be significant due to the shifting in the peaks and valleys in each graph derived through the same combined primer schemes {[(A)+(B)] and [(C+O)+(D+P)]} from two separate RT-PCR reactions against random hexamers with the same tenfold dilutions (1×10^{-5} and 1×10^{-6}) of TBEV-Hochosterwitz RNA (Figure A.3, and Figure A.5, Annexure I).

3.13.3 Quality of reconstructed TBEV-Hochosterwitz genomes by overlapping primer pair schemes

The number of reads mapped, percentage of identical sites, percentage of pairwise identity, average depth of sequencing, mean coverage and average breadth of sequencing above threshold of 20 reads were calculated (Table A.4, Annexure I) for each of the Oxford Nanopore sequenced data. This was generated by the PCR reactions against each primer scheme at different dilution levels of TBEV-Hochosterwitz cDNA. Through the tabulated data, the highest calculated average percentage of breadth of sequencing above the 20 reads value (66.8%) received for SAM from the PCR products amplified against the combined primer scheme [(M)+(N)] derived through the tenfold dilution (1×10^{-4}) of the TBEV-Hochosterwitz cDNA sample (Figure 3.9).

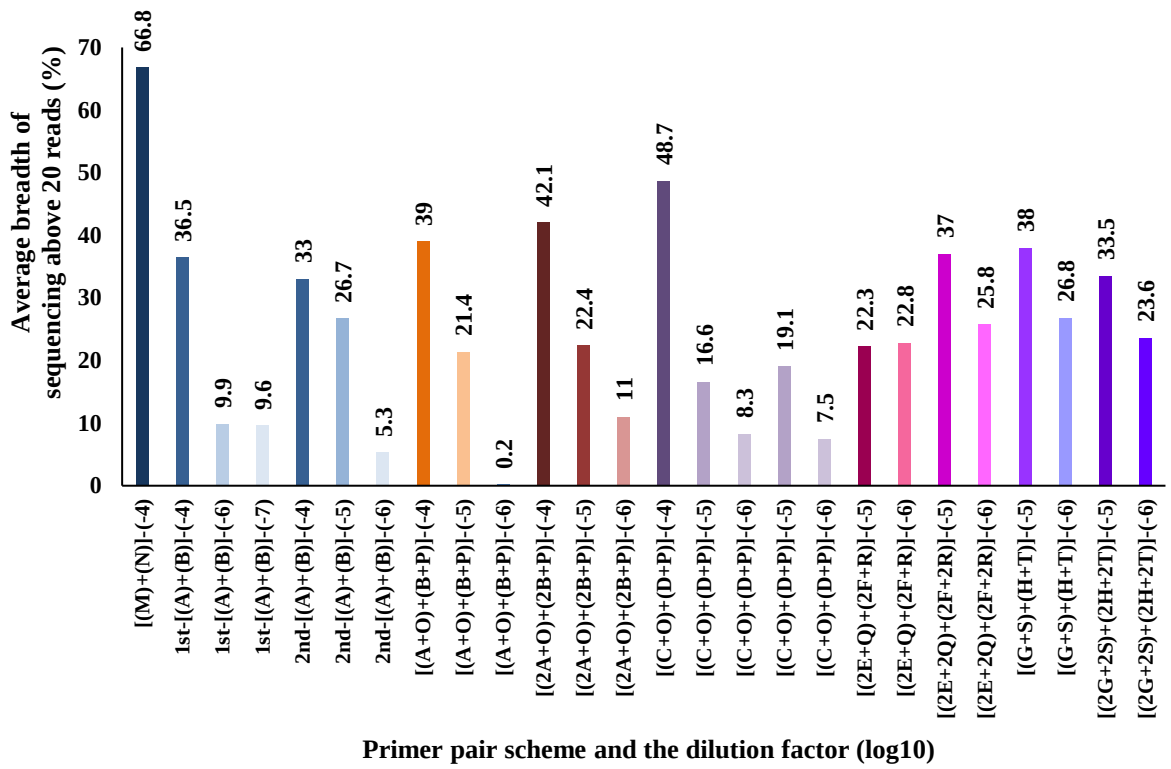


Figure 3.9. Percentage of average breadth of sequencing above 20 reads in each SAMs derived through the reference-based mapping of Oxford Nanopore sequenced data from each PCR product amplified against nine different combine primer pair schemes of TBEV-Hochosterwitz cDNA at four different tenfold dilutions (1×10^{-4}, 1×10^{-5}, 1×10^{-6} and , 1×10^{-7}).

Among the SAMs derived through the (A) and (B) primer schemes and subsequent derivatives of the (A) and (B) primer schemes, the highest calculated average percentage of breadth of sequencing above the 20 reads value at the tenfold dilution level (1×10^{-4}), exhibited in the SAM derived through the primer scheme [(C+O)+(D+P)] with the value of 48.7% (Figure 3.9). For the tenfold dilutions at 1×10^{-5} and 1×10^{-6}, the highest calculated average percentage of breadth of sequencing above the 20 reads values were exhibited in the SAMs derived through the combine primer pair scheme [(G+S)+(H+T)] with the corresponding values of 38.0% and 26.8%, respectively (Figure 3.9).

To understand the relative performance of each new primer pair scheme compared to the already existing primer pair schemes, relative difference in the percentage of average breadth of sequencing above 20 reads were determined (Figure 3.10). The results showed that the highest relative performance exhibited by the combine primer pair scheme of [(G+H)+(S+T)] with the value of 21.5%, but the increase in the concentration of the same combined primer scheme showed a negative effect by reducing the relative percentage by 3.2%. The second highest relative performance was exhibited by the combine primer pair scheme

[(2E+2F)+(2Q+2R)] by having 20.5% relative average breath of sequencing above 20 reads. Which is the same primer scheme of [(2E+2F)+(Q+R)] with the increase of the concentration in the part of the combine primer scheme (Q+R). The least difference in the relative performance was observed with related to the combine primer pair scheme of [(C+D)+(O+P)] giving negative values at higher tenfold dilutions, compared to the lower tenfold dilutions of the template TBEV-Hochosterwitz cDNA.

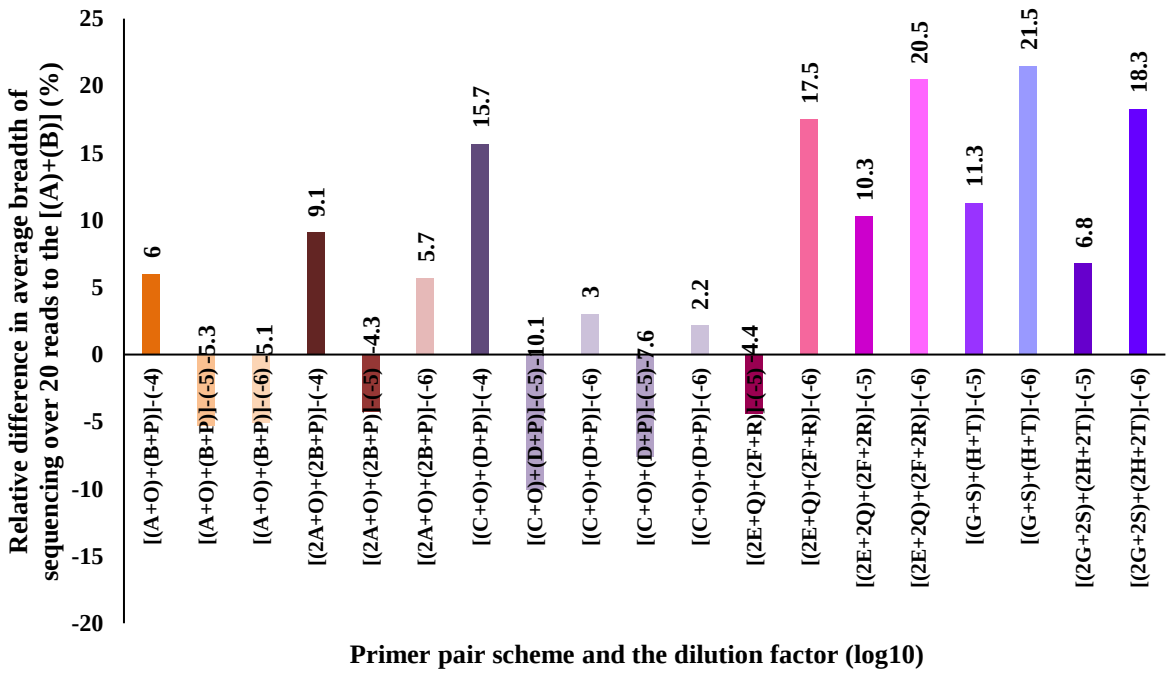


Figure 3.10. Difference in the percentage of average breadth of sequencing above 20 reads compared to the already existing primer schemes in each SAMs derived through the reference-based mapping of Oxford Nanopore sequenced data from each PCR products amplified against eight different combined new primer pair schemes of TBEV-Hochosterwitz cDNA at three different tenfold dilutions (1×10^{-4}, 1×10^{-5} and 1×10^{-6}).

To identify the best performing primer scheme which would be able to amplify the whole genome of the TBEV at low viraemic samples, in addition to the determination of average breadth of sequencing data derived from each of the combined primer pair schemes, percentage of coverage and the pairwise sequence identity to the TBEV-Hochosterwitz reference sequence (MT311861) were determined. The highest percentage of identity (99.5%) exhibited by the consensus sequence derived through the combine primer pair schemes [(A)+(B)] at the dilution (1×10^{-4}) with the percentage of coverage of 97.7%, with the increase in the dilution from 1×10^{-5} to 1×10^{-6}, the percentage of identity (90.4% to 89.0%) and the coverage (99.9% to 83.0%) decreased, respectively. The percentage of reduction in coverage and the identity were significantly higher when the dilution level increased from the 1×10^{-5} to 1×10^{-6}.

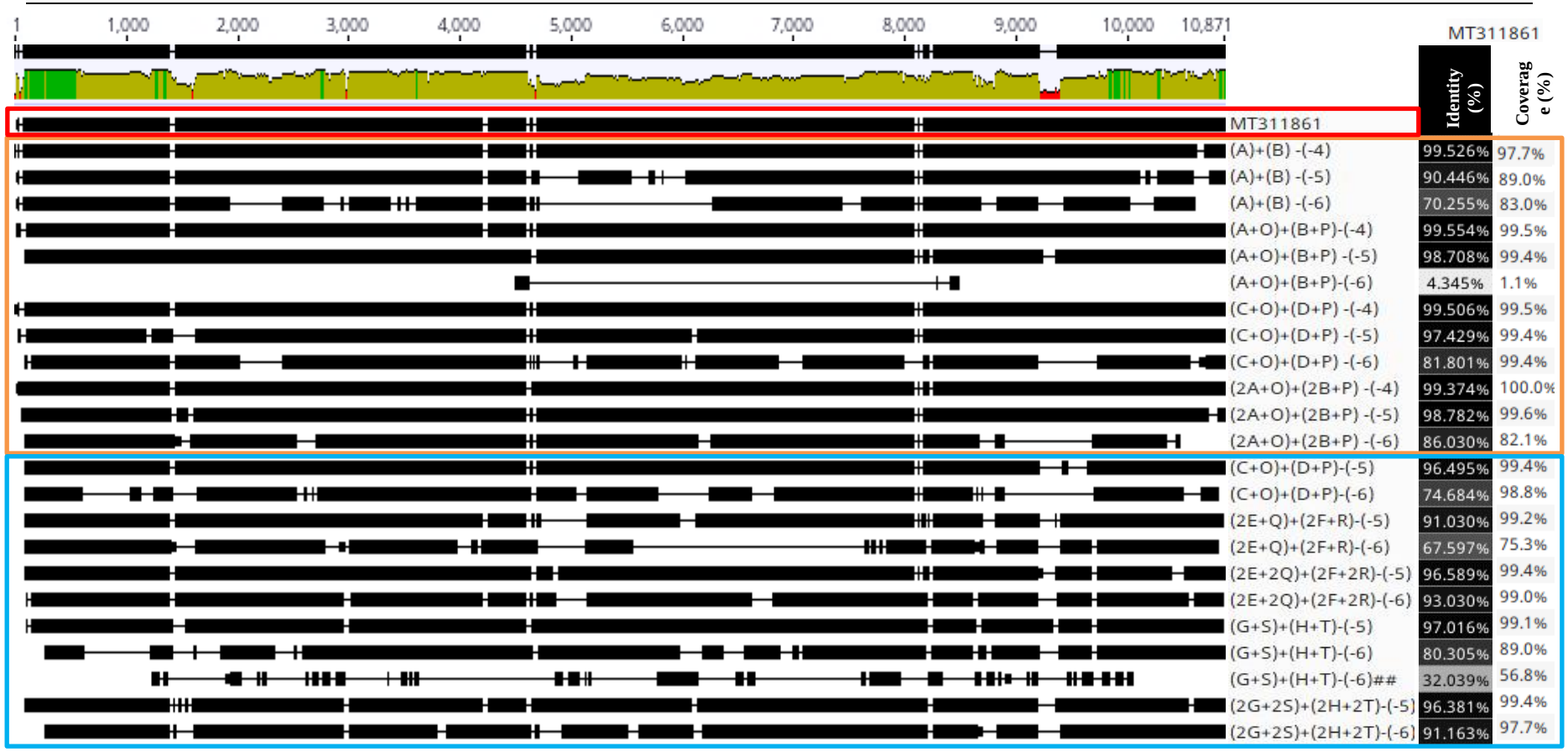


Figure 3.10. Identified gaps in the consensus sequences derived through the Oxford Nanopore sequenced data from the amplified PCR products of the TBEV-Hochosterwitz cDNA against eight different combined primer schemes {[(A+O)+(B+P)],[(C+O)+(D+P)], [(2A+O)+(2B+P)], [(A)+(B)], [(2E+Q)+(2F+R)], [(2E+2Q)+(2F+2R)], [(2G+2S)+(2H+2T)], [(G+S)+(H+T)] } at 63 ˚C with four different dilutions (1×10^{-4}, 1×10^{-5}, 1×10^{-6} & 1×10^{-7}). Multiple sequence alignment [MAFFT Alignment: FFT-NS-i×1000), 1PAM/k=2, Geneious Prime® (version: 2022.0.1)] carried out for each consensus sequence (Threshold = 20 per-base depth) derived through the reference-based mapping [Minimap2 plugin in the Geneious Prime® (version: 2022.0.1)] of Oxford Nanopore sequenced data, against the reference sequence of TBEV-Hochosterwitz strain (MT311861, Indicated in the Red coloured box). ##: Reverse transcription of the TBEV-Hochosterwitz RNA was carried out against combined mixture of all the TBEV secific reverse peirmers. For all the other PCR were carried out using the reverse trancribed TBEV-Hochosterwitz RNA agaisnt rnadom hexamers. Coloured frames repersent two separate attempts of sequencing experments (Orange: sequencing experiment No.2 & Blue: sequencing experiment No.3).

The highest percentages in the coverage and the identity in the dilution level at 1×10^{-5}, exhibited in the consensus sequence derived through the PCR amplification of TBEV-Hochosterwitz cDNA against the combine primer scheme [(2A+O)+(2B+P)] with the value of 99.6% and 98.8%, respectively. Similarly, the highest corresponding values at the dilution level of 1×10^{-6}, exhibited by the combine primer pair scheme [(2E+2Q)+(2F+2R)], with values of 99.0% and 93.0%, respectively. The percentage of identity and the coverage values greatly reduced in the combined primer scheme of [(G+H)+(S+T)] at the dilution level of 1×10^{-6}, where the PCR amplification was carried out with the cDNA derived through the TBEV specific priers instead of the random hexamers, indicating the poor performance during the RT-PCR when employee the TBEV specific primers.

In addition to the quality of the reconstructed TBEV-Hochosterwitz genomes, the amount of unspecify amplifications through the percentage of number of unmapped sequences were evaluated (Figure A.9, Annexure I). Through the tabulated data, with the dilution factor increases, the percentage of unmapped sequences in the constructed DNA library increases nearly exponential levels. The highest percentage of unspecific amplifications found in the combined primer scheme [(A+O)+(B+P)] at 1×10^{-6} dilution giving 68.3% and the lowest level of unspecific amplifications were identified in the combined primer scheme [(A)+(B)] at the same dilution level giving 10.64% . The combined primer schemes [(G+S)+(H+T)] and the [(2G+2S)+(2H+2T)] also exhibited the lowest values similar to the combined primer scheme [(A)+(B)] giving 15.0% and 13.5% respectively at 1×10^{-6} dilution level.

3.14 Identified TBEV positive tick samples by the RT-qPCR methods

Two separate semiquantitative real-time PCR analysis of total RNA extracts of nymph tick pools (76), adult females (46) and adult male tick samples (50) collected from the Larvik region of Norway were carried out for the detection of envelope coding region of the TBEV RNA genome using the Taq-man probe assays method I & method II, (Table A.14, Annxure). Using both techniques total of 26 positive tick samples were identiefied within the 48 cycles of PCRs (Figure 3.11). Corresponding Ct values of TBEV positive samples were tabulated to understand the relative quantities of TBEV viraemic content in each tick sample by the Ct values compared to the TBEV-Hochosteriwtz tenfold dilutions (Figure 3.11). A total of 16 tick samples were identified by the semiquantitative real-time PCR method I with two of 46

adult female tick samples (prevalence 4.3%), six of 50 adult male tick samples (12%) and eight of 76 nymph pools (1.1 ± 0.4%) with the Ct values ranging between 30.72 to 40.75 range (Figure 3.11) (method I). Total of six out of 46 adult female tick samples (13%), four of 50 adult male tick samples (6%) and three of 76 samples of nymph pools (0.4%) were identified as TBEV positive by following the semiquantitative RT-qPCR method II with the Ct values ranging from 6.36 to 40.72. Comparatively higher amount of TBEV vreamic content exhibited in one adult female tick sample [SL1(21)-13F] and in two adult male tick samples [SL1(21)-03M] and [SL1(21)-04M] with the Ct values of 6.36, 6.73 and 6.88, respectively though the semiquantitative real-time PCR carried out by mehtod II, but in the method I, corresponding samples exhibited negative for TBEV. Both semiquantitative real-time PCR analysis methods did not exhibit any false positive results for three NTC parallelly carried out with the other samples and also gave true positive results for all the TBEV-Hochosterwitz dilutions analysed ranging from 1×10^{-1} to 1×10^{-8} with the consistent difference in the Ct values recorded in two methods. Out of 22 TBEV positive samples, only four samples gave Ct values by both methods and the corresponding Ct values exhibited with differences between two methods, rest of the eighteen samples were identified either by one the semiquantitative real-time PCR analysis methods (Figure 3.11).

To understand the is there any significant difference between two methods of analysis, single factor ANOVA test was carried out for the Ct values obtained for each tick sample by both methods. The resulting P value ($P = 0.003$) < 0.05 indicated a significant difference in method I and method II of testing TBEV in tick samples (Table A.14, Annexure I). To understand any significant difference in the number of TBEV infected male ticks, female ticks and pooled nymph tick samples, single factor ANOVA was carried out for the results obtained from the method I with the confidence level of 0.05% and the results indicated ($P = 0.3378$) > 0.05 with no significant difference in the number of TBEV infected male ticks, female ticks and the pooled nymph tick samples. Single factor ANOVA analysis was carried out with the method II, gave the similar results by indicating lack of significant differnece ($P = 0.7016$) > 0.05 in the number of TBEV infected male ticks, female ticks and the pooled nymph tick samples.

Through the Ct values received for the 26 individuals TBEV positive tick samples by semiquantitative real-time PCR techniques, the relative quantities of viraemic content in each was estimated using the Ct values received for the tenfold dilutions of TBEV-Hochosterwitz strain. The Ct values revied for each of the TBEV positive tick samples exhibited more similar to Ct values received for the tenfold dilutions ranging from log10: -6 - -7, except for the

exceptionally lower values where the relative comparisons were unable to compare since the Ct values were out of the range of the Ct values obtained for the tenfold dilutions of TBEV-Hochosterwitz strain.

3.15 Oxford Nanopore sequence data analysis and the reconstructed TBEV genomes from tick samples

Concentrations of the purified PCR products derived from the 22 different tick samples amplified against six different primer schemes were analysed and tabulated (Table A.11, Annexure I) to identify the best possible PCR products for the Oxford Nanopore sequencing. Through the dsDNA concentration values, PCR products derived through the primer scheme [(G+S)+(H+T)] exhibited the highest concentration of dsDNA compared to the other PCR primer schemes used in the analysis and selected for the Oxford Nanopore sequencing. Reference based sequence mapping was carried out for the resulted Oxford Nanopore reads against TBEV-Hochosterwitz sequence (MT311861) and through the SAMs, quality of the reconstructed TBEV genomes were analysed by tabulating the data corresponding to the average depth, breadth of sequencing and number of reads corresponding to each TBEV positive tick sample (Sequencing experiment No. 4 & 5, Table A.4, Annexure I). Through the tabulated data and the consensus sequences generated by each reference-based mapping, TBEV genome sequence fragments containing ten different tick samples were identified and confirmed as the TBEV positive samples and selected for the further analysis.

Ten PCR products amplified against [(2C+O)+(2D+P)] from the confirmed TBEV positive samples through the above analysis were selected and subjected to Oxford Nanopore sequencing and the resulting data were analysed for the quality of the reconstructed TBEV genomes (Sequencing experiment No.6, Table A.4, Annexure I) and found out all the ten tick samples analysed were positive for the TBEV and also the number of reads generated after the reference based mapping was found to be less and the quality of the reconstructed genomes were poor for further analysis. The highest number of mapped TBEV genome sequence fragments exhibited in the nymph tick pool (SL1(21)-17N) with the average sequencing breadth above 20 reads was found to be 2.3%.

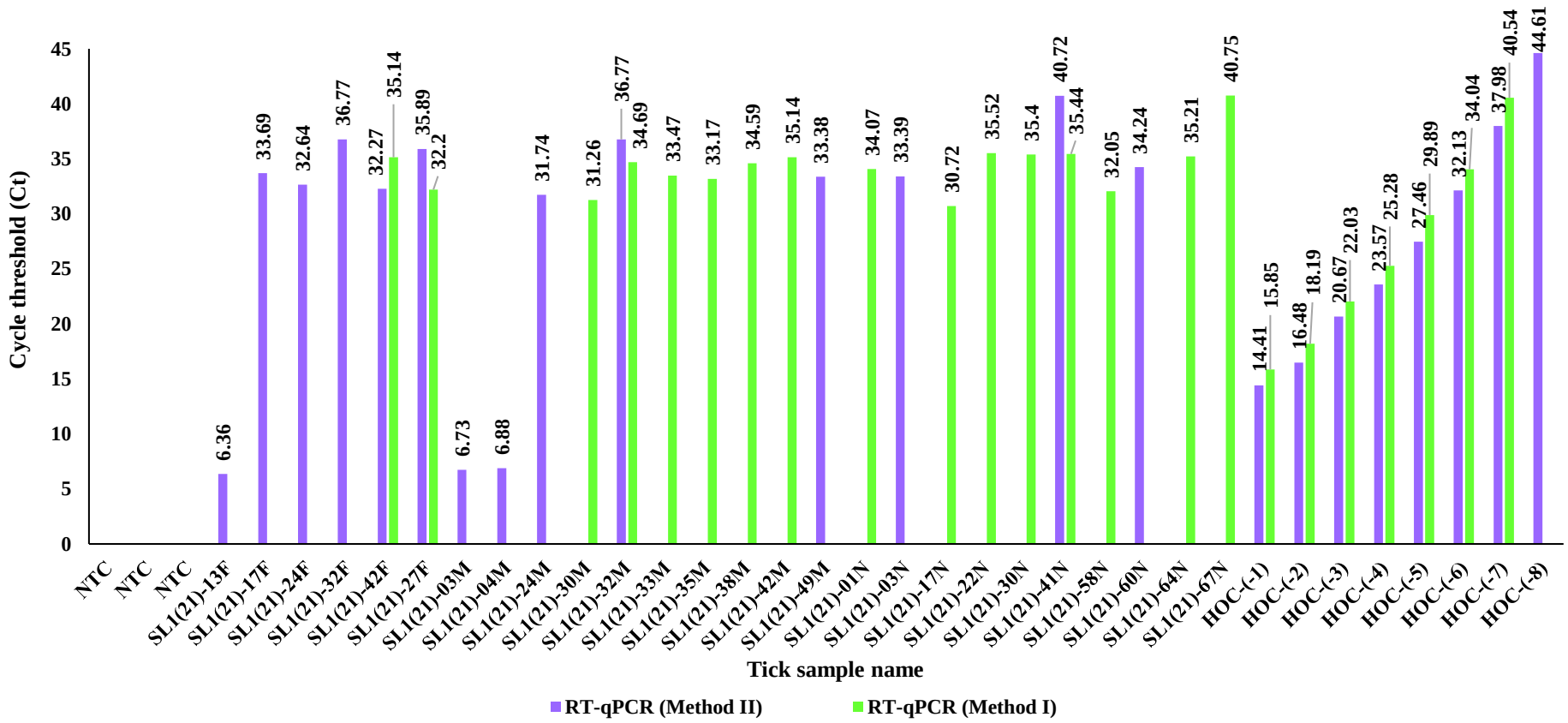


Figure 3.11 Cycle threshold values received for the TBEV positive tick samples and for the dilution series (log 10) of TBEV-Hochosterwitz (Sample No. 07, Table 2.1), through two similar semiquantitative real-time PCR method I & II. NTC: Non templated controls. F: Female adult tick, M: male adult tick, N: Nymph pools (Included ten nymphs of a sample).

Further optimizations were carried out by increasing the volume of the total RNA (8 µL) used in the RT-PCR, increasing the cDNA volume (7 µL) used in the PCR amplifications and also the PCR amplifications were carried out against four different primer schemes to receive two different combined PCR products. Concentration of dsDNA of the purified PCR products were determined and combined to increase the best possible outcome and finally subjected to Oxford Nanopore sequencing followed by reference-based mapping against TBEV-Hochosterwitz sequence (MT311861). The resulting SAMs were evaluated for the determination of the quality of the reconstructed genomes (Sequencing experiment No.7, Table A.4, Annexure I).

After the PCR optimizations, the improvements in each reconstructed TBEV sequences were compared with the data from the previous attempts and found out the increase in the percentage in the average breadth of sequencing above 20 reads in some of the TBEV genomes derived from the total RNA extracts of the tick samples (Figure 3.12). Considerable improvements in the percentage of average breadth of sequencing were exhibited related to seven tick samples, but the rest of the five samples did not exhibit any improvements over 20 reads depths after the optimizations (Figure 3.12).

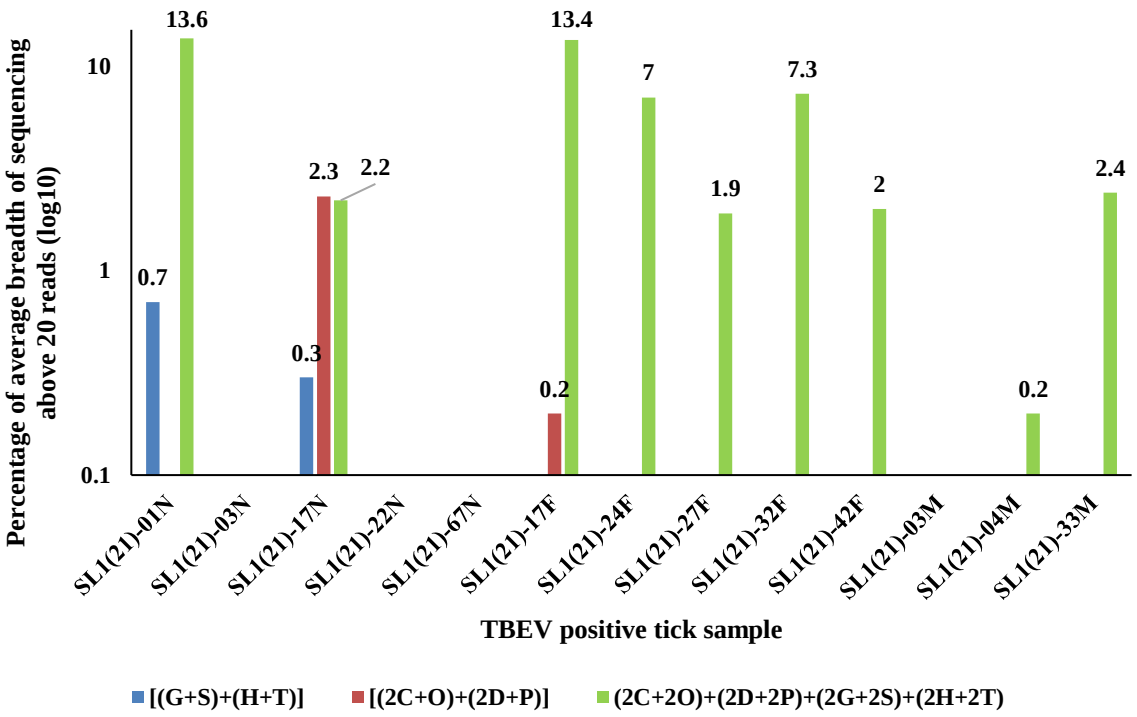


Figure 3.12. Comparison of percentage of average breadth of sequencing above 20 reads in each SAMs of TBEV positive tick samples, after the optimization in the RT-PCR and PCR steps by the increase of volume of the template RNA, template cDNA and the combination of PCR products from four different primer pair schemes [(2C+2O)+(2D+2P) +(2G+2S)+(2H+2T)].

3.16 Identified TBEV viral subtypes

Oxford Nanopore reads derived from each of the whole genome PCR amplifications of the 12 different tick samples with against the combine primer scheme [(2C+2O)+(2D+2P) +(2G+2S)+(2H+2T)], were separately mapped against 250 different TBEV sequences. The resulting SAMs were sorted according to the highest number reads mapped against and selected top ~10 consensus sequences (Over 20 per-base sequencing depth) for each SAMs corresponding to each TBEV positive tick sample. Corresponding ~10 consensus sequences belonging to the same TBEV positive tick sample were subjected to multiple sequence alignment and the consensus sequence was obtained. Finally, each consensus sequence was subjected to multiple sequence alignment against 250 TBEV sequence variants followed by the phylogenetic tree construction for the identification of best possible closely related TBEV subtypes with minimum reference sequence bias.

Through the constructed phylogenetic tree, found out the reconstructed TBEV genome sequence fragments derived from two adult male tick samples SL1(21)-33M, SL1(21)-04M, one adult female tick sample SL1(21)-17F and one nymph tick pool sample SL1(21)-17N were exhibited with the minimum number of substitutions to the European subtype of TBEV Zmeinogorsk-5 strain (NCBI Acc. No. KY069125) isolated from the Alti region of the Western Siberia in Russia (1986) from a tick sample (*Ixodes persulcatus*) (Through the NCBI records, NCBI Acc. No. KY069125). Three other reconstructed TBEV genome sequence fragments derived from the three adult female tick samples SL1(21)-27F, SL1(21)-32F and SL1(21)-24F were exhibited with the minimum number of substitutions to the TBEV-Hochosterwitz strain (NCBI Acc. No. MT311861) collected near the Hochosterwitz region in Austria (1971) from a tick sample (*Ixodes ricinus*) (K. M. Paulsen *et al.*, 2021). Another adult female tick sample SL1(21)-42F exhibited with minimum number of substitutions to the European subtype of TBEV KEM-1 strain (NCBI Acc. No. MW256716) isolated from a tick sample of *Ixodes ricinus* in Hungary (1952) (Egyed et al., 2021b). The remaining one nymph tick pool sample identified to be exhibited with the minimum number of substitutions to the European subtype of the TBEV HB_IF06_8040 strain (NCBI Acc. No. MK922616) from Lower Saxony region in Germany (2018) (Boelke et al., 2019). During the phylogenetic tree construction, the identified TBEV sequences were considered for the evolutionary relationship to the Mandal sequence (KF991107) and found no close evolutionary relationships.

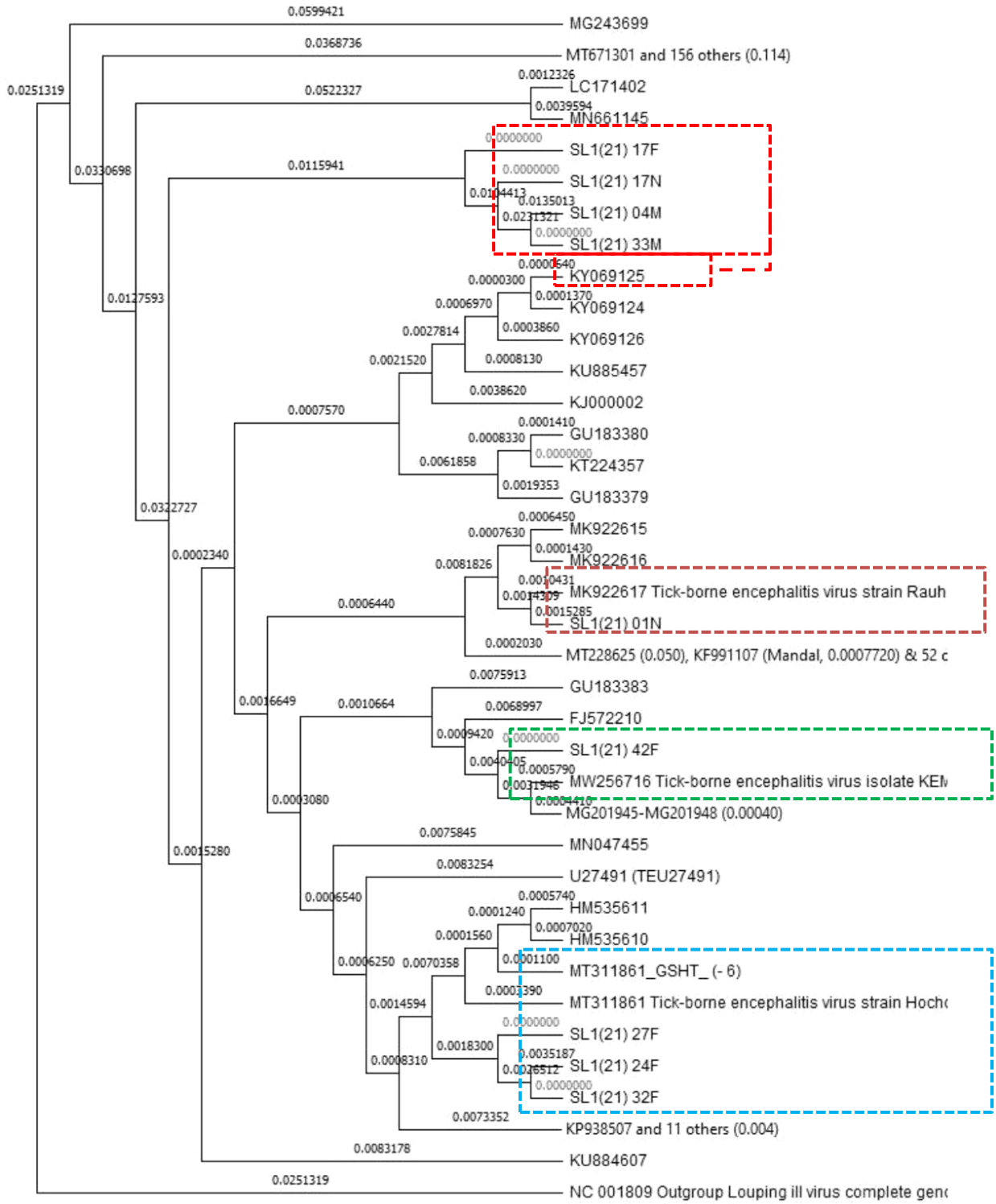


Figure 3.13. Identified closely related TBEV subtypes through the phylogenetic tree analysis of the consensus sequences derived from the reference-based mapping of Oxford Nanopore sequence data by the PCR amplification of nine different TBEV positive tick samples. Phylogenetic tree was constructed using PhyML tool in the Geneious Prime® (version: 2022.0.1) with the following parameters: K80-Kimura, Bootstrap:1000, Substitution rate: 4, Optimized based on substitution rate and all the other parameters by the estimations. Each individual-coloured frames represent the closely related subtypes for each reconstructed TBEV genome sequence fragments from the TBEV positive tick samples with the minimum substitutions.

Accuracy of the method used in the phylogenetic tree construction was evaluated though the poorly reconstructed TBEV-Hochosterwitz strain sequence [MT311861_GSHT_(-6) and its positioning in the phylogenetic tree and fond out, it was positioned very close to the reference TBEV-Hochosterwitz strain (NCBI Acc. No. MT311861), indication the correct MSA and correct phylogenetic tree construction.

Oxford Nanopore sequence data received from each of the TBEV positive tick sample were separately mapped against the phylogenetically identified evolutionary closes TBEV variants. The resulting consensus sequences above the 20 per-base depth value from the SAMs were obtained and multiple sequence alignment was carryout against each of the respective reference sequences identified in the phylogenetic tree analysis (Figure 3.13). The maximum sequence coverage was exhibited in the TBEV sequence derived from an adult female tick sample [SL1(21)-17F] with 50.56% and also exhibited the sequence identity (55.27%). A reconstructed TBEV sequence obtained from the nymph tick pool sample [SL1(21)-17N] contained the sequence coverage (19.24%) and with the sequence identity (20.29%) and the rest of the TBEV sequences exhibited sequence coverage between 12.51 – 0.68% range (Figure 3.14).

Figure 3.14. Multiple sequence alignment exhibiting the regions of sequence gaps in the consensus sequence derived from SAMs with minimum sequencing depth of 20 reads in the reconstructed TBEV partial genomes from two nymph tick pools, five adult female tick samples and two male tick samples against the reference TBEV variants. The number of identical sites to the corresponding reference sequences are also listed.(Complete MSA is available in Figure A3.1 Annexure III).

3.17 Species level of identification of TBEV positive tick species

Oxford Nanopore reads from each of the PCR amplified products of the reverse transcribed total RNA extracts of TBEV positive tick samples, were separately mapped against an 18S-rRNA gene region of the *Ixodes ricinus* (NCBI Acc. No. Z74479.1) (Table A.15, Annexure I). Resulting consensus sequences were subjected to multiple sequence alignment followed by phylogenetic tree analysis with 18S-rRNA gene regions from 22 different tick species. This was carried out to identify the tick species which have been identified as TBEV positive in the RT-qPCR analysis. Phylogenetic tree analysis reviled all the tick samples closest evolutionary relationship to the *Ixodes ricinus* (NCBI Acc. No. GU074656.1) by exhibiting a minimum number of substitutions per site (2.21×10^{-4}) with the 18S-rRNA gene region. The rest of the other 18S-rRNA gene sequences from different tick species exhibited a higher number of substitutions per site against the evaluated sequences derived from each of the total RNA extract of TBEV positive tick samples. The highest number of substitutions per site values exhibited against the outgroup sequence of 18S-rRNA region of *Argas lahorensis* (NCBI Acc. No. AGG18SRA) with compared to all the other sequences (Figure 3.15).

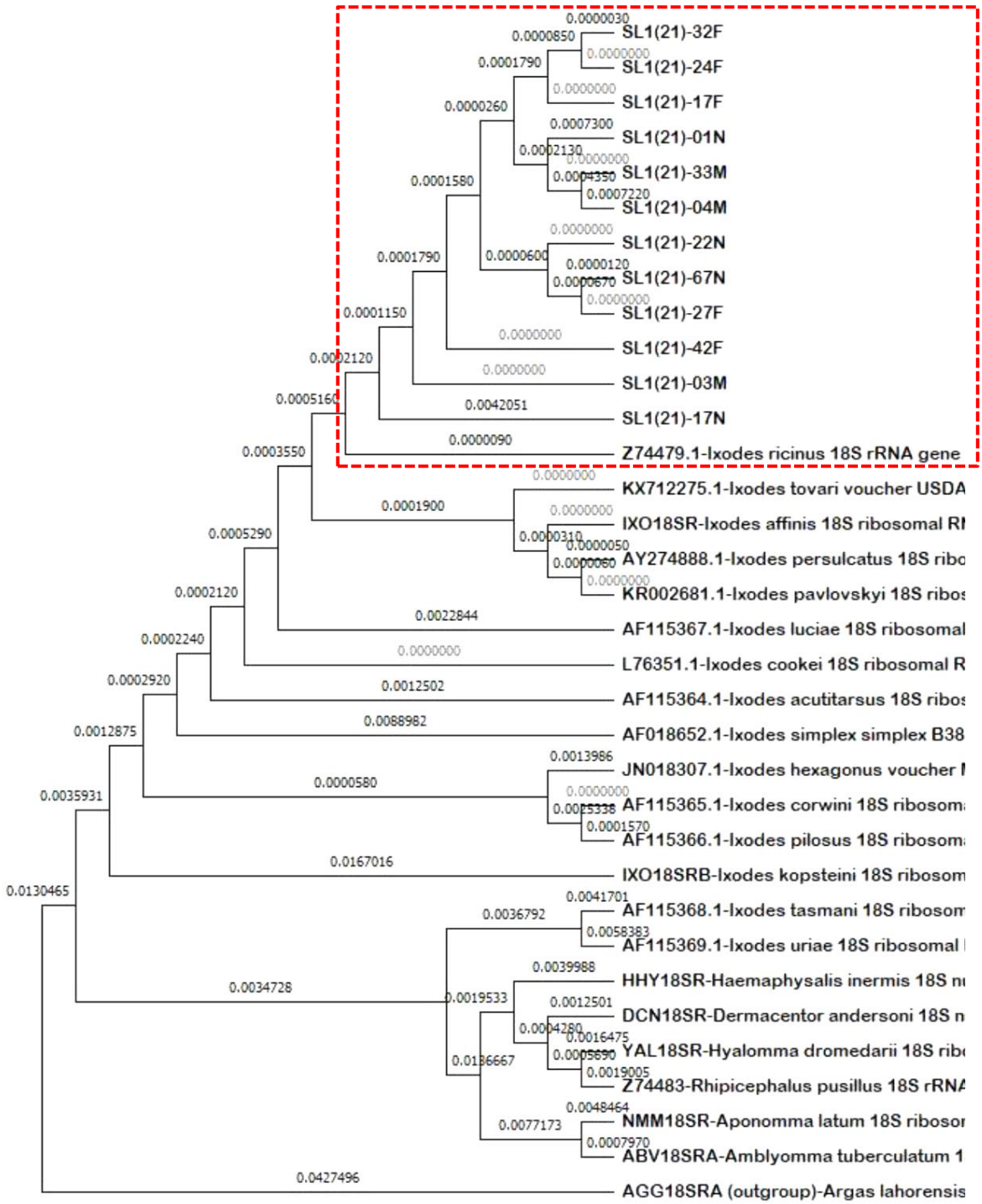


Figure 3.15. A rooted phylogenetic tree constructed using the PHYML method, predicting the closest evolutionary relationship through the substitution per site values of the reconstructed 18S-rRNA gene regions from 12 different TBEV positive tick samples [(SL1(21)~]. Phylogenetic tree was constructed against the 18S-rRNA gene sequence regions from 22 different tick species obtained from the NCBI nucleotide database and an outgroup sequence of 18S-rRNA gene region from the Argas lahorensis (NCBI Acc. No. AGG18SRA).

4. Discussion

4.1 The limitations associated with the already existing PCR methods at NIPH for the amplification of the whole genome of TBEV

The analysed Oxford Nanopore sequence data derived by the combined primer schemes of [(A)+(B)] and [(M)+(N)] at four tenfold serial dilutions (1×10^{-4}, 1×10^{-5},1×10^{-6}, and 1×10^{-7}), were able to identify the capabilities of each of the primer pools to generate whole genome of the TBEV-Hochosterwitz strain with the sequencing depth above 20 reads. Analysed data showed none of the combined primer schemes had the ability to PCR amplify the whole genome of TBEV-Hochosterwitz strain, at any of the tenfold dilutions considered in this study. The combined primer scheme [(M)+(N)] exhibited a sequence gap at 10493-10708 nt region corresponding to the JK-7-F/R primer pair. Previous studies carried out with the combined primer scheme [(M)+(N)] at NIPH, have been identified poor performance at the lower concentrations of the TBEV-Hochosterwitz RNA (Verbal communications). This was also confirmed by the poor amplification performance observed after the tenfold dilution at (1×10^{-3}) level during the PCR amplification followed by the visualization of the electroporated agarose gels of "JK" primer pairs (Section 3.4).

Therefore, this study mainly focused on the improvement of the primer schemes [(A)+(B)] and the sequencing experiments to evaluate their performance. Through the analysed sequence data, considerable reduction in per-base sequencing depth in the regions with below 20 reads identified. The regions with poor average sequencing depths were consistent among the tenfold dilutions and found in the genome sequence regions, 693-1607, 2770-2985, 4605-6500,7378-7579 and 8303-10871 nt (Figure A.3, Annexure I). Primer pairs corresponding to the lower sequencing depths were identified (Eur_1-F/R, Eur_3-F/R to Eur_6-F/R, Eur_11-F/R, Eur_14-F/R, Eur_15-F/R, Eur_17-F/R, Eur_18-F/R, Eur_20-F/R to Eur_22-F/R, Eur_29-F/R, Eur_33-F/R, Eur_34-F/R and Eur_36-F/R and Eur_39-F/R) (Table 3.2) and found to be consistent among the tenfold dilutions. Average sequencing depths generated by the rest of the primer pairs exhibited over the 20 reads in the evaluated tenfold dilutions (1×10^{-4} to 1×10^{-7}) for TBEV-Hochosterwitz strain.

Changes in the average breadth of sequencing above the 20 reads against the tenfold dilutions of the TBEV-Hochosterwitz strain were evaluated and the highest reduction (16.8%) was found between the tenfold dilutions (1×10^{-5} and 1×10^{-6}). The reduction in percentage of average breadth of sequencing between the tenfold dilutions (1×10^{-4} and 1×10^{-5}) and between (1×10^{-6} and 1×10^{-7}) were 9.8% and 0.3%, respectively. This indicates the combined primer scheme [(A)+(B)] is only capable of amplifying the higher concentration of TBEV-Hochosterwitz cDNA at the dilutions of 1×10^{-5} level with minimum of three gaps in the TBEV-Hochosterwitz genome at the sequencing depth over 20 reads (Figure A.3, Annexure I).

In addition, possible reasons for the poor performance of the primer scheme were identified to be the low percentage of primer identity to TBEV sequence variants and the poor primer positional accuracy values. Three primer pairs were identified to have a poor primer positional accuracy (Eur_1-F/R, Eur_29-F/R, Eur_38-F/R and Eur_39-F/R) giving amplifications at unspecific regions of the TBEV genome leading to the creations of gaps in the reconstructed genome. Out of 78 primers 37 exhibited primer identity below 100%. Per-base sequencing depth data shown at higher dilutions of the template TBEV cDNA, sequencing depths fell below 20 reads when the primer sequence identity was below 100% (Table A.2, Annexure I). This is very clear with the comparison of the data from the Table 3.2 with lower primer identity values vs. the per-base sequencing depth values of the Figure A.3, Annexure I. In addition, the percentage of TBEV sequences (out of 250 TBEV variants) with 100% identity to the primer sequences was found to be 23.68% (Data is not shown). Therefore, the whole genome amplifications of TBEV variants other than TBEV-Hochosterwitz is most likely not possible at low TBEV viraemic content using the already existing combined primer schemes [(A)+(B) and (M)+(N)].

4.2 Effects of temperature and the primer concentration on the PCR amplification of whole genome of TBEV

The best temperature for the whole genome amplification of TBEV using PCR was determined using two different approaches. The first method was by the PCR amplification of individual primer pairs followed by the visualization of the PCR products through the agarose gel electrophoresis. Electroporated PCR products shown that, 63 °C provide better amplifications with much higher yields and prone to have more of the unspecific amplifications compared to the PCR amplification at 65 °C. At 65 °C the PCR amplifications tend to form lower amounts of unspecific amplifications with lower amounts of desired PCR products as well (Section 3.5).

The second method was through the data generated by the sequence aligned maps (SAMs) of whole genome amplification of TBEV-Hochosterwitz strain against the combined primer scheme [(A)+(B)] at two different temperatures (63 & 65 °C). The resulting data were evaluated by comparing the percentage of average breadth of sequencing over 20 reads in the sequencing experiments one and two (Table A.4, Annexure I). The PCR amplifications carried out at 63 °C exhibited lower percentage of average breadth of sequencing over 20 reads than the PCR amplification carried out at 65 °C indicating better PCR amplification at 65 °C. But the total average breadth of sequencing was higher in the PCR reactions carried out at 63 °C compared to the PCR reaction carried out at 65 °C. This indicates, at 63 °C PCR was able to cover most of the regions of the TBEV genome, but the depth of the sequencing is not optimal as 65 °C. In addition, the number of gaps observed in the consensus sequence derived through the PCR amplification at 63 °C found to be less compared to the PCR carried out at the 65 °C (Figure 3.6 & 3.10). The main focus in the improvement was to achieve maximum coverage at first and the second to improve its depth of sequencing, therefore based on this assumption, the rest of PCR amplifications were carried out at 63 °C, instead of carrying it at 65 °C by giving more priorities to get more coverage over the whole genome to have fewer sequencing gaps at higher dilutions.

The effect of primer concentrations (0.15, 0.5 ad 1.0 mM) on the quality of the reconstructed whole genome of TBEV-Hochosterwitz strain was evaluated. Through the analysed sequencing depth and breadth values found out at higher primer scheme concentrations (0.5 and 1.0 mM), giving better sequencing depth and breadth values compared to the lower primer scheme concentration (0.15 mM). By the data tabulated (Figures 3.2, 3.9 and 3.10), it was very

clear when the primer scheme concentration is between 0.5 - 1.0 mM, average breadth and the average depth of sequencing increased by nearly 8 - 66% in some of the primer schemes {[(A+O)+(B+P)] and [(E+Q)+(F+R)]}, respectively. But one primer scheme [(G+S)+(H+T)] exhibited reduction in average breadth and depth values indicating which is not always consistent with changes in the primer compositions in each primer scheme. When considering the coverage of the reconstructed TBEV-Hochosterwitz genome, it is very clear that the concentration at 1.0 mM giving better coverage over the 0.5 mM concentration of primer schemes (Figure 3.10).

4.3 Effects of total RNA from tick samples and the dilution factor on unspecific amplifications

The effect of total RNA extracts from tick samples and their effect on PCR amplification of whole genome was evaluated in two separate experiments. One experiment was carried out by the PCR amplification of TBEV-Hochosterwitz cDNA (Dilution range: 10^{-4} - 10^{-8}) against individual “Eur_1-F/R” and “JK_1-F/R” primer pairs in the presence and the absence of a total RNA extract from a tick samples. In this experiment, the effect of PCR inhibitors from the blood meals accumulated in the tick samples was evaluated. It has been identified in several studies, proteins present in the mammalian blood, such as haemoglobin, immunoglobulin G (IgG) and heparin known to interfere the RT-PCR reactions by binding to the RNA (C. Schrader *et al.*, 2012). Through the analysed data no considerable inhibitions were observed by both observations in the electroporated agarose gel bands and the Oxford Nanopore sequenced data. Only the distinguishable significant difference was observed in the percentage of unmapped reads (unspecific amplifications) with the increase of 7.58% for “Eur_1-F/R” primer pair and 3.26% increase for the “JK_1-F/R” primer pair in the presence of total RNA from the tick sample compared to the in the absence of total RNA from the tick sample.

The unspecific amplifications and their percentage of variations depending on the level of dilution in the absence of the total RNA from tick sample were also compared against the primer schemes [(M+N) & (A+B)]. Through the visualized data (Figure A.9, Annexure I), with the increase in the dilution factor of the TBEV-Hochosterwitz RNA, the percentage of unspecific amplifications had increased exponentially. This led to a lower percentage of TBEV specific DNA products with a poor reconstruction of TBEV whole genome at higher dilutions.

This effect was very characteristic with the Oxford Nanopore sequence data derived from the TBEV positive tick samples where the percentage of unspecific reads nearly between 81-99.9% in most of the Oxford Nanopore sequenced data. This indicates the reduction in the targeted template cDNA of TBEV in the PCR mixture for the amplifications of whole TBEV genome leading to amplify the near exponential amount of increase of the unspecific by-products. Therefore, the importance of primer specificity towards the TBEV is crucial for the whole genome PCR amplification to achieve a properly mapped whole genome from a sample with low TBE viraemic content.

The unmapped reads derived though each of the Oxford Nanopore sequenced data of TBEV positive tick species were subjected to NCBI BLASTn analysis (Data is not shown, NCBI BLAST parameters: BLASTn, E value: 1×10^{-5}, NCBI BLAST database 2014). The data has shown, Majority of them were related to the transcripts sequence fragments from prokaryotic organisms and from the vector tick species. The main possible source of prokaryotic transcripts was through the prokaryotes associated with tick species. The transcripts related to the vector tick species were mainly from the 18S-rRNA ribosomal gene region and the percentages were ranging from 10.33-0.11% to the total number of Oxford Nanopore sequences derived from each of the TBEV positive tick species (Table A.15, Annexure I).

The reason for the amplification of 18S-rRNA region was not clearly identified but most probably due to the higher number of 18S-rRNA available in the cells may have played a critical role. In addition, primer sequence identities to the 18S-rRNA regions of *I. ricinus* (NCBI ACC. No: Z74479.1) were evaluated and did not find any sequence region exhibited with primer sequence identity over 50%, indicating least likely to be amplified by the primers used in the study. Further analysis with the TBEV-Hochosterwitz strain and the 18S-rRNA sequence of the *I. ricinus* (NCBI ACC. No: Z74479.1) using pairwise alignment found, three regions with full length un-gapped alignments with the pairwise identity ranging from 31.1 - 55.9%. In addition, the NCBI BLASTn analysis carried out with the 18S-rRNA sequence of the *I. ricinus* (NCBI ACC. No: Z74479.1) against TBEV (taxid: 11084) and found out the 18S-rRNA region (833-876 nt) exhibited 82% sequence identity to the TBEV strains (NCBI Acc. no. MN615728.1 & MN615726.1) in the region (10413-10454 nt). The higher number of 18S-rRNA sequence regions most probably through the annealing of the PCR products of TBEV sequences to the complementary regions of the 18S-rRNA sequences and finally amplifying the chimeric sequences.

New primer sequences were designed by screening against the human transcriptome sequence database obtained from the NCBI human RefSeq Transcripts (https://ftp.ncbi.nlm.nih.gov/refseq/H_sapiens/annotation/GRCh38_latest/refseq_identifiers/GRCh38_latest_rna.fna.gz). The sequence identity was maintained over 70% by facilitating to use the same primer pairs for the detection of TBEV in human clinical blood serum samples. But did not perform against the transcriptome library of *I. ricinus* due to limited time availability. This could have been able to avoid most of the unspecific amplifications further by the transcripts from ticks. In addition, a large majority of unspecific amplifications were also derived from the procaryotic origin and this exemplifies the sheer complexity and the difficulty of designing very specific primers to perform highly sensitive PCR at very low viraemic content to construct whole genome of TBEV.

4.4 Challenges in whole genome mapping of TBEV to Oxford Nanopore sequence data

The accuracy of the Oxford Nanopore sequencing technique is ranging from 86% to 95%. The rapid barcoding sequencing kit (SQK-RBK004) used in this study consisted with 95% accuracy since it is $1D^2$ type Nanopore sequencing technique (Wang *et al.*, 2021). This is also considerably a large value when it comes to a whole genome sequencing. Therefore, required to have higher sequencing depth to achieve the accurate results with compared to Illumina (Error frequency: 1×10^{-3}) and other sequencing techniques (El-Metwally *et al.*, 2014). In low viraemic samples, higher sequencing depth is very hard to achieve, since the initial number of template sequences are very low and due to the higher unspecific amplifications leading to lower the desired PCR products in the final DNA library prepared for the sequencing. In Oxford Nanopore sequencing techniques, there is no subsequent secondary amplifications as in Illumina technique. To achieve acceptable accuracy over the Oxford Nanopore sequence data generated in this study, sequencing depth was maintained over 20 reads for each of the reconstructed TBEV genomes by following the recommendations provided by the Radhakrishnan *et al.*, 2019. This led to have higher sequencing gaps with low amount of coverage in each of the TBEV genomes constructed from the TBEV positive tick samples and higher accuracy over the resulting consensus genome sequence fragments derived through the reference-based mapping.

The reference-based mapping of Oxford Nanopore reads against 250 different TBEV variants were carried out using Minimap2 tool. Some of the reference TBEV sequence variants (NCBI Acc. No: U27495, NC_001672, KC835595) contained poly-A regions stretching from 3-54 nt in length. During the reference-based mapping, it has been identified non-TBEV (poly-A regions containing transcripts fragments from bacteria and ticks) contaminants were mapped against the poly-A regions of the TBEV sequences in the sequence alignment map. Those regions were manually identified and removed only to preserve the sequence reads derived only from the TBEV with the help of NCBI BLAST analysis. To minimize the mapping of contaminants during the reference-based mapping the contigs length for the overlapping regions was increased over 20 nt length. This caused to reduce the number of mapped reads greatly. This issue was not observed with the *de novo* assembly techniques with trinity, where only the most accurate reads got assembled by avoiding the poly-(A) regions, this is mainly because of the lack of reference template for the assembling and only guiding with the overlapping regions. The main issue with the *de novo* assembly was the lower coverage of the mapped sequences derived from the reference-based mapping of the TBEV-Hochosterwitz strain due to the poor PCR amplifications of some of the regions in the TBEV genome and the lower number of overlapping regions to assemble.

The existence of quasispecies in TBEV with varying length of poly-(A) region revealed in a study carried out by the Asghar et al. in 2014. Through the critical analysis of Oxford Nanopore sequences of the TBEV-Hochosterwitz strain derived in this study and the analysed Illumina sequences from a previous study, revealed the presence of quasispecies in the TBEV-Hochosterwitz strain with the varying number of lengths of the poly-(A) tail ranging from 0-104 nt in length (Data not shown). Through the data generated by the reference-based mapping and also the de novo sequence assembly, these valuable information does not take into account to produce a final consensus sequence and it masks the true nature of the viral genome and its complete diversity in a viral focus. In addition, when the individual Oxford Nanopore sequence reads from TBEV-Hochosterwitz strain were BLAST against NCBI nucleotide archive, individual sequences gave higher percentage identity towards various TBEV variants other than the TBEV-Hochosterwitz strain, indicating even in a single viral culture, the existence of co-variants in very low percentages is possible as it has previously discovered by Helmová *et al.,* 2020.

In the reference-based assembly of the TBEV sequences generated by the Oxford Nanopore sequencing, two main possible issues were identified such as the reference sequence bias on

the newly constructed sequences and the percentage of sequencing error (Martiniano *et al.*, 2019). When the percentage of error in the sequencing is high, there is a possibility to increase the sequencing bias since the higher percentage of error can alter the conserved nucleotide position along the column of aligned reads. This leads to more bias towards the reference sequence and the outcome will be more similar to the reference sequence. This reference bias can be eliminated by *de novo* assembly rather than the use of reference-based mapping. But the main drawback of the *de novo* assembly is the formation of chimeric assembled sequences due to the similar overlapping regions. This would be more common with the low viraemic samples when the number of reads may be limited to calculate proper statistics by the *de novo* assemblers. In addition, there is a possibility to become more complex when it comes to the existence of co-variants of TBEV in the same sample and finally leading to a chimeric and new sequence variant, which may be totally different from the previously known sequence variants. In addition, repeated sequence regions (Found to have 211-318 repeated regions with the 9-12 nt lengths) along the TBEV genome (data are not shown) give rise to the chimeric sequences when the *de novo* sequence assembly is performed with lower k-mer length to generate maximum coverage of the assembled sequences. Therefore, the complexity could arise even more with low sequence coverage.

4.5 Relative accuracy and the reliability of the semi quantitative real-time PCR techniques employed in this study

The identification and the semi quantitative determination of TBEV positive samples of ticks were carried out by two separate semiquantitative real-time PCR techniques method I & method II. Together with two semi quantitative real-time PCR techniques, 26 samples were identified as TBEV positive from a total of 172 tick samples analysed. The semiquantitative real-time PCR analysis method I alone, identified 16 of TBEV positive tick samples, whereas 14 TBEV positive tick samples were identified by the semiquantitative real-time PCR method II. By both methods, only four samples of TBEV positive tick samples were identified and the rest of the samples were identified by only one of the methods employed in the study. The data generated by both methods were analysed by single factor ANOVA test and found out the results generated by those two techniques exhibited significant difference.

The confirmations of the analysed data were done using the Oxford Nanopore sequencing of individual tick samples identified as TBEV positive by both semiquantitative real-time PCR techniques. Thirteen out of 26 samples were confirmed to be true TBEV positive and the rest of the 13 tick samples were confirmed as false positive. Through the semiquantitative Realtime PCR method I alone, nine out of 172 samples (5.2%) were found to be false positive and six out of 172 samples (3.5%) were found to be false negative for TBEV. Whereas the accuracy of the semiquantitative Realtime PCR method II was found to be six out of 172 samples (3.5%) were false positive and five out of 172 samples (2.9%) were false negative for TBEV. Both test methods had equal number of true positive samples with the value of seven out of 172 samples (4.1%), but the true positive samples were different from each test.

As a comparison, the RT-semiquantitative real-time PCR method II, identified to be slightly more accurate (~2.3%) than the semiquantitative Realtime PCR method I for the detection of TBEV in tick samples. The evaluated test methods can have higher false negative TBEV percentage value. This is mainly because the false negative values were calculated only through the TBEV positive tick samples generated from each test, but the rest of the samples might have a higher number of false negative samples and those were unable to identify through the data generated by the test methods. Therefore, these numerical values can be considered as relative accuracies generated comparatively to each test method of testing for the detection of TBEV from tick samples. To get more accurate evaluations towards each test method used for the TBEV detection in tick samples, need to have standard known positive and negative tick samples confirmed for TBEV using much more accurate methods.

4.6 Effect of Random primers and TBEV specific primers on RT-PCR and subsequent whole genome amplification of TBEV

The effect of random primers and the TBEV specific primers (a pool of all the TBEV specific reverse primers were used in this study) on the RT-PCR of TBEV-Hochosterwitz strain were evaluated by conducting two separate RT-PCR reactions against TBEV-Hochosterwitz RNA at the tenfold dilution level (1×10^{-6}). The resulting cDNA was subjected to PCR amplification with the combined primer scheme [(G+H)+(S+D)] and the purified PCR products were subjected to Oxford Nanopore sequencing and the quality of the reconstructed sequences were

evaluated using reference-based mapping against TBEV-Hochosterwitz strain (NCBI Acc. No. MT311861) as the procedures mentioned in the sections 2.12 and 2.13.

Through the reference-based mapping of both Oxford Nanopore sequence data, found that only the RT-PCR carried out with random hexamers, performed well by giving higher number of mapped sequences (reads: 20054) in the reference-based mapping of the Oxford Nanopore sequence data. The RT-PCR carried out with a pool of all the TBEV specific reverse primers shown to have performed poorly by giving a lower number of mapped sequences (reads:48). But theoretically RT-PCR carried out with a pool of all the TBEV specific reverse primers expected to be performed well with the subsequent PCR amplification compared to the random hexamers. This is mainly due to the expected low level of unspecific amplifications, also due to the higher specificity of the TBEV primers towards the TBEV RNA sequences and the proper long linear cDNA synthesis than the random hexamers.

Even though the number of mapped sequences were found different, the number of Oxford Nanopore reads were found to be considerably similar in both cases (reads: 44-48,000). To understand the reason to not to have the expected outcome, all the Oxford Nanopore reads derived though a pool of all the TBEV specific reverse primers were subjected to MegaBLAST ($e = 1\times10^{-5}$) and identified 4966 reads with TBEV genome fragments ranging from 138-855 bases in lengths. Those identified reads were subjected to NCBI BLASTn analysis and though the sequence alignments, found that each sequence consists with several TBEV genome sequence fragments, arranged in a different order (Chimeric sequences) than the reference TBEV-Hochosterwitz sequence. Due to this, the reference-based mapping by the Minimap2, did not perform to map the sequences to the reference TBEV-Hochosterwitz genome. Further analysis found, repeated sequence elements (There are nearly 211-318 repeated elements in the TBEV genome sequence variants, data is not shown) found in the TBEV genome has involved in the development of chimeric sequences. This gives an idea about the native tertiary folding pattern of the TBEV RNA and which is governed by the repeated elements of the RNA sequence to generate quasi degenerate sequences in a viral focus. To confirm this hypothesis, it requires further studies with a different approach to an experiment other than the focus of this study.

4.7 Poorly amplified regions of TBEV-Hochosterwitz genome

Per-base sequencing depth graphs of each of the reconstructed TBEV-Hochosterwitz genome were evaluated for the identification of poorly amplified regions before and after the introduction of newly designed primer pairs. Through the data, it was found sequencing depth of specific regions (1200-1560, 4600-6500, 7900-8300, 8400-8550 and 9200-9520 nt, a reference to the TBEV-Hochosterwitz sequence NCBI Acc. No. MT311861) of the TBEV genome did not improve after the changing of the primer pairs. Corresponding poorly performed primer pairs (Eur_6-F/R, CJB_6-F/R, Eur_18-23-F/R, Eur_34-36-F/R, Eur_39-F/R, CJB_18-23-F/R & CJB_33-36-F/R) were identified and found to have no exact reason to perform poorly during the PCR with considering the primer identity and the melting temperature of the designed primers. The main possibilities to have poor PCR amplification in the identified regions may be based on cross interactions among the other primes in the same primer pool, by the existence of highly repeated regions, the secondary stable secondary structure formed or by the degradation of template TBEV RNA. To improve the PCR amplification ability in the sequence region (9200-9520 nt) of the TBEV genome, three different primer pairs were designed (Eur_34-F/R, CJB_34-F/R and CJB_34-NF/R), and found, none of the primer pairs were capable enough to properly amplify the region at very low viraemic content above the 4th tenfold dilution. All the primer pairs consisted with 100% sequence identity to the template reference sequence and therefore the unspecific amplifications can be ruled out. To amplify the 1200-1550 and 4600-8550 nt regions of the TBEV-Hochosterwitz genome, four sets of primer pairs (Eur_6-F/R, CJB_6-F/R, Eur_18-33-F/R and CJB_18-33-F/R) were tested, but none of the primer sets were able to amplify the corresponding regions above the 4th tenfold dilution.

4.8 Calculated of primer pair efficiencies and the method developed

Determination of the performance of each primer pair used in this study was carried to identify the how well each primer pair was performing under different dilution levels during the PCR amplification of TBEV-Hochosterwitz genome. Primer pair efficiency was determined using SAMs generated from each of the Oxford Nanopore sequenced data and identifying the sequence reads generated by each of the primer pair separately, plotting them against each concentration level and identifying the graphs intercept through the regression analysis. A threshold value was defined to categorize and identify the poorly performing primer pairs from the better performing primer pairs by assigning the value as below zero. This helped to isolate the better performing primers from the poorly performing primers. To calculate the primer performance accurately, minimum of three SAMs at three different concentrations required as necessary for the calculation of the intercept of the graph equation.

In the first two whole genome sequencing experiments carried out in this study contained three different concentration levels, but the third experiment was carried out with only two experiments land which led to poor results and giving higher number of poorly performing primer pairs (Table A.13, Annexure I). Through the results generated by the first two sequencing experiments, identified poorly performing primer pairs and their calculated primer efficiencies were found to be well correlated with the issues identified with the primer identities and inaccurate positioning in the TBEV-Hochosterwitz genome. This indicates the accuracy of the method used to calculate the primer pair efficiencies using the SAMs files generated through the reference-based assembly of Oxford Nanopore sequence reads of TBEV genome at three different dilution levels. This method can be used as a better optimization technique for whole genome sequencing applications to find out poorly performing primer pair, to replace them with the better performing primer pairs. Finally can be optimized to develop a reliable and accurate method for a development of high efficiency PCR techniques for whole genome sequencing.

4.9 Prevalence, viraemic content and the sequence diversity of TBEV in tick samples from Larvik

Through the data generated by both semiquantitative real-time PCR analysis and the Oxford Nanopore sequencing, a total of 13 tick samples were confirmed to have genome sequence fragments of TBEV. The composition of TBEV positive tick samples were, five nymph tick pools, five adult female ticks and three adult male ticks, giving a prevalence of 0.68 ± 0.30%, 10.90% and 6.00%, respectively. The prevalence values calculated by only the semiquantitative real-time PCR techniques found to be much higher (Pooled nymph tick samples: 1.40 ± 0.40%, adult female ticks: 13.04% and adult male ticks: 20.00%). The disparity over two methods were due to the false positive results generated by the semiquantitative real-time PCR techniques in this study. The error associated with the false negative results of the semiquantitative real-time PCR techniques in this study can have considerable effect on the accuracy of the prevalence of TBEV in the study site, but with the available data it is not possible to calculate.

A comparison can be made between the prevalence values received in this study and also with the reported studies carried out in Norway in 2011-2014. A previous study carried out in 2011 to 2014 for Vestfold County in Norway where the current study site (Larvik) is located, the prevalence of TBEV in adult ticks and nymph ticks were reported to be 5.5% mean MIR (Minimum Infection Rate) (range: 1.9-20.0%) and 0.4 mean MIR (range: 0-0.8%), respectively (Vikse *et al.*, 2020). The prevalence data reported in this study is more or less within the prevalence range given in the previous study data and hard to evaluate any changes in the prevalence, whether the prevalence of TBEV has increased or decreased. This is mainly because of the previous finding and prevalence values were given as a broad range and but not as an individual value to compare with the current values. When the prevalence values compared with the other regions in the Europe (In adult ticks: 1.3-5.8%, in nymph ticks: 0.5-14%,) the prevalence values calculated in this study also the similar in range (Gerhard *et al.*, 2021, Gresíková & Noseck, 1966).

Through the Ct values received for each of the TBEV positive tick sample analysed in this study, the relative quantities of viraemic content in each TBEV positive tick samples was estimated. This was done by comparing the Ct values received for the tenfold dilutions of TBEV-Hochosterwitz strain and the tick samples by the semiquantitative real-time PCR data. The data has shown the viraemic content of tick samples were equivalent to the tenfold dilution

range of log10: -6 - -7 of TBEV-Hochosterwitz RNA extracts. Only a single tick sample contained exceptionally high amount of TBEV viraemic content, giving very low Ct value (6.73) in the semiquantitative real-time PCR carried out using the method II. But the average breadth and depth of sequencing contained comparatively lower values to the other TBEV positive tick samples analysed in this study, indicating some lower Ct values probably due to the contaminants present in the sample.

Higher percentage of an average breadth of sequencing above 20 reads was found in both SL1(21)-01N and SL1(21)-17F samples giving 13.6% and 13.4%, respectively and their corresponding Ct values by the semiquantitative real-time PCR were found to be 34.07 and 33.69. The rest of the confirmed TBEV positive tick samples exhibited lower average percentage of breadth of sequencing above 20 reads ranging from 7.3 – 0.0%. The corresponding Ct values given by the semiquantitative real-time PCR were ranging from 6.88-36.77. Some of the TBEV positive samples exhibited lower Ct values, but the average breadth of sequencing was not high as expected by comparison to the other samples. This might be due to some other factors influencing the sequencing depth, mostly due to the TBEV sequence subtypes which are not giving 100% identity to the primer sequences used for the PCR in the whole genome amplifications of TBEV, amount of tick RNA present in the sample or else by the factors influencing the real-time PCR.

To understand the TBEV sequence diversity in the study site Larvik near the southern coastal region of Norway, phylogenetic tree analysis was carried out for each reconstructed TBEV genome fragments from tick samples. Through the phylogenetic tree analysis based on the substitution per site values, identified four different TBEV sequence variants previously categorized as belonging into the TBEV-Eur subtype (Boelke et al., 2019, Egyed et al., 2021b, Paulsen et al., 2021). The origin of each reference strain found to be far from Norway and the possibility of finding all four different strains in a single site is highly unlikely. In addition, none of the sequences had any close relationship to the TBEV-Mandal sequence (NCBI Acc. No. KF991107) reported in Norway through a previous study carried out by Asghar et.al 2014 using the samples collected in year 2009. To confirm the accuracy of the phylogenetic tree construction, poorly reconstructed genome sequence of TBEV-Hochosterwitz strain (MT311861_DHSD_-6) was also incorporated into the MSA and phylogenetic tree construction. The introduced sequence was positioned closer to the reference TBEV-Hochosterwitz strain sequence (NCBI Acc. No. MT311861), indicating the accuracy of the phylogenetic tree analysis carried out in this study.

In addition, all three of the TBEV strains (SL1(21)_17F, SL1(21)_17N and SL1(21)_33M) were identified to be evolutionary closer to the TBEV strain (NCBI Acc. No. KY069125) and clustered into a single clade in the phylogenetic tree. Those sequences were also found to have similar PCR amplified regions in the multiple sequence alignment (Figure 3.13), indicating sequence similarity and the accuracy of the method. The TBEV-KY069125 sequence deposited in the NCBI database is through an unpublished work and the TBEV strain was isolated from a *Ixodes persulcatus* in Altai region, Western Siberia.

The TBEV genome sequence segments obtained from the tick sample [SL1(21)_01N], found to be evolutionary closely related to the TBEV strain (NCBI Acc. no. MK922617) reported from the Saxony region in Germany in this study by the phylogenetic tree analysis. The TBEV-MK922617 strain is known to infect humans, leading to cause clinical TBE disease, associated and isolated from *I. ricinus* tick species. This TBEV strain know to be closely evolutionary related to the TBEV strain isolated from Poland in 1971 and a known common TBEV variant circulating between the east-west regions in Europe (Boelke et al., 2019). The possibility of existing in southern coastal region in Norway can be considered based on the proximity to the previously reported areas in western Europe.

Three other genome sequence fragments obtained from tick samples [SL1(21)_27F, SL1(21)_24F and SL1(21)_32F] were identified to be closely associated with the TBEV-Hochosterwitz strain. The association between the TBEV-Hochosterwitz strain and the newly discovered TBEV genome sequence fragments from the tick samples may leading to a suspicion of a possible contamination with the reference standard TBEV-Hochosterwitz strain RNA sample used in this study for the optimization of the PCR technique for the whole genome sequencing of low viraemic TBEV samples. The sequencing experiments were carried out in separate and the possibility of contaminations were minimum. But this could be attributed to the close sequence similarity between the European subtypes and the possibility of the generation of quasi degenerative subtypes.

The other TBEV variant identified in this study was from an adult female tick sample [SL1(21)_42F] and found out evolutionary closer to the TBEV strain (NCBI Acc. no. MW256716) isolated from Hungary in 1952. The TBEV strain (MW256716) known to be phylogenetically closer to the Neudorfl and Czech isolates from 1953 (Egyed et al., 2021b). Therefore, the close geographically distribution throughout the Europe is more reasonable for the possible existence in southern coastal region in Norway.

The higher sequence diversity found in this study may be due to the poor sequencing depth and also due to the poor sequencing coverage of the reconstructed sequences. In addition, quasi degenerative sequence variants generation ability of the TBEV genome might have considerably affected the existence of higher sequence diversity of the tick samples collected from the same study site. This effect could be through the amplification of very low viraemic sub-sequences which may present in the sample and to finally give rise to a different chimeric sequence during the PCR amplification and during the reference-based mapping.

The quasi-degenerative ability of the TBEV genome was also found in this study with TBEV-Hochosterwitz strain, when it was PCR amplified with two different primer schemes. With combined overlapping primer scheme (M+N), the reconstructed TBEV-Hochosterwitz genome contained a gap generated by a deletion of the genome region (10,493-10,709 nt) and it was confirmed by analysing individually mapped sequences. The identified gap was not present when the combined overlapping primer scheme (A+B) was used to amplify the whole genome of TBEV-Hochosterwitz strain. This indicates, primers and their annealing abilities to specific quasi-degenerate sequence variants and the PCR amplification efficiencies of each can have considerable bias on the final outcome of the TBEV subtype identification. This effect could enhance in low viraemic samples, where the availability of the template RNA sequence fragments might be limited for the cDNA synthesis and subsequent PCR steps leading to chimeric sequences to generate final whole genome and due to that finally identify as a new TBEV variant. These quasi species and their effects on the genetic and phenotypic variants have been thoroughly studied in Růžek *et al.*, 2008. In addition to the quasi-degenerate sequence variants, the error associated with the Oxford Nanopore sequencing accompanied with a lower average percentage of sequencing depth could have contributed to generate more of the sequence variants during the phylogenetic tree analysis.

The quasi-degenerative sequence generation ability and the co-existence of TBEV variants are well known and have been studied previously (Asghar et al., 2016). The capability of development of quasi species and the existence of co-variants in TBEV, has given capability to generate extensively diverse host-pathogen interactions to evade the immune responses, invade different tissue types and organs in diverse set of host species (Gäumann et al., 2011). These capabilities of TBEV have been given excellent survival and persistent strategies withing the wild tick populations with the interactions among the mammals, birds and reptiles as the major TBEV incubating reservoirs (Kupča et al., 2010). The persistence of the natural foci is basically depend on the pathogen-vector-host interactions and the adaptability to

various hosts gives and added advantage to increase the survivability and the persistence of the TBEV foci in nature (Holzmann et al., 1997). The mechanism involving in the deletion and rearrangement of the TBEV RNA genome is yet to be unravelled. The probability of finding higher number of TBEV variants might be due to the quasi degenerative ability and also the existence of co-variants in the natural foci and that could be the reasons to identify higher number TBEV variants in this study.

4.10 Neuroinvasive ability of the identified TBEV strains

As envisioned, the development of PCR method to identify the TBEV viral subtype to predict the neuroinvasive ability and the disease severity of the strain through the whole genome sequencing. The identification of the mutation associated with the NS3 protein coding sequence to replace the A→S_{73} amino acid change, proximity to the NS2B-NS3 protease active site is one key determinants of the neuroinvasive ability demonstrated in mouse (Růžek et al., 2008). The corresponding sequence regions in the reconstructed TBEV genomes from the tick samples were unable to be evaluated because of the sequencing gaps created on the reconstructed TBEV genome sequence fragments from the tick samples due to the poor sequencing depths and coverage (Figure A.12, Annexure I).

A study carried out by Akiko et al. in 2003 by a comparison of attenuated neuroinvasive properties of two TBEV strains, corresponding changes in the nucleotides for the loss of neuroinvasive abilities have been characterized. The changes in the nucleotide sequence, A→G (1,579), A→G (10,228) and C→T (10,796) have been reported to be related to the loss of neuroinvasive abilities of the neuroinvasive TBEV variant (NCBI Acc. No. AB062063). The mutation at (A)-1579 nt support the higher neuroinvasive ability leading to an aspartate residue (Asp-483) (Nucleotide and amino acid residue numbering are according to the NCBI Acc. No. AB062063). This mutational changes have been identified to enhances the poor affinity towards the host antibodies to evade the immune system and infect the brain tissue in a study carried out with mice (Goto et al., 2003). Comparison to the corresponding region in the reconstructed TBEV genome sequence fragments obtained from the tick samples in this study and found out those regions were not recovered during the reconstruction of the TBEV genomes due to the sequencing gaps created during the whole genome amplification (Figure A.12, Annexure I).

In addition, the neuroinvasive virulence factor at the 10,227-10,229 nt region due to Lysine residue (3366) in the (Nucleotide and amino acid residue numbering are according to the NCBI Acc. No. AB062063) TBEV genome was evaluated. Only two reconstructed TBEV sequence fragments from the SL1(21)_17N and SL1(21)_42F tick samples had the corresponding genomic regions and found out which is associated with the Lysine residue in the corresponding positioning and indicating higher possibility to exist neuroinvasive ability by the mutational changes. The rest of the neuroinvasive virulence factors related to the TBEV strains from the tick samples [SL1(21)_17N and SL1(21)_42F] were not able to be determined due to the low sequencing depths and poor coverage over the corresponding regions (Figure A.12, Annexure I).

Accuracy of the evaluated data were low due to the poor sequencing depths (Average sequencing depth was below 10 reads) and considerably large number of gaps were found in the reconstructed TBEV genomes. Therefore, further analysis required to confirm the accuracy of the findings more precisely regarding the neuroinvasive abilities of the TBEV genomes identified in this study.

4.11 Vector tick species identified in Larvik

Species level of identification of TBEV positive tick species was carried out using the Oxford Nanopore sequence data. It gave an idea about the common tick species of the study site, Larvik in Norway. The identification of vector tick species was made possible to achieve mainly through the higher number of unspecifically amplified 18S-rRNA gene sequence fragments found in Oxford Nanopore reads from each of the PCR amplification of TBEV positive tick samples. Though the constructed phylogenetic tree it was identified as *Ixodes ricinus* through the minimum substitution per site values given for each of the tick samples. Therefore, *Ixodes ricinus* can be considered as the main vector tick species found in the study site, Larvik in Norway and responsible for the TBEV transmissions. This has been previously confirmed in many studies carried out related to the Vestfold County in Norway and this funding will support it further (Vikse *et al.*, 2020).

Initially, it was planned to identify both the vector tick species and the tick bearing host species through the mRNA from the host species preserved in the blood meals inside the ticks. After

BLAST data analysis, 18S-rRNA fragments from vector tick species were able to identify, but none of the samples consisted with RNA sequence fragments related to the tick bearing host species such as mammals, reptiles and birds. The possibility of finding host RNA in Nymph tick seems to unrealistic due to the smaller body size and the very small amount of host blood accumulated inside the body but, it might be possible to identify it though the adult tick samples if the studied adult ticks contained undigested blood meals in their digestive tracks. In addition to the RNA sequence fragments from tick species, higher number of unspecifically amplified sequences from prokaryotic origin were identified. Further evaluation though the implemented Bioinformatics workflow in this study, the sequence data derived from the Oxford Nanopore sequencing of tick samples can be used to identify the pathogenic bacteria and bacteria associated with the ticks for preliminary rapid screening purposes.

4.12 Reported studies on successful attempts of whole genome sequencing of TBEV and a comparison

There have been numerous studies of the sequencing of TBEV viral genomes using various sequencing techniques. Among them one recent study brough the attention mainly due to the success in the whole genome sequencing of TBEV directly taken from tick samples and TBEV infected cow and goat milk and cheese samples by the Gonzalez *et al.*, 2022 and their colleges. As they have described in the publication, they have used both adult ticks and nymph tick pools with 30 individual nymph ticks, TBEV contaminated milk and cheese samples for their studies. They were able to reconstruct whole genomes of TBEV from tick samples and milk samples. But they haven’t specified the whole TBEV genome was constructed from the individual adult tick sample or by the nymph pools with 30 nymphs in each. The viraemic content in the nymph pools in Gonzalez’s study suspected to be high and the hgiher TBE vireamic content probably facilitated to generate better and complete whole TBEV genome. In addition to that, the possibility to have TBEV positive nymph ticks in the nymph pools collected closer to the study site in Gonzale’s study might be really high. This is mainly due to the collected location where the recent incident of an outbreak of TBEV reported in the particular farm. In addition the viraemic content in the individual adult ticks might be really high mainly because of the recent TBEV incidents in the particular area. Therefore the vireamic contnet and the quality of the viral RNA found in the tick samples might be

considerably higher. In addition to these facotrs, they have specifically stated the minimal detection of viral particles with known content in the goat milk and the cheese samples were found to be 1×10^4 and 1×10^5 for tissue culture infectious dose at 50% (TCID50)/ml, respectively. For the ticks, the reported mean viral copy number were between $2\times10^2 – 4.8\times10^3$ RNA copies per sample and the corresponding Ct value obtained for the reconstructed TBEV samples from ticks, milk and cheese were sated as (Ct = 34) in Gonzale's study (Gonzalez *et al.,* 2022).

A comparison can be made by considering the TBEV viraemic content of the tick samples analysed in this study. Nymph pool tick samples used in this study contained only 10 nymphs and the amount was one third of the number of ticks in the Gonzalez study. The foci formation capability of the viral samples used to extract the TBEV-Hochosterwitz RNA was found to be 6.5×10^6 mL^{-1} (Determined in an another study carried out in NIPH, K. M. Paulsen *et al.,* 2021) and the foci formation values, can be assumed as equivalent to the number of viral RNA in the sample. Ten fold dilution series prepared in this experiement contained approximately the number of viral RNA sequences between 4.16×10^6 to 4.16×10^{-3} mL^{-1} (At the 0th and 8th tenfold dilution). The lowest detection limit of both semiquantitative real-time PCR techniques employed in this study were at the dilution levels of log10: -7 and -8, giving Ct values between 37-44 for TBEV-Hochosterwitz RNA. The dilution factors at log10: -6, the calculated TBEV sequence copy number approximately around 4.15 mL^{-1} and the corresponding Ct values obtained for the sample by the semiquatitative real-time PCR analyssis were between Ct: 32.13-34.04. The Ct values received by the semiquantitative real-time PCR analysis of TBEV positive tick samples in this study found to be between Ct: 30.72-36.77 and the corresponding TBEV viraemic copy number calculated to be approximately between 41.6 – 4.16 in each RT-PCR reaction.

The number of copies of TBEV RNA present in the tick samples studied in Gonzalez's study ($2\times10^2 – 4.8\times10^3$) and this sutdy [41.6 – 4.16 /(8 µL)] were found to be considerably different and the amount was found to be really less in this study than the study carried out by the Gonzalez and their colleges. In addition, the Ct value that they have reported (Ct: 34) for the tick samples and this might be due to the low sensititvity of the RT-qPCR that they have used to detect the NS5 region of the TBEV over the envelope protein coding region of TBEV used in this study.

In this study the samples were collected about 9 months before conducting the experiments from a random site near a previously reported area and stored in a fridge. Storing, freeze and thawign might have degraded the quality of the TBEV viral RNA up to a certain extent. In addition the viraemic content of the collected tick samples might be really less compared to the site of collection carried out by the Gonzalez and colleges near the farm with acute TBEV cases were reported. In addition to the quality of the RNA collected might be higher due to the analyssis carried out closer to the collection date and the higher viral circulation frequency in the study site.

By considering all the information and the comparisons between current study and the reported study by the Gonzalez and colleges, the main reason to have incomplete reconstructed TBEV viral genome in the present study was most probably due to the very low vireamic content in the tick samples analyzed and the success in the reported Gonzalez's study might be due to the higher TBEV viraemic content in ticks, milk and cheese samples that they have analyzed.

4.13 Improved performance, limitations and the further developments in PCR

The initial primer schemes [(M+N) and (A+B)] used in this study were capable of amplifying whole genome of TBEV at the higher viraemic content. At low viraemic content whole genome amplification ability of the initial primer schemes became poor and lead to sequencing gaps in the reconstructed TBEV genomes. The TBEV positive tick samples and the whole genome amplification ability directly from the samples found to be impossible to achieve with the initial primer schemes due to the poor performance in the RT-PCR and PCR amplification stages. Therefore, the new primers to replace poorly performing primers were introduced to improve the whole genome amplification of TBEV.

The performance was compared and found to have a better improvement in the whole genome amplification ability at low TBEV viraemic content, directly from the tick samples nearly at the below Ct = 30 value of semiquantitative real-time PCR equivalent to the tenfold dilution at log10: -6 with TBEV-Hochosterwitz strain. With the best modified primer scheme [(G+S)+(H+T)] the percentage of average sequencing breadth was increased by 11.3% and 21.5% at both dilutions log10: -5 and -6, respectively. In addition to the sequencing coverage

was extended beyond the capabilities had with the initial primer schemes and found out extended sequence region from 10,871 to 11,141 nt of TBEV-Hochosterwitz genome (NCBI Acc. no. MT311861) (Figure A.10-11, Annexure I) by the introduction of primer pairs (CJB_37-F/R to CJB_39-F/R) during the PCR reaction. Therefore, with the new changes in the primer compositions, PCR amplification ability towards the sequencing coverage of TBVE genome has increased to cover the full length of the 3'-UTR complexity region.

The newly designed primer pairs of the optimized primer schemes are capable of amplifying majority of the TBEV subtypes and it is an added advantage to overcome the genome sequence variations commonly associated with the TBEV genome. By utilizing newly designed optimized primer schemes, the PCR amplification of the whole genome of TBEV-Eur, TBEV-Sib and TBEV-Far in moderately high viraemic content samples would be able to achieve. This will facilitate to sequence and detect randomly distributed TBEV subtypes away from the known geographical regions.

The primer sequence positions in the TBEV genome were carefully designed to avoid the palindromic sequence regions, to avoid the stem-loop elements of the TBEV cDNA. Moreover, the primer sequences of the optimized primer schemes consist with best suitable primers for the TBEV genome amplifications in blood serum samples from human patients, since the primers were designed to avoid the possible unspecific amplifications could arise with the human genome and transcriptome sequences. Therefore, the application of optimized primer schemes for the detection of TBEV in human serum samples would be a possibility with low unspecific amplifications.

The tick samples evaluated in this study were found to be well below the lowest margin of the capability of the best primer scheme to amplify the whole genome of TBEV without any sequencing gaps in the final reconstructed genome. A notable difference was observed during the 2nd sequencing experiment where all the primer schemes found be performed well compared to the other two attempts of whole genome sequencing experiments with TBEV-Hochosterwitz RNA. The most probable reason to achieve this unexpected comparatively higher whole genome sequencing coverage and depth values were mainly due to the increase in the amount of reverse transcriptase used during the cDNA synthesis. The quality of the cDNA has a great impact in the subsequent PCR and the DNA library development. Inability

to synthesis cDNA to cover the whole genome of the TBEV will lead to sequencing gaps in the end. But further analysis and replicated experiments are needed to evaluate and to clarify this effect and the reproducibility.

To improve the quality of the RT-PCR and the PCR techniques employed in this study to achieve the whole genome sequencing of TBEV directly from the tick samples with low viraemic content, the following changes are recommended. The improvement in the specific amplifications, reduction of unspecific amplifications, the removal of mRNA and excessively remaining rRNA from the total RNA extract of tick samples is a critical factor. Which may increase the number of amplified TBEV specific PCR products in an exponential manner at lower dilutions. To achieve the removal of mRNA, several well-established methods can be utilized before the cDNA synthesis or during the cDNA synthesis. One such method is to capture and remove the poly-(A) regions containing mRNA before the cDNA synthesis by using mRNA capture techniques. The second method is to minimize the cDNA produced by the mRNA by minimizing the possibilities of amplifying poly-(A) regions by removing oligo (dT) primer sequences employed during the cDNA synthesis (Zhao *et al.,* 2014). But these techniques can have some drawbacks of having a low yield of recovered viral RNA, leading to the poor cDNA synthesis and the loss of sample, but these effects can be eliminated by optimizing the conditions and the parameters used in the process.

The TBEV specific primers and their applicability in cDNA synthesis of TBEV RNA were tested in this study and found incapable to enhance only the cDNA synthesis towards the TBEV. This is due to the complex chimeric sequence products formed. In addition, the use of a mixture of TBEV specific primers with random hexamers with appropriate ratios might be a good solution to enhance the TBEV specific cDNA synthesis and to minimize the unspecific amplifications. In this study, only carried out the cDNA synthesis with a mixture of TBEV specific RNA and random hexamers together and evaluated only the dsDNA concentrations and a found considerably higher amount of dsDNA in the purified PCR products. Unfortunately, resources were limited to conduct and to evaluate the quality of the sequences generated by the changes of the cDNA synthesis through the combination of both TBEV specific primers and random hexamers during the cDNA synthesis.

In addition to the above-mentioned methods to remove mRNA from the total RNA extracts from ticks, the need of removal of rRNA is found to be critically important before the cDNA synthesis. This will reduce the amplification of rRNA regions from samples and enhance the

TBEV specific cDNA synthesis. To achieve this, application of RNase H-based ribosomal RNA depletion methods are much more efficient and need to be applied with the specifically designed primers by targeting the tick rRNA (Duan *et al.,* 2020). This method can be designed to remove rRNA from the blood serum samples from mammals as well. By utilizing these techniques, the elimination of unspecific amplification during the PCR amplification stage would be possible to a greater extent.

In addition to the above viral RNA enrichment strategies, the use of RNA-dependent RNA polymerase in a cell free system to amplify the viral RNA sequences is a novel proposed approach which can be introduced here. In this proposed method the purified active NS5 protein (RNA-dependent RNA polymerase) can be introduced to a cell free reaction system to amplify the TBEV genome by specifically selected over the other RNA as it is happening naturally inside the host cells. This will amplify the TBEV RNA by the RNA dependent RNA polymerase activity in a single reaction before the cDNA synthesis. The subsequent cDNA synthesis and PCR stages will generate the complete whole genome due to the higher number of initial viral RNA, even though the quantity of TBEV RNA in the original sample is less significant to be reconstructed by the conventional PCR alone. This proposed method can be utilized to evaluate and identify other Flaviviruses and the other positive stand RNA viruses at very low viraemic content, where the detection and whole genome construction is impossible by the conventional RT-PCR and PCR techniques directly from the samples.

To improve the efficiency of the PCR stage of the whole genome amplification, several suggestions can be made regarding to avoid the gaps in the reconstructed whole genomes of TBEV. To increase the coverage and to reduce the sequencing gaps, the avoidance of the heterodimer formations among the primer pairs is important. The heterodimer formation ability of each primer sequence can be evaluated by a calculation and can predict the possible heterodimer formations in each primer pair. The other important factors to consider avoiding the possible sequencing gaps in the reconstructed genomes are mainly through the selection of highly conserved regions for the primer designing and also to consider the sequence diversity of the subtypes. In here the primer sequences need to be designed to accommodate higher numbers of TBEV subtypes found closer to specific geographical region. In this study, the evaluated TBEV sequences probably contained a higher sequence diversity due to the quasi degenerate and quasi species found in TBEV. This may create an additional challenge in the better primer designing to identify and amplify the whole genome of TBEV in a particular study site.

In addition to avoid the unspecific amplifications with tick rRNA and other genomic and transcriptomic regions of thick, MegaBLAST analysis can be performed against designed primers. By doing this would be possible to remove the possible primer sequences, which might amplify the tick cDNA instead of amplifying TBVE sequence regions. In addition to avoid the sequencing gaps and also to increase the sequencing breath, the number of primer pair schemes can be increased by only accommodating a fewer number of primer pairs for each primer scheme and carry out the PCRs and finally combine them together to create a complete DNA library. This will eventually enhance the quality of the reconstructed genomes by reducing heterodimer formations and unspecific amplifications.

The changes introduced in this study with the increase in the RNA amount, the cDNA amounts used in RT-PCR and PCR stages and DNA library construction have made a great impact on the average breadth of sequencing by increasing the quality of the final reconstructed TBEV genome. As changes in the RT-PCR, amount of template RNA was increased by eight times, the amount of cDNA increased by 3.5 times in the PCR. The volume of purified DNA used in the DNA library construction for the Oxford Nanopore sequencing, was increased by nearly three times than the recommended amounts to increase the sequencing breadth and the coverage. As a result, the total amount of percentage average sequencing breadth was increased by 67.5% in some of the TBEV positive tick samples, indicating the importance of optimization in the volumes of template RNA, cDNA and the DNA used in the RT-PCR, PCR and DNA library development stages.

5. Conclusion

The limitations of initial primer schemes [(M+N) & (A+B)] were identified and found to be incapable of whole genome PCR amplification of TBEV-Hochosterwitz strain at the tenfold dilution (1×10^{-4}). Twenty different poorly performing primer pairs were identified in the combined primer scheme (A+B) by the novel method established to determine each primer pair efficiency. Two modified primer schemes {[(2E+2Q)+(2F+2R)] and [(G+S)+(H+T)]} gave the best PCR amplification of whole genome of TBEV-Hochosterwitz cDNA at low viraemic content (at 6th tenfold dilution) giving maximum average breadth of sequencing (25.8 & 26.8%) over 20 read depth and the coverage (99.0 & 89.0%), respectively.

A total of 13 (Five nymph tick pools, five female adult ticks and three adult male ticks) out of 172 tick samples collected from Larvik at the southern coastal region in Norway were confirmed to be positive for TBEV. The prevalence of TBEV in tick samples were 0.68 ± 0.30% (Nymph tick pools), 10.90% (Adult female ticks) and 6.00% (Adult male ticks). The viraemic content of tick samples were equivalent between the tenfold dilutions (log10: -6 and -7) of TBEV-Hochosterwitz RNA sample.

The accuracy of two similar Semiquantitative real-time PCR method were compared and the method II found to be slightly more accurate than the method I and both methods exhibited false positive results (3.5 & 5.2%), respectively. The maximum average percentage of sequencing breadth over 20 reads and the percentage of coverage of the consensus sequence generated by the two best combined overlapping primer schemes during the whole genome sequencing against the TBEV positive tick samples fell below 13.6 & 50.56%, respectively. The closest evolutionary relationships of the TBEV genome sequence fragments from the TBEV positive tick samples were found to be clustered into four different clades within the European subtype and the TBEV positive ticks were identified as the *Ixodes ricinus*.

Further optimization of RT-PCR and PCR are necessary to increase the average breadth and coverage of whole genome sequencing at low viraemic TBEV positive tick samples to increase the accuracy of the determination of neuroinvasive abilities and the identification of TBEV sequence variants.

6. References

Achazi, K., Nitsche, A., Patel, P., Radonić, A., Mantke, O. D., & Niedrig, M. (2011). Detection and differentiation of tick-borne encephalitis virus subtypes by a reverse transcription quantitative real-time PCR and pyrosequencing. *Journal of Virological Methods*, *171*(1), 34–39. https://doi.org/10.1016/j.jviromet.2010.09.026

Afgan, E., Baker, D., Batut, B., van den Beek, M., Bouvier, D., Čech, M., Chilton, J., Clements, D., Coraor, N., Grüning, B. A., Guerler, A., Hillman-Jackson, J., Hiltemann, S., Jalili, V., Rasche, H., Soranzo, N., Goecks, J., Taylor, J., Nekrutenko, A., & Blankenberg, D. (2018). The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. *Nucleic Acids Research*, *46*(W1), W537–W544. https://doi.org/10.1093/nar/gky379

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, *215*(3), 403–410. https://doi.org/10.1016/S0022-2836(05)80360-2

Anaconda Software Distribution (Vers. 2-2.4.0). (2020). Anaconda Inc. https://docs.anaconda.com/

Andreassen, A., Jore, S., Cuber, P., Dudman, S., Tengs, T., Isaksen, K., Hygen, H. O., Viljugrein, H., Ånestad, G., Ottesen, P., & Vainio, K. (2012). Prevalence of tick borne encephalitis virus in tick nymphs in relation to climatic factors on the southern coast of Norway. *Parasites and Vectors*, *5*(1), 1–12. https://doi.org/10.1186/1756-3305-5-177

Asghar, N., Lee, Y. P., Nilsson, E., Lindqvist, R., Melik, W., Kröger, A., Överby, A. K., & Johansson, M. (2016). The role of the poly(A) tract in the replication and virulence of tick-borne encephalitis virus. *Scientific Reports*, *6*(December), 1–13. https://doi.org/10.1038/srep39265

Asghar, N., Lindblom, P., Melik, W., Lindqvist, R., Haglund, M., Forsberg, P., Överby, A. K., Andreassen, Å., Lindgren, P. E., & Johansson, M. (2014). Tick-borne encephalitis virus sequenced directly from questing and blood-feeding ticks reveals quasispecies variance. *PLoS ONE*, *9*(7), 1–8. https://doi.org/10.1371/journal.pone.0103264

Asghar, N., Pettersson, J. H. O., Dinnetz, P., Andreassen, Å., & Johansson, M. (2017). Deep sequencing analysis of tick-borne encephalitis virus from questing ticks at natural foci reveals similarities between quasispecies pools of the virus. *Journal of General Virology*, *98*(3), 413–421. https://doi.org/10.1099/jgv.0.000704

Bandara, C. J. (2021). *Conserved regions identifier for Primer designing in PCR, RT-PCR, qPCR, RT-qPCR and NGS library development purposes using Python*. ResearchGate. https://doi.org/10.13140/RG.2.2.27207.21924/3

Bastakoti, S. (2019). *Master ' s Degree Thesis Surveillance and identification of tick-borne encephalitis virus (TBEV) in Scandinavia : Next generation sequencing (NGS) and Bioinformatics Master ' s Degree in Applied and Commercial*.

Beati, L., & Klompen, H. (2019). Phylogeography of ticks (Acari: Ixodida). In *Annual Review of Entomology*. https://doi.org/10.1146/annurev-ento-020117-043027

Belikov, S. I., Kondratov, I. G., Potapova, U. V., & Leonova, G. N. (2014). The relationship between the structure of the tick-borne encephalitis virus strains and their pathogenic properties. *PLoS ONE*. https://doi.org/10.1371/journal.pone.0094946

Biernat, B., Karbowiak, G., Werszko, J., & Stańczak, J. (2014). Prevalence of tick-borne encephalitis virus (TBEV) RNA in Dermacentor reticulatus ticks from natural and urban environment, Poland. *Experimental and Applied Acarology*. https://doi.org/10.1007/s10493-014-9836-5

Blankenberg, D., Gordon, A., Von Kuster, G., Coraor, N., Taylor, J., Nekrutenko, A., & Team, G. (2010). Manipulation of FASTQ data with Galaxy. *Bioinformatics*, *26*(14), 1783–1785. https://doi.org/10.1093/BIOINFORMATICS/BTQ281

Boelke, M., Bestehorn, M., Marchwald, B., Kubinski, M., Liebig, K., Glanz, J., Schulz, C., Dobler, G., Monazahian, M., & Becker, S. C. (2019). First Isolation and Phylogenetic Analyses of Tick-Borne Encephalitis Virus in Lower Saxony, Germany. *Viruses*, *11*(5), 462. https://doi.org/10.3390/v11050462

Brinton, M. A., & Dispoto, J. H. (1988). Sequence and secondary structure analysis of the 5′-terminal region of flavivirus genome RNA. *Virology*. https://doi.org/10.1016/0042-6822(88)90468-0

Carpi, G., Bertolotti, L., Rosati, S., & Rizzoli, A. (2009). Prevalence and genetic variability of tick-borne encephalitis virus in host-seeking Ixodes ricinus in northern Italy. *Journal of General Virology*, *90*(12), 2877–2883. https://doi.org/10.1099/vir.0.013367-0

Černý, J., Selinger, M., Palus, M., Vavrušková, Z., Tykalová, H., Bell-Sakyi, L., Štěrba, J., Grubhoffer, L., & Růžek, D. (2016). Expression of a second open reading frame present in the genome of tick-borne encephalitis virus strain Neudoerfl is not detectable in infected cells. *Virus Genes*. https://doi.org/10.1007/s11262-015-1273-y

Chen, Y., Ye, W., Zhang, Y., & Xu, Y. (2015). High speed BLASTN: an accelerated MegaBLAST search tool. *Nucleic Acids Research*, *43*(16), 7762–7768. https://doi.org/10.1093/nar/gkv784

Chene, P. (2012). NS3 Helicases as Drug Targets. *Current Chemical Biology*. https://doi.org/10.2174/2212796810903030334

Coffin J.M., Hughes S.H., V. H. E. (Ed.). (1997). Biochemistry of Reverse Transcription. In *Retroviruses*. Cold Spring Harbor Laboratory Press. https://www.ncbi.nlm.nih.gov/books/NBK19435/

Conze, T. M., Bagó, Z., Revilla-Fernández, S., Schlegel, J., Goehring, L. S., & Matiasek, K. (2021a). Tick-borne encephalitis virus (Tbev) infection in two horses. In *Viruses* (Vol. 13, Issue 9). https://doi.org/10.3390/v13091775

Conze, T. M., Bagó, Z., Revilla-Fernández, S., Schlegel, J., Goehring, L. S., & Matiasek, K. (2021b). Tick-Borne Encephalitis Virus (TBEV) Infection in Two Horses. *Viruses*, *13*(9), 1775. https://doi.org/10.3390/v13091775

Danecek, P., Bonfield, J. K., Liddle, J., Marshall, J., Ohan, V., Pollard, M. O., Whitwham, A., Keane, T., McCarthy, S. A., Davies, R. M., & Li, H. (2021). Twelve years of SAMtools and BCFtools. *GigaScience*, *10*(2), 1–4. https://doi.org/10.1093/gigascience/giab008

Daniel, M., Danielová, V., Fialová, A., Malý, M., Kříž, B., & Nuttall, P. A. (2018). Increased relative risk of tick-borne encephalitis in warmer weather. *Frontiers in Cellular and Infection Microbiology*, *8*(MAR), 1–11. https://doi.org/10.3389/fcimb.2018.00090

De Coster, W., D'Hert, S., Schultz, D. T., Cruts, M., & Van Broeckhoven, C. (2018).

NanoPack: visualizing and processing long-read sequencing data. *Bioinformatics*, *34*(15), 2666–2669. https://doi.org/10.1093/BIOINFORMATICS/BTY149

Deviatkin, A. A., Karganova, G. G., Vakulenko, Y. A., & Lukashev, A. N. (2020). TBEV Subtyping in Terms of Genetic Distance. *Viruses*, *12*(11), 1240. https://doi.org/10.3390/v12111240

Dobler, G., Hufert, F., Pfeffer, M., & Essbauer, S. (2011). Progress in Parasitology. In H. Mehlhorn (Ed.), *Progress in Parasitology*. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-21396-0

Duan, Y., Sun, Y., & Ambros, V. (2020). RNA-seq with RNase H-based ribosomal RNA depletion specifically designed for C. elegans. *MicroPublication Biology*, *2020*, 22–25.

Egyed, L., Biksi, I., Varga, T., Zöldi, V., & Dán. (2021a). Analysing the genomes of two tick-borne encephalitis viruses isolated in Hungary in 1952 and 2019. *Ticks and Tick-Borne Diseases*, *12*(6). https://doi.org/10.1016/j.ttbdis.2021.101806

Egyed, L., Biksi, I., Varga, T., Zöldi, V., & Dán, Á. (2021b). Analysing the genomes of two tick-borne encephalitis viruses isolated in Hungary in 1952 and 2019. *Ticks and Tick-Borne Diseases*, *12*(6), 101806. https://doi.org/10.1016/j.ttbdis.2021.101806

El-Metwally, S., Ouda, O. M., & Helmy, M. (2014). *Novel Next-Generation Sequencing Applications* (pp. 61–70). https://doi.org/10.1007/978-1-4939-0715-1_7

Estrada-Peña, A., & De La Fuente, J. (2014). The ecology of ticks and epidemiology of tick-borne viral diseases. *Antiviral Research*, *108*(1), 104–128. https://doi.org/10.1016/j.antiviral.2014.05.016

European Centre for Disease Prevention and Control. (2021). Tick-borne encephalitis. In: ECDC. Annual epidemiological report for 2019. Stockholm: ECDC; 2021. In *European Centre for Disease Prevention and Control* (Issue March). https://www.ecdc.europa.eu/sites/default/files/documents/AER-TBE-2019.pdf

Fitzpatrick, A. H., Rupnik, A., O'Shea, H., Crispie, F., Keaveney, S., & Cotter, P. (2021). High Throughput Sequencing for the Detection and Characterization of RNA Viruses. *Frontiers in Microbiology*, *12*(February). https://doi.org/10.3389/fmicb.2021.621719

Frey, S., Mossbrugger, I., Altantuul, D., Battsetseg, J., Davaadorj, R., Tserennorov, D., Buyanjargal, T., Otgonbaatar, D., Zöller, L., Speck, S., Wölfel, R., Dobler, G., & Essbauer, S. (2012). Isolation, preliminary characterization, and full-genome analyses of tick-borne encephalitis virus from Mongolia. *Virus Genes*, *45*(3), 413–425. https://doi.org/10.1007/s11262-012-0795-9

Füzik, T., Formanová, P., Růžek, D., Yoshii, K., Niedrig, M., & Plevka, P. (2018). Structure of tick-borne encephalitis virus and its neutralization by a monoclonal antibody. *Nature Communications*, *9*(1), 1–11. https://doi.org/10.1038/s41467-018-02882-0

Galgani, I., Bunge, E. M., Hendriks, L., Schludermann, C., Marano, C., & De Moerlooze, L. (2017). Systematic literature review comparing rapid 3-dose administration of the GSK tick-borne encephalitis vaccine with other primary immunization schedules. *Expert Review of Vaccines*, *16*(9), 919–932. https://doi.org/10.1080/14760584.2017.1358620

Gäumann, R., Mühlemann, K., Strasser, M., & Beuret, C. M. (2010). High-Throughput Procedure for Tick Surveys of Tick-Borne Encephalitis Virus and Its Application in a National Surveillance Study in Switzerland. *Applied and Environmental Microbiology*, *76*(13), 4241–4249. https://doi.org/10.1128/AEM.00391-10

Gäumann, R., Mühlemann, K., Strasser, M., Beuret, C. M., Gäumann, R., Mühlemann, K., Strasser, M., Beuret, C. M., Gäumann, R., Mühlemann, K., Strasser, M., & Beuret, C. M. (2010). High-Throughput Procedure for Tick Surveys of Tick-Borne Encephalitis Virus and Its Application in a National Surveillance Study in Switzerland. *Applied and Environmental Microbiology*, *76*(13), 4241–4249. https://doi.org/10.1128/AEM.00391-10

Gäumann, R., Růžek, D., Mühlemann, K., Strasser, M., & Beuret, C. M. (2011). Phylogenetic and virulence analysis of tick-borne encephalitis virus field isolates from Switzerland. *Journal of Medical Virology*, *83*(5), 853–863. https://doi.org/10.1002/jmv.21993

Gerhard, D., Wilheim, E., Michael, B., & Heinz-Josef, S. (Eds.). (2021). *The TBE Book* (4th ed.). Global Health Press Pte Ltd.

Gonzalez-Perez, I., Cayarga, A. A., Hernández, Y. P., de la Rosa, I. G., González, Y. J. G., León, C. S., & Álvarez, R. R. (2010). Long-term conservation of HCV RNA at 4°C using a new RNA stabilizing solution. *Journal of Virological Methods*, *168*(1–2), 207–211.

https://doi.org/10.1016/j.jviromet.2010.05.015

Gonzalez, G., Bournez, L., Moraes, R. A., Marine, D., Galon, C., Vorimore, F., Cochin, M., Nougairède, A., Hennechart-Collette, C., Perelle, S., Leparc-Goffart, I., Durand, G. A., Grard, G., Bénet, T., Danjou, N., Blanchin, M., Lacour, S. A., Franck, B., Chenut, G., … Lecollinet, S. (2022). A One-Health Approach to Investigating an Outbreak of Alimentary Tick-Borne Encephalitis in a Non-endemic Area in France (Ain, Eastern France): A Longitudinal Serological Study in Livestock, Detection in Ticks, and the First Tick-Borne Encephalitis Virus . *Frontiers in Microbiology*, *13*. https://doi.org/10.3389/fmicb.2022.863725

Gorbalenya, A. E., & Lauber, C. (2017). Phylogeny of Viruses ☆. In *Reference Module in Biomedical Sciences* (Issue January). Elsevier. https://doi.org/10.1016/B978-0-12-801238-3.95723-4

Goto, A., Hayasaka, D., Yoshii, K., Mizutani, T., Kariwa, H., & Takashima, I. (2003). A BHK-21 cell culture-adapted tick-borne encephalitis virus mutant is attenuated for neuroinvasiveness. *Vaccine*, *21*(25–26), 4043–4051. https://doi.org/10.1016/S0264-410X(03)00269-X

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., Di Palma, F., Birren, B. W., Nusbaum, C., Lindblad-Toh, K., … Regev, A. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology 2011 29:7*, *29*(7), 644–652. https://doi.org/10.1038/nbt.1883

Gresíková, M., & Noseck, J. (1966). Isolation of tick-borne encephalitis virus from Haemaphysalis inermis ticks. *Acta Virologica*.

Gritsun, T. ., Nuttall, P. ., & Gould, E. A. (2003). Tick-borne Flaviviruses. In *Advances in Virus Research* (Vol. 61, pp. 317–371). https://doi.org/10.1016/S0065-3527(03)61008-0

Gritsun, T. S., & Gould, E. A. (2007). Origin and evolution of flavivirus 5′UTRs and panhandles: Trans-terminal duplications? *Virology*, *366*(1), 8–15. https://doi.org/10.1016/j.virol.2007.04.011

Guglielmone, A. A., Apanaskevich, D. A., Estrada-Peña, A., Robbins, R. G., Petney, T. N., & Horak, I. G. (2014). The hard ticks of the world: (Acari: Ixodida: Ixodidae). In *The Hard Ticks of the World: (Acari: Ixodida: Ixodidae)*. https://doi.org/10.1007/978-94-007-7497-1

Harvey, E., & Holmes, E. C. (2022). Diversity and evolution of the animal virome. *Nature Reviews Microbiology*, *0123456789*. https://doi.org/10.1038/s41579-021-00665-x

Hayasaka, D., Aoki, K., & Morita, K. (2013). Development of simple and rapid assay to detect viral RNA of tick-borne encephalitis virus by reverse transcription-loop-mediated isothermal amplification. *Virology Journal*, *10*(1), 68. https://doi.org/10.1186/1743-422X-10-68

Heinz, F. X., Stiasny, K., Auer, G. P., Holzmann, H., Allison, S. L., Mandl, C. W., & Kunz, C. (1994). Structural changes and functional control of the tick-borne encephalitis virus glycoprotein E by the heterodimeric association with protein prM. *Virology*. https://doi.org/10.1006/viro.1994.1013

Helmová, R., Hönig, V., Tykalová, H., Palus, M., Bell-Sakyi, L., & Grubhoffer, L. (2020). Tick-Borne Encephalitis Virus Adaptation in Different Host Environments and Existence of Quasispecies. *Viruses*, *12*(8), 902. https://doi.org/10.3390/v12080902

Hjetland, R., Henningsson, A. J., Vainio, K., Dudman, S. G., Grude, N., & Ulvestad, E. (2015a). Seroprevalence of antibodies to tick-borne encephalitis virus and Anaplasma phagocytophilum in healthy adults from western Norway. *Infectious Diseases*, *47*(1), 52–56. https://doi.org/10.3109/00365548.2014.959044

Hjetland, R., Henningsson, A. J., Vainio, K., Dudman, S. G., Grude, N., & Ulvestad, E. (2015b). Seroprevalence of antibodies to tick-borne encephalitis virus and Anaplasma phagocytophilum in healthy adults from western Norway. *Infectious Diseases*, *47*(1), 52–56. https://doi.org/10.3109/00365548.2014.959044

Hofhuis, A., van den Berg, O. E., Meerstadt-Rombach, F. S., van den Wijngaard, C. C., Chung, N. H., Franz, E., & Reimerink, J. H. J. (2021). Exposure to tick-borne encephalitis virus among nature management workers in the Netherlands. *Ticks and Tick-Borne Diseases*, *12*(5), 101762. https://doi.org/10.1016/j.ttbdis.2021.101762

Holding, M., Dowall, S. D., Medlock, J. M., Carter, D. P., McGinley, L., Curran-French, M., Pullan, S. T., Chamberlain, J., Hansford, K. M., Baylis, M., Vipond, R., & Hewson, R. (2019). Detection of new endemic focus of tick-borne encephalitis virus (TBEV), Hampshire/Dorset border, England, September 2019. *Eurosurveillance*, *24*(47), 1–5. https://doi.org/10.2807/1560-7917.ES.2019.24.47.1900658

Holzmann, H., Stiasny, K., Ecker, M., Kunz, C., & Heinz, F. X. (1997). Characterization of monoclonal antibody-escape mutants of tick-borne encephalitis virus with reduced neuroinvasiveness in mice. *Journal of General Virology*, *78*(1), 31–37. https://doi.org/10.1099/0022-1317-78-1-31

Jääskeläinen, A. E., Sironen, T., Murueva, G. B., Subbotina, N., Alekseev, A. N., Castrén, J., Alitalo, I., Vaheri, A., & Vapalahti, O. (2010). Tick-borne encephalitis virus in ticks in Finland, Russian Karelia and Buryatia. *Journal of General Virology*, *91*(11), 2706–2712. https://doi.org/10.1099/vir.0.023663-0

Jansen, S., Smlatic, E., Copmans, D., Debaveye, S., Tangy, F., Vidalain, P. O., Neyts, J., & Dallmeier, K. (2021). Identification of host factors binding to dengue and Zika virus subgenomic RNA by efficient yeast three-hybrid screens of the human ORFeome. *RNA Biology*, *18*(5), 732–744. https://doi.org/10.1080/15476286.2020.1868754

Katargina, O., Russakova, S., Geller, J., Kondrusik, M., Zajkowska, J., Zygutiene, M., Bormane, A., Trofimova, J., & Golovljova, I. (2013). Detection and Characterization of Tick-Borne Encephalitis Virus in Baltic Countries and Eastern Poland. *PLoS ONE*, *8*(5), e61374. https://doi.org/10.1371/journal.pone.0061374

Kerlik, J., Avdičová, M., Musilová, M., Bérešová, J., & Mezencev, R. (2022). Breast Milk as Route of Tick-Borne Encephalitis Virus Transmission from Mother to Infant. *Emerging Infectious Diseases*, *28*(5), 1060–1061. https://doi.org/10.3201/eid2805.212457

Klaus, C., Hoffmann, B., Hering, U., Mielke, B., Sachse, K., Beer, M., & Süss, J. (2010). Tick-borne encephalitis (TBE) virus prevalence and virus genome characterization in field-collected ticks (Ixodes ricinus) from risk, non-risk and former risk areas of TBE, and in ticks removed from humans in Germany. *Clinical Microbiology and Infection*, *16*(3), 238–244. https://doi.org/10.1111/j.1469-0691.2009.02764.x

Kono, N., & Arakawa, K. (2019). Nanopore sequencing: Review of potential applications in

functional genomics. *Development, Growth & Differentiation*, *61*(5), 316–326. https://doi.org/10.1111/dgd.12608

Korenberg, E. I. (2000). Seasonal population dynamics of Ixodes ticks and tick-borne encephalitis virus. *Experimental and Applied Acarology*, *24*(9), 665–681. https://doi.org/10.1023/A:1010798518261

Kovalev, S. Y., Kokorev, V. S., & Belyaeva, I. V. (2010). Distribution of Far-Eastern tick-borne encephalitis virus subtype strains in the former Soviet Union. *Journal of General Virology*, *91*(12), 2941–2946. https://doi.org/10.1099/vir.0.023879-0

Kozlova, I. V., Demina, T. V., Tkachev, S. E., Doroshchenko, E. K., Lisak, O. V., Verkhozina, M. M., Karan, L. S., Dzhioev, Y. P., Paramonov, A. I., Suntsova, O. V., Savinova, Y. S., Chernoivanova, O. O., Ruzek, D., Tikunova, N. V., & Zlobin, V. I. (2018). CHARACTERISTICS OF THE BAIKAL SUBTYPE OF TICK-BORNE ENCEPHALITIS VIRUS CIRCULATING IN EASTERN SIBERIA. *Acta Biomedica Scientifica (East Siberian Biomedical Journal)*, *3*(4), 53–60. https://doi.org/10.29413/ABS.2018-3.4.9

Kuno, G., Chang, G.-J. J., Tsuchiya, K. R., Karabatsos, N., & Cropp, C. B. (1998). Phylogeny of the Genus Flavivirus. *Journal of Virology*, *72*(1), 73–83. https://doi.org/10.1128/JVI.72.1.73-83.1998

Kunze, M., Banović, P., Bogovič, P., Briciu, V., Čivljak, R., Dobler, G., Hristea, A., Kerlik, J., Kuivanen, S., Kynčl, J., Lebech, A., Lindquist, L., Paradowska-Stankiewicz, I., Roglić, S., Smíšková, D., Strle, F., Vapalahti, O., Vranješ, N., Vynograd, N., … Erber, W. (2022). Recommendations to Improve Tick-Borne Encephalitis Surveillance and Vaccine Uptake in Europe. *Microorganisms*, *10*(7), 1283. https://doi.org/10.3390/microorganisms10071283

Kupča, A. M., Essbauer, S., Zoeller, G., de Mendonça, P. G., Brey, R., Rinder, M., Pfister, K., Spiegel, M., Doerrbecker, B., Pfeffer, M., & Dobler, G. (2010). Isolation and molecular characterization of a tick-borne encephalitis virus strain from a new tick-borne encephalitis focus with severe cases in Bavaria, Germany. *Ticks and Tick-Borne Diseases*, *1*(1), 44–51. https://doi.org/10.1016/j.ttbdis.2009.11.002

Labuda, M., Austyn, J. M., Zuffova, E., Kozuch, O., Fuchsberger, N., Lysy, J., & Nuttall, P.

A. (1996). Importance of localized skin infection in tick-borne encephalitis virus transmission. *Virology*. https://doi.org/10.1006/viro.1996.0261

Larsen, A. L., M, A. K. Ø., Rland, M. B. J. Ø., Andreassen, Å. S., Soleng, A., Vene, S., & Dudman, S. G. (2014). *Detection of specifi c IgG antibodies in blood donors and tick-borne encephalitis virus in ticks within a non-endemic area in southeast Norway*. *August 2013*, 1–4. https://doi.org/10.3109/00365548.2013.865140

Leal, É. de S., & Zanotto, P. M. de A. (2000). Viral diseases and human evolution. *Memórias Do Instituto Oswaldo Cruz*, *95*(suppl 1), 193–200. https://doi.org/10.1590/S0074-02762000000700033

Lee, J. M., Crooks, A. J., & Stephenson, J. R. (1989). The synthesis and maturation of a non-structural extracellular antigen from Tick-borne encephalitis virus and its relationship to the intracellular NS1 protein. *Journal of General Virology*. https://doi.org/10.1099/0022-1317-70-2-335

Li, Z., Malla, S., Shin, B., & Li, J. M. (2014). Battle against RNA oxidation: molecular mechanisms for reducing oxidized RNA to protect cells. *Wiley Interdisciplinary Reviews: RNA*, *5*(3), 335–346. https://doi.org/10.1002/wrna.1214

Liebig, K., Boelke, M., Grund, D., Schicht, S., Bestehorn-Willmann, M., Chitimia-Dobler, L., Dobler, G., Jung, K., & Becker, S. C. (2021). The stable matching problem in TBEV enzootic circulation: How important is the perfect tick-virus match? *Microorganisms*. https://doi.org/10.3390/microorganisms9010196

Mackay, I. M., Arden, K. E., & Nitsche, A. (2002). Real-time PCR in virology. *Nucleic Acids Research*, *30*(6), 1292–1305. https://doi.org/10.1093/nar/30.6.1292

Mangold, A. J., Bargues, M. D., & Mas-Coma, S. (1997). 18S rRNA gene sequences and phylogenetic relationships of European hard-tick species (Acari: Ixodidae). *Parasitology Research*, *84*(1), 31–37. https://doi.org/10.1007/s004360050352

Marston, D. A., McElhinney, L. M., Ellis, R. J., Horton, D. L., Wise, E. L., Leech, S. L., David, D., de Lamballerie, X., & Fooks, A. R. (2013). Next generation sequencing of viral RNA genomes. *BMC Genomics*, *14*(1). https://doi.org/10.1186/1471-2164-14-444

Martiniano, R., Garrison, E., Jones, E. R., Manica, A., & Durbin, R. (2019). Removing reference bias in ancient DNA data analysis by mapping to a sequence variation graph. *BioRxiv*, 1–18. https://doi.org/10.1101/782755

Mazeaud, C., Freppel, W., & Chatel-Chaix, L. (2018). The Multiples Fates of the Flavivirus RNA Genome During Pathogenesis. *Frontiers in Genetics*, *9*(December), 1–19. https://doi.org/10.3389/fgene.2018.00595

McNeely, J. A. (2021). Nature and COVID-19: The pandemic, the environment, and the way ahead. *Ambio*, *50*(4), 767–781. https://doi.org/10.1007/s13280-020-01447-0

Michelitsch, A., Wernike, K., Klaus, C., Dobler, G., & Beer, M. (2019). Exploring the Reservoir Hosts of Tick-Borne Encephalitis Virus. *Viruses*, *11*(7), 669. https://doi.org/10.3390/v11070669

Nanda, S., Jayan, G., Voulgaropoulou, F., Sierra-Honigmann, A. M., Uhlenhaut, C., McWatters, B. J. P., Patel, A., & Krause, P. R. (2008). Universal virus detection by degenerate-oligonucleotide primed polymerase chain reaction of purified viral nucleic acids. *Journal of Virological Methods*, *152*(1–2), 18–24. https://doi.org/10.1016/j.jviromet.2008.06.007

Nicholson, W. L., Sonenshine, D. E., Noden, B. H., & Brown, R. N. (2018). Ticks (ixodida). In *Medical and Veterinary Entomology*. https://doi.org/10.1016/B978-0-12-814043-7.00027-3

Paulsen, K. M., Das Neves, C. G., Granquist, E. G., Knut Madslien, |, Stuen, S., Benedikte, |, Pedersen, N., Vikse, | Rose, Rocchi, M., Laming, E., Stiasny, K., Åshild, |, & Andreassen, K. (2019). *Cervids as sentinel-species for tick-borne encephalitis virus in Norway-A serological study*. https://doi.org/10.1111/zph.12675

Paulsen, K. M., Lamsal, A., Bastakoti, S., Pettersson, J. H. O., Pedersen, B. N., Stiasny, K., Haglund, M., Smura, T., Vapalahti, O., Vikse, R., Alfsnes, K., & Andreassen, Å. K. (2021). High-throughput sequencing of two European strains of tick-borne encephalitis virus (TBEV), Hochosterwitz and 1993/783. *Ticks and Tick-Borne Diseases*, *12*(1), 101557. https://doi.org/10.1016/j.ttbdis.2020.101557

Paulsen, K., Vikse, R., Soleng, A., S. Edgar, K., Dudman, S., Sofie Wiklund, B., & K.

Andreassen, Å. (2020). TBE in Norway. In Gerhard Dobler, W. Erber, M. Broker, & H.-J. Schmitt (Eds.), *Tick-borne encephalitis - The Book* (5th ed.). https://doi.org/10.33442/978-981-14-0914-1_12b-24

Paulsen, Lamsal, A., Bastakoti, S., Pettersson, J. H. O., Pedersen, B. N., Stiasny, K., Haglund, M., Smura, T., Vapalahti, O., Vikse, R., Alfsnes, K., & Andreassen, Å. K. (2021). High-throughput sequencing of two European strains of tick-borne encephalitis virus (TBEV), Hochosterwitz and 1993/783. *Ticks and Tick-Borne Diseases*, *12*(1), 101557. https://doi.org/10.1016/j.ttbdis.2020.101557

Pfeffer, M., & Dobler, G. (2010). Emergence of zoonotic arboviruses by animal trade and migration. In *Parasites and Vectors*. https://doi.org/10.1186/1756-3305-3-35

Pfeifer, S. P. (2017). From next-generation resequencing reads to a high-quality variant data set. *Heredity*, *118*(2), 111–124. https://doi.org/10.1038/hdy.2016.102

Pierson, T. C., & Diamond, M. S. (2020). The continued threat of emerging flaviviruses. *Nature Microbiology*, *5*(6), 796–812. https://doi.org/10.1038/s41564-020-0714-0

Proutski, V., Gould, E. A., & Holmes, E. C. (1997). Secondary structure of the 3' untranslated region of flaviviruses: Similarities and differences. *Nucleic Acids Research*. https://doi.org/10.1093/nar/25.6.1194

Puchhammer-Stöckl, E., Kunz, C., Mandl, C. W., & Heinz, F. X. (1995). Identification of tick-borne encephalitis virus ribonucleic acid in tick suspensions and in clinical specimens by a reverse transcription-nested polymerase chain reaction assay. *Clinical and Diagnostic Virology*, *4*(4), 321–326. https://doi.org/10.1016/0928-0197(95)00022-4

Quick, J., Grubaugh, N. D., Pullan, S. T., Claro, I. M., Smith, A. D., Gangavarapu, K., Oliveira, G., Robles-Sikisaka, R., Rogers, T. F., Beutler, N. A., Burton, D. R., Lewis-Ximenez, L. L., de Jesus, J. G., Giovanetti, M., Hill, S. C., Black, A., Bedford, T., Carroll, M. W., Nunes, M., … Loman, N. J. (2017). Multiplex PCR method for MinION and Illumina sequencing of Zika and other virus genomes directly from clinical samples. *Nature Protocols*, *12*(6), 1261–1266. https://doi.org/10.1038/nprot.2017.066

Radhakrishnan, G. V., Cook, N. M., Bueno-Sancho, V., Lewis, C. M., Persoons, A., Mitiku, A. D., Heaton, M., Davey, P. E., Abeyo, B., Alemayehu, Y., Badebo, A., Barnett, M.,

Bryant, R., Chatelain, J., Chen, X., Dong, S., Henriksson, T., Holdgate, S., Justesen, A. F., … Saunders, D. G. O. (2019). MARPLE, a point-of-care, strain-level disease diagnostics and surveillance tool for complex fungal pathogens. *BMC Biology*, *17*(1), 65. https://doi.org/10.1186/s12915-019-0684-y

Roychoudhury, S., Das, A., Sengupta, P., Dutta, S., Roychoudhury, S., Choudhury, A. P., Ahmed, A. B. F., Bhattacharjee, S., & Slama, P. (2020). Viral Pandemics of the Last Four Decades: Pathophysiology, Health Impacts and Perspectives. *International Journal of Environmental Research and Public Health*, *17*(24), 9411. https://doi.org/10.3390/ijerph17249411

Rudenko, N., Golovchenko, M., Cihlářova, V., & Grubhoffer, L. (2004). Tick-borne encephalitis virus-specific RT-PCR - A rapid test for detection of the pathogen without viral RNA purification. *Acta Virologica*, *48*(3), 167–171.

Růžek, D., Gritsun, T. S., Forrester, N. L., Gould, E. A., Kopecký, J., Golovchenko, M., Rudenko, N., & Grubhoffer, L. (2008). Mutations in the NS2B and NS3 genes affect mouse neuroinvasiveness of a Western European field strain of tick-borne encephalitis virus. *Virology*. https://doi.org/10.1016/j.virol.2008.01.010

Schrader, C., Schielke, A., Ellerbroek, L., & Johne, R. (2012). PCR inhibitors - occurrence, properties and removal. *Journal of Applied Microbiology*, *113*(5), 1014–1026. https://doi.org/10.1111/j.1365-2672.2012.05384.x

Schrader, Christina, & Süss, J. (1999). A nested RT-PCR for the detection of tick-borne encephalitis virus (TBEV) in ticks in natural foci. *Zentralblatt Fur Bakteriologie*. https://doi.org/10.1016/S0934-8840(99)80069-3

Schwaiger, M., & Cassinotti, P. (2003). Development of a quantitative real-time RT-PCR assay with internal control for the laboratory detection of tick borne encephalitis virus (TBEV) RNA. *Journal of Clinical Virology*, *27*(2), 136–145. https://doi.org/10.1016/S1386-6532(02)00168-3

Sergeant. (2018). Epitools Epidemiological Calculators. *ESG*. http://epitools.ausvet.com.au

Smura, T., Tonteri, E., Jääskeläinen, A., von Troil, G., Kuivanen, S., Huitu, O., Kareinen, L., Uusitalo, J., Uusitalo, R., Hannila-Handelberg, T., Voutilainen, L., Nikkari, S., Sironen,

T., Sane, J., Castrén, J., & Vapalahti, O. (2019). Recent establishment of tick-borne encephalitis foci with distinct viral lineages in the Helsinki area, Finland. *Emerging Microbes and Infections*, *8*(1), 675–683. https://doi.org/10.1080/22221751.2019.1612279

Süss, J. (2011). Tick-borne encephalitis 2010: Epidemiology, risk areas, and virus strains in Europe and Asia—An overview. *Ticks and Tick-Borne Diseases*, *2*(1), 2–15. https://doi.org/10.1016/j.ttbdis.2010.10.007

Tamura, K., Stecher, G., & Kumar, S. (2021). MEGA11: Molecular Evolutionary Genetics Analysis Version 11. *Molecular Biology and Evolution*, *38*(7), 3022–3027. https://doi.org/10.1093/molbev/msab120

Thortveit, E. T., Aase, A., Petersen, L. B., Lorentzen, Å. R., Mygland, Å., & Ljøstad, U. (2020). Human seroprevalence of antibodies to tick-borne microbes in southern Norway. *Ticks and Tick-Borne Diseases*, *11*(4), 101410. https://doi.org/10.1016/j.ttbdis.2020.101410

Tkachev, S. E., Babkin, I. V., Chicherina, G. S., Kozlova, I. V., Verkhozina, M. M., Demina, T. V., Lisak, O. V., Doroshchenko, E. K., Dzhioev, Y. P., Suntsova, O. V., Belokopytova, P. S., Tikunov, A. Y., Savinova, Y. S., Paramonov, A. I., Glupov, V. V., Zlobin, V. I., & Tikunova, N. V. (2020). Genetic diversity and geographical distribution of the Siberian subtype of the tick-borne encephalitis virus. *Ticks and Tick-Borne Diseases*, *11*(2), 101327. https://doi.org/10.1016/j.ttbdis.2019.101327

Traavik, T. (2009). Antibodies to Tick-Borne Encephalitis Virus in Human Sera from the Western Coast of Norway. *Acta Pathologica Microbiologica Scandinavica Section B Microbiology*, *87B*(1–6), 9–13. https://doi.org/10.1111/j.1699-0463.1979.tb02396.x

Uchil, P. D., & Satchidanandam, V. (2003). Architecture of the Flaviviral Replication Complex. *Journal of Biological Chemistry*, *278*(27), 24388–24398. https://doi.org/10.1074/jbc.M301717200

Uzcátegui, N. Y., Sironen, T., Golovljova, I., Jääskeläinen, A. E., Välimaa, H., Lundkvist, Å., Plyusnin, A., Vaheri, A., & Vapalahti, O. (2012). Rate of evolution and molecular epidemiology of tick-borne encephalitis virus in Europe, including two isolations from the same focus 44 years apart. *Journal of General Virology*, *93*(4), 786–796.

https://doi.org/10.1099/vir.0.035766-0

Varlacher, J.-F., HÄGGLUND, S., JUREMALM, M., BLOMQVIST, G., RENSTRÖM, L., ZOHARI, S., LEIJON, M., & CHIRICO, J. (2015). Tick-borne encephalitis. *Revue Scientifique et Technique de l'OIE*, *34*(2), 453–466. https://doi.org/10.20506/rst.34.2.2371

Veje, M., Studahl, M., Norberg, P., Roth, A., Möbius, U., Brink, M., & Bergström, T. (2014). Detection of Tick-Borne encephalitis virus RNA in urine. *Journal of Clinical Microbiology*. https://doi.org/10.1128/JCM.02428-14

Vikse, R., Paulsen, K. M., Edgar, K. S., H.-O. Pettersson, J., Ottesen, P. S., Okbaldet, Y. B., Kiran, N., Lamsal, A., Lindstedt, H. E. H., Pedersen, B. N., Soleng, A., & Andreassen, Å. K. (2020). Geographical distribution and prevalence of tick-borne encephalitis virus in questing Ixodes ricinus ticks and phylogeographic structure of the Ixodes ricinus vector in Norway. *Zoonoses and Public Health*, *67*(4), 370–381. https://doi.org/10.1111/zph.12696

Wang, Y., Zhao, Y., Bollas, A., Wang, Y., & Au, K. F. (2021). Nanopore sequencing technology, bioinformatics and applications. *Nature Biotechnology*, *39*(11), 1348–1365. https://doi.org/10.1038/s41587-021-01108-x

Wick, R. (2017). *Porechop* (0.2.4). GitHub repository. https://github.com/rrwick/Porechop

Xu, Y., Lewandowski, K., Lumley, S., Pullan, S., Vipond, R., Carroll, M., Foster, D., Matthews, P. C., Peto, T., & Crook, D. (2018). Detection of Viral Pathogens With Multiplex Nanopore MinION Sequencing: Be Careful With Cross-Talk. *Frontiers in Microbiology*, *9*(SEP), 1–7. https://doi.org/10.3389/fmicb.2018.02225

Ytrehus, B., Vainio, K., Dudman, S. G., Gilray, J., & Willoughby, K. (2013). Tick-Borne Encephalitis Virus and Louping-Ill Virus May Co-Circulate in Southern Norway. *Vector-Borne and Zoonotic Diseases*, *13*(10), 762–768. https://doi.org/10.1089/vbz.2012.1023

Yun, S.-M., Kim, S. Y., Ju, Y. R., Han, M. G., Jeong, Y. E., & Ryou, J. (2011). First complete genomic characterization of two tick-borne encephalitis virus isolates obtained from wild rodents in South Korea. *Virus Genes*, *42*(3), 307–316. https://doi.org/10.1007/s11262-011-0575-y

Zhao, W., He, X., Hoadley, K. A., Parker, J. S., Hayes, D. N., & Perou, C. M. (2014). Comparison of RNA-Seq by poly (A) capture, ribosomal RNA depletion, and DNA microarray for expression profiling. *BMC Genomics*, *15*(1), 419. https://doi.org/10.1186/1471-2164-15-419

7. Annexure I

A.1. High-Capacity cDNA reverse transcription kit protocol

Reverse transcription was conducted using the High-Capacity cDNA reverse transcription kit (Applied Biosystems™, Thermo Fisher Scientific, Cat. number: 4368813), according to the following procedure given by the manufacturer. The total volume of the individual RT-PCR reaction was maintained at 20.0 µL by combining, 10X RT random hexamers (2.0 µL), 10X RT buffer (2.0 µL), 25X dNTP mix (0.80 µL, 100 mM), MultiScribe™ Reverse Transcriptase (1.0 µL, 50 U/µL), RNase inhibitor (1.0 µL, 40 U/µL), nuclease-free H_2O (8.20 µL).

A.2. Semiquantitative real-time PCR method I

Semiquantitative real-time PCR was carried out according to the procedure already established method by Andreassen *et al.,* 2012. TaqMan probe was designed with 5'-Fluorescein amidite (6-FAM™) labeled -3'-minor groove binding protein (MGB) anchored (TBE339-probe, FAM-5'-AACGCAGAAAGAC-3'-MGB), to target the TBEV envelope protein coding region of the RNA [1647-1659 nt (reference to the TBEV genome, Acc. No. KF991107, NCBI)]. Semiquantitative real-time PCR reaction was carried out by adding the cDNA (3.0 µL) the reaction mixture (22.0 µL) composed with, TBE 339 probe (0.30 µL, 25 µM), Platinum-Taq™ DNA polymerase (0.19 µL, 5U/µL, Invitrogen™, Thermo Fisher Scientific, Cat. number: 15966005), $MgCl_2$ (2.50 µL, 50 mM), dNTPs (0.20 µL, 25 mM), forward primer [TBE320F, 0.25 µL, 25 µM, 5'-GGGAGCGCAAAACTGGAA-3', Acc.no (NCBI): KF991107: 1628-1645 nt], reverse primer (Anchored with biotin at the 3' end) [TBE373R, 0.25 µL 25 µM, 5'-TGAGGAGCCCCAAATTCAAC-3', Acc.no (NCBI): KF991107: 1681-1662 nt], AB buffer {10X, 2.50 µL, [HEPES potassium salt (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid), 150 mM), KCl (1.1 M) $MgCl_2$ (50.0 mM), EDTA (Ethylenediaminetetraacetic acid , 1.0 mM), DTT (Dithiothreitol, 25.0 mM), Sodium metabisulfite (12.5 mM), PMSF (Phenylmethanesulfonylfluoride, 2.5 mM) and Benzamidine (12.5 mM)]} and RNase-free H_2O (16.06 µL).

A.3 Semiquantitative real-time PCR method II

The semi-quantitative RT-qPCR developed by Ryan Easterday and Kristian Alfsnes was separately carried out for each total RNA extracts of nymph ticks (76), total RNA extracts of female ticks (46) and total RNA extracts of male ticks (50), separately with individual reaction consists with total tick(s) RNA (1.0 µL), TaqPath™ 1-Step RT-qPCR Master Mix, CG (2.50 µL, Cat. No. A15299, applied biosystems®, life technologies™, Thermo Fisher Scientific), forward primer [TBE320F, 0.60 µL, 25 µM, 5'-GGGAGCGCAAAACTGGAA-3', Acc.no (NCBI): KF991107: 1628-1645 nt], reverse primer [TBE373R, 0.60 µL 25 µM, 5'-TGAGGAGCCCCAAATTCAAC-3', Acc.no (NCBI): KF991107: 1681-1662 nt], probe [0.13 µL, TaqMan probe was designed with 5'-Fluorescein amidite (6-FAM™) labeled-3'-minor groove binding protein-eclips (MGB-eclips) anchored (TBE339-probe, FAM-5'-AACGCAGAAAGAC-3'-MGB-eclips), to target the TBEV envelope protein coding region of the RNA [1647-1659 nt (reference to the TBEV genome, Acc. No. KF991107, NCBI)] and the PCR grade water (5.17 µL). The actual real-time PCR was carried out on the Corbett Rotor-gene™ 6000 (Qiagen Corbett, Hilden, Germany) with the initial reverse transcription related random hexamer primer annealing at 25 °C for 10 minutes, the reverse transcription and the cDNA synthesis at 50 °C for 15 minute, the inactivation of Reverse transcriptase, denaturation of DNA and the TaqPath™ enzyme activation at 95 °C for 2 minutes and followed by 40 cycles: 95 °C for 3 seconds, 60 °C for 30 seconds for the denaturation, annealing of the primers and for the extension; at the end of the cycles storing at 4 °C. Fluorescent data acquired by the instrument software Rotor-Gene Q Software (Version: 2.3.1.49), at the very beginning and also at each cycle in the primer annealing and the extension stages were used to determine the cycle threshold (Ct) value of each sample by keeping the normalized fluorescence threshold cutoff value (0.02).

A.4 RT-PCR protocol for SuperScript™ IV First-Strand synthesis system

RT-PCR using SuperScript™ IV First-Strand synthesis system, (Thermo Fisher Scientific, Cat. number: 18091050) procedure for the initial primer annealing was separately carried out by heating each template RNA (1.0 µL) with the 1st mixture: [DEPC-treated water (10.0 µL), dNTP mix (1.0 µL, 10 mM), 5X SSIV buffer (4.0 µL), random hexamers (1.0 µL, 50 ng/ µL) /or/ template specific primers (1.0 µL, 10 µM) at 65 °C for 5 min and followed by 1 min incubation on ice. The reverse transcription reaction was carried out separately for each sample by adding the 2nd mixture: [5X SSIV buffer (4.0 µL), DTT (1.0 µL, 100 mM), Ribonuclease inhibitor (1.0 µL, 40 U/µL) and SuperScript™ IV reverse transcriptase (1.0 µL, 200 U/µL)]. Thermal cycler profile was maintained for the cDNA synthesis by the initial primer annealing at 23 °C for 10 min and RT-polymerization at 55 °C for 10 min, followed by quenching at 80 °C for 10 min and finally holding at 4 °C in a thermal cycler [Thermal Cycler (2720), Thermo Fisher Scientific].

A.5 PCR reaction protocol

PCR amplification was carried out using Q5® High-Fidelity PCR kit (New England BioLabs® Inc., Cat. No. E0555L). Total volume (25.0 µL) of each reaction was maintained by adding, nuclease free water (11.25 µL), 5X Q5 reaction buffer (5.0 µL), dNTPs (0.50 µL, 10 mM), corresponding forward primer (0.25 µL), corresponding reverse primer (0.25 µL), 5X Q5 High GC enhancer (5.0 µL), Q5 High-Fidelity DNA polymerase (0.25 µL, 100 U/µL) and the template cDNA (2.50 µL).

A.6 PCR products purification protocol

DNA products derived through the PCR amplifications were separately subjected to PCR products purifications using Agencourt® AMPure® XP beads (Product number: A63881, Beckman Coulter). PCR products purification was carried out by adding 60% v/v of homogenized AMPure® XP beads (0.6X based size cut clean-up) containing liquid to each PCR products containing PCR tubes and homogenized by gentle shaking and followed by incubation (15 min) to bind the DNA to the magnetic beads. After the incubation, each mixture was transferred to a 96 semi-skirted PCR plate (Eppendorf® Twin.tec microbiology PCR plate, 96 wells, Cat. number: EP0030129300) placed on a 96 well magnet (96S Super Magnet, Cat. number: A001322, Alpaqua®) using a multichannel pipette and left until to form a clear solution by separating out the magnetic beads from the solution. Once the clear solution is formed, supernatant was discarded using a multichannel micro pipette and supplemented with aqueous Ethanol (80%, v/v, 180 µL) followed by 30 S incubation on the magnetic tray. After the incubation, the supernatant was removed and supplemented with the same amount of aqueous Ethanol (80%, v/v) and followed by 30 S incubation and kept for the evaporation of the remaining amount of ethanol after pipetting out the supernatant. The DNA bound to magnetics beads containing 96 well PCR plate was taken out from the magnetic tray and supplemented with the elution buffer (Buffer EB, Qiagen, Cat. number: 19065) for each well, followed by incubation for 2 min with gentle mixing. Finally, the 96 well plate was placed on the 96 well magnet again and left to separate the magnetic beads from the solution and transferred the clear supernatant containing DNA to new PCR tubes for further analysis.

A.7 Flow-cell priming, DNA library preparation and Nanopore sequencing protocol

The Oxford Nanopore flow cell (R9.4.1, Oxford Nanopore Technologies) was used for the nanopore sequencing with the GridION MK1 (GXB03008, Oxford Nanopore Technologies) instrument. First the flow cell was taken from the storage (4-8 °C) and mounted on the instrument port to equilibrate at 37 °C. After the temperature equilibration, using the MinKNOW software (Oxford Nanopore Technologies) the number of pores available for the sequencing were confirmed to have above the minimum requirements.

Library preparation was carried out according to the following procedure. Firs the reagents of Rapid Barcoding Sequencing kit: SQK-RBK004 (Oxford Nanopore Technologies) were allowed to equilibrate at room temperature and spun down. Purified DNA (7.50 μL) from each selected 12 different samples were separately transferred to thin wall PCR tubes and each sample was supplemented with one fragmentation mix of 12 [2.50 μL, Fragmentation Mix RB 01-12, Rapid Barcoding Sequencing kit: SQK-RBK004 (Oxford Nanopore Technologies)], mixed gently, spun down and incubated at 30 °C for 1 minute with subsequent heating at 80 °C for 1 minute followed by cooling at 4 °C using the thermal cycler (T100™ Thermal cycler, Bio-Rad). The resulting barcoded fragments were combined with the volume (10.0 μL) from each into a new Eppendorf DNA LoBind tube, combined with the resuspended AMPure XP beads [120 μL, Rapid Barcoding Sequencing kit: SQK-RBK004 (Oxford Nanopore Technologies)] and incubated on a rotator mixture for 5 minutes at room temperature. After the incubation, sample containing PCR tube was transferred on to a magnetic tray and supernatant was pipetted out and washed the beads using freshly prepared aqueous ethanol (200 μL, 70% v/v) solution for two times and allowed to dry to evaporate any residual ethanol remained in the tube. The magnetic beads containing tube was taken out of the magnet and resuspended with Tris-HCl with NaCl [10.0 μL, Tris-HCl (10 mM), NaCl (50 mM), pH 7.5-8.0] [Rapid Barcoding Sequencing kit: SQK-RBK004 (Oxford Nanopore Technologies] and incubated (2 min.) at room temperature. After the incubation, sample containing PCR tube was kept on a magnet until it became a clear solution and transferred the supernatant to Eppendorf DNA LoBind tube. From the Rapid adapter [12.0 μL, Rapid Barcoding Sequencing kit: SQK-RBK004 (Oxford Nanopore Technologies)] was added to the barcoded DNA and mixed by pipetting up and down and finally spun down and incubated for 5 min at room temperature. After the priming of the flow cell according to the procedure given by the manufacturer, above prepared DNA library (11.0 μL) was transferred into a new PCR tube and mixed with sequencing buffer (34.0 μL), homogenized loading Beads (25.5 μL) and nuclease free water (4.50 μL). Finally, the prepared library (75.0 μL) was transferred into the sample port dropwise manner and closed the port with the cover. Using the MinKNOW software and the protocol, data acquisition was carried out for 72 hours in a GridION MK1 device.

A.8 Total RNA extraction by RNeasy mini kit

Total RNA extraction of tick samples was carried out using the RNeasy mini kit (Cat. No. 74106, QIAGEN, Hilden, Germany) according to the manufacture's protocol. The main steps of the protocol are shown in the figure A.1. The complete protocol is available at https://www.qiagen.com/us/resources/resourcedetail?id=14e7cf6e-521a-4cf7-8cbc-bf9f6fa33e24&lang=en (viewed: 11/09/2022).

Figure A.1. Total RNA extraction procedure using RNeasy mini kit (Cat. No. 74106, QIAGEN, Hilden, Germany).

Table A.1. Primers used for synthesis of cDNA from seven different total RNA extracts of cell culture samples incubated with TBEV-Hochosterwitz strain and three NTC

Sample number	RNA extract	Primer utilized in each cDNA synthesis (Separately)				
		Random hexamers#	JK1-F#	JK7-R#	Eur_39-R#	A mixture of primers*#
1	Hochosterwitz_Trizol_11-04-2016	+	+	+	+	+
2	Hochosterwitz_Qiagen_11-04-2016	+	+	+	+	+
3	Hochosterwitz_RLT_28-04-2016	+	+	+	+	+
4	Hochosterwitz_250416_26-04-2016	+	+	+	+	+
5	Hochosterwitz_2_AVL_28-04-2016	+	+	+	+	+
6	Hochosterwitz_extracted_05-04-2016_isolated_17-06-2021_MAJD	+	+	+	+	+
7	Hochosterwitz_2_extracted_25-04-2016-isolated-26-04_2016	+	+	+	+	+
8	NTC	+	+	+	+	+
9	NTC	+	+	+	+	+
10	RNA extract from Hela cell culture* (Negative control)	+	+	+	+	+

** A mixture of primers: composition [JK1-F (3.3 µM), JK7-R (3.3 µM), Eur-39-R (3.3 µM)], # Primer sequences and their properties are listed in the Table 1 in the Annexure I . * Provided with the SuperScript™ IV First-Strand synthesis system, Thermo Fisher Scientific, Catalog number: 18091050.*

Table A.2. Primers, primer schemes and their physical properties

Primer Name	Sequence	Primer scheme and the composition								Length (b)	Tm* (°C)	GC (%)	Primer position#		Sequence identity##
		A	B	C	D	E	F	G	H				Start	End	
Eur_1-F	AGCATTAGCAGCGGTTGGTTTG	X				X		2X		22	69.0	50.0	37	59	21
Eur_1-R	CCAACCATTAGGGCACTCACTG	X				X		2X		22	69.0	54.6	378	400	100
Eur_2-F	CGAGAAACCCCGTATTGAAGGC		X		X		X		2X	22	69.0	54.6	297	319	95
Eur_2-R	ACGTCAACAGGCTCTTCTCCTT		X		X		X		2X	22	69.0	50.0	648	670	100
Eur_3-F	GAATGGCACCTGTGTGATCCTG	X								22	69.0	54.6	565	587	100
Eur_3-R	AACACAGGAGCACAACCAGAAC	X								22	69.0	50.0	938	960	100
Eur_4-F	GGATGGGTTTGGAAGAACAGGC		X							22	69.0	54.6	851	873	90
Eur_4-R	TGGTGTTCTTCAGCCAAAGTGG		X							22	69.0	50.0	1215	1237	100
Eur_5-F	GCTTGACGCCATTTACCAGGAG	X								22	68.0	54.6	1108	1130	95
Eur_5-R	AACTGTGAAGGATGCCGTCTTC	X								22	68.0	50.0	1456	1478	100
Eur_6-F	ACGACGCCAACAAAATAGTGTACA		X						2X	24	67.0	41.7	1371	1395	95
Eur_6-R	TCTGGTCTCCGAGGTTGTACAC		X		X		X		2X	22	67.0	54.6	1736	1758	100
Eur_7-F	GAAACATGAGGGAGCGCGAAAC	X		X		X		2X		22	69.0	54.6	1657	1679	95
Eur_7-R	ACATCTGGAGATCCATGTGCCA	X		X		X		2X		22	69.0	50.0	2011	2033	100
Eur_8-F	AGAGAGCTCCAACAGACAGTGG		X		X		X		X	22	70.0	54.6	1921	1943	95
Eur_8-R	CCCCCGAAGATGCTGTTGAAAG		X		X		X		X	22	70.0	54.6	2308	2330	100
Eur_9-F	ACTGACAGTGATAGGGGAGCAC	X		X		X		X		22	69.0	54.6	2213	2235	95
Eur_9-R	GTGTCTCCGGGTAGTAGGCATA	X		X		X		X		22	69.0	54.6	2559	2581	100
Eur_10-F	GTTTTGGCCATGACCCTTGGAG		X		X		X		X	22	70.0	54.6	2436	2458	90
Eur_10-R	GCTCCAGATCATTGAATGGCCC		X		X		X		X	22	70.0	54.6	2822	2844	100
Eur_11-F	TTTGACCCCACTGACTACCGAG	X		X						22	69.0	54.6	2742	2764	100
Eur_11-R	ATGAGCAGTTCCTCAGGTCAGT	X		X						22	69.0	50.0	3121	3143	100

Table A.2. Primers, primer schemes and their physical properties cont..

Primer Name	Sequence	Primer scheme and the composition								Length (b)	Tm* (°C)	GC (%)	Primer position#		Sequence identity##
		A	B	C	D	E	F	G	H				Start	End	
Eur_12-F	GCATGGCAATCCACACAGATCA		X		X		X		2X	22	69.0	50.0	3037	3059	100
Eur_12-R	GAACGTCACTGGTGGCATTGTG		X		X		X		2X	22	69.0	54.6	3417	3439	100
Eur_13-F	CGCCAAGTGTGACAAAAGAGGA	X		X		X		X		22	69.0	50.0	3333	3355	100
Eur_13-R	AGTAGCATCAAGGCCACGATCT	X		X		X		X		22	69.0	50.0	3743	3765	95
Eur_14-F	TCGTTCTCGCTTTGCTTGTCAC		X		X					22	69.0	50.0	3644	3666	86
Eur_14-R	GATGCACAGTCACCACATCCTG		X		X					22	69.0	54.6	4012	4034	95
Eur_15-F	GCTGATATTCAGGGCTTGCACG	X						X		22	69.0	54.6	3939	3961	95
Eur_15-R	CATCTTTCTTGTCCCCAGCACC	X		X		X		X		22	69.0	54.6	4335	4357	100
Eur_16-F	GACATTGGCTAGCGGCATGATG		X		X		X		2X	22	69.0	54.6	4251	4273	86
Eur_16-R	AGACACCGTCCTTAACCTCGAAA		X		X		X		2X	23	69.0	47.8	4656	4679	100
Eur_17-F	GACTGTGGACGCTCACGAAAAT	X		X						22	69.0	50.0	4565	4587	95
Eur_17-R	GTATTGCCCCAAGCTTCCTTCC	X		X						22	69.0	54.6	4972	4994	95
Eur_18-F	GTACAGGTTCATGCCTTCCCAC		X						2X	22	69.0	54.6	4894	4916	95
Eur_18-R	TCAATGCATTGGCGAATGAGCT		X		X		X		2X	22	69.0	45.4	5240	5262	100
Eur_19-F	GGCTGGACATCAAAGGGTCAGA	X		X		X		2X		22	70.0	54.6	5164	5186	100
Eur_19-R	CAGGAGGTGTCGCTGTCATCAA	X		X		X		2X		22	70.0	54.6	5561	5583	100
Eur_20-F	TGGACTGACCCTCACAGCATAG		X							22	69.0	54.6	5486	5508	81
Eur_20-R	CAGTCGTCACCACAAAGTCAGG		X							22	69.0	54.6	5828	5850	95
Eur_21-F	GCTCGCACCTTGAGACAGAAAG	X						X		22	69.0	54.6	5738	5760	90
Eur_21-R	GCATCTTGTCCTGTTCTGGTCC	X						X		22	69.0	54.6	6140	6162	100
Eur_22-F	GGATTAGTGCAATGGAAGGAGGC		X							23	69.0	52.2	6062	6085	91
Eur_22-R	ACGCCACGAACTCTTTGATGTC		X							22	69.0	50.0	6431	6453	100
Eur_23-F	GATCGAAGCTGGACATGGGAAG	X		X		X		X		22	69.0	54.6	6287	6309	100
Eur_23-R	TGATTGAAGTCCGGACGACGAA	X		X		X		X		22	69.0	50.0	6681	6703	95

Table A.2. Primers, primer schemes and their physical properties cont..

Primer Name	Sequence	Primer scheme and the composition								Length (b)	Tm* (°C)	GC (%)	Primer position#		Sequence identity##
		A	B	C	D	E	F	G	H				Start	End	
Eur_24-F	GCAATGAGAATGGCGGAGAGAG		X		X		X		2X	22	69.0	54.6	6585	6607	95
Eur_24-R	GTCTGCCTTGGTCTTCTCCAGA		X		X		X		2X	22	69.0	54.6	6932	6954	100
Eur_25-F	GCAGTGATGACAACAAACTGGC	X		X		X		X		22	68.0	50.0	6850	6872	95
Eur_25-R	TAGCCCCAATCAGTGACACCAC	X		X		X		X		22	68.0	54.6	7212	7234	95
Eur_26-F	AACAACTTGTCAACAGTGCCGT		X		X		X		2X	22	68.0	45.4	7108	7130	95
Eur_26-R	TCTGTTATGGAGGCCACCGTTC		X		X		X		2X	22	68.0	54.6	7483	7505	100
Eur_27-F	TGGTGGATGGGGATGTCATCAA	X		X		X		X		22	69.0	50.0	7363	7385	100
Eur_27-R	AGAATTCCTCCCTGGTGCAGTT	X		X		X		X		22	69.0	50.0	7725	7747	100
Eur_28-F	AGACTCTGGCTTCGAGCTTCTG		X		X		X		2X	22	69.0	54.6	7641	7663	100
Eur_28-R	CAAGTTCCAACCCAGGCTTGTT		X		X		X		2X	22	69.0	50.0	8021	8043	100
Eur_29-F	GCTGGTCCTATTATGCGGCATC	X		X						22	69.0	54.6	7930	7952	100
Eur_29-R	TTCCCAGTGACAGCTGTTGAGT	X		X						22	69.0	50.0	8338	8360	100
Eur_30-F	CACTGCACAGATTCCAACTGCA		X		X		X		2X	22	69.0	50.0	8257	8279	100
Eur_30-R	ATGGCCAGCTGAGAAGTTTCAC		X		X		X		2X	22	69.0	50.0	8649	8671	100
Eur_31-F	CGTACCGGACATGGCAGTATTG	X		X		X		X		22	69.0	54.6	8572	8594	90
Eur_31-R	TCCTCTACAGCCTCTCTTGCAC	X		X		X		X		22	69.0	54.6	8938	8960	100
Eur_32-F	GCATGTGCAGCAGAGAGGAATT		X		X		X		2X	22	68.0	50.0	8851	8873	95
Eur_32-R	AGAAGAGTCCTCCATTCAGGGT		X		X		X		2X	22	68.0	50.0	9246	9268	100
Eur_33-F	TTGGGCCTCTAGAGAGTCCAGT	X		X						22	70.0	54.6	9164	9186	95
Eur_33-R	CCTTCCCCTTCCATCATTCGGA	X		X						22	70.0	54.6	9538	9560	95
Eur_34-F	AGAGATCAAAGAGGTTCGGGCC		X							22	69.0	54.6	9465	9487	95
Eur_34-R	TCCGTCCTTCATCACTAGCTCG		X							22	69.0	54.6	9818	9840	100
Eur_35-F	ACCAGGAAGGACATTGGGGAAT	X		X		X		2X		22	69.0	50.0	9735	9757	100
Eur_35-R	TCCAGCATGTCTTCTGTGGTCA	X		X		X		2X		22	69.0	50.0	10087	10109	100

Table A.2. Primers, primer schemes and their physical properties cont..

Primer Name	Sequence	Primer scheme and the composition								Length (b)	Tm* (°C)	GC (%)	Primer position#		Sequence identity##
		A	B	C	D	E	F	G	H				Start	End	
Eur_36-F	CATCAACTCAGCAGTGCCTGTC		X							22	69.0	54.6	10010	10032	100
Eur_36-R	TTGCTCTGTCACAGTCTGGGTT		X							22	69.0	50.0	10386	10408	100
Eur_37-F	GGGCCTGAAAAGTTCAAGGACT	X								22	68.0	50.0	10299	10321	95
Eur_37-R	GCCGCATCATTCCATGATCTGT	X								22	68.0	50.0	10743	10765	95
Eur_38-F	ACATGACAAAGTAAAGAGGCTGAGC		X							25	68.0	44.0	10635	10660	
Eur_38-R	ATTTCTCTCTTCCCTCCTCCCG		X							22	68.0	54.6	11064	11086	
Eur_39-F	AGCAATACTTCGTGAGACCCCC	X								22	68.0	54.6	10815	10837	
Eur_39-R	ACCTCCTTGTCAGACTATCTGTGT	X								24	68.0	45.8	11183	11207	

Primer Name	Sequence	Primer scheme and the composition		Length (b)	Tm* (°C)	GC (%)	Primer position#		Sequence identity##
		M	N				Start	End	
JK_1-F	AGCATTAGCAGCGGTTGGTTTG	X		22	70.0	50.0	37	59	63
JK_1-R	GACTGGGATCCTACAGGGCTTT	X		22	70.0	54.2	1982	2004	95
JK_2-F	CGGAGACCAGACTGGAGTGTTA		X	22	69.0	54.5	1747	1769	100
JK_2-R	AACACAGCCTGGAGTAGCATCA		X	22	69.0	22.0	3755	3777	100
JK_3-F	TTGCGGACAACGGTGAATTACT	X		22	68.0	45.5	3518	3540	95
JK_3-R	GAACCTGACCCGTTTCCCATTC	X		22	68.0	54.5	5329	5351	100
JK_4-F	ATGAGACCTACGTCAGCAGCAT		X	22	69.0	50.0	5084	5106	100
JK_4-R	CATCTCATTGGCTGCAACCAGT		X	22	69.0	50.0	6905	6927	100
JK_5-F	CTTCGTCGTCCGGACTTCAATC	X		22	69.0	54.5	6680	6702	95
JK_5-R	GGCCAGCTGAGAAGTTTCACAA	X		22	69.0	50.0	8647	8669	100
JK_6-F	ACTTTTGGCTCGGTTTGGAGAC		X	22	69.0	50.0	8393	8415	100
JK_6-R	CCCAGATGTTCTTGGCCCATTC		X	22	69.0	54.5	10248	10270	100
JK_7-F	GCACAAACAATTGGCAACCACA	X		22	68.0	45.5	9356	9378	100
JK_7-R	ATTTCTCTCTTCCCTCCTCCCG	X		22	68.0	54.5	11064	11086	

Table A.2. Primers, primer schemes and their physical properties cont..

Primers designed to replace the poorly performing primers identified in this study														
Primer Name	**Sequence**	**Primer scheme and the composition**						**Length (b)**	**Tm* (°C)**	**GC (%)**	**Primer position#**		**Sequence identity##**
		O	**P**	**Q**	**R**	**S**	**T**				**Start**	**End**	
CJB_1-F	CAGCTTAGGAGAACAAGAGCTG	X		X		X		22	65.0	50.0	109	131	100
CJB_1-R	CCCCAACAGAGTAATGACCAGC	X		X		X		20	66.0	50.0	596	616	100
CJB_3-F	CCTATGAGTGTGTGACCATAGA	X		X		X		22	63.0	45.5	624	646	100
CJB_3-R	GTCAAGCCACACATCCATTGA	X		X		X		21	65.0	47.6	1094	1115	95
CJB_4-F	ACYATAACAGCTGAGGGGAAG		X		X		X	21	63-67	47.6	1070	1091	100
CJB_4-R	GTGGCTTTCTTTTTTGCCTCACA		X		X		X	23	66.0	43.5	1337	1360	100
CJB_5-F	CAGAGTGATCGAGGCTGGGG	X		X		X		20	70.0	65.0	1262	1282	100
CJB_5-R	GTGAAGGATGCCGTCTTCCT	X		X		X		20	67.0	55.0	1454	1474	100
CJB_6-F	GCAAAAAAGAAAGCCACAGGAC		X		X		2X	22	65.0	45.5	1343	1365	100
CJB_15-F	CATGGCTCTCATGACACAGC	X		X				20	66.0	55.0	3993	4013	100
CJB_17-F	GTGATGGGACTGTGGACGCT	X						20	70.0	60.0	4558	4578	100
CJB_17-R	TGAGCAATGCTGCTGACGTA	X						20	67.0	50.0	5092	5112	100
CJB_18-F	CAGGTCCATGCCTTCCCACC		X		X		2X	20	71.0	65.0	4897	4917	100
CJB_20-F	TGGGAGGTGGCAATCATGGA		X		X		X	20	69.0	55.0	5456	5476	100
CJB_20-R	CATCACACTGTCCAGAGTATATGT		X		X		X	24	63.0	41.7	6027	6051	95
CJB_21-F	AGAGTGAGGGATGAGAAGCC	X		X		2X		20	66.0	55.0	5810	5830	100
CJB_21-R	TGCTTTCTTTTCTCTTCAGTGAGTC	X		X		2X		25	65.0	40.0	6183	6208	100
CJB_22-F	GCTTTCTTTTCTCTTCAGTGAGTC		X		X		X	20	65.0	50.0	6142	6162	100
CJB_22-R	CGAACTCTTTGATGTCACGTCC		X		X		X	22	66.0	50.0	6425	6447	100
CJB_34-F	GGAGGCTGCATCATGGATGT		X		X		X	20	68.0	55.0	9435	9455	100
CJB_34-R	AAGTAGCTCAGCAGCCACAT		X					20	66.0	50.0	9957	9977	100
CJB_36-F	TGCCGAGACCAAGATGAACT		X		X		X	20	66.0	50.0	9858	9878	100
CJB_36-R	AGCTCCCAGTGCAGGTCATG		X		X		X	20	70.0	60.0	10342	10362	100
CJB_37-F	GGGCCAAGAACATCTGGGGA	X		X		X		20	70.0	60.0	10252	10272	100

Table A.2. Primers, primer schemes and their physical properties cont..

Primer Name	Sequence	Primer scheme and the composition						Length (b)	Tm* (°C)	GC (%)	Primer position#		Sequence identity##
		O	P	Q	R	S	T				Start	End	
CJB_37-R	ATGATGCAGGGGGTCTCACA	X						20	69.0	55.0	10992	11012	100
CJB_38-F	TCCCGACGTAGGGCACTCTG		X		X		X	20	72.0	65.0	10963	10983	100
CJB_38-R	AGGAGGAGGAAAAATCCTGAAG		X				X	22	64.0	45.5	11096	11118	100
CJB_39-F	TGCATCATGATAAGGCCGAAC	X		X		X		21	65.0	47.6	11004	11025	100
CJB_39-R	CTGTGTCTGGGTGATGGTGG	X		X		X		20	68.0	60.0	11168	11188	100
CJB_11-F	CTCGAGATGGCCATGTGGAG			X		X		20	68	60	2653	2672	100
CJB_11-R	GCAGCTCCCATCACTCCTGT			X		X		20	70	60	3007	3026	100
CJB_14-F	CTTGTTCGCTCAATGGTGGT				X		X	20	66	50	3500	3519	100
CJB_14-R	GCTGTGTCATGAGAGCCAT				X		X	19	65	53	3995	4014	100
CJB_17-NF	GGTGGAGAGGTTAGCCTGCG			X		X		20	71	65	4415	4434	100
CJB_17-NR	GGTGGGAAGGCATGGACCTG			X		X		20	71	65	4898	4917	100
CJB_29-F	TGCTCAGAAGAGGAGAGACCAA			X		X		22	67	60	7802	7823	100
CJB_29-R	CAGTTGGAATCTGTGCAGTG			X		X		20	64	50	8258	8277	100
CJB_33-F	AACATGATGGGCAAGAGAGA			X		X		20	63	45	9034	9053	100
CJB_33-R	ACCTTTATGTTGGTGAGGGT			X		X		20	63	45	9511	9530	100
CJB_34-NR	TGCTCCCATTCCCCAATGTC				X		X	20	68	55	9745	9764	100
CJB_37-NR	CACCATGTTCGGCCTTATCATG			X		X		22	66	50	11010	11031	

** Calculation of melting temperature (Tm) was carried out using the online Tm calculator from the NEW ENGLAND BioLabs® inc., according to the Q5® High-Fidelity PCR kit (New England BioLabs® Inc., Cat. No. E0555L) requirements (https://tmcalculator.neb.com/#!/main); # Reference to the consensus sequence derived from the multiple sequence alignment of 249 different TBEV strains obtained from NCBI nucleotide archive; ## Sequence identity reference to the Hochosterwitz stain (NCBI. Acc. No: MT311861); Blank (" "): Impossible to calculate the primer identity, due to the lack of corresponding position in the reference genome sequence of TBEV Hochosterwitz stain (NCBI. Acc. No: MT311861);"X": Included into the corresponding primer scheme with an average volume (1.0 µL). JK primer sequences obtained from the previous study carried out at NIPH,* (Paulsen et al., 2021).

Table A.3. Primer compositions, PCR conditions maintained for the DNA library developments and Oxford Nanopore sequencing

Sequencing experiment number	DNA library Barcode number	Template RNA volume (µL)	RT-PCR primer composition	Template cDNA volume (µL) and the dilution (-log10)	PCR primer extension temperature (°C)	Primer scheme and the concentration (X = 0.5 µM, 1.25 µL; 2X = 1.0 µM, 2.50 µL)				
						A	B	M	N	Other primers
1	1	1.00	Rand-Hex	2.50 (4)	65.0			X		
1	2	1.00	Rand-Hex	2.50 (4)	65.0				X	
1	3	1.00	Rand-Hex	2.50 (4)	65.0	X				
1	4	1.00	Rand-Hex	2.50 (4)	65.0		X			
1	5	1.00	Rand-Hex	2.50 (6)	65.0	X				
1	6	1.00	Rand-Hex	2.50 (6)	65.0		X			
1	7	1.00	Rand-Hex	2.50 (7)	65.0	X				
1	8	1.00	Rand-Hex	2.50 (7)	65.0		X			
1	9	1.00	Rand-Hex	2.50 (4)	65.0					JK_1-F/R
1	10	1.00	Rand-Hex	2.50 (4)	65.0					Eur_1-F/R
1	11	1.00	Rand-Hex	2.50 (4)	65.0					JK_1-F/R
1	12	1.00	Rand-Hex	2.50 (4)	65.0					Eur_1-F/R

Table A.3. Primer compositions, PCR conditions maintained for the DNA library developments and Oxford Nanopore sequencing cont.

Sequencing experiment number	DNA library Barcode number	Template RNA volume (µL)	RT-PCR primer composition	Template cDNA volume (µL) and the dilution (-log10)	PCR primer extension temperature (°C)	Primer scheme and the concentration (X = 0.5 µM, 1.25 µL; 2X = 1.0 µM, 2.50 µL)					
						A	B	C	D	O	P
2	1	1.00	Rand-Hex	2.50 (4)	65.0	X				X	
2	1	1.00	Rand-Hex	2.50 (4)	65.0		X				X
2	2	1.00	Rand-Hex	2.50 (5)	65.0	X				X	
2	2	1.00	Rand-Hex	2.50 (5)	65.0		X				X
2	3	1.00	Rand-Hex	2.50 (6)	65.0	X				X	
2	3	1.00	Rand-Hex	2.50 (6)	65.0		X				X
2	4	1.00	Rand-Hex	2.50 (4)	65.0			X		X	
2	4	1.00	Rand-Hex	2.50 (4)	65.0				X		X
2	5	1.00	Rand-Hex	2.50 (5)	65.0			X		X	
2	5	1.00	Rand-Hex	2.50 (5)	65.0				X		X
2	6	1.00	Rand-Hex	2.50 (6)	65.0			X		X	
2	6	1.00	Rand-Hex	2.50 (6)	65.0				X		X
2	7	1.00	Rand-Hex	2.50 (4)	65.0	2X				X	
2	7	1.00	Rand-Hex	2.50 (4)	65.0		2X				X
2	8	1.00	Rand-Hex	2.50 (5)	65.0	2X				X	
2	8	1.00	Rand-Hex	2.50 (5)	65.0		2X				X
2	9	1.00	Rand-Hex	2.50 (6)	65.0	2X				X	
2	9	1.00	Rand-Hex	2.50 (6)	65.0		2X				X
2	10	1.00	Rand-Hex	2.50 (4)	65.0	X					
2	10	1.00	Rand-Hex	2.50 (4)	65.0		X				
2	11	1.00	Rand-Hex	2.50 (5)	65.0	X					
2	11	1.00	Rand-Hex	2.50 (5)	65.0		X				
2	12	1.00	Rand-Hex	2.50 (6)	65.0	X					
2	12	1.00	Rand-Hex	2.50 (6)	65.0		X				

Table A.3. Primer compositions, PCR conditions maintained for the DNA library developments and Oxford Nanopore sequencing cont.

Sequencing experiment number	DNA library Barcode number	Template RNA volume (µL)	RT-PCR primer composition	Template cDNA volume (µL) and the dilution (-log10)	PCR primer extension temperature (°C)	Primer scheme and the concentration (X = 0.5 µM, 1.25 µL; 2X = 1.0 µM, 2.50 µL)												
						C	D	E	F	G	H	O	P	Q	R	S	T	Other primers
3	1	1.00	Rand-Hex	2.50 (5)	63.0	X						X						
3	1	1.00	Rand-Hex	2.50 (5)	63.0		X						X					
3	2	1.00	Rand-Hex	2.50 (6)	63.0	X						X						
3	2	1.00	Rand-Hex	2.50 (6)	63.0		X						X					
3	3	1.00	Rand-Hex	2.50 (5)	63.0			2X						X				
3	3	1.00	Rand-Hex	2.50 (5)	63.0				2X						X			
3	4	1.00	Rand-Hex	2.50 (6)	63.0			2X						X				
3	4	1.00	Rand-Hex	2.50 (6)	63.0				2X						X			
3	5	1.00	Rand-Hex	2.50 (5)	63.0			2X						2X				
3	5	1.00	Rand-Hex	2.50 (5)	63.0				2X						2X			
3	6	1.00	Rand-Hex	2.50 (6)	63.0			2X						2X				
3	6	1.00	Rand-Hex	2.50 (6)	63.0				2X						2X			
3	7	1.00	Rand-Hex	2.50 (5)	63.0					X						X		
3	7	1.00	Rand-Hex	2.50 (5)	63.0						X						X	
3	8	1.00	Rand-Hex	2.50 (6)	63.0					X						X		
3	8	1.00	Rand-Hex	2.50 (6)	63.0						X						X	
3	9	1.00	Rand-Hex	2.50 (5)	63.0					2X						2X		
3	9	1.00	Rand-Hex	2.50 (5)	63.0						2X						2X	
3	10	1.00	Rand-Hex	2.50 (6)	63.0					2X						2X		
3	10	1.00	Rand-Hex	2.50 (6)	63.0						2X						2X	
3	11	1.00	Rand-Hex	2.50 (5)	63.0													CJB_37-F/R
3	11	1.00	Rand-Hex	2.50 (5)	63.0													CJB_38-F/R

Table A.3. Primer compositions, PCR conditions maintained for the DNA library developments and Oxford Nanopore sequencing cont.

Sequencing experiment number	**DNA library Barcode number**	**Template RNA volume (µL)**	**RT-PCR primer composition**	**Template cDNA volume (µL) and the dilution (-log10)**	**PCR primer extension temperature (°C)**	**Other primers**
3	11	1.00	Rand-Hex	2.50 (5)	63.0	CJB_39-F/R
3	12	1.00	Rand-Hex	2.50 (5)	65.0	(CJB_37-F + Eur_37-R)
3	12	1.00	Rand-Hex	2.50 (5)	63.0	(CJB_37-R + Eur_37-F)

Table A.3. Primer compositions, PCR conditions maintained for the DNA library developments and Oxford Nanopore sequencing cont.

Sequencing experiment number	DNA library Barcode number	Template RNA volume (µL)	RT-PCR primer composition	Template cDNA volume (µL)	PCR primer extension temperature (°C)	Primer scheme and the concentration (X = 0.5 µM, 1.25 µL; 2X = 1.0 µM, 2.50 µL)			
						G	H	S	T
4	1	1.00	Rand-Hex	2.50	63.0	X		X	
4	1	1.00	Rand-Hex	2.50	63.0		X		X
4	2	1.00	Rand-Hex	2.50	63.0	X		X	
4	2	1.00	Rand-Hex	2.50	63.0		X		X
4	3	1.00	Rand-Hex	2.50	63.0	X		X	
4	3	1.00	Rand-Hex	2.50	63.0		X		X
4	4	1.00	Rand-Hex	2.50	63.0	X		X	
4	4	1.00	Rand-Hex	2.50	63.0		X		X
4	5	1.00	Rand-Hex	2.50	63.0	X		X	
4	5	1.00	Rand-Hex	2.50	63.0		X		X
4	6	1.00	Rand-Hex	2.50	63.0	X		X	
4	6	1.00	Rand-Hex	2.50	63.0		X		X
4	7	1.00	Rand-Hex	2.50	63.0	X		X	
4	7	1.00	Rand-Hex	2.50	63.0		X		X
4	8	1.00	Rand-Hex	2.50	63.0	X		X	
4	8	1.00	Rand-Hex	2.50	63.0		X		X
4	9	1.00	Rand-Hex	2.50	63.0	X		X	
4	9	1.00	Rand-Hex	2.50	63.0		X		X
4	10	1.00	Rand-Hex	2.50	63.0	X		X	
4	10	1.00	Rand-Hex	2.50	63.0		X		X
4	11	1.00	Rand-Hex	2.50	63.0	X		X	
4	11	1.00	Rand-Hex	2.50	63.0		X		X
4	12	1.00	Rand-Hex	2.50	63.0	X		X	

Table A.3. Primer compositions, PCR conditions maintained for the DNA library developments and Oxford Nanopore sequencing cont.

Sequencing experiment number	DNA library Barcode number	Template RNA volume (µL)	RT-PCR primer composition	Template cDNA volume (µL)	PCR primer extension temperature (°C)	Primer scheme and the concentration (X = 0.5 µM, 1.25 µL; 2X = 1.0 µM, 2.50 µL)																Other primers
						A	B	C	D	E	F	G	H	M	N	O	P	Q	R	S	T	
5	1	1.00	Rand-Hex	2.50	63.0								X								X	
5	2	1.00	Rand-Hex	2.50	63.0							X								X		
5	2	1.00	Rand-Hex	2.50	63.0								X								X	
5	3	1.00	Rand-Hex	2.50	63.0							X								X		
5	3	1.00	Rand-Hex	2.50	63.0								X								X	
5	4	1.00	Rand-Hex	2.50	63.0							X								X		
5	4	1.00	Rand-Hex	2.50	63.0								X								X	
5	5	1.00	Rand-Hex	2.50	63.0							X								X		
5	5	1.00	Rand-Hex	2.50	63.0								X								X	
5	6	1.00	Rand-Hex	2.50	63.0							X								X		
5	6	1.00	Rand-Hex	2.50	63.0								X								X	
5	7	1.00	Rand-Hex	2.50	63.0							X								X		
5	7	1.00	Rand-Hex	2.50	63.0								X								X	
5	8	1.00	Rand-Hex	2.50	63.0							X								X		
5	8	1.00	Rand-Hex	2.50	63.0								X								X	
5	9	1.00	Rand-Hex	2.50	63.0							X								X		
5	9	1.00	Rand-Hex	2.50	63.0								X								X	
5	10	1.00	Rand-Hex	2.50	63.0							X								X		
5	10	1.00	Rand-Hex	2.50	63.0								X								X	
5	11	1.00	Rand-Hex	2.50	63.0							X								X		
5	11	1.00	Rand-Hex	2.50	63.0								X								X	
5	12	1.00	Rand-Hex	2.50	63.0							X								X		
5	12	1.00	Rand-Hex	2.50	63.0								X								X	

Table A.3. Primer compositions, PCR conditions maintained for the DNA library developments and Oxford Nanopore sequencing cont.

Sequencing experiment number	DNA library Barcode number	Template RNA volume (µL)	RT-PCR primer composition	Template cDNA volume (µL)	PCR primer extension temperature (°C)	Primer scheme and the concentration (X = 0.5 µM, 1.25 µL; 2X = 1.0 µM, 2.50 µL)															
						A	B	C	D	E	F	G	H	M	N	O	P	Q	R	S	T
6	1	1.00	Rand-Hex	2.50	63.0				2X								X				
6	2	1.00	Rand-Hex	2.50	63.0			2X								X					
6	2	1.00	Rand-Hex	2.50	63.0				2X								X				
6	3	1.00	Rand-Hex	2.50	63.0			2X								X					
6	3	1.00	Rand-Hex	2.50	63.0				2X								X				
6	4	1.00	Rand-Hex	2.50	63.0			2X								X					
6	4	1.00	Rand-Hex	2.50	63.0				2X								X				
6	5	1.00	Rand-Hex	2.50	63.0			2X								X					
6	5	1.00	Rand-Hex	2.50	63.0				2X								X				
6	6	1.00	Rand-Hex	2.50	63.0			2X								X					
6	6	1.00	Rand-Hex	2.50	63.0				2X								X				
6	7	1.00	Rand-Hex	2.50	63.0			2X								X					
6	7	1.00	Rand-Hex	2.50	63.0				2X								X				
6	8	1.00	Rand-Hex	2.50	63.0			2X								X					
6	8	1.00	Rand-Hex	2.50	63.0				2X								X				
6	9	1.00	Rand-Hex	2.50	63.0			2X								X					
6	9	1.00	Rand-Hex	2.50	63.0				2X								X				
6	10	1.00	Rand-Hex	2.50	63.0			2X								X					
6	10	1.00	Rand-Hex	2.50	63.0				2X								X				
6	11	1.00	Rand-Hex	2.50	63.0							X								X	
6	11	1.00	Rand-Hex	2.50	63.0								X								X
6	12	1.00	TBEV	2.50	63.0							X								X	
6	12	1.00	reverse primers[#]	2.50	63.0								X								X

Table A.3. Primer compositions, PCR conditions maintained for the DNA library developments and Oxford Nanopore sequencing cont.

Sequencing experiment number	DNA library Barcode number	Template RNA volume (µL)	RT-PCR primer composition	Template cDNA volume (µL)	PCR primer extension temperature (°C)	Primer scheme and the concentration (X = 0.5 µM, 1.25 µL; 2X = 1.0 µM, 2.50 µL)															
						A	B	C	D	E	F	G	H	M	N	O	P	Q	R	S	T
7	1	8.00	Rand-Hex	7.00	63.0				2X								2X				
7	1	8.00	Rand-Hex	7.00	63.0							2X								2X	
7	1	8.00	Rand-Hex	7.00	63.0								2X								2X
7	2	8.00	Rand-Hex	7.00	63.0			2X								2X					
7	2	8.00	Rand-Hex	7.00	63.0				2X								2X				
7	2	8.00	Rand-Hex	7.00	63.0							2X								2X	
7	2	8.00	Rand-Hex	7.00	63.0								2X								2X
7	3	8.00	Rand-Hex	7.00	63.0			2X								2X					
7	3	8.00	Rand-Hex	7.00	63.0				2X								2X				
7	3	8.00	Rand-Hex	7.00	63.0							2X								2X	
7	3	8.00	Rand-Hex	7.00	63.0								2X								2X
7	4	8.00	Rand-Hex	7.00	63.0			2X								2X					
7	4	8.00	Rand-Hex	7.00	63.0				2X								2X				
7	4	8.00	Rand-Hex	7.00	63.0							2X								2X	
7	4	8.00	Rand-Hex	7.00	63.0								2X								2X
7	5	8.00	Rand-Hex	7.00	63.0			2X								2X					
7	5	8.00	Rand-Hex	7.00	63.0				2X								2X				
7	5	8.00	Rand-Hex	7.00	63.0							2X								2X	
7	5	8.00	Rand-Hex	7.00	63.0								2X								2X
7	6	8.00	Rand-Hex	7.00	63.0			2X								2X					
7	6	8.00	Rand-Hex	7.00	63.0				2X								2X				
7	6	8.00	Rand-Hex	7.00	63.0							2X								2X	
7	6	8.00	Rand-Hex	7.00	63.0								2X								2X

Table A.3. Primer compositions, PCR conditions maintained for the DNA library developments and Oxford Nanopore sequencing cont.

Sequencing experiment number	DNA library Barcode number	Template RNA volume (µL)	RT-PCR primer composition	Template cDNA volume (µL)	PCR primer extension temperature (°C)	Primer scheme and the concentration (X = 0.5 µM, 1.25 µL; 2X = 1.0 µM, 2.50 µL)															
						A	B	C	D	E	F	G	H	M	N	O	P	Q	R	S	T
7	7	8.00	Rand-Hex	7.00	63.0			2X								2X					
7	7	8.00	Rand-Hex	7.00	63.0				2X								2X				
7	7	8.00	Rand-Hex	7.00	63.0							2X								2X	
7	7	8.00	Rand-Hex	7.00	63.0								2X								2X
7	8	8.00	Rand-Hex	7.00	63.0			2X								2X					
7	8	8.00	Rand-Hex	7.00	63.0				2X								2X				
7	8	8.00	Rand-Hex	7.00	63.0							2X								2X	
7	8	8.00	Rand-Hex	7.00	63.0								2X								2X
7	9	8.00	Rand-Hex	7.00	63.0			2X								2X					
7	9	8.00	Rand-Hex	7.00	63.0				2X								2X				
7	9	8.00	Rand-Hex	7.00	63.0							2X								2X	
7	9	8.00	Rand-Hex	7.00	63.0								2X								2X
7	10	8.00	Rand-Hex	7.00	63.0			2X								2X					
7	10	8.00	Rand-Hex	7.00	63.0				2X								2X				
7	10	8.00	Rand-Hex	7.00	63.0							2X								2X	
7	10	8.00	Rand-Hex	7.00	63.0								2X								2X
7	11	8.00	Rand-Hex	7.00	63.0			2X								2X					
7	11	8.00	Rand-Hex	7.00	63.0				2X								2X				
7	11	8.00	Rand-Hex	7.00	63.0							2X								2X	
7	11	8.00	Rand-Hex	7.00	63.0								2X								2X
7	12	8.00	Rand-Hex	7.00	63.0			2X								2X					
7	12	8.00	Rand-Hex	7.00	63.0				2X								2X				
7	12	8.00	Rand-Hex	7.00	63.0							2X								2X	

**** Each row represents single PCR reaction, can have more than one primer scheme***

Table A.4. Overview of the sequenced data mapped against the reference TBEV-Hochosterwitz strain (NCBI Acc. No. MT311861)

Sequencing experiment number	Sample name	DNA library barcode number	Primers used in cDNA synthesis	Primer scheme composition*	Assembly or mapping technique	Number of sequences			Identical sites (%)	Pairwise identity (%)	Average depth of sequencing	Average breadth of sequencing above 20 reads (%) #	Mean Coverage
						Total	Mapped	Unmapped					
1	HOC-(-4)	1	Rand-Hex	(M)	Minimap2	211791	207262	4529	94.3	94.4	4553.0	64.2	27468.9
1	HOC-(-4)	2	Rand-Hex	(N)	Minimap2	212124	208440	3684	98.0	98.1	4080.2	55.2	28408.2
1	HOC-(-4)	3	Rand-Hex	(A)	Minimap2	58903	53344	5559	99.1	99.2	1362.5	14.7	1829.4
1	HOC-(-4)	4	Rand-Hex	(B)	Minimap2	139872	135745	4127	97.9	97.9	2962.8	29.5	5381.5
1	HOC-(-6)	5	Rand-Hex	(A)	Minimap2	23113	20759	2354	98.9	99.0	438.3	8.4	922.8
1	HOC-(-6)	6	Rand-Hex	(B)	Minimap2	26513	23376	3137	99.6	99.6	491.7	7.5	1099.2
1	HOC-(-7)	7	Rand-Hex	(A)	Minimap2	17701	11118	6583	97.8	97.9	222.3	10.0	626.1
1	HOC-(-7)	8	Rand-Hex	(B)	Minimap2	25737	22690	3047	87.6	87.6	469.7	8.2	1223.5
1	HOC-(-4)	9	Rand-Hex	JK_1-F/R	Minimap2	312283	306670	5613					
1	HOC-(-4)	10	Rand-Hex	Eur_1-F/R	Minimap2	45717	43621	2096					
1	HOC-(-4) + Tick RNA	11	Rand-Hex	JK_1-F/R	Minimap2	455272	431691	23581					
1	HOC-(-4) + Tick RNA	12	Rand-Hex	Eur_1-F/R	Minimap2	80031	70064	9967					
1		Unspecific			Minimap2	120496	115549	4947	99.0	99.0	3002.6		14328.4

Table A.4. Overview of the sequenced data mapped against the reference TBEV-Hochosterwitz strain (NCBI Acc. No. MT311861) cont.

Sequencing experiment number	Sample name	DNA library barcode number	Primers used in cDNA synthesis	Primer scheme composition*	Assembly or mapping technique	Number of sequences			Identical sites (%)	Pairwise identity (%)	Average depth of sequencing	Average breadth of sequencing above 20 reads (%) #	Mean Coverage
						Total	Mapped	Unmapped					
1	HOC-(-4)	1&2	Rand-Hex	(M)+(N)	Minimap2	417963	410236	7727	94.3	97.9	7117.6	66.8	34708.1
1	HOC-(-4)	3&4	Rand-Hex	(A)+(B)	Minimap2	198347	189204	9143	79.9	99.7	3964.0	36.5	6072.1
1	HOC-(-6)	5&6	Rand-Hex	(A)+(B)	Minimap2	49600	44319	5281	90.3	99.3	959.3	9.9	1642.1
1	HOC-(-7)	7&8	Rand-Hex	(A)+(B)	Minimap2	43416	33948	9468	70.0	95.5	714.7	9.6	1714.5
1	HOC-(-4)	1	Rand-Hex	(M)	Trinity > Minimap2	211791			76.9	77.0			
1	HOC-(-4)	2	Rand-Hex	(N)	Trinity > Minimap2	212124			99.4	99.4			
1	HOC-(-4)	3	Rand-Hex	(A)	Trinity > Minimap2	58903			94.4	94.6			
1	HOC-(-4)	4	Rand-Hex	(B)	Trinity > Minimap2	139872			95.5	95.5			
1	HOC-(-6)	5	Rand-Hex	(A)	Trinity > Minimap2	23113			96.5	96.6			
1	HOC-(-6)	6	Rand-Hex	(B)	Trinity > Minimap2	26513			96.0	96.0			
1	HOC-(-7)	7	Rand-Hex	(A)	Trinity > Minimap2	17701			93.7	93.9			
1	HOC-(-7)	8	Rand-Hex	(B)	Trinity > Minimap2	25737			71.5	71.6			
1	HOC-(-4)	9	Rand-Hex	(A)	Trinity > Minimap2	312283			90.5	90.5			
1	HOC-(-4)	10	Rand-Hex	(B)	Trinity > Minimap2	45717			99.7	99.7			
1	HOC-(-4)	11	Rand-Hex	(A)	Trinity > Minimap2	455272			68.0	68.6			
1	HOC-(-4)	12	Rand-Hex	(B)	Trinity > Minimap2	80031			81.7	81.9			
1		Unspecific			Trinity > Minimap2	120496			99.0	99.0			

Table A.4. Overview of the sequenced data mapped against the reference TBEV-Hochosterwitz strain (NCBI Acc. No. MT311861) cont.

Sequencing experiment number	Sample name	DNA library barcode number	Primers used in cDNA synthesis	Primer scheme composition*	Assembly or mapping technique	Number of sequences			Identical sites (%)	Pairwise identity (%)	Average depth of sequencing	Average breadth of sequencing above 20 reads (%) #	Mean Coverage
						Total	Mapped	Unmapped					
1	HOC-(-4)	1&2	Rand-Hex	(M)+(N)	Trinity > Minimap2	120496			97.7	97.7			
1	HOC-(-4)	3&4	Rand-Hex	(A)+(B)	Trinity > Minimap2	417963			99.4	99.5			
1	HOC-(-6)	5&6	Rand-Hex	(A)+(B)	Trinity > Minimap2	198347			96.8	96.8			
1	HOC-(-7)	7&8	Rand-Hex	(A)+(B)	Trinity > Minimap2	49600			96.6	97.0			
2	HOC-(-4)	1	Rand-Hex	(A+O)+(B+P)	Minimap2	312631	286085	26546	99.4	99.9	4236.6	39.0	7706.7
2	HOC-(-5)	2	Rand-Hex	(A+O)+(B+P)	Minimap2	191118	148108	43010	98.6	99.8	2284.8	21.4	4339.5
2	HOC-(-6)	3	Rand-Hex	(A+O)+(B+P)	Minimap2	613	197	419	5.0	65.9	3.5	0.2	3.5
2	HOC-(-4)	4	Rand-Hex	(C+O)+(D+P)	Minimap2	526840	472964	53876	99.5	99.9	5262.2	48.7	12684.2
2	HOC-(-5)	5	Rand-Hex	(C+O)+(D+P)	Minimap2	124924	91759	33165	96.4	99.9	1707.9	16.6	2277.1
2	HOC-(-6)	6	Rand-Hex	(C+O)+(D+P)	Minimap2	71874	40533	31341	81.4	98.2	677.4	8.3	1117.3
2	HOC-(-4)	7	Rand-Hex	(2A+O)+(2B+P)	Minimap2	561778	519470	42308	99.3	99.9	4581.6	42.1	15546.7
2	HOC-(-5)	8	Rand-Hex	(2A+O)+(2B+P)	Minimap2	191725	171721	20004	9.3	99.8	4252.8	22.4	5564.5
2	HOC-(-6)	9	Rand-Hex	(2A+O)+(2B+P)	Minimap2	92990	55331	37659	85.9	98.7	942.4	11.0	1771.8
2	HOC-(-4)	10	Rand-Hex	(A)+(B)	Minimap2	225221	214118	11103	97.9	99.9	3585.58	33.0	6532.4
2	HOC-(-5)	11	Rand-Hex	(A)+(B)	Minimap2	249050	172742	76308	98.6	99.9	2840.0	26.7	4278.5
2	HOC-(-6)	12	Rand-Hex	(A)+(B)	Minimap2	24064	20325	3739	72.2	93.3	373.0	5.3	682.7
2		Unspecific			Minimap2	150218	129524	20694	99.9	99.9			2995.2

Table A.4. Overview of the sequenced data mapped against the reference TBEV-Hochosterwitz strain (NCBI Acc. No. MT311861) cont.

Sequencing experiment number	Sample name	DNA library barcode number	Primers used in cDNA synthesis	Primer scheme composition*	Assembly or mapping technique	Number of sequences			Identical sites (%)	Pairwise identity (%)	Average depth of sequencing	Average breadth of sequencing above 20 reads (%) #	Mean Coverage
						Total	Mapped	Unmapped					
3	HOC-(-5)	1	Rand-Hex	(C+O)+(D+P)	Minimap2	160135	131534	28601	99.0	99.4	2034.3	19.1	2705.0
3	HOC-(-6)	2	Rand-Hex	(C+O)+(D+P)	Minimap2	60012	31280	28732	98.6	97.7	589.0	7.5	608.4
3	HOC-(-5)	3	Rand-Hex	(2E+Q)+(2F+R)	Minimap2	128080	116778	11302	96.8	98.2	2218.6	22.3	2671.8
3	HOC-(-6)	4	Rand-Hex	(2E+Q)+(2F+R)	Minimap2	140032	93449	46583	94.4	93.0	1645.9	22.8	1943.8
3	HOC-(-5)	5	Rand-Hex	(2E+2Q)+ (2F+2R)	Minimap2	432293	370515	61778	97.1	97.7	3957.5	37.0	8434.5
3	HOC-(-6)	6	Rand-Hex	(2E+2Q)+ (2F+2R)	Minimap2	228021	177662	50359	98.3	96.4	2549.4	25.8	4065.0
3	HOC-(-5)	7	Rand-Hex	(G+S)+(H+T)	Minimap2	576125	521922	54203	97.0	97.9	4019.5	38.0	11425.0
3	HOC-(-6)	8	Rand-Hex	(G+S)+(H+T)	Minimap2	216060	183669	32391	95.8	96.2	2268.1	26.8	3877.4
3	HOC-(-5)	9	Rand-Hex	(2G+2S)+ (2H+2T)	Minimap2	308067	261885	46182	97.4	98.6	3543.3	33.5	5827.1
3	HOC-(-6)	10	Rand-Hex	(2G+2S)+ (2H+2T)	Minimap2	176038	152204	23834	98.0	97.5	2285.7	23.6	5464.6
3	HOC-(-5)	11	Rand-Hex	CJB_37-F/R + CJB_38-F/R + CJB_39-F/R	Minimap2	612082	186637	425445					
3	HOC-(-5)	12	Rand-Hex	(CJB_37-F + Eur_37-R), (CJB_37-R + Eur_37-F)	Minimap2	260101	253201	6900					
3		Unspecific			Minimap2	224068	182512	41556	98.3	98.7			4315.56

Table A.4. Overview of the sequenced data mapped against the reference TBEV-Hochosterwitz strain (NCBI Acc. No. MT311861) cont.

Sequencing experiment number	Sample name	DNA library barcode number	Primers used in cDNA synthesis	Primer scheme composition*	Assembly or mapping technique	Number of sequences			Identical sites (%)	Pairwise identity (%)	Average depth of sequencing	Average breadth of sequencing above 20 reads (%) #	Mean Coverage
						Total	Mapped	Unmapped					
4	SL1(21)-01N	1	Rand-Hex	(G+S)+(H+T)	Minimap2	34709	288	34421	36.5	91.9	3.7	0.7	3.8
4	SL1(21)-03N	2	Rand-Hex	(G+S)+(H+T)	Minimap2	19745	1	19744	96.9	96.9	0.0	0.0	0.0
4	SL1(21)-17N	3	Rand-Hex	(G+S)+(H+T)	Minimap2	12060	113	11947	57.3	91.7	2.0	0.3	2.1
4	SL1(21)-22N	4	Rand-Hex	(G+S)+(H+T)	Minimap2	45553	1	45552	92.2	92.2	0.1	0.0	0.1
4	SL1(21)-30N	5	Rand-Hex	(G+S)+(H+T)	Minimap2	49437	0	49437	0.0	0.0	0.0	0.0	0.0
4	SL1(21)-41N	6	Rand-Hex	(G+S)+(H+T)	Minimap2	67407	0	67407	0.0	0.0	0.0	0.0	0.0
4	SL1(21)-58N	7	Rand-Hex	(G+S)+(H+T)	Minimap2	102282	0	102282	0.0	0.0	0.0	0.0	0.0
4	SL1(21)-60N	8	Rand-Hex	(G+S)+(H+T)	Minimap2	90320	0	90320	0.0	0.0	0.0	0.0	0.0
4	SL1(21)-64N	9	Rand-Hex	(G+S)+(H+T)	Minimap2	114068	0	114068	0.0	0.0	0.0	0.0	0.0
4	SL1(21)-67N	10	Rand-Hex	(G+S)+(H+T)	Minimap2	80670	1	80669	95.8	95.8	0.0	0.0	0.0
4	SL1(21)-27F	11	Rand-Hex	(G+S)+(H+T)	Minimap2	214249	2	214247	93.3	93.3	0.0	0.0	0.0
4	SL1(21)-49M	12	Rand-Hex	(G+S)+(H+T)	Minimap2	19521	0	19521	0.0	0.0	0.0	0.0	0.0
4		Unspecific			Minimap2	45367	27	45340	85.7	92.9	0.4		0.4

Table A.4. Overview of the sequenced data mapped against the reference TBEV-Hochosterwitz strain (NCBI Acc. No. MT311861) cont.

Sequencing experiment number	Sample name	DNA library barcode number	Primers used in cDNA synthesis	Primer scheme composition*	Assembly or mapping technique	Number of sequences			Identical sites (%)	Pairwise identity (%)	Average depth of sequencing	Average breadth of sequencing above 20 reads (%) #	Mean Coverage
						Total	Mapped	Unmapped					
5	SL1(21)-03M	1	Rand-Hex	(G+S)+(H+T)	Minimap2	30088	1	30087	86.1	86.1	0.0	0.0	0.0
5	SL1(21)-04M	2	Rand-Hex	(G+S)+(H+T)	Minimap2	48674	1	48673	98.8	98.8	0.1	0.0	0.1
5	SL1(21)-24M	3	Rand-Hex	(G+S)+(H+T)	Minimap2	12556	0	12556	0.0	0.0	0.0	0.0	0
5	SL1(21)-33M	4	Rand-Hex	(G+S)+(H+T)	Minimap2	72373	2	72371	91.1	91.1	0.0	0.0	0.0
5	SL1(21)-35M	5	Rand-Hex	(G+S)+(H+T)	Minimap2	59405	0	59405	0.0	0.0	0.0	0.0	0
5	SL1(21)-38M	6	Rand-Hex	(G+S)+(H+T)	Minimap2	44492	0	44492	0.0	0.0	0.0	0.0	0
5	SL1(21)-30M	7	Rand-Hex	(G+S)+(H+T)	Minimap2	62429	0	62429	0.0	0.0	0.0	0.0	0
5	SL1(21)-42M	8	Rand-Hex	(G+S)+(H+T)	Minimap2	24746	0	24746	0.0	0.0	0.0	0.0	0
5	SL1(21)-32M	9	Rand-Hex	(G+S)+(H+T)	Minimap2	34329	0	34329	0.0	0.0	0.0	0.0	0
5	SL1(21)-13F	10	Rand-Hex	(G+S)+(H+T)	Minimap2	45492	0	45492	0.0	0.0	0.0	0.0	0
5	SL1(21)-42F	11	Rand-Hex	(G+S)+(H+T)	Minimap2	115845	2	115843	94.5	94.5	0.1	0.0	0.1
5	SL1(21)-24F	12	Rand-Hex	(G+S)+(H+T)	Minimap2	30541	10	30531	87.0	95.8	0.1	0.0	0.1
5		Unspecific			Minimap2	31950	2	31948	93.5	93.5	0.0		0.0

Table A.4. Overview of the sequenced data mapped against the reference TBEV-Hochosterwitz strain (NCBI Acc. No. MT311861) cont.

Sequencing experiment number	Sample name	DNA library barcode number	Primers used in cDNA synthesis	Primer scheme composition*	Assembly or mapping technique	Number of sequences			Identical sites (%)	Pairwise identity (%)	Average depth of sequencing	Average breadth of sequencing above 20 reads (%) #	Mean Coverage
						Total	Mapped	Unmapped					
6	SL1(21)-32F	1	Rand-Hex	(2C+O)+(2D+P)	Minimap2	12001	27	11974	90.7	92.2	0.4	0.0	0.4
6	SL1(21)-17F	2	Rand-Hex	(2C+O)+(2D+P)	Minimap2	28001	35	27966	79.9	88.0	0.5	0.2	0.5
6	SL1(21)-42F	3	Rand-Hex	(2C+O)+(2D+P)	Minimap2	32000	8	31992	93.3	93.3	0.2	0.0	0.2
6	SL1(21)-27F	4	Rand-Hex	(2C+O)+(2D+P)	Minimap2	52000	9	51991	93.1	93.6	0.1	0.0	0.1
6	SL1(21)-03M	5	Rand-Hex	(2C+O)+(2D+P)	Minimap2	36000	12	35988	88.1	89.9	0.2	0.0	0.2
6	SL1(21)-04M	6	Rand-Hex	(2C+O)+(2D+P)	Minimap2	28000	17	27983	87.8	91.1	0.3	0.0	0.3
6	SL1(21)-01N	7	Rand-Hex	(2C+O)+(2D+P)	Minimap2	56000	22	55978	93.7	94.4	0.3	0.0	0.3
6	SL1(21)-17N	8	Rand-Hex	(2C+O)+(2D+P)	Minimap2	48001	1262	46739	47.7	90.1	23.3	2.3	23.8
6	SL1(21)-22N	9	Rand-Hex	(2C+O)+(2D+P)	Minimap2	28000	11	27989	87.5	91.4	0.2	0.0	0.2
6	SL1(21)-67N	10	Rand-Hex	(2C+O)+(2D+P)	Minimap2	24000	18	23982	90.7	91.1	0.2	0.0	0.2
6	HOC-(-6)	11	Rand-Hex	(G+S)+(H+T)	Minimap2	48009	20054	27955	13.9	89.1	319.7	6.11	351.2
6	HOC-(-6)	12	All the TBEV specific reverse primers*	(G+S)+(H+T)	Minimap2	44001	48	43953	88.4	92.2	0.6	0.0	0.7
6		Unspecific			Minimap2	28001	846	27155	55.0	88.7	15.4		15.8

Table A.4. Overview of the sequenced data mapped against the reference TBEV-Hochosterwitz strain (NCBI Acc. No. MT311861) cont.

Sequencing experiment number	Sample name	DNA library barcode number	Primers used in cDNA synthesis	Primer scheme composition*	Assembly or mapping technique	Number of sequences			Identical sites (%)	Pairwise identity (%)	Average depth of sequencing	Average breadth of sequencing above 20 reads (%) #	Mean Coverage
						Total	Mapped	Unmapped					
7	SL1(21)-01N	1	Rand-Hex	(2C+2O)+(2D+2P) +(2G+2S)+(2H+2T)	Minimap2	164001	2682	161319	37.5	85.3	57.9	13.6	60.6
7	SL1(21)-17N	2	Rand-Hex	(2C+2O)+(2D+2P) +(2G+2S)+(2H+2T)	Minimap2	280006	2780	277226	20.7	87.2	46.1	2.2	48.1
7	SL1(21)-22N	3	Rand-Hex	(2C+2O)+(2D+2P) +(2G+2S)+(2H+2T)	Minimap2	164001	30	163971	54.5	72.7	0.7	0.0	0.8
7	SL1(21)-67N	4	Rand-Hex	(2C+2O)+(2D+2P) +(2G+2S)+(2H+2T)	Minimap2	124001	46	123955	54.7	73.0	0.7	0.0	0.9
7	SL1(21)-17F	5	Rand-Hex	(2C+2O)+(2D+2P) +(2G+2S)+(2H+2T)	Minimap2	256005	47025	208980	4.2	88.5	758.9	13.4	901.9
7	SL1(21)-24F	6	Rand-Hex	(2C+2O)+(2D+2P) +(2G+2S)+(2H+2T)	Minimap2	132001	3164	128837	27.7	86.0	59.1	7.0	61.6
7	SL1(21)-27F	7	Rand-Hex	(2C+2O)+(2D+2P) +(2G+2S)+(2H+2T)	Minimap2	152001	434	151567	42.2	82.9	7.8	1.9	8.2
7	SL1(21)-32F	8	Rand-Hex	(2C+2O)+(2D+2P) +(2G+2S)+(2H+2T)	Minimap2	128002	5586	122416	23.5	87.6	101.4	7.3	105.2
7	SL1(21)-42F	9	Rand-Hex	(2C+2O)+(2D+2P) +(2G+2S)+(2H+2T)	Minimap2	68001	610	67391	39.1	83.9	11.9	2.0	12.5
7	SL1(21)-03M	10	Rand-Hex	(2C+2O)+(2D+2P) +(2G+2S)+(2H+2T)	Minimap2	64001	52	63949	68.4	86.6	0.6	0.0	0.7
7	SL1(21)-04M	11	Rand-Hex	(2C+2O)+(2D+2P) +(2G+2S)+(2H+2T)	Minimap2	132001	63	131938	53.9	81.7	0.9	0.2	1.0
7	SL1(21)-33M	12	Rand-Hex	(2C+2O)+(2D+2P) +(2G+2S)+(2H+2T)	Minimap2	160001	1529	158472	29.3	87.4	26.7	2.4	27.7
7		Unspecific			Minimap2	108001	4306	103695	13.2	87.1	82.6		85.7

Table A.4. Overview of the sequenced data mapped against the reference TBEV-Hochosterwitz strain (NCBI Acc. No. MT311861) cont.

Sequencing experiment number	Sample name	DNA library barcode number	Primers used in cDNA synthesis	Primer scheme composition*	Assembly or mapping technique	Number of sequences			Identical sites (%)	Pairwise identity (%)	Average depth of sequencing	Average breadth of sequencing	Mean Coverage
						Total	Mapped	Unmapped					
8###	HOC	X	Rand-Hex	(A)+(B)	Minimap2	159161	132443	26718	98.0	98.5	4531.2	41.7	12144.9
8###	HOC	X	Rand-Hex	(A)+(B)	Trinity > Minimap2	159161			87.5	87.5			
9###	HOC	Y	Rand-Hex	(A)+(B)	Minimap2	2837342	2746650	90692	100.0	100.0	7932.4	73.0	72713.0
9###	HOC	Y	Rand-Hex	(A)+(B)	Trinity > Minimap2				99.9	99.9			

*HOC: Hochosterwitz strain (NCBI Acc. No. MT311861); HOC-**(-#)**, (-#): Tenfold dilution factor, Rand-Hex: Random hexamers; "+": Combined PCR products from different primer schemes; Trinity > Minimap2: 1st assembled through Trinity tool from the Galaxy online server (version: 21.09, "https://usegalaxy.org.au/", Afgan et al., 2018) and followed by mapping through Minimap2 plugin of Geneious Prime® (Geneious Prime®, 2022.0.1); Blank (" "): No data /or/ Not applicable /or/ No significant meaning; Identical sites (%), Pairwise identity (%) and Mean coverage: Values were derived through the Minimap2 plugin of Geneious Prime® (Geneious Prime®, 2022.0.1); Mean coverage: Calculated by SAMtools (Danecek et al., 2021, Version: 1.3.1+htslib-1.3.1) installed with an IDE (Anaconda Software Distribution, 2020, Version: 4.11.0) in a Linux Mint (Version: 20.3); Reference based mapping based on the reference TBEV sequence (Hochosterwitz strain [NCBI Acc. No. MT311861]); * Primer schemes according to the primer scheme table (Table A.2. , Annexure I); ###: Data from a previous study carried out with the TBEV-Hochosterwitz strain with no dilutions at NIPH (8###: Oxford Nanopore sequence data, 9###: Illumina sequence data).*

Table A.5. TBEV nucleotide sequences downloaded from the NCBI nucleotide archive for the construction of consensus sequences in the new primers designing

Acc. No. (NCBI)	**Country**	**Strain**	**Description**
AB062063.2			Tick-borne encephalitis virus gene, complete cds, strain:Oshima 5-10
AB062064.1			Tick-borne encephalitis virus gene, complete cds, strain:Sofjin-HO
AB753012.1	Japan:Hokkaido, Hokuto		Tick-borne encephalitis virus RNA, complete genome, strain: Oshima 08-As
AF069066.1		Infectious clone	Tick-borne encephalitis virus strain Vasilchenko, complete genome
AF527415.1		Siberian	Tick-borne encephalitis virus strain Zausaev polyprotein gene, complete cds
AM600965.1	Germany:Karlsruhe		Tick-borne encephalitis virus gene for polyprotein, strain K23, genomic RNA
AY169390.3	Russia		Tick-borne encephalitis virus strain Primorye-332, complete genome
AY182009.1	China		Tick-borne encephalitis virus strain Senzhang polyprotein gene, complete cds
AY217093.1	China		Tick-borne encephalitis virus strain MDJ-01 nonfunctional polyprotein gene, complete sequence
DQ401140.3	Sweden: Toro		Tick-borne encephalitis virus isolate Toro-2003, complete genome
DQ862460.1	Russia: Primorsky krai		Tick-borne encephalitis virus strain Glubinnoe/2004 polyprotein mRNA, complete cds
DQ989336.1	Russia		Tick-borne encephalitis virus strain 205 polyprotein mRNA, complete cds
EF469661.1	Russia: Irkutsk region		Tick-borne encephalitis virus strain 178-79, complete genome
EF469662.1	Russia: Irkutsk region		Tick-borne encephalitis virus strain 886-84, complete genome
EU816450.1	Russia		Tick-borne encephalitis virus strain Primorye-212, complete genome
EU816451.1	Russia		Tick-borne encephalitis virus strain Primorye-253, complete genome
EU816452.1	Russia		Tick-borne encephalitis virus strain Primorye-270, complete genome
EU816453.1	Russia		Tick-borne encephalitis virus strain Primorye-69, complete genome
EU816454.1	Russia		Tick-borne encephalitis virus strain Primorye-94, complete genome
EU816455.2	Russia		Tick-borne encephalitis virus strain Primorye-86, complete genome
FJ402885.1	Russia		Tick-borne encephalitis virus strain Kavalerovo, complete genome
FJ402886.1	Russia		Tick-borne encephalitis virus strain Dalnegorsk, complete genome
FJ572210.1	Germany	Central European	Tick-borne encephalitis virus strain Salem, complete genome
FJ906622.1	Russia		Tick-borne encephalitis virus strain Primirye-89, complete genome
FJ968751.1	Russia	Siberian	Tick-borne encephalitis virus isolate Kolarovo-2008, complete genome
FJ997899.1	Russia		Tick-borne encephalitis virus strain Primorye-90, complete genome
GQ228395.1	Russia		Tick-borne encephalitis virus strain Primorye-18, complete genome
GQ266392.1	Germany: Amberg		Tick-borne encephalitis virus isolate AS33, complete genome
GU121642.1	Russia		Tick-borne encephalitis virus strain Svetlogorie, complete genome

GU183379.1	Finland		Tick-borne encephalitis virus strain Kumlinge 25-03 polyprotein gene, complete cds
GU183380.1	Finland		Tick-borne encephalitis virus strain Kumlinge A52 polyprotein gene, complete cds
GU183381.1	Finland		Tick-borne encephalitis virus strain Joutseno polyprotein gene, complete cds
GU183382.1	Latvia		Tick-borne encephalitis virus strain Latvia-1-96 polyprotein gene, complete cds
GU183383.1	Estonia		Tick-borne encephalitis virus strain Est3476 polyprotein gene, complete cds
GU183384.1	Estonia		Tick-borne encephalitis virus strain Est54 polyprotein gene, complete cds
HM120875.1	Russia: Altay		Tick-borne encephalitis virus strain 84.2 polyprotein gene, complete cds
HM535610.1	South Korea	Western	Tick-borne encephalitis virus strain KrM 213, complete genome
HM535611.1	South Korea	Western	Tick-borne encephalitis virus strain KrM 93, complete genome
HM859894.1	Russia	Far Eastern	Tick-borne encephalitis virus strain Primorye-633, complete genome
HM859895.1	Russia	Far Eastern	Tick-borne encephalitis virus strain Primorye-2239, complete genome
HQ201303.1	Russia		Tick-borne encephalitis virus strain Primorye-92, complete genome
HQ901366.1	Russia	Far Eastern	Tick-borne encephalitis virus strain Primorye-1153, complete genome
HQ901367.1	Russia	Far Eastern	Tick-borne encephalitis virus strain Primorye-501, complete genome
JF316707.1	China: Mudanjiang	Far Eastern	Tick-borne encephalitis virus strain MDJ-02, complete genome
JF316708.1	China: Mudanjiang	Far Eastern	Tick-borne encephalitis virus strain MDJ-03, complete genome
JF819648.2	Russia: Far East,	Far Eastern	Tick-borne encephalitis virus strain SofjinKSY, complete genome
JN003205.1	Russia	Far Eastern	Tick-borne encephalitis virus strain Irkutsk-1861 polyprotein gene, complete cds
JN003206.1	Russia	Siberian	Tick-borne encephalitis virus strain Aina polyprotein gene, complete cds
JN003207.1	Russia	Siberian	Tick-borne encephalitis virus strain Cht-653 polyprotein gene, complete cds
JN003208.1	Russia	Siberian	Tick-borne encephalitis virus strain Cht-22 polyprotein gene, complete cds
JN003209.1	Russia	Siberian	Tick-borne encephalitis virus strain Irkutsk-12 polyprotein gene, complete cds
JN229223.1	Russia	Far Eastern	Tick-borne encephalitis virus strain Sofjin-Ru polyprotein gene, complete cds
JQ650522.1	China		Tick-borne encephalitis virus isolate MDJ01, complete genome
JQ650523.1	China		Tick-borne encephalitis virus isolate Senzhang, complete genome
JQ654701.1	Slovenia		Tick-borne encephalitis virus strain Ljubljana I polyprotein gene, complete cds
JQ825144.1	Russia	Far Eastern	Tick-borne encephalitis virus strain Primorye-828 polyprotein gene, complete cds
JQ825145.1	Russia	Far Eastern	Tick-borne encephalitis virus strain Primorye-895 polyprotein gene, complete cds
JQ825146.1	Russia	Far Eastern	Tick-borne encephalitis virus strain Kiparis-94 polyprotein gene, complete cds
JQ825147.1	Russia	Far Eastern	Tick-borne encephalitis virus strain Shkotovo-94 polyprotein gene, complete cds
JQ825148.1	Russia	Far Eastern	Tick-borne encephalitis virus strain Primorye-82 polyprotein gene, complete cds
JQ825149.1	Russia	Far Eastern	Tick-borne encephalitis virus strain Primorye-87 polyprotein gene, complete cds
JQ825150.1	Russia	Far Eastern	Tick-borne encephalitis virus strain Primorye-91 polyprotein gene, complete cds
JQ825151.1	Russia	Far Eastern	Tick-borne encephalitis virus strain Spassk-72 polyprotein gene, complete cds
JQ825152.1	Russia	Far Eastern	Tick-borne encephalitis virus strain Primorye-75 polyprotein gene, complete cds

JQ825153.1	Russia	Far Eastern	Tick-borne encephalitis virus strain Primorye-183 polyprotein gene, complete cds
JQ825154.1	Russia	Far Eastern	Tick-borne encephalitis virus strain Primorye-52 polyprotein gene, complete cds
JQ825155.1	Russia	Far Eastern	Tick-borne encephalitis virus strain Primorye-196 polyprotein gene, complete cds
JQ825156.1	Russia	Far Eastern	Tick-borne encephalitis virus strain Primorye-739 polyprotein gene, complete cds
JQ825157.1	Russia	Far Eastern	Tick-borne encephalitis virus strain Primorye-202 polyprotein gene, complete cds
JQ825158.1	Russia	Far Eastern	Tick-borne encephalitis virus strain Primorye-208 polyprotein gene, complete cds
JQ825159.1	Russia	Far Eastern	Tick-borne encephalitis virus strain Primorye-274 polyprotein gene, complete cds
JQ825160.1	Russia	Far Eastern	Tick-borne encephalitis virus strain Primorye-320 polyprotein gene, complete cds
JQ825161.1	Russia	Far Eastern	Tick-borne encephalitis virus strain Primorye-345 polyprotein gene, complete cds
JQ825162.1	Russia	Far Eastern	Tick-borne encephalitis virus strain Primorye-437 polyprotein gene, complete cds
JQ825163.1	Russia	Far Eastern	Tick-borne encephalitis virus strain Primorye-750 polyprotein gene, complete cds
JQ825164.1	Russia	Far Eastern	Tick-borne encephalitis virus strain Primorye-823 polyprotein gene, complete cds
JX534167.1	China		Tick-borne encephalitis virus isolate Xinjiang-01, complete genome
JX968560.1	Russia	Far Eastern	Tick-borne encephalitis virus strain Zabaikalye-1-98 polyprotein gene, complete cds
KC414090.1	Russia	Siberian	Tick-borne encephalitis virus isolate Zabaikalye 11-99, complete genome
KC422663.2	Russia	Siberian	Tick-borne encephalitis virus strain Zabaikalye 68B-00 polyprotein gene, complete cds
KC422667.2	Russia	Far Eastern	Tick-borne encephalitis virus strain Zabaikalye 30-00 polyprotein gene, complete cds
KC835595.1	Slovakia: Plastovce	Central European	Tick-borne encephalitis virus strain 114, complete genome
KC835596.1	Slovakia: Malacky	Central European	Tick-borne encephalitis virus strain 285, complete genome
KC835597.1	Slovakia: Zahorska Ves	Central European	Tick-borne encephalitis virus strain CGl223, complete genome
KF151173.1	Austria	Central European	Tick-borne encephalitis virus strain A104, complete genome
KF823822.1	Russia	Siberian	Tick-borne encephalitis virus strain Irkutsk_BR_683-11 polyprotein gene, complete cds
KF951037.1	Russia	Far Eastern	Tick-borne encephalitis virus strain 4072, complete genome
KF991106.1	Sweden: Saringe	Western European	Tick-borne encephalitis virus strain Saringe-2009 polyprotein gene, complete cds
KF991107.1	Norway: Mandal	Western European	Tick-borne encephalitis virus strain Mandal-2009 polyprotein gene, complete cds
KJ000002.1	Russia: Karelia	Western	Tick-borne encephalitis virus strain Absettarov, complete genome
KJ626343.1	USSR: Kyrgyzstan	Siberian	Tick-borne encephalitis virus strain Buzuuchuk, complete genome
KJ633033.1	Russia: Irkutsk region		Tick-borne encephalitis virus strain 886-84, complete genome
KJ701416.1	Russia: Novosibirsk	Siberian	Tick-borne encephalitis virus strain Lesopark 11, complete genome
KJ739729.1	Russia: Novosibirsk		Tick-borne encephalitis virus strain Novosibirsk-L2008, complete genome
KJ739730.1	Russia: Novosibirsk		Tick-borne encephalitis virus strain Tomsk-K6, complete genome
KJ739731.1	Russia: Tomsk		Tick-borne encephalitis virus strain Tomsk-M83, complete genome
KJ744034.1	Russia: Khabarovsk		Tick-borne encephalitis virus strain Malishevo polyprotein precursor, gene, complete cds
KJ755186.1	China	Far Eastern	Tick-borne encephalitis virus strain WH2012, complete genome
KJ914682.1	Russia: Tomsk region	Far Eastern	Tick-borne encephalitis virus strain Tomsk-PT14, complete genome

KJ914683.1	Russia: Tomsk region	Far Eastern	Tick-borne encephalitis virus strain Tomsk-M202, complete genome
KJ922512.1	Czech Republic		Tick-borne encephalitis virus strain Kubinova, complete genome
KJ922513.1	Czech Republic		Tick-borne encephalitis virus strain Petracova, complete genome
KJ922514.1	Czech Republic		Tick-borne encephalitis virus strain Skrivanek, complete genome
KJ922515.1	Czech Republic		Tick-borne encephalitis virus strain Tobrman, complete genome
KJ922516.1	Czech Republic		Tick-borne encephalitis virus strain Vlasaty, complete genome
KM019545.1	Russia: Tomsk region	Siberian	Tick-borne encephalitis virus strain Tomsk-PT122, complete genome
KM019546.1	Russia: Tomsk region	Far Eastern	Tick-borne encephalitis virus strain Tomsk-PT12, complete genome
KP331441.1	Russia: Eastern Siberia	European	Tick-borne encephalitis virus strain IrkutskBR_99-08, complete genome
KP331442.1	Russia: Eastern Siberia	European	Tick-borne encephalitis virus strain IrkutskBR_1434-09, complete genome
KP331443.1	Russia	European	Tick-borne encephalitis virus strain IrkutskBR_1456-09, complete genome
KP345889.1	China	Siberian	Tick-borne encephalitis virus isolate Sib-XJ-X5 polyprotein gene, complete cds
KP644245.1	Russia: Novosibirsk		Tick-borne encephalitis virus strain C11-13, complete genome
KP716971.1			Tick-borne encephalitis virus clone Vs[Hypr_E], complete genome
KP716972.1			Tick-borne encephalitis virus clone Vs[Hypr_prM-E], complete genome
KP716973.1			Tick-borne encephalitis virus clone Vs[Hypr_str], complete genome
KP716974.1		West European	Tick-borne encephalitis virus clone Hypr_IC, complete genome
KP716975.1		West European	Tick-borne encephalitis virus clone Hypr_IC_[short_3'UTR], complete genome
KP716976.1			Tick-borne encephalitis virus clone Hypr[Vs_E], complete genome
KP716977.1			Tick-borne encephalitis virus clone Hypr[Vs_prM-E], complete genome
KP716978.1			Tick-borne encephalitis virus clone Hypr[Vs_str], complete genome
KP844724.1	Russia: Khabarovsk	Far Eastern	Tick-borne encephalitis virus strain Chichagovka 1222, complete genome
KP844725.1	Russia: Khabarovsk	Far Eastern	Tick-borne encephalitis virus strain Chichagovka 1223, complete genome
KP844726.1	Russia: Birobidzhan	Far Eastern	Tick-borne encephalitis virus strain Birobidzhan 1354, complete genome
KP844727.1	Russia: Birobidzhan	Far Eastern	Tick-borne encephalitis virus strain Birobidzhan 1357, complete genome
KP869172.1	Russia: Khabarovsk	Far Eastern	Tick-borne encephalitis virus strain Nikolaevsk 855, complete genome
KP938507.1	Russia: Eastern Siberia	Western	Tick-borne encephalitis virus strain Sorex 18-10, complete genome
KT001070.1	Russia: Khabarovsk	Far Eastern	Tick-borne encephalitis virus strain Khekhtzir 9-13, complete genome
KT001071.1	Russia:Khabarovsk	Far Eastern	Tick-borne encephalitis virus strain Khekhtzir 10-13, complete genome
KT001072.1	Russia: Khabarovsk	Far Eastern	Tick-borne encephalitis virus strain Khekhtzir 17-13, complete genome
KT001073.1	Russia: Khabarovsk	Far Eastern	Tick-borne encephalitis virus strain Lazo MP36, complete genome
KT069219.1	Russia		Tick-borne encephalitis virus isolate Primorye-1001, complete genome
KT224352.1	Russia: Arkhangelsk		Tick-borne encephalitis virus strain LEIV-13652Ar polyprotein precursor, gene, complete cds
KT224353.1	Russia: Altai		Tick-borne encephalitis virus strain LEIV-10133Al polyprotein precursor, gene, complete cds
KT224357.1	Finland		Tick-borne encephalitis virus strain Kumlinge polyprotein precursor, gene, complete cds

KT321430.1	Russia	Siberian	Tick-borne encephalitis virus strain Konst-14 polyprotein gene, complete cds
KU761567.1	Russia		Tick-borne encephalitis virus isolate Golubnichiy, complete genome
KU761568.1	Russia		Tick-borne encephalitis virus isolate Primorye-512, complete genome
KU761569.1	Russia		Tick-borne encephalitis virus isolate Primorye-696, complete genome
KU761570.1	Russia		Tick-borne encephalitis virus isolate Primorye-949, complete genome
KU761571.1	Russia		Tick-borne encephalitis virus isolate Primorye-1001, complete genome
KU761572.1	Russia		Tick-borne encephalitis virus isolate Primorye-1035, complete genome
KU761573.1	Russia		Tick-borne encephalitis virus isolate Primorye-1056, complete genome
KU761574.1	Russia		Tick-borne encephalitis virus isolate Primorye-1284, complete genome
KU761575.1	Russia		Tick-borne encephalitis virus isolate Primorye-1285, complete genome
KU761576.1	Russia: Primorskiy kray		Tick-borne encephalitis virus isolate Sofjin-1953, complete genome
KU884607.1	Germany		Tick-borne encephalitis virus isolate Leila-BH95/15, partial genome
KU885457.1	Russia: Leningrad region	Western	Tick-borne encephalitis virus strain Absettarov, complete genome
KX268728.1	Germany		Tick-borne encephalitis virus strain MucAr-HB-171/11, complete genome
KX966398.1	Sweden	European	Tick-borne encephalitis virus strain JP-296, complete genome
KX966399.1	Sweden	European	Tick-borne encephalitis virus strain JP-554, complete genome
KY069119.1	Russia: Irkutsk region	European	Tick-borne encephalitis virus strain 1G-98 polyprotein gene, complete cds
KY069120.1	Russia: Irkutsk region	European	Tick-borne encephalitis virus strain 118-71 polyprotein gene, complete cds
KY069121.1	Russia: Irkutsk region	European	Tick-borne encephalitis virus strain 163-74 polyprotein gene, complete cds
KY069122.1	Russia: Irkutsk region	European	Tick-borne encephalitis virus strain 262-74 polyprotein gene, complete cds
KY069123.1	Russia: Irkutsk region	European	Tick-borne encephalitis virus strain 126-71 polyprotein gene, complete cds
KY069124.1	Russia: Altai region	European	Tick-borne encephalitis virus strain Zmeinogorsk-1 polyprotein gene, complete cds
KY069125.1	Russia: Altai region	European	Tick-borne encephalitis virus strain Zmeinogorsk-5 polyprotein gene, complete cds
KY069126.1	Russia: Altai region	European	Tick-borne encephalitis virus strain Zmeinogorsk-9 polyprotein gene, complete cds
L40361.3			Tick-borne encephalitis virus polyprotein gene, complete cds
LC017691.1	Russia		Tick-borne encephalitis virus RNA, complete genome, strain: IR99-22f7
LC017692.1	Mongolia		Tick-borne encephalitis virus RNA, complete genome, strain: MGL-Selenge-13-12
LC017693.1	Mongolia		Tick-borne encephalitis virus RNA, complete genome, strain: MGL-Selenge-13-14
LC171402.1	Netherlands		Tick-borne encephalitis virus genomic RNA, complete genome, strain: NL
LC440459.1	Japan: Sapporo		Tick-borne encephalitis virus Sapporo-17-Io1 RNA, complete genome
LC440460.1	Japan:Nanporo		Tick-borne encephalitis virus Nanporo-18-44 RNA, complete genome
MF043953.1	Russia		Tick-borne encephalitis virus strain C11-13PAK, complete genome
MF043954.1	Russia		Tick-borne encephalitis virus strain C11-13-293, complete genome
MF043955.1	Russia		Tick-borne encephalitis virus strain C11-13-NEU, complete genome
MF374487.1	Japan	Far eastern	Tick-borne encephalitis virus strain Oshima 5.10 polyprotein gene, complete cds

MF774565.1	Russia: Novosibirk region	Siberian	Tick-borne encephalitis virus isolate TBEV-2871, complete genome
MG210945.1	Hungary	European	Tick-borne encephalitis virus isolate KEM-118, complete genome
MG210946.1	Hungary	European	Tick-borne encephalitis virus isolate KEM-125, complete genome
MG210947.1	Hungary	European	Tick-borne encephalitis virus isolate KEM-127, complete genome
MG210948.1	Hungary	European	Tick-borne encephalitis virus isolate KEM-168, complete genome
MG243699.1	Austria		Tick-borne encephalitis virus strain N5-17/chamois/Austria/2017, complete genome
MG589937.1	Finland	European	Tick-borne encephalitis virus strain Kuutsalo_Human_Cerebellum-2015, complete genome
MG589938.1	Finland	European	Tick-borne encephalitis virus strain Kuutsalo-14_Ixodes_ricinus_Finland-2017, complete genome
MG589939.1	Finland	Siberian	Tick-borne encephalitis virus strain Kuutsalo_2_Human_Cerebellum -2015, complete genome
MG589940.1	Finland	Siberian	Tick-borne encephalitis virus strain Kotka-18_Ixodes_ricinus_Finland-2011, complete genome
MG599476.1	China		Tick-borne encephalitis virus isolate Himalaya-1, complete genome
MG599477.1	China		Tick-borne encephalitis virus isolate Himalaya-2, complete genome
MH021184.1	Netherlands		Tick-borne encephalitis virus strain NL/UH 2016 polyprotein gene, complete cds
MH094241.1	Estonia		Tick-borne encephalitis virus strain Ek-328c, complete genome
MH481364.1	Russia		Tick-borne encephalitis virus strain 110-01 polyprotein gene, complete cds
MH481365.1	Russia		Tick-borne encephalitis virus strain 3094-29_10 polyprotein gene, complete cds
MH481366.1	Russia		Tick-borne encephalitis virus strain 3033-1_10 polyprotein gene, complete cds
MH481367.1	Russia		Tick-borne encephalitis virus strain 3094-9_10 polyprotein gene, complete cds
MH490796.1	Russia		Tick-borne encephalitis virus strain 43-99 polyprotein gene, complete cds
MH490797.1	Russia		Tick-borne encephalitis virus strain 3094-18_10 polyprotein gene, complete cds
MH645612.1	Russia: Irkutsk Region	Siberian	Tick-borne encephalitis virus strain 3869-03, complete genome
MH645613.1	Russia: Irkutsk Region	Siberian	Tick-borne encephalitis virus strain Baikal-3, complete genome
MH645614.1	Russia: Irkutsk Region	Siberian	Tick-borne encephalitis virus strain HimDym-6, complete genome
MH645615.1	Russia: Irkutsk Region	Siberian	Tick-borne encephalitis virus strain Lukovka-3, complete genome
MH645616.1	Bosnia and Herzegovina	Siberian	Tick-borne encephalitis virus strain Bosnia-3, complete genome
MH645617.1	Russia: Western Siberia	Siberian	Tick-borne encephalitis virus strain TBEV-370, partial genome
MH645618.1	Russia: Western Siberia	Siberian	Tick-borne encephalitis virus strain TBEV-2836, complete genome
MH645619.1	Russia: Western Siberia	Siberian	Tick-borne encephalitis virus strain TBEV-2922, complete genome
MK560446.1	Russia: Irkutsk region	European	Tick-borne encephalitis virus strain 172-68, complete genome
MK562430.1	Russia	European	Tick-borne encephalitis virus isolates 214-67, complete genome
MK801803.1	Finland	European	Tick-borne encephalitis virus isolate Isosaari-5-Finland-2005, complete genome
MK801804.1	Finland	European	Tick-borne encephalitis virus isolate Sipoo-4-Finland-2013, complete genome
MK801805.1	Finland	European	Tick-borne encephalitis virus isolate Sipoo-8-Finland-2013, complete genome
MK801806.1	Finland	European	Tick-borne encephalitis virus isolate Sipoo-12-Finland-2013, complete genome
MK801807.1	Finland	European	Tick-borne encephalitis virus isolate Sipoo-22-Finland-2013 polyprotein gene, complete cds

MK801808.1	Finland	European	Tick-borne encephalitis virus isolate Sipoo-23-Finland-2013, complete genome
MK801809.1	Finland	European	Tick-borne encephalitis virus isolate E266-Espoo-Finland-2017, complete genome
MK801810.1	Finland	European	Tick-borne encephalitis virus isolate K12-Espoo-Finland-2018 polyprotein gene, complete cds
MK801811.1	Finland	European	Tick-borne encephalitis virus isolate K13-Espoo-Finland-2018 polyprotein gene, complete cds
MK801812.1	Finland	European	Tick-borne encephalitis virus isolate K14-Espoo-Finland-2018 polyprotein gene, complete cds
MK801813.1	Finland	European	Tick-borne encephalitis virus isolate K15-Espoo-Finland-2018, complete genome
MK801814.1	Finland	European	Tick-borne encephalitis virus isolate K16-Espoo-Finland-2018 polyprotein gene, complete cds
MK922615.1	Germany	European	Tick-borne encephalitis virus strain HB_IF06_8033 polyprotein gene, complete cds
MK922616.1	Germany	European	Tick-borne encephalitis virus strain HB_IF06_8040 polyprotein gene, complete cds
MK922617.1	Germany	European	Tick-borne encephalitis virus strain Rauher_BuschP19_S40 polyprotein gene, complete cds
MN047455.1	Finland	European	Tick-borne encephalitis virus strain A10-Jollas-Finland-2016 polyprotein gene, complete cds
MN114635.1	Russia	Siberian	Tick-borne encephalitis virus strain TSA-18, complete genome
MN114636.1	Russia	Siberian	Tick-borne encephalitis virus strain 1512-18, complete genome
MN114637.1	Russia	Siberian	Tick-borne encephalitis virus strain 1827-18, complete genome
MN115817.1	Russia	Far Eastern	Tick-borne encephalitis virus strain 1020-69, complete genome
MN115818.1	Russia	Siberian	Tick-borne encephalitis virus strain 41-65, complete genome
MN115819.1	Russia	Siberian	Tick-borne encephalitis virus strain 163-64, complete genome
MN115820.1	Russia	Siberian	Tick-borne encephalitis virus strain 518-66, complete genome
MN128700.1	United Kingdom	Western European	Tick-borne encephalitis virus isolate UK-Thetford Forest 2018, complete genome
MN520110.1	Russia	Siberian	Tick-borne encephalitis virus strain 506, complete genome
MN520111.1	Russia	Siberian	Tick-borne encephalitis virus strain 206, complete genome
MN520112.1	Russia	Siberian	Tick-borne encephalitis virus strain 210, complete genome
MN520113.1	Russia	Siberian	Tick-borne encephalitis virus strain 562, complete genome
MN520114.1	Russia	Siberian	Tick-borne encephalitis virus strain 253, complete genome
MN661145.1	United Kingdom	Western European	Tick-borne encephalitis virus isolate UK-Hampshire 2019, complete genome
MN735988.1	Denmark		Tick-borne encephalitis virus isolate DEN19_S3A_Tisvilde, complete genome
MN735989.1	Denmark		Tick-borne encephalitis virus isolate DEN19_S3_Tisvilde, complete genome
MN735990.1	Denmark		Tick-borne encephalitis virus isolate DEN19_S4A_Tisvilde, complete genome
MN735991.1	Denmark		Tick-borne encephalitis virus isolate DEN09_Tokkekoeb, complete genome
MT228625.1			Tick-borne encephalitis virus clone Hypr_variant_40_IRE, complete genome
MT228626.1			Tick-borne encephalitis virus clone Hypr_variant_LT_IRE, complete genome
MT228627.1			Tick-borne encephalitis virus clone parental_Hypr_strain_0, complete genome
MT228628.1			Tick-borne encephalitis virus clone Hypr_variant_40_PS, complete genome
MT311861.1	Austria		Tick-borne encephalitis virus strain Hochosterwitz polyprotein gene, complete cds
MT581212.1	Sweden	Western European	Tick-borne encephalitis virus strain 93/783, complete genome

MT671300.1	Russia	Far Eastern	Tick-borne encephalitis virus strain 95, complete genome
MT671301.1	Russia	Far Eastern	Tick-borne encephalitis virus strain 1024, complete genome
MT671302.1	Russia	Far Eastern	Tick-borne encephalitis virus strain ChB, complete genome
MT974474.1	Mongolia	Siberian	Tick-borne encephalitis virus isolate 92M, complete genome
MW256716.1	Hungary	European	Tick-borne encephalitis virus isolate KEM-1, complete genome
NC_001672.1			Tick-borne encephalitis virus, complete genome
U39292.1			Tick-borne encephalitis virus Hypr polyprotein gene, complete cds
U27491.1			Tick-borne encephalitis virus 263 polyprotein, complete cds
U27495.1			Tick-borne encephalitis virus Neudoerfl polyprotein, complete cds
U39292.1			Tick-borne encephalitis virus Hypr polyprotein gene, complete cds
NC_001809_Outgroup			Louping ill virus, complete genome

Blank (“ “): Data is not available at the NCBI database.

Table A.6. List of 18S-rRNA gene sequence regions of Tick species obtained from NCBI nucleotide database for the MSA and phylogenetic tree construction to identify the Tick species of the analysed tick samples

Acc. No. (NCBI)	Species name	Sequence length	Description
ABV18SRA	*Amblyomma tuberculatum*	793	18S ribosomal RNA (18S rRNA) gene
Z74483	*Rhipicephalus pusillus*	1812	18S rRNA gene
AF018652.1	*Ixodes simplex*	793	B387 18S ribosomal RNA gene, partial sequence
AF115364.1	*Ixodes acutitarsus*	1747	18S ribosomal RNA gene, complete sequence
AF115365.1	*Ixodes corwini*	1747	18S ribosomal RNA gene, complete sequence
AF115366.1	*Ixodes pilosus*	1748	18S ribosomal RNA gene, complete sequence
AF115367.1	*Ixodes luciae*	1786	18S ribosomal RNA gene, complete sequence
AF115368.1	*Ixodes tasmani*	1787	18S ribosomal RNA gene, complete sequence
AF115369.1	*Ixodes uriae*	1785	18S ribosomal RNA gene, complete sequence
AY274888.1	*Ixodes persulcatus*	1786	18S ribosomal RNA gene, partial sequence
DCN18SR	*Dermacentor andersoni*	1784	18S nuclear ribosomal RNA (18S rRNA) gene
HHY18SR	*Haemaphysalis inermis*	1771	18S nuclear ribosomal RNA (18S rRNA) gene
IXO18SR	*Ixodes affinis*	1766	18S ribosomal RNA (18S rRNA) gene
IXO18SRB	*Ixodes kopsteini*	800	18S ribosomal RNA (18S rRNA) gene
JN018307.1	*Ixodes hexagonus*	1772	Voucher MNHN-JAA3 18S ribosomal RNA gene, partial sequence
KR002681.1	*Ixodes pavlovskyi*	1643	18S ribosomal RNA gene, partial sequence
KX712275.1	*Ixodes tovari*	1787	Voucher USDA:NVSLcase_T06_740 18S ribosomal RNA gene, partial sequence
L76351.1	*Ixodes cookei*	1783	18S ribosomal RNA (18S rRNA) gene
NMM18SR	*Aponomma latum*	336	18S ribosomal RNA (18S rRNA) gene
YAL18SR	*Hyalomma dromedarii*	1786	18S ribosomal RNA (18S rRNA) gene
Z74497.1	*Ixodes ricinus*	1813	18S ribosomal RNA (18S rRNA) gene
AGG18SRA	*Argas lahorensis*	1789	18S ribosomal RNA (18S rRNA) gene

Table A.7. Semi-quantitative real-time PCR analysis of cDNA from seven different RNA extracts amplified using five different primer compositions with two NTCs and two negative controls (Hela cell culture RNA and RNA from sample number 6) (Threshold value: 0.02)

Sample number	**Sample name**	**Template for the RT-qPCR**	**Ct value of RNA samples**	**Ct values of the cDNA products derived through SuperScript™ IV first strand synthesis system with different primer composition**				
				Random Hexamers	**JK_1-F#**	**JK_7-R#**	**Eur_39-R#**	**Mix of primers*#**
1	Hochosterwitz_Trizol_11-04-2016	cDNA		14.89	13.81	14.43	12.60	20.30
2	Hochosterwitz_Qiagen_11-04-2016	cDNA		16.41	17.31	14.71	14.02	20.10
3	Hochosterwitz_RLT_28-04-2016	cDNA		14.56	15.89	13.94	13.73	16.53
4	Hochosterwitz_250416_26-04-2016	cDNA		12.22	13.07	11.99	10.79	9.99
5	Hochosterwitz_2_AVL_28-04-2016	cDNA		11.22	12.56	12.51	11.45	10.07
6	Hochosterwitz_17-06-2021_MAJD	cDNA		10.31	10.63	9.50	8.84	7.67
7	Hochosterwitz_2_extracted_25-04-2016	cDNA		10.88	11.79	11.96	10.85	9.98
8	NTC	cDNA		--	--	--	--	18.74
9	NTC	cDNA		--	--	--	16.20	26.93
10	NTC	cDNA		--	--	--	--	--
11	Hela RNA	cDNA		--	--	--	--	--
12	TICOVAC-vaccine, Pfizer Inc. (5.0 µL)			32.41				
13	Hochosterwitz_ 17-06-2021_MAJD	RNA	--					

**Mix of primers [JK_1-F (3.3 µM), JK_7-R (3.3 µM), Eur_39-R (3.3 µM)],#Detailed description of the primers used in the study are tabulated in the Table A.2, Annexure I, TBE-320F [0.3 µM, 1628-1645 position at TBEV genome (NCBI) Acc.No:KF991107], NTC: Nontemplate control samples (--): no Ct value, Blank (" "): Wasn't performed or not applicable.*

Table A.8. Concentrations of purified DNA amplified against different concentrations of primer pairs using TBEV-Hochosterwitz cDNA

Composition of the template RNA in the RT-PCR reaction	Primer concentration (µM)	Primer pair used in the PCR	JK_1-F/R	JK_3-F/R	JK_7-F/R	Eur_1-F/R	Eur_19-F/R	Eur_39-F/R
		Amplified region#	**38-1998**	**3511-5343**	**9447-11197**	**35-395**	**5157-5575**	**10721-11076**
		Amplified fragment size (b)	**1961**	**1833**	**1852**	**361**	**419**	**365**
		Dilution factor (log10)	**dsDNA concentration (ng/µL)**					
TBEV-HOC (1.0 µL)	0.15	0	3.720	5.630	3.400	2.290	2.140	0.950
TBEV-HOC (1.0 µL)	0.15	-1	3.940	4.200	3.220	2.870	1.650	0.560
TBEV-HOC (1.0 µL)	0.15	-2	3.370	2.000	1.930	2.010	2.130	0.050
TBEV-HOC (1.0 µL)	0.15	-3	0.718	0.083	0.082	0.178	0.117	0.700
TBEV-HOC (1.0 µL)	0.15	-4	1.630	0.928	0.518	1.880	0.785	0.750
TBEV-HOC (1.0 µL)	0.15	-5	0.780	1.050	0.274	1.010	0.698	1.070
TBEV-HOC (1.0 µL)	0.15	-6	0.257	0.884	0.110	0.533	1.150	0.933
TBEV-HOC (1.0 µL)	0.15	-7	0.429	0.572	0.214	0.762	0.085	0.845
TBEV-HOC (1.0 µL)	0.15	-8	0.288	0.397	0.074	0.095	0.117	0.404
TBEV-HOC (1.0 µL)	0.15	-9	0.000	0.123	0.204	0.137	0.205	0.334
TBEV-HOC (1.0 µL)	0.15	-10	0.330	0.130	0.077	0.034	0.960	0.315
NTC	0.15	H_2O	0.000	0.000	0.061	0.000	0.275	0.077
TBEV-HOC (1.0 µL)	0.5	-2	26.203	23.901	16.701		23.814	10.412
TBEV-HOC (1.0 µL)	0.5	-3	31.901	23.12	14.001		18.802	8.14
TBEV-HOC (1.0 µL)	0.5	-4	14.711	4.161	7.691		18.303	2.15
TBEV-HOC (1.0 µL)	0.5	-5	8.514	1.634	5.593		15.629	0.412
TBEV-HOC (1.0 µL)	0.5	-6	0.321	0.982	1.814		6.192	0.364
NTC	0.5	H_2O	0.104	0.091	0.000		0.000	0.000

(“‘) Blank: Did not perform, #: Reference to the TBEV-Hochosterwitz sequence (NCBI Acc. No. MT311861), TBEV—HOC: TBEV-Hochosterwitz strain

Table A.9. Concentrations of purified DNA amplified against different primer pairs using cDNA of TBEV-Hochosterwitz RNA reverse transcribed with the addition of total RNA extract from tick samples

Composition of the template RNA in the RT-PCR reaction	**Primer concentration (µM)**	**Primer pair used in the PCR**	**JK_1-F/R**	**JK_3-F/R**	**JK_7-F/R**	**Eur_1-F/R**	**Eur_19-F/R**	**Eur_39-F/R**
		Amplified region#	**38-1998**	**3511-5343**	**9447-11197**	**35-395**	**5157-5575**	**10721-11076**
		Amplified fragment size (b)	**1961**	**1833**	**1852**	**361**	**419**	**365**
		Dilution factor (log10)	**dsDNA concentration (ng/µL)**					
TBEV-HOC (1.0 µL)	0.50	-4	20.900	12.500	8.320	2.860	6.620	6.250
TBEV-HOC (1.0 µL)	0.50	-5	0.063	2.520	5.330	2.590	3.410	2.510
TBEV-HOC (1.0 µL)	0.50	-6	0.000	0.062	2.690	1.840	0.556	0.203
TBEV-HOC (1.0 µL)	0.50	-7	0.051	0.059	0.000	0.562	0.296	0.162
TBEV-HOC (1.0 µL)	0.50	-8	0.202	0.085	0.053	0.000	0.059	0.058
NTC	0.50	-	0.000	0.000	0.000	0.000	0.065	0.090
TBEV-HOC (1 µL) + TICK$^{(A)}$ RNA (1.0 µL)	0.50	-4	15.700	12.300	11.300	20.400	15.800	3.970
TBEV-HOC (1 µL) + TICK$^{(A)}$ RNA (1.0 µL)	0.50	-5	7.640	5.220	7.740	11.300	4.190	2.500
TBEV-HOC (1 µL) + TICK$^{(A)}$ RNA (1.0 µL)	0.50	-6	0.914	0.116	4.070	9.130	1.250	0.739
TBEV-HOC (1 µL) + TICK$^{(A)}$ RNA (1.0 µL)	0.50	-7	5.270	0.159	0.772	1.580	0.168	0.485
TBEV-HOC (1 µL) + TICK$^{(A)}$ RNA (1.0 µL)	0.50	-8	0.209	0.091	2.210	0.421	0.083	0.054
TICK$^{(A)}$ RNA (1.0 µL)	0.50	-	0.208	0.087	2.800	0.267	0.242	0.087
TICK$^{(B)}$ RNA (1.0 µL)	0.50	-	18.000	0.076	0.789	1.420	0.264	0.117

(-): No dilutions, Tick$^{(A)}$ RNA: RNA extract from a tick sample considered to be negative for TBEV infections. Tick$^{(B)}$ RNA: RNA extract from a tick sample considered to be suspected for TBEV infections, NTC: Nontemplate control. #: Reference to the TBEV-Hochosterwitz sequence (NCBI Acc. No. MT311861), TBEV—HOC: TBEV-Hochosterwitz strain

Table A.10. Concentrations of dsDNA in purified PCR products amplified against different primer schemes using cDNA of TBEV-Hochosterwitz strain

Sequencing experiment No.	Primer scheme	Dilution factor (log10)	-3	-4	-5	-6	-7	-8	NTC
		cDNA composition	dsDNA concentration (ng/µL)						
1	(M)	TBEV-HOC		11.9	0.153				0
1	(N)	TBEV-HOC		15.4	0.057				0
1	(A)	TBEV-HOC		17	6.63	2.09	0.501		0
1	(B)	TBEV-HOC		23.7	5.47	1.76	0.975		0.304
2	(A+O)	TBEV-HOC	10.8	2.9	1.23	0.335	0.448	0.314	0.26
2	(B+P)	TBEV-HOC	18.2	3.82	3.75	1.14	0.508	0.361	0.373
2	(C+O)	TBEV-HOC	11.8	3	2.03	0.5	0.372	0.247	0.413
2	(D+P)	TBEV-HOC	15.7	4.23	1.73	0.829	0.579	0.313	0.43
2	(2A+O)	TBEV-HOC	5.68	2.61	1.38	0.897	0.333	0.269	0.219
2	(2B+P)	TBEV-HOC	7.39	13.9	4.51	0.582	0.527	0.307	0.469
3	(A)	TBEV-HOC	5.29	3.68	1.24	0.241	0.096	0.101	0.114
3	(B)	TBEV-HOC	8.11	1.7	0.896	0.271	0.123	0.098	0.122
3	(2E+Q)	TBEV-HOC		24.4	10.1	0.838			0.435
3	(2F+R)	TBEV-HOC		38.4	20	4.14			0.82
3	(2E+2Q)	TBEV-HOC		29.6	5.95	3.45			0.273
3	(2F+2R)	TBEV-HOC		57	22.3	8.52			3.12
3	(G+S)	TBEV-HOC		23.8	10.9	2.48			0.533
3	(H+T)	TBEV-HOC		39.7	18.9	6.65			1.06
3	(2G+2S)	TBEV-HOC		23.4	9.36	3.52			0.449
3	(2H+2T)	TBEV-HOC		37.7	12.2	9.91			1.23
6	(G+S)[##]	TBEV-HOC			5.24	7.85			1.18
6	(H+T)[##]	TBEV-HOC			3.40	3.80			0.817

[#:] Reference to the TBEV-Hochosterwitz sequence (NCBI Acc. No. MT311861), (" ") Blank: Not done, TBEV—HOC: TBEV-Hochosterwitz strain

Table A.11. Concentrations of dsDNA purified PCR products, amplified against different primer schemes using cDNA of tick samples in the sequencing experiment number 4, 5 & 6

Sample	Primer scheme	dsDNA concentration (ng/µL)	Sample	Primer scheme	dsDNA concentration (ng/µL)	Sample	Primer scheme	dsDNA concentration (ng/µL)
SL1(21)-13F	(A)	0.574	SL1(21)-04M	(2C+O)	1.14	SL1(21)-03N	(G+S)	1.16
SL1(21)-13F	(B)	0.66	SL1(21)-04M	(2D+P)	1.81	SL1(21)-03N	(H+T)	2.03
SL1(21)-13F	(2C+O)	1.02	SL1(21)-04M	(G+S)	1.12	SL1(21)-17N	(A)	0.446
SL1(21)-13F	(2D+P)	4.07	SL1(21)-04M	(H+T)	2.18	SL1(21)-17N	(B)	0.639
SL1(21)-13F	(G+S)	0.724	SL1(21)-24M	(A)	0.342	SL1(21)-17N	(2C+O)	1.1
SL1(21)-13F	(H+T)	2.99	SL1(21)-24M	(B)	0.493	SL1(21)-17N	(2D+P)	1.42
SL1(21)-42F	(A)	0.647	SL1(21)-24M	(2C+O)	1.03	SL1(21)-17N	(G+S)	1.11
SL1(21)-42F	(B)	0.879	SL1(21)-24M	(2D+P)	1.12	SL1(21)-17N	(H+T)	1.94
SL1(21)-42F	(2C+O)	1.17	SL1(21)-24M	(G+S)	0.983	SL1(21)-22N	(A)	0.491
SL1(21)-42F	(2D+P)	1.75	SL1(21)-24M	(H+T)	1.35	SL1(21)-22N	(B)	0.531
SL1(21)-42F	(G+S)	0.755	SL1(21)-33M	(A)	0.833	SL1(21)-22N	(2C+O)	1.1
SL1(21)-42F	(H+T)	1.84	SL1(21)-33M	(B)	0.451	SL1(21)-22N	(2D+P)	2.16
SL1(21)-24F	(A)	0.419	SL1(21)-33M	(2C+O)	1.05	SL1(21)-22N	(G+S)	1.26
SL1(21)-24F	(B)	0.425	SL1(21)-33M	(2D+P)	0	SL1(21)-22N	(H+T)	1.9
SL1(21)-24F	(2C+O)	1.15	SL1(21)-33M	(G+S)	0.624	SL1(21)-30N	(A)	2.71
SL1(21)-24F	(2D+P)	2.04	SL1(21)-33M	(H+T)	1.15	SL1(21)-30N	(B)	0.594
SL1(21)-24F	(G+S)	1.01	SL1(21)-35M	(A)	0.31	SL1(21)-30N	(2C+O)	4.55
SL1(21)-24F	(H+T)	1	SL1(21)-35M	(B)	0.706	SL1(21)-30N	(2D+P)	2.19
SL1(21)-32F	(A)	0.861	SL1(21)-35M	(2C+O)	1.64	SL1(21)-30N	(G+S)	5.39
SL1(21)-32F	(B)	0.33	SL1(21)-35M	(2D+P)	1.5	SL1(21)-30N	(H+T)	1.69
SL1(21)-32F	(2C+O)	1.64	SL1(21)-35M	(G+S)	0.397	SL1(21)-41N	(A)	1.71
SL1(21)-32F	(2D+P)	0.786	SL1(21)-35M	(H+T)	1.63	SL1(21)-41N	(B)	0.587
SL1(21)-32F	(G+S)	2.19	SL1(21)-38M	(A)	0.494	SL1(21)-41N	(2C+O)	7.17
SL1(21)-32F	(H+T)	0.648	SL1(21)-38M	(B)	0.54	SL1(21)-41N	(2D+P)	1.57
SL1(21)-17F	(A)	0.699	SL1(21)-38M	(2C+O)	1.57	SL1(21)-41N	(G+S)	3.38

Table A.11. Concentrations of dsDNA purified PCR products, amplified against different primer schemes using cDNA of tick samples in the sequencing experiment number 4, 5 & 6, cont.

Sample	Primer scheme	dsDNA concentration (ng/µL)	Sample	Primer scheme	dsDNA concentration (ng/µL)	Sample	Primer scheme	dsDNA concentration (ng/µL)
SL1(21)-17F	(B)	1.61	SL1(21)-38M	(2D+P)	1.96	SL1(21)-41N	(H+T)	1.72
SL1(21)-17F	(2C+O)	1.05	SL1(21)-38M	(G+S)	0.973	SL1(21)-58N	(A)	1.28
SL1(21)-17F	(2D+P)	1.39	SL1(21)-38M	(H+T)	1.16	SL1(21)-58N	(B)	0.585
SL1(21)-17F	(G+S)	0.833	SL1(21)-30M	(A)	0.193	SL1(21)-58N	(2C+O)	7.58
SL1(21)-17F	(H+T)	1.08	SL1(21)-30M	(B)	0.461	SL1(21)-58N	(2D+P)	1.94
SL1(21)-32M	(A)	0.397	SL1(21)-30M	(2C+O)	1.24	SL1(21)-58N	(G+S)	6.45
SL1(21)-32M	(B)	0.289	SL1(21)-30M	(2D+P)	1.61	SL1(21)-58N	(H+T)	1.3
SL1(21)-32M	(2C+O)	0.9	SL1(21)-30M	(G+S)	0.84	SL1(21)-60N	(A)	1.88
SL1(21)-32M	(2D+P)	1.11	SL1(21)-30M	(H+T)	1.62	SL1(21)-60N	(B)	0.671
SL1(21)-32M	(G+S)	1.3	SL1(21)-42M	(A)	0.472	SL1(21)-60N	(2C+O)	6.32
SL1(21)-32M	(H+T)	1.17	SL1(21)-42M	(B)	0.477	SL1(21)-60N	(2D+P)	1.6
SL1(21)-42M	(A)	0.472	SL1(21)-42M	(2C+O)	1.32	SL1(21)-60N	(G+S)	6.69
SL1(21)-42M	(B)	0.477	SL1(21)-42M	(2D+P)	1.36	SL1(21)-60N	(H+T)	1.59
SL1(21)-42M	(2C+O)	1.32	SL1(21)-42M	(G+S)	1.18	SL1(21)-64N	(A)	2.23
SL1(21)-42M	(2D+P)	1.36	SL1(21)-42M	(H+T)	2.1	SL1(21)-64N	(B)	0.456
SL1(21)-42M	(G+S)	1.18	SL1(21)-01N	(A)	0.392	SL1(21)-64N	(2C+O)	3.54
SL1(21)-42M	(H+T)	2.1	SL1(21)-01N	(B)	0.663	SL1(21)-64N	(2D+P)	1.24
SL1(21)-03M	(A)	0.592	SL1(21)-01N	(2C+O)	0.853	SL1(21)-64N	(G+S)	6.68
SL1(21)-03M	(B)	0.374	SL1(21)-01N	(2D+P)	1.92	SL1(21)-64N	(H+T)	1.78
SL1(21)-03M	(2C+O)	0.977	SL1(21)-01N	(G+S)	0.925	SL1(21)-67N	(A)	1.67
SL1(21)-03M	(2D+P)	1.34	SL1(21)-01N	(H+T)	2.11	SL1(21)-67N	(B)	0.488
SL1(21)-03M	(G+S)	0.918	SL1(21)-03N	(A)	0.415	SL1(21)-67N	(2C+O)	4.46
SL1(21)-03M	(H+T)	1.65	SL1(21)-03N	(B)	0.569	SL1(21)-67N	(2D+P)	1.72
SL1(21)-04M	(A)	0.388	SL1(21)-03N	(2C+O)	0.89	SL1(21)-67N	(G+S)	4.12
SL1(21)-04M	(B)	0.536	SL1(21)-03N	(2D+P)	2.24	SL1(21)-67N	(H+T)	1.4

Table A.11. Concentrations of dsDNA purified PCR products, amplified against different primer schemes using cDNA of tick samples in the sequencing experiment number 7

Sample	**Primer scheme**	**dsDNA concentration (ng/µL)**	**Primer scheme**	**dsDNA concentration (ng/µL)**	**Primer scheme**	**dsDNA concentration (ng/µL)**	**Primer scheme**	**dsDNA concentration (ng/µL)**
SL1(21)-01N	(2C+2O)	1.73	(2D+2P)	1.96	(2G+2S)	1.07	(2H+2T)	1.55
SL1(21)-17N	(2C+2O)	1.52	(2D+2P)	4.86	(2G+2S)	1.00	(2H+2T)	1.67
SL1(21)-22N	(2C+2O)	0.75	(2D+2P)	1.96	(2G+2S)	0.66	(2H+2T)	1.56
SL1(21)-67N	(2C+2O)	1.00	(2D+2P)	1.78	(2G+2S)	0.82	(2H+2T)	1.15
SL1(21)-17F	(2C+2O)	1.52	(2D+2P)	1.71	(2G+2S)	1.12	(2H+2T)	2.28
SL1(21)-24F	(2C+2O)	0.81	(2D+2P)	1.92	(2G+2S)	0.77	(2H+2T)	0.937
SL1(21)-27F	(2C+2O)	1.11	(2D+2P)	1.47	(2G+2S)	1.06	(2H+2T)	1.5
SL1(21)-32F	(2C+2O)	1.05	(2D+2P)	2.03	(2G+2S)	0.90	(2H+2T)	1.29
SL1(21)-42F	(2C+2O)	2.44	(2D+2P)	1.16	(2G+2S)	0.74	(2H+2T)	1.10
SL1(21)-03M	(2C+2O)	0.73	(2D+2P)	1.20	(2G+2S)	0.69	(2H+2T)	1.01
SL1(21)-04M	(2C+2O)	0.72	(2D+2P)	0.72	(2G+2S)	0.50	(2H+2T)	1.32
SL1(21)-33M	(2C+2O)	0.56	(2D+2P)	4.99	(2G+2S)	2.91	(2H+2T)	0.64
NTC	(2C+2O)	0.39	(2D+2P)	0.35	(2G+2S)	0.71	(2H+2T)	0.61
NTC	(2C+2O)	0.32	(2D+2P)	0.41	(2G+2S)	0.51	(2H+2T)	.054

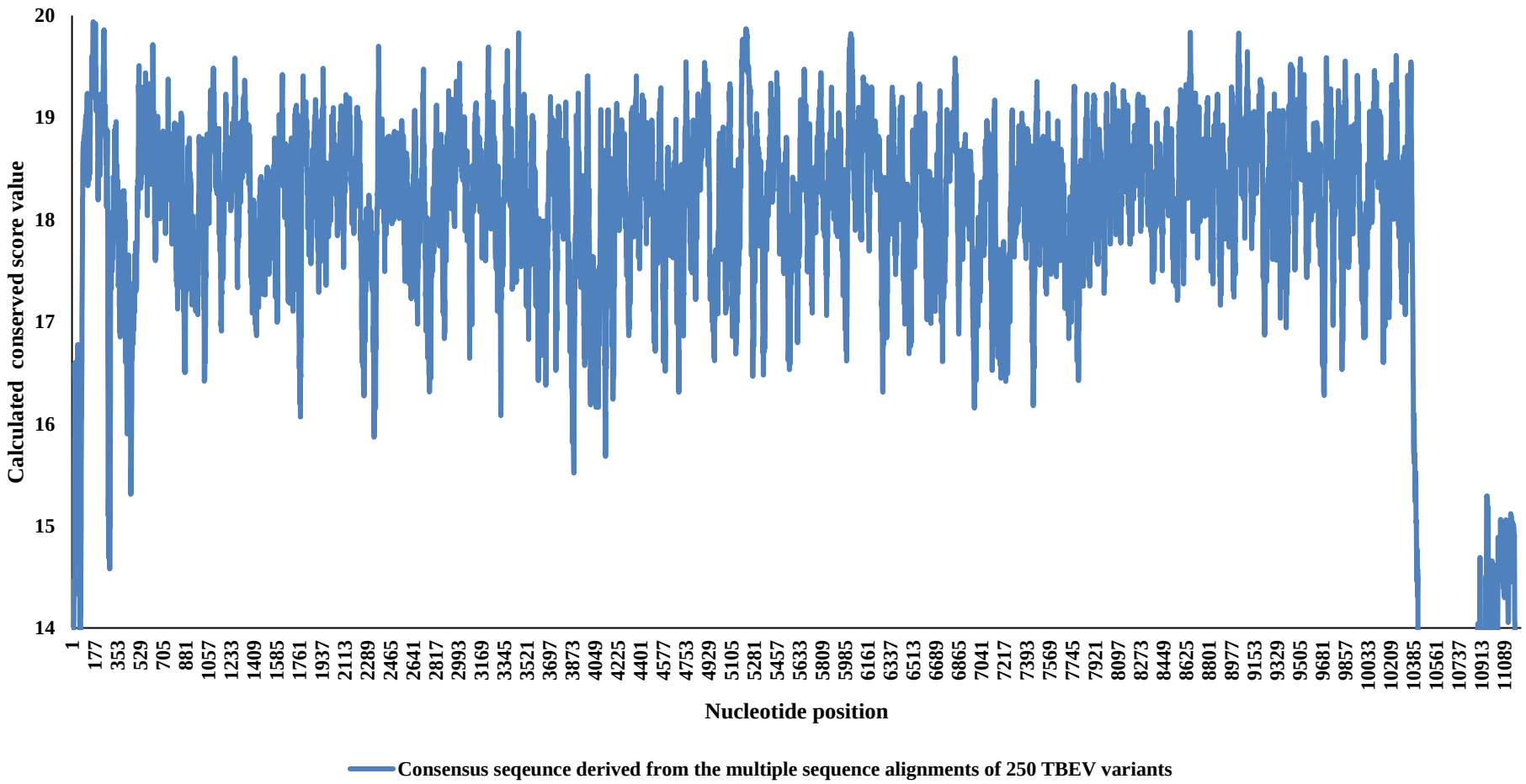


Figure A.2. Variations in the calculated conserved score values (Fragments length: 20 b) of the consensus sequence derived from the multiple sequence alignment of 250 different variants of TBEV.

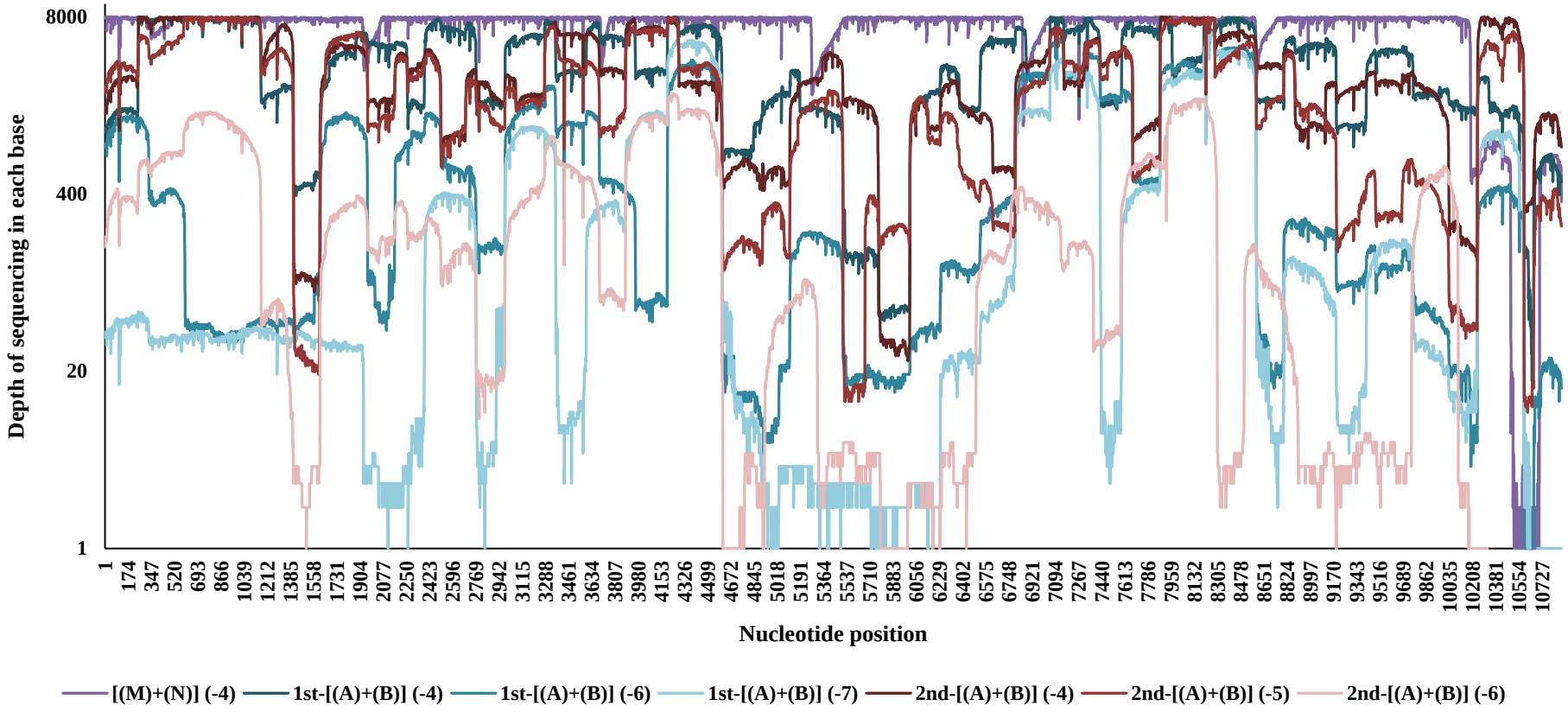


Figure A.3. Sequencing depth of each base of TBEV-Hochosterwitz strain calculated using SAMtools (Danecek *et al.*, 2021) *for two different combined primer schemes at three different dilutions in two replicated experiments.(1st and 2nd are the sequencing experiment numbers and the (-4—7) are the dilution factors.*

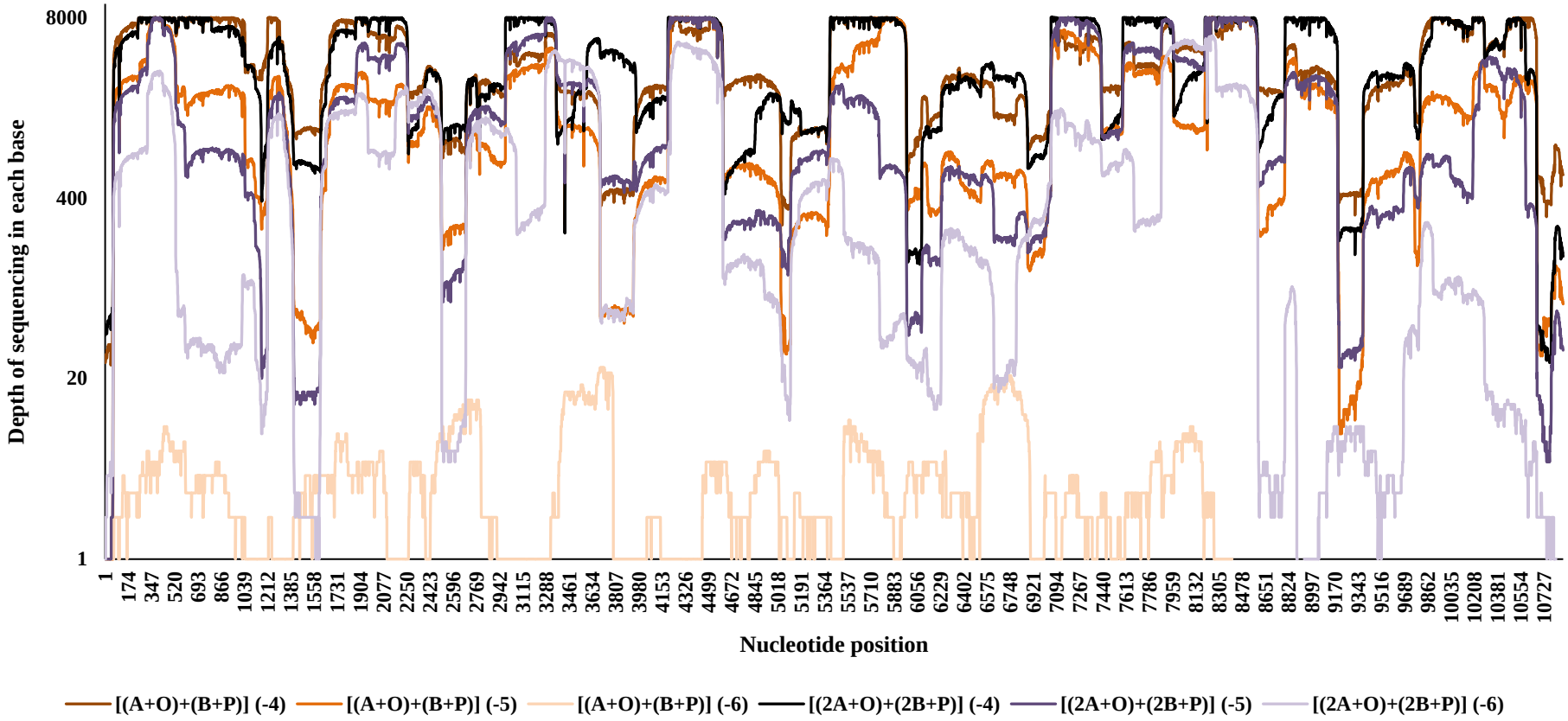


Figure A.4 Sequencing depth of each base of TBEV-Hochosterwitz strain calculated using SAMtools (Danecek *et al.*, 2021) *for a combined primer schemes at two different primer concentrations in each dilution levels of initial template RNA.*

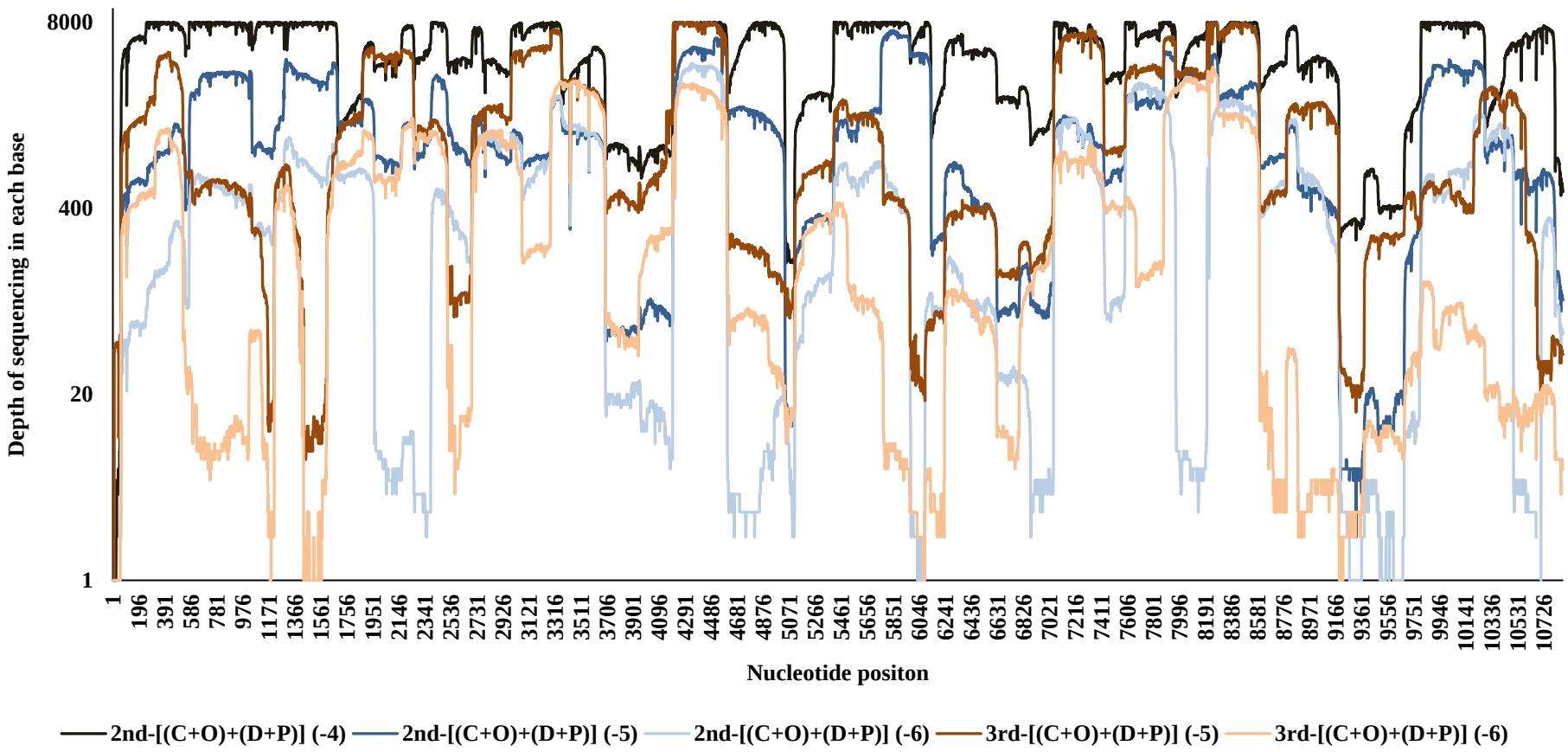


Figure A.5. Sequencing depth of each base of TBEV-Hochosterwitz strain calculated using SAMtools (Danecek *et al.*, 2021) *for a combined primer schemes in two replicated experiments. (2nd and 3rd are the sequencing experiment numbers and (-4—6) are the dilution factors.*

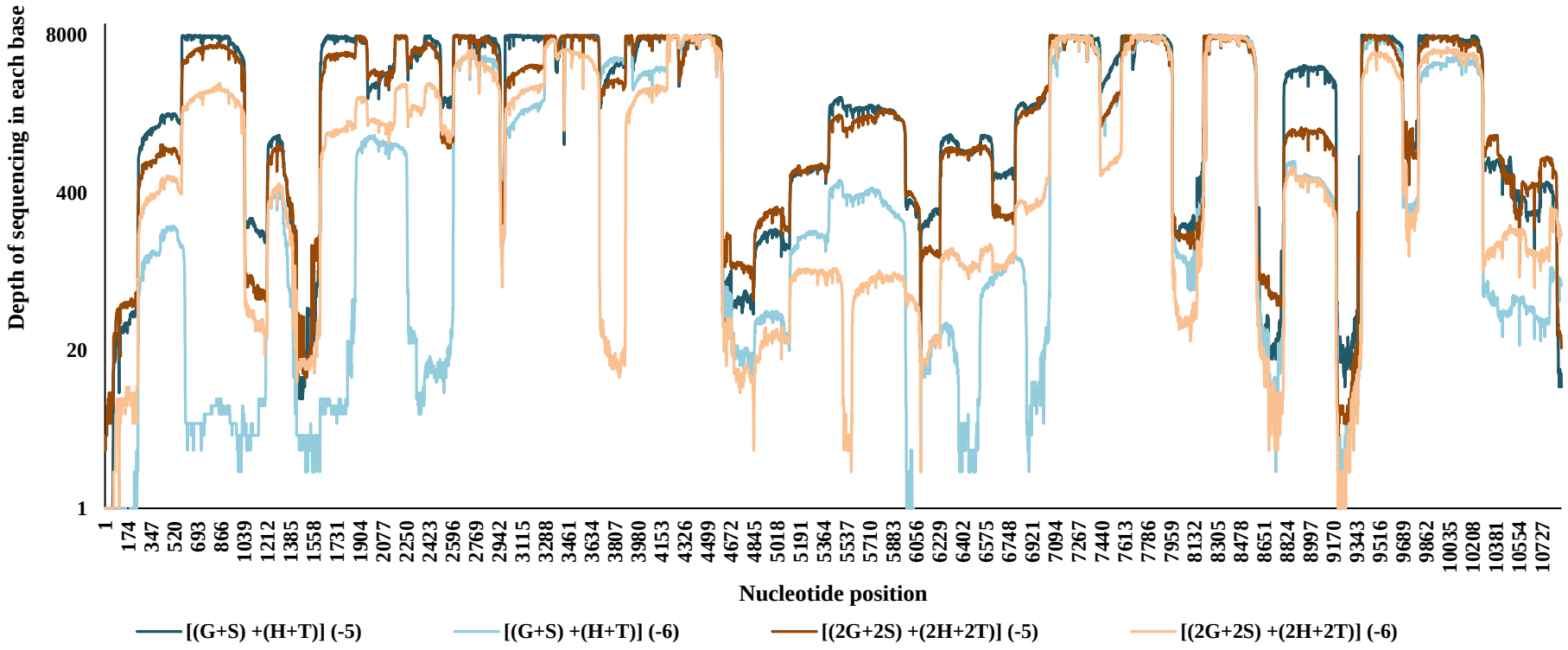


Figure A.6. Sequencing depth of each base of TBEV-Hochosterwitz strain calculated using SAMtools (Danecek *et al.*, 2021) *for a combined primer schemes at two different primer concentrations in each dilution levels of initial template RNA.*

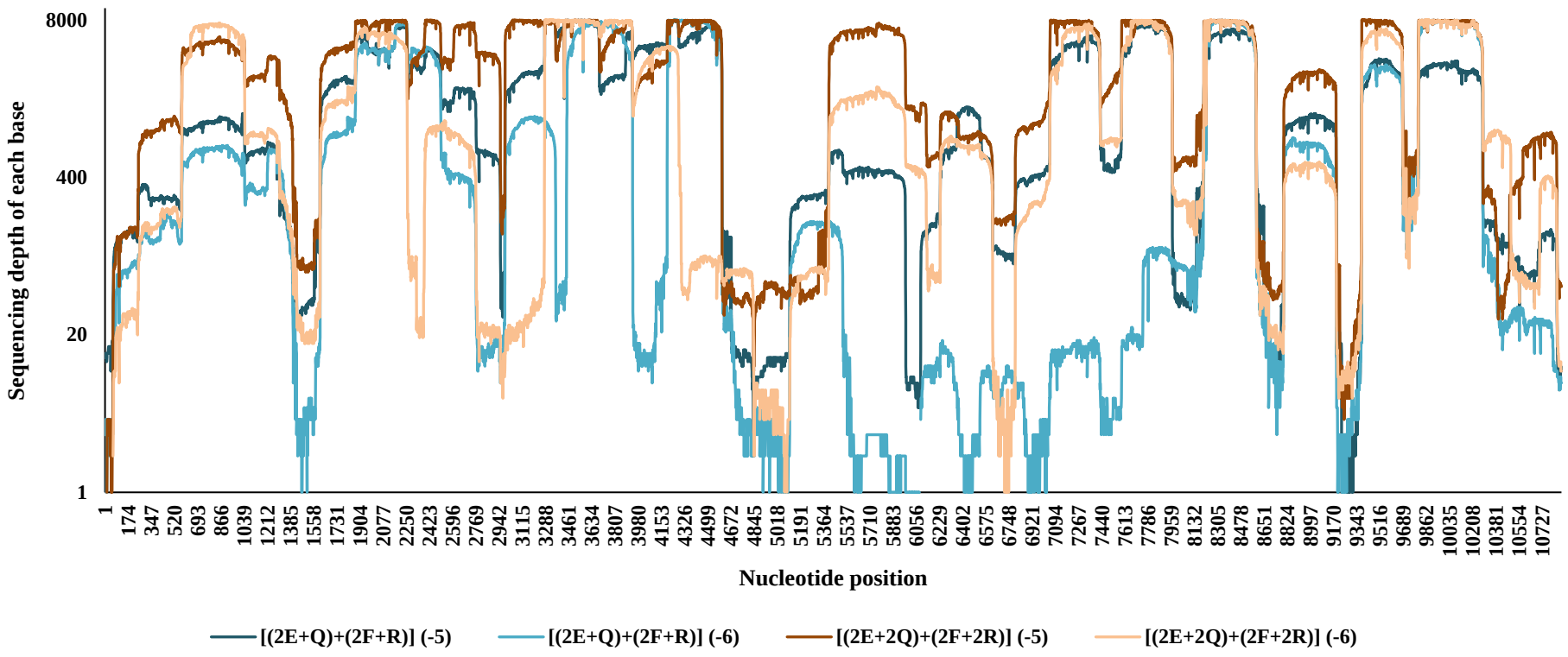


Figure A.7. Sequencing depth of each base of TBEV-Hochosterwitz strain calculated using SAMtools (Danecek *et al.*, 2021) *for a combined primer schemes at two different primer concentrations in each dilution levels.*

Table A.12. Expected PCR fragments length of each primer pair used in the study

Primer Name	Primer position#		Length of the amplified region	Primer Name	Primer position#		Length of the amplified region	Primer Name	Primer position#		Length of the amplified region
	Start	End			Start	End			Start	End	
Eur_1-F/R/R	37	400	363	Eur_24-F/R	6585	6954	369	CJB_1-F/R	109	616	507
Eur_2-F/R/R	297	670	373	Eur_25-F/R	6850	7234	384	CJB_3-F/R	624	1115	491
Eur_3-F/R/R	565	960	395	Eur_26-F/R	7108	7505	397	CJB_4-F/R	1070	1360	290
Eur_4-F/R/R	851	1237	386	Eur_27-F/R	7363	7747	384	CJB_5-F/R	1262	1474	212
Eur_5-F/R	1108	1478	370	Eur_28-F/R	7641	8043	402	CJB_6-F/R	1343	1365	22
Eur_6-F/R	1371	1758	387	Eur_29-F/R	7930	8360	430	CJB_15-F/R	3993	4357	364
Eur_7-F/R	1657	2033	376	Eur_30-F/R	8257	8671	414	CJB_17-F/R	4558	5112	554
Eur_8-F/R	1921	2330	409	Eur_31-F/R	8572	8960	388	CJB_18-F/R	4897	5262	365
Eur_9-F/R	2213	2581	368	Eur_32-F/R	8851	9268	417	CJB_20-F/R	5456	6051	595
Eur_10-F/R	2436	2844	408	Eur_33-F/R	9164	9560	396	CJB_21-F/R	5810	6208	398
Eur_11-F/R	2742	3143	401	Eur_34-F/R	9465	9840	375	CJB_22-F/R	6142	6447	305
Eur_12-F/R	3037	3439	402	Eur_35-F/R/R	9735	10109	374	CJB_34-F/R	9435	9977	542
Eur_13-F/R	3333	3765	432	Eur_36-F/R/R	10010	10408	398	CJB_36-F/R	9858	10362	504
Eur_14-F/R	3644	4034	390	Eur_37-F/R/R	10299	10765	466	CJB_37-F/R	10252	11012	760
Eur_15-F/R	3939	4357	418	Eur_38-F/R/R	10635	11086	451	CJB_38-F/R	10963	11118	155
Eur_16-F/R	4251	4679	428	Eur_39-F/R	10815	11207	392	CJB_39-F/R	11004	11188	184
Eur_17-F/R	4565	4994	429	JK_1-F/R	37	2004	1967	CJB_11-F/R	2653	3026	373
Eur_18-F/R	4894	5262	368	JK_2-F/R	1747	3777	2030	CJB_14-F/R	3500	4014	514
Eur_19-F/R	5164	5583	419	JK_3-F/R	3518	5351	1833	CJB_17-NF	4415	4917	502
Eur_20-F/R	5486	5850	364	JK_4-F/R	5084	6927	1843	CJB_29-F/R	7802	8277	475
Eur_21-F/R	5738	6162	424	JK_5-F/R	6680	8669	1989	CJB_33-F/R	9034	9530	496
Eur_22-F/R	6062	6453	391	JK_6-F/R	8393	10270	1877	CJB_34-NR	9745	11031	1286
Eur_23-F/R	6287	6703	416	JK_7-F/R	9356	11086	1730				

Table A.13. Calculated primer pair efficiency in each primer scheme derived through the Oxford Nanopore sequence data

Sequencing experiment No.	**2**				**3**				
	Overlapping primer pair scheme								
Primer pair	**(A+O) +(B+P)**	**(C+O) +(D+P)**	**(2A+O) +(2B+P)**	**(A)+(B)**	**(C+O) +(D+P)**	**(2E+Q) +(2F+R)**	**(2E+2Q) +(2F+2R)**	**(G+S) +(H+T)**	**(2G+2S) +(2H+2T)**
	Primer pair efficiencies								
Eur_1_L/R	240.0		348.4	575.2		66.9	22.8	32.8	103.1
Eur_2_L/R	374.2	-127.1	377.9	1863.0	97.0	103.9	67.9	-5.6	263.7
Eur_3_L/R	375.6		103.1	1000.2					
Eur_4_L/R	68.8		3.9	678.2					
Eur_5_L/R	177.7		130.5	1684.0					
Eur_6_L/R	212.3		316.4	535.7				30.2	97.2
Eur_7_L/R	122.2	483.3	697.4	1243.7	396.0	493.0	936.2	-818.2	568.7
Eur_8_L/R	600.1	406.7	1490.7	991.9	374.2	4112.8	4552.4	361.4	1304.7
Eur_9_L/R	414.8	317.6	1678.9	1510.3	947.9	4181.9	36.1	-724.8	1372.0
Eur_10_L/R	463.8	1034.4	1021.4	1201.2	358.4	1386.7	501.8	-598.6	1561.0
Eur_11_L/R	388.2	972.6	1249.0	896.3	941.6				
Eur_12_L/R	1184.5	409.9	1363.5	910.2	-55.6	640.4	-778.6	283.8	1886.2
Eur_13_L/R	1164.2	1461.1	4129.4	2907.4	2506.2	-877.4	10975.8	8655.0	4147.7
Eur_14_L/R	60.1	175.4	751.8	669.6	636.1				827.4
Eur_15_L/R	3.2		63.4	3027.4				2384.6	1269.3
Eur_16_L/R	7713.8	2796.3	11473.4	3203.2	834.0	9750.4	-1777.6	16755.1	17324.0
Eur_17_L/R	1194.5	949.6	2745.8	749.4	236.9				
Eur_18_L/R	20.6		32.2	14.8				21.3	-5.8
Eur_19_L/R	61.1	40.1	381.4	342.8	115.9	95.7	46.6	93.3	27.6
Eur_20_L/R	1191.8		549.2	362.6				196.8	-97.6
Eur_21_L/R	553.4		-48.1	181.7				68.3	-1.1
Eur_22_L/R	106.8		21.7	804.7					
Eur_23_L/R	198.1	147.8	315.3	349.4	79.9	-50.4	739.2	-93.3	40.1
Eur_24_L/R	183.6	-11.9	48.8	126.7	25.4	-7.4	78.4	8.7	48.9
Eur_25_L/R	61.3	-7.9	141.5	947.1	84.1	-34.2	87.4	-147.9	132.9
Eur_26_L/R	2594.4	1107.6	2947.0	2511.9	374.4	-372.3	2148.4	4279.2	6680.9
Eur_27_L/R	798.7	318.9	1373.8	1680.2	267.0	-190.7	1562.2	4037.7	3306.2
Eur_28_L/R	1229.0	1862.3	1624.3	753.7	-158.3	-438.0	3763.8	3128.7	5607.7
Eur_29_L/R	794.2	1310.8	4374.9	3138.6	2023.6				
Eur_30_L/R	3476.4	1450.7	5522.9	3473.3	1021.2	6839.0	11569.8	14771.0	7640.6

Table A.13. Calculated each primer pair efficiency in each primer scheme derived through the Oxford Nanopore sequence data cont..

Sequencing experiment No.	**2**				**3**				
	Overlapping primer pair scheme								
Primer pair	**(A+O) +(B+P)**	**(C+O) +(D+P)**	**(2A+O) +(2B+P)**	**(A)+(B)**	**(C+O) +(D+P)**	**(2E+Q) +(2F+R)**	**(2E+2Q) +(2F+2R)**	**(G+S) +(H+T)**	**(2G+2S) +(2H+2T)**
	Primer pair efficiencies								
Eur_31_L/R	1258.4	604.1	1646.1	1154.1	-17.6	1506.1	2341.6	3157.1	2133.8
Eur_32_L/R	1677.7	1000.3	736.0	1552.9	-157.4	716.7	274.1	347.4	511.9
Eur_33_L/R	125.4	19.6	88.9	52.6	-22.8				
Eur_34_L/R	159.4		-6.7	140.9					
Eur_35_L/R	114.8	23.8	152.9	189.9	-6.1	113.1	82.8	175.4	115.7
Eur_36_L/R	137.3		-121.9	15.9					
Eur_37_L/R	601.6		1317.8	1824.1					
Eur_38_L/R	622.9		505.7	1934.9					
Eur_39_L/R	122.4		-8.1	106.2					
CJB_1_F/R	923.0	-4.7	1057.2		254.7	58.6	11.1	-2.0	3.9
CJB_3_F/R	412.4	820.7	-252.6		-38.8	487.2	3528.4	-1319.9	1731.4
CJB_4_F/R	234.7	8.1	134.5		35.3	278.7	746.3	-45.6	50.8
CJB_5_F/R	823.4	204.3	1516.9		490.4	515.6	207.2	302.6	350.6
CJB_6_F/R	442.1	1127.4	605.2		166.9	90.9	39.2	45.9	170.8
CJB_15_F/R	84.7	2.3	412.7		139.9	-524.1	5892.2		1694.7
CJB_17_F/R	1226.3	975.2	2830.7		238.8				
CJB_18_F/R	21.5	24.7	31.0		7.4	1.2	0.0	21.3	-5.8
CJB_20_F/R	1037.4	425.9	542.4		15.1	-3.6	949.9	173.8	-85.1
CJB_21_F/R	1739.3	1969.3	-159.2		-25.0	-6.3	540.9	37.8	-17.3
CJB_22_F/R	175.8	318.2	26.2		-2.0	-4.6	40.8	-8.7	2.9
CJB_34_F/R	135.4	-18.9	-3.3		-10.8				
CJB_36_F/R	659.5	1027.6	190.3		75.1	7980.9	8372.9	3218.7	4891.4
CJB_37_F/R	399.4	1176.1	405.9		-111.1	1986.1	1782.1	733.0	886.8
CJB_38_F/R	15.0	148.5	17.5		-3.0	-0.2	38.8	80.6	249.8
CJB_39_F/R	12.4	72.8	17.3		-3.8	1.2	-0.8	83.8	250.1
CJB_11_NF/NR						58.6	-239.2	3375.7	3288.4
CJB_14_NF/NR						4803.1	20501.1	8945.3	488.9
CJB_17_NF/NR						4493.2	-864.1	5822.3	6406.6
CJB_29_NF/NR						-101.7	3434.9	3728.4	5042.1
CJB_33_NF/NR						237.2	58.8	35.3	74.8
CJB_34_NF/NR						2943.3	4590.1	5036.1	4002.3
CJB_37_F/NR						1986.1	1782.1	733.0	886.8

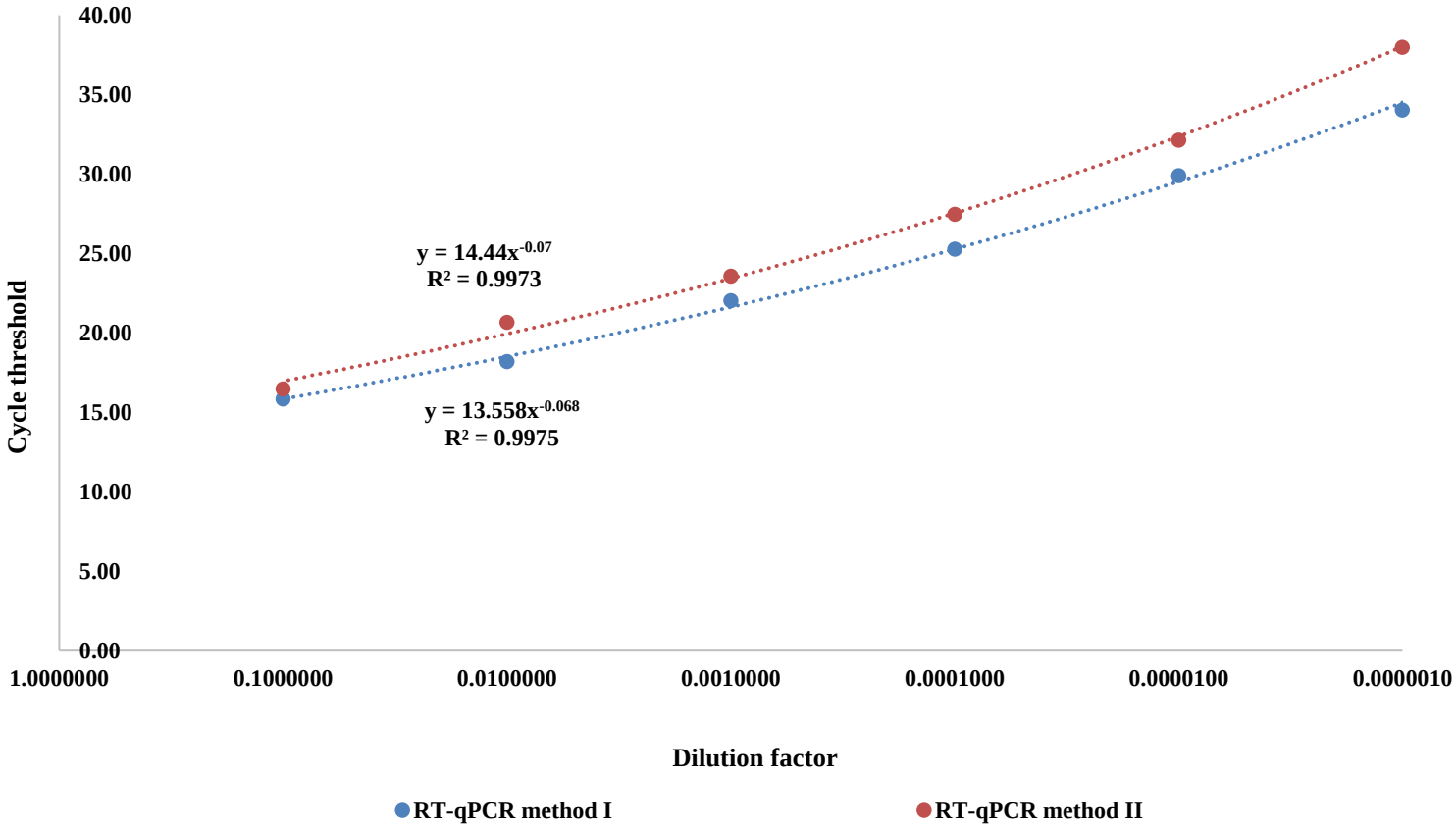


Figure A.8. Cyclic threshold values received by each semiquantitative real-time-PCR reaction method I and method II by the tenfold dilutions of TBEV-Hochosterwitz RNA.

Table A.14. Cycle threshold (Ct) values received for the total RNA extracts of tick samples (172) by two similar semiquantitative real-time PCR analysis and the single factor ANOVA analysis for the determination of variance between two methods

Sample name	Method II	Method I	ANOVA: Single Factor (Summery)				
	Ct	Ct	Groups	Count	Sum	Average	Variance
SL1(21)-1F	NEG	NEG	Row 1	2	0	0	0
SL1(21)-2F	NEG	NEG	Row 2	2	0	0	0
SL1(21)-3F	NEG	NEG	Row 3	2	0	0	0
SL1(21)-4F	NEG	NEG	Row 4	2	0	0	0
SL1(21)-5F	NEG	NEG	Row 5	2	0	0	0
SL1(21)-6F	NEG	NEG	Row 6	2	0	0	0
SL1(21)-7F	NEG	NEG	Row 7	2	0	0	0
SL1(21)-8F	NEG	NEG	Row 8	2	0	0	0
SL1(21)-9F	NEG	NEG	Row 9	2	0	0	0
SL1(21)-10F	6.36	NEG	Row 10	2	6.36	3.18	20.2248
SL1(21)-11F	NEG	NEG	Row 11	2	0	0	0
SL1(21)-12F	NEG	NEG	Row 12	2	0	0	0
SL1(21)-13F	NEG	NEG	Row 13	2	0	0	0
SL1(21)-14F	33.69	NEG	Row 14	2	33.69	16.845	567.5081
SL1(21)-15F	NEG	NEG	Row 15	2	0	0	0
SL1(21)-16F	NEG	NEG	Row 16	2	0	0	0
SL1(21)-17F	NEG	NEG	Row 17	2	0	0	0
SL1(21)-18F	NEG	NEG	Row 18	2	0	0	0
SL1(21)-19F	NEG	NEG	Row 19	2	0	0	0
SL1(21)-20F	NEG	NEG	Row 20	2	0	0	0

Table A.14. Cycle threshold (Ct) values received for the total RNA extracts of tick samples (172) by two similar semiquantitative real-time PCR analysis and the single factor ANOVA analysis for the determination of variance between two methods cont.

Sample name	Method II	Method I	ANOVA: Single Factor (Summery)				
	Ct	Ct	Groups	Count	Sum	Average	Variance
SL1(21)-21F	32.64	NEG	Row 21	2	32.64	16.32	532.6848
SL1(21)-22F	NEG	NEG	Row 22	2	0	0	0
SL1(21)-23F	NEG	NEG	Row 23	2	0	0	0
SL1(21)-24F	35.89	NEG	Row 24	2	35.89	17.945	644.0461
SL1(21)-25F	NEG	NEG	Row 25	2	0	0	0
SL1(21)-26F	NEG	NEG	Row 26	2	0	0	0
SL1(21)-27F	NEG	NEG	Row 27	2	0	0	0
SL1(21)-28F	NEG	NEG	Row 28	2	0	0	0
SL1(21)-29F	NEG	NEG	Row 29	2	0	0	0
SL1(21)-30F	NEG	NEG	Row 30	2	0	0	0
SL1(21)-31F	NEG	NEG	Row 31	2	0	0	0
SL1(21)-32F	NEG	NEG	Row 32	2	0	0	0
SL1(21)-33F	NEG	NEG	Row 33	2	0	0	0
SL1(21)-34F	NEG	NEG	Row 34	2	0	0	0
SL1(21)-35F	NEG	NEG	Row 35	2	0	0	0
SL1(21)-36F	NEG	NEG	Row 36	2	0	0	0
SL1(21)-37F	NEG	NEG	Row 37	2	0	0	0
SL1(21)-38F	NEG	NEG	Row 38	2	0	0	0
SL1(21)-39F	NEG	NEG	Row 39	2	0	0	0
SL1(21)-40F	NEG	NEG	Row 40	2	0	0	0
SL1(21)-41F	NEG	NEG	Row 41	2	0	0	0

Table A.14. Cycle threshold (Ct) values received for the total RNA extracts of tick samples (172) by two similar semiquantitative Realtime PCR analysis and the single factor ANOVA analysis for the determination of variance between two methods cont.

Sample name	Method II	Method I	ANOVA: Single Factor (Summery)				
	Ct	Ct	Groups	Count	Sum	Average	Variance
SL1(21)-42F	NEG	35.14	Row 42	2	35.14	17.57	617.4098
SL1(21)-43F	NEG	NEG	Row 43	2	0	0	0
SL1(21)-44F	NEG	NEG	Row 44	2	0	0	0
SL1(21)-45F	NEG	NEG	Row 45	2	0	0	0
SL1(21)-46F	NEG	NEG	Row 46	2	0	0	0
SL1(21)-47F	NEG	NEG	Row 47	2	0	0	0
SL1(21)-48F	35.89	32.2	Row 48	2	68.09	34.045	6.80805
SL1(21)-1M	NEG	NEG	Row 49	2	0	0	0
SL1(21)-2M	NEG	NEG	Row 50	2	0	0	0
SL1(21)-3M	6.73	NEG	Row 51	2	6.73	3.365	22.64645
SL1(21)-4M	6.88	NEG	Row 52	2	6.88	3.44	23.6672
SL1(21)-5M	NEG	NEG	Row 53	2	0	0	0
SL1(21)-6M	NEG	NEG	Row 54	2	0	0	0
SL1(21)-7M	NEG	NEG	Row 55	2	0	0	0
SL1(21)-8M	NEG	NEG	Row 56	2	0	0	0
SL1(21)-9M	NEG	NEG	Row 57	2	0	0	0
SL1(21)-10M	NEG	NEG	Row 58	2	0	0	0
SL1(21)-11M	NEG	NEG	Row 59	2	0	0	0
SL1(21)-12M	NEG	NEG	Row 60	2	0	0	0
SL1(21)-13M	NEG	NEG	Row 61	2	0	0	0
SL1(21)-14M	NEG	NEG	Row 62	2	0	0	0

Table A.14. Cycle threshold (Ct) values received for the total RNA extracts of tick samples (172) by two similar semiquantitative Realtime PCR analysis and the single factor ANOVA analysis for the determination of variance between two methods cont.

Sample name	Method II	Method I	ANOVA: Single Factor (Summery)				
	Ct	Ct	Groups	Count	Sum	Average	Variance
SL1(21)-15M	NEG	NEG	Row 63	2	0	0	0
SL1(21)-16M	NEG	NEG	Row 64	2	0	0	0
SL1(21)-17M	NEG	NEG	Row 65	2	0	0	0
SL1(21)-18M	NEG	NEG	Row 66	2	0	0	0
SL1(21)-19M	NEG	NEG	Row 67	2	0	0	0
SL1(21)-20M	NEG	NEG	Row 68	2	0	0	0
SL1(21)-21M	NEG	NEG	Row 69	2	0	0	0
SL1(21)-22M	NEG	NEG	Row 70	2	0	0	0
SL1(21)-23M	NEG	NEG	Row 71	2	0	0	0
SL1(21)-24M	31.74	NEG	Row 72	2	31.74	15.87	503.7138
SL1(21)-25M	NEG	NEG	Row 73	2	0	0	0
SL1(21)-26M	NEG	NEG	Row 74	2	0	0	0
SL1(21)-27M	NEG	NEG	Row 75	2	0	0	0
SL1(21)-28M	NEG	NEG	Row 76	2	0	0	0
SL1(21)-29M	NEG	31.26	Row 77	2	31.26	15.63	488.5938
SL1(21)-30M	NEG	NEG	Row 78	2	0	0	0
SL1(21)-31M	NEG	33.47	Row 79	2	33.47	16.735	560.1205
SL1(21)-32M	36.77	34.69	Row 47	2	71.46	35.73	2.1632
SL1(21)-33M	NEG	33.17	Row 81	2	33.17	16.585	550.1245
SL1(21)-34M	NEG	NEG	Row 82	2	0	0	0
SL1(21)-35M	NEG	NEG	Row 83	2	0	0	0

Table A.14. Cycle threshold (Ct) values received for the total RNA extracts of tick samples (172) by two similar semiquantitative Realtime PCR analysis and the single factor ANOVA analysis for the determination of variance between two methods cont.

Sample name	Method II	Method I	ANOVA: Single Factor (Summery)				
	Ct	Ct	Groups	Count	Sum	Average	Variance
SL1(21)-36M	NEG	34.59	Row 84	2	34.59	17.295	598.2341
SL1(21)-37M	NEG	NEG	Row 85	2	0	0	0
SL1(21)-38M	NEG	NEG	Row 86	2	0	0	0
SL1(21)-39M	NEG	NEG	Row 87	2	0	0	0
SL1(21)-40M	NEG	NEG	Row 88	2	0	0	0
SL1(21)-41M	NEG	NEG	Row 89	2	0	0	0
SL1(21)-42M	NEG	NEG	Row 90	2	0	0	0
SL1(21)-43M	NEG	35.14	Row 91	2	35.14	17.57	617.4098
SL1(21)-44M	NEG	NEG	Row 92	2	0	0	0
SL1(21)-45M	NEG	NEG	Row 93	2	0	0	0
SL1(21)-46M	NEG	NEG	Row 94	2	0	0	0
SL1(21)-47M	NEG	NEG	Row 95	2	0	0	0
SL1(21)-48M	NEG	NEG	Row 96	2	0	0	0
SL1(21)-49M	NEG	NEG	Row 97	2	0	0	0
SL1(21)-50M	33.38	NEG	Row 98	2	33.38	16.69	557.1122
SL1(21)-51M	NEG	NEG	Row 99	2	0	0	0
SL1(21)-52M	NEG	NEG	Row 100	2	0	0	0
SL1(21)-53M	NEG	NEG	Row 101	2	0	0	0
SL1(21)-54M	NEG	NEG	Row 102	2	0	0	0
SL1(21)-1N	NEG	34.07	Row 103	2	34.07	17.035	580.3825
SL1(21)-2N	NEG	NEG	Row 104	2	0	0	0

Table A.14. Cycle threshold (Ct) values received for the total RNA extracts of tick samples (172) by two similar semiquantitative Realtime PCR analysis and the single factor ANOVA analysis for the determination of variance between two methods cont.

Sample name	Method II	Method I	ANOVA: Single Factor (Summery)				
	Ct	Ct	Groups	Count	Sum	Average	Variance
SL1(21)-3N	33.39	NEG	Row 105	2	33.39	16.695	557.4461
SL1(21)-4N	NEG	NEG	Row 106	2	0	0	0
SL1(21)-5N	NEG	NEG	Row 107	2	0	0	0
SL1(21)-6N	NEG	NEG	Row 108	2	0	0	0
SL1(21)-8N	NEG	NEG	Row 109	2	0	0	0
SL1(21)-9N	NEG	NEG	Row 110	2	0	0	0
SL1(21)-10N	NEG	NEG	Row 111	2	0	0	0
SL1(21)-11N	NEG	NEG	Row 112	2	0	0	0
SL1(21)-12N	NEG	NEG	Row 113	2	0	0	0
SL1(21)-13N	NEG	NEG	Row 114	2	0	0	0
SL1(21)-14N	NEG	NEG	Row 115	2	0	0	0
SL1(21)-15N	NEG	NEG	Row 116	2	0	0	0
SL1(21)-16N	NEG	NEG	Row 117	2	0	0	0
SL1(21)-17N	NEG	30.72	Row 118	2	30.72	15.36	471.8592
SL1(21)-18N	NEG	NEG	Row 119	2	0	0	0
SL1(21)-19N	NEG	NEG	Row 120	2	0	0	0
SL1(21)-20N	NEG	NEG	Row 121	2	0	0	0
SL1(21)-21N	NEG	NEG	Row 122	2	0	0	0
SL1(21)-22N	NEG	35.52	Row 123	2	35.52	17.76	630.8352
SL1(21)-23N	NEG	NEG	Row 124	2	0	0	0
SL1(21)-24N	NEG	NEG	Row 125	2	0	0	0

Table A.14. Cycle threshold (Ct) values received for the total RNA extracts of tick samples (172) by two similar semiquantitative Realtime PCR analysis and the single factor ANOVA analysis for the determination of variance between two methods cont.

Sample name	Method II	Method I	ANOVA: Single Factor (Summery)				
	Ct	Ct	Groups	Count	Sum	Average	Variance
SL1(21)-25N	NEG	NEG	Row 126	2	0	0	0
SL1(21)-26N	NEG	NEG	Row 127	2	0	0	0
SL1(21)-27N	NEG	NEG	Row 128	2	0	0	0
SL1(21)-28N	NEG	NEG	Row 129	2	0	0	0
SL1(21)-29N	NEG	NEG	Row 130	2	0	0	0
SL1(21)-30N	NEG	35.4	Row 131	2	35.4	17.7	626.58
SL1(21)-31N	NEG	NEG	Row 132	2	0	0	0
SL1(21)-32N	NEG	NEG	Row 133	2	0	0	0
SL1(21)-33N	NEG	NEG	Row 134	2	0	0	0
SL1(21)-34N	NEG	NEG	Row 135	2	0	0	0
SL1(21)-35N	NEG	NEG	Row 136	2	0	0	0
SL1(21)-36N	NEG	NEG	Row 137	2	0	0	0
SL1(21)-37N	NEG	NEG	Row 138	2	0	0	0
SL1(21)-38N	NEG	NEG	Row 139	2	0	0	0
SL1(21)-39N	NEG	NEG	Row 140	2	0	0	0
SL1(21)-40N	NEG	NEG	Row 141	2	0	0	0
SL1(21)-41N	40.72	35.44	Row 142	2	76.16	38.08	13.9392
SL1(21)-42N	NEG	NEG	Row 143	2	0	0	0
SL1(21)-43N	NEG	NEG	Row 144	2	0	0	0
SL1(21)-44N	NEG	NEG	Row 145	2	0	0	0
SL1(21)-45N	NEG	NEG	Row 146	2	0	0	0

Table A.14. Cycle threshold (Ct) values received for the total RNA extracts of tick samples (172) by two similar semiquantitative Realtime PCR analysis and the single factor ANOVA analysis for the determination of variance between two methods cont.

Sample name	Method II	Method I	ANOVA: Single Factor (Summery)				
	Ct	Ct	Groups	Count	Sum	Average	Variance
SL1(21)-46N	NEG	NEG	Row 147	2	0	0	0
SL1(21)-47N	NEG	NEG	Row 148	2	0	0	0
SL1(21)-48N	NEG	NEG	Row 149	2	0	0	0
SL1(21)-49N	NEG	NEG	Row 150	2	0	0	0
SL1(21)-50N	NEG	NEG	Row 151	2	0	0	0
SL1(21)-51N	NEG	NEG	Row 152	2	0	0	0
SL1(21)-52N	NEG	NEG	Row 153	2	0	0	0
SL1(21)-53N	NEG	NEG	Row 154	2	0	0	0
SL1(21)-54N	NEG	NEG	Row 155	2	0	0	0
SL1(21)-55N	NEG	NEG	Row 156	2	0	0	0
SL1(21)-56N	NEG	NEG	Row 157	2	0	0	0
SL1(21)-57N	NEG	NEG	Row 158	2	0	0	0
SL1(21)-58N	NEG	32.05	Row 159	2	32.05	16.025	513.6013
SL1(21)-59N	NEG	NEG	Row 160	2	0	0	0
SL1(21)-60N	34.24	NEG	Row 161	2	34.24	17.12	586.1888
SL1(21)-61N	NEG	NEG	Row 162	2	0	0	0
SL1(21)-62N	NEG	NEG	Row 163	2	0	0	0
SL1(21)-63N	NEG	NEG	Row 164	2	0	0	0
SL1(21)-64N	NEG	35.21	Row 165	2	35.21	17.605	619.8721
SL1(21)-65N	NEG	NEG	Row 166	2	0	0	0
SL1(21)-66N	NEG	NEG	Row 167	2	0	0	0

Table A.14. Cycle threshold (Ct) values received for the total RNA extracts of tick samples (172) by two similar semiquantitative Realtime PCR analysis and the single factor ANOVA analysis for the determination of variance between two methods cont.

Sample name	Method II	Method I	ANOVA: Single Factor (Summery)				
	Ct	Ct	Groups	Count	Sum	Average	Variance
SL1(21)-68N	NEG	NEG	Row 169	2	0	0	0
SL1(21)-69N	NEG	NEG	Row 170	2	0	0	0
SL1(21)-70N	NEG	NEG	Row 171	2	0	0	0
SL1(21)-71N	NEG	NEG	Row 172	2	0	0	0
SL1(21)-72N	NEG	NEG	Row 173	2	0	0	0
SL1(21)-73N	NEG	NEG	Row 174	2	0	0	0
SL1(21)-74N	NEG	NEG	Row 175	2	0	0	0
SL1(21)-75N	NEG	NEG	Row 176	2	0	0	0
SL1(21)-76N	NEG	NEG	Row 177	2	0	0	0

ANOVA single factor, between each sample (Rows)

Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	16479.836	175.000	94.170	1.519	0.003	1.283
Within Groups	10913.171	176.000	62.007			
Total	27393.007	351.000				

SUMMARY — **ANOVA: Single Factor, between two methods**

Groups	Count	Sum	Average	Variance
Column 1	176.000	368.320	2.093	66.045
Column 2	176.000	508.070	2.887	90.169

Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	55.483	1.000	55.483	0.710	0.400	3.868
Within Groups	27337.524	350.000	78.107			
Total	27393.007	351.000				

Table A.15. Cycle threshold (Ct) values received for the tenfold serial dilutions of TBEV-Hochosterwitz RNA by two similar semiquantitative real-time PCR analysis and the single factor ANOVA analysis for the determination of variance between two methods cont.

TBEV-Hochosterwitz RNA dilution factor (-log10)	**Method II**	**Method I**	**ANOVA: Single Factor (Summery)**				
	Ct	**Ct**	**Groups**	**Count**	**Sum**	**Average**	**Variance**
1	14.41	15.85	1	2	30.26	15.13	1.0368
2	16.48	18.19	2	2	34.67	17.335	1.46205
3	20.67	22.03	3	2	42.7	21.35	0.9248
4	23.57	25.28	4	2	48.85	24.425	1.46205
5	27.46	29.89	5	2	57.35	28.675	2.95245
6	32.13	34.04	6	2	66.17	33.085	1.82405
7	37.98	40.54	7	2	78.52	39.26	3.2768

ANOVA single factor, between each sample (Rows)

Source of Variation	**SS**	**df**	**MS**	**F**	**P-value**	**F crit**
Between Groups	898.9004	6	149.8167	81.05086	4.12E-06	3.865968853
Within Groups	12.939	7	1.848429			
Total	911.8394	13				

SUMMARY	**ANOVA: Single Factor, between two methods**			
Groups	**Count**	**Sum**	**Average**	**Variance**
Column 1	0.1	6	0.011111	0.001852
Column 2	14.41	6	158.29	26.38167

Source of Variation	**SS**	**df**	**MS**	**F**	**P-value**	**F crit**
Between Groups	3004.147	2	1502.074	35.0279	2.23E-06	3.682320344
Within Groups	643.233	15	42.8822			
Total	3647.38	17				

Table A.16. Overview of the sequenced data mapped against the reference sequence of Ixodes ricinus 18S-rRNA gene (NCBI Acc. No. Z74479.1) for the identification of TBEV positive tick species

Sequencing experiment number	Sample name	DNA library barcode number	Primers used in cDNA synthesis	Primer scheme composition*	Assembly or mapping technique	Number of sequences			Pairwise identity (%)	Mean Coverage above 20 reads
						Total	Mapped	Unmapped		
7	SL1(21)-01N	1	Rand-Hex	(2C+2O)+(2D+2P)+(2G+2S)+(2H+2T)	Minimap2	164001	3804	160197	97.9	515.5
7	SL1(21)-17N	2	Rand-Hex	(2C+2O)+(2D+2P)+(2G+2S)+(2H+2T)	Minimap2	280006	19314	260692	97.7	2801.0
7	SL1(21)-22N	3	Rand-Hex	(2C+2O)+(2D+2P)+(2G+2S)+(2H+2T)	Minimap2	164001	5416	158585	99.7	945.0
7	SL1(21)-67N	4	Rand-Hex	(2C+2O)+(2D+2P)+(2G+2S)+(2H+2T)	Minimap2	124001	7678	116323	99.9	1224.1
7	SL1(21)-17F	5	Rand-Hex	(2C+2O)+(2D+2P)+(2G+2S)+(2H+2T)	Minimap2	256005	5925	250080	99.9	961.9
7	SL1(21)-24F	6	Rand-Hex	(2C+2O)+(2D+2P)+(2G+2S)+(2H+2T)	Minimap2	132001	6051	125950	99.9	910.4
7	SL1(21)-27F	7	Rand-Hex	(2C+2O)+(2D+2P)+(2G+2S)+(2H+2T)	Minimap2	152001	3634	148367	99.9	586.6
7	SL1(21)-32F	8	Rand-Hex	(2C+2O)+(2D+2P)+(2G+2S)+(2H+2T)	Minimap2	128002	10513	117489	97.6	1752.5
7	SL1(21)-42F	9	Rand-Hex	(2C+2O)+(2D+2P)+(2G+2S)+(2H+2T)	Minimap2	68001	3614	64387	99.9	684.6
7	SL1(21)-03M	10	Rand-Hex	(2C+2O)+(2D+2P)+(2G+2S)+(2H+2T)	Minimap2	64001	180	63821	56.3	23.2
7	SL1(21)-04M	11	Rand-Hex	(2C+2O)+(2D+2P)+(2G+2S)+(2H+2T)	Minimap2	132001	13644	118357	97.1	2972.7
7	SL1(21)-33M	12	Rand-Hex	(2C+2O)+(2D+2P)+(2G+2S)+(2H+2T)	Minimap2	160001	182	159819	47.4	20.6

* *Composition of the primer schemes according to the table A.2, Annexure I.*

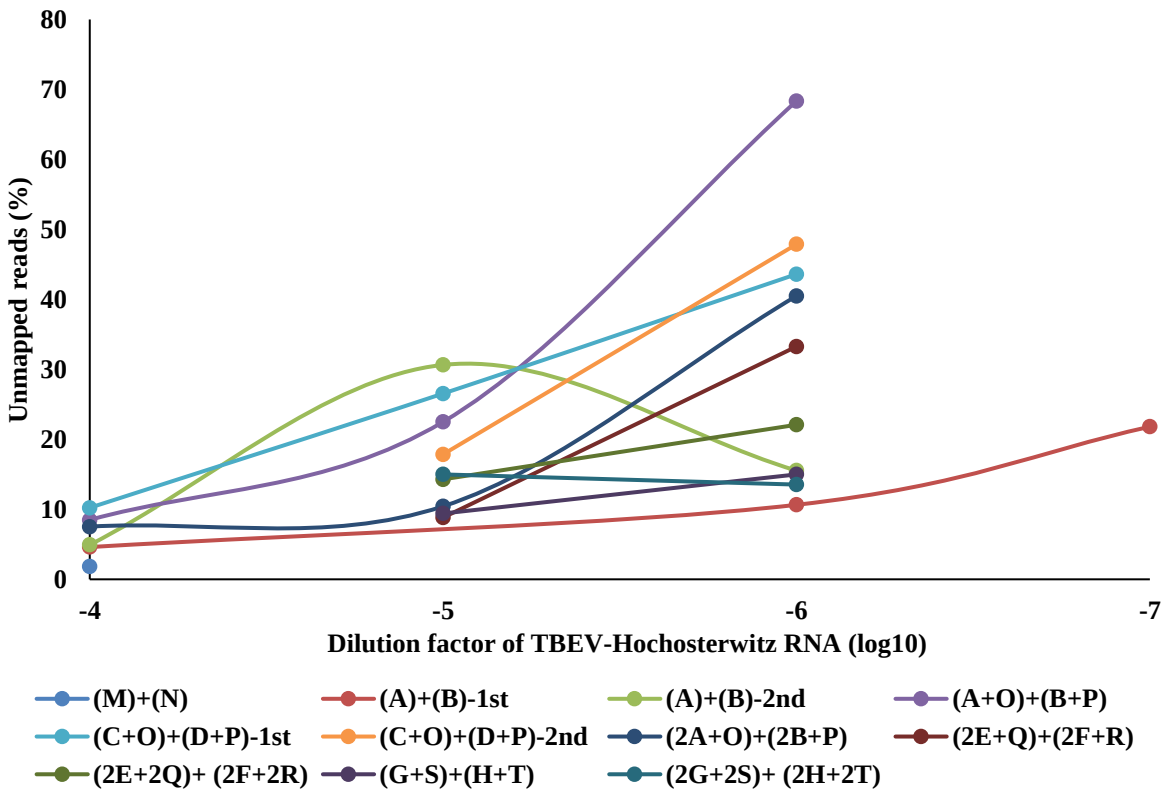
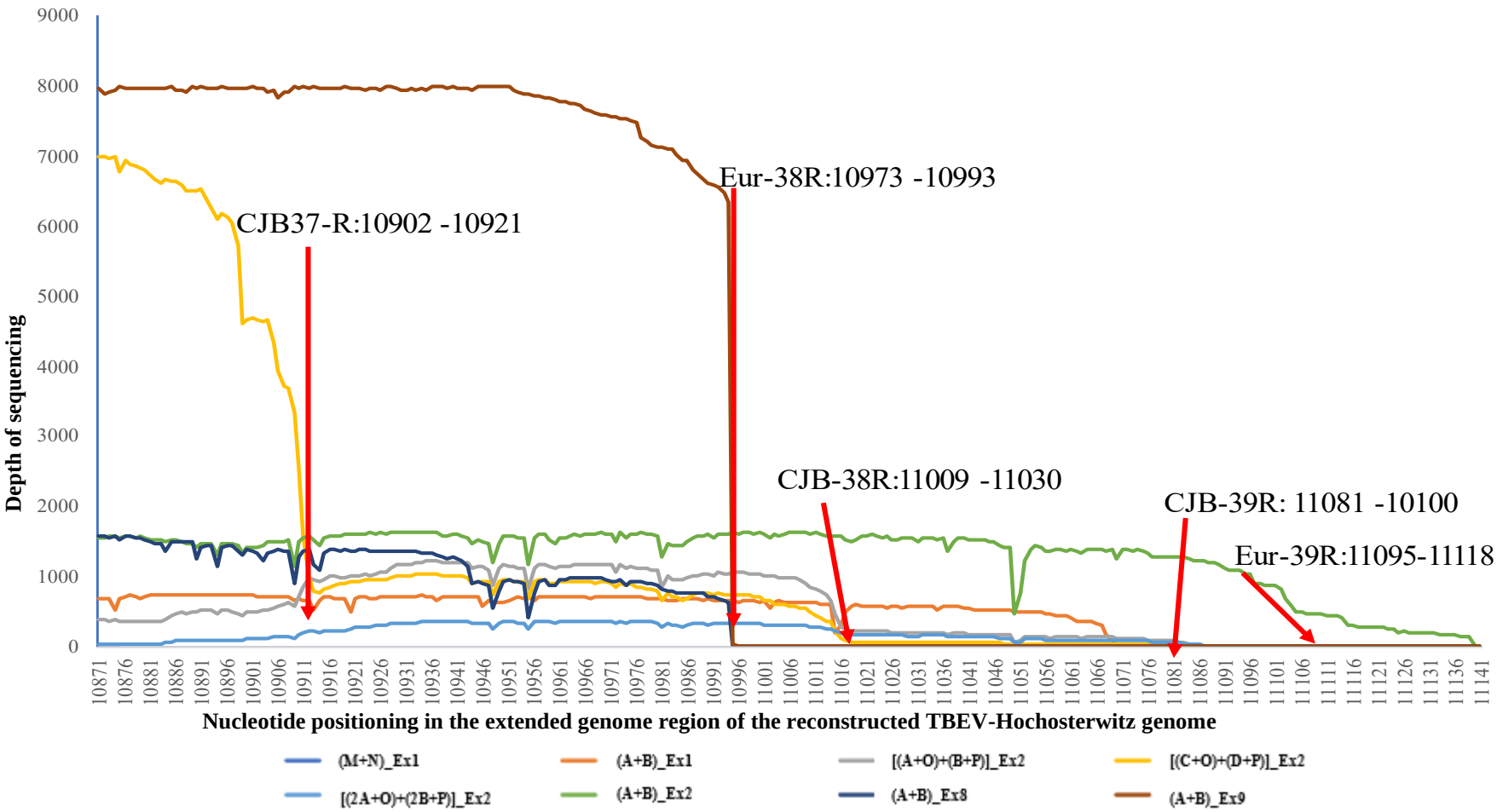
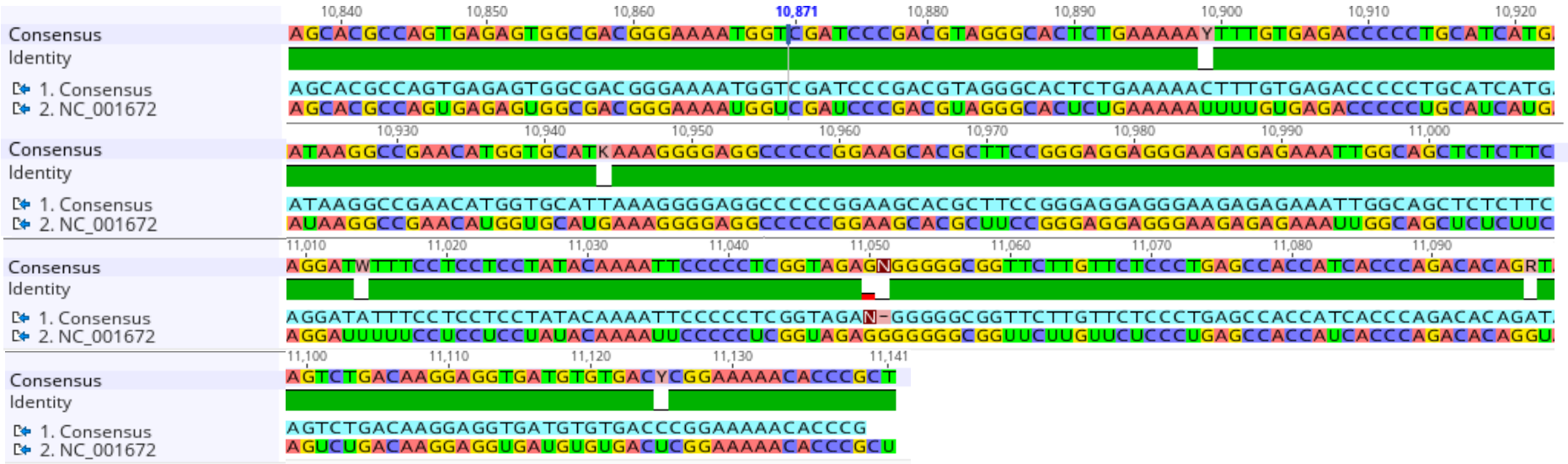


Figure A.9. Percentage of unmapped reads of reference-based mapping of Oxford Nanopore sequence data for each primer scheme against the dilution levels of TBEV-Hochosterwitz RNA used for whole genome amplification of TBEV-Hochosterwitz genome (Data source: Table A.4, Annexure I).

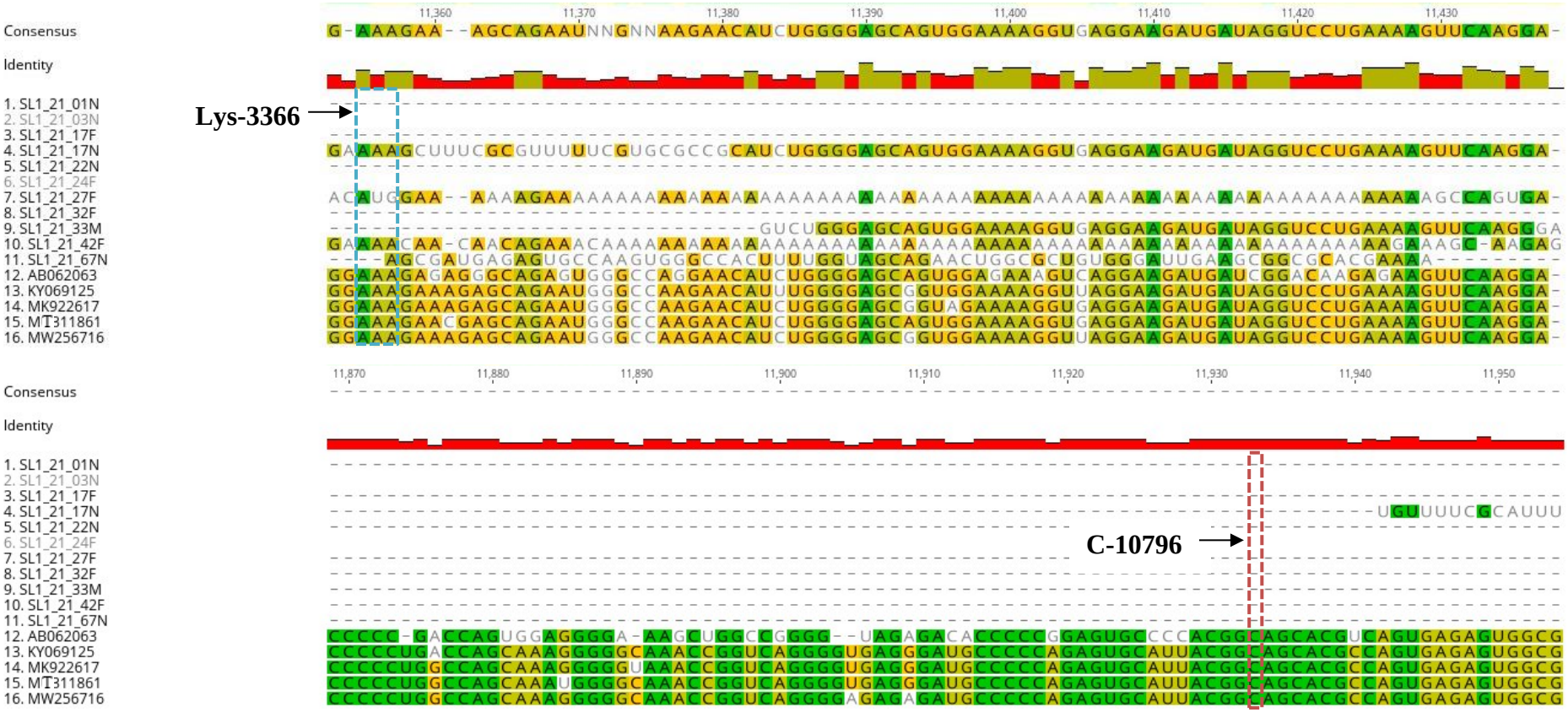


A.10. Extended genomic 3'-UTR region of the TBEV-Hochosterwitz RNA, reconstructed through the use of new (CJB) and old (Eur) primer pairs under different sequencing experiments [Ex: Sequencing experiment number (Table A.4, Annexure I)]. Arrowheads indicate the genomic sequence position where the ending of each PCR, amplified by the final pair of primers.



A.11. Extended 3'-UTR region of the TBEV-Hochosterwitz genome (1. Consensus sequence in the figure), reconstructed by the reference based (NCBI Acc. no. NC_001672) assembling of Oxford Nanopore sequence reads derived through the PCR amplification of TBEV-Hochosterwitz cDNA (tenfold dilution: log10: -4) by using the combination of primer pairs [Eur_37-F/R, CJB_37-F/R, Eur_39-F/R, CJB_37-F/R, CJB_38-F/R](Table A.4, Sequencing experiment no. 3, Annexure).

A.12. Neuroinvasive virulence factors [NS3 Ser-73 (Růžek et al., 2008), Asp-483, Lys-3366 and C-10796(NCBI Acc. no. AB062063) (Goto et al., 2003)]and their corresponding positions in the TBEV genome are highlighted in a MSA of TBEV genome sequences segments from the tick samples (Consensus sequences were obtained with the minimum reads depth = 10) with corresponding evolutionary closely related TBEV reference genome sequences (Identified through the phylogenetic tree analysis, Figure 3.13) and a neuroinvasive TBEV sequence (NCBI Acc. no. AB062063, Goto et al., 2003). Only one neuroinvasive virulence factor (Lys-3366) was identified in two separate reconstructed genome sequence segements from the tick samples (SL1_21-17N & SL1_21-42F). Complete multiple sequence alignment is attached as a separate pdf file (Figure A2.1 Annexture II).



A.12. Neuroinvasive virulence factors [NS3 Ser-73 (Růžek et al., 2008), Asp-483, Lys-3366 and C-10796(NCBI Acc. no. AB062063) (Goto et al., 2003)]and their corresponding positions in the TBEV genome are highlighted in a MSA of TBEV genome sequences segments from the tick samples (Consensus sequences were obtained with the minimum reads depth = 10) with corresponding evolutionary closely related TBEV reference genome sequences (Identified through the phylogenetic tree analysis, Figure 3.13) and a neuroinvasive TBEV sequence (NCBI Acc. no. AB062063, Goto et al., 2003). Only one neuroinvasive virulence factor (Lys-3366) was identified in two separate reconstructed genome sequence segements from the tick samples (SL1_21-17N & SL1_21-42F). Complete multiple sequence alignment is attached as a separate pdf file (Figure A2.1 Annexture II). Cont..

8. Annexure II

Content related to the Annexure II is attached as a separate below and the respective figure caption is given here.

Figure A2.1. MSA of TBEV genome sequence segments from the tick samples (Consensus sequences were obtained with the reads depth = 10) with corresponding evolutionary closely related TBEV reference genome sequences (Identified through the phylogenetic tree analysis, Figure 3.13) and a neuroinvasive TBEV sequence (NCBI Acc. no. AB062063, Goto *et al., 2003).*

9. Annexure III

Content related to the Annexure III is attached as a separate below and the respective figure caption is given here.

Figure A3.1: Multiple sequence alignment exhibiting the regions of sequence gaps in the consensus sequence derived from SAMs with minimum sequencing depth of 20 reads in the reconstructed TBEV partial genomes from two nymph tick pools, five adult female tick samples and two male tick samples against the reference TBEV variants identified in the phylogenetic tree construction (Figure 3.13).