Application of Next Generation Phage Display technology to develop virus specific peptide-based serology diagnostics for mosquito-borne and tick-borne Flaviviruses.

Anitha Ann Varghese

Thesis submitted for the degree of Doctor of Philosophy

Primary Supervisor: Professor Janet Daly

Secondary supervisors: Professor Kevin Gough, Dr Mara Rocchi

School of Veterinary Medicine and Science,

University of Nottingham, United Kingdom

15 May 2023

Declaration

I hereby declare that this thesis is all my own work, except as indicated in the text.

Signed:

Date: 24 November 2021

Resubmission: 10February 2023

Abstract

Flaviviruses have been intensively studied since the early 20th century as a direct result of the major impacts they have on human and animal health globally. Most flaviviruses are vectored by hematophagous arthropods and the disease distribution and seasonality matches that of the specific vector and its lifecycle. Vector competence in transmitting multiple arboviruses have resulted in some flaviviruses co-circulating. Flaviviruses are structurally similar, and antibodies generated during flavivirus infections to the envelope protein are highly cross-reactive especially if the viruses belong to the same sero-group. These antibodies can also lead to negative outcomes for the host when the immune system next encounters a secondary flavivirus infection for example, in the case of Dengue Haemorrhagic Fever. Cross-reactive antibodies have also been a major challenge to developing accurate serological methods of detecting specific flaviviruses.

In this thesis, phage biopanning was used to identify virus specific peptide mimotopes using flavivirus IgG positive serum samples. The specific phage library used was a phagemid library expressing a random 16-mer peptide on the major coat protein of M13 phage. Over several rounds of biopanning, a sub-library of phage that are mimics of the target serum antibody are enriched. The strategy used included subtractive biopanning where the phage library was incubated with the control sera and the bound phage was depleted from the library and positive biopanning where binders to the target were rescued and amplified as a sub-library.

Following the phage biopanning, the peptide mimotopes from the sub-libraries are identified through sequencing. The traditionally used method of clone picking, and the modern route of next generation sequencing were used to analyse the peptide mimotopes enriched towards the target.

A novel method of data analysis, originally used to analyse peptide microarray data, was applied to the Next Generation Phage Display (NGPD) datasets from flavivirus biopanning experiments to identify target specific peptide mimotopes. Machine learning concepts of data training and testing were applied to generate a list of peptides (immunosignature) enriched only in the target sub-libraries in

comparison to the control. The immunosignature was then tested against target and control samples. Area Under the Curve (AUC) values from Receiver Operating Characteristic (ROC) curve was used as the statistical method used to evaluate the sensitivity and specificity of the immunosignature as an *in-silico* assay.

The aim of the first study where this analysis was applied was to identify a peptide mimotope panel unique to Zika virus (ZIKV) IgG from human sera (also positive for DENV IgG). A panel of human sera from mosquito-borne flavivirus endemic region was used in the biopanning experiment. The control sera were seropositive for dengue (DENV) IgG only. NGPD was applied to develop a ZIKV immunosignature with an AUC value of 0.922 and the *in-silico* assay was 82.1% sensitive and 69.0% specific when tested against additional sub-libraries generated by re-screening the phage sub-library with sera not used in biopanning.

Immunosignature analysis was also applied to identify virus specific peptide mimotopes for LIV and tick-borne encephalitis viruses (TBEV) from ovine sera in a method similar to that followed in the first study. The AUC for the TBEV immunosignature was 0.930 and for 0.78 for the LIV immunosignature. Due to a smaller number of samples available, the sensitivity and specificity values against independent samples not used in the generation of the library.

The second output of this thesis was the development of a proof of principle phage peptide ELISA to detect LIV IgG from ovine sera. Currently there are no ELISA based serology assays for detecting LIV IgG. A phage sub-library enriched with binders to LIV IgG was raised by biopanning the peptide phage library with serum samples positive for LIV IgG. Following traditional phage display techniques, phage clones were picked, grown, Sanger sequenced and tested on ELISA for their ability to distinguish between LIV IgG positive and LIV IgG negative ovine sera. A phage ELISA method was developed, and the clones were screened against a panel of sera not used in the biopanning process. Statistical analysis of the results showed that one peptide had an AUC value over 0.9 and four others had AUC values >0.80. To improve the assay further, it has been proposed that the best performing clones be combined to increase sensitivity and specificity. The approaches developed in this project can be applied to other

flaviviruses and potentially lead to the development of a flavivirus pan-species multiplex assay.

Acknowledgements

Now, our God, we give you thanks, and praise your glorious name.

1 Chronicles 29:13

I would like to thank my supervisors Professor Janet Daly, Professor Kevin Gough, and Dr Mara Rocchi for their knowledge, expertise, support, and time during my PhD journey. I want to specifically thank Janet for being a brilliant primary supervisor. You have encouraged and mentored me through this process, which has been one of huge personal growth for me.

I want to thank the University of Nottingham and Moredun Research Institute for the funding for this project and our collaborators in the UK, Estonia and Norway for the resources that were provided for this project. I want to express my gratitude to the 'Return to Research' and technical team at SVMS who made it possible for me to return to labs and complete my experimental work safely during the disruption caused by the pandemic. Thank you, Jon, Maria, and Rose, for the help you provided on my phage display learning curve, Emma for your advice on maths-y stuff, the One Virology group for their encouragement.

Thank you to my friends and colleagues in the Vet school. Friends in A43: Ge, Apaa, Meshach, Tosin (for helping me with washing all the ELISA plates), Rachael, Steve, Tosh and Rob. Friends in B64: Claire, Keith, Mary, Britney, Kieran. Thank you, Juan, Kate, Sukanya, Alice, Beth, Megan, Jemma, Hayley, Caleb, Mirelle, and all the other vet school friends for the memes, flowers, bakes and chats.

Writing this acknowledgement section made me reflect on the path that has led to this point. Part of my childhood was spent in Kerala, where tropical diseases, especially mosquito-borne diseases such as dengue fever are still a scary reality. I want to thank Dr Kiki Bodman-Smith (University of Surrey), who I credit to having shaped the path by encouraging me to continue my interest in tropical diseases by studying at the London School of Hygiene and Tropical Medicine.

Thank you to my friends and colleagues at UKHSA for all the encouragement that preceded the PhD journey. Thank you, Deborah Wright, Professor Roger

Hewson, Professor Miles Carroll, Dr Ed Newman, Dr Yper Hall, Dr Richard Vipond, Dr Stuart Dowall and friends from V+P, Hannah, Adrian, Kim, Nicola, Sudhish and Sarah. The scientific and life-skills I gained have helped me navigate my PhD, especially the challenges brought on by the pandemic.

Thank you to the Carter family, Judd, Owens and Varughese families for your kindness over the years. A huge thank-you to Val and John, Vee, Dave and Di, Jerry, Ash, Kav, Jal, Mona, and Ammu. I thank my friends at St Paul's Salisbury, SB Baptist, and the wider church family for spiritual guidance and practical support especially during the isolating times during the lockdowns. I thank and give thanks for friendship circle, too many to name. So, for brevity's sake if you read this, it means you have been a source of encouragement and I am grateful for your friendship, encouragement, and support.

Now, to thank the cheerleaders who have been there right from the beginning. Thank you, mum, for all that you have done for me, for encouraging, believing and for the freedom to make decisions even when it was counter-culture. Thank you Agith, for iron sharpens iron. Thank you David uncle, my family in Kerala especially Mon acha, Reji uncle for encouraging, challenging and cheering me on.

Now to thank the brightest stars: my father K. U Varghese, my grandparents especially Itty M John and Annamma Ulahannan for the path they set me on. My grandmother was unable to pursue an education beyond primary schooling due to the cultural expectations of her gender roles. Her thirst for education has been a strong inspiration for me and this achievement reflects her vision and encouragement.

Saving the best for last, I am so thankful for you Dan. As the cliché goes, you wait for a bus for ages and two come at once. You and this PhD! You encouraged me to embark on this journey since the day I saw the advert for this PhD and you have cheered me on through all of it. You brighten my day 😊

Conference/talks associated with this thesis

AVTRW Annual Conference 2021 talk: Application of next generation phage display to develop new serology diagnostics for Flaviviruses.

AVTRW Annual Conference 2020 (Poster prize runner up): Application of next generation phage display to develop a new serology assay for detection of louping ill virus (LIV) IgG

Microbiology Society Annual Conference 2019 poster: Application of next generation phage display to identify peptide mimotopes from Zika and Dengue IgG positive serum.

The UK Hepacivirus and Flavivirus Meeting, Rydal Hall talk 2019: Applying Next Generation Phage Display technology to identify immunoreactive peptides for use in Flavivirus sero-diagnostics.

Contents

Table of Tables

Table of Figures

Abbreviation	Definition
ACDP	Advisory Committee on Dangerous Pathogen
ADE	Antibody-dependent Enhancement
BC	Barcode
CFU	Colony forming units
CL	Containment level
DENV	Dengue virus
DEPC	Deionized, diethylpyrocarbonate
DHF	Dengue haemorrhagic fever
DIVA	Differentiating Infected from Vaccinated Animals
DNA	Deoxyribonucleic acid
DSS	Dengue shock syndrome
EDTA	Ethylenediaminetetraacetic acid
ELISA	Enzyme-linked immunosorbent assay
HA	Haemagglutination assay
HAI/HI	Haemagglutination inhibition assay
HRP	Horseradish peroxidase
ICTV	The International Committee on Taxonomy of Viruses
IFA	Immunofluorescence assay
Ig	Immunoglobulin
LIV	Louping ill virus
MOI	multiplicity of infection
NGPD	Next generation phage display
NGS	Next generation sequencing
NS	Non-structural protein
OD	Optical density
PCR	Polymerase chain reaction
PEG	Polyethylene glycol
PRNT	Plaque reduction neutralization test
ROC	Receiver operating characteristic
RPL	Random peptide library
SD/STDEV	Standard deviation
SNT	Serum neutralisation test
TAE	Tris-acetate-EDTA
TBE	Tris/Borate/EDTA
TBEV	Tick-borne encephalitis virus
TMB	3,3’,5,5’-Tetramethylbenzidine
ZIKV	Zika virus

Chapter 1 **Introduction**

1.1 Background to the project

Flaviviruses are arboviruses that are maintained in a complex cycle of infection involving arthropod vectors and a range of mammalian and avian hosts. Some flaviviruses mainly cause disease in humans, such as Zika virus (ZIKV), dengue virus serotypes 1-4 (DENV) and tick-borne encephalitis virus (TBEV) and some mainly cause disease in susceptible animals and birds, such as louping ill virus (LIV). Flaviviruses have been studied for over a century because of the impact they have had on human and animal health (Holbrook, 2017). Most of the global burden of flaviviruses are centred around the tropics, due to the mosquito vector while the tick vectored flaviviruses are seen in the temperate regions.

Arthropod vectors and flaviviruses associated with them have been introduced to new territories due to a mix of natural and environmental factors such as migratory birds, increasing global temperatures and human factors such as travel and commercial shipping. Examples include the emergence and subsequent global spread of ZIKV (Musso, Nilles and Cao-Lormeau, 2014), discovery of TBEV in the UK (Holding *et al.*, 2020), and West Nile virus (WNV) in Austria (Wodak *et al.*, 2011) to name a few.

1.2 Introduction to Flaviviruses

Flaviviridae are a family of 89 single stranded, positive sense, enveloped RNA viruses that are sub-divided into four genera: Hepacivirus, Pegivirus, Pestivirus and the Flavivirus genus. *Flaviviridae* are one of the many virus families whose members can infect arthropods which are also known as arboviruses (arthropod-borne viruses).

1.2.1 Virus morphology

Flavivirus virions are spherical in structure and measure 50 nm in diameter. The surface glycoprotein exhibits icosahedral symmetry, and the mature virion has a smooth surface (Kaufmann and Rossmann, 2011). Two forms of the virion have been distinguished. The immature form is non-infectious and generally not seen outside the host cell whereas the mature form is present in the extracellular matrix. The mature virion is infectious (Figure 1-1) and enters the host cell through receptor mediated endocytosis and the low pH in the endosome triggers the viral and host cell membrane to fuse, which releases the viral genome (Bressanelli *et al.*, 2004).

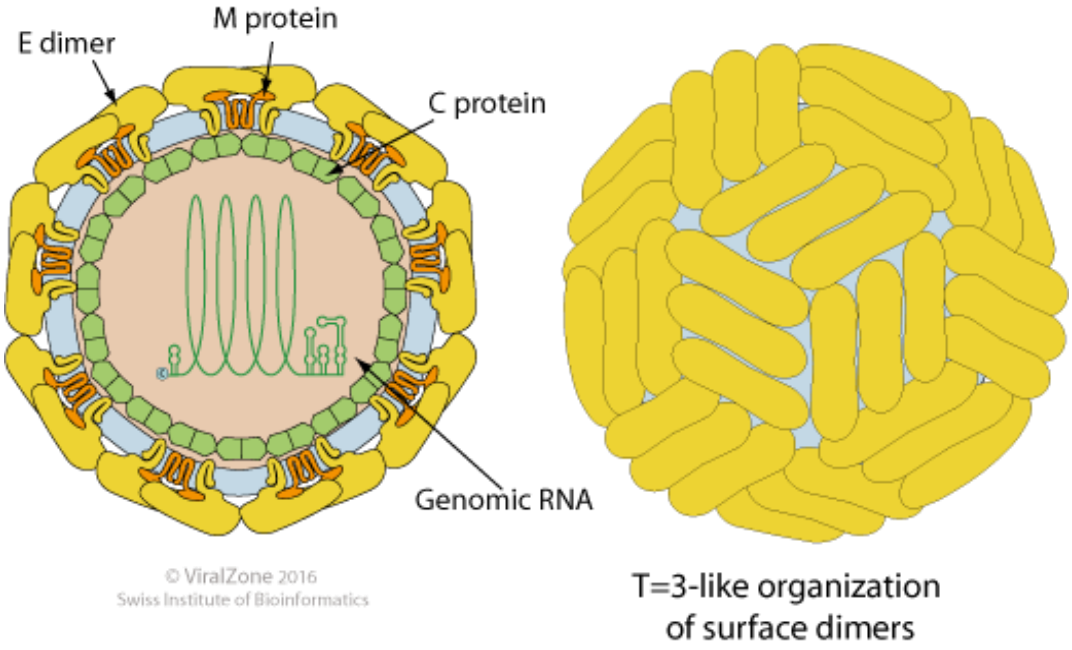


Figure 1-1: Flavivirus structure.
Source: Viral Zone: SIB Swiss Institute of Bioinformatics.

The viral genome is 9.2–11 kb in size and is a non-segmented, linear, single stranded positive sense RNA genome. The viral genome is translated into a single viral polyprotein which is glycosylated and cleaved into 3 structural and 7 non-structural viral proteins through the action of

viral and host proteases (Lindenbach, Thiel and Rice, 2011). The three structural proteins are namely envelope protein (E), membrane protein (M) and capsid protein (C) (Figure 1-1). The seven non-structural proteins are NS1, NS2a, NS2b, NS3, NS4a, NS4b, and NS5 (Kaufmann and Rossmann, 2011) and their functions are shown in Table 1-1.

Table 1-1: Functions of Flavivirus non-structural proteins

Non-structural proteins	Function
NS1	Plays an early role in replication
NS2a	Undefined but essential role in RNA accumulation
NS2b	Co-factor to NS3
NS3	Along with its co-factor NS2b, forms the main viral protease. NS3 also has helicase and nucleoside triphosphatase properties which are essential for viral replication
NS4a	Undefined but essential role in RNA accumulation
NS4b	Undefined but essential role in RNA accumulation
NS5	RNA dependent RNA polymerase

The immature and non-infectious virion assembly occurs in the host cell endoplasmic reticulum. The viral C protein forms a nucleocapsid core with the viral RNA in a host-derived lipid bilayer surrounded by dimers of E and the precursor form of M (PrM) proteins (Fernandez-Garcia *et al.*, 2009). Following viral assembly, the immature virion undergoes a maturation step where E protein is structurally re-organised and prM is cleaved by a protease, which allows the infectious virion to exit the cell through exocytosis (Stadler *et al.*, 1997; Elshuber *et al.*, 2003). Cryo-electron micrographs of DENV virions showed a structural difference in how the envelope proteins are arranged on the surface of the immature and mature forms of the virion. The E dimers are arranged in a herringbone fashion on the mature virion surface and there is icosahedral symmetry, as shown in the image on the right in Figure 1-1.

In the immature form, the conserved regions on the surface structure are exposed due to the configuration of E and prM proteins (Rey *et al.*, 2018). The antigenic regions of flaviviruses include the non-structural protein NS1, and from the structural proteins, E-domain dimers (EDE) of the

surface envelope glycoprotein, fusion loop epitope (FLE), part of the conserved region of the E protein are immunogenic.

1.3 Taxonomy of flaviviruses

Flaviviruses can be taxonomically classified on a range of criteria such as nucleotide and deduced amino acid sequence data, antigenic characteristics, geographic association, vector association, host association, disease association and ecological characteristics (Simmonds *et al.*, 2017). At present, a combination of these attributes is used to classify flaviviruses. Figure 1-2 is a radial format phylogenetic tree generated using reference polyprotein regions and it shows two distinct clades of tick-borne flaviviruses and mosquito-borne flaviviruses.

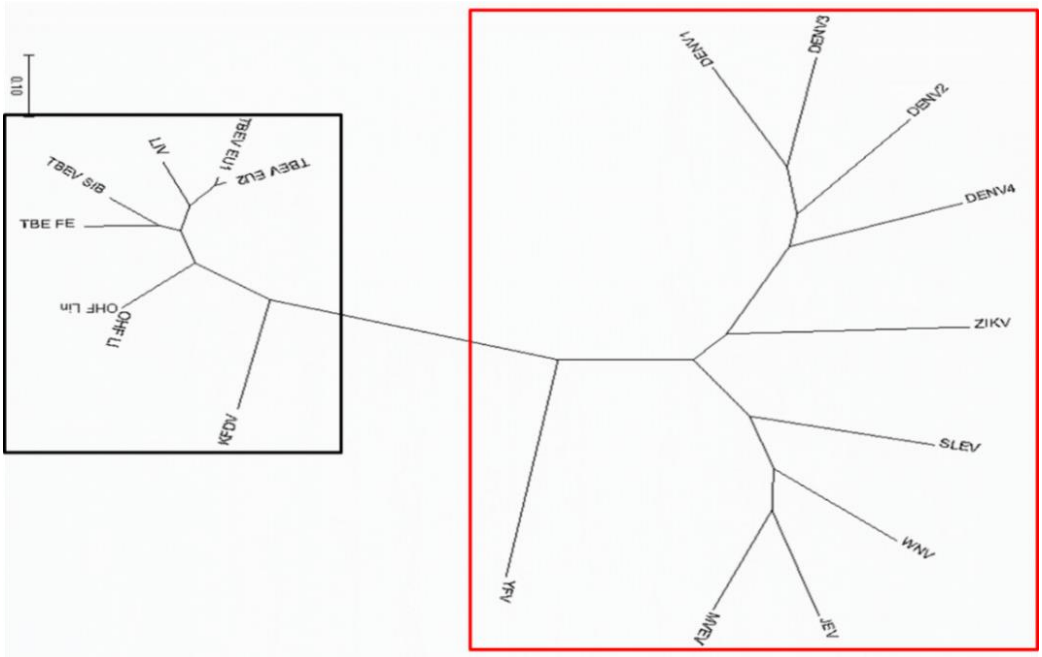


Figure 1-2: Arboviral clades of Flavivirus.
A radial format phylogenetic tree using reference genomes of key mosquito and tick-borne flaviviruses generated in MEGA 11. The complete genome nucleotide sequences of the NCBI reference genomes for each species were aligned in MEGA using Clustal W and the phylogeny was inferred using the default neighbour joining method. Grouped in the red box are the mosquito-borne flaviviruses and in the black box are the tick-borne flaviviruses.

Members of the mosquito-borne flavivirus group include four different DENV serotypes, ZIKV, Japanese encephalitis virus (JEV), WNV and Yellow fever virus (YFV). Members of the tick-borne flavivirus group include TBEV, LIV, Spanish Goat Encephalitis Virus (SGEV) and, Kyasanur Forest disease virus (KFDV).

The complexity of flavivirus classification can be demonstrated using the tick-borne flaviviruses LIV and TBEV. LIV and TBEV are serologically cross-reactive, and LIV is antigenically closer to the European serotype of TBEV (TBEV-Eur) than TBEV-EU is to the Far-Eastern (FE) and Siberian (SIB) TBEV serotypes. LIV and TBEV-Eur also share a common arthropod vector. The differentiating characteristics between LIV and TBEV is the host susceptibility where LIV is mainly a veterinary disease while TBEV is a human disease. Geographical differences are also important in taxonomic classification and LIV is mainly found in the British Isles while TBEV has only recently been discovered in the UK, but it is endemic in many European countries (Jeffries *et al.*, 2014; Simmonds *et al.*, 2017; Holding *et al.*, 2019).

1.4 Flavivirus vectors and host species

As mentioned earlier, medically important arboviral flaviviruses are classified into two groups based on the hematophagous vector. Figure 1-3 illustrates the complex vector—host relationship that is involved in the transmission of mosquito-borne and tick-borne flaviviruses.

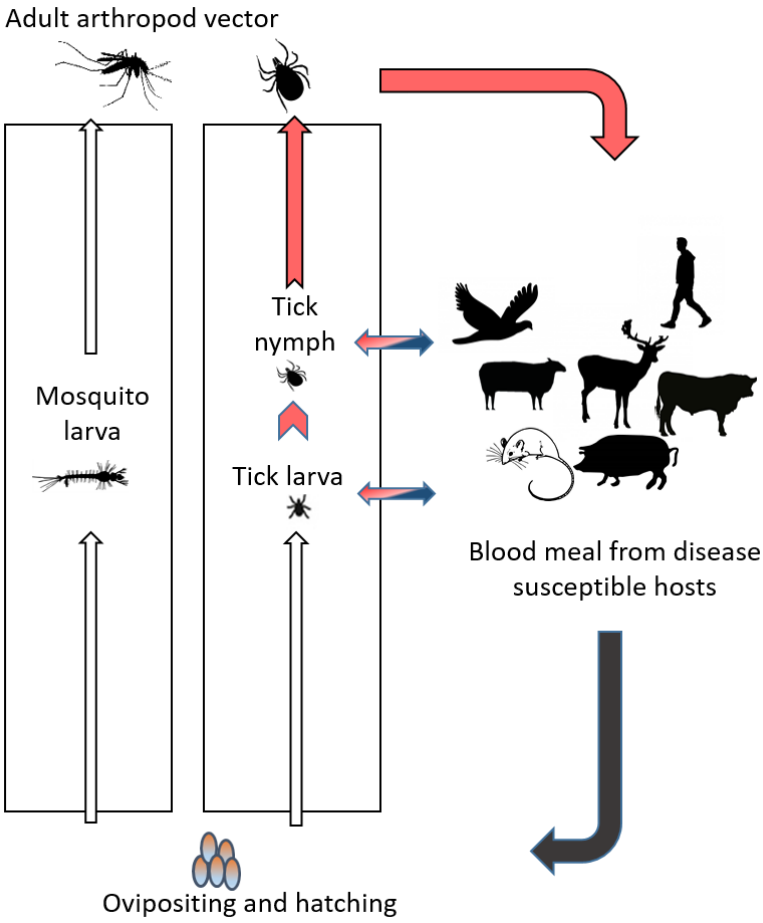


Figure 1-3: Illustration to show the complex inter-relationship between flaviviruses, their insect vectors and the host species.
Own image. The female mosquito vector requires a blood meal before ovipositing. In contrast, the male and female tick vectors require a blood meal prior to moulting and before the female lays her eggs (Földvári, 2016; Tahir *et al.*, 2020). The nymphs and larval ticks can acquire infections through blood feeding which are then transmitted to susceptible animal hosts during its next blood meal (blue/red arrow). The red arrows indicate adult vectors becoming infected following a blood meal and vertical transmission from female to her eggs (Lai *et al.*, 2020).

1.4.2 Mosquitoes as flavivirus vector

The earliest flavivirus transmission studies which hypothesised the link between mosquito and virus transmission were by J.C. Nott and Carlos Finlay during the YFV outbreak in Cuba in 1887 and this link was confirmed by Walter Reed (Lindenbach, Thiel and Rice, 2006; Holbrook, 2017).A majority of mosquito-borne flaviviruses are transmitted by members of the *Culicinae* subfamily, and to a smaller degree by members of the *Anophilinae* subfamily. The mosquitoes blood feed from a vertebrate host because a blood meal is required for them to lay their eggs. The two main mosquito vectors associated with flaviviruses are *Aedes albopictus* (*Ae. albopictus*) and *Aedes aegypti* (*Ae. aegypti*).

Evolutionary and modelling studies along with historical accounts shows that the ancestor species of *Ae. aegypti* was a tree dwelling species from northern Africa (*Ae. aegypti formosus*) and as the species moved closer to human settlements following changes in climate diverged into the anthrophilic subspecies *Ae. aegypti aegypti* (Aaa). The female Aaa mosquitoes preferred to take blood meals from humans rather than birds or other animals and adapted themselves to breeding in small water containers near human settlements. The species was subsequently transported out of Africa during the slave trade and established itself in ecologically suitable regions. The species is currently present in tropical, sub-tropical and temperate regions across the world (Mattingly, 1957; Tabachnick, 1991; Brown *et al.*, 2014; Kraemer *et al.*, 2015; Lwande *et al.*, 2020).

Ae. albopictus commonly known as the Asian tiger mosquito is native to the forests of Southeast Asia and through the international trade of used tyres and exotic house plants such as lucky bamboo in the 1970s, this species has spread globally. It is listed as one of the top 100 invasive species on the Global Invasive Species Database. This species is known for its biological and ecological plasticity, as it is able to feed on multiple

animal and bird species and breed in densely vegetated regions as well as water receptacles near human settlements ranging from bird baths, flowerpots to tyres. It has been able to replace *Ae. aegypti* in some regions. The eggs of this species are laid just above the water line and are desiccation proof. In Europe where the winters are colder, the eggs hibernate (diapause) over winter. The link between this species and DENV was made during an outbreak of DENV in Japan in 1942 as this occurred in the absence of *Ae. aegypti.* It is vector competent for up to 22 arboviruses including other flaviviruses namely YFV and ZIKV (Gratz, 2004; Global Invasive Species Database, 2009; Kraemer *et al.*, 2015; European Centre for Disease Prevention and Control (ECDC), 2016; Leta *et al.*, 2018). The spread of vectors globally has also resulted in the expansion of the associated flaviviruses. Both mosquito species are vector competent for transmitting ZIKV and DENV, the two mosquito-borne flaviviruses of focus in this study.

1.4.3 Ticks as flavivirus vector

The second group of arboviral flaviviruses are those transmitted by ticks. Members of the *Ixodes* genus and to smaller extent *Dermacentor* and *Haemophysalis* genus of the subfamily *Ixodidae* are flavivirus vectors. Members of the Ixodes genus *Ixodes ricinus* (*I. ricinus*), and *Ixodes perseculatus* (*I. perseculatus*) are the main vectors of the tick-borne viruses TBEV and LIV (Michelitsch *et al.*, 2019), which are the focus of this work. Other members known to transmit flaviviruses include *Ixodes scapularis* (*I. scapularis*) and *Ixodes cookei* (*I.cookei*) and these ticks transmit Powassan virus (POWV) (Corrin *et al.*, 2018; Cumbie, Whitlow and Eastwood, 2022), a tick-borne flavivirus which is found in North America and the Far Eastern region of Russia.

I. ricinus (castor bean tick/sheep tick) is vector competent for both TBEV and LIV in Europe and is known as a three host tick. *Ixodes* ticks can live for many years with an average life span between 2 to 3 years. Its life

cycle begins with the tick larva, which requires a blood meal before moulting to form the nymph and a blood meal is required before moulting to the adult. Female and male ticks blood feed before mating. As ticks are susceptible to desiccation in extreme temperature conditions, they are most active during spring and autumn and the rest of the year they enter a stage of diapause (Földvári, 2016).

In all three stages, the tick climbs up vegetation and waits for a vertebrate host to brush past and attaches to take a blood meal. This behaviour is known as questing. Virus transmission can occur from vector to host within hours of attachment and the duration of the blood meal can vary from 2 days to up to 12 days. Virus transmission studies have shown that an infected tick can transmit TBEV to an uninfected nymph while co-feeding on an uninfected host animal (Labuda *et al.*, 1993). In this study the host animal remained non-viraemic while 55% of the nymphs acquired the virus. The virus infects the tick in the midgut and passes through the hemocoel into the salivary glands. The tick then transmits the virus during its next blood meal. (Nuttall and Labuda, 2003). Flaviviruses can also be transmitted vertically from the female tick to its eggs (Nuttall *et al.*, 1994).

1.4.4 Non-vectored transmission

Non-vector routes of disease transmission do exist for some flaviviruses (Chen and Wilson, 2016), such as sexual transmission of ZIKV (Atkinson *et al.*, 2016; D'Ortenzio *et al.*, 2016), and DENV transmission through infected blood or organs (Levi *et al.*, 2015). LIV infections have been reported in people working closely with livestock such as shepherds and veterinarians. Laboratory acquired cases of LIV have also been reported where lab personnel became infected through exposure to the virus while slaughtering experimentally infected animals (Davidson, Williams and Macleod, 1991).

1.4.5 Host animals

Host population dynamics for flavivirus infections are influenced by a range of conditions such as geographical distribution of the vector, climatic conditions, the preferred host for the arthropod for its blood meal. Table 1-2 provides a list of flavivirus species, arthropod vector, their preferred animal/bird host, and geographic distribution.

Table 1-2: Mosquito-borne and Tickborne flaviviruses of importance to human and veterinary disease along with the virus' preferred host, mode of transmission, outcome of severe disease and general geographic distribution. Table from Veterian Key, 2016.

Virus	Host of Concern (Reservoir Host)	Arthropod Host (Mode of Transmission)	Disease in Domestic Animals (or Humans)	Geographic Distribution
(i) Mosquito-borne flaviviruses				
Dengue viruses 1, 2, 3, and 4	Humans (humans and monkeys)	Mosquitoes: Aedes aegypti, other Aedes spp.	(Fever and rash, arthralgia, myalgia, hemorrhagic fever)	Africa, tropical areas of Asia, Oceania, Australia, and the Americas (Central and South America)
Japanese encephalitis virus	Swine, humans, horses (birds)	Mosquitoes: Culex tritaeniorhynchus, other Culex spp.	Abortion, neonatal disease (encephalitis)	Asia
Murray Valley encephalitis virus	Humans (birds)	Mosquitoes: Culex annulirostris	(Encephalitis)	Australia, New Guinea
St. Louis encephalitis virus	Humans (birds)	Mosquitoes: Culex tarsalis, Cx. pipiens	(Encephalitis)	United States, Canada, Central and South America
Usutu virus	Birds and humans (birds)	Mosquitoes: Culex spp. and Aedes spp.	Encephalitis, myocardial necrosis, hepato- and splenomegaly	Africa and Europe
Wesselsbron virus	Sheep	Mosquitoes: Aedes spp.	Generalized infection, abortion	Africa
West Nile virus	Humans, horses, birds (birds)	Mosquitoes: Culex spp. (rarely ticks)	Fever, generalized disease, and encephalomyelitis	Africa, Middle East, North America, Central America, South America, India
Yellow fever virus	Humans (humans and monkeys)	Mosquitoes: Aedes aegypti, Aedes albopictus and Haemagogus spp.	(Yellow fever)	Africa, Central America and South America
(ii) Tick-borne encephalitis viruses				
Tick-borne encephalitis virus	Humans (rodents, birds, ruminants)	Ticks: Ixodes spp. and via ingestion of raw milk	Encephalitis	Europe, Russia, Asia
Kyasanur Forest disease virus	Humans (monkeys, rodents)	Ticks: Haemaphysalis spp.	(Haemorrhagic fever, encephalitis)	India (Mysore)
Louping ill virus	Sheep, horses, humans	Ticks: Ixodes ricinus	Encephalitis	Europe
Omsk haemorrhagic fever virus	Humans (muskrats)	Ticks: Dermacentor spp.	(Haemorrhagic fever, gastrointestinal disease)	Central Siberia, Confederation of Independent States
Powassan virus	Small mammals, humans, possibly horses	Ticks: Ixodes spp.	Encephalitis	Canada, United States, Russia

1.5 Disease pathogenesis and host immune response

Disease pathogenesis caused by flaviviruses is influenced by the susceptibility of the animal host. Some flaviviruses are maintained in the one animal host (reservoir) without the virus causing symptomatic disease and yet in other hosts (susceptible) the virus can be of high consequence. Some flaviviruses such as DENV, TBEV, YFV are human pathogens and others such as LIV mainly affect animals and are of economic importance. Acute flavivirus disease presentation can range from asymptomatic or mild self-limiting febrile illness to severe and life-threatening diseases shown in Table 1-2 in both humans and animals.

Flaviviruses exhibit tissue trophism and this can be seen in disease pathology that accompanies severe disease (Table 1-2). Viscerotropic flaviviruses such as DENV can cause shock syndrome, hypotension, vascular leakage, and haemorrhage. Infection with neurotropic flaviviruses such as TBEV or LIV can cause encephalitis, seizures, cognitive impairment, paralysis, and long term neurological damage.

Flavivirus infection during pregnancy can lead to loss of pregnancy and birth defects in the foetus. In most human cases of ZIKV infection, the virus is thought to cause asymptomatic or mild disease but in pregnant women, the virus can cross the placental barrier and infect the developing nervous system of the fetus. This can lead to congenital birth defects such as microcephaly, other neurological abnormalities and can also lead to fetal death (Antoniou *et al.*, 2020).

1.5.1 Immune response to flaviviruses

In the animal host, the flavivirus infection is known to be modulated by the proteins present in the arthropod's saliva, thus allowing the virus to evade the antigen presenting cells (APC) such as macrophages, neutrophils and dendritic cells (DC). The immune system is activated when the virus is recognised by pathogen recognition receptors (PRR)

including toll like receptors (TLR). The cytokine pathway is activated and proinflammatory cytokines including type 1 interferon, (IFN), and interleukins (IL-6, IL-12) are upregulated which activates the T-cell response. Cytokines such as IL-1 and tumour necrosis factor-alpha (TNF-a) induce DC migration to the lymph node for antigen presenting and induces the Th-1, Th-2 and cytotoxic killer cell response. The upregulation of major histocompatibility complex (MHC) class 1 molecules activates the CD8 T cell response and this in turn activates the B-cell response and production of antibodies (Müllbacher and Lobigs, 1995; Nuttall and Labuda, 2003; Devadiga *et al.*, 2020). IgM is produced if the host has not encountered the virus previously and is short lived whereas the production of IgG provides long-term protection from a secondary infection.

1.6 Diagnosing flavivirus infections

Identifying the specific pathogen is important for reasons such as ensuring the right treatment can be administered in a timely manner as the outcome of infection can be life altering (eg ZIKV infection during early pregnancy), or there may be several pathogens which result in a similar disease presentation in the early days of infection (eg CHIKV, ZIKV, DENV), and in veterinary disease accurate diagnostics could help the farmer make decisions which have an economic impact on their flock management strategy (eg LIV). There are two main approaches of clinically diagnosing an infection. One is direct detection to the pathogen through molecular or cellular assays and the other, indirect detection of the antibody response caused by the pathogen (Figure 1-4).

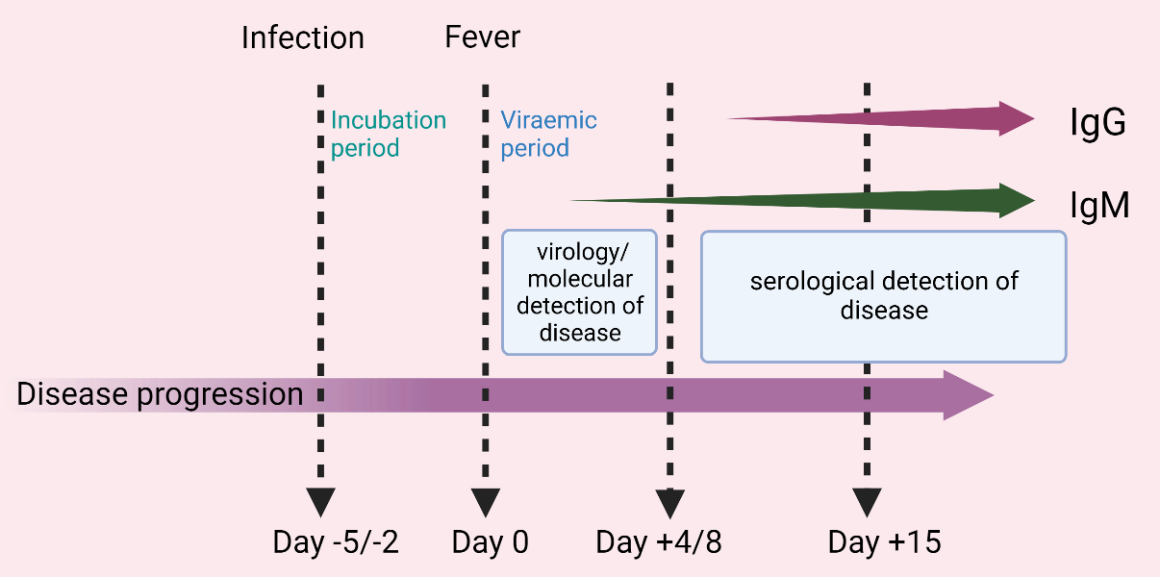


Figure 1-4: The infection stages and the best method of detection for identifying active infection and/or previous exposure.

The viraemic period of disease is where the virus is present in body fluids. In flavivirus infections, this stage lasts for 4-8 days and during this period, disease can be diagnosed using molecular methods such as PCR and direct virus detection methods. The antibodies produced in response to the infection can be detected during and post viraemia. The early immune response to infection is detected through the IgM antibody levels and the later immune response detects IgG antibody levels. For serosurveillance studies, IgG response would be the target as these antibodies are longer lasting than IgM. Figure adapted from (Musso and Desprès, 2020).

1.6.1 Virological methods

Yellow fever was the disease that sparked flavivirus research and the disease burden caused by YFV and other flaviviruses led to major discoveries that have established virology research. Virology was an unknown concept in the 1880's to the extent that in 1896, Giuseppe Sanarelli postulated that yellow fever was caused by a bacterium. The application of Koch's postulates was the gold standard for demonstrating the relationship between a pathogen and the disease it caused. The discovery of viruses as an entity, arthropod transmission of viruses, discovery of disease transmission link between the arthropod vector,

reservoir host and susceptible host were all developed during studies into yellow fever, and the Yellow Fever Commission and the West Africa Yellow Fever Commission played a critical role in these discoveries (Walter Reed *et al.*, 1900; Stokes, Bauer and Hudson, 1928; Sawyer and Frobisher, 1929; Bryan *et al.*, 1997; Reed, Carroll and Agramonte, 2001; Clements and Harbach, 2017).

During the very early days of arbovirus research, observational studies were carried out on patients and transmission studies were conducted by infecting arthropod vectors with biological material from patients. James Carroll of the Yellow Fever Commission disproved the theory that yellow fever disease was a bacterial infection by isolating the virus by filtering sera from YFV infected patients through a Berkefeld N filter and subsequently inoculating it into an immune naïve individual who went on to develop yellow fever disease (Bryan *et al.*, 1997; Clements and Harbach, 2017). In the following years, Berkefeld filters were used in other studies to isolate YFV from arthropod vectors, patients and animal models and subsequently used in transmission experiments to demonstrate the casual relationship between the virus and the disease it caused (Stokes, et al., 1928; Sawyer and Frobisher, 1929). Human volunteer infection studies were also carried out for DENV pathogenesis studies. A range of animals including monkeys, chimpanzees, and guinea pigs were used for experimental transmission studies (Stokes, et al., 1928).

The advent of in-vitro cell culture in the 1920s made it easier to study flaviviruses using continuous cell lines of mammalian and insect cell origin. Cell culture is still used routinely for the propagation of viruses; to study pathogenesis mechanisms such as persistence of viral infection in the arthropod vector and to test compounds for antiviral activity (Casals and Brown, 1953; Ludwig, 1993).

A classic virology technique which uses cell culture is the plaque assay which was originally developed for quantifying bacteriophage and subsequently modified use in quantifying pathogenic viruses by Dulbecco in 1952 (Renato Dulbecco, 1952; Cooper, 1962). Flaviviruses can form plaques and the cytopathic effect (CPE) of the virus of interest is quantified using a permissive cell monolayer and overlay method. In brief, a serial dilution of the virus is made, and a set volume of the dilution is added to each well containing a monolayer of a mammalian cell line (eg Vero E6). The plate is incubated at 37°C for an hour for the virus to infect the cells and after the incubation, the excess virus is removed. A suitable overlay is added to all the wells and the plate is incubated for as many days as it takes for the virus to exhibit CPE in the form of plaques. Following the incubation, the cell monolayer is fixed, and the plaques are counted. This provides a quantification for the number of infective virions. One of the drawbacks of this technique is that it is a slow assay. A modern variant of the plaque assay known as fluorescent focus assay (FFA) has been developed, one of the benefits of this assay being a shorter incubation period of 2 days required before the plate can be read. Traditional flavivirus plaque assays have an incubation period of 2-11 days (Payne *et al.*, 2006).

A similar assay to the plaque assay is the median tissue culture infectious dose assay ($TCID_{50}$) is an end point dilution assay which relies on quantifying CPE in the cell monolayer (Payne *et al.*, 2006). In the $TCID_{50}$ assay, the Reed-Muench method is used to calculate the percentage of infected cells at a given dilution and the virus dilution at which 50% of the cell layer is infected (Smither *et al.*, 2013). Microscopy techniques including fluorescence based super resolution microscopy and electron microscopy (scanning electron microscopy, transmission electron microscopy and cryo-electron microscopy), have been used to study various aspects of flavivirus virology including understanding the viral particle structures, pathogen entry mechanisms and protein interactions

(Chong *et al.*, 2014). The drawback of microscopy is that it can be a very expensive and niche tool which is not suited for routine diagnostic use (Chong *et al.*, 2014).

1.6.2 Molecular methods

In clinical human cases, molecular testing can be used to accurately diagnose infection if the sample was drawn during the viraemic phase. The advent of polymerase chain reaction (PCR) assays and next generation sequencing has revolutionised flavivirus diagnostics. Molecular detection methods are used during the short viraemic phase of infection. Molecular detection methods are also used to detect the presence of the virus in the arthropod vector. The commonly used method of molecular detection is the polymerase chain reaction (PCR) assay, which relies on the amplification of a specific region of the viral nucleic matter. A range of PCR assays are currently used, these include reverse transcription PCR (RT-PCR) and quantitative real time RT-PCR. Other methods used for molecular detection of flaviviruses are loop mediated isothermal amplification assays (RT-LAMP) (Lopez-Jimena *et al.*, 2018) and sequence independent single primer amplification (SISPA) which is a meta-transcriptomic method (Xiao *et al.*, 2018).

PCR can be optimised to detect viral nucleic matter from a range of biological matrices such as blood, cerebrospinal fluid, urine, and semen (Atkinson *et al.*, 2016). The pros of using PCR detection methods include excellent sensitivity, specificity, and the speed at which results are made available especially in a diagnostic setting. The ability to detect and accurately distinguish different flaviviruses through pan-flavi PCR assays gives this technique a significant advantage as it bypasses antibody-cross reactivity, a significant hurdle in flavivirus serology especially in regions where there is co-circulation of flaviviruses. The main drawback of molecular methods is that it is only successful if viraemia is present, as seen in

Figure 1-4. A positive result on PCR does not always indicate active infection as non-infectious virus particles can circulate in biological fluids just after viraemia has ended.

Molecular methods have also been used to classify flaviviruses. The first flavivirus genome to be sequenced was that of YFV in 1985 (Rice *et al.*, 1985). Flaviviruses were grouped in the *Togaviridae* family till 1984, when 65 viruses were separated into the *Flaviviridae* family based on molecular as well as structural and morphological differences (Westaway *et al.*, 1985). The advent of next generation sequencing technologies and bioinformatics have made it easier to group viruses based on phylogenetics. Sequencing data can be shared with other researchers through public repositories such as the NCBI database which are useful in cataloguing emerging and re-emerging viruses.

Sequencing data from full length viral genomes or proteins etc can be analysed using software such as Clustal W and Molecular Evolutionary Genetics Analysis (MEGA) to make phylogenetic trees. Flavivirus phylogenetic trees (Figure 1-2) constructed using the full length flavivirus polypeptide has shown that flavivirus clusters are organised by their antigenic relationship with each other as well as the common vector. The classification reported by (Rathore and st. John, 2020) gave 12 distinct flavivirus complexes and these are listed in Table 1-3.

Table 1-3: Phlyogenetic classification of flaviviruses

Phylogenetic classification
Mammalian tick-borne virus complex
Dengue virus complex
Spondweni virus complex
Japanese Encephalitis virus complex
Yellow fever virus complex
Ntaya virus complex
Seabird tick-borne virus complex
Modoc virus complex
Rio Bravo virus complex
Kokobera virus complex

Aroa virus complex
Entebbe bat virus complex

One of the main drawbacks of molecular detection methods is that the assays do not work in the absence of viraemia as shown in Figure 1-4. Post viraemia, serological assays provide confirmation of current infection or previous infection.

1.6.3 Serological methods

Antibodies are raised to the antigenic regions on the virus, and although the main antigenic regions are the envelope proteins, some antibodies are also raised to the non-structural proteins such as NS1. Serology assays detect the presence of antigen, or the antigen-antibody complex formed when cross-reactive or neutralising antibodies are present in patient serum Figure 1-4. Primary infection is detected by the presence of IgM in a samples and IgG antibody presence shows a convalescing response or previous exposure.

Flaviviruses can cause haemagglutination, and this characteristic was utilised to develop the early serological assays (Casals and Brown, 1953). Complement fixation (CF) assay was the earliest serology assay used to demonstrate antigenic cross reactivity in flavivirus species. The principle behind the assay is to detect neutralising antibody to a pathogen using complement proteins which detect antibody-antigen complexes. Serum is heat treated to remove existing complement proteins and incubated with antigen. After the incubation, complement is added to the sample and incubated once again. Finally, RBC is added to the sample and a final incubation is carried out after which the test is read. In the presence of neutralising antibody to the antigen, the complement does not lyse the RBC as it binds to the antibody-antigen complex. The RBC settles at the base of the well and forms a red dot. CFA was used by Jordi Casals and his team in 1943 to demonstrate the pattern of antigenic cross-reactivity

with a panel of neurotropic flaviviruses that included LIV, Russian Spring Summer Encephalitis virus (RSSV), WNV, SLE (Casals, 1944).

Hallauer used the Hirst assay, now known as hemagglutination assay (HA) and reported on the specific haemagglutination seen in viscerotropic but not neurotropic arboviruses (Hallauer, 1946). Sabin and colleagues used these neutralisation assays to identify different types and strains of flaviviruses. The cross-reactivity among JEV, YFV, DENV1 and 2, and WNV was reported by Sabin in 1950 (Sabin, 1950). Most of Sabin's work was carried out on YFV and DENV (Sabin, 1952). In the HA assay, the virus is serially diluted and added to 'U' or 'V' wells of a microtitre assay plate. A standard volume of red blood cells (RBC) is added to the viral dilutions and the plate is incubated for 30 minutes. If a diffuse shield is formed in the well as a result of a lattice structure forming as a result of crosslinking of the virus with the RBCs, it is indicative of a positive reaction. If there is no crosslinking, the RBC settles at the base of the well and forms a red dot and this is reported as a negative result. The assay can be used to test live and inactivated virus.

A modification to the HA led to the development of the haemagglutination inhibition assay (HAI), an assay used to quantify neutralising antibody to the virus. In HAI, the test serum sample is serially diluted and added to microtitre wells along with a standard concentration of the test virus. An incubation period allows the serum antibodies to interact with the virus particles. Following the incubation, a standard volume of prepared RBCs is added to the mix and incubated. Like in the CF assay, neutralising, or cross-reactive antibody to the virus results in the formation of antigen-antibody complex and this allows the RBC to settle to the base of the well and present as a red dot. A negative result would present as a diffuse shield due to the haemagglutination of the virus and the RBC.

Casals and Brown established the haemagglutination inhibition (HI) assays to study cross-reactivity of flaviviruses. They used CF, HA and HAI

to classify flaviviruses into group B arboviruses. They conducted HA and HAI studies using a panel of 15 different arboviruses and tested each virus against a panel of 22 sera. Casals also defined the term 'antigenic group' to describe groups of viruses which cross-react in serological assays because they share common antigenic regions (Casals and Brown, 1954; Porterfield, 1986).

Cross-neutralisation tests with 66 arboviruses against polyclonal sera demonstrated higher cross-reactivity among some viruses and a lesser level of cross-reactivity between different groups. Flaviviruses were assigned to broadly 8 antigenic complexes (Table 1-4). This study did not group ZIKV and 16 other flaviviruses as cross-reactivity was not observed (Theiler and Downs, 1973; Calisher *et al.*, 1989).

Table 1-4: Flavivirus antigenic complexes and the viruses that belong to each complex

Antigenic complex	**Viruses**
Tick-borne encephalitis	(Russian spring–summer encephalitis, Central European encephalitis), Omsk haemorrhagic fever, louping ill, Kyasanur Forest disease, (Langat, Phnom Penh bat, Carey Island), Negishi, Powassan, Karshi, Royal Farm
Rio Bravo	Rio Bravo, Entebbe bat, Dakar bat, Bukalasa bat, Saboya, Apoi
Japanese encephalitis	Japanese encephalitis, Murray Valley encephalitis, Kokobera, Alfuy, Stratford, St Louis encephalitis, Usutu, West Nile, Kunjin, Koutango
Tyuleniy	Tyuleniy, Saumarez Reef, Meaban
Ntaya	Ntaya, (Tembusu, Yokose), (Israel turkey meningoencephalitis, Bagaza)
Uganda S	Uganda S, Banzi, Bouboui, Edge Hill
Dengue	Dengue 1, dengue 2, dengue 3, dengue 4
Modoc	Modoc, Cowbone Ridge, Jutiapa, Sal Vieja, San Perlita

More recent antigenic cartography studies using vaccinated human sera (TBEV, JEV and YFV) against a panel of 5 flaviviruses (WNV, TBEV, JEV, LIV, YFV, and DENV-2) demonstrated individual variations in antibody responses. This study demonstrated strong cross-reactivity indicating close antigenic relationship between LIV and TBEV and weaker cross-reactivity with LIV and the mosquito-borne flaviviruses (Mansfield *et al.*, 2011).

Following the emergence of ZIKV as a disease as human public health concern in DENV endemic regions, studies conducted using monoclonal antibodies to DENV showed that DENV IgG was cross-reactive for ZIKV but they were mostly non-neutralising (Priyamvada *et al.*, 2016; Collins *et al.*, 2017; Rathore and st. John, 2020).

Plaque reduction neutralisation tests (PRNT) is the most sensitive of the flavivirus serology assays and is still considered the gold standard method for detecting and quantifying neutralising antibody to a specific virus. It is a modification of the plaque assay in which a dilution series of the test serum is made and incubated with the viral suspension to allow neutralising antibody if present to neutralise the virus. This method is also used for differential diagnosis of flaviviruses as there are co-circulation of flaviviruses (eg ZIKV and DENV).

1.6.4 The challenges posed by flavivirus antibody cross-reactivity

Cross-reactive antibodies to closely related flaviviruses is a growing public health concern due to the emergence of competent vector species into new geographical regions as a result of trade and changes to weather patterns (Pierson and Diamond, 2020). The broad cross-reactivity of flaviviruses has been traced to the E and prM proteins and virus specific serology diagnosis relies on the viral protein NS1 (Datta and Wattal, 2010; Lima *et al.*, 2019).

As mentioned earlier in the chapter, the major antigenic regions on flaviviruses are the envelope proteins and due to the structural similarity among flaviviruses the antibodies are cross-reactive. Virus specific neutralising antibodies are produced against the conserved proteins such as NS1. Pre-existing antibodies to one flavivirus can lead to exacerbation of disease and this is known as antibody-mediated enhancement of infection (ADE) and is seen with DENV infection with a different serotype. Since the emergence of ZIKV in DENV endemic region, studies have shown that pre-existing immunity to one DENV can enhance the disease

upon infection with ZIKV and vice-versa(Bonheur *et al.*, 2021). ADE was reported following natural infection in children vaccinated with the DENV vaccine (Dengvaxia) (Slon Campos, Mongkolsapaya and Screaton, 2018)

Some flavivirus diseases have been brought under control using vaccines. The first and the most successful flavivirus vaccine is the YFV vaccine developed by Max Theiler. Max Theiler passaged the virulent Asibi strain in chicken embryo which had the brain and spinal cord removed. The 17D variant that emerged between the 89th and 114th passage did not have the viscerotropic or neurotropic properties of the parent strain but was immunogenic (Theiler and Smith, 1937). The vaccine was very successful and is still produced using the same method. Max Theiler was awarded the 1951 Nobel Prize in Physiology or Medicine for this work (Norrby, 2007).

1.7 Viruses in this study

The mosquito-borne flaviviruses of focus in this thesis are DENV and ZIKV. The tick-borne flaviviruses of focus are TBEV and LIV. Mosquito-borne flaviviruses are mainly found in tropical and sub-tropical regions of the world whereas tick-borne flaviviruses are present in temperate regions.

1.7.1 Dengue virus

A disease with the symptoms associated with what we now know is caused by DENV has been documented since the 17th century, with outbreaks reported in the 18th and 19th centuries from countries and regions across the globe starting with Java, USA, Zanzibar, India, Hong Kong, with outbreaks reported from Greece, USA, Australia, and Japan. Single serotype outbreaks of DENV in central America and the Caribbean have caused multiple epidemics since the 1950s. Global conflict caused by the World Wars had an impact on disease pathogenesis as the different

serotypes of DENV started circulating in the same regions (Henchal and Putnak, 1990; Holbrook, 2017).

Following the confirmation of the role of the mosquito vector in transmitting YFV, this route of transmission was confirmed for DENV also (Holbrook, 2017). A study conducted by the US army during an outbreak in the Philippines conducted human experiments where volunteers were exposed to mosquitoes which had previously fed on patients and to fomites from the patients to induce disease (Ashburn and Craig, 1907). Studies showed that multiple mosquito vectors were capable of virus transmission in the different geographical locations, mainly *A. aegypti* in the Americas and *A. albopictus* in Southeast Asia (Holbrook, 2017). In 1943 during an outbreak in Japan, Dr Hotta isolated DENV-1 from a patient sample and passaged the virus in suckling mouse brain (Konishi and Kuno, 2013).

The relationship between dengue viruses and other flaviviruses were confirmed through serological work carried out by Sabin, Casals and Brown and Henchal and colleagues. The four serotypes namely DENV-1, DENV-2, DENV-3 and DENV-4 were described and, using neutralisation tests, all dengue viruses were grouped into a single serogroup by Smithburn in 1954 (Henchal and Putnak, 1990). The presence of a specific NS1 common to all 4 serotypes was demonstrated by immunoprecipitation studies (Russell, Chiewsilp and Brandt, 1970). In 2013, a new serotype, DENV-5, was reported to have been identified in Malaysia (Vasilakis *et al.*, 2011; Ortiz García, 2013; Mustafa *et al.*, 2015).

A pattern of DENV outbreak every 50 years used to be observed but with the expansion of the competent mosquito vectors, the disease has become endemic in over 100 countries (Gubler, 1998; Bhatt *et al.*, 2013). Since 2010, autochthonous dengue infections have been reported in parts of Europe including France, Croatia, and Madeira (Marchand *et al.*, 2013; Wilder-Smith *et al.*, 2014).

There were large DENV outbreaks globally affecting more than 2 million cases in 2016 but there was a 70% drop in cases in America during 2017 and severe dengue cases were down 50% in 2017. The reduction has been linked to the rapid expansion of ZIKV in 2016. The number of cases increased over the following years, and in 2019 case numbers peaked globally. High case numbers have been reported in subsequent years from countries in Southeast Asia, Africa and Central America (World Health Organisation, 2021).

The virus has an incubation period of 5–8 days before the onset of fever which coincides with viraemia which persists for up to 6 days. Under guidelines issued by WHO in 2009, dengue is defined by a combination of ≥2 clinical features in febrile patients with travel history to or living in an endemic region. Clinical signs include widespread rash, anorexia, positive tourniquet test, leukopenia, severe headache and retro-orbital pain, arthralgia, and myalgia. There are warning signs which can predict severe dengue, and these include persistent vomiting, mucosal bleeding, abdominal tenderness, and oedema (World Health Organization, 2009).

Disease pathogenesis following DENV infection is complex and is mediated by several host factors such as ADE, memory cross-reactive T cells, anti-DENV NS1 antibodies, previous exposure to a different DENV serotype and autoimmunity. Viral factors such as NS1 protein, sub genomic flavivirus RNA (sfRNA) and genomic variant of DENV also play a role in the severity of disease (Gould *et al.*, 1990; Bhatt *et al.*, 2021). The link between ADE and enhanced infection was observed for mainly DENV serotypes. ADE can lead to severe dengue syndrome which includes dengue haemorrhagic fever (DHF) and dengue shock syndrome (DSS). Symptoms seen include shock or oedema due to severe plasma leakage, severe bleeding or impairment to organs including heart or impaired consciousness. In pregnant women, perinatal transmission of DENV has been reported and this can cause symptomatic infection including haemorrhage in the new-born (Henchal and Putnak, 1990; Gubler, 1998).

Following recovery, the individual develops lifelong immunity to the serotype they were infected with. There are no specific treatments for the disease other than supportive care to alleviate symptoms. Virology and molecular methods of detection are successful during the viraemic period.

In 2009, the WHO published a guideline which had the current landscape of assays used (World Health Organization, 2009). Virus isolation from samples using mosquito and mammalian cell-lines, antigen detection using immunofluorescence assay (IFA) or other staining methods. Molecular detection methods of reverse transcriptase PCR (RT-PCR), single plex and multiplex quantitative (RT-qPCR) are used and isothermal amplification methods of nucleic acid sequence based amplification (NASBA).

Serology assays including IgM antibody-capture enzyme-linked immunosorbent assay (MAC-ELISA) which produces good distinction when used in differential diagnosis but not specific to detect between DENV serotypes. Other serology assays are IgG capture ELISA (GAC-ELISA) for detecting IgG, HI assays which are cross-reactive for other flaviviruses.

New methods of serology detection are in development including the application of mass spectrometry(Wee *et al.*, 2019), microarrays, and microsphere immunoassays. Antigen-specific detection methods have been developed using the NS1 as a specific antigen, which has been used to investigate cases where viraemia has elapsed (Datta and Wattal, 2010).

1.7.2 Zika virus

ZIKV is a member of the Spondweni serocomplex and as a mosquito-borne flavivirus, it is also closely related to JEV, WNV and DENV 1-4 serogroup (Lanciotti *et al.*, 2008; Haddow *et al.*, 2016). It is transmitted by female *A. aegypti* and *A. albopictus* mosquitoes. ZIKV was first isolated from a sentinel Rhesus monkey in the Zika forest in Uganda (Dick,

Kitchen and Haddow, 1952) and in *Aedes africanus* mosquitoes the following year. The first human case was reported in 1964 and the virus caused sporadic cases in parts of Africa and southeast Asia until 2007 (Kucharski and Riley, 2016). Genotyping confirmed two lineages (Asian and African) of ZIKV (Faye *et al.*, 2014).

A major outbreak was reported in 2007 from Yap Island in the Federated States of Micronesia which affected 73% of the population. Between 2013 and 2014, the virus was causing outbreaks in neighbouring islands (Musso, Nilles and Cao-Lormeau, 2014; Roth *et al.*, 2014). In mid-2015, ZIKV was reported from Brazil and led to a large epidemic which then spread to other countries in South America, the Caribbean and parts of the US. ZIKV infections peaked in 2016. Other routes of ZIKV transmission include infected donated blood products and sexual transmission of the virus which was reported in 2008 (Foy *et al.*, 2011). The sexual transmission of ZIKV has been linked to the epidemiological pattern where more women were infected during the outbreak compared to men.

The disease symptoms are usually mild and can include fever, rash, arthralgia, and conjunctivitis. ZIKV can result in neurological disorders such as Guillan–Barre syndrome, myelitis, and neuropathy. Although it has been widely reported that most ZIKV infections are asymptomatic with only 20% infections resulting in symptomatic infection, more high quality data is required to ascertain accurately determine the proportion of asymptomatic and symptomatic cases (Haby *et al.*, 2018).

In pregnant women, ZIKV infection can lead to congenital abnormalities including microcephaly (Calvet *et al.*, 2016; de Araújo *et al.*, 2018) in the developing fetus and can result in loss of pregnancy or pre-term birth. The severity of the outbreak led WHO to issue alerts regarding the association of the virus with microcephaly (WHO, 2016). Studies have shown that the rate of birth defects is higher in women who were

symptomatic and in those who were exposed to the virus during the first two trimesters (Lessler *et al.*, 2016). The virus has been isolated from a range of biological specimens including blood, semen, urine, amniotic fluid, saliva.

ZIKV cases are diagnosed molecularly using RT-PCR and results are more accurate when samples are tested after day 7. ELISA is used for serological diagnosis, but previous exposure to DENV can result in cross-reactive response on ZIKV ELISA.

1.7.3 Louping Ill virus

LIV is a tick-borne flavivirus and a member of the TBEV serogroup. There are four lineages of LIV in Britain, and they are geographically separate with genotype 1 present in England and Scotland, genotype 2 only in Scotland, genotype 3 only in Wales, and genotype 4 only present in Ireland (Gao *et al.*, 1997). Geographically, the virus is mainly found in the uplands of the British Isles, but it has also been reported elsewhere in Europe such as Norway, Spain, and Turkey.

It was the only known flavivirus endemic to the British Isles and is classified as an ACDP hazard group 3 pathogen, due to its zoonotic properties. LIV is very closely related to TBEV. The virus gets its name from a classic sign exhibited by the infected animal which appears to leap, "to loup" is Scottish for to leap.

The virus is transmitted to its vertebrate hosts through the tick vector *I. ricinus*, (

Figure 1-3) and disease transmission of LIV follows the activity pattern of ticks, which are most active in autumn and spring, and the virus is able to replicate in the vector and survive the transitions between moulting stages (Hudson *et al.*, 1995). The tick larva and nymph usually feed on smaller mammals and birds whereas the adult life stage feeds on larger animals such as sheep, deer, cattle and less frequently dogs and humans.

Apart from LIV, this vector also transmits bacterial pathogens *Anaplasma phagocytophilum* (*A. phagocytophilum*) the causative agent of tick-borne fever, and *Borrelia burgdorferi* (*B. burgdorferi*), which causes Lyme disease. The bacterial co-infection is a factor in enhancing viral disease in the vertebrate host as infection with two pathogens can dampen the immune response. The existing research carried out on LIV has been extensively covered by two reviews (Jeffries *et al.*, 2014; Buxton and Reid, 2017).

LIV is neurotropic and causes louping ill disease, a viral encephalomyelitis which mainly affects sheep in Britain. The disease has an economic impact due to the loss of sheep and red grouse in endemic areas. LIV can infect other species including hare, deer, cattle, horses, dogs, alpacas, and occasionally humans as well (Davidson, Williams and Macleod, 1991).

The disease has been known as early as 1879 due to the losses incurred to sheep farmers. The virus was identified in 1930 (Pool, Brownlee and Wilson, 1930). Other seminal works include disease progression of LIV in sheep and red grouse, exacerbation of disease with co-infection with Anaplasma, vector competence of *I. ricinus*, vaccine development to name a few were carried out in the 1930s and 1970s (Brownlee and Wilson, 1932; Gordon *et al.*, 1932; Macleod and Gordon, 1932; Brotherston and Boyce, 1970; Doherty and Reid, 1971; Reid, 1975).

Infected sheep and red grouse have elevated virus titres (viraemia) for approximately 5 days after infection, which is sufficient for passing the infection to a naïve vector during blood meals thus sustaining the infection cycle. Amplifying hosts implicated in the infection cycle of LIV include mountain hare, deer and red grouse. Red grouse chicks are highly susceptible to the virus with high mortality rates in young chicks (Reid, 1975; Reid *et al.*, 1978), and this has an economic impact on estates which offer commercial grouse shooting.

In sheep, the progression of disease is biphasic, and starts with an incubation period of 6–18 days accompanied by viraemia, and initial clinical signs include fever, depression, and anorexia. In the second phase, clinical signs include neurological signs such as muscular trembling, unsteady gait, leaping, followed by seizures, paralysis and death. Sheep that survive the neurological disease phase can be left with neurological impairment. Severe clinical disease is generally (but not necessarily) observed in animals co-infected with *A. phagocytophilum,* which acts as an immunosuppressant agent. Sheep with asymptomatic disease develop sterile immunity, which protects them from subsequent exposure. Lambs born to ewes with pre-existing immunity to the virus are protected for the first few months of their life until the maternal antibodies have waned, and they encounter the virus for the first time when they start living on hill pasture.

Disease confirmation can be carried out using molecular and serological methods. molecular detection uses RT-PCR (Marriott *et al.*, 2006). Serology for LIV uses techniques including HI, serum neutralisation tests (SNT), PRNT. Some of the serological assays (SNT and PRNT) and the molecular assays require handling of virus, organs or tissues at CL3 and these specialised facilities are not available to all diagnostic centres.

Surveillance and diagnosis in sheep, hare and grouse is currently carried out using HI assay, which has lower containment requirements than the other assays but is still very time consuming. ELISA-based high throughput assays which can be carried out on the bench in a smaller lab are not available at present.

LIV vaccines have been developed and used to protect sheep that live in LIV endemic areas. The initial vaccine was made using sheep neural tissue suspension in formalin and was in use since the 1930s, which was then later replaced with an inactivated vaccine strain of the virus prepared in an oil emulsion and used until 2017 when the vaccine

production was halted due to production difficulties (reference: personal communication with Dr Mara Rocchi).

A recombinant vaccine made using Semliki Forest Virus (SFV) particles expressing the LIV NS1 and prME proteins was protective in mouse challenge studies (Fleeton *et al.*, 1999) Currently there are vaccine trials in progress (Game and Wildlife Conservation Trust, 2021).

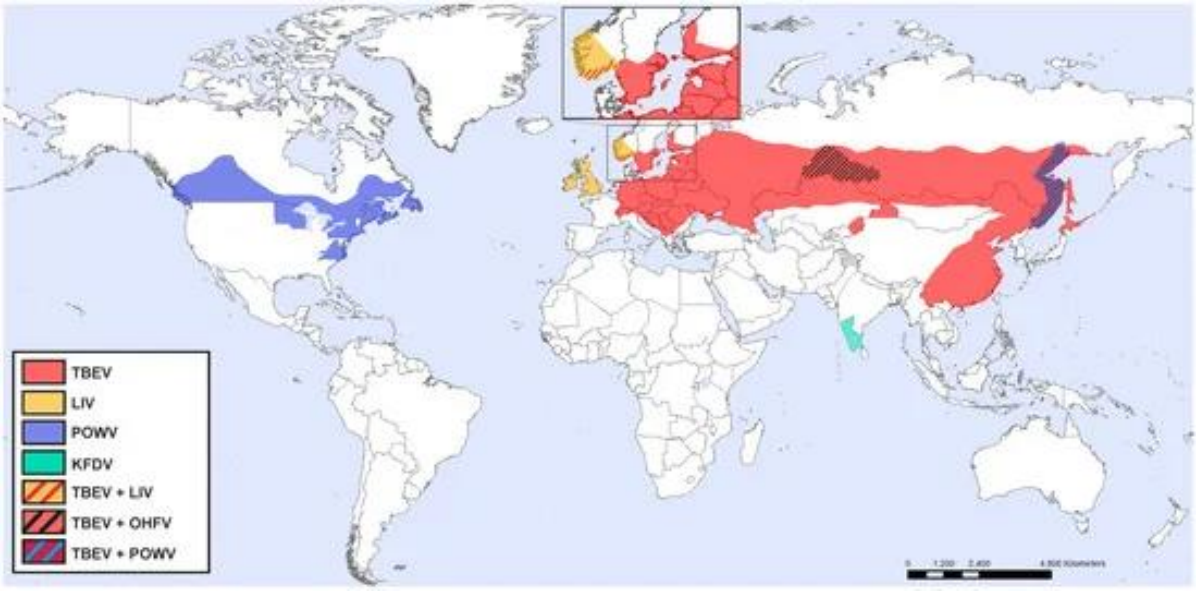


Figure 1-5: Map showing geographical distibution of LIV, TBEV and other tick-borne flaviviruses pathogenic to humans.
Figure from (Lindqvist, Upadhyay and Överby, 2018).

1.7.4 Tick-borne encephalitis virus

TBEV is a tick-borne flavivirus and belongs to the tick-borne serogroup. The earliest reports of severe TBE disease in humans were in 1932 and they were reported from Far-Eastern region of current Russia. Upon investigation by the authorities, the disease was linked to a tick vector- *I. persulcatus*. The virus was discovered in 1937 and named Far-Eastern encephalitis virus (formerly known as Russian Spring Summer encephalitis virus). A milder form of the disease linked to consuming

unpasteurised milk was detected in Eastern Europe and Western Russia and named Western Encephalitis (biphasic milk fever). The associated tick vector in this case was *I. ricinus*. Subsequently a Siberian variant of the encephalitis virus was discovered, vectored by *I. persulcatus.* Following establishment of their similarity in disease presentation and antibody cross-reactivity, they were discovered to be variants of the same virus and arranged in the order of disease severity, the three variants were TBEV-Eu, TBEV-Sib and TBEV FE. TBEV-Eu is most common type seen in western Europe (Holbrook, 2017). A recent global distribution of tick-borne flaviruses can be seen in Figure 1-5.

Although TBEV can infect animals and birds, it is only known to cause symptomatic disease in humans. Serum antibodies for TBEV have been detected from domestic animals such as sheep, goats, dogs, and wild animals such as small rodents (voles, mice), deer, wild boar, and migratory birds. Occasional clinical disease has been reported in horses, sheep, goat and monkey (Mansfield *et al.*, 2009; Klaus *et al.*, 2014; Michelitsch *et al.*, 2019).

Disease transmission occurs when people are exposed to questing ticks when out exploring forests and other tick infested habitats (Mansfield *et al.*, 2009). Transmission through unpasteurised milk and milk products (Caini *et al.*, 2012; Brockmann *et al.*, 2018; Wallenhammar *et al.*, 2020) also occurs. More recent sub-typing using molecular methods has revealed that there are 7 strains of TBEV (Deviatkin *et al.*, 2020)

The disease manifestation can be dependent on the strain and host factors including inoculation dose, age, and overall health. Disease can be asymptomatic or sub-clinical, in TBEV-Eu the percentage of asymptomatic infection is between 70-90% and for TBEV-Sib, up to 98% cases are asymptomatic (Růžek, Dobler and Mantke, 2010). In symptomatic cases, TBEV is neurotropic, and TBEV-FE and TBEV-Sib infections are mainly monophasic with symptoms including fever, meningitis, encephalitis with

long term neurological damage, chronic infection, with a case fatality ratio of 35% and 2%, respectively.

TBEV-Eu causes a biphasic disease with an average incubation period of 7—14 days. Initial symptoms include fever, fatigue, malaise, leukocytopenia, gastrointestinal symptoms, elevated liver enzymes. The symptoms subside for a week before the second phase presents with high fever, severe headaches and neurological symptoms of meningitis and acute encephalitis and fatality in 2% cases. In 50% cases, there are long-term sequelae including severe neurological disorders and spinal paralysis (Růžek, Dobler and Mantke, 2010).

The first vaccine was made in Russia and used to protect workers in endemic regions against TBEV-FE. This was an inactivated virus vaccine. Two vaccines were developed for TBEV-Eu, the first made using strain Neudörfl was named FSME-IMMUN® (Baxter AG, Vienna, Austria) and the second was made using K23 strain and named ENCEPUR® (Novartis/GlaxoSmithKline, Germany). Both vaccines are still in use and have high immunogenicity and provide good protection (Holbrook, 2017). There are two Russian made vaccines as well. Even with the availability of 4 successful vaccines, there are 12,000 cases reported from Europe annually with more cases reported from countries with lower vaccine rates (Amicizia *et al.*, 2013).

Phylogenetic analysis using NS3 has shown that TBEV-Sib and TBEV-Fe are closely related to each other, and TBEV-Eu is more closely related to LIV (Figure 1-2) than either of the other TBEV subtypes (Mansfield *et al.*, 2009).

TBEV-Eu is predominantly found in Europe and Russia. TBEV is endemic in at least 27 European countries but had never been reported in the UK until 2019 when it was detected in deer serum from East Anglia and Hampshire (Holding *et al.*, 2019). A probable human case of TBEV has

also been reported in a visitor to the Hampshire focus (Kreusch *et al.*, 2019).

The detection of TBEV in the UK is very important for public health. The right tick vector is already present in the UK and amplifying hosts are also present. Sero-surveillance is an important tool to identify virus establishment especially as TBEV is highly pathogenic human disease which is a notifiable disease in many countries where it is endemic. A serology assay which can distinguish between the two viruses in single infected or co-infected serum sample will be required in the future.

1.7.5 Developing virus-specific diagnostics

New diagnostic assays which can distinguish between closely related flaviviruses circulating in the same geographical regions has become a priority. One approach used for discovering virus specific antigenic epitopes is by interrogating the antibody population in serum. The antibodies raised against the flavivirus envelope proteins are cross-reactive, but those raised against the non-structural proteins are virus specific. Several methods are available to study the antibody-antigen interaction and most of them rely on interacting the antibodies with the antigen (epitope mapping) or by using peptides that mimic the epitopic region (mimotope discovery) (Abbott, Damschroder and Lowe, 2014).

Advances in computation along with peptide based techniques such as peptide array and next generation phage display have been used to discover virus specific epitopes, some which are only present in certain host populations such as pregnant women (Datta and Wattal, 2010; Kam et al., 2019; Clapham et al., 2021; Fumagalli, Figueiredo and Aquino, 2021; Yap et al., 2021).

In peptide microarrays, a library of peptides of overlapping or non-overlapping segments of the target polypeptide region of the virus is bound to a chip. The target serum antibodies are incubated with the

peptides. The antibody binding to the antigenic region is recognised by a colorimetric or fluorescent reporter and this response is read computationally. The computational element enabled quicker screening and identification of binders. This technique has been used for flavivirus epitope mapping studies (Bergamaschi *et al.*, 2019; Clapham *et al.*, 2021).

Phage display meanwhile uses a library of phage expressing the peptide on its coat protein. This technology has been used since the 1980's for peptide and antibody discovery for a range of applications. The benefits of phage display are that the library is a biological resource and can be grown from a stock. It is also a cheaper method than peptide microarray, has a large diversity of clones and recent advances in the technique such as computationally identifying clones of interest has made it the choice for this experiment.

1.8 Phage display

1.8.1 Bacteriophage

Bacteriophage (phage) are one of the largest family of viruses and as their name suggests are viruses that infect bacteria. They were independently described for the first time by Fredrick Twort in 1915 (Twort, 1915) and Felix D'Hérelle in 1917 (D'Hérelle, 1917) after their presence was associated with an unknown cause of antibacterial activity observed in the late 19th century and early 20th century. D'Hérelle chose the name bacteriophage as he observed the bacteria being 'eaten' by the virus. Due to their ability to infect bacteria, phage have been studied extensively for a range of applications including phage therapy as an alternative to antibiotics and phage display, which is the application that will be the focus here.

Phage families are organised based on their structure and the phage used in this report, M13 bacteriophage, belongs to the *Inoviridae* family, is a filamentous coliphage and infects their host bacteria through the F pilus.

1.8.2 M13 phage

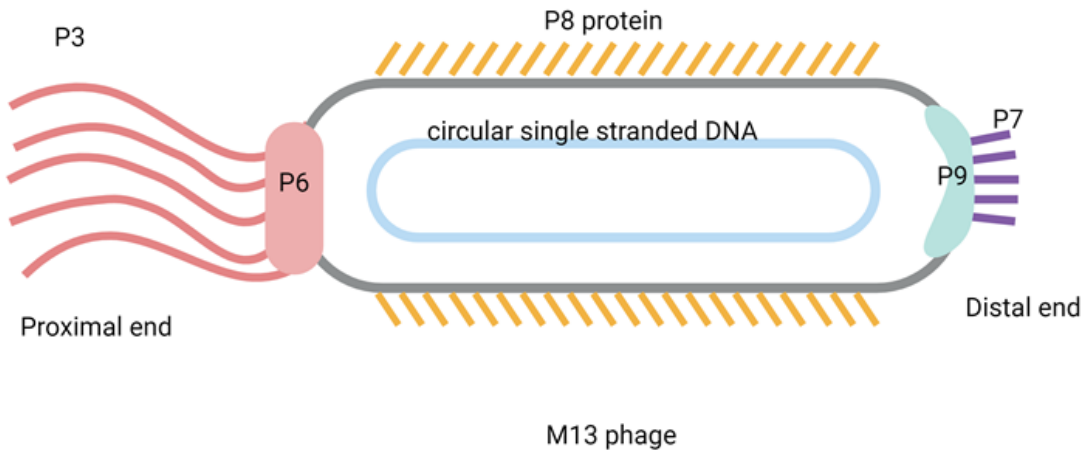


Figure 1-6: Illustration of M13 bacteriophage re-drawn from (Stopar *et al.*, 2003).

There are two forms of viral replication: lysogenic and lytic. Lysogenic phage genome integrates with the host cell genome or exists as a replicon in the cell cytoplasm (prophage). The host cell becomes a virus production unit without its membrane or cell wall being destroyed for the phage progeny to exit the host cell as would be the case with infections with lytic viruses. During bacterial cell division, the prophage is transmitted to the daughter cells.

M13 is a lysogenic phage which only infects *Escherichia coli* (E. coli) that possess F1 pili (Ray, Bscheider and Hofschneider, 1966) and is therefore an F-specific filamentous phage (Ff phage). It possesses a circular single stranded DNA (ssDNA) genome which is 6.4kb in length and encodes for 11 genes. The ssDNA is encased in a thin cylinder which is covered by the major coat protein P8, encoded by gene 8 (VIII). P8 is about 50 amino acids long (Henry and Pratt, 1969) and there are around 2700 copies of the protein. At one end of the phage, there are proteins P3 (III) and P6 (VI) and at the opposite end of the phage there are proteins P7 (VII) and P9 (IX) as seen in Figure 1-6. The viral protein P2 is associated with the translation and transcription process of the phage genome. Protein P3 attaches to the bacterial pilus and the retraction of the pilus into the cell

allows phage DNA to be injected into the cell (Tzagoloff and Pratt, 1964) which starts the phage replication cycle.

In M13 phage infection, the positive sense DNA from the phage enters the host cell cytoplasm. The complimentary strand is synthesised by the bacterial enzymes. This double stranded phage DNA forms a parental replicative form (RF) in the host. The phage proteins P2, P5 and P10 are involved in the replication of phage DNA and the initiation of single stranded viral DNA for creating the progeny phage. Due to the lysogenic process, the phage progeny are able to leak out of the bacteria without disrupting the cell wall and membrane (Karimi *et al.*, 2016).

1.9 Phage display

Phage display was first described by GP Smith in 1985 (Smith, 1985) and was awarded the Nobel Prize for Chemistry in 2018. Foreign DNA fragments were cloned into gene III of M13, and the resultant polypeptide was displayed on the coat protein of the phage, which retained its infectivity. This fusion phage could be significantly enriched over the wild-type phage by affinity purification using antibodies directed against the expressed foreign sequence in a process called biopanning.

There are mainly two types of libraries used in phage display: peptide display and antibody display. Out of all the M13 coat proteins, P3 and P8 are mainly used for phage display. The phage library used in this project was a P8 linear random peptide library expressing 16-mer peptides on the major coat protein, P8. P3 allows for larger peptides >100 amino acid residues (Iannolo *et al.*, 1995) at a lower abundance of 3–5 copies of peptides per phage, and phage libraries expressing CDR3 region of antibodies usually use P3 expression In comparison, P8 is used for displaying smaller peptides of 6-22 amino acids (aa) but P8 has a higher abundance of over 2000 copies in wild–type (wt) phage, of which, 200-300 are transgenic in a peptide library constituent.

A major limiting factor in phage display is the loss of coat protein assembly when trying to display larger peptides and this is resolved by using the phagemid display system as described below (Smith and Petrenko, 1997; Tikunova and Morozova, 2009; Pande, Szewczyk and Grover, 2010; Bazan, Całkosiński and Gamian, 2012).

In type 8+8 phagemid system (Figure 1-7), two copies of gene VIII on separate genomes are used. The wt gene is on a 'helper phage' which contains a defective origin of replication and the other being a plasmid containing both a bacterial and a phage origin of replication, hence the term phagemid (Qi *et al.*, 2012). The phagemid also contains a pelB leader sequence, and this allows the expressed peptide to be transported across the bacterial inner membrane to the periplasmic space.

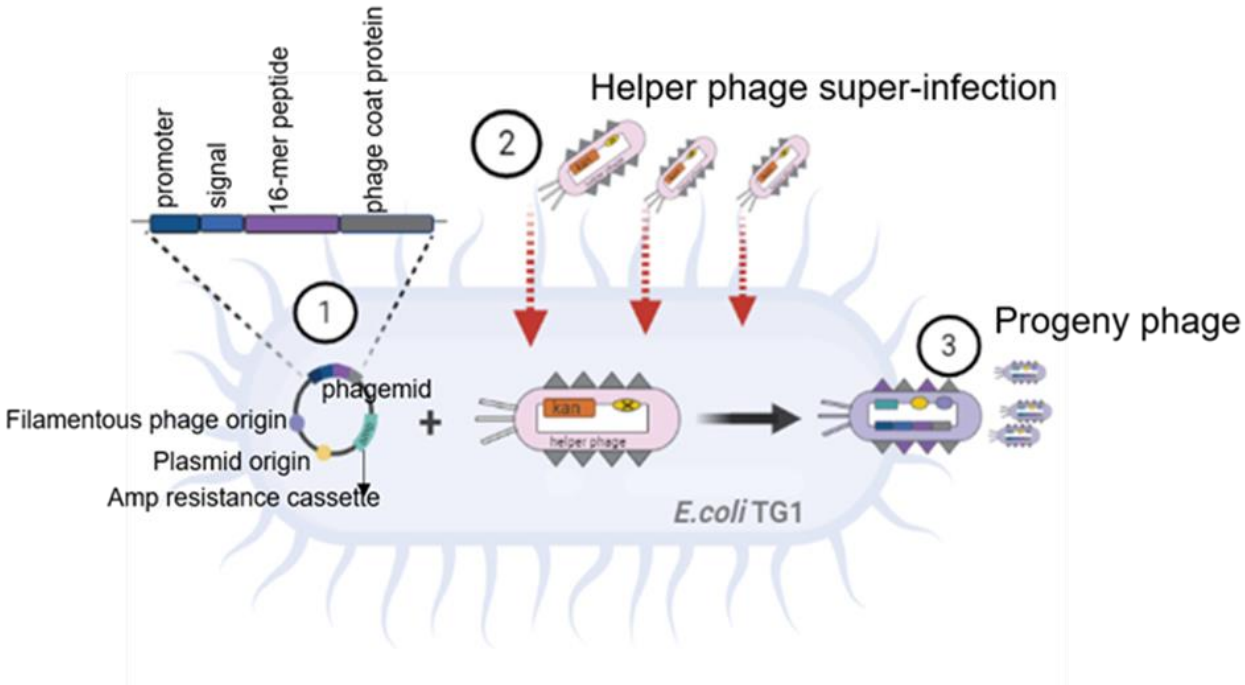


Figure 1-7: P8+8 phage display system.
1 shows the phagemid plasmid with the foreign sequence, wt phage origin of replication transfected into the *E. coli*. **2** helper phage infection which provides the phage genes necessary for progeny phage production. The helper phage has a defective origin of replication which makes it inefficient in self packaging. **3** progeny phage displaying the foreign peptide.

The helper phage provides the structural proteins required for generating a complete virion for the phagemid. Phage replication protein on the helper phage starts the production of progeny by switching on the phage origin in both the helper phage and the phagemid. The resultant progeny has a mosaic coat composed of recombinant and wt. P8. The defective origin of replication in the helper phage suppresses the incorporation of helper phage genome into packaged phage.

Some of the benefits of using the phagemid system compared to phage display are that the genome of phagemid is smaller, allowing the insertion of larger fragments for display (Breitling *et al.*, 1991) and higher transformation efficiency of the phagemid allows higher library diversity (Qi *et al.*, 2012).

In the random peptide library (RPL), the exogenous inserts are made up of chemically synthesised 'degenerate' oligonucleotides. The degenerate sequence [NNK] x 16 is used where each N is an equal mixture of the nucleotides A, C, T or G and the K is an equal mixture of nucleotides G and T. Every NNK triplet is a mix that includes codons for all the natural amino acids and the entire sequence is an equimolar mixture that can cover for a wide range of peptides across different species (Smith and Petrakos, 1997). There are theoretically 20×10^{16} potential combinations of 20 amino acids in a 16-mer peptide library. Despite this the diversity of a typical library is in the order of 1×10^{8} to 1×10^{9}, which would still be expected to contain several potential binders against a particular target.

Phage libraries are subdivided further based on how the peptides are display by the library. In a linear peptide library, the peptide insert is not expressed as a three dimensional structure. Whereas in a constrained library, the peptides are presented conformationally through the addition of constraints such as disulphide bonds between 2 cysteine residues. The benefit of having a constrained library is that the peptide might be

accessed better on the active site of the target surface than one expressed by a linear library.

1.10 Biopanning

Biopanning is the term coined for the process of affinity purifying a peptide expressed on the phage coat to the antigen of interest (Parley and Smith, 1988).

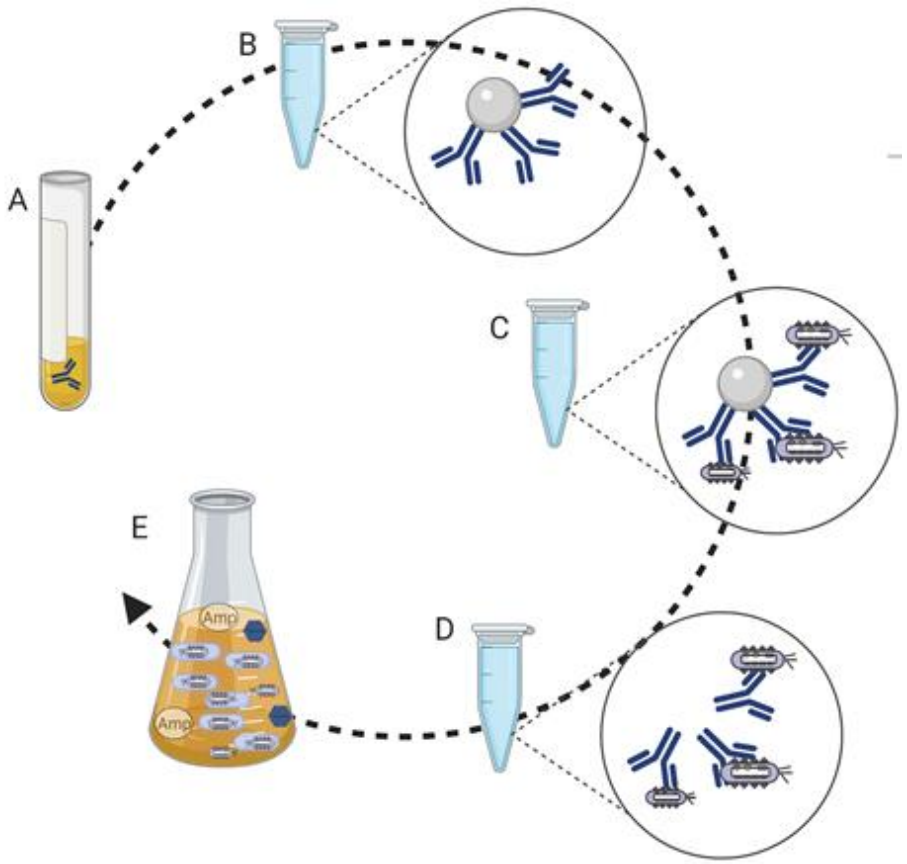


Figure 1-8: Biopanning process.
[A] shows the target antibody for biopanning, [B] target bound to solid surface. [C] phage library incubated with the immobilised target, [D] eluting the target following biopanning. [F] growing the rescued phage

In each round of biopanning (Figure 1-8), the target, bound to a solid support, is incubated with the phage library. The phage which are bound

to the target are eluted using a shift in pH, rescued by infection into a suitable host bacterium displaying an F' pilus for M13 phage, and more phage subsequently produced by superinfection with helper phage.

This process is repeated for several rounds and phage that exhibit peptides which mimic the epitope (mimotopes) of the target are enriched, causing a reduction in the original library diversity between the first and second round of biopanning. The term mimotope was coined by Mario Geysen in the 1980s to refer to small peptides that bind to a receptor binding site and mimic the epitope region on the antigen. The mimotopes will probably not have the same amino acid sequence as the epitope but are useful in applications such as epitope mapping and diagnostic assay development (Smith and Petrenko, 1997).

The sub-population that are enriched through all the rounds are expected to have been enriched as a result of their capacity to bind the target antibody. The "fittest" clones are not necessarily specific to the target, especially if the target is in a matrix such as serum, as this contains a range of proteins to which the phage peptides could interact. To reduce enrichment of the non-specific clones, strategies to remove non-specific phage are used, where binders to reagents e.g. sera, plastic ware, etc. are removed by subtractive panning. Here non-specifically binding phage are taken out of the library prior to biopanning by introducing the phage library to the non-specific target and collecting the resulting supernatant which has been depleted of binders (Smith and Petrenko, 1997).

Following biopanning, the enriched sub-population of phage that had bound to the target is rescued and bacterial glycerol stocks are made. There are mainly two ways to identify phage clones of interest: traditional phage display and next generation sequencing the output library.

The biopanning experiments for flaviviruses have used antibody libraries being panned against specific proteins such as NS1 or biopanning with

specific peptide libraries generated against the virus of interest (Lebani *et al.*, 2017; Shriver-Lake *et al.*, 2018; Rajan *et al.*, 2021).

1.10.2.1 Traditional phage display

Bacterial glycerol stock of the phage output library is serially diluted and plated onto selective growth media. Individual bacterial colonies are picked and as each bacterial cell will only have one phagemid, the 16-mer peptide encoded in the phagemid can be identified through Sanger sequencing the bacterial colony. This individual phage clone can be grown in the same process used to grow the phage library, by superinfecting with helper phage, and the phage supernatant or the precipitated form is used in ELISA against the target. Multiple phage clones are picked and the success of the (phage) clone lies in the ability of the peptide to detect the target in a specific and sensitive manner. This process is known as traditional phage display.

As the bacterial colony will only have one phagemid which creates an individual phage particle expressing the same peptide on its fusion coat, the terms phage clone/peptide can be used interchangeably.

While traditional phage display is a good method to identify phage clones, the process of picking clones, Sanger sequencing, testing on ELISA can be rate limiting and inefficient if the biopanning sample matrix is complex.

1.10.2.2 Next generation sequencing of the output phage library

The reduction in cost of next generation sequencing (NGS) has made it possible to sequence all output phage clones from both the final panning round as well as monitoring enrichment of a particular peptide sequence through the panning rounds. The advent of next generation sequencing technologies has changed the landscape of peptide/antibody discovery using phage display(Dias-Neto *et al.*, 2009).Next generation sequencing

of output phage libraries is termed **'next-generation phage display'** (NGPD) (Ryvkin *et al.*, 2012).

Next generation sequencing the output phage has made it easier to identify clones that are under-represented and to adapt panning strategies to allow for maximum diversity of clones (Matochko *et al.*, 2012). It is also a very useful method for identifying peptide motifs specific to an organism (Hurwitz *et al.*, 2017).

Of the several NGS platforms available, Ion Torrent sequencing platform (Life Technologies) requires very small sample volume, is very affordable whilst giving good depth of reads (Quail *et al.*, 2012). It uses semiconductor technology to identify the nucleotides being added during the sequencing run. Whilst the Illumina platform has a lower error rate, it is much more expensive. The workflow had been optimised for use with the Ion Torrent platform(Naqid, et al., 2016).

The phage DNA samples are 'barcoded' whereby a unique nucleotide sequence is added to each sample, and this allows for differentiation between samples after NGS. The Ion Torrent sequencing uses a two-step process of sequencing and in the first step, the barcoded sample is amplified on beads known as Ion Sphere Particles by emulsion PCR and sequenced by synthesis on a silicon wafer chip with proton sensors. Nucleotides are introduced one at a time and the chip detects the protons which are released during the addition of the complementary nucleotide.

This method of barcoding allows simultaneous analysis of multiple library outputs which are individually barcoded with a DNA barcode, thus being an economical approach in the field of NGS given the depth of reads provided by an Ion Torrent S5 machine 540 chip used in these experiments vastly exceeds the number of reads necessary to identify enrichment within any one particular experiment.

The data is generated as a FASTQ file, which is analysed to identify peptide mimotopes of interest. A typical analysis pipeline would begin with separating individual DNA sequence files (de-multiplexing) from the single DNA sequence file using the individual barcoded samples, reverse complementing the samples (library dependent), translation of DNA to protein to identify the correct reading frame, identification of amino acid sequences which flank the library variable region, quantification of this region relative to the control panning experiment or library, binning the peptide reads into individual files and also including quality control actions such as omitting sequences that do not match the expected sequence length (Dias-Neto *et al.*, 2009; Ryvkin *et al.*, 2012; Ravn *et al.*, 2013; Rebollo *et al.*, 2014)

The data sets are then analysed to identify enriched peptides, for example using Z-score analysis (Naqid, et al., 2016) .This approach allows for a data led identification of peptides of interest which can then be synthesised and used for developing immunoassays. This approach was successfully used in the group to develop a DIVA assay, traditionally used in veterinary medicine, to distinguish antibody responses between vaccinated and animals exposed to wild type infection (Naqid, et al., 2016; Naqid, et al., 2016b). This approach could also be applied to identify virus specific peptide mimotopes by biopanning serum IgG positive for singular or multiple flaviviruses.

1.11 Aims and objectives

The aim of this project was to apply traditional and next generation phage display techniques to develop highly specific flavivirus serological tests from human and animal sera.

Objectives

- Biopan a peptide phage library with serum IgG to generate a phage sub-library enriched with peptide mimotopes to the target.

- Develop a novel method of analysing NGS data from biopanning studies in order to generate a target specific peptide immunosignature using closely related pairs of flaviviruses namely ZIKV and DENV, and LIV and TBEV
- Apply this immunosignature as an *in-silico* assay to calculate sensitivity and specificity when tested against target and control sub-libraries not used in generating the immunosignature.
- Develop a phage ELISA to detect LIV antibodies in infected sheep serum samples.

1.11.1 Ethical approval for study

Ethical approval was granted for this project (2764 190603) by the School of Veterinary Medicine and Science's ethical review board, and it covered the use of serum from various sources.

Chapter 2 **Materials and Methods**

2.1 Phage Display

The schematic below shows how the phage (P8 library phage) used in the biopanning experiments was generated. This method was also used to generate individual phage clones of interest.

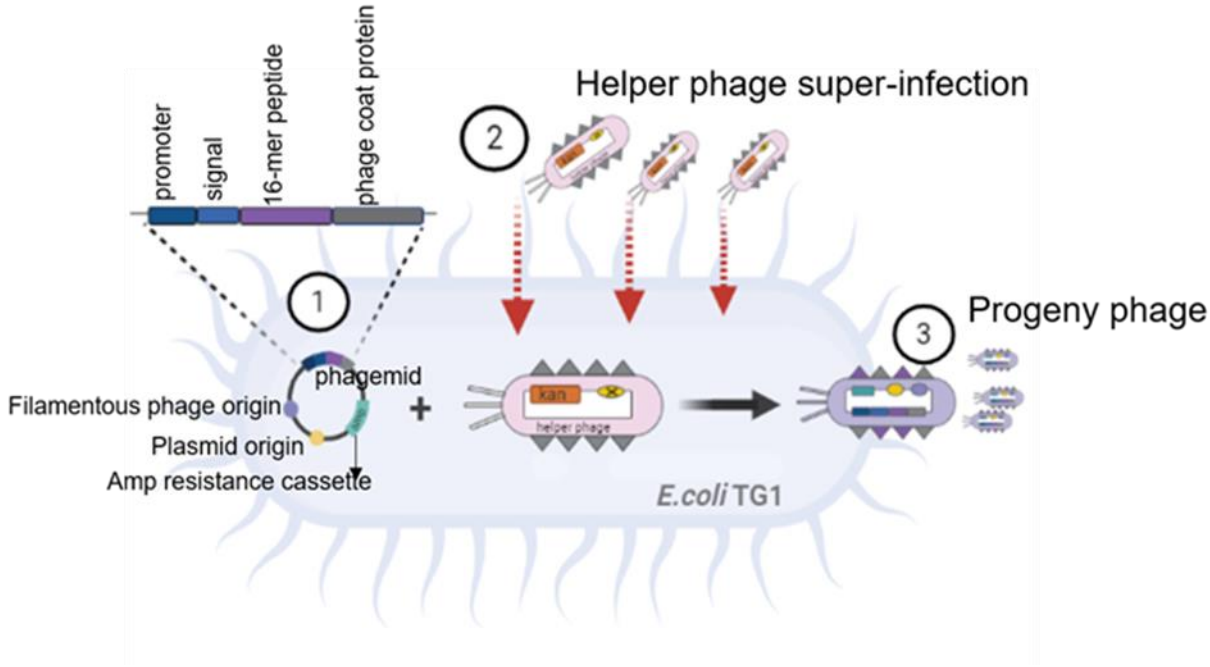


Figure 2-1: Schematic to show how the library phage is produced.
1: Bacterial cell transformed with a single phagemid vector, with ampicillin resistance cassette, plasmid origin of replication, and the components required for the library phage production such as the region coding for the 16-mer peptide that will be expressed on the phage coat and the phage origin of replication, **2**: Bacteria is superinfected with helper phage, shown in pink. Helper phage triggers the phagemid origin of replication to make progeny library phage preferentially over its own progeny as it has a defective origin of replication. **3** Library phage progeny expressing the 16-mer peptide on its coat protein, is released from the bacteria cell. Image made in Biorender.

2.1.1 Phagemid library

The phagemid library (Figure 2-1[1]) used in the biopanning experiments was provided by Maria Tsoumpeli (Tsoumpeli *et al.*, 2022). The library is based on the pc89 PVIII phage library (Felici *et al.*, 1991). It is a 16-mer random peptide library and the peptides were expressed on the major coat protein, P8, of the progeny phage. The diversity range of the in-house library was ~$1x10^9$ plaque forming units (pfu/mL), but it was used in the individual biopanning experiments at a titre of ~$1x10^{12}$ - $1x10^{14}$ pfu/mL.

The library was made by cloning a random 16-mer peptide insert into the vector 'pc89_new vector' through inverse PCR using degenerate primers. An NNK codon was used for the purpose of introducing diversity in the library. The library was then transformed into *Escherichia coli* (*E.coli*) TG1 bacteria and was maintained under ampicillin antibiotic selection as the phagemid provides the TG1 (bacterial) cells with an ampicillin resistance gene.

The library was sequenced using Sanger sequencing and NGS (Ion Torrent) for quality control and to confirm the depth of diversity. Analysis demonstrated that over 80% of the sequenced clones were in the correct frame and half of them contained an amber (UAG) stop codon.

The phagemid P8 library requires the wt M13 genes provided by helper phage to produce progeny phage and M13KO7 was the helper phage used in all the phage production experiments.

2.1.2 TG1 bacteria

The phagemid library is maintained in the electrocompetent strain of *E. coli* TG1 (Lucigen) is the bacterial strain that the phagemid library was maintained in. TG1 is F^+ and an amber suppressor strain (supE).

Mutations in the coding regions of genomes resulting in any of the three stop codons can lead to premature termination of the translation process. As the termination efficiency of the nonsense codons is not 100% and there are alternate functions assigned to the stop codons, it is possible to read through the stop codons. There are three types of stop codon suppressors namely amber, ochre and opal. Amber suppressors can read through the stop codon UAG *supE* is an amber suppressor tRNA and TG1 are a strain that have this allele. The suppression efficiency of amber strains can be as high as 35% (Singaravelan, Roshini and Munavar, 2010) compared to opal and ochre. In TG1, a glutamine (Q) is incorporated into the translated peptide sequence instead of the peptide sequence being terminated.

Naïve TG1 bacteria which were not infected with phage were maintained on Minimal agar plates (Table 2-1). The 5x M9 salts used in making Minimal Agar was made as shown in Table 2-2. Minimal agar was supplemented with thiamine to select for bacteria which possess F' pilus. F' pilus is required for successful phage infection as the M13 bacteriophage P3 protein attaches to the F' pilus of the TG1 cell during phage infection. TG1 bacteria grown from a single colony were streaked onto freshly made Minimal Agar plates and following a 24-hour incubation at 37°C, the plates were stored at 4°C for a period of up to 3 weeks.

Table 2-1: Recipe for Minimal agar

Minimal agar	100 mL for 4 plates
Agar	1.5 g
Water	75 mL
Autoclave agar solution, allow to cool to 50°c, then add the following	
5 x M9 salts	25 mL
25% Glucose	1.6 mL
$MgCl_2$	100 µL
Thiamine Hydrochloride	50 µL

Table 2-2: Recipe for 5x M9 salt solution. Solution was autoclaved before use.

5x M9 salts	500 mL
$Na_2HPO_4.7H_2O$ (Anhydrous)	32 g
KH2PO4 Potassium Phosphate Monobasic	7.5 g
NH4Cl Ammonium Chloride	2.5 g
NaCl Sodium Chloride	1.25 g
Distilled water	500 mL

Naïve TG1 colonies are sensitive to antibiotics. The two antibiotics used in the biopanning experiments were ampicillin (*Amp*) and kanamycin (*Kan*). Infection with the phage library provides the naïve TG1 bacteria with Amp resistance and the superinfection with the helper phage confers the TG1 with Kan resistance.

TG1 naïve colonies or the frozen glycerol stock of the phage library were grown at 37°C, 220 rpm on a shaking incubator in baffled flasks, 100 mL media in 250 mL flask for a small volume culture or 1 L media in 2.5 L flask for a larger culture. TG1 culture was used at OD_{600} 0.4–0.6 for phage infection, just before log phase, to ensure optimum library growth. For growing phage library after super-infection with helper phage at the optimal conditions (30°C and 180 rpm) for overnight growth.

2YT media and agar

The growth media used in all the steps of the biopanning experiments was 2YT media. The media (Table 2-3) was made and autoclaved before use. The agar of the same media was used to titrate the library phage and helper phage. The same recipe was followed with the added ingredient of 7.5 gm agar to 500 mL 2YT solution and autoclaved.

2YT media and agar were supplemented with 1% (w/v) glucose and antibiotics depending on the stage of the biopanning process (Table 2-5).

Table 2-3: 2YT media. Media was autoclaved before use

2YT Media	500 mL
Tryptone	8 g
Yeast	5 g
Sodium Chloride	2.5 g
Water	500 mL

2.1.3 M13KO7 helper phage

M13KO7 helper phage (Figure 2-1[2]) was used in the biopanning experiments to provide the phagemid vector the additional genes required to produce progeny. M13KO7 is a derivative of M13 bacteriophage which has an origin of replication from P15A and a kanamycin resistance cassette. Like other M13 bacteriophage, M13KO7 infects TG1 bacteria through the F-pilus. The process of growing M13KO7, extraction and calculating virus titre are outlined below. A stock of helper phage was produced and aliquoted into individual portions which were used for all phage generation experiments.

2.1.3.1 Infection of TG1

TG1 *E. coli* at optical density (OD_{600nm}) 0.4—0.6 was infected with the helper phage M13KO7 at a multiplicity of infection (MOI) of 10. Following the addition of helper phage, the bacterial culture was left static for 30 minutes at 37°C, for the helper phage to infect the TG1 culture.

To make a stock of helper phage, 500 mL TG1 culture was used, and the culture was grown in baffled 2 L conical flasks. Following static incubation, the culture was transferred to a shaking incubator for 30 minutes shaking at 220 rpm at 37°C for the bacteria to multiply following infection by the helper phage. A further 500 mL media was added to the culture. Kanamycin was added at a concentration of 50 µg/ml and the culture was divided into two 2 L flasks and left to grow overnight in a shaking incubator at 30°C and 180 rpm.

2.1.3.2 PEG precipitation

The overnight culture of the helper phage was centrifuged at 3000 x *g* for 20 minutes and the phage supernatant was collected into a fresh 2 L conical flask. The phage was precipitated by the addition 250 mL 20% (w/v) PEG-8000/2.5 M NaCl (Appendix 1, Table A1.2) for every litre and incubated on ice for 1–2 hours. Following the precipitation, the phage was centrifuged at 8000 *x g* for 20 minutes. The supernatant was decanted without disturbing the phage pellet. To remove residual PEG/NaCl, the pots were centrifuged a further two times at 8000 *x g* for one minute and any liquid remaining at the bottom of the centrifuge pot after each centrifugation was removed using a P1000 pipette.

The phage pellet was re-suspended in 30 mL of PBS and left on a roller mixer for the PBS to soak into the precipitated pellet for 5–10 minutes. Undissolved pellet was suspended into the PBS solution by gentle pipetting and the phage suspension was then left rotating on the roller at room temperature for one hour. The helper phage solution was then transferred into 1.5 mL micro centrifuge tubes and centrifuged at 13,000 *x g* for 5 minutes to pellet any remaining bacterial debris. The phage supernatant was transferred into a fresh 50 mL centrifuge tube and glycerol was added to make a 20% (v/v) glycerol stock. The phage glycerol stock was stored at $^{-}$80°C in 500 µL aliquots for long term storage. For short-term storage (up to 2 weeks), the phage solution was stored at 4°C.

2.1.3.3 Titration to quantify the concentration of M13KO7 helper phage

Two methods were used to calculate the titre of the M13KO7 helper phage. The two methods used were Kanamycin 2YT agar plate titration and the more sensitive 'Top Agar' titration method. Uninfected TG1 control was also tested along with the titration experiments to check for contamination.

Kanamycin 2YT agar plate titration

The helper phage was serially diluted in PBS and infected into naïve TG1 culture as explained above in 2.1.3.1 at 37°C for 45 minutes. After incubation, 100 µL of each serial dilution was plated on 2YT agar plates containing 50 µg/mL kanamycin (2YT/*Kan*) and left to grow overnight at 37°C. The negative control plate was set up by plating 100 µL of uninfected TG1 cells. Colonies that grew on the 2YT/Kan plates were counted and the phage titre using this method was $5.8x10^{10}$ cfu/mL.

Top Agar Titration (H-top titration)

Top agar (Table 2-4) which has a low percentage agar, was made and kept molten at 42°C in a water bath whilst 2YT agar plates without antibiotics or glucose were kept pre-warmed at 37°C. Helper phage was serially diluted in PBS as described previously. Hundred and eighty microlitres (180 µL) of TG1 bacteria at OD_{600nm} 0.2 was mixed with 20 µL helper phage dilution and added to 3 mL of molten top agar.

Table 2-4: Top agar for helper phage titration

Top Agar (100 mL)	Weight (g)
Agar	0.7
NaCl	0.8
Tryptone	1

This mixture was poured on the pre-warmed 2YT agar plates and left to set before incubating overnight at 37°C. The plaques were counted the next day and the titre of the helper phage stock was calculated. The titre calculated from the H-top agar titration was 4.2×10^{12} pfu/mL, which was closer to the titre expected. The stock of M13KO7 helper phage was divided into 500 µL aliquots and stored at -80°C till use.

Table 2-5: TG1 culture media for the the different stages of the biopanning process.

Ingredients	Amplifying the phage library stock to OD 0.4	Phage production step post helper phage infection	Overnight colony pick from minimal agar plate	Amp control for overnight colony pick from minimal agar plate
2YT media	500 mL	500 mL	10 mL	10 mL
Ampicillin (150 µg/mL) (Amp)	1.5mL	1.5mL	n/a	30 µL
Kanamycin (Kan) (50 µg/mL)	n/a	500 µL	n/a	n/a
25% Glucose used at 1% (w/v) final volume	20 mL	n/a	n/a	n/a

2.1.4 P8 library phage production

This step of the process describes amplifying the glycerol stock of the 16-mer random peptide P8 library (provided by Maria Tsoumpeli) from a 500 µL aliquot stored at -80°C. The library was thawed on ice and mixed with 500 mL pre-warmed 2YT media containing *Amp* and 1% (w/v) glucose solution in a 2 L conical flask and grown to an OD_{600nm} of 0.4–0.8 at 37°C and shaking at 220 rpm. Table 2-5 second column details the ratio of glucose and antibiotic added to the media for growing the library.

Helper phage M13KO7 was then added at an MOI of 10 and the culture was left to incubate at 37°C in static conditions for a minimum period of 30 minutes followed by for 30 minutes at 37°C in a shaking incubator (180 rpm). The culture was then centrifuged at 3000 *x g* for 10 minutes. The culture supernatant was discarded, and the bacterial pellet was re-suspended in 500 mL fresh media supplemented with *Amp* and *Kan* (Column 3, Table 2-5) in a 2 L conical flask The culture was left to grow in a shaking incubator set to 180rpm and 30°C to produce phage progeny overnight ([3]Figure 2-1).

The following day, the culture was centrifuged at 12,000 *x g* for 20 minutes. The phage supernatant was decanted into fresh 250 mL pots and PEG precipitated. For every litre of phage library, 250 mL PEG-NaCl solution was used. The same principle used in PEG precipitating the helper phage was followed.

After PEG precipitation and centrifugation, the library phage was resuspended in 2.5 mL PBS per 250 mL original culture and left to mix on a rotating mixer for an hour. The dissolved phage library was then pooled and to remove any remaining bacterial debris, the phage library was transferred in 1 mL volumes into 1.5 mL microcentrifuge tubes and centrifuged at 12,000 *x* g for 2 minutes. The concentrated phage library was re-pooled into a single 50 mL centrifuge tube, titrated and stored at 4°C until it was used for the first round of biopanning.

There are 3 key terms that will be used in describing types of biopanning that were carried out with the P8 library phage and these terms are defined in Table 2-6.

Table 2-6: Definition of biopanning terms used

Biopanning key terms	
Positive biopanning	The bead-antibody-phage complex is rescued and progresses to the next step of the experiment
Subtractive biopanning	The bead-antibody-phage complex is discarded from the experiment
Screening	This is usually the final biopanning round where the phage library is positively biopanned against the test and control set of sera and output libraries from both sets are sequenced or used downstream.

2.1.5 P8 phage library titration

The PEG precipitated phage library was titrated on 2YT/*Amp*/Gluc agar plates (Table 2-7) and a TG1 control plate was also plated to ensure the naïve TG1 cells were *Amp* sensitive. The process described for the 2YT/*Kan* agar titration was followed, with the phage serially diluted in PBS buffer as shown in Table A- and the phage dilution infected into naïve TG1 bacteria at OD ^{600nm} 0.4–0.8 as shown in Table A- and left to infect for 30 to 45 minutes at 37°C static conditions. After incubation, 100 µL of each dilution were plated and incubated at 37°C overnight.

The next day if the *Amp* control plate was free of colonies, the titre of the library was calculated from the colony counts from the plates plated with the phage dilutions.

Table 2-7: 2YT/*Amp*/Gluc agar plates for titrating phage library

2YT/Amp/Gluc agar plates for titrating phage library. This volume makes approximately 12 plates	
2YT agar	250 mL
25% Glucose (1% (w/v))	10 mL
Ampicillin (150 µg/mL)	750 µL

2.1.6 Biopanning process

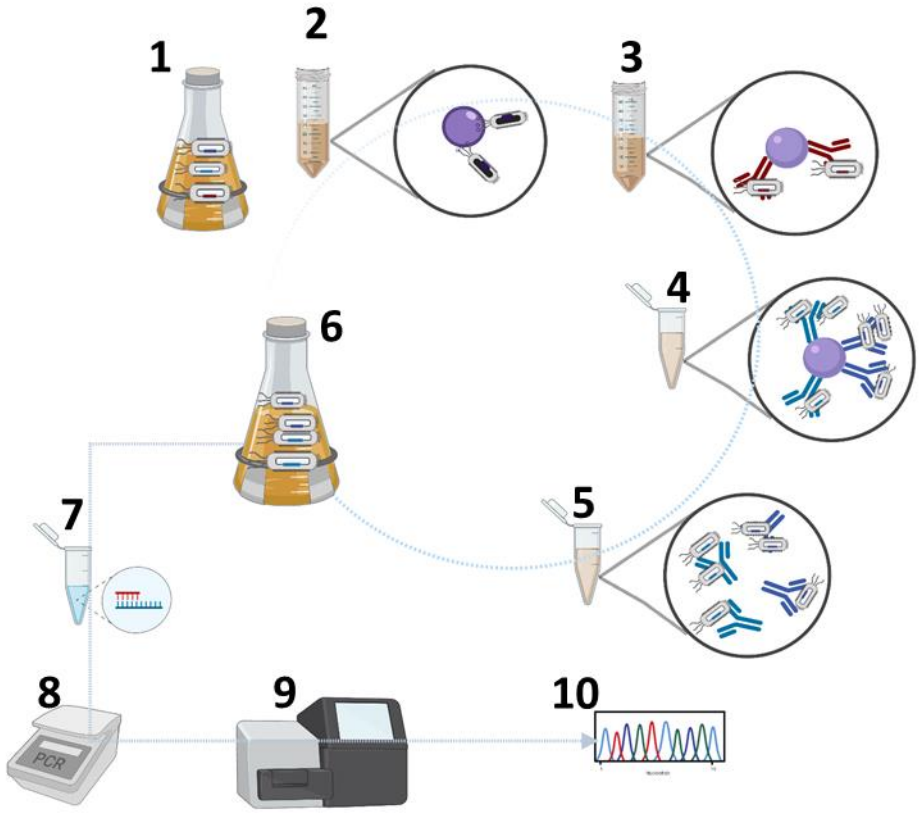


Figure 2-2: Schematic to show phage biopanning.
Steps are numbered **1**: The starter phage library is grown from bacterial glycerol stock. **2**: P8 phage library is processed to remove bacterial remnant, PEG precipitated and a subtractive biopanning is carried out to remove phage binders to plastic and Protein G agarose beads, **3**: subtractive biopanning where phage binders to the negative control (red IgG) are removed. **4**: Positive biopanning, with target antibody in serum (blue IgG) using individual sera. **5**: rescuing the individual output phage sub-library from the overnight biopanning. **6**: the rescued sub-library is infected into TG1 bacteria and glycerol stocks are made. **7-10:** processing the individual sub-library for NGS sequencing and NGS data analysis. Figure made in Biorender.

2.1.7 Biopanning with protein G agarose beads

The biopanning methods detailed in this section are the general methods used in all the experimental chapters. The specific target (described in the chapter) and the control serum antibody (described in the chapter) differs in each result chapter and the detail is provided in each result chapter. All the biopanning experiments carried out in this thesis used Protein G agarose beads (Pierce™ Protein G Agarose beads, Catalogue number 20399, Thermo Scientific™) as the solid support for the phage-antibody complex.

The agarose bead slurry volume used was dependent on individual experiments. a minimum volume of 50 µL bead slurry was used for biopanning 10 samples, bead slurry was pipetted into 1.5 mL microcentrifuge tubes, the supernatant removed once the beads had settled, and the beads washed twice in 1 mL PBS, centrifuging at 2500 x g to pellet the beads between washes.

The beads were resuspended in 300 µL PBS. All the biopanning experiments had a minimum of 10 samples and to 1.5 mL microcentrifuge tube, 30-50 µL of resuspended beads were added for biopanning. All protein G agarose beads were processed in this way before use.

2.1.8 Preparing phage library for biopanning

Once the phage library concentration was ascertained and the sensitivity of the naïve TG1 cells to ampicillin confirmed, the library was ready for biopanning experiments. To avoid non-specific phage binding to the block solution, plasticware and protein G agarose beads, being enriched in the biopanning process, the phage library was subtractively panned to these before biopanning with serum samples.

The precipitated phage library was blocked using milk block (Appendix 1, Table A-1.2) solution (final concentration 3% v/v). Protein G agarose beads were also blocked in 3% milk block for 1 hour and added to the

blocked phage library. This mixture was left to incubate on a rotating mixer at room temperature for 30 minutes after which it was centrifuged at 2,500 x *g* for 2 minutes to pellet the beads. The phage supernatant was carefully transferred to a fresh 50 mL tube.

2.1.9 Coupling serum antibody to Protein G agarose beads

Thirty microlitres of washed and resuspended protein G agarose beads and 2.5—5 µL of individual serum sample were combined in 1.5 mL tubes and left to couple for 30—45 minutes on a rotator mixer at room temperature. After incubation, 1 mL of 3% milk block was added to each tube and incubated for a further 30 minutes. The tubes were centrifuged at 2,500 x *g* for 1 minute to pellet the beads and the supernatant removed. The serum antibody coupled agarose beads were ready for use in biopanning.

2.1.10 Subtractive biopanning

The experimental aim in this thesis is to identify virus specific peptide mimotopes from biopanning the phage library against serum samples. The subtractive biopanning is a control measure which removes non-specifically binding phage. For the subtractive biopanning, control serum samples (2.5µL each) were bound to beads and blocked as described in section 2.1.9 and then transferred to a 50 mL centrifuge tube with the blocked and subtractively panned (against beads, Section 2.1.8) phage library. The subtractive biopanning reaction was carried out at room temperature on a roller mixer for up to 2 hours. The phage library was centrifuged to pellet the beads and the supernatant containing the subtracted library was transferred to a clean tube ready for positive biopanning.

2.1.11 Overnight biopanning

Equal volumes of the phage library (Sections 2.1.8 and 2.1.10) and the pelleted target antibody coupled beads were mixed into microcentrifuge

tubes and gently agitated to ensure the beads were in contact with the phage library. The reaction was left to incubate overnight at 4°C on a rocking platform.

2.1.12 Rescue of output phage

During the overnight incubation, the phage displaying mimotopes recognised by serum antibody (IgG) will have bound to the antibody-bead scaffold and could then be rescued.

The microcentrifuge tubes containing the biopanned libraries were initially centrifuged at 2,500 *x g* for 2 minutes to pellet the protein G agarose beads. The supernatant was discarded, and the beads washed three times with 1 mL PBS-0.1% Tween 20 buffer, and then a further three times with 1 mL PBS buffer.

The output phage library was eluted from the beads by increasing the pH through the addition of 250 µL 100 mM Triethylamine solution to each sample tube. The shift in pH elutes the phage from the Protein G agarose bead scaffold. The beads were gently agitated to mix and incubated for 10 minutes at room temperature. The tubes were centrifuged at 2,500 *x g* briefly to pellet the beads and the eluted phage supernatant was transferred to fresh tubes containing 250 µL 1M Tris-HCl buffer to neutralise the pH. The rescued phage samples were stored at 4°C.

2.1.13 Titrating and making bacterial glycerol stocks of output phage

Bacterial glycerol stocks were made to store the rescued phage long term at $^{-}$80°C. This was done by infecting 250 µL of the rescued output phage into 10 mL TG1 bacterial culture at OD_{600} 0.4–0.8. The culture was left to incubate at 37°C static for 30 to 45 minutes for phage infection.

A small volume of culture (100 µL) was taken from each of the TG1 infections, pooled to make the starting point for the titration dilution

series. The pooled library for the target and control output libraries were kept separate.

Four ten-fold serial dilution were produced using the pooled phage infected TG1 cells and 20 µL of each diluted sample and 20 µL of the undiluted infected stock were added to 180 µL of media. The dilutions were mixed by pipetting and 100 µL from each was plated on 2YT/*Amp*/Gluc agar plates and incubated overnight at 37°C. A control plate of uninfected TG1 bacteria were also set. The titres were calculated from the colony counts the next day.

The rest of the phage infected TG1 cultures were centrifuged at 3,000 *x g* for 10 minutes; media supernatant was discarded, and the bacterial pellets were resuspended in 10 mL 2YT/Amp/Gluc media and left to grow overnight in a shaking incubator at 220 rpm at 37°C.

The next day, the bacterial cultures were centrifuged at 5000 *x g* for 10 minutes, the supernatant was removed, and pellets were resuspended in 2 mL freezing media (Appendix 1, Table A-1.2). The output sub-libraries were stored in 500 µL aliquots of individual samples and as pooled output sub-libraries from each round of biopanning. The pooled output library is the enriched sub-library of phage binders to the target from each round of biopanning and is used as the starting library for the subsequent round of biopanning.

Dependent on the experimental plan, the pooled libraries were used as the input library for the subsequent round of biopanning.

2.2 Next Generation Sequencing

The higher throughput method chosen to identify phage expressing peptides of interest enriched through the biopanning process was to sequence the output libraries on a next generation sequencing platform. In this thesis the output libraries were sent for NGS at the University of

Pennsylvania, to be sequenced on the Ion S5 sequencer using Ion 540™ Chip (Thermo Scientific).

2.2.1 DNA extraction

Phage DNA from individual sub-libraries were extracted using GenElute Plasmid Mini Prep Kit (Sigma) (for section 3.2.3) and Qiaprep Spin miniprep kit (Qiagen 27106) (for sections 3.2.7,4.2.6 and 5.2.6). DNA was eluted into 30 µL of NE buffer and quantified using a Nanodrop Spectrophotometer (Thermo Scientific). DNA was further diluted to a concentration of 10 ng/µL for use as PCR template.

Two stages of PCR amplification (Figure 2-3) were carried out with the diluted DNA samples to amplify the gene segment encoding the peptides displayed on the phage and to add the linker and DNA barcode required for Ion Torrent sequencing.

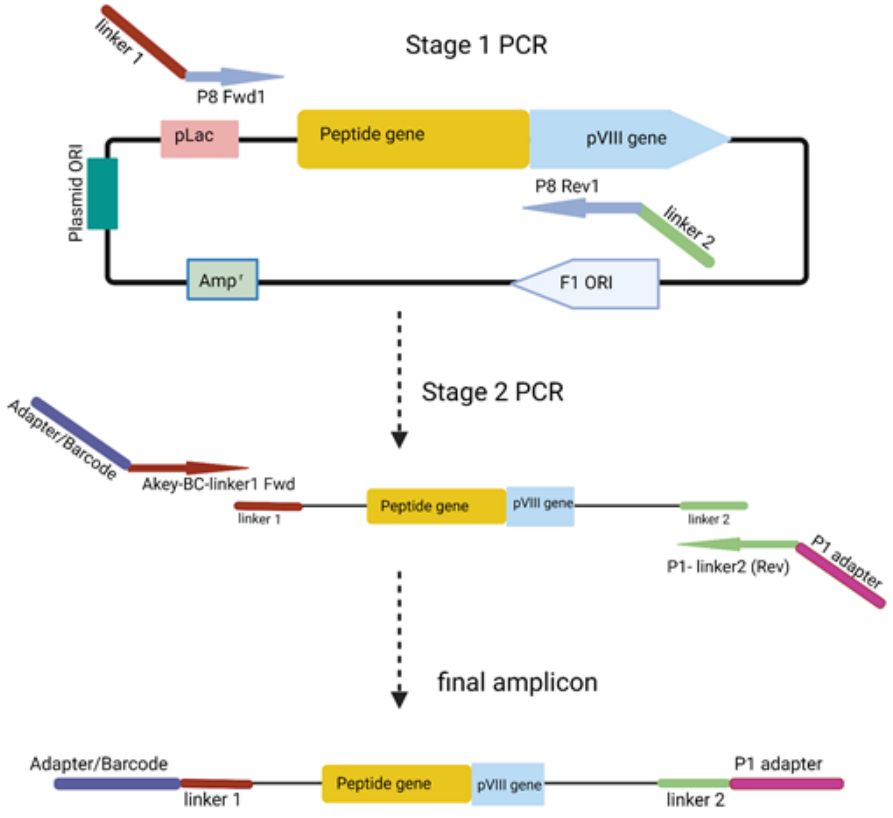


Figure 2-3: The section that is amplified during the two stages of PCR amplification and the final PCR DNA amplicon.

2.2.2 Stage 1 PCR (R1 PCR)

The PCR reagents and conditions for the first round of PCR are shown in Table 2-8. PCR reagents used were Q5® Hot Start High-Fidelity DNA Polymerase and Deoxynucleotide (dNTP) Solution Mix (New England Biolabs) and DEPC-Treated Water (Ambion, Invitrogen™). The first round PCR primers were library specific P8 forward and reverse primers seen in Table 2-9.

A master mix was made and 49 µL was added to each tube with 1 µL template DNA. The first round PCR incorporated linkers for the primers used in the second round of PCR. Negative control with no template was also run.

Table 2-8: Stage 1 PCR reagents and PCR cycle conditions

Stage 1 PCR reaction mix	Volume (µL)
DEPC water	25.5
Q5 buffer	10
GC enhancer	10
dNTP	1
Forward primer (10 nM/ml)	1
Reverse primer (10 nM/ml)	1
Q5 enzyme	0.5
Template (10 ng/uL)	1

Thermocycler parameters	
95°C for 3 min	
95°C for 30s	x 30
60°C for 30s	
68°C for 30s	
72°C for 5 mins	

Table 2-9: Stage 1 PCR primers. Linker region is highlighted.

Stage 1 PCR primers	
P8 FWD	GTAATCCTTGTGGTATCGGATGCTGTCTTTCGCTGC
P8 REV	CTAGAACATTTCACTTACGGTTTTCCCAGTCACG

2.2.3 Gel electrophoresis and PCR amplicon clean-up

Gel electrophoresis of the amplified PCR reaction was carried out using 10 µL of PCR product mixed with 8 µL DEPC water and 3 µL loading dye. The PCR products were loaded on a 2% (w/v) agarose/Tris acetate EDTA (TAE) gel stained with Nancy-520 and run for up to an hour at 100 V and visualised on the Image Quant 300 (Amersham Biosciences) or the ChemiDoc™ MP (BioRad) imaging systems.

Ready-to-use MassRuler Low Range DNA Ladder (Thermo Scientific™) kit which includes a DNA ladder and loading dye was used as a size guide on the gels. Upon confirmation of the presence of the correct sized band (typically 200 — 250 bp) in the samples and no bands in the negative control, all the PCR amplicons including the negative control were purified using the Machery-Nagel Nucleospin Gel and PCR purification kit

(catalogue number 740609.250) and eluted in 20 µL of NE elution buffer (pre-warmed to 65°C) which was provided in the kit.

2.2.4 Stage 2 PCR (R2 PCR): barcoding

The second stage of PCR used primer P1 plus the individual barcode primers of a possible 96 (Appendix 1, Table 1-3). Barcode 42 was not used for barcoding any samples as it was unavailable. Apart from the different template, the reaction was set up as described for stage 1 PCR (Table 2-8). The PCR cycle parameters were different and can be seen below (Table 2-10).

Primer P1 links the sample DNA to the sequencing bead in the Ion Torrent sequencing process. The NGS barcodes are used to distinguish the individual samples apart as the samples get pooled after the second round of PCR.

Table 2-10: Stage 2 PCR thermocycler parameter

Thermocycler parameters	
95°C for 3 min	
95°C for 30s	
63°C for 30s	x 12
68°C for 30s	
72°C for 5 mins	

The PCR amplicon size was checked by gel electrophoresis as described in section (2.2.3) End product size was ~350 bp. The PCR amplicons were purified after quantification and pooling.

2.2.5 Quantifying DNA

Qubit™ dsDNA BR Assay Kit (Invitrogen™, Thermo Fisher Scientific) was used to quantify the amplicons from the R2 PCR. A reagent master mix was prepared with 1 µL component A (Qubit® dsDNA BR Reagent) and 199 µL of component B (Qubit® dsDNA BR Buffer). The standard curve was plotted on the analyser using the two standards 0 ng/µL and 100 ng/µL provided in the kit. Ten microlitres of each standard was mixed with 190 µL master mix, incubated for five minutes, and read on the Qubit analyser.

Once the standard curve had been set, 2 µL of each barcoded sample was mixed with 198 µL of the reagent master mix and read on the analyser. The DNA concentration was expressed in ng/µL units. The DNA from all the sub-libraries were pooled together by combining the same volume of DNA (100 ng) from each sample.

2.2.6 Purifying pooled DNA

The pooled sub-library amplicons were gel purified on high resolution agarose to ensure purity and presence of a single band. The samples were run on 3% MetaPhor™ Agarose (Lonza) gel made in TAE buffer and stained with Nancy 520. The gel was cast with wide tooth combs the day before running it and stored at 4°C.

The pooled DNA was diluted 1:6 with loading dye and up to 45 µL loaded per well. The gel was run for 3 hours at 50 V to ensure separation of the bands. The gel was visualised on an imager and the 350 bp band excised from the gel and purified using the Machery-Nagel clean up kit as described previously.

The DNA pool was purified twice using the magnetic bead based Agencourt AMPure XP kit (Beckman Coulter). The pure amplicon was eluted into 20 µL elution buffer, and the DNA was quantified on the Qubit

analyser. A confirmatory gel was run with 50 ng DNA to confirm for presence and purity of the pooled DNA. The instructions provided by the sequencing company requested 50 µl DNA at a concentration of 4 ng/µl volume in a 0.5 mL DNA LoBind (Eppendorf) tube which was sent for Ion Torrent sequencing at the commercial sequencing service provided by the University of Pennsylvania.

2.3 Colony picking of phage clones, Sanger sequencing and phage ELISA

In traditional phage display experiments, colony picking followed by Sanger sequencing is used to identify the phage peptide of interest. The phage sub-library of interest was plated on 2YT/Amp/Gluc plates in serial dilutions as described previously.

Individual colonies were picked and inoculated into 600 µL 2YT/*Amp*/Gluc media (Table 2-3, Table 2-5) in deep well plates (2.2 mL well, BRAND™ Deep Well Plates, 10089910, Fisher Scientific) in duplicate. The cultures were grown at 37°C, 220 rpm in a shaking incubator. Control wells with naïve TG1 colonies were used to measure OD_{600}nm values. The cultures were grown to OD_{600} 0.8 — 1 and centrifuged at 3,500 x *g* to pellet the cultures. The supernatants were removed, and the pellets resuspended in 200 µL freezing media (Table A-) and stored at -80°C.

One plate was used to make glycerol stocks for long-term storage. The second plate was used as the 'working' plate and 250 µL from each well (phage clone) was used for DNA extraction using the Qiaprep Spin miniprep kit (Qiagen 27106) as per manufacturer's instructions. The remaining volume of culture in this plate was centrifuged at 3,500 x *g* to remove the supernatant. The pellets were resuspended in freezing media, and of 50 µL from was used as starter for growing individual phage clones.

Phage ELISA was carried out to test the phenotypic properties of the isolated clones. A series of experiments were conducted to identify the best protocol, and this is described in detail in Phage ELISA development (Chapter 4). The reagents and conditions used in growing the P8 library phage was used for growing the phage clones (Section 2.1.4)

Sanger sequencing

DNA extracted from the bacterial colonies was quantified using the NanoDrop (Thermo Fisher Scientific) platform and Sanger sequencing was used to identify the 16-mer peptide expressed by the particular phage clone. A volume of 5 µL DNA (100 ng) per sample was submitted to Source Bioscience for Sanger sequencing along with the unidirectional primer 'pc89_qPCR_universal', 19-mer sequence 'ATTACCCCAGCTGGCGAAA-3' (Eurofins).

2.4 NGS Data analysis pipelines

Data analysis of the deep sequenced libraries was carried out on the operating system Linux which has the benefit of having open-source software which can be used for automation of processing large datasets. The Ubuntu Linux virtual machine (Ubuntu 64-bit Oracle VM VirtualBox) installed in Microsoft Windows was used for running scripts for analysing NGS data.

2.4.1 Processing Ion Torrent sequencing file

The sequenced data file was supplied as zipped FASTQ files with the extension *'gz'*. The file was downloaded, and the software 'unpigz' was used to unzip the FASTQ file. Perl scripts written by Professor Richard Emes (Advanced Data Analysis Centre, University of Nottingham) and edited to fit the individual phage library by Dr Jon Owens (ADAS, UK) were applied to the NGS data sets in this thesis. The scripts and the files associated with it are accessible from the Git hub repository using the link https://github.com/UoN-ADAS-KGough/NGS_Scripts01.

The perl script was used to convert the FASTQ file to FASTA format, de-multiplexing and binning data into files according to the barcode, translating the DNA sequence into peptide sequences in all three reading frames. The 16-mer peptide is presented in the following format 'AEGEF—(X)16—DPAKA', where AEGEF and DPAKA are fixed N and C terminal amino acid sequences from the vector that flanks the randomised peptide '(X)16', which was enriched in the biopanning process. The script was able to identify the known flanking regions and extract just the peptide sequence of interest. The peptides were arranged according to frequency of occurrence.

A program named GNU Parallel was used to speed up the data processing of the files by maximising the processing speed of the computer. The NGS files were named with a prefix of the experiment they originated from and

the barcode to track the peptide libraries back to the original phage sub-library it originated from. Additional details on the data processing are provided in Appendix 1 (A-1.4).

2.4.2 Data analysis

Two different approaches were used to identify peptides enriched to the target compared to comparator samples, seen in Figure 2-4. The Z-score analysis has been used for analysing previous NGPD projects from the group and the data will also be analysed using immunosignature analysis, a novel machine learning informed method of identifying peptides of interest.

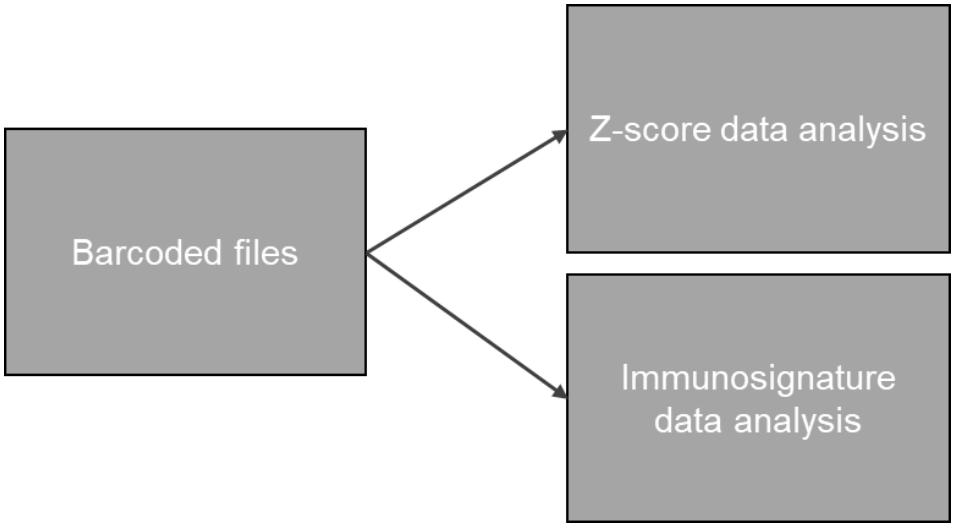


Figure 2-4: Two data analysis pipelines used

2.4.3 Analysing NGS data using Z-score test

Two proportion Z-score test has been used to identify peptides of interest from NGPD libraries and the methodology used in this thesis was published by the Gough group (Naqid, et al., 2016b). Every peptide in every target sample was compared to all the peptides combined from the control samples. The Z-score test was used to compare the frequency of occurrence of each peptide in the target sample, the ratio (of peptide enriched in the target: comparator) and the total number of peptides in each barcode. This generates a ranked list of peptides by relative

statistical importance. An arbitrary cut-off is then applied to select the peptides enriched above the threshold. The script used to generate the Z-scores are provided in the Github repository. Table 2-11 shows the first four lines from a file as an example.

Table 2-11: Example of a Z-score output file

Sequence	Test sample	Control pool	Z-score
LKTDGVNFNTPAVPTN	7497	21	88.24
NRIPTLFTNAGAVHPV	2244	356	47.28
NKYVDKSAPNNWNATP	1227	5	35.13
KGDRFTSRHTETLQVA	1190	626	33.73

Peptides with a higher the Z-score value will have a higher enrichment value in the target sample comparatively than in all the pooled peptides in the control. Generally, the arbitrary Z-score cut off value for picking peptides of interest is 2, but in a complex assay this value can be made more stringent. It should also be noted that if a very high stringency value is used, there might not be many peptides available for downstream testing.

2.4.4 Immunosignature data analysis

Immunosignature data analysis is a novel method of analysing the NGS output data. It uses the machine learning method of 'Training and Testing' data, where a subset of the dataset from both target and comparator files are used to generate data (immunosignature) specific to the target. This immunosignature is subsequently tested against the remaining target and comparator files.

This analysis uses frequency of occurrence (peptide frequency) of the peptides in each barcode. For each barcode, the list of peptides and their corresponding peptide frequency were arranged by descending order of frequency. This data was transferred to Microsoft Excel spreadsheets for the data training and testing. Figure 2-5 explains the 'Training and

Testing' concept used to identify peptide immunosignature specific to the target.

2.4.4.1 Data Training

From the NGS data set, a subset of the target and comparator data files were trained to refine a list of peptides seen only in the biopanning target files. This was done by filtering for peptides unique to the target and removing any peptides that were common to both sets. The trained peptide list is the immunosignature. The immunosignature can be trained and retrained to ensure the list of peptides is as specific to the target as possible.

2.4.4.2 Data Testing

The normalised peptide frequencies for each peptide in all barcodes used for Data testing were calculated in Excel using equation shown below.

$$normalised\ peptide\ frequency = \frac{peptide\ frequency\ of\ unique\ 16-mer\ peptide}{count\ of\ 16-mer\ peptides\ in\ barcode} \times 100$$

Once the peptide immunosignature for the target was finalised, it was then 'Tested' using the remaining data files (target and comparator). Microsoft Excel formula 'VLOOKUP' was used to identify and pull out the peptide sequences and the corresponding normalised peptide frequencies from the Test files if they matched the immunosignature peptide list. The normalised peptide frequencies for each file were added together. The sum of normalised peptide frequencies the test and comparator files of the immunosignature test files was used to generate data for the sensitivity/specificity of the immunosignature to the target. A subset of the samples in Data Testing was used to measure the sensitivity and specificity of the immunosignature, using ROC analysis on GraphPad Prism software (versions 8 and 9).

Statistical analysis

GraphPad Prism software was used for statistical analysis of the data. The immunosignatures were plotted on heatmaps and statistical analysis was carried out using binomial regression analysis to determine sensitivity and specificity to the target samples.

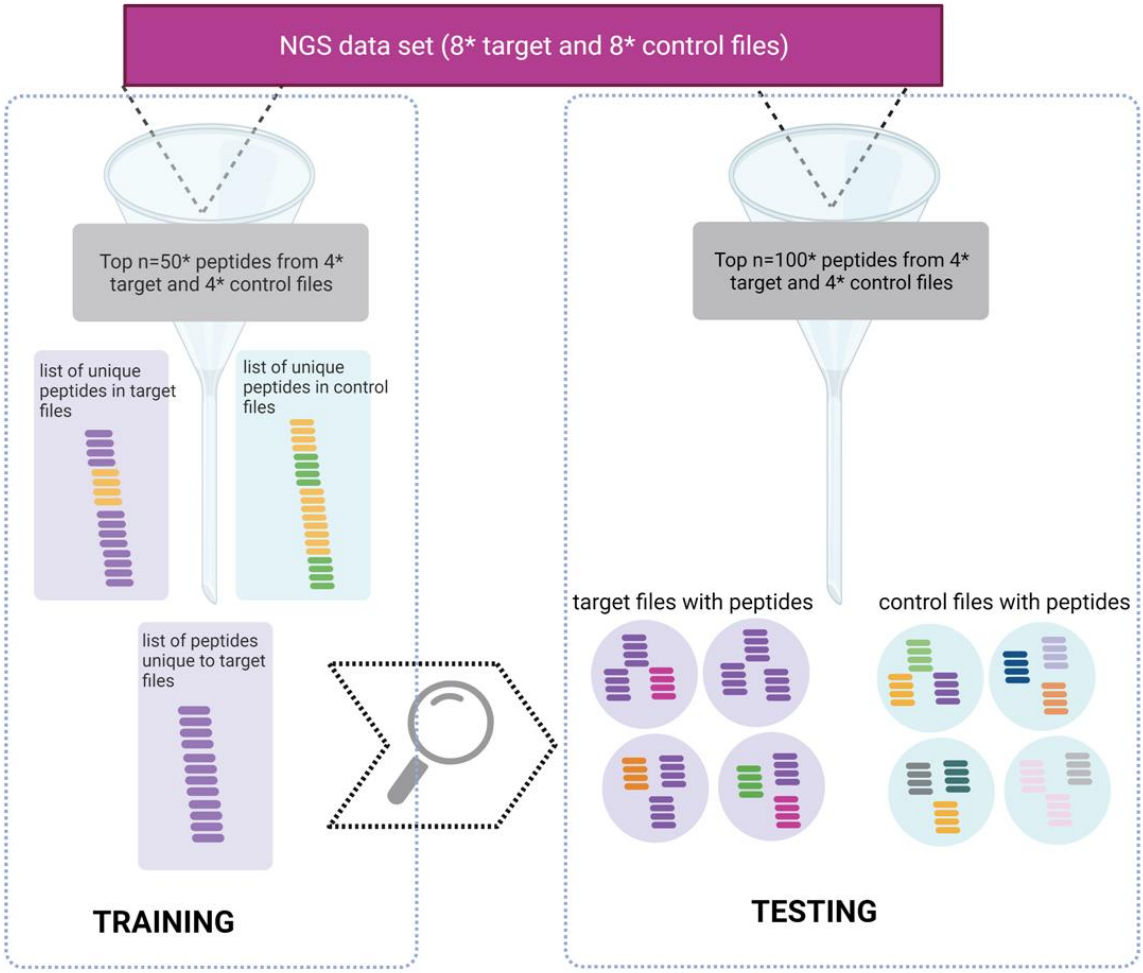


Figure 2-5: Schematic to explain training and test method to identify patterns in immunosignature data analysis.
Peptides unique to the target data are identified through a process of training x number of peptides from a proportion of target and control NGS data files. The inter and intra group duplicates are removed and peptides unique to the target are identified. These are then 'tested' with x number of peptides from each target and control file not used in the training. The number of peptides seen enriched is represented as a percentage of the total peptide reads in the testing files. *denotes example number of data files

2.5 ELISA

2.5.1 Commercial ELISA

Commercial ELISA kits were used for detecting serum IgG to Dengue (Abcam ab108728) and Zika (Abcam ab221844) in human sera as outlined in Chapter 2 and TBEV IgG as discussed in Chapter 5. They were carried out and results analysed according to manufacturer's instructions. ELISA plates were read on LT-400 Microplate Reader.

2.5.2 LIV IgG ELISA

Phage clones identified through the biopanning process were tested in sandwich ELISA format, where individual (monoclonal) phage clones were used as coating reagent as outlined in Section 4.2.4.

Diluted serum samples with LIV IgG bind to those phage clones which are mimotopes of the antigen. The antibody-phage complex was detected using a secondary antibody and substrate. The reaction was stopped, and absorbance readings were acquired once ELISA plates were read on LT-400 Microplate Reader using Manta (Dazdaq Solutions Ltd) software. Method is detailed in 4.2.4

Chapter 3 **Application of next generation phage display to identify species-specific peptide mimotopes from serum samples antibody positive for Zika and Dengue viruses**

3.1 Introduction

ZIKV and DENV are two members of the mosquito-borne flavivirus group. DENV has been a major cause of infectious disease in humans and is currently endemic in over 100 countries including several in Europe. This is due to factors such as global trade, change in weather patterns and vector competence (Bhatt *et al.*, 2013).

Unlike DENV, ZIKV cases in humans had been sporadic since it was discovered in 1964 (Kucharski and Riley, 2016)till a large outbreak in the Yap Islands in 2007 (Lanciotti *et al.*, 2008). Since then, there have been multiple outbreaks starting in 2015-2016 and the virus has spread to distant geographic regions such as the Caribbean Islands and the Americas. ZIKV and DENV are transmitted by the mosquito vectors *A.aegypti* and *A.albopictus* (Kraemer *et al.*, 2015). There have been concurrent outbreaks of DENV and ZIKV in the same geographical areas. Enhancement of disease in ZIKV infections due to pre-existing DENV antibodies and the reverse has been well documented (Priyamvada *et al.*, 2017; Bonheur *et al.*, 2021). There is 43% similarity in the amino acid profiles of ZIKV and DENV across the viral polyprotein and envelope protein (Lazear and Diamond, 2016) . Antibodies to FLE and the three E protein domains and the E-dimer region are conserved in both viruses and are responsible for the cross-reactive antibody response. MAC-ELISA can be used for accurate diagnosis but is only possible to use when IgM levels are high and the NS1 based ELISAs are very expensive. The Euroimmun IgG and IgM NS1 ELISA used by the Rare and Imported Pathogens Laboratory, UKHSA was costed at around £1000 for a 96 well assay plate.

This is the first experimental chapter of this thesis. In this chapter, a panel of human serum originating from the Caribbean was used to biopan a peptide phage library to generate a sub-library of peptides which are mimotopes to ZIKV IgG compared to DENV IgG.

The antibodies produced during Zika virus (ZIKV), and dengue virus (DENV) infections are highly cross-reactive, which makes serological diagnosis challenging. In the first study presented in this thesis, phage panning was used to enrich for virus-specific mimotopes using serum samples from clinical sera from a region where both viruses are currently present.

Objectives

1. Use next generation phage display (NGPD) to identify peptide mimotopes specific to ZIKV IgG over four rounds of biopanning.
2. Develop an *in-silico* assay to identify ZIKV infection in serum with pre-existing antibodies to DENV.

3.2 Materials and Methods

3.2.1 Serum samples

The serum samples used in this experiment were provided by Dr Mark Page (NIBSC). They were human serum samples collected from the Caribbean during the 2016 ZIKV outbreak, had been tested by the Caribbean Public Health Agency (CARPHA) for ZIKV via RT-PCR (real time PCR) and were surplus to requirement. Ethics approval to use these samples was sought and granted by the School of Veterinary Medicine and Science (SVMS) ethics committee. The samples when provided were labelled as either ZIKV+ (n=61) or DENV+ (n=50). A further 10 samples were provided at a later date. A small proportion of the samples had been tested at NIBSC using their in-house ZIKV IgG ELISA. A database

provided mentioned an unspecified IgM ELISA and RT-PCR were used for serological and molecular detection of DENV.

3.2.1.1 Diagnostic ELISA

All the serum samples were tested for the presence of ZIKV and DENV IgG using commercially available ELISA test kits: human anti-Dengue virus IgG ELISA Kit (Abcam ab108728) and anti-Zika virus IgG ELISA kit (Abcam ab221844). Manufacturer's instructions were followed during the execution of the tests and for data analysis.

3.2.2 Biopanning strategy and method

Subtractive and positive biopanning (Table 2-6) were used in this experiment. The first step in all the rounds was subtractive biopanning to remove non-specific binders to plastic, milk etc. The second step was subtractive biopanning using pooled DENV sera, and the third step was to biopan the library to enrich for binders to ZIKV IgG. The phage binders eluted from the biopanning round were infected into TG1 at OD 0.5, grown overnight and stored as bacterial glycerol at -80°C. The titre of the pooled sub-library was calculated on 2YT/*Amp*/Gluc agar plates as described in Chapter 2. Biopanning output sub-libraries from each serum was stored individually and as a pooled output from that round.

The specific biopanning strategy used in this experiment is shown in Figure 3-1. In rounds 1-3, the phage library was subtractively biopanned with 20 pooled DENV sera and then positively biopanned with 10 ZIKV sera panel. In the 4th round, the 3rd round output library was subtractively biopanned with the pooled DENV sera and then 'screened' against all sera from both ZIKV+ (20) and DENV+ (20) sera panels and all 40 output phage libraries were sequenced on the Ion Torrent sequencing platform, data was analysed and to test for reproducibility of the results, the 4th round biopanning was repeated at a later date (Figure 3-3). In the data analysis of the Ion Torrent NGS data, the files related to the DENV+ sub-

libraries form the control dataset and those from the ZIKV +sub-libraries form the test data.

Sequencing all the output libraries allows for a clear picture of the peptide mimotopes that were enriched preferentially to ZIKV+ IgG. It also allows for non-specific and common peptides to both sets of data to be identified and discarded. The Material and Methods chapter sections 2.1.6 — 2.1.13 gives more detail on the general biopanning process (Figure 3-2).

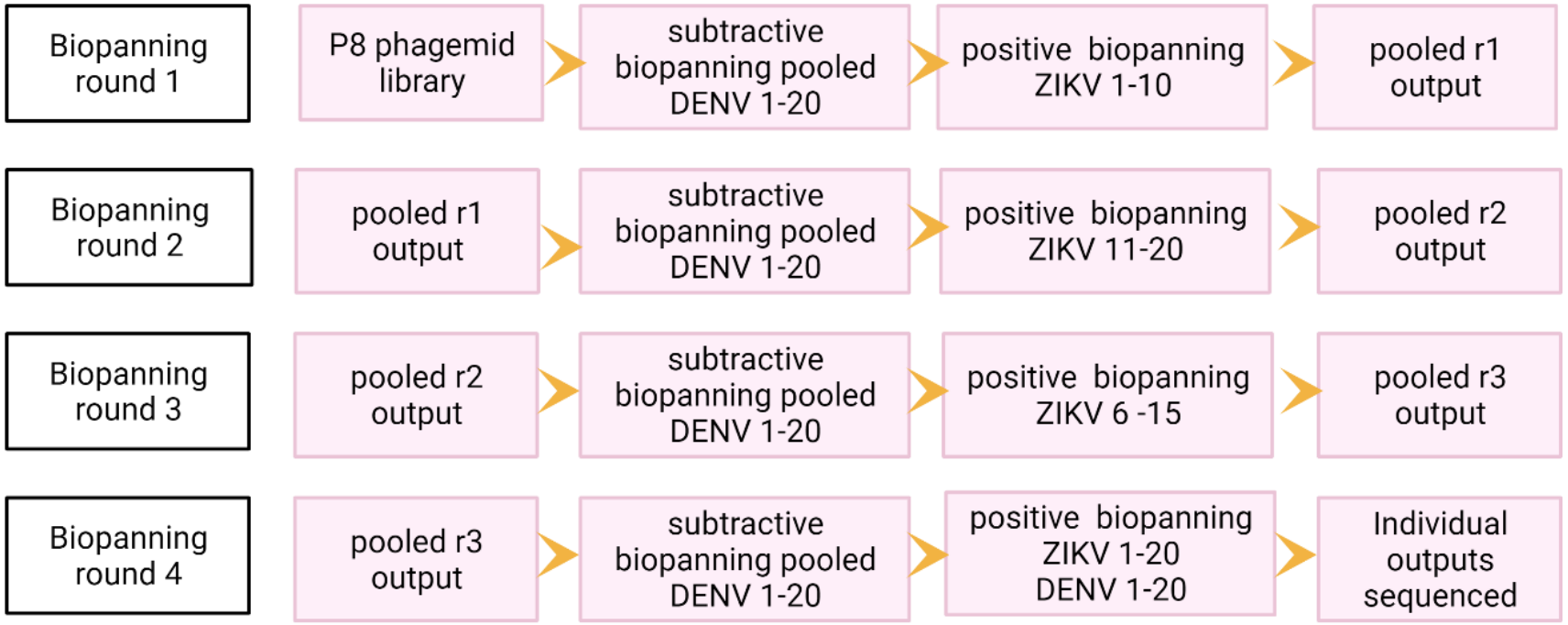


Figure 3-1: Biopanning strategy in this study.

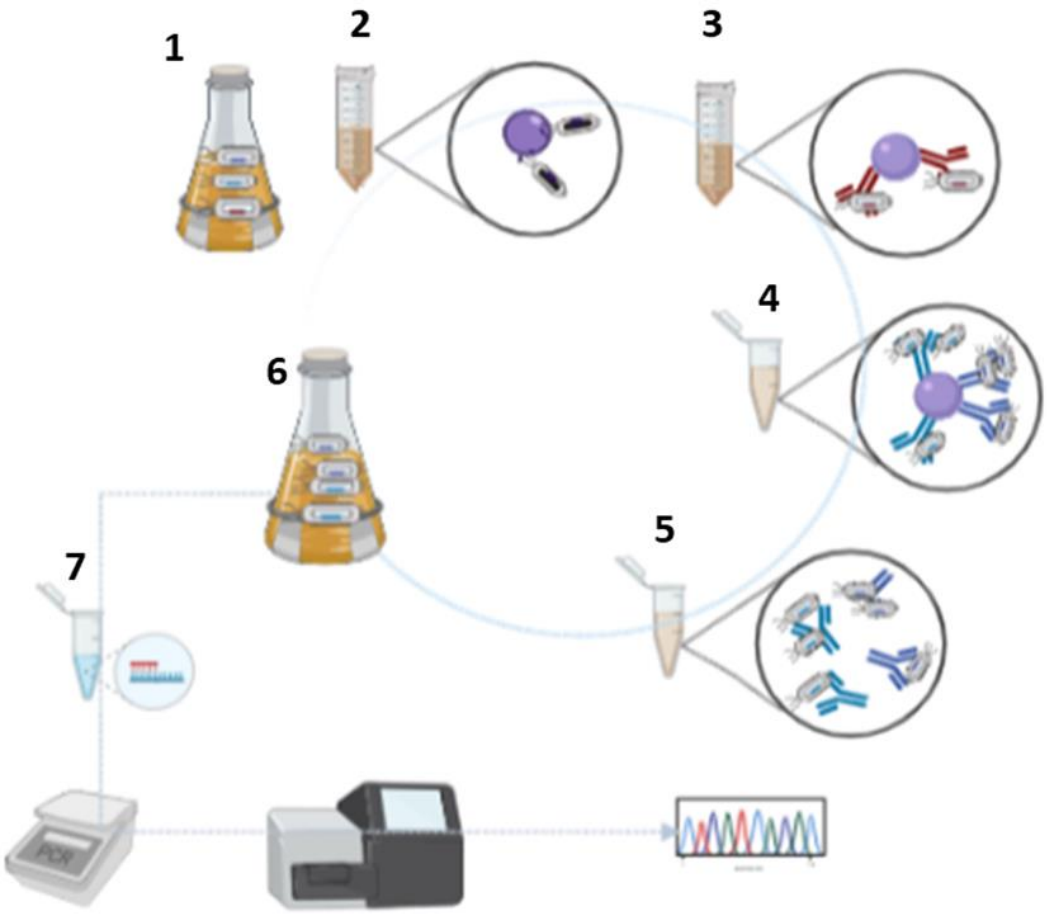


Figure 3-2: Biopanning schematic.
1 shows the input phage library, steps 2—3 shows the subtractive biopanning process with the red antibody representing the DENV+ IgG. Steps 4—6 shows positive biopanning with ZIKV+ IgG represented with the blue antibody. The phage sub-library from each sample is rescued and infected into TG1 bacteria to make glycerol stock of the sub-library. Step 7 shows that the DNA is extracted from the glycerol sub-libraries, PCR amplified and barcoded for Ion Torrent sequencing. The NGS data is then analysed to understand whether the biopanning experiment was able to identify specific peptide motifs to the target, which in this experiment is ZIKV IgG. Image created in Biorender.

In this experiment, 12 mL of PEG precipitated P8 library was incubated with 50 µL Protein G agarose beads bead slurry which had been washed and resuspended in PBS buffer. The P8 library was blocked in 30 minutes and the phage library was moved to a clean 50 mL falcon tube without disturbing the bead pellet.

Subtractive biopanning P8 library with DENV+ sera

Step 3 in the illustration shows the subtractive panning to pooled DENV sera. A 50 µL pool of sera was made by aliquoting 2.5 µL sera from all 20 samples of the DENV panel. The beads and serum were incubated together for 30 minutes and 1 mL of 3% M-PBS (Table A-) was added, and the reaction was incubated for a further 30 minutes on a rotating shaker.

The block and serum were removed after the incubation by centrifuging the reaction mix at 2,500 x *g* and the beads were added to the P8 phage library and incubated for 1 hour at room temperature on a rotating mixer. After the incubation, the beads were pelleted by centrifugation at 2,500 x *g* and the phage supernatant was removed to a fresh tube ready for the positive panning.

Positive panning of phage library

In rounds 1–3, ten sera from the ZIKV panel were used for positive biopanning, as shown in Table 3-4. Thirty microlitres of bead slurry was used for each reaction tube and 2.5 µL of serum sample was added and incubated for up to an hour on a rotating platform after which 1 mL 3% M-PBS was added and incubated for a further 30 minutes. The tubes were centrifuged at 2,500 x *g* to pellet the beads and remove the supernatant and 1 mL of the subtractively panned P8 library was added to the beads and left to biopan overnight on a rocking platform at 4°C

Processing output sub-libraries

Phage rescue steps outlined in 2.1.12 — 2.1.13 were followed. The output phage from the overnight biopanning was eluted from the Protein G agarose beads, rescued, and grown in *Amp* sensitive TG1 cultures to be stored. The sub-libraries from each biopanning round were stored as glycerol stocks individually and as pooled ZIKV output library at $^{-}$80°C. The pooled output library was the input library for the following round of biopanning.

3.2.3 Next generation sequencing

All 40 sub-libraries from the fourth round of biopanning were sequenced on the Ion Torrent sequencing platform. The method is detailed in section 2.2. Briefly, phage DNA was extracted using the GenElute Plasmid Mini Prep Kit (Sigma) and quantified on the Nanodrop and 1 µL DNA at concentration 10 ng/µL was used for the first round of PCR. This round used library specific primers that added linkers for the second round of PCR.

First round PCR products were cleaned and 1 µL of the PCR product was used as template for the second round of PCR, which added the P1 linker for the sequencing run and the individual DNA barcodes that were the identifiers for the samples.

The barcoded amplicons were cleaned and quantified on the Qubit analyser and 100 ng/µL DNA from each sample was pooled to make the sequencing pool. The pooled DNA was gel purified and the single band of DNA was excised, and the DNA was extracted and further purified using Ampure beads and 50 µl of pooled DNA in 4 ng/µl volume was shipped to the commercial Ion Torrent sequencing facility at the University of Pennsylvania.

3.2.4 NGS data analysis

The sequenced data files were downloaded and analysed using existing Linux and Perl scripts written by the ADAC group and Jon Owens (ADAS). The files were converted from FASTQ to FASTA format. They were then de-multiplexed into individual sample files using standard Ion Torrent barcodes and translated from DNA to protein sequence. ZIKV serum samples 1–20 were barcoded with barcodes 1–20 and DENV 1–20 samples were barcoded with barcodes 21–40. These samples will be identified by their barcodes for the rest of this chapter. The 16-mer peptides in each file were identified using the flanking regions 'AEGEF–(X) 16–DPAKA'.

Two types of data analysis were performed on the NGS data: two-proportion Z-score analysis, which was the established data analysis pipeline and a novel approach based on training and testing the NGS data to identify a specific peptide profile referred to as an 'immunosignature' for ZIKV.

3.2.5 Two proportion Z-score analysis

The two proportion Z-score test is a statistical method of comparing two populations. Here the peptides enriched from individual ZIKV samples were compared to the peptides enriched from DENV samples. The Z-score of a peptide reflected the ratio of ZIKV samples vs DENV samples it was seen in, the absolute frequency and its relative statistical importance. The threshold score for Z-score was set to 5 to limit non-specific phage binders.

Workflow:

i) Peptides from individual ZIKV+ samples were compared against pooled peptides from DENV+ samples using Linux scripts. This comparison generated the ranked Z-score.

iii) Those peptides seen enriched above an arbitrary Z-score were further ranked by the number of ZIKV+ samples the peptide was seen in compared to the pool of DENV+ samples.

3.2.6 Peptide immunosignature

The NGS data sets were analysed using the machine learning 'train and test approach' to identify a peptide immunosignature that would be specific for ZIKV. In this approach, a subset of the NGS files from both target (ZIKV) and control (DENV) were '***trained'*** to generate an immunosignature specific for the target (ZIKV). These are peptides which are only enriched in the files of the target and not in the DENV (control).

Once the immunosignature set of peptides were identified, they were interrogated ***'(tested)'*** against the remaining target (ZIKV) and comparator (DENV) data to ascertain whether the biopanning experiment was able to enrich a set of peptides unique to the target (ZIKV) compared to the control (DENV).

The trained ZIKV immunosignature peptide list was compared with the top 200 unique peptides from the Test set files. This part of the analysis was carried out in Microsoft Excel. The list of top 200 peptides and frequency reads for each sample in the immunosignature test cohort were listed in rows. A formula was set up using 'VLOOKUP' function which enabled the filtering of peptides common to the immunosignature peptide list and the frequency to be printed into adjacent rows.

Training set: the top 50 peptides by frequency from 10 ZIKV files (ZIKV 1–10) and 10 DENV (DENV 1–10) files were pooled into two groups. Inter–group and intra–group peptide duplicates were removed, and the remaining unique peptides in the ZIKV peptide set was the ZIKV immunosignature.

Test set: ZIKV 11–20 and DENV 11–19 formed the test set. The top 200 unique peptides ranked by their absolute frequency from each sample

made the 19 test files. The immunosignature analysis was used for two purposes.

I. The frequency of occurrence of the peptides common to the immunosignature peptide list and each test sample were used to generate a heatmap to show the comparative frequency of occurrence of the peptide in the ZIKV (target) and DENV (comparator).

II. the possibility of using the immunosignature method as an *in-silico* diagnostic assay for distinguishing between ZIKV and DENV was also investigated.

3.2.7 Re-panning third round output library

Following the results of the immunosignature data analysis, the serum samples that formed the 'immunosignature test cohort' ZIKV+ 11–20 and DENV+11–19, were biopanned with the round 3 output library to see if the results would be replicated. An additional 10 DENV and 9 ZIKV sera were biopanned against this sub-library (Figure 3-3). The sub-libraries were sequenced, and the data was used to test the ZIKV immunosignature that was previously defined.

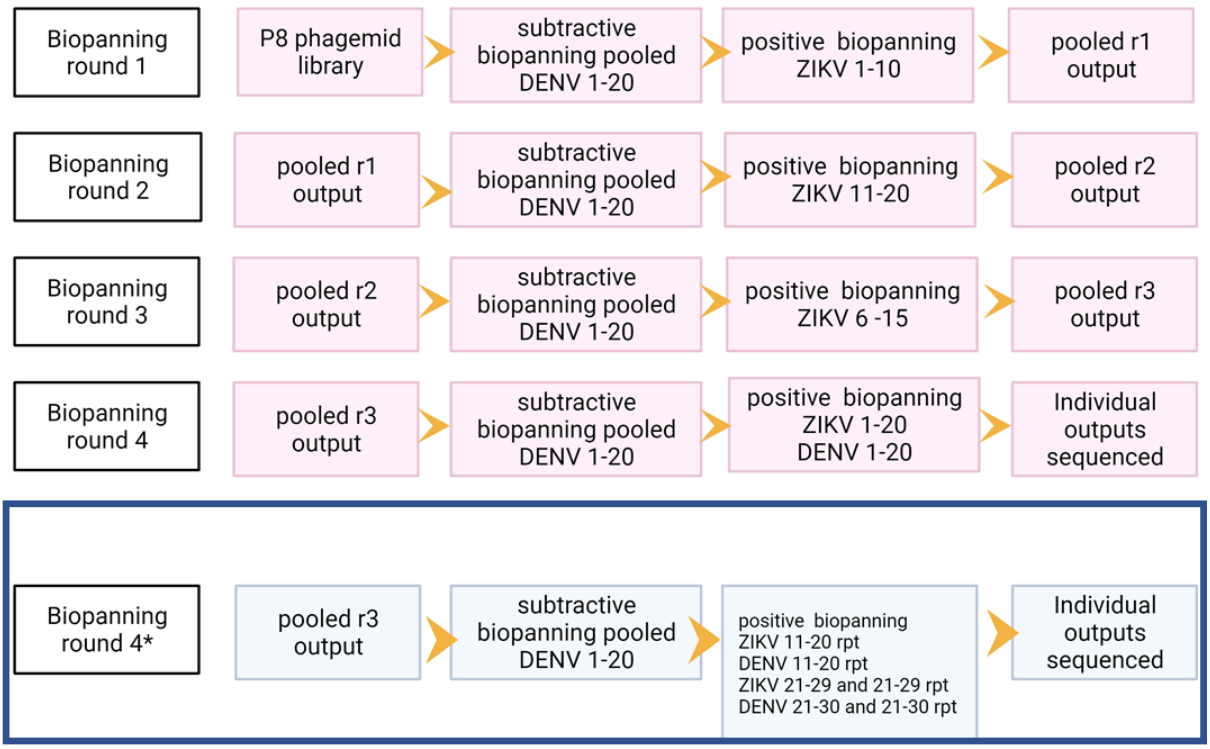


Figure 3-3: Schematic of all biopanning experiments conducted in this study. Highlighted in the blue box are the re-panning experiments.

3.3 Results

3.3.8 ELISA

Abcam DENV IgG ELISA: Using the Abcam Dengue ELISA, 111 out of 121 samples tested positive for Dengue IgG.

Abcam ZIKV IgG ELISA: Using the Abcam Zika IgG ELISA, 34 out of 121 samples tested positive for ZIKV IgG.

The results from the Abcam ELISA were compared with Zika IgG ELISA results conducted on the NIBSC in-house ZIKV ELISA. Twenty-nine samples were positive in both tests, and these were chosen for the ZIKV+ panel. These samples were also positive for DENV IgG on the ELISA.

Sera that became the DENV+ panel were picked if they had a negative Zika IgG result but a positive DENV IgG result, 44 samples met these criteria.

From the two panels, twenty samples each were taken forward for biopanning and these were randomly allocated into their groups using an online random number generator (https://www.random.org/lists/) to avoid sample bias. Table 3-1 and Table 3-2 shows the Abcam ELISA absorbance readings of the serum samples used for biopanning.

Table 3-1: ELISA absorbance values of serum samples positive for both DENV and ZIKV IgG on the Abcam ELISA and also positive for ZIKV IgG on the NIBSC ELISA.
These samples formed the ZIKV biopanning panel.

	ZKV+		
	sample ID	Absorbance values DENV IgG ELISA	Absorbance values ZIKV ELISA
1	752	2.62	0.38
2	472	2.03	3.98
3	834	2.59	2.83
4	292	2.07	1.04
5	803	1.92	0.95
6	523	2.54	2.17
7	570	2.00	0.45
8	836	2.64	0.53
9	525	1.95	2.33
10	410	2.18	1.94
11	751	2.14	0.41
12	770	2.70	2.67
13	833	2.54	1.12
14	775	2.18	1.46
15	810	2.09	4.76
16	319	2.15	3.70
17	369	2.16	1.93
18	786	1.92	1.96
19	749	2.59	1.14
20	568	2.03	1.78

Table 3-2: ELISA absorbance values of serum samples that tested positive for Abcam DENV IgG ELISA but negative on the Abcam ZIKV ELISA. These sera formed the DENV+ sera panel used in biopanning.

	DENV+		
	Sample Id	Absorbance values DENV IgG ELISA	Absorbance values ZIKV ELISA
1	741	2.5	0.003
2	577	2.468	0.148
3	511	2.53	0.008
4	753	2.651	0.053
5	447	2.565	0.085
6	476	2.509	0.032
7	768	2.424	0.047
8	593	2.457	0.034
9	409	2.483	0
10	763	2.427	0.026
11	582	2.5	0.118
12	734	2.612	0.009
13	759	2.322	0
14	798	0.313	0.074
15	571	1.721	0
16	806	2.651	0.041
17	795	2.468	0.071
18	479	2.55	0.056
19	743	2.544	0.105
20	559	0.298	0.009

Table 3-3: ELISA absorbance values of serum samples that were used in the re-screening of the round 3 output phage library

ZIKV+ samples		Absorbance value ZIKV IgG	Absorbance value DENV IgG
21	568	1.78	2.03
22	570	0.45	2.00
23	751	0.41	2.14
24	775	1.46	2.18
25	786	1.96	1.92
26	790	0.97	1.91
27	803	0.95	1.92
28	810	4.76	2.09
29	841	0.56	1.99
DENV+ samples			
21	593	0.034	2.457
22	594	0.079	2.519
23	599	0.096	2.596
24	724	0.108	2.553
25	725	0.067	2.745
26	734	0.009	2.612
27	741	0.003	2.5
28	743	0.105	2.544
29	753	0.053	2.651
30	759	0	2.322

3.3.9 Sera used for biopanning

Sera that were IgG positive for just DENV will be referred to as 'DENV' and sera that were IgG positive for both DENV and ZIKV will be referred as 'ZIKV'.

The serum samples and the panning sample number assigned to them in the biopanning experiments are shown in Table 3-4 along with the rounds of biopanning they were used in. The different rounds of biopanning used a different set of sera to ensure diversity of mimotopes being enriched over the rounds of biopanning. The serum samples will be referred to by their biopanning sample number.

Table 3-4: Samples used in biopanning

Panning sample number	ZIKV+	DENV+	Round of biopanning
1	752	741	1 and 4
2	472	577	1 and 4
3	834	511	1 and 4
4	292	753	1 and 4
5	803	447	1 and 4
6	523	476	1,3 and 4
7	570	768	1,3 and 4
8	836	593	1,3 and 4
9	525	409	1,3 and 4
10	410	763	1,3 and 4
11	751	582	2,3 and 4
12	770	734	2,3 and 4
13	833	759	2,3 and 4
14	775	798	2,3 and 4
15	810	571	2,3 and 4
16	319	806	2 and 4
17	369	795	2 and 4
18	786	479	2 and 4
19	749	743	2 and 4
20	568	559	2 and 4

3.3.10 Titre of P8 library phage

The P8 library was titrated before biopanning commenced and the library titre was 6.4 x 10^{12} pfu/mL in the phage library which was dissolved into 10 mL PBS buffer. The titre of the pooled output sub-library was calculated after every round (Table 3-5).

Table 3-5: Output titre of pooled phage library

Biopanning round	Titre (cfu/mL)
R1 ZIKV+	4 x 10^5
R2 ZIKV+	2.78 x 10^5
R3 ZIKV+	2.32 x 10^6
R4 ZIKV+	3.32 x 10^7
R4 DENV+	1.84 x 10^7

3.3.11 Preparing samples for Ion Torrent sequencing

Two stages of PCR were carried out as previously described in Sections 2.2.2 and 2.2.4, with the first stage of PCR being specific for the P8 library and the second stage for adding the barcoding primers so that the samples can be pooled together for Ion Torrent sequencing. The gel images in Figure 3-4 and Figure 3-5 show the single band at 250 bp and 350bp respectively which corresponds to the PCR product expected. Figure 3-6 shows the single band, which is the pooled and cleaned product that was sent for Ion Torrent sequencing.

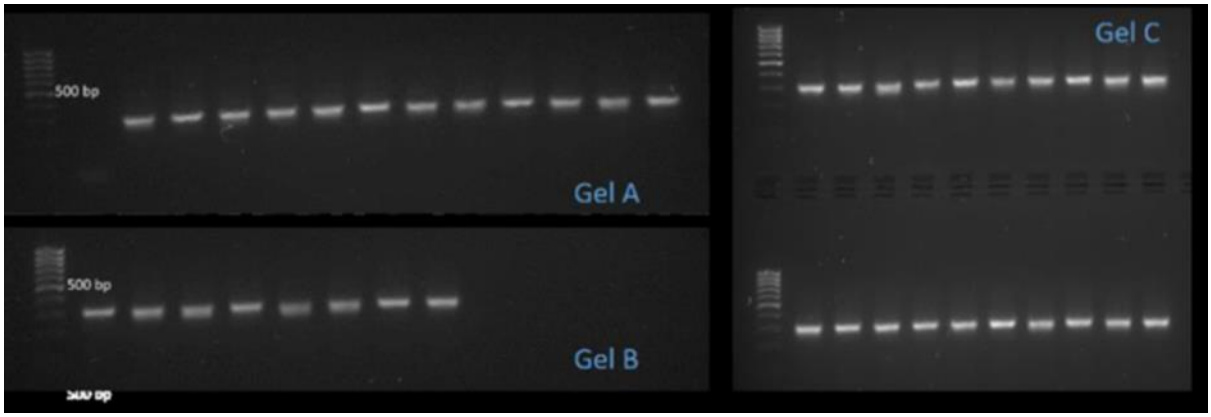


Figure 3-4: Gel images A-C shows round 1 PCR products seen just below 250 bp in all 3 images. Gels A and B are DENV+ R4 library outputs and gel C is ZIKV+ R4 outputs. The negative control was in the well adjacent to the ladder in gel A.

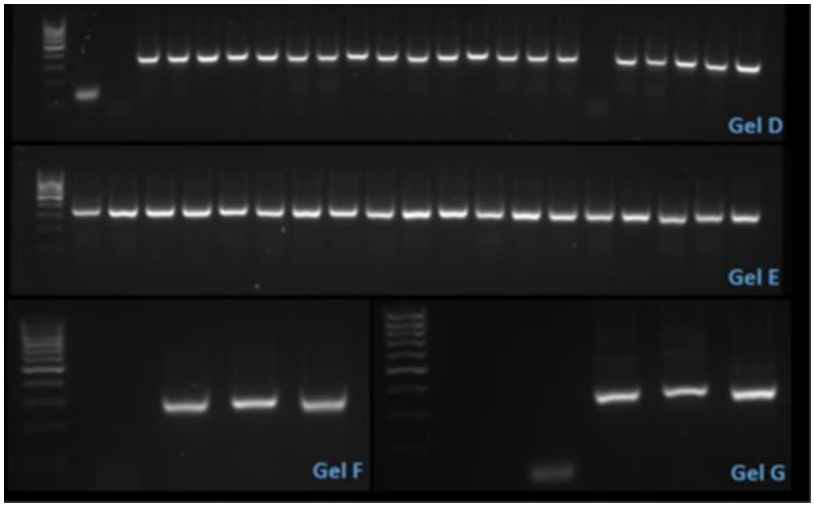


Figure 3-5: Gel images D and E show amplified PCR products following second round of PCR. PCR product of interest is seen just below 400 bp in comparison to the ladder. The two wells adjacent to the ladder of gel D are negative PCR controls from R1 (showing primer dimers only) and R2. The samples were loaded in the following orientation: ZIKV 1–20 plus DENV 1 R2 PCR products on gel D and DENV 2–20 R2 PCR products on gel E. The missing sample in gel D is sample 16. Gels F and G show the repeated R1 and R2 PCR for samples ZIKV 15, 16 and 17.

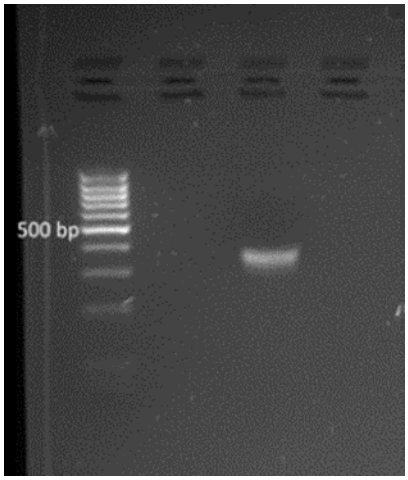


Figure 3-6: :Gel showing the single band of pooled DNA that was sent for sequencing.

3.3.12 NGS data analysis

The round 4 output libraries from 20 ZIKV+ samples were barcoded with barcodes 1–20. The sample lists were cross checked against the ELISA results before NGS data analysis and DENV 20 was discovered to have been negative for both DENV and ZIKV IgG and so the associated sub-library was discarded from the data analysis. The remaining 19 DENV+ libraries were barcoded with barcodes 21–39. After the initial data mining using scripts mentioned in section 2.4.1, Microsoft Excel and GraphPad Prism were used to analyse the data from the NGS files.

3.3.12.1 Z-score analysis

The Z-score of 5 was chosen for stringency to allow for the cross-reactivity expected from ZIKV and DENV being very closely related. At this Z-score, 1765 peptide sequences were seen enriched in ZIKV+ compared to DENV+.

Table 3-6 shows the peptides that were enriched in most ZIKV samples compared to the DENV pool at Z-score 5. The top 3 peptides were enriched in 8 out of 20 ZIKV+ files and the next best 6 peptide sequences were enriched in 7 individual ZIKV+ samples compared to the total pooled DENV peptides.

Table 3-6: Peptides enriched in individual ZIKV files compared to pooled peptides from all peptides in DENV 1–19 files

16-mer peptide	Number of ZIKV samples peptide seen in
QMPPPEFNGALVSGPS	8
PWNAKGLVVAPSTGGM	8
PIPIPNQRPWEPSSPL	8
SRNRPPPGRMPTLTPL	7
SPNSPRKILLPHENPT	7
RRPPSGSRHHNTHTTS	7
RRPPAGRSSSAPNTDP	7
KRPPPGMSPHTQLGTS	7
HPSPIQNSHRWELPLQ	7

3.3.12.2 Peptide immunosignature specific to ZIKV IgG

The NGS data was also analysed using the peptide immunosignature method described in section 2.4.4

Training data: the top 50 peptides by percent frequency were chosen to make two training sets using of ZIKV+ 1–10 (BC1–10) and DENV+ 1–10 (BC 21–30) each with 500 peptides. After removing peptides common to both data sets and removing duplicates in the ZIKV+ set, 145 peptides remained. This set of peptides was termed the 'ZIKV immunosignature'. Table of immunosignature peptides can be seen in Table A-.

Test data: the top 200 unique peptides by percent frequency from the remaining ten ZIKV+ and nine DENV+ files were individually tabulated. The samples were not pooled to enable mapping the peptides enriched in individual samples.

In GraphPad Prism, the 145 ZIKV immunosignature peptides were plotted against the top 200 peptides from test files into a heatmap using the number of times each peptide occurs in the data set for each test serum (Figure 3-7). More of the bands in the ZIKV test samples were brighter, indicating that there were peptides common to the ZIKV test samples and the immunosignature. There were some bands in the DENV test samples as well, but these were faint.

To account for the variations in depths of sequencing in the NGS data files, the peptide frequencies for individual peptides were converted from absolute counts to normalised peptide frequencies using the formula shown below.

The data was replotted into a new heat map (Figure 3-8) after converting the absolute counts of peptides in the test data to a percentage of total peptide count as mentioned above. This normalises the counts between samples. A baseline value of 0.5% was set to reduce background. A similar pattern seen in the first heatmap was seen here. This heatmap showed that DENV 19 had a similar profile to that seen in the ZIKV test samples.

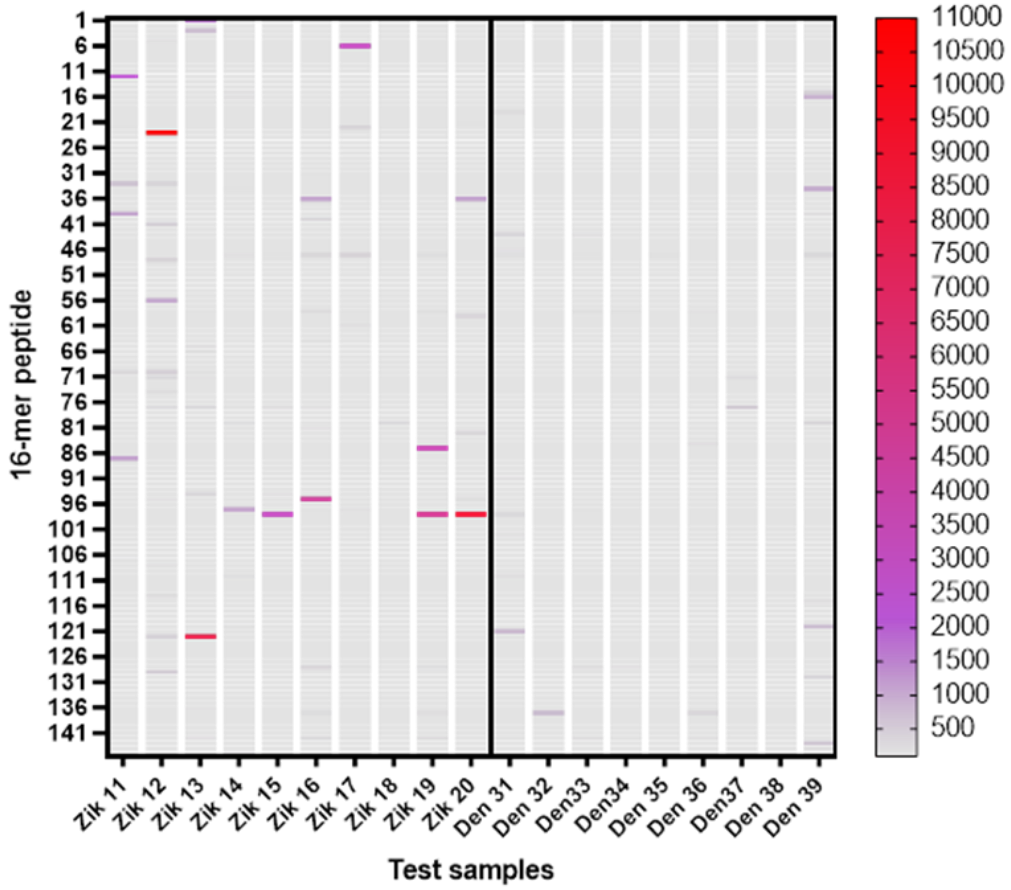


Figure 3-7: Heat map showing the ZIKV+ immunosignature peptides on the Y-axis and the test samples on the X-axis.

The colour intensity shows the number of reads of the peptide in each sample. The major value was set at an arbitrary value of 11,000 and baseline at 2,000 for the heatmap. The immunosignature peptides are shown on the left Y-axis as numbers due to there being 145 peptides.

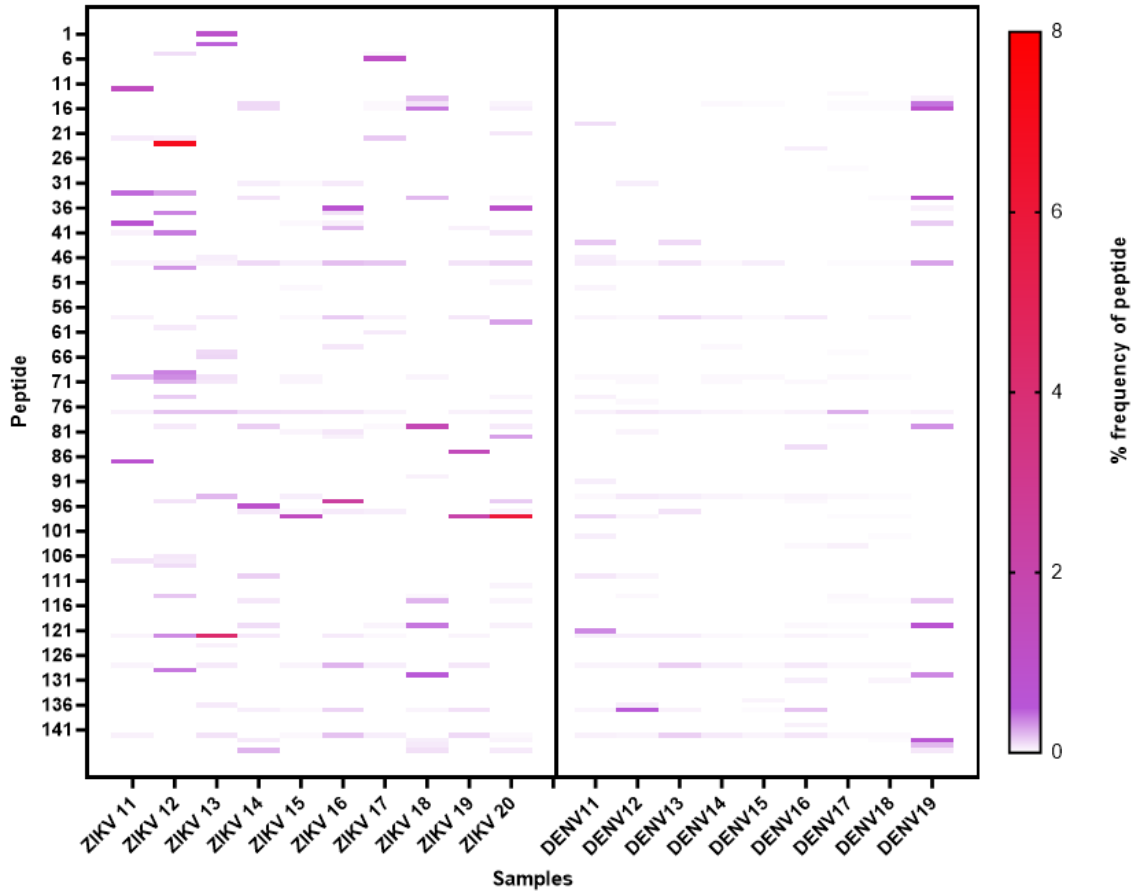


Figure 3-8: Heat map showing frequency (as a percentage of total reads) of the 145 immunosignature peptides in ZIKV and DENV test samples.
The baseline value shown in purple was set at 0.5 and the largest value was set at 8.

3.3.12.3 *In-silico* diagnostic using immunosignature method

The data from the ZIKV immunosignature was used to develop an *in-silico* ZIKV diagnostic assay. The sum of the normalised peptide frequencies in each sample in the immunosignature Test group was calculated (Table 3-7). This data was statistically analysed using unpaired t-test with Welch's correction. The P value is 0.0026 and therefore the immunosignature is statistically significant in discriminating between ZIKV/DENV IgG positive samples over those just positive for DENV IgG. The data from the table is presented graphically in Figure 3-9.

Table 3-7: Test samples and the sum of normalised peptide frequencies common to ZIKV immunosignature

Test samples	sum of all peptide frequencies
ZIKV 11	3.75
ZIKV 12	11.89
ZIKV 13	6.76
ZIKV 14	2.38
ZIKV 15	1.76
ZIKV 16	4.75
ZIKV 17	1.83
ZIKV 18	3.99
ZIKV 19	4.07
ZIKV 20	8.31
DENV11	1.34
DENV12	1.05
DENV13	0.85
DENV14	0.35
DENV15	0.24
DENV16	0.79
DENV17	0.6
DENV18	0.22
DENV19	5.04

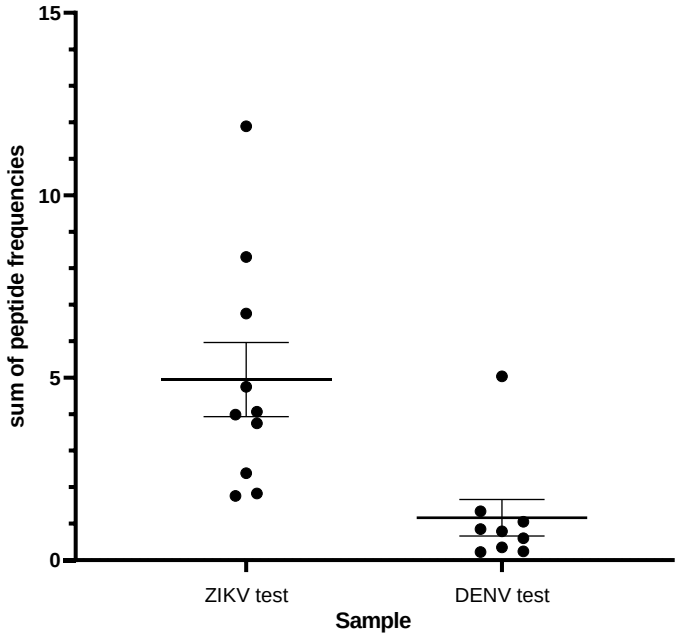


Figure 3-9: In-silico diagnostic using ZIKV+ immunosignature to test ZIKV and DENV test samples.

Horizontal bars are set at mean plus standard error of the mean.

3.3.12.4 Re-screening round 3 output sub-library

The results from the *in silico* analysis showed that the ZIKV+ immunosignature was able to distinguish between the ZIKV+ and DENV+ samples. To confirm this, the serum samples that formed the ZIKV and DENV test cohort were re-screened against the round 3 output sub-library. These were ZIKV 11–20 and DENV 11–19.

Table 3-8: Additional samples that were biopanned against the round 3 output library

Panning sample number	ZIKV+	DENV+	Round of biopanning
21	800	801	4*
22	841	433	4*
23	402	506	4*
24	406	485	4*
25	822	563	4*
26	790	799	4*
27	760	393	4*
28	555	832	4*
29	354	599	4*
30		594	4*

Additional sera (Table 3-8) were included in the re-screening process to test the robustness of the model if it were to be used as a diagnostic. The serum selection for biopanning stipulated ZIKV+ samples to be ZIKV IgG positive on both Abcam and NIBSC ELISA and also positive for DENV IgG on the Abcam ELISA. DENV+ samples were negative on both ZIKV ELISAs and positive for DENV on the Abcam ELISA. There were 9 ZIKV+ sera (ZIKV 21–29) and 10 DENV+ sera (DENV 21–30) and they were screened in duplicate.

ZIKV Immunosignature

The 145 peptides from the ZIKV immunosignature was applied to the new test sets ZIKV11r–20r/DENV11r–19r; ZIKV21–29/DENV21–30; and ZIKV21r–29r/DENV21r–30r. Heatmaps (Figure 3-9, Figure 3-11) were plotted using the settings used previously (Figure 3-8). The heatmaps of the replicate data are shown next to each other for comparison.

Although the peptide profiles changed slightly in the replicate test samples and there was a difference in the peptide profiles seen in the additional test samples, one particular peptide was seen enriched in the ZIKV data sets compared to the DENV data sets. This was peptide 97, 'PWHLLYATHPTAQLET'.

This peptide was seen highly enriched in ZIKV test samples 22, 28 and their replicates. The percent peptide frequency was 10.8% and 41.95% in 22 and 28 and 4% and 6% in the respective replicates. Data point ZIKV 28 is indicated by a red asterisk in Figure 3-11B as the percent peptide frequency was over the higher cut off value set for the heatmap.

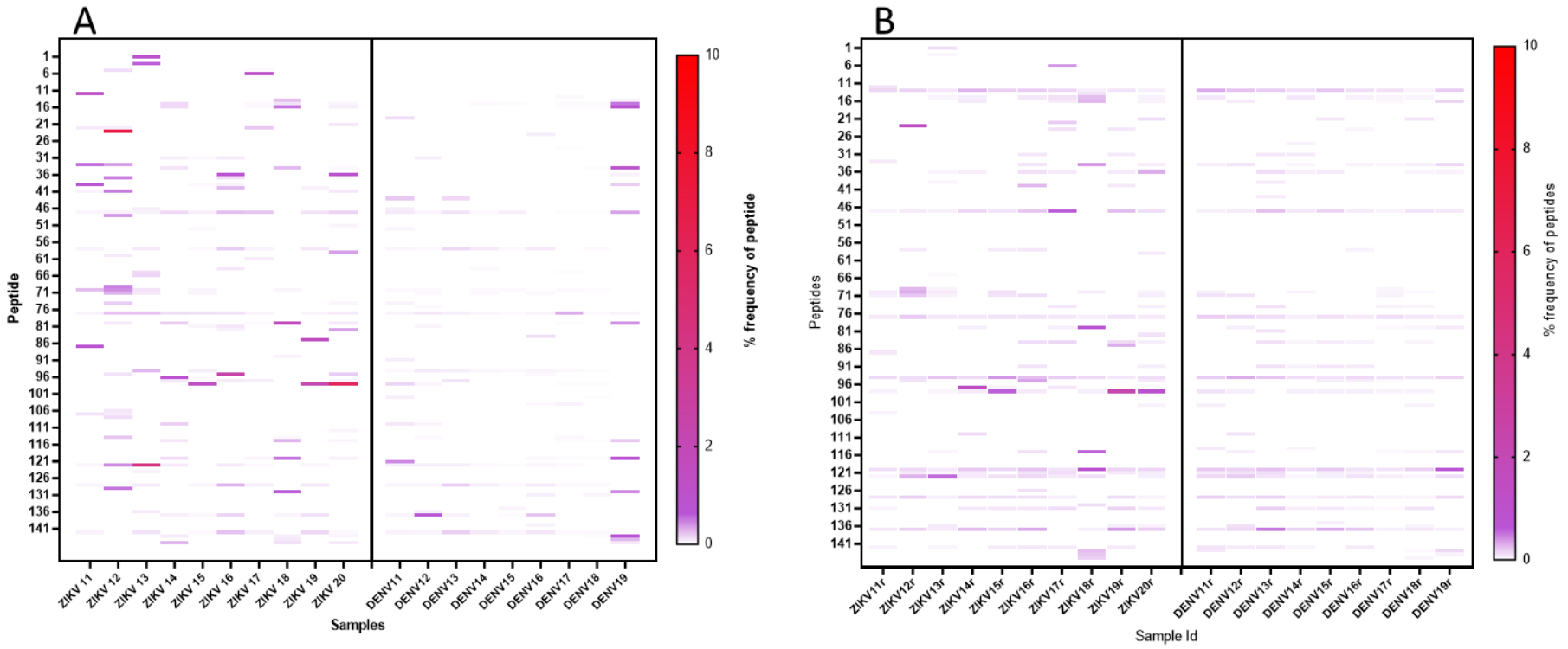


Figure 3-10: Heatmap A (also shown in Figure 3-8) shown for reference.

Heatmap B shows the top 200 peptides from test samples ZIKV 11r–20r and DENV 11r–19r fitted with the ZIKV immunosignature. Base value was set at 0.5%. There are brighter bands in the ZIKV test samples compared to the DENV test samples. Peptide 97 is consistently bright in the ZIKV test samples in both heatmaps. In the case of DENV 19 and DENV 19r, although the band pattern is present, the brightness of bands is not replicated in heatmap B.

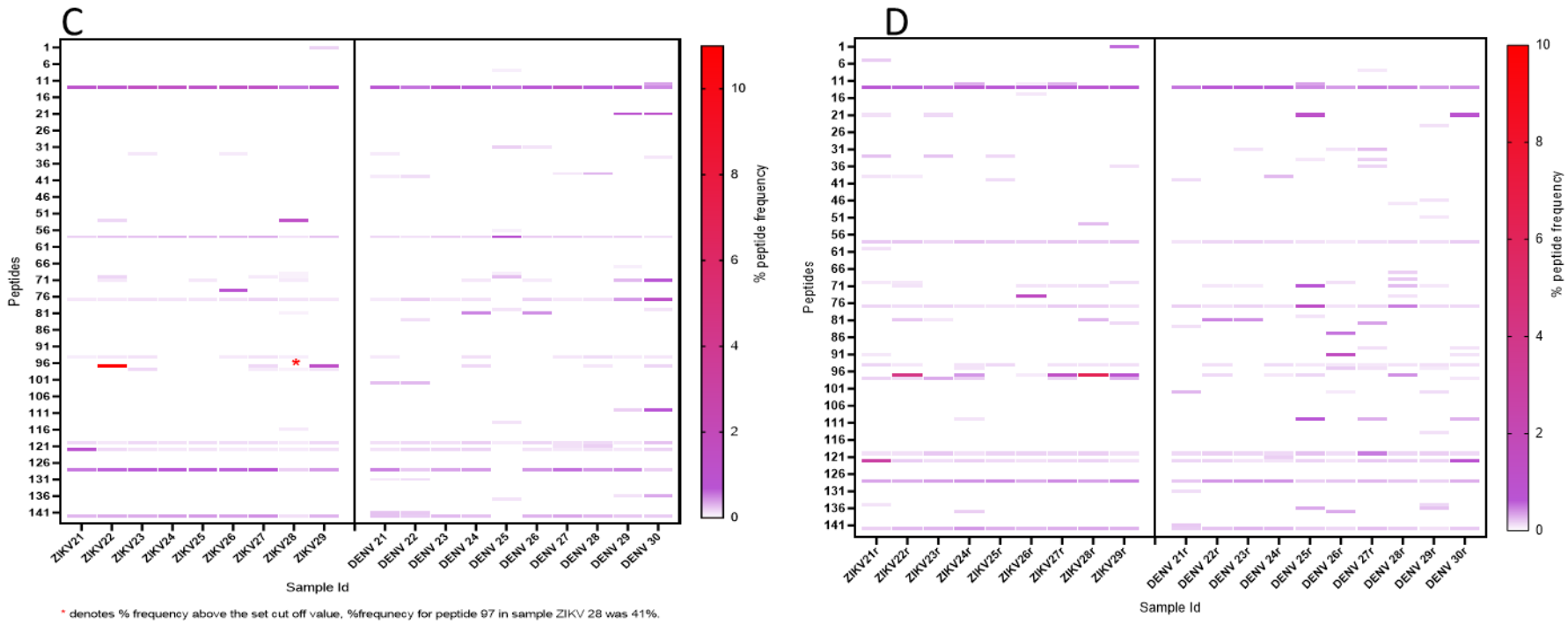


Figure 3-11: Heatmaps C and D shows the ZIKV immunosignature fitted with samples ZIKV 21–29/DENV 21–30 and their replicates ZIKV 21r–29r/DENV 21r–30r.

Base value for the heatmap was set at 0.5%. Intense bands corresponding to peptide 97 can be seen in multiple ZIKV samples. Asterisk in heatmap C corresponds to ZIKV 28 to indicate peptide 97.

3.3.12.5 Total peptide frequencies

The sum of percent frequency values of peptides common to the immunosignature was calculated (Table 3-10) and plotted as dot plots (Figure 3-12) for the data sets ZIKV11r–20r/DENV11r–19r; ZIKV21–29/DENV21–30; and ZIKV21r–29r/DENV21r–30r. As mentioned previously, due to peptide 97, the sum of peptide frequencies for ZIKV 22 and ZIKV 28 was much higher than the sum of peptides in other samples.

Figure 3-12B was plotted to show the data sets without the influence of ZIKV 28. The highest value on Y axis when including ZIKV 28 is 50 and by showing the data without the influence of peptide 97 in this sample, the highest value on the Y axis drops to 15. The data point for that sample was set to 0, highlighted as the red data point on the graph. The sum of percent frequencies for majority of the remaining samples were below 10.

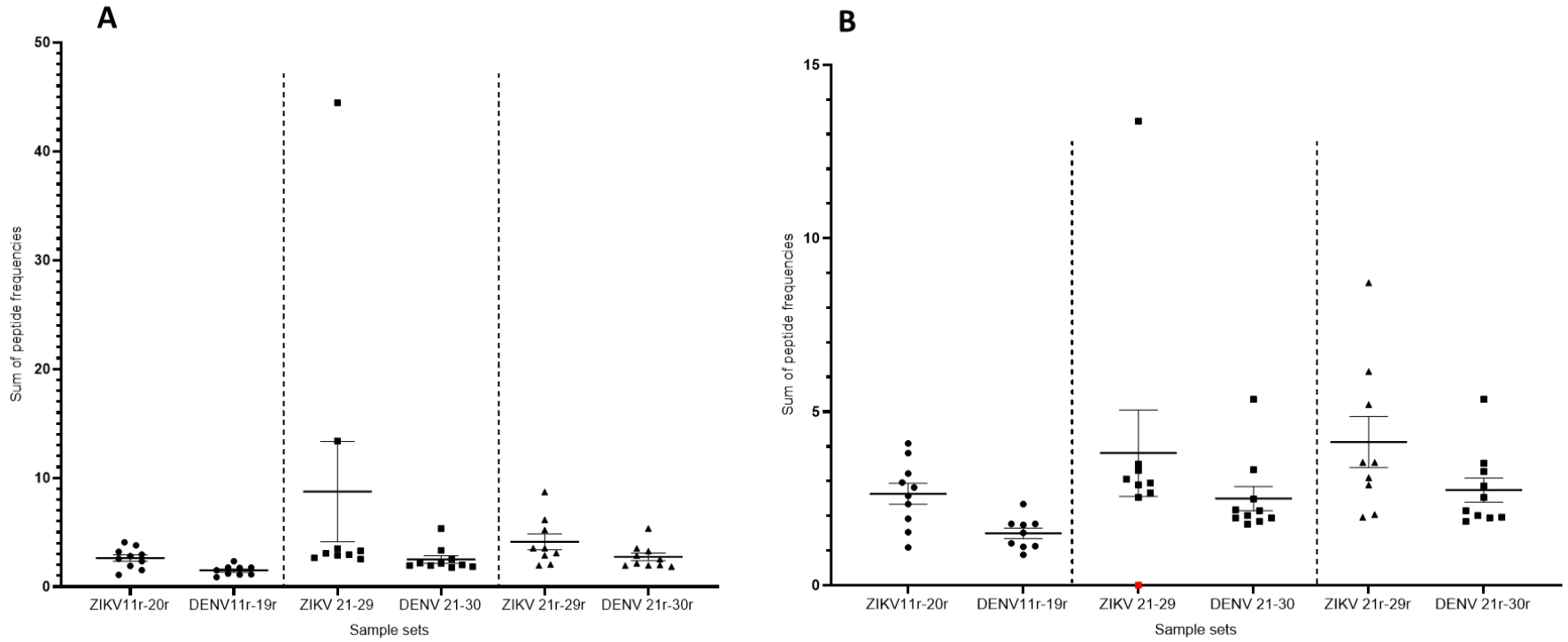


Figure 3-12: Plots A and B shows the sum of percent peptide frequency values in the test samples.
Plot B shows the data without ZIKV 28 which is the highest data point in plot A, the scale of the y-axis on the plots reflects this. Sample 28 was included in data set B, in the second panel as the red data point with an assigned value of 0.

3.3.12.6 Statistical analysis of the ZIKV immunosignature

The literature published on peptide array immunosignature data analysis has used a range of statistical analysis (Brown *et al.*, 2011) to analyse the data and ascertain whether the immunosignature has diagnostic capability in distinguishing between the target and control samples in the Immunosignature Test group. One of the methods is logistic regression and binomial logistic regression analysis on the software GraphPad Prism version 9.2 was used in this experiment. Binomial regression analysis was chosen as it offers a dichotomous outcome for the immunosignature, and it enabled for a subset of the immunosignature test group to be used to set the parameters for an *in-silico* diagnostic assay.

The sum of peptide frequencies from the immunosignature test samples were the variable used for analysing the immunosignature. A Receiver Operating Curve (ROC) plot was generated based on the values assigned to the target and control samples in this subset. The target samples (ZIKV 11-20) were assigned a value of 1 and the controls (DENV 11-19) were assigned a value of 0. The ROC curve generated by the software is shown in Figure 3-13. Table 3-9 shows the outcome of the ROC curve generated by the software.

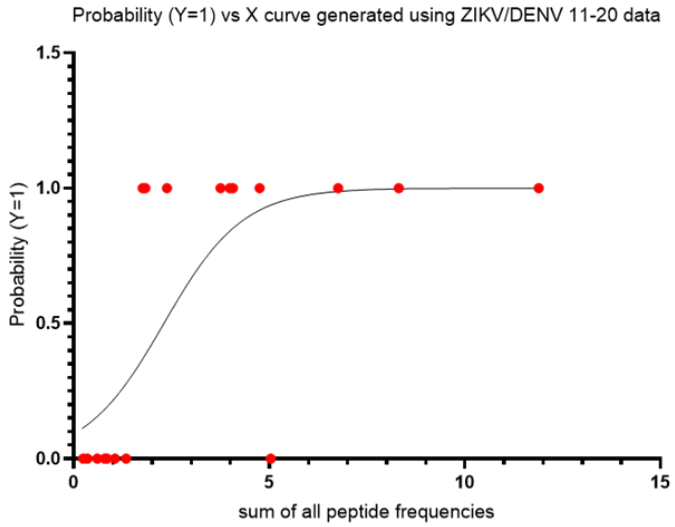


Figure 3-13: ROC plot generated as part of the binomial regression analysis of the ZIKV immunosignature. On the Y-axis, values at 1 were ZIKV 11—20 and values at 0 were DENV 11-19.

Table 3-9: The results from the ROC curve generated for the ZIKV immunosiganture. The positive/negative cut off value is 2.30.

The ROC analysis parameters calculated for the ZIKV immunosignature	
Best-fit values	
β0	-2.28
β1	0.990
X at 50%	**2.30**
Std. Error	
β0	1.09
β1	0.450
X at 50%	0.671
Area under the ROC curve	
Area	0.922
Std. Error	0.0766
95% confidence interval	0.772 to 1.00
P value	0.0019

In the best fit values of the ROC curve, β0 is the term given for the intercept and β1 is the coefficient for 'X', which here is the sum of all peptide frequencies. The model calculated that the 50% value for X was 2.3. Anything above this value would be classed as positive and anything below would be classed as negative.

The area under the curve (AUC) from the ROC analysis is a commonly used measure of accuracy for diagnostic assays. It informs on the sensitivity (ability of the test to detect true positives) and specificity (ability of the test to distinguish between true and false positives). The ZIKV immunosignature ROC analysis gave an AUC value of 0.922. Once the software plotted the ROC, it then calculated the predictive probability value of the remaining samples using X at 50% on the 'Best fit curve'. The P value of the ROC plot was 0.0019.

The rest of the test samples were fitted to the model and the probability values of all samples are shown in Table 3-10. The samples fitted to the ROC plot were ZIKV11r–20r/DENV11r–19r; ZIKV21–29/DENV21–30; and ZIKV21r–29r/DENV21r–30r. Samples where the sum of peptide frequencies was ≥2.3 would be considered positive and are highlighted in green.

Table 3-10: The ZIKV immunosignature *in-silico* analysis using sum of peptide frequencies, and predicted peptide probability values from the binomial regression model. Green shading indicates positive result for ZIKV (≥2.3) immunosignature.

Sample ID	Sum of all peptide frequencies	Assigned value	Predicted probability value
Samples in the 'immunosignature test' group assigned the value of 1 for the binomial regression model			
ZIKV 11	3.75	1	0.81
ZIKV 12	11.89	1	1
ZIKV 13	6.76	1	0.99
ZIKV 14	2.38	1	0.52
ZIKV 15	1.76	1	0.37
ZIKV 16	4.75	1	0.92
ZIKV 17	1.83	1	0.38
ZIKV 18	3.99	1	0.84

ZIKV 19	4.07	1	0.85
ZIKV 20	8.31	1	1
Samples in the 'immunosignature test' group assigned the value of 0 for the binomial regression model			
DENV11	1.34	0	0.28
DENV12	1.05	0	0.22
DENV13	0.85	0	0.19
DENV14	0.35	0	0.13
DENV15	0.24	0	0.11
DENV16	0.79	0	0.18
DENV17	0.6	0	0.16
DENV18	0.22	0	0.11
DENV19	5.04	0	0.94
Samples in the 'Immunosignature test' fitted to the simple logistic regression model			
ZIKV11r	1.09	-	0.23
ZIKV12r	3.22	-	0.71
ZIKV13r	1.53	-	0.32
ZIKV14r	2.96	-	0.66
ZIKV15r	1.92	-	0.41
ZIKV16r	2.82	-	0.63
ZIKV17r	2.34	-	0.51
ZIKV18r	3.81	-	0.82
ZIKV19r	4.09	-	0.85
ZIKV20r	2.58	-	0.57
DENV11r	1.74	-	0.36
DENV12r	1.77	-	0.37
DENV13r	2.34	-	0.51
DENV14r	1.13	-	0.24
DENV15r	1.51	-	0.31
DENV16r	1.21	-	0.25
DENV17r	0.88	-	0.2
DENV18r	1.11	-	0.23
DENV19r	1.77	-	0.37
ZIKV21	2.89	-	0.64
ZIKV22	13.38	-	1
ZIKV23	2.95	-	0.65
ZIKV24	2.66	-	0.59
ZIKV25	2.53	-	0.56
ZIKV26	3.49	-	0.76
ZIKV27	3.06	-	0.68
ZIKV28	44.46	-	1
ZIKV29	3.31	-	0.73
DENV 21	2.17	-	0.47
DENV 22	1.84	-	0.39

DENV 23	1.76	-	0.37
DENV 24	2.15	-	0.46
DENV 25	1.94	-	0.41
DENV 26	2.01	-	0.43
DENV 27	2.49	-	0.55
DENV 28	1.94	-	0.41
DENV 29	3.33	-	0.73
DENV 30	5.36	-	0.95
ZIKV21r	5.21	-	0.95
ZIKV22r	6.16	-	0.98
ZIKV23r	2.04	-	0.44
ZIKV24r	2.89	-	0.64
ZIKV25r	1.96	-	0.42
ZIKV26r	3.1	-	0.69
ZIKV27r	3.54	-	0.77
ZIKV28r	8.72	-	1
ZIKV29r	3.54	-	0.77
DENV 21r	1.84	-	0.39
DENV 22r	2.15	-	0.46
DENV 23r	2.01	-	0.43
DENV 24r	1.94	-	0.41
DENV 25r	5.36	-	0.95
DENV 26r	3.52	-	0.77
DENV 27r	2.86	-	0.63
DENV 28r	2.53	-	0.56
DENV 29r	1.96	-	0.42
DENV 30r	3.28	-	0.72

3.3.12.7 Sensitivity and specificity

The sensitivity and specificity of the model to distinguish between ZIKV and DENV samples were tested based on the predicted probability score (Table 3-10). The sensitivity and specificity of the model was calculated (Table 3-11) using the three immunosignature test data sets (ZIKV11r–20r/DENV11r–19r; ZIKV21–29/DENV21–30; and ZIKV21r–29r/DENV21r–30r). The sensitivity was 82.1% and specificity was 69.0%.

Table 3-11: Table A shows the formula used to calculate sensitivity and specificity values. Table B is the sensitivity and specificity values calculated for the ZIKV immunosignature *in-silico* diagnostic for the test samples fitted to the simple logistic regression model.

Positive result (Number)	A (true pos)	B (false pos)
Negative result (number)	C (false neg)	D (true neg)
sensitivity	A/(A+C)*100	
specificity	D/(D+B)*100	

A

	True positive (A)	False positive (B)	False negative ('C)	True negative (D)
ZIKV	23.0	0.0	5.0	0.0
DENV	0	9	0	20
Sensitivity	82.1			
Specificity	69.0			

B

3.4 Discussion

In this study, serum samples positive for antibodies of very closely related flaviviruses, ZIKV and DENV, were biopanned with a 16-mer peptide phage library to identify peptide mimotopes specific to ZIKV IgG. The NGS data from the biopanning output libraries were used to develop an immunosignature assay for ZIKV which has a sensitivity of 82.1% and specificity of 69%.

3.4.1 Serum samples

The sample cohort were collected during the ZIKV outbreak in the Caribbean in 2016. They were surplus to requirement clinical samples from patients diagnosed using molecular methods by CARPHA and were collected by NIBSC. The samples were tested for ZIKV IgG at NIBSC. Prior to 2015/2016, DENV was the most prominent flavivirus circulating in the Caribbean.

DENV and ZIKV antibody levels were determined in this sample cohort using commercially available DENV IgG and ZIKV IgG ELISAs as the aim of the biopanning experiment was to enrich peptide mimotopes to serum IgG. The ELISA results revealed that most of the serum samples had DENV IgG and a smaller proportion of samples were ZIKV IgG positive.

The almost universal presence of DENV IgG in the sample cohorts led the decision to use seropositivity for DENV IgG the baseline. The results of both ZIKV ELISAs were compared and sera that tested positive for both sets were chosen to form the ZIKV panel.

The serum samples were not tested for serum antibody to any of the other flaviviruses that are known to be present in the region such as yellow fever virus (YFV), West-Nile virus (WNV) and Saint Louis encephalitis virus (SLEV). The effect on the current results is negligible because the samples were submitted for clinical testing during an

outbreak of ZIKV, but if this experiment were to be repeated the additional information will be useful to understand results that stand out as was seen in NGS data point DENV 19.

Serum sample 'DENV 20' was removed from the data analysis because it was realised after panning that this serum sample was IgG negative for both DENV and ZIKV.

3.4.2 Phage library and biopanning strategy

In this study, a random peptide phage library expressing 16-mer peptides was used to generate peptide mimotopes that can distinguish between ZIKV and DENV IgG, which would be the antigenic regions of the virus. Existing literature on phage display in virology applications have used phage libraries generated specifically to the pathogen of interest (Fuentes *et al.*, 2016; Ravichandran *et al.*, 2019) or the phage library is biopanned with specific monoclonal antibodies to study the antigenic regions of the virus (Tarr *et al.*, 2006).

The positive panning sera being both ZIKV and DENV antibody positive, added an additional layer of complexity. Therefore, the control panel of sera used DENV IgG positive status as the baseline instead of using human sera negative for antibodies to both pathogens (ZIKV and DENV). Four rounds of biopanning were chosen to give the phage output library the chance to become more specific towards ZIKV IgG binders. The pitfall associated with this strategy is potential loss of phage that may be discriminatory for ZIKV but get depleted over the rounds of biopanning because they are not as competitive as a phage clone with less specificity for the biopanning target. The potential loss of good binders can be tracked by sequencing the pooled ZIKV libraries 1–4 and comparing the peptide diversity seen through the rounds of biopanning. Care was taken to reduce any bias towards a particular set of samples by biopanning with all 20 sera in the ZIKV set in different permutations in rounds 1–3.

3.4.3 NGS Data analysis

The first set of NGS data from this experiment was analysed on a Z-score NGS data analysis pipeline that was already established (Naqid, et al., 2016b). In this specific method of data analysis, the frequency (number of reads) of peptides enriched in individual ZIKV samples were compared to all the peptides seen in all the DENV samples (DENV 1–19). For a more stringent data analysis, Z-score was set to 5. Peptides were enriched in multiple ZIKV libraries with a maximum of 8 samples out of 20 over the DENV pool.

3.4.3.1 Establishing a peptide immunosignature for ZIKV from NGS data

The NGS data was also analysed using the 'Immunosignature' method as described in Section 2.4.4. A ZIKV specific peptide immunosignature comprised of 145 16-mer peptides was generated using the top 50 peptides from ten target (ZIKV 1-10) and control (DENV 1–10) samples. The immunosignature was then assessed against a test set of samples which were ZIKV 11-20 and DENV 11-19. The top 200 peptides from these samples arranged in order of frequency from were used for this.

3.4.3.2 Heatmaps as an early indicator of the success of ZIKV immunosignature analysis

The first method of data visualisation used for the ZIKV immunosignature results was in the form of a two tone heatmap generated on GraphPad Prism software. in the peptide array immunosignature data analysis, heatmaps are a common way to express experimental outcomes for each variable/condition (Stafford *et al.*, 2012). In this experiment, heatmaps were used to get an idea of the results especially as the data analysis methodology had not already been established.

Immunosignature data were plotted as heatmaps with the lowest threshold set to 0.5% peptide frequency of the DENV samples to reduce

background noise. The heatmap results showed that brighter bands were present in ZIKV test samples compared to most of the DENV test samples, except for DENV19. The brightness of bands in DENV 19 was similar to the profiles seen in ZIKV test samples This led to questions regarding the authenticity of this sample being truly positive for DENV IgG only, even though it was negative in the screening ELISA for ZIKV IgG.

The heatmaps that were generated following the rescreening of the phage sub-library showed that although the peptide profiles were not identical among the replicates of ZIKV 11–20/DENV 11–19 and ZIKV 21–29/DENV 21–30, there were some peptides especially Peptide 97 'PWHLLYATHPTAQLET' which was detected in more abundance in the ZIKV samples compared to the DENV samples in all the screens. This could possibly be controlled by using purified IgG instead of serum IgG in the biopanning rounds. The result from this chapter shows that it is useful in early visualisation of data.

3.4.3.3 Binomial regression analysis and using ROC plots to assess the sensitivity and specificity of the immunosignature.

A more robust method of understanding the outcome of the biopanning experiment using the immunosignature assay was required. The sum of all peptide frequencies in the test samples of peptides common to the ZIKV immunosignature were calculated. Statistical analysis of the first immunosignature analysis used unpaired t-test and this showed that the sum of peptide frequency in ZIKV test samples was significantly higher than DENV test samples, except for one DENV sample which had sum of peptides above the cut off (DENV19). This result mirrored the heatmap profiles seen for these samples.

To test the robustness of the ZIKV peptide immunosignature, the R3 ZIKV sub-library was re-panned with the samples that formed the immunosignature test cohorts. Additional ZIKV and DENV sera that had

not yet been used in the biopanning were also screened against this library. The new test data sets were challenged with the ZIKV peptide immunosignature.

As the immunosignature has the potential for use as an *in-silico* assay, the sum of peptide frequencies common to the established ZIKV immunosignature peptide list was added up and this was the basis of the data analysis. As explained earlier in the results section, binomial regression analysis was used as the statistical test used to analyse the robustness of the immunosignature based on the literature on peptide array immunosignature assay (Brown *et al.*, 2011).

The sum of peptide frequencies from the initial immunosignature test data was used in the binomial regression model. Using a subset of the samples in the immunosignature test subset, a ROC plot was drawn. The AUC for the plot was 0.92 and this falls under 'outstanding' discrimination between the target and control criteria (Mandrekar, 2010). The sum of peptide frequencies in the remaining samples in the immunosignature test cohort (including the re-screened samples and the additional samples that were screened) were plotted against the ROC plot by the software. The sensitivity and specificity of the ZIKV immunosignature was calculated to 82% and 69% respectively.

The current commercial assays where recombinant ZIKV NS1 has been used with 100% sensitivity and 94% specificity in ZIKV samples and 0.7% cross-reactivity in acute DENV samples (Euroimmun IgG ELISA). The Euroimmun assay for ZIKV IgG used to cost around £1000 for a 96 well plate assay and this might not be suitable for diagnostic labs in low and middle income countries operating on smaller budgets. The immunosignature peptides with the best discriminatory power in a serology assay could be used to create a cheaper assay with the same level of sensitivity and specificity.

A NS1 rapid diagnostic test (RDT) in development reported sensitivity/specificity for ZIKV IgG at 99/99.3% and cross reactivity to DENV at 16% (Kim *et al.*, 2018). Other NS1 assays in development have reported sensitivity/specificity of 67/96% in a combined assay of RDT and ELISA (Yap *et al.*, 2021).

ZIKV NS1 plant expressed ELISA.

The serum samples used in biopanning, ZIKV 1–20 and DENV 1–20 were tested on a ELISA under development where the coating antigen was plant expressed ZIKV NS1 (unpublished data). In the ELISA using plant expressed NS1, three of the DENV samples tested positive and these samples were DENV 2, 5 and 19. For confirmation of antibody status, these cross-reactive sera should be tested using virus neutralisation assays for ZIKV and DENV but weren't in this study.

3.4.4 Future work

The results of this biopanning experiment have shown that through the application of NGPD, peptides or a peptide profile specific to one flavivirus can be detected over a closely related and cross-reactive flavivirus in an individual serum sample. The biopanning outputs were stored as individual libraries and as pooled libraries, so there is potential to return to analyse peptide profiles across each specific round.

The data set that was generated for this experiment can be further explored to develop a better immunosignature model using methods described for peptide array immunosignatures. Following further refinement of the ZIKV immunosignature, the next steps would be to synthesise the peptides and test them against the independent panel of sera as a serology assay. Peptide 97 should be investigated further for its ability to detect ZIKV IgG in a serology assay.

3.4.5 Concluding remarks

This chapter outlined the first experiment that was performed, and it was also the most complex of the three experiments in this thesis. In this chapter, the first instance of using Immunosignaturing for NGPD data sets has been outlined. The statistical analysis of the ZIKV immunosignature showed promise in differentiating between the antibodies raised to two flaviviruses from the same sample.

Chapter 4 **Identifying peptide mimotopes that are discriminatory for serum antibodies to louping ill virus compared to healthy ovine sera**

4.1 Introduction

Louping ill virus (LIV) is a tick-borne flavivirus that causes louping ill disease and is endemic to the British Isles with disease distribution concentrated mainly in the uplands of northern England, Scotland, and Wales. Serological evidence of LIV has been reported from willow ptarmigan samples collected in Norway (Ytrehus *et al.*, 2021).

LIV has an economic impact as sheep and red grouse, which are both farmed, are susceptible to the virus. The vaccine used for disease control in sheep has been unavailable since 2017 due to manufacturing issues. At present a vaccine is currently in development (Game and Wildlife Conservation Trust, 2021). Since 2018, there have been 47 clinically diagnosed cases of LIV in the UK according to data reported in the sheep disease surveillance dashboard and, so far, in 2021, there have been 7 cases of LIV in the UK, 4 in Scotland and 3 in England (APHA and SRUC, 2021). Weaned lambs are moved to graze on the upland hill pastures where the tick population is very high and due to waning immunity provided by maternal antibodies, they are at a higher risk of infection and potentially dying from the infection (Jeffries *et al.*, 2014). The flocks are left to graze on the hills and aren't monitored closely and so the actual number of sheep that die from the disease are potentially higher than that reported.

Serum antibodies to LIV are detected using haemagglutination inhibition test (HAI) and serum neutralisation test (SNT). However, these assays require containment level 3 facilities for antigen preparation for HI and for growing the virus for SNT, the latter also being time consuming and

requiring highly skilled operators. An ELISA-based diagnostic would be beneficial for a quicker turnaround time, automation and to screen larger sample numbers and reduction in the requirement for containment level 3 facilities.

The aim of this chapter was to apply traditional phage display techniques to identify a panel of peptides that can detect LIV IgG and distinguish between seropositive and seronegative serum. In traditional phage display, the phage sub-library following biopanning are infected into bacteria which are plated onto selective media. Bacterial colonies which represent the individual phage clones are picked, grown and the phage clones are tested using ELISA format. Traditional phage display has been applied to identify mimotopes of viral antigen epitopes from biopanning serum antibodies, which were then characterised for their application in immunotherapy and vaccine design (Zhang et al., 2012).

The P8 random peptide library expressing 16-mer peptides introduced in the previous chapter was used in this study. Phage expressing peptide mimotopes that bound to the target, here LIV IgG, were positively selected for by consecutive rounds of biopanning with sheep serum. In this experiment, three rounds of biopanning were carried out rather than the four rounds reported for the previous chapter. Although the phage library is being enriched in the multiple rounds of biopanning, there is a chance that during progressive rounds the diversity of binding clones reduces and so this experiment had only three rounds of biopanning.

Objectives:

1. To biopan the P8 phage library with LIV IgG positive and healthy ovine sera (LIV IgG negative sera) to select for phage displaying peptides that selectively bind to LIV IgG.
2. To select a panel of individual phage clones from the LIV IgG enriched phage library and develop a phage ELISA.

3. To test the peptide panel with independent ovine serum samples for discrimination between LIV IgG positive and negative sera.

4.2 Materials and methods

4.2.1 Serum samples

Forty two ovine serum samples that were positive for LIV IgG were provided by Dr Mara Rocchi at the Moredun Research Institute (MRI). These samples were submitted to the MRI between 2014 and 2017 for laboratory confirmation of LIV infection and had tested positive for LIV IgG and IgM (total Ig) using the haemagglutination inhibition (HI) assay. Heat treatment of an aliquot of serum at 64.5°C for 30 minutes was undertaken to denature the IgM molecules allowing quantification of the remaining immunoglobulin subclasses.

Table 4-1 provides the details of the LIV IgG positive ovine sera used in the biopanning process, the total immunoglobulin titre and specifically IgG titres. The animal ID in the table reflects the year the sample was taken in and the animal details at the time of submission.

The LIV vaccination status of these animals was unknown. Animals with a total Ig titre equal or marginally higher than the corresponding IgG titre (less than 4 fold difference) are more likely to have pre-existing immunity from the vaccine or prior exposure. The sera that were chosen for the biopanning had a higher total Ig/IgG difference (also called IgM predominance) and most likely came from unvaccinated sheep that were recently infected. Sera with a minimum IgG titre of 1/40 and above were selected to provide the best chance of enriching for binders to LIV IgG through the biopanning.

MRI also provided 41 'uninfected' sera which were from animals whose health/flock history was known. The sera were collected in 2016 and

2017. Table 4-2 provides information of the healthy sheep sera used in the biopanning experiment. Sera from different animals were used to ensure the enriched sub-library is still diverse and not enriched towards a single animal.

Table 4-1: LIV IgG positive ovine serum samples used in biopanning

Panning sample number	Animal Id	IgG only	Total Ig	round of biopanning
LIV pos 1	14/0647/1	1/320	1/1280	1 and 3
LIV pos 2	14/0724/2	1/640	1/2560	1 and 3
LIV pos 3	14/0724/4	1/1640	1/2560	1 and 3
LIV pos 4	15/0441	1/640	1/1280	1 and 3
LIV pos 5	15/0675	1/320	1/5120	1 and 3
LIV pos 6	16/0383/1	1/160	1/2560	1 and 3
LIV pos 7	16/0562	1/640	1/2560	1 and 3
LIV pos 8	16/0704	1/1280	1/5120	1 and 3
LIV pos 9	17/0433	1/160	1/2560	1 and 3
LIV pos 10	17/0504/1	1/640	1/5120	1 and 3
LIV pos 11	14/0859	1/160	1/5120	2
LIV pos 12	14/0916	1/40	1/320	2
LIV pos 13	14/0938/2	1/640	1/5120	2
LIV pos 14	16/0358	1/320	1/2560	2
LIV pos 15	16/0409/2	1/160	1/2560	2
LIV pos 16	16/0741/2	1/160	1/640	2
LIV pos 17	16/0797	1/160	1/320	2
LIV pos 18	16/0860	1/160	1/2560	2
LIV pos 19	17/0224	1/160	1/1280	2
LIV pos 20	17/0319	1/320	1/1280	2

Table 4-2: LIV IgG negative serum ovine samples used in biopanning

Panning sample number	Animal Id	round of biopanning
LIV neg 1	16/0092/1	2
LIV neg 2	16/0092/3	2
LIV neg 3	16/0092/4	2
LIV neg 4	16/0092/5	2
LIV neg 5	16/0319/2	2
LIV neg 6	16/0319/4	2
LIV neg 7	16/0543/1	2
LIV neg 8	16/0543/7	2
LIV neg 9	17/0209/3	2
LIV neg 10	17/0209/5	2
LIV neg 11	16/0092/7	3
LIV neg 12	16/0092/8	3
LIV neg 13	16/0319/6	3
LIV neg 14	16/0319/7	3
LIV neg 15	16/0543/9	3
LIV neg 16	16/0543/10	3
LIV neg 17	16/0886/2	3
LIV neg 18	16/0886/3	3
LIV neg 19	17/0209/1	3
LIV neg 20	17/0209/2	3

4.2.2 Biopanning strategy and methodology

The P8 random peptide phage library was grown, PEG precipitated and titrated before use in the biopanning experiment. The phage library culture, precipitation and titration processes have been described in chapter 2. The details provided in this section reflect the specific methods used in this experiment.

Initially, the phage library was blocked with a milk block solution (3% M-PBS Table A-) and was subtractively biopanned with 30 µL of protein G agarose beads to remove phage that bind to naïve protein G agarose beads. Following this stage, the P8 phage library was biopanned as shown in the biopanning strategy used for this experiment shown in Figure 4-1. The general biopanning methods described in Section 2.1.6—2.1.13 were followed.

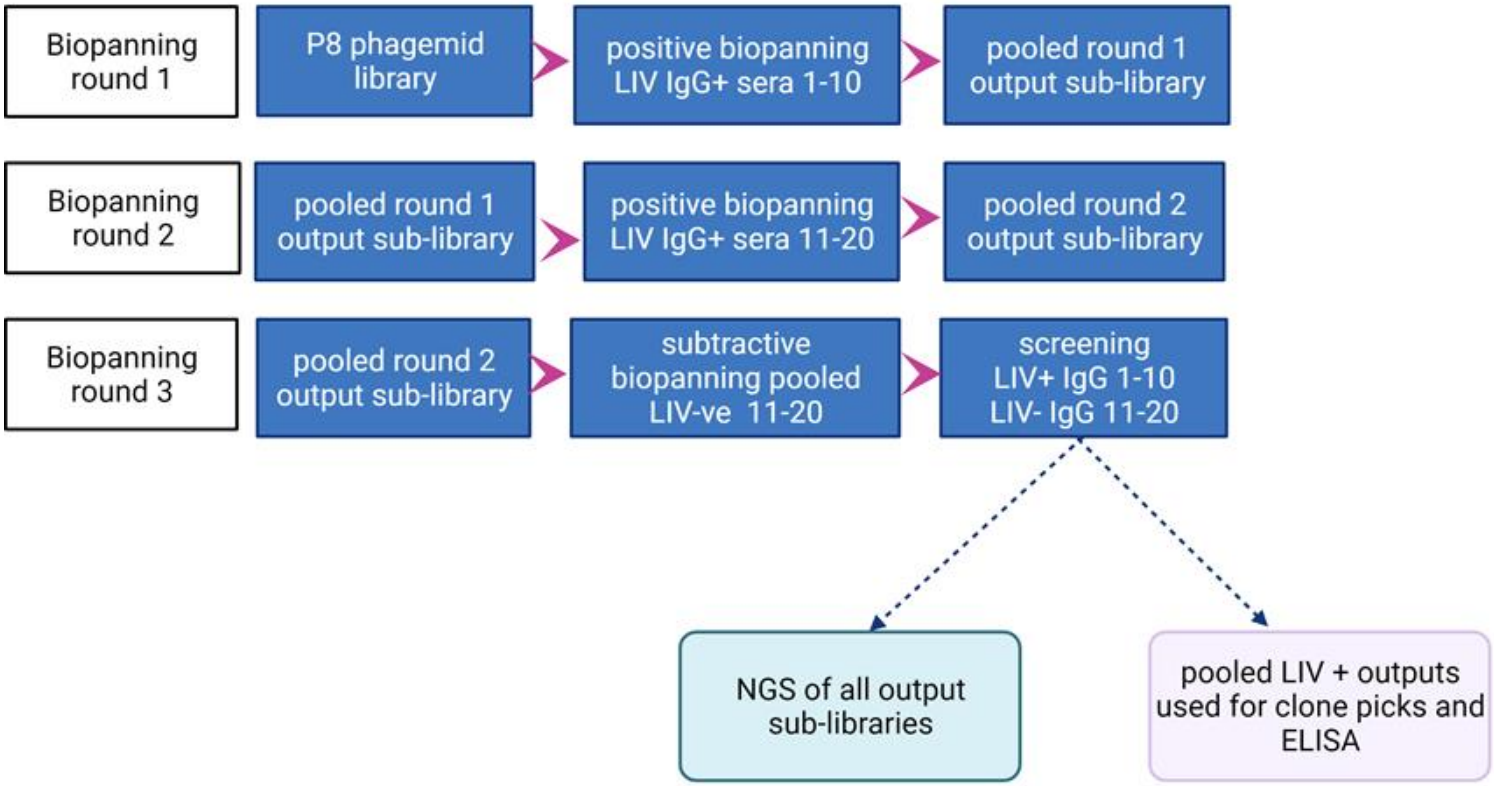


Figure 4-1: Biopanning strategy used to enrich phage binders to LIV IgG.

In the first two rounds, there were positive biopanning with LIV IgG+ sera. In the third round of biopanning, subtractive biopanning with LIV IgG-ve sera (11—20) was carried out and the phage library was screened with LIV IgG+ and LIV IgG-ve sera. From the third-round sub-libraries, all the sub-libraries were sequenced on the Ion Torrent platform. Pooled LIV IgG+ve phage sub-library was used for random colony picking and characterisation of the colony picks.

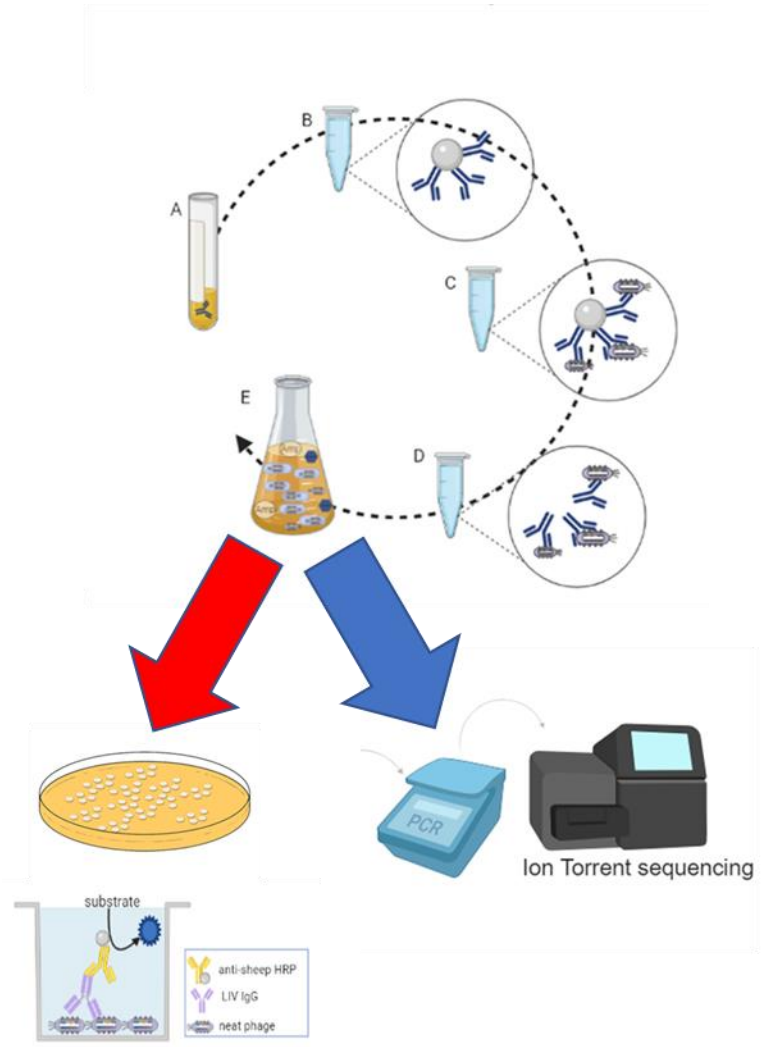


Figure 4-2: Schematic showing biopanning strategy.
A-E shows the biopanning process. In the final round of biopanning, two methods of detecting phage mimotopes to LIV IgG were carried out. The red arrow indicates clone picking. The pooled LIV R3 library was plated on selective media, individual clones were picked and Sanger sequenced and subsequently used for developing an ELISA to detect LIV IgG in ovine serum. The blue arrow indicates that individual output libraries from the third round were sequenced using the Ion Torrent platform, thus sequencing as much of the phage clones as possible.

4.2.1 The biopanning process

In the positive biopanning experiments, 5 µL sera was incubated with 30 µL washed protein G agarose beads. The beads were incubated with the serum for a minimum period of 30 minutes and blocked with 3% M-PBS (Table A-) for a further 30 minutes after which the microcentrifuge tubes

were spun at 2,500 x *g* to pellet the agarose beads with serum IgG. The supernatant was removed and 500 µL phage library was added to each tube and incubated (biopanned) overnight at 4°C. The output phage were rescued the following day.

Positive biopanning rounds 1 and 2

The first two rounds of biopanning were positive biopanning rounds (Figure 4-1). The output sub-libraries were stored individually and as pooled output from that round of biopanning. In round 1 (R1), the P8 phagemid library was biopanned against LIV+ 1-10. In the second round of biopanning (R2), the pooled R1 output phage sub-library was biopanned against samples LIV+11–20. The output sub-libraries were stored individually and also as pooled R2 output phage sub-library.

Subtractive and positive biopanning in round 3

In the third and final round of biopanning, the R2 sub-library was negatively biopanned with LIV-11—20. Five microlitres of serum from each LIV negative sera 11–20 was pooled and incubated with the Protein G agarose beads for 30 minutes and blocked with 3% M-PBS for a further 30 minutes. After removing the supernatant, the agarose beads were incubated with the R2 phage sub-library to remove non-specific phage binders to ovine serum. Following this subtraction, the non-bound sub-library was transferred to a clean 50 mL tube, ready for the positive biopanning step.

The phage library was then 'screened' (positively biopanned) as shown in Figure 4-1 with LIV IgG+ sera 1–10 and LIV IgG- sera 11–20. The rescued sub-libraries were stored as individual samples and pooled sub-libraries for the LIV +and LIV-sera were also made. The round 3 (R3) LIV output phage library was made by pooling the individual LIV+ sub-libraries.

Rescuing the output library and bacterial stocks

The phage sub-libraries from the overnight biopanning were rescued as explained in section 2.1.12. The titre of the pooled output libraries were calculated and bacterial glycerol stocks of the output libraries were prepared for longer-term stability as explained in section as explained in section 2.1.13.

4.2.2 Colony picking (selection)

The bacterial glycerol stock of R3 LIV phage sub-library was used for the colony picking step. A dilution series was made by initially adding, 100 µL of phage bacterial glycerol stock to 9.990 mL of 2YT media to make dilution A (1:1000). Subsequently, 100 µL dilution A was mixed with 900 µL 2YT media to make dilution B (1:100) and this was repeated till dilution E. From each dilution, 100 µL was plated on 2YT/*Amp*/Gluc agar plates (Table 2-3,Table 2-5) and incubated at 37°C overnight.

The plates were examined by eye and after identifying single colonies, 80 colonies (phage clones) were picked and grown in duplicate in 2 x 96 deep-well plates in 600 µL 2YT/*Amp*/Gluc broth (Section 2.3). A control well was set up as baseline to measure the OD of the culture. The cultures were grown at 37°C in a shaking incubator at 220 rpm. When the colony picks reached 0.5 at OD_{600nm}, they were superinfected with the helper phage M13KO7 and left to incubate for 30 to 45 minutes to allow helper phage superinfection. This was followed by centrifugation to pellet the bacteria and removal of the old media; the phage clones were grown overnight in 2YT/*Amp*/*Kan* at 30°C and 180 rpm in a shaking incubator.

The next day, the plates were centrifuged to pellet the bacteria and the phage supernatant from each well was transferred to a clean deep well plate and PEG precipitated and used as coating antigen for ELISA. The PEG precipitation was added to reduce the background in ELISA.

The replicates in the second deep well plate were grown to a high density OD_{600nm} 0.8-1; after centrifugation to remove the supernatant, the bacterial pellets were resuspended in 200 µL of freezing media and stored at $^{-}$80°C.

Scaling up individual phage clone cultures

The culture volume of the clones increased from 600 µL volume in deep well plates to 10 mL cultures in 50 mL centrifuge tubes and finally 50–100 mL cultures grown in 250 mL flasks as the clones moved down the phage ELISA testing pipeline.

4.2.3 Sanger sequencing phage clones

Phagemid DNA was extracted from all 80 clones using plasmid DNA miniprep kit (Qiagen) and the concentration of DNA was quantified on the Nanodrop (Thermo Fisher Scientific). A volume of 5 µL per sample was submitted to Source Bioscience for Sanger sequencing along with the unidirectional primer pc_89, 19-mer sequence 'ATTACCCCAGCTGGCGAAA-3' (Eurofins).

4.2.4 LIV Phage ELISA method development

Initially 20 of the 80 selected clones were used to set the assay parameters for a proof of principle LIV IgG phage ELISA. The serum samples used in the biopanning rounds were pooled and used for the phage ELISA. The LIV IgG positive sera (Table 4-1) were pooled to make an LIV IgG serum pool and the same was done with the LIV IgG negative sera (Table 4-2). During the first three ELISA method development studies shown in Table 4-4, the phage clones were tested against the positive sera pools only and from the fourth ELISA onwards, the phage clones were also tested against the LIV IgG negative sera pool.

In the first development ELISA, serum was coated on the microplate wells and tested with the phage clones. Detection antibody was anti-fd

bacteriophage alkaline phosphatase (AP). From the second developmental ELISA onwards the phage clone was used as the coating antigen, as seen in

Figure 4-3. Table 4-4 shows the steps carried out to refine the LIV IgG phage ELISA method. Helper phage M13KO7 was introduced as a negative phage control in the 2nd ELISA. The final method is shown in Appendix 3 (Section A.3.4).

Briefly, PEG precipitated phage clones (dissolved in PBS) were coated on wells of a 96 well plate (Nunc MaxiSorp®) in duplicate and incubated overnight at 4°C. The excess coating antigen was removed, and the wells were washed with PBS-Tween 20 (0.1%v/v) buffer. The wash buffer was removed, plates were patted to remove excess liquid, then wells were blocked with 200 µL 3% M-PBS for an hour at 37°C. The block solution was discarded, and 50-100 µL serum diluted 1:200 in 3% M-PBS was added and incubated for an hour at 37°C. After the incubation, the wells were washed four times with PBS-(0.1%) Tween buffer. Rabbit anti-sheep IgG horse radish peroxidase (HRP) antibody diluted 1:5000 was added to the wells in 50–100 µL and incubated at 37°C for 60 minutes.

The wells were washed and dried after incubation. Substrate, 3,3',5,5'-Tetramethylbenzidine (Thermo Scientific™ 1-Step™ Slow TMB-ELISA Substrate Solution) was added to wells (50–100 µL) and incubated at room temperature for 25 minutes. The reaction was stopped by the addition of 50–100 µL 1 M sulphuric acid. The OD of the wells was measured at wavelength 450nm on a microplate reader within 30 minutes of the reaction being stopped. There were 4 separate negative controls set (Table 4-3). Control 1 and 2 were coated with helper phage and tested with pooled serum and 3% M-PBS respectively. Control 3 was coated with PBS and tested with pooled serum and control 4 was uncoated well tested with 3% M-PBS.

Table 4-3: ELISA control wells used in the LIV IgG phage ELISA

ELISA control	Coating antigen	Test sera
Control 1	Helper phage	Pooled serum dilution
Control 2	Helper phage	Milk block
Control 3	PBS	Pooled serum dilution
Control 4	uncoated	Milk block

Once the ELISA method was established, the phage clones were tested against pooled LIV positive and negative sera that had already been used in the biopanning. The phage clones preferentially reacting to LIV IgG positive pooled sera were tested against the individual members of the biopanning sera panels. Clones which gave the best discrimination between the LIV IgG+ sera and LIV IgG – sera used in the biopanning process were then tested against a panel of 20 LIV IgG+ and 20 LIV IgG- sera not previously used in the biopanning rounds (independent sera). The phage clones which preferentially recognised LIV IgG+ sera over LIV IgG -sera in each stage of the ELISA screening experiments were termed ‘best performing clones’.

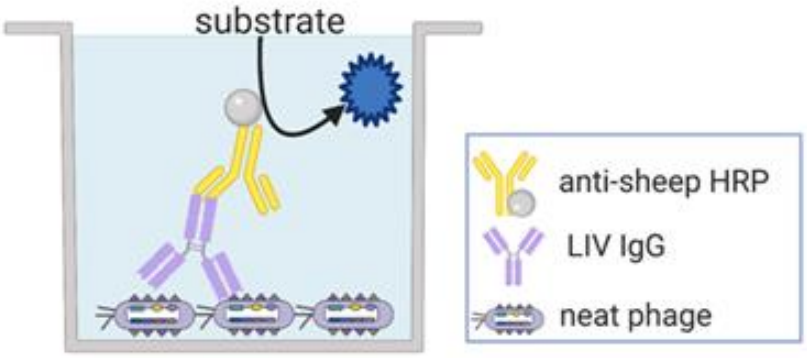


Figure 4-3: Finalised ELISA format using neat phage as coating antigen and serum samples were added to the phage clone coated wells. Anti-sheep HRP was used as detection antibody.

4.2.5 ELISA data analysis

The raw OD ELISA data was captured by the Manta software (Dazdaq Solution Ltd) and exported into Microsoft Excel. An Excel template spreadsheet was set up and this automated data processing. The automated data processing included subtracting the Control 4's OD value from all the test and remaining control wells; calculating mean OD values and setting cut offs.

In developmental ELISAs 1-6, the cut off values were the value above 2xSD+mean of control 1. In assays 7 and 8, the cut off values were 3XSD+mean of control 1. ROC analysis and dot plots were generated for the independent sera ELISA results for the individual phage clones using GraphPad Prism software.

4.2.6 Ion Torrent sequencing

Round 3 sub-libraries, 10 LIV positive and 10 LIV negative, from the biopanning were sequenced on the Ion Torrent sequencing platform using the method that described in Chapter 2.2.

Table 4-4: LIV IgG ELISA development experiments

Experiment	W8ell coating	Target	Detection antibody	Temperature
1	Pooled sera at 1:100 and 1:1000 dilution	Phage supernatant	rabbit anti-fd bactetiophage AP and anti-rabbit-AP	Room temperature
2	Phage supernatant 1:100, 1:1000 dilution	Pooled serum 1:50	Anti-sheep IgG HRP	Room temperature
3	PEG precipitated phage 1:100, 1:1000 dilution	Pooled sera 1:50	Anti-sheep IgG HRP	Room temperature
4	Neat PEG precipitated phage	Pooled sera 1:500	Anti-sheep IgG HRP	Room temperature
5	Neat PEG precipitated phage	Pooled sera 1:500	Anti-sheep IgG HRP	Room temperature
6	Neat PEG precipitated phage	Pooled sera 1:500	Anti-sheep IgG HRP	37°C
7	Neat PEG precipitated phage	Pooled sera 1:500	Anti-sheep IgG HRP	37°C
8	Neat PEG precipitated phage	Pooled sera 1:500	Anti-sheep IgG HRP. Slow reacting TMB	37°C
9	Neat PEG precipitated phage	Pooled sera 1:500	Anti-sheep IgG HRP. Slow reacting TMB	37°C

4.3 Results

4.3.1 Phage titre

The titre of the P8 phage library that was used in the first round of biopanning was 3.08 x 10^{14} cfu/mL. The output titre of the library for the pooled LIV positive library and the pooled control library are shown in Table 4-5

Table 4-5: Phage titre of pooled sub-libraries following each round of biopanning

Biopanning round	Phage library	Titre (cfu/mL)
0	P8	3.08 x 10^{14}
1	Pooled LIV positive	3.3 x 10^{5}
2	Pooled LIV positive	3.62 x 10^{5}
3	Pooled LIV positive	3.56 x 10^{7}
3	Pooled LIV negative	3.5 x 10^{5}

4.3.2 Sanger sequencing of the colony picks

The 80 phage clone picks underwent Sanger sequencing and the 16-mer peptide corresponding to the bacterial clone is listed in Table 4-6. The flanking regions were used to identify the peptide from the sequence and SnapGene Viewer was used to translate the DNA sequence to amino acids. The flanking regions are shown below:

GCTGAGGGTGAATTC: AEGEF

GATCCCGCAAAAGCGGCCTTTG: DPAKA

Seventy-eight unique clones were isolated from the colony picks: clones 39 and 45 and clones 42 and 76 represented the same peptide. Two

clones were of different length to the expected 16-mer peptide. Clone 11 had poor sequencing quality and the peptide sequence was not identified.

Table 4-6: 16-mer peptides following Sanger sequencing the phage clones

Number	Peptide	Comments
1	RRRTPVREIFNNGAPN	
2	TNFKSETKPTVTTTVA	
3	KQGSDNFPPAPTSAST	
4	TKDMTLWSRPVVPTNA	
5	QKNQGLASFAKPVIHT	
6	YLKQWIEQRNPTDGTL	
7	GPMPAQMWYKSDHAPT	
8	AYFAETISIDQHSTTP	
9	NADRKLAHQTKMWPVP	
10	YNQRPSTTQQSYVPTA	
11	no sequence	poor sequence quality
12	RNEKLQRYAQLTFNER	
13	FNAIDHKVHAGPSPSV	
14	ALARAARQVNKALNTP	
15	TTLQSSYMPPRDVQAA	
16	YRSSWEKIAPSSTPSS	
17	KSPGRSMLTMAPSTTH	
18	PKGHANSGNSYQSPNP	
19	LNMQYHPADTLRSAFT	
20	SKFVYSLKAKTDSHID	
21	PRGYTPYSQNKPFGHA	
22	NWPTRANTPYNAPQFV	
23	AHPNIDFKSRKEILAM	
24	SRFHSYMLGNAPKPAP	
25	ASLARAEVDNMKPLPTD	
26	LHWPTSMRRNPPVTSS	
27	PWPTKASTLAREPIST	
28	RMLTSRHRSLPETEMG	
29	YRNSSIYNKKHPMTPL	
30	RKTHAKAWGGQIPTSD	
31	SHWDKPYIPQMIPIDN	
32	KYHPRDRVTTTPTSNT	
33	PPYLPKHAHSNKETWP	
34	ILPDTQFAPRYNTVGL	
35	NRPTKVPPRTSLEVPL	
36	AYGNLTFASLHDTLAL	
37	YGKYRVNGGPPSVNAA	

38	ATVRKGSHQMNKPVTI	
39	TGAPIRAGVPLTHAKP	duplicate colony 39&45
40	NWPSAVNKPTAERLNA	
41	KKTNSGTYTATEKPQP	
42	PGPMSNWNKMESANAS	duplicate colony 42& 76
43	NDNHKVRSSTEAIPPW	
44	PHRGVKTYANAERAEW	
45	TGAPIRAGVPLTHAKP	duplicate colony 39&45
46	SHWDKPYIPQMIPIDN	
47	PSTRQPSSLAREYSHI	
48	TDMVAELTHLAPLERQ	
49	WRSNSMNKPSHLWMTA	
50	LKSTHSHNWPRSMNMS	
51	GVFAASTLHQSNNTGI	
52	LHDSRSSTSKMVSWLE	
53	SAANSKQHTAHVNDAV	
54	TPLSSAGRSTKPLTSI	
55	GKYAARRTATNTSELR	
56	KWFAASTLLATNATAH	
57	SMPSKNWPTPMGKCAP	
58	KTLLNSSIDGGTPRST	
59	FDVSSLFASVHQDARM	
60	SWYGYSSDKPNIPGPS	
61	KKTIVTFAKGPPDARA	
62	SSLRDFVDRVSRDNGV	
63	RPSYNPNDPKYHDKPI	
64	TKQHTHQYYNSRMTTW	
65	YMPLDTLRNSTTSGPF	
66	YWPNDVVNYTSSSLAN	
67	YSYTNPGVSKPWRNSG	
68	IRRLRPKLTSDVENRL	
69	RTRSQADKPWNPSNQM	
70	FSRFSKPAPPGTSMGP	
71	AYYFPIDALPDPIPAP	
72	PWKTMEKTSAPHAVNA	
73	NKINKLMPHP	*10 AA*
74	WWLSVLAHLASLQPTS	
75	NEPWKPSLEHLPTTSNNLA	*19 AA*
76	PGPMSNWNKMESANAS	duplicate colony 42&76
77	LYNLKYEKLSDNPPYT	
78	KRSTRSTEVFPIPSPL	
79	TTHSSGKNLHMTDQPR	
80	PMQSNSASKLLNHHKP	

4.3.3 Phage ELISA development

Phage clones 1–20 were used to develop the method for the phage ELISA. Once the method was finalised, the remaining clones were tested. Method development and refinement were carried out over a series of ELISA assays as described below.

4.3.3.1 Developmental ELISA 1

The format used in this ELISA was phage antigen capture, where LIV IgG positive sera from the biopanning panel was pooled and diluted to 1:500 in PBS buffer. This was used to coat the ELISA plate which was then incubated with the phage supernatant from phage clones 1–20. The detection antibody was anti-fd bacteriophage, which binds to the coat protein of M13 phage. The output from this assay was poor due to extremely high background and the format was changed for the next ELISA.

4.3.3.2 Developmental ELISA 2 and 3

The second ELISA results are shown in Table 4-7. The phage supernatants were PEG precipitated and coated on microplate wells. To identify optimal coating concentrations, the clones were coated on the wells at 1:100 and 1:1000 dilutions. The OD readings indicated that the ELISA was successful as some of the phage clones bound strongly to the pooled LIV IgG positive sera. The phage clones were tested against LIV positive serum only. For the data analysis, the mean of the OD values from the duplicate wells were calculated and the OD values of plate blank (uncoated/sera pool) were subtracted from the rest of the wells.

In the third developmental ELISA, M13KO7 was introduced as the phage (negative) control for the ELISA, as shown in Table 4-8. The LIV positive sera should not react to this phage. For the data analysis, the mean of the OD values from the duplicate wells were calculated and the OD values of plate blank (uncoated well/sera pool) were subtracted from the rest of

the wells. OD values higher than the phage control were considered positive.

Table 4-7: Absorbance values at OD450 from ELISA 2. Yellow shaded wells have OD readings after extraction of plate blank (0.186).

ELISA 2	Phage dilution coated on wells	
	1:100	1:1000
LIV clone 1	0.197	0.188
LIV clone 2	0.097	0.151
LIV clone 3	0.353	0.355
LIV clone 4	0.218	0.186
LIV clone 5	0.373	0.403
LIV clone 6	0.075	0.112
LIV clone 7	0.105	0.193
LIV clone 8	0.112	0.172
LIV clone 9	0.213	0.257
LIV clone 10	0.133	0.203
LIV clone 11	0.253	0.185
LIV clone 12	0.085	0.125
LIV clone 13	0.181	0.233
LIV clone 14	0.046	0.106
LIV clone 15	0.096	0.148
LIV clone 16	0.231	0.183
LIV clone 17	0.121	0.167
LIV clone 18	0.307	0.219
LIV clone 19	0.265	0.223
LIV clone 20	0.621	0.300
average plate blank	0.186	n/a

Table 4-8: Absorbance values at OD450 from ELISA 3. Results in yellow are the values that are higher than those of the phage control (helper phage M13KO7). The phage clones were only tested against LIV IgG positive serum pool in this experiment

ELISA 3	phage dilution coated on wells	
Phage coated on wells	1:100 dilution	1:1000 dilution
LIV clone 1	0.389	0.108
LIV clone 2	0.027	0.098
LIV clone 3	0.530	0.293
LIV clone 4	0.414	0.296
LIV clone 5	0.613	0.405
LIV clone 6	0.116	0.029
LIV clone 7	0.203	0.060
LIV clone 8	0.149	0.062
LIV clone 9	0.435	0.081
LIV clone 10	0.108	0.102
LIV clone 11	0.169	0.098
LIV clone 12	0.084	0.053
LIV clone 13	0.202	0.053
LIV clone 14	0.043	0.011
LIV clone 15	0.153	0.039
LIV clone 16	0.596	0.079
LIV clone 17	0.206	0.043
LIV clone 18	0.312	0.028
LIV clone 19	0.446	0.116
LIV clone 20	1.235	0.109
negative: M13KO7	0.775	n/a
no phage	0.401	n/a

4.3.3.3 Developmental ELISA 4 to 6

Results from these assays are shown in Figure 4-4 and the tabulated data is presented in Appendix 3 (Table A3-3). Neat PEG precipitated phage clones were coated on wells and tested against pooled LIV positive (LIV IgG+ 1—20) and pooled LIV negative (LIV IgG-1—20) sera at 1:500 dilutions. The plates were incubated at room temperature for assay 4 and 5. The positive/negative cut off value for this assay was based on the value from control 1 (helper phage/pooled LIV neg sera).

Assay 5 (Figure 4-4B) was a direct repeat of assay 4 yet gave much higher signals and background. To control fluctuations of room

temperature, the ELISA plates were incubated at 37°C in a dry heat incubator from assay 6 onwards.

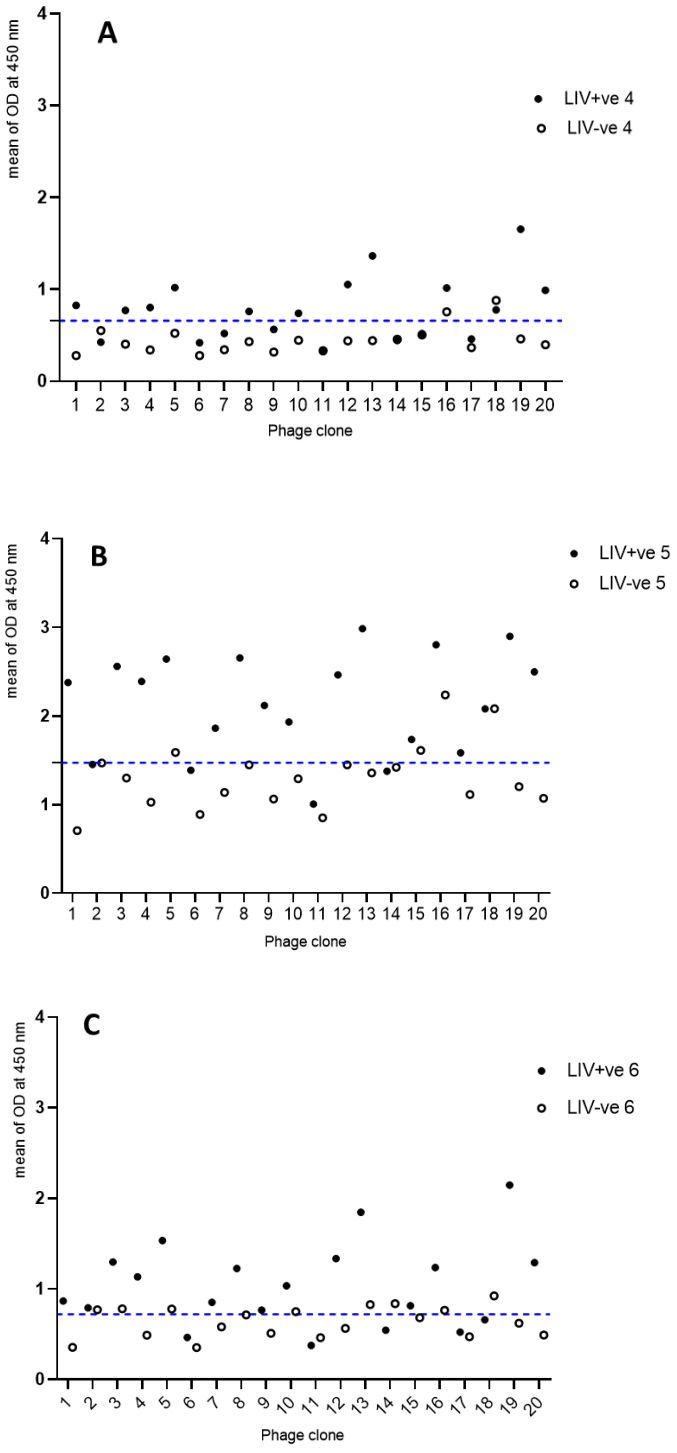


Figure 4-4: ELISA results of the individual phage clones 1—20 tested against pooled LIV IgG positive and negative sera (1:500) from experiments 4 (A), 5 (B) and 6 (C). Experiments 4 and 5 were replicates tested on separate days, but there were fluctuations in the OD readings which were attributed to change in room temperature. The following developmental ELISA was repeated with all the same conditions as the previous 2 assays but with incubations at 37°C. The samples were plotted against the mean OD value at 450nm for all three experiments. The blue dotted line in the plots

shows the positive/negative cut off based on control 1 (phage control/pooled LIV negative sera) OD value. The cut off value for plot A was 0.66, plot B was 1.5 and plot C was 0.72.

4.3.3.4 Developmental ELISA 7 and 8

In these two ELISAs, phage clones 21–40 were tested using the method refined by assay 6. The phage clones were tested against pooled LIV positive and negative sera at 1:500 dilution and incubations were carried out at 37°C in a dry heat incubator. Figure 4-5 shows the data distribution plotted with the cut off value indicated; the tabulated version is presented in Appendix 3 (Table A3-4).

The positive/negative cut off values were set based on the values of the pooled LIV negative sera for all the 20 test clones, mean+3xSTDEV of the OD for LIV negative serum.

In experiment 7 the cut off was calculated to be 3.31 whereas in experiment 8, the cut off was 0.368. The assays were carried out on separate days, the change in OD values in assay 8 is most likely due to the change made to the substrate solution when it was changed from 1-Step™ Ultra TMB-ELISA Substrate Solution (assay 7) to 1-Step™ Slow TMB-ELISA Substrate Solution (assay 8).

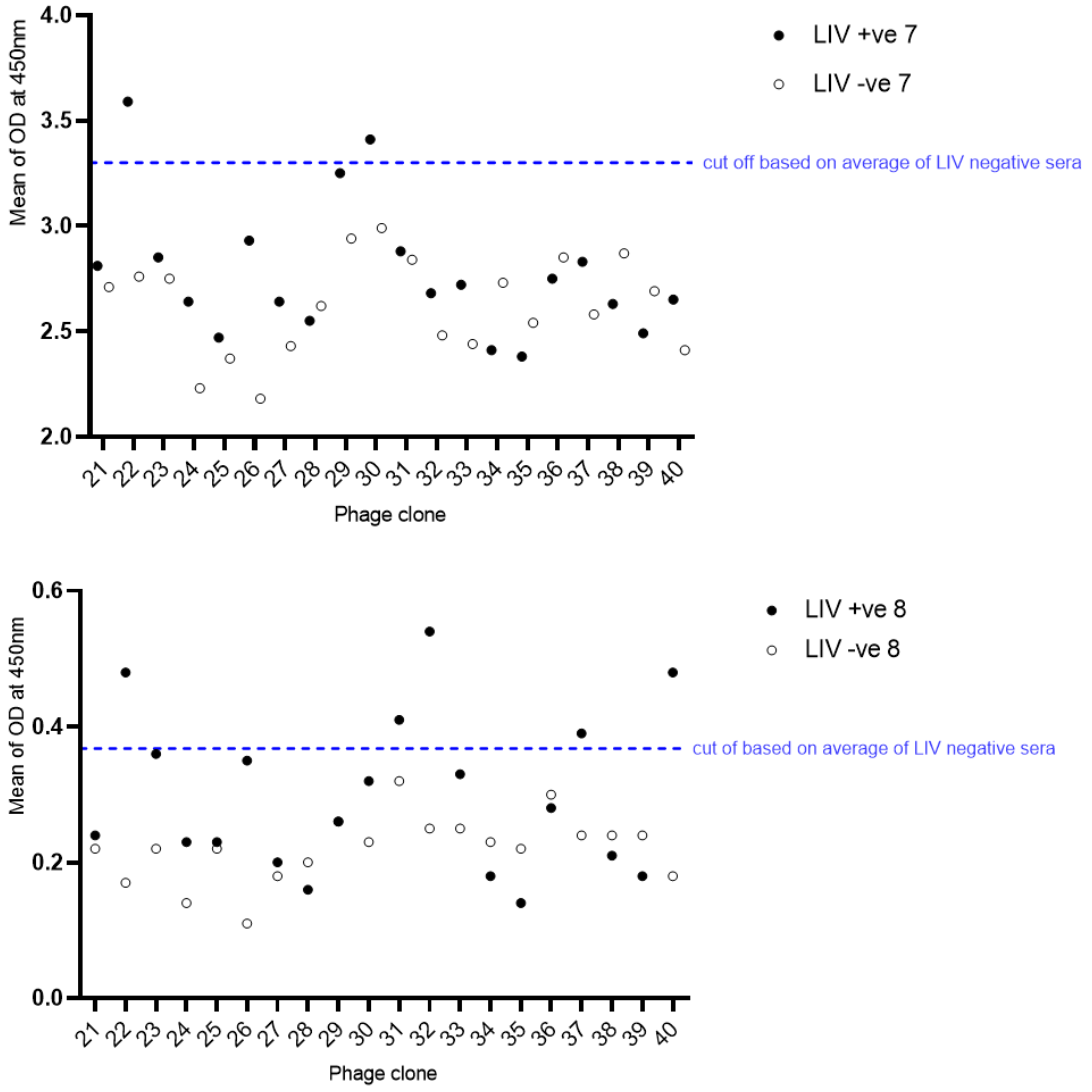


Figure 4-5: Dot plot distribution of the analysed ELISA data showing clones 21–40 tested against pooled LIV IgG positive and pooled LIV negative sera.

The change in the results between assay 7 and 8 is reflected by the change in protocol made by replacing the substrate solution from 1-Step™ Ultra TMB-ELISA Substrate Solution to 1-Step™ Slow TMB-ELISA Substrate Solution. A different cut off was applied in assays 7 and 8 where the cut off was calculated on the mean+3xSTDEV of the OD values for LIV negative serum, rather than control 1 (helper phage/LIV negative serum pool). The cut off was 3.3 for ELISA 7 and 0.37 for ELISA 8.

4.3.4 Pooled serum ELISA assay for all phage clones

Once the ELISA protocol was standardised, all the phage clones were tested in a pooled serum ELISA. Table 4-9 shows the results. The cut-off values for the assays were set on the mean OD plus three standard deviations of the LIV negative sera values for each plate, this was to factor in the individual plate variations.

The cut-off values were set high to identify the best performing clones in the pooled serum ELISA as these clones would be tested against the individual sera that made up the serum panels in the next step.

Twenty-one clones tested positive in both replicate assays and were selected to be tested with individual serum samples. Eight phage clones tested positive on the second set of ELISA but not on the first, therefore these were not taken forward, but they remain of interest.

Table 4-9: Table showing results for all 80 clones in the pooled serum ELISAs. The red shaded cells had values above the cut off and are considered positive and able to detect LIV IgG preferentially. The cut offs were calculated on the mean plus 3*SD of the LIV IgG neg serum pool.

	repeat 1			repeat 2		
Phage	LIV pos pool	LIV neg pool	cut off (mean+3*SD LIV neg pool)	LIV pos pool	LIV neg pool	cut off (mean+3*SD LIV neg pool)
Clone 1	0.22	0.09	0.24	0.2	0.07	0.19
Clone 2	0.14	0.17		0.13	0.12	
Clone 3	0.23	0.15		0.22	0.1	
Clone 4	0.22	0.11		0.22	0.06	
Clone 5	0.3	0.18		0.28	0.13	
Clone 6	0.12	0.08		0.1	0.05	
Clone 7	0.17	0.1		0.14	0.06	
Clone 8	0.27	0.17		0.25	0.14	
Clone 9	0.17	0.12		0.18	0.11	
Clone 10	0.14	0.16		0.13	0.11	
Clone 11	0.08	0.07		0.06	0.04	
Clone 12	0.23	0.12		0.21	0.07	
Clone 13	0.33	0.11		0.3	0.07	
Clone 14	0.09	0.1		0.06	0.06	
Clone 15	0.14	0.11		0.1	0.07	
Clone 16	0.27	0.13		0.28	0.14	
Clone 17	0.11	0.09		0.11	0.04	
Clone 18	0.18	0.22		0.17	0.15	
Clone 19	0.32	0.13		0.33	0.06	
Clone 20	0.25	0.1		0.23	0.06	
Clone 21	0.17	0.16	0.27	0.15	0.11	0.22
Clone 22	0.33	0.11		0.4	0.1	
Clone 23	0.25	0.14		0.25	0.11	
Clone 24	0.14	0.08		0.15	0.07	
Clone 25	0.18	0.15		0.15	0.12	
Clone 26	0.22	0.07		0.21	0.04	
Clone 27	0.11	0.09		0.09	0.06	
Clone 28	0.13	0.13		0.09	0.1	
Clone 29	0.18	0.19		0.16	0.16	
Clone 30	0.24	0.17		0.21	0.13	
Clone 31	0.28	0.21		0.31	0.17	
Clone 32	0.41	0.18		0.43	0.14	
Clone 33	0.27	0.19		0.25	0.14	
Clone 34	0.14	0.16		0.11	0.12	
Clone 35	0.09	0.14		0.07	0.11	

Clone 36	0.23	0.21		0.18	0.16	
Clone 37	0.23	0.16		0.23	0.12	
Clone 38	0.13	0.17		0.16	0.14	
Clone 39	0.1	0.17		0.08	0.13	
Clone 40	0.31	0.12		0.28	0.13	
Phage	LIV pos pool	LIV neg pool	cut off (mean+3*SD LIV neg pool)	LIV pos pool	LIV neg pool	cut off (mean+3*SD LIV neg pool)
Clone 41	0.27	0.17		0.24	0.14	
Clone 42	0.21	0.17		0.2	0.14	
Clone 43	0.15	0.17		0.13	0.12	
Clone 44	0.14	0.15		0.12	0.1	
Clone 46	0.37	0.27		0.31	0.19	
Clone 47	0.11	0.15		0.09	0.12	
Clone 48	0.21	0.15		0.19	0.1	
Clone 49	0.22	0.18		0.18	0.14	
Clone 50	0.43	0.11		0.38	0.08	
Clone 51	0.31	0.18	0.26	0.25	0.13	0.2
Clone 52	0.13	0.16		0.1	0.1	
Clone 53	0.14	0.19		0.11	0.12	
Clone 54	0.21	0.17		0.2	0.12	
Clone 55	0.33	0.13		0.28	0.08	
Clone 56	0.24	0.08		0.19	0.05	
Clone 57	0.44	0.16		0.38	0.09	
Clone 58	0.19	0.13		0.18	0.09	
Clone 59	0.16	0.14		0.12	0.09	
Clone 60	0.19	0.15		0.16	0.08	
Clone 61	0.36	0.14		0.3	0.07	
Clone 62	0.22	0.15		0.18	0.11	
Clone 63	0.14	0.14		0.1	0.1	
Clone 64*	0.25	0.17		0.18	0.12	
Clone 65	0.41	0.12		0.36	0.09	
Clone 66	0.47	0.15		0.47	0.11	
Clone 67	0.13	0.1		0.1	0.07	
Clone 68	0.08	0.13	0.24	0.05	0.07	0.18
Clone 69	0.17	0.13		0.12	0.1	
Clone 70	0.16	0.09		0.09	0.07	
Clone 71	0.4	0.16		0.37	0.14	
Clone 72	0.09	0.1		0.06	0.08	
Clone 73	0.1	0.12		0.06	0.09	
Clone 74	0.13	0.13		0.12	0.13	
Clone 75	0.09	0.1		0.07	0.08	

Clone 76	0.24	0.13		0.22	0.09	
Clone 77	0.11	0.1		0.07	0.08	
Clone 78	0.09	0.11		0.07	0.1	
Clone 79	0.14	0.11		0.12	0.1	
Clone 80	0.21	0.15		0.19	0.14	

4.3.5 Phage ELISA with individual sera from biopanning panel

The best performing 21 clones selected from the pooled sera ELISA were tested against the individual members of the serum panels. Clone 64 was missed from this selection due to human error. In the ELISA, serum samples were diluted to 1:200 dilution rather than 1:500 as the sera were not pooled.

Table 4-10 shows the ELISA results for the 21 selected clones against the individual LIV IgG positive sera and Table 4-11 shows the results with the LIV negative sera panel and the cut off value, which was 0.32. The cut off value was calculated on the mean+2xSTDEV of control 1 in LIV negative for all the phage clones. The stringency was lowered to ensure phage clones were not being discarded from the process but was still stringent enough to remove phage clones which were not LIV IgG specific.

Seven phage clones (Table 4-12) were identified as the best performing clones in the individual sera ELISA. Two of the 20 LIV positive serum samples were not detected by any of the clones tested whereas 3 LIV negative sera tested positive with more than 5 phage clones.

The seven phage clones that had the best results were taken forward for testing against a panel of sera not used in the biopanning process.

Table 4-10: Phage clones tested against LIV positive sera panel, the red boxes highlight values above the cut off

	LIV positive sera panel that made up LIV pos pool.																			
	14/0647/1	14/0724/2	14/0724/4	15/0441	15/0675	16/0383/1	16/0562	16/0704	17/0433	17/0504/1	14/0859	14/0916	14/0938/2	16/0358	16/0409/2	16/0741/2	16/0797	16/0860	17/0224	17/0319
phage 13	0.24	0.09	0.08	0.35	0.28	0.29	0.35	0.28	0.19	0.05	0.27	0.11	0.57	0.40	0.13	0.34	0.15	0.32	0.17	0.06
phage 19	0.19	0.14	0.12	0.29	0.48	0.10	0.18	0.27	0.09	0.05	0.16	0.12	0.17	0.18	0.11	0.27	0.07	0.30	0.29	0.10
phage 22	0.34	0.30	0.24	0.46	0.45	0.40	0.41	0.26	0.23	0.00	0.40	0.22	0.26	0.37	0.15	0.44	0.27	0.39	0.32	0.14
phage 32	0.23	0.17	0.15	0.29	0.51	0.11	0.23	0.26	0.15	0.07	0.17	0.12	0.16	0.21	0.07	0.23	0.19	0.25	0.25	0.10
phage 50	0.41	0.20	0.26	0.44	0.34	0.36	0.38	0.25	0.20	0.05	0.51	0.25	0.19	0.34	0.15	0.45	0.15	0.31	0.43	0.08
phage 55	0.41	0.32	0.35	0.52	0.26	0.25	0.41	0.40	0.14	0.05	0.34	0.26	0.27	0.35	0.14	0.28	0.29	0.30	0.40	0.06
phage 57	0.17	0.13	0.18	0.35	0.23	0.38	0.40	0.27	0.04	0.03	0.34	0.24	0.22	0.36	0.23	0.38	0.18	0.26	0.34	0.09
phage 61	0.40	0.10	0.13	0.42	0.26	0.27	0.42	0.27	0.57	0.07	0.50	0.44	0.18	0.49	0.30	0.39	0.29	0.25	0.37	0.11
phage 65	0.27	0.20	0.14	0.32	0.36	0.12	0.30	0.35	0.10	0.05	0.20	0.16	0.11	0.31	0.30	0.35	0.09	0.23	0.32	0.10
phage 66	0.33	0.13	0.17	0.39	0.35	0.23	0.29	0.32	0.23	0.09	0.27	0.17	0.22	0.35	0.19	0.33	0.15	0.29	0.36	0.11
phage 71	0.25	0.16	0.17	0.35	0.58	0.21	0.26	0.28	0.16	0.08	0.24	0.12	0.25	0.25	0.15	0.30	0.19	0.29	0.33	0.09
phage 5	0.08	0.26	0.31	0.37	0.34	0.13	0.33	0.25	0.15	0.05	0.20	0.08	0.08	0.14	0.17	0.23	0.15	0.25	0.23	0.03
phage 8	0.09	0.08	0.09	0.35	0.26	0.07	0.26	0.16	0.07	0.05	0.11	0.09	0.10	0.12	0.18	0.21	0.05	0.31	0.08	0.04
phage 16	0.16	0.13	0.16	0.32	0.36	0.33	0.41	0.31	0.30	0.16	0.35	0.21	0.24	0.32	0.34	0.32	0.23	0.29	0.27	0.01
phage 20	0.19	0.06	0.08	0.38	0.27	0.22	0.35	0.16	0.02	0.21	0.32	0.18	0.09	0.30	0.14	0.31	0.08	0.28	0.43	0.00
phage 31	0.11	0.10	0.12	0.31	0.20	0.22	0.30	0.26	0.16	0.07	0.25	0.14	0.12	0.20	0.21	0.27	0.09	0.22	0.13	0.03
phage 40	0.21	0.18	0.23	0.47	0.33	0.40	0.40	0.27	0.35	0.12	0.40	0.20	0.26	0.33	0.20	0.35	0.27	0.37	0.38	0.07
phage 41	0.22	0.12	0.16	0.38	0.24	0.23	0.44	0.29	0.23	0.05	0.24	0.21	0.10	0.35	0.25	0.38	0.17	0.42	0.15	0.08
phage 46	0.29	0.19	0.23	0.42	0.37	0.40	0.38	0.32	0.36	0.14	0.34	0.21	0.21	0.34	0.27	0.31	0.23	0.31	0.23	0.09
phage 51	0.21	0.11	0.15	0.47	0.23	0.14	0.24	0.26	0.17	0.05	0.18	0.15	0.19	0.34	0.12	0.22	0.10	0.46	0.27	0.06
phage 76	0.25	0.19	0.18	0.35	0.56	0.29	0.36	0.23	0.26	0.09	0.27	0.15	0.18	0.30	0.26	0.25	0.20	0.29	0.16	0.07

Table 4-11: Phage clones tested against the LIV negative sera panel, the red boxes highlight values above the cut off. The blue box shows the cut off value. LIV negative sera 2, 11 and 16 tested positive with more than 3 phage clones.

	LIV negative sera panel that made up LIV neg pool.																				
	16/0092/1	16/0092/3	16/0092/4	16/0092/5	16/0319/2	16/0319/6	16/0543/1	16/0543/7	17/0209/3	17/0209/5	16/0092/1	16/0092/3	16/0092/4	16/0092/5	16/0319/4	16/0319/4	16/0543/1	16/0543/7	17/0209/3	17/0209/5	control 1
phage 13	0.13	0.26	0.13	0.24	0.11	0.14	0.02	0.07	0.10	0.14	0.20	0.17	0.07	0.17	0.07	0.21	0.07	0.09	0.09	0.05	0.19
phage 19	0.12	0.28	0.19	0.27	0.15	0.18	0.06	0.10	0.08	0.10	0.27	0.18	0.11	0.20	0.04	0.29	0.13	0.07	0.22	0.09	0.18
phage 22	0.16	0.32	0.16	0.23	0.17	0.18	0.04	0.06	0.17	0.15	0.21	0.11	0.09	0.24	0.02	0.21	0.13	0.15	0.21	0.09	0.20
phage 32	0.18	0.32	0.25	0.30	0.20	0.24	0.10	0.13	0.11	0.17	0.33	0.18	0.20	0.26	0.08	0.33	0.06	0.06	0.31	0.17	0.19
phage 50	0.10	0.27	0.18	0.23	0.13	0.18	0.08	0.08	0.09	0.16	0.30	0.13	0.09	0.17	0.03	0.25	0.00	0.09	0.08	0.08	0.20
phage 55	0.21	0.32	0.29	0.31	0.18	0.24	0.09	0.10	0.14	0.17	0.39	0.22	0.16	0.28	0.06	0.30	0.10	0.17	0.19	0.14	0.27
phage 57	0.14	0.26	0.33	0.31	0.15	0.24	0.13	0.13	0.12	0.14	0.34	0.20	0.19	0.30	0.07	0.35	0.02	0.07	0.13	0.11	0.23
phage 61	0.20	0.25	0.25	0.28	0.17	0.24	0.12	0.10	0.14	0.12	0.32	0.18	0.14	0.29	0.08	0.29	0.08	0.08	0.15	0.12	0.21
phage 65	0.19	0.26	0.30	0.32	0.15	0.24	0.13	0.26	0.11	0.18	0.55	0.31	0.13	0.37	0.12	0.43	0.02	0.08	0.36	0.12	0.16
phage 66	0.24	0.29	0.29	0.30	0.22	0.29	0.13	0.11	0.21	0.17	0.41	0.21	0.17	0.39	0.07	0.32	0.22	0.17	0.38	0.29	0.30
phage 71	0.21	0.33	0.30	0.32	0.23	0.26	0.12	0.13	0.18	0.17	0.41	0.22	0.20	0.33	0.08	0.35	0.19	0.22	0.35	0.15	0.30
phage 5	0.04	0.20	0.10	0.14	0.13	0.06	0.07	0.13	-0.04	0.01	0.22	0.06	0.04	0.10	0.03	0.24	-0.01	0.18	0.16	0.03	0.09
phage 8	0.05	0.18	0.11	0.17	0.19	0.09	0.04	0.06	-0.01	0.02	0.23	0.08	0.03	0.13	0.02	0.25	-0.01	0.06	0.07	0.05	0.13
phage 16	0.11	0.30	0.19	0.25	0.16	0.21	0.26	0.33	0.10	0.16	0.25	0.16	0.14	0.21	0.17	0.27	0.19	0.21	0.19	0.13	0.12
phage 20	0.03	0.19	0.10	0.16	0.14	0.03	0.01	0.03	0.13	0.07	0.22	0.05	0.07	0.02	0.00	0.14	0.01	0.07	0.04	0.10	0.08
phage 31	0.03	0.28	0.16	0.24	0.12	0.11	0.10	0.17	-0.01	0.05	0.29	0.13	0.10	0.16	0.06	0.26	0.08	0.09	0.09	0.05	0.09
phage 40	0.13	0.31	0.19	0.24	0.16	0.18	0.06	0.08	0.06	0.08	0.31	0.15	0.14	0.20	0.07	0.26	0.05	0.07	0.09	0.08	0.18
phage 41	0.15	0.35	0.27	0.26	0.19	0.18	0.10	0.18	0.01	0.09	0.36	0.19	0.16	0.23	0.08	0.31	0.02	0.10	0.13	0.19	0.20
phage 46	0.21	0.41	0.28	0.34	0.24	0.25	0.15	0.22	0.18	0.13	0.37	0.22	0.25	0.30	0.12	0.30	0.30	0.13	0.15	0.11	0.25
phage 51	0.12	0.30	0.18	0.23	0.17	0.16	0.05	0.08	0.07	0.07	0.30	0.14	0.13	0.19	0.05	0.27	0.04	0.07	0.10	0.08	0.21
phage 76	0.23	0.37	0.23	0.27	0.23	0.23	0.09	0.16	0.19	0.10	0.32	0.18	0.18	0.24	0.06	0.30	0.11	0.09	0.11	0.08	0.22

Avg (Con 1) LIV neg (phage 13-76)	0.19
STDEV	0.06
2*STDEV	0.12
cut off	**0.32**

Table 4-12: Phage clones tested against the individual sera panel; the number of sera that tested positive are shown in the LIV positive column and the false positive in the LIV negative column.

The peptides in red detected more LIV IgG positive sera than negative sera which were subsequently tested with the independent panel of sera.

Phage clone	Peptide sequence	LIV positive sera	LIV negative sera
13	FNAIDHKVHAGPSPSV	6	0
19	LNMQYHPADTLRSAFT	1	0
22	NWPTRANTPYNAPQFV	10	1
32	KYHPRDRVTTTPTSNT	1	3
50	LKSTHSHNWPRSMNMS	9	0
55	GKYAARRTATNTSELR	9	2
57	SMPSKNWPTPMGKCAP	7	3
61	KKTIVTFAKGPPDARA	9	1
65	YMPLDTLRNSTTSGPF	5	5
66	YWPNDVVNYTSSSLAN	7	4
71	AYYFPIDALPDPIPAP	3	6
5	QKNQGLASFAKPVIHT	3	0
8	AYFAETISIDQHSTTP	1	0
16	YRSSWEKIAPSSTPSS	8	1
20	SKFVYSLKAKTDSHID	4	0
31	SHWDKPYIPQMIPIDN	0	0
40	NWPSAVNKPTAERLNA	10	0
41	KKTNSGTYTATEKPQP	5	2
46	SHWDKPYIPQMIPIDN	8	3
51	GVFAASTLHQSNNTGI	3	0
76	PGPMSNWNKMESANAS	3	2

4.3.6 Phage ELISA with independent sera panels

4.3.6.1 Individual clone performance in ELISA

The 7 phage clones that showed the best discrimination between LIV positive and LIV negative sera (Table 4-12) were tested on the phage ELISA against an independent panel of 20 LIV IgG positive and 20 LIV negative sera. The results are shown in Table 4-13. The positive/negative cut off for the ELISA was set on control 1 (M13KO7 phage tested with pooled LIV negative sera) The data for each phage clone was plotted as a dot plot and receiver operating characteristic curves (ROC) were generated to show the result for each clone against the positive and negative sera (Figure 4-6—Figure 4-7).

Sensitivity and specificity of each clone was calculated (Table 4-14) and clone 50 had the best sensitivity and specificity values (90%/85%) of all the phage clones. Clones 13 and 16 had poor sensitivity but high specificity values. Clones 22 and 40 had sensitivity values of 85% but slightly lower specificity (65%) and clones 55 and 61 had lower sensitivity (55%) but high specificity (90 and 85%). The p-value for each clone was calculated from the ROC analysis and all peptides had p value <0.05. Clones 22, 40 and 50 had the lowest p values which indicate that they had the best sensitivity/specificity profile in detecting LIV IG+ sera over LIV IgG -ve sera among all the clones tested.

Table 4-13: Phage ELISA testing the clones with LIV positive sera and LIV negative sera not used in the initial biopanning. The positive/negative cut off values were based on mean plus 3xSD of control 1 (yellow boxes) for each assay.
The green boxes indicate a positive result. Control 1 was represented by the helper phage coated and tested with pooled LIV negative sera at 1:500 dilution.

phage clone 13		phage clone 16		phage clone 22		phage clone 40		phage clone 50		phage clone 55		phage clone 61	
LIV +ve sera	LIV - sera	LIV +ve sera	LIV - sera	LIV +ve sera	LIV - sera	LIV +ve sera	LIV - sera	LIV +ve sera	LIV - sera	LIV +ve sera	LIV - sera	LIV +ve sera	LIV - sera
0.084	-0.010	0.200	0.196	**0.220**	0.034	**0.323**	**0.147**	**0.193**	-0.013	0.068	-0.078	**0.108**	-0.007
0.018	0.053	0.104	0.141	0.025	0.053	0.074	0.053	0.051	0.017	0.015	0.018	0.006	-0.009
0.075	0.032	0.195	0.139	**0.304**	0.111	**0.283**	0.129	**0.256**	0.056	**0.181**	0.076	**0.114**	0.007
0.138	0.055	0.144	0.128	**0.163**	**0.128**	**0.207**	0.126	**0.165**	0.055	**0.284**	0.098	**0.103**	0.020
0.155	0.077	**0.230**	0.107	**0.244**	0.047	**0.329**	0.097	**0.263**	0.004	**0.147**	-0.005	**0.147**	0.012
0.118	0.065	0.195	0.118	**0.158**	0.055	**0.226**	**0.135**	**0.145**	0.024	0.084	0.024	0.088	0.018
0.098	0.089	0.155	0.117	**0.130**	0.050	0.117	0.125	**0.176**	0.049	0.040	0.029	0.017	0.037
0.088	0.060	**0.216**	0.146	**0.305**	0.084	**0.302**	0.100	**0.280**	0.064	**0.268**	0.072	**0.155**	0.031
0.133	0.017	**0.259**	**0.229**	**0.255**	**0.148**	**0.321**	**0.203**	**0.227**	0.058	**0.139**	-0.003	**0.184**	0.016
0.073	0.041	**0.223**	0.110	**0.211**	**0.131**	**0.243**	0.087	**0.235**	0.041	**0.191**	0.103	**0.191**	0.007
0.066	0.103	0.155	0.115	**0.303**	0.072	**0.280**	0.101	**0.253**	0.058	**0.155**	**0.124**	**0.186**	0.040
0.104	0.054	0.177	0.098	**0.120**	0.105	**0.182**	**0.144**	**0.101**	0.026	0.093	0.036	0.059	-0.012
0.069	**0.175**	0.148	**0.257**	0.071	**0.158**	**0.160**	**0.175**	0.037	**0.121**	0.047	0.092	0.053	**0.102**
0.105	0.121	0.180	0.191	**0.241**	**0.123**	**0.281**	**0.162**	**0.239**	0.085	**0.172**	0.056	0.082	**0.100**
0.030	**0.191**	**0.205**	**0.216**	0.094	**0.166**	0.132	**0.225**	**0.099**	**0.176**	0.088	**0.168**	0.079	**0.188**
0.086	0.124	0.090	0.148	**0.263**	**0.125**	**0.254**	**0.138**	**0.230**	**0.128**	0.097	0.112	**0.111**	0.081
0.125	-0.031	**0.201**	0.125	**0.187**	0.024	**0.267**	0.040	**0.161**	-0.043	0.062	-0.067	**0.140**	-0.015
0.072	-0.019	0.149	0.142	**0.137**	-0.011	**0.162**	0.023	**0.133**	-0.031	**0.272**	-0.006	-0.001	0.009
0.134	0.018	0.143	0.102	**0.232**	0.006	**0.245**	0.049	**0.213**	-0.004	**0.255**	0.011	0.067	-0.006
0.083	0.059	0.134	0.092	**0.250**	0.025	**0.210**	0.062	**0.182**	0.018	**0.248**	0.044	**0.109**	0.005
0.13		**0.20**		**0.11**		**0.13**		**0.09**		**0.12**		**0.09**	

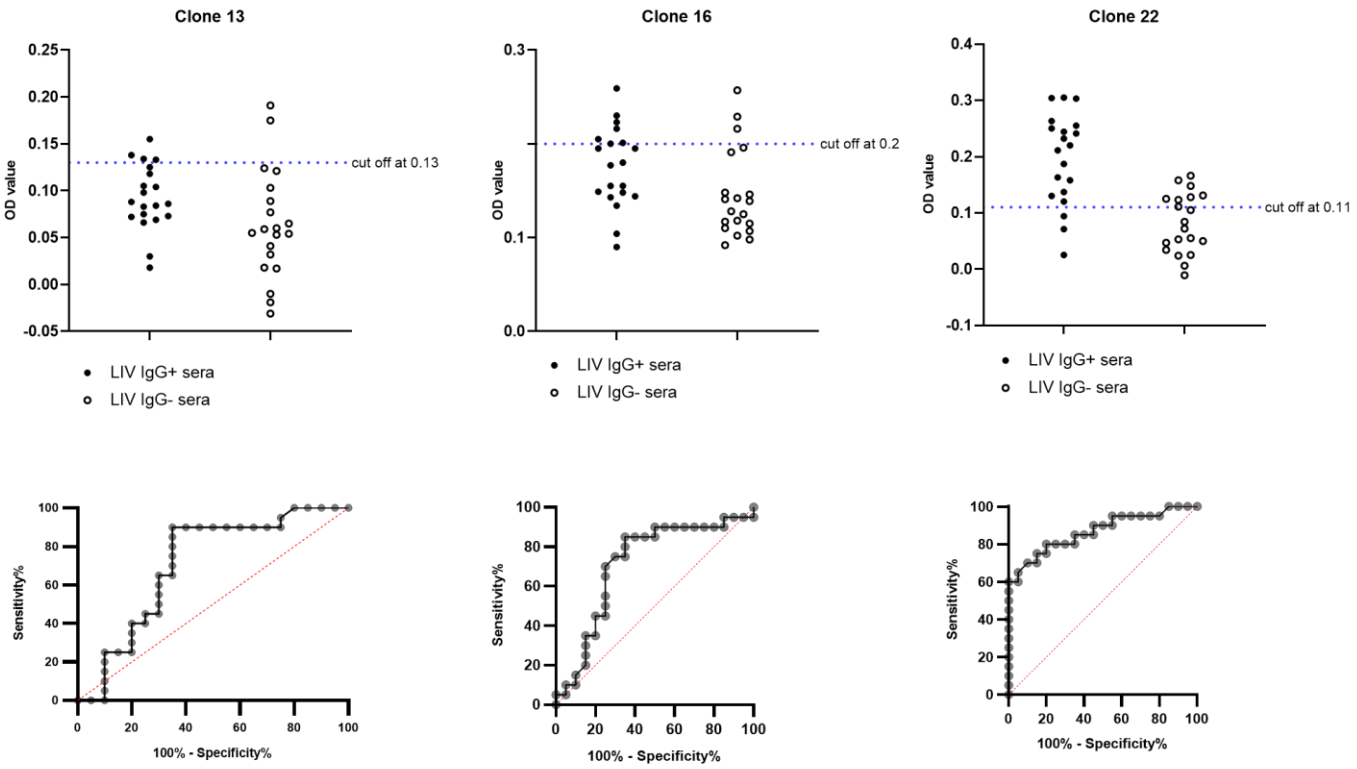


Figure 4-6: Dot plots and ROC graphs of phage clones 13, 16 and 22 tested against individual sera panels.
The filled dots indicating LIV IgG +ve sera and the unfilled dots indicating the LIV IgG-ve sera. The blue line indicates the positive/negative cut off value set. on control 1. The cut offs were calculated on 3xSD+mean of control 1 (M13KO7 helper phage plus pooled LIV-ve sera). On the y axis, is the average OD values. The phage clones were tested against the same sera panels and the dot plot for phage clone 22 shows that it was able to distinguish the most LIV+ve sera out of all three.

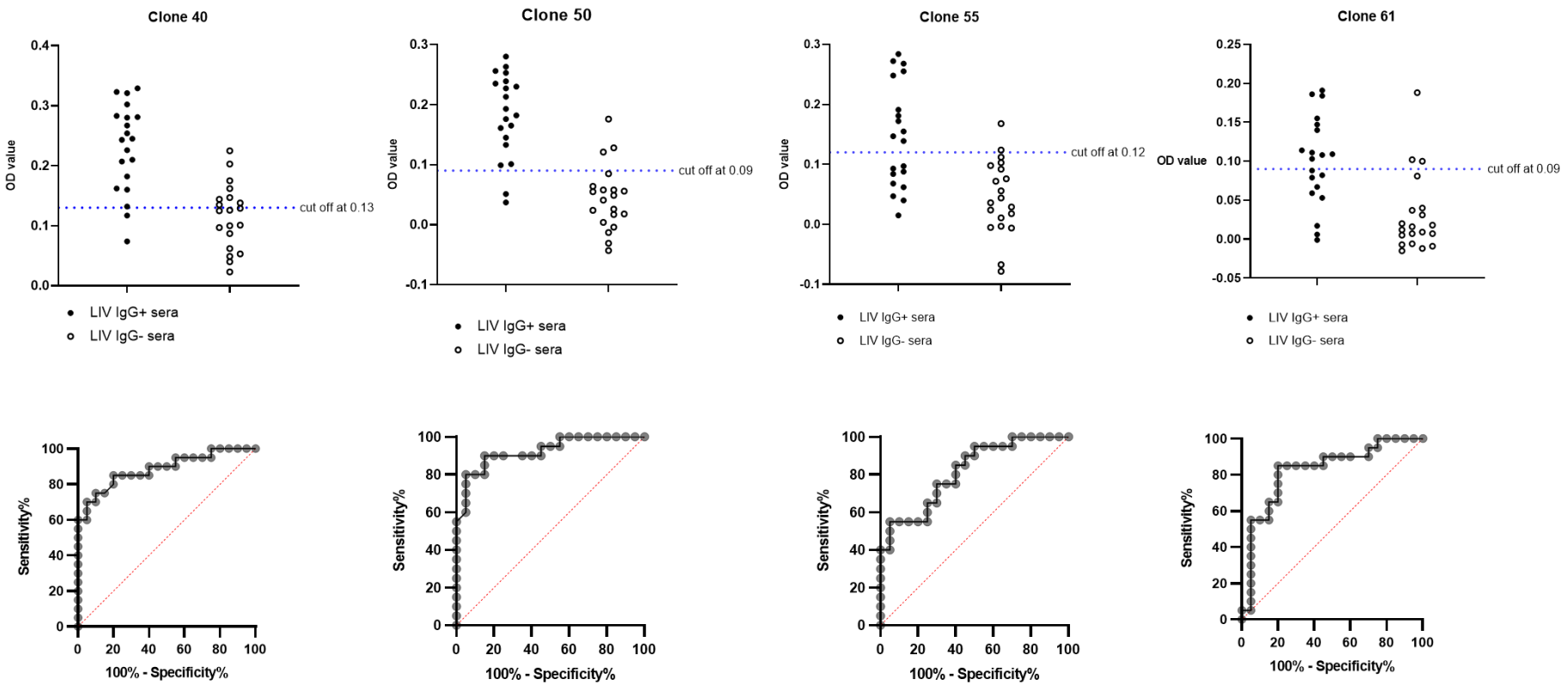


Figure 4-7: Dot plots and ROC curves of phage clones 40, 50, 55 and 61 tested against individual sera.
The filled dots indicating LIV IgG +ve sera and the unfilled dots indicating the LIV IgG-ve sera. The blue line indicates the positive/negative cut off value set on control 1 for each assay. The cut offs were calculated on 3xSD+mean of control 1 (M13KO7 helper phage plus pooled LIV-ve sera). On the y axis, is the mean OD values. The phage clones were tested against the same sera from the independent sera panel.

Table 4-14: This table shows the ELISA results of phage clones with sera (20 LIV IgG+ and 20 LIV IgG-) not used in the biopanning process (independent sera).

Similarity between the peptide sequences are highlighted in blue. The third column in the table shows the number of LIV IgG + sera detected by each clone (true positive) and the fourth column shows the number of LIV IgG-ve serum samples detected (false positive). These values were used to calculate the sensitivity in the fifth column and specificity values in the sixth column. The values in the seventh and eight columns are statistical analysis values generated from ROC analysis of the data.

Phage clone	Peptide sequence	LIV positive sera detected	LIV negative sera detected	sensitivity	specificity	AUC values of ROC plots	P value for ROC plots
13	FNAIDHKVHAGPSPSV	4	2	20	90	0.7	0.0239
16	YRSSWEKIAPSSTPSS	6	3	30	85	0.71	0.02
22	NWPTRANTPYNAPQFV	17	7	85	65	0.86	<0.0001
40	NWPSAVNKPTAERLNA	18	8	85	60	0.88	<0.0001
50	LKSTHSHNWPRSMNMS	18	3	90	85	0.92	<0.0001
55	GKYAARRTATNTSELR	11	2	55	90	0.81	0.0007
61	KKTIVTFAKGPPDARA	11	3	55	85	0.83	0.0004

4.3.6.2 Analysing the phage ELISA data for peptides as a panel

The positive/negative cut off value calculated on 2xSTDEV plus mean of control 1 for LIV negative sera and this was 0.135, as seen in Table 4-15. When a stricter cut off based on mean plus 3xSD for the same control (0.150) was applied, as seen in Table 4-16, this resulted in a reduction in false positives.

Table 4-15: Phage ELISA testing the clones with LIV positive sera and LIV negative sera not used in the initial biopanning.

The positive/negative cut off values were based on mean plus 2xSD of control1 with LIV negative sera from all 7 assays. Control 1 was represented by the helper phage coated and tested with pooled LIV negative sera at 1:500 dilution.

	LIV IgG+ve sera not used in biopanning																			
	LIV +21	LIV +22	LIV +23	LIV +24	LIV +25	LIV +26	LIV +27	LIV +28	LIV +29	LIV +30	LIV +31	LIV +32	LIV +33	LIV +34	LIV +35	LIV +36	LIV +37	LIV +38	LIV +39	LIV +40
phage 13	0.084	0.018	0.075	0.138	0.155	0.118	0.098	0.088	0.133	0.073	0.066	0.104	0.069	0.105	0.030	0.086	0.125	0.072	0.134	0.083
phage 16	0.200	0.104	0.195	0.144	0.230	0.195	0.155	0.216	0.259	0.223	0.155	0.177	0.148	0.180	0.205	0.090	0.201	0.149	0.143	0.134
phage 22	0.220	0.025	0.304	0.163	0.244	0.158	0.130	0.305	0.255	0.211	0.303	0.120	0.071	0.241	0.094	0.263	0.187	0.137	0.232	0.250
phage 40	0.323	0.074	0.283	0.207	0.329	0.226	0.117	0.302	0.321	0.243	0.280	0.182	0.160	0.281	0.132	0.254	0.267	0.162	0.245	0.210
phage 50	0.193	0.051	0.256	0.165	0.263	0.145	0.176	0.280	0.227	0.235	0.253	0.101	0.037	0.239	0.099	0.230	0.161	0.133	0.213	0.182
phage 55	0.068	0.015	0.181	0.284	0.147	0.084	0.040	0.268	0.139	0.191	0.155	0.093	0.047	0.172	0.088	0.097	0.062	0.272	0.255	0.248
phage 61	0.108	0.006	0.114	0.103	0.147	0.088	0.017	0.155	0.184	0.191	0.186	0.059	0.053	0.082	0.079	0.111	0.140	-0.001	0.067	0.109

	LIV IgG -ve sera not used in biopanning																			
	LIV -21	LIV -22	LIV -23	LIV -24	LIV -25	LIV -26	LIV -27	LIV -28	LIV -29	LIV -30	LIV -31	LIV -32	LIV -33	LIV -34	LIV -35	LIV -36	LIV -37	LIV -38	LIV -39	LIV -40
phage 13	-0.010	0.053	0.032	0.055	0.077	0.065	0.089	0.060	0.017	0.041	0.103	0.054	0.175	0.121	0.191	0.124	-0.031	-0.019	0.018	0.059
phage 16	0.196	0.141	0.139	0.128	0.107	0.118	0.117	0.146	0.229	0.110	0.115	0.098	0.257	0.191	0.216	0.148	0.125	0.142	0.102	0.092
phage 22	0.034	0.053	0.111	0.128	0.047	0.055	0.050	0.084	0.148	0.131	0.072	0.105	0.158	0.123	0.166	0.125	0.024	-0.011	0.006	0.025
phage 40	0.147	0.053	0.129	0.126	0.097	0.135	0.125	0.100	0.203	0.087	0.101	0.144	0.175	0.162	0.225	0.138	0.040	0.023	0.049	0.062
phage 50	-0.013	0.017	0.056	0.055	0.004	0.024	0.049	0.064	0.058	0.041	0.058	0.026	0.121	0.085	0.176	0.128	-0.043	-0.031	-0.004	0.018
phage 55	-0.078	0.018	0.076	0.098	-0.005	0.024	0.029	0.072	-0.003	0.103	0.124	0.036	0.092	0.056	0.168	0.112	-0.067	-0.006	0.011	0.044
phage 61	-0.007	-0.009	0.007	0.020	0.012	0.018	0.037	0.031	0.016	0.007	0.040	-0.012	0.102	0.100	0.188	0.081	-0.015	0.009	-0.006	0.005

LIV negative Control 1 average OD	0.102
STDEV	0.017
2*STDEV	0.033
Cut off (Avg+2*STDEV)	**0.135**

Table 4-16: Phage ELISA testing the clones with LIV positive sera and LIV negative sera not used in the initial biopanning.

The positive/negative cut off values were based on mean plus 3SD of control 1 with LIV negative sera from all 7 assays. Control 1 was represented by the helper phage coated and tested with pooled LIV negative sera at 1:500 dilution.

	LIV IgG +ve sera not used in biopanning																			
	LIV +21	LIV +22	LIV +23	LIV +24	LIV +25	LIV +26	LIV +27	LIV +28	LIV +29	LIV +30	LIV +31	LIV +32	LIV +33	LIV +34	LIV +35	LIV +36	LIV +37	LIV +38	LIV +39	LIV +40
phage 13	0.084	0.018	0.075	0.138	0.155	0.118	0.098	0.088	0.133	0.073	0.066	0.104	0.069	0.105	0.030	0.086	0.125	0.072	0.134	0.083
phage 16	0.200	0.104	0.195	0.144	0.230	0.195	0.155	0.216	0.259	0.223	0.155	0.177	0.148	0.180	0.205	0.090	0.201	0.149	0.143	0.134
phage 22	0.220	0.025	0.304	0.163	0.244	0.158	0.130	0.305	0.255	0.211	0.303	0.120	0.071	0.241	0.094	0.263	0.187	0.137	0.232	0.250
phage 40	0.323	0.074	0.283	0.207	0.329	0.226	0.117	0.302	0.321	0.243	0.280	0.182	0.160	0.281	0.132	0.254	0.267	0.162	0.245	0.210
phage 50	0.193	0.051	0.256	0.165	0.263	0.145	0.176	0.280	0.227	0.235	0.253	0.101	0.037	0.239	0.099	0.230	0.161	0.133	0.213	0.182
phage 55	0.068	0.015	0.181	0.284	0.147	0.084	0.040	0.268	0.139	0.191	0.155	0.093	0.047	0.172	0.088	0.097	0.062	0.272	0.255	0.248
phage 61	0.108	0.006	0.114	0.103	0.147	0.088	0.017	0.155	0.184	0.191	0.186	0.059	0.053	0.082	0.079	0.111	0.140	-0.001	0.067	0.109

	LIV IgG -ve sera not used in biopanning																			
	LIV -21	LIV -22	LIV -23	LIV -24	LIV -25	LIV -26	LIV -27	LIV -28	LIV -29	LIV -30	LIV -31	LIV -32	LIV -33	LIV -34	LIV -35	LIV -36	LIV -37	LIV -38	LIV -39	LIV -40
phage 13	-0.010	0.053	0.032	0.055	0.077	0.065	0.089	0.060	0.017	0.041	0.103	0.054	0.175	0.121	0.191	0.124	-0.031	-0.019	0.018	0.059
phage 16	0.196	0.141	0.139	0.128	0.107	0.118	0.117	0.146	0.229	0.110	0.115	0.098	0.257	0.191	0.216	0.148	0.125	0.142	0.102	0.092
phage 22	0.034	0.053	0.111	0.128	0.047	0.055	0.050	0.084	0.148	0.131	0.072	0.105	0.158	0.123	0.166	0.125	0.024	-0.011	0.006	0.025
phage 40	0.147	0.053	0.129	0.126	0.097	0.135	0.125	0.100	0.203	0.087	0.101	0.144	0.175	0.162	0.225	0.138	0.040	0.023	0.049	0.062
phage 50	-0.013	0.017	0.056	0.055	0.004	0.024	0.049	0.064	0.058	0.041	0.058	0.026	0.121	0.085	0.176	0.128	-0.043	-0.031	-0.004	0.018
phage 55	-0.078	0.018	0.076	0.098	-0.005	0.024	0.029	0.072	-0.003	0.103	0.124	0.036	0.092	0.056	0.168	0.112	-0.067	-0.006	0.011	0.044
phage 61	-0.007	-0.009	0.007	0.020	0.012	0.018	0.037	0.031	0.016	0.007	0.040	-0.012	0.102	0.100	0.188	0.081	-0.015	0.009	-0.006	0.005

LIV negative Control 1 average OD	0.10
STDEV of LIV -ve	0.02
3*STDEV	0.05
avg+3*STDEV	**0.15**

4.3.6.3 Phage ELISA sensitivity and specificity

The sensitivity and specificity scores were calculated for the lower and higher cut off values.

Table 4-17: Formula used to calculate sensitivity and specificity scores.

	LIV pos sera (number)	LIV neg sera (number)
Positive result (Number)	A (true pos)	B (false pos)
Negative result (number)	C (false neg)	D (true neg)
sensitivity	A/(A+C)*100	
specificity	D/(D+B)*100	

Lower cut off

Table 4-15 shows the data for the lower cut off values. Sensitivity and specificity (formula used shown in Table 4-17) calculations for the clones are shown in Table 4-18. In the lower cut off, individual phage clones reacted with at least two LIV positive samples and as a panel, they detected 19 out of 20 LIV positive sera.

Clones 16 and 40 detected 17 of the 20 sera and therefore had the highest sensitivity (85%). Three clones (50, 55 and 61) had 95% specificity, and if the results of clones 50 and 55 were combined, a sensitivity of 80% and specificity of 95% is achieved.

Looking at the negative sera results at the lower cut off, there were 11 false positives with clone 16 and 40 contributing to most of these. These 2 clones therefore had the highest sensitivity (85%) and the lowest specificity (50% and 65%). Serum sample 35 in the LIV negative panel tested positive for all the clones.

Table 4-18: This table shows the sensitivity and specificity of each phage clone tested with independent sera panel at the lower cut off.

Clone	Peptide sequence	True positives detected	False positives detected	Sensitivity	Specificity
13	FNAIDHKVHAGPSPSV	2	2	10	90
16	YRSSWEKIAPSSTPSS	17	10	85	50
22	NWPTRANTPYNAPQFV	15	3	75	85
40	NWPSAVNKPTAERLNA	17	7	85	65
50	LKSTHSHNWPRSMNMS	15	1	75	95
55	GKYAARRTATNTSELR	11	1	55	95
61	KKTIVTFAKGPPDARA	6	1	30	95

Higher cut off

Table 4-16 shows the higher cut off data. The sensitivity and specificity scores are shown in Table 4-19. Looking at the LIV positive serum panel results, all 7 clones reacted with at least one LIV positive sample. The combined data for the LIV positive sera shows that all the clones detected 19 out of 20 positive LIV sera. Clone 40 reacted with 17 of the 20 sera and had the highest sensitivity (85%) the next best performing clones were 22 and 50 which detected 14 of the 20 sera and the sensitivity was calculated to be 70%.

Looking at the LIV negative serum panel, there were only 5 false positives compared to the 11/20 seen when the cut off value was lower. The specificity value of clones 16 and 40 increased to 75% and 80% compared to what it was at the lower cut off. All 7 phage clones reacted positively with LIV IgG-ve 35. Clones 16, 40 and 50 have a combined sensitivity of 95% and specificity of 75% at the higher cut off.

Table 4-19: Sensitivity and specificity of the clones tested with independent sera panel at the higher cut off.

Clone	Peptide sequence	True positives detected	False positives detected	Sensitivity	Specificity
13	FNAIDHKVHAGPSPSV	1	2	5	90
16	YRSSWEKIAPSSTPSS	13	5	65	75
22	NWPTRANTPYNAPQFV	14	2	70	90
40	NWPSAVNKPTAERLNA	17	4	85	80
50	LKSTHSHNWPRSMNMS	14	1	70	95
55	GKYAARRTATNTSELR	9	1	45	95
61	KKTIVTFAKGPPDARA	4	1	20	95

4.3.7 Ion Torrent sequencing

The 20 sub-libraries (LIV+1—10, LIV-11—20) from round 3 biopanning were deep sequenced on the Ion Torrent platform to identify the phage clones being preferentially enriched towards LIV IgG. The sequencing data was analysed to determine the presence and percent frequency of the 7 phage clones tested with the independent sera panel and a heatmap of this data,

Figure 4-8, was generated.

Four out of the seven clones tested in the independent sera ELISA were enriched enough to be present in the top 250 peptides arranged by peptide frequency in 7 LIV positive sub-libraries. Clones 13, 40, 55 and 61 combined were seen in seven out of the ten LIV positive sub libraries. Two clones, 40 and 61 combined were present in four of the ten LIV negative sub libraries.

Clones 16, 22 and 50, all which were effective in ELISA, were not identified in the top 250 peptides in either the positive or negative sub-libraries. This could be because these are rare clones which have a phenotypic advantage as they grew on culture, were picked, grown, and tested on ELISA. They are potentially in the NGS sub library at a much lower frequency than the filter used in this analysis. In total four of the seven clones had been enriched in the data set.

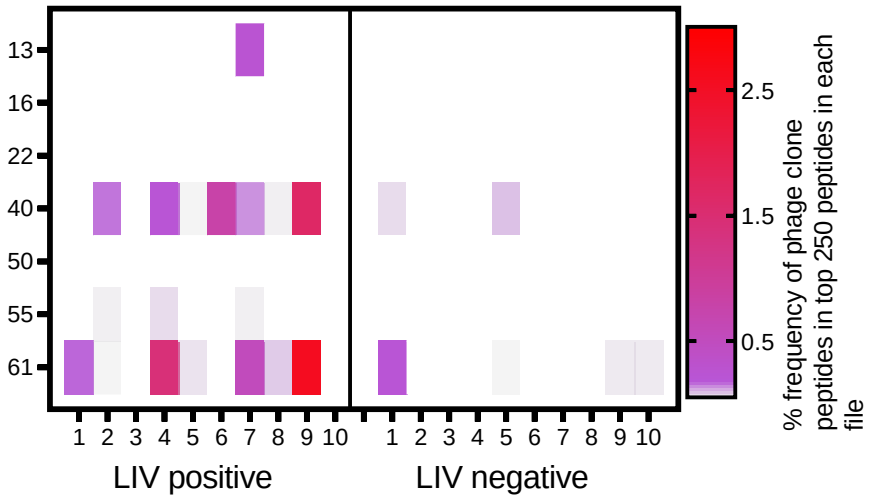


Figure 4-8: Heatmap showing the % peptide frequency of the 7 phage clones (tested with the independent sera panel in ELISA) compared to the NGS data of the top 250 peptides by percent frequency in the round 3 sub-libraries.

The sum percent peptide frequencies for each peptide (Table 4-20) were calculated from data used to generate the heatmap. Clones 40 and 61 were enriched more than any of the other peptides. There were similarities in the peptide sequences (NWP, NTP, NKP), and these have been highlighted in the table.

Table 4-20:This table shows the peptides from the clone pick ELISA and the sum of peptide frequencies in the top 250 peptides in the NGS sub-libraries (LIV pos 1-10 and LIV neg 1-10). The yellow shaded rows reflect the peptides that were seen in the top 250 peptides in the NGS data set.

Clone	Peptide	Sum of % peptide frequencies in positive sub-libraries	Sum of % peptide frequencies in negative sub-libraries
13	FNAIDHKVHAGPSPSV	0.26	0.00
16	YRSSWEKIAPSSTPSS	0.00	0.00
22	NWPTRANTPYNAPQFV	0.00	0.00
40	NWPSAVNKPTAERLNA	3.16	0.25
50	LKSTHSHNWPRSMNMS	0.00	0.00
55	GKYAARRTATNTSELR	0.21	0.00
61	KKTIVTFAKGPPDARA	4.98	0.41

4.4 Discussion

The objectives of this study were to biopan a random peptide phage library with ovine sera positive for LIV IgG to enrich for phage mimotopes to LIV; pick phage clones and test these in ELISA to establish their ability to distinguish between LIV+ and LIV-ve sera.

4.4.1 Biopanning

There is an absence of literature where biopanning and phage display have been used for LIV epitope mapping. The biopanning process in this experiment was straightforward as it was enriching for mimotopes bound by serum antibodies from a set of samples that were positive for the target viral infection whereas the control set were negative for the target viral infection.

Subtractive biopanning was carried out to reduce non-specific binders to ovine polyclonal serum IgG. The phage output titres calculated from the pooled LIV IgG positive sub-libraries indicate that there was an enrichment compared to the control. The volume of serum used in the biopanning rounds was higher (5 µL) than in the previous study reported in this thesis. This was done as an improvement on the method in order to maximise enrichment of peptide mimotopes for LIV IgG.

4.4.2 ELISA method development

As mentioned above, the approach used to identify peptide mimotopes to LIV which were then used to develop a diagnostic assay has not been reported before, but phage display and biopanning have been applied to identify peptides with diagnostic capability for other pathogens, such as HIV (Zhang *et al.*, 2012) and bacterial pathogens such as *Salmonella typhimurium* (Agrawal et al., 2016).

A process of method development and finessing was carried out in this study to develop a phage ELISA to detect ovine sera positive for LIV IgG.

The finalised ELISA method was used to test the phage clones with sera used in biopanning experiments. Initially the phage clones were tested against pooled sera: LIV IgG +ve 1–20 and LIV IgG-ve 1–20. The positive/negative cut off values were set on the double negative control wells (Control 1) in the assay which was wells coated with the helper phage M13KO and tested with pooled LIV IgG -ve sera 1–20. The value of the positive/negative cut off was the mean OD plus 3xSTDEV of Control 1.

Clones that tested positive in the replicate assays were taken forward for testing against individual members of the pooled serum samples. The mean of the control 1 values of the replicate assays was used as the final cut off value. This led to the loss of eight clones from the pipeline as seven clones tested positive only in one set of data and one (clone 64) was missed out due to human error even though it had tested positive in both sets of data. Ideally these clones would be tested in the future to see whether less stringent cut off values at the pooled ELISA stage is better than removing them at such an early stage. The cut off values ranged between 0.18 and 0.27 and an improvement on the selection criterion would have been to calculate the mean cut off value from all 8 ELISA plates (0.22).

Phage clones tested with individual members of serum pools.

The phage clones (n=21) that progressed through the pooled serum ELISA were tested with the individual members that formed the serum pools. The positive/negative cut off value was set on control 1 (M13KO7/pooled LIV sera 1—20). The individual OD values for control 1 ranged from 0.08 to 0.3 and to standardise the cut off values, the mean of all 21 Control 1 values was calculated. The common cut off value was mean plus 2xSD of control 1 for all 21 clones (0.32). The cut off was set high enough to reduce phage with high false positive results. There were three serum samples (LIV IgG-ve 2, 11 and 16) which tested false positive by more than 3 phage clones.

The clones which had the best true positive results, and the lowest false positive (n=7) results were taken forward for testing with independent sera panel comprised of 20 LIV IgG positive and 20 LIV IgG negative sera that had not been used in the biopanning process.

Phage clones tested with independent sera.

The peptides that were tested in ELISA against the independent sera panel were analysed for their individual ability (singleplex) to distinguish between LIV positive and negative sera as well as the combined performance of the peptides as a panel with the best sensitivity and specificity. One of the LIV -ve samples (LIV-35) was detected as positive by all 7 clones.

For the individual analysis, 3xSD plus mean of control 1 was calculated for all the clones. Looking at the specificity and sensitivity values from the ELISA results, the clones were either very specific (90%) with a lower sensitivity (20%) or highly sensitive (90%) with a lower specificity (85%). The best singleplex clone in the series, with the best overall sensitivity (90%) and specificity (85%) values was clone 50.

Statistical analysis of the individual clone's performance was evaluated by the AUC values generated following ROC analysis. ROC analysis was chosen as it is the commonly chosen statistical analysis for evaluating diagnostic accuracy of assays. An assay with very high sensitivity and good specificity would have an AUC value close to 1 and a value of 0.5 suggests that there is no discrimination between the positive and negative sample and a value of 0 means it is completely inaccurate. Clone 50 had the highest AUC value (0.92) which falls under the 'outstanding' region (Mandrekar, 2010). Clone 40 (AUC 0.88), clone 22 (AUC 0.86), clone 61 (AUC0.83) and clone 55 (AUC 0.81) were the next best and clones 13 (AUC 0.70) and 16 (AUC 0.71) scored the least. This tallies with the sensitivity and specificity values calculated using the cut off values.

The specificity and sensitivity of the ELISA could be improved by using a multi-peptide approach as it would cover more of the antigenic regions that the mimotopes were enriched to. The ELISA data from the singleplex assays were collated into a table and the sensitivity/specificity values were calculated based on a lower and higher cut-off values set on control 1. Combining the results of specific clones had an impact on the sensitivity and specificity profiles at the higher cut off. Combining clones 16, 40, 50 and 55 and masking the results of LIV negative serum 35, the sensitivity remained at 95% but the specificity improved to 78%. Another combination of peptides, also masking LIV neg 35, clones 22, 40, 50, 55 and 61 gave a better combined specificity at 91% but the sensitivity dropped to 84%.

The final assay will have to consider whether the goal of this ELISA is to be used in clinic as a diagnostic tool or whether it will be a research tool to study seroprevalence of LIV IgG. A diagnostic assay will require high sensitivity and high specificity especially as other tick-borne flaviviruses might also be circulating in the same geographic locations. If the assay is to be used as a research tool to understand the seroprevalence of LIV IgG in flocks across the country, higher sensitivity might be preferred over a very specific assay. In this instance, the multi-peptide assay at a higher cut off would be preferred as it currently has the best profile.

The phage clones have demonstrated that they have the potential to be used as singleplex or multiplex coating antigen in ELISA. As all 7 clones reacted with sample 35 from the LIV negative panel, this sample will be investigated further with an existing LIV serology assay. The peptides reacted positively for this serum because they are mimotopes to the antigenic region identified by the serum antibodies present in the serum. This could be IgG to LIV or another closely related virus. The next steps for the phage clones would be to synthesise the peptides and test these against the independent serum panel. The peptides should also be tested against TBEV sera to study if there is cross-reactivity observed.

4.4.3 NGS analysis

The individual sub-libraries from R3 were sequenced to identify the 16-mer peptides enriched. The pooled R3 sub-library was where the phage clones were picked from. By combining the ELISA and NGS data, the results show that the best performing clones are not always enriched in the top 250 peptides of the NGS data. It also shows that the most enriched peptides in the NGS data sets might not be represented on the culture plate.

Further work would include data mining the NGS data sets to identify peptides enriched preferentially in LIV IgG positive sub-libraries and tracing the 3 peptide clones which were not represented in the top 250 peptides but had good ELISA profiles.

4.4.4 Conclusion

To conclude, peptide mimotopes to LIV were discovered by biopanning serum IgG with a phage peptide library. The peptides were able to discriminate between LIV positive and negative serum with five of the peptides showing AUC values >0.80. Masking LIV negative serum 35 results, two peptide combinations for a multiplex ELISA have been suggested based on the results from the independent sera tested against of the phage clones. One panel had a sensitivity/specificity value of 95 and 78% respectively while the other had 91% and 84 % respectively.

The ELISA presented in this study is a proof-of-concept assay for detecting LIV IgG in ovine sera. With further refinements to the assay such as testing on more samples and screening some more of the peptides identified in the NGS data sets, this ELISA could be used for studying sero-prevalence of LIV and potentially be introduced as a serology diagnostic used routinely to detect LIV in veterinary samples. Finally, this work shows that combining traditional clone picking and deep

sequencing the output phage libraries, it is possible to identify a large peptide panel which are mimotopes for LIV IgG phenotypically.

Chapter 5 **Application of next generation phage display technology to identify 16-mer peptide mimotopes specific to Tick-borne encephalitis virus and louping ill virus serum antibodies.**

5.1 Introduction

Tick-borne encephalitis virus (TBEV) and louping ill virus (LIV) are flaviviruses and members of the TBEV serocomplex which are present in the UK. European subtype (TBEV-Eu) of TBEV is the TBEV subtype that is most closely related to LIV. TBEV and LIV are transmitted by the tick species *Ixodes ricinus* as they blood feed from a range of animals and birds during their life cycle and experimental work in goats has shown that LIV is also transmissible in milk (Reid *et al.*, 1984)

LIV is endemic in the UK (Jeffries *et al.*, 2014; Gilbert, 2016). TBEV was not reported from the UK until 2019 when it was detected in two regions of the UK from a sentinel deer population and in ticks collected from the area (Holding *et al.*, 2019, 2020). The locations where TBEV was reported does not overlap with the locations from where LIV has been reported previously. Although both viruses can infect a range of mammals and birds, TBEV is mainly associated with human disease but can also cause disease in animals such as dogs and sheep (Böhm *et al.*, 2017). LIV is mainly associated with a fatal disease in sheep and red grouse (Gilbert, 2016). Both viruses are neurotropic and disease outcome can be fatal.

Currently, there are several human vaccines available for disease control in regions where the virus is endemic (Kubinski *et al.*, 2020) but veterinary vaccines for TBEV do not currently exist. In the case of LIV, as mentioned in the previous chapter, a vaccine is currently in development (Game and Wildlife Conservation Trust, 2021).

As is the case with other closely related flaviviruses, the antibodies to these two viruses are cross-reactive (Klaus *et al.*, 2014). This is an issue when it comes to serological detection of the two viruses especially now that TBEV is also present in the UK. Both viruses can infect animals of economic importance and in the case of TBEV, it is also highly pathogenic to humans. A serological diagnostic assay which is specific to distinguish between the two viruses is currently lacking. This chapter looks at whether next generation phage display technology (NGPD) and virus specific peptide immunosignature to serum IgG would be able to provide a solution to the problem.

In 4.2.2, a biopanning experiment was described with biopanning the P8 16-mer random peptide phage library to identify a panel of phage peptides, subsequently tested in an ELISA, and were able to distinguish ovine sera positive for LIV IgG from a panel of LIV IgG positive and healthy serum.

Can phage biopanning enrich for peptide mimotopes to serum IgG which will then be able to distinguish between two closely related flaviviruses in a panel of sheep sera which have sera positive for TBEV IgG or LIV IgG?

The objectives of the chapter were:

1. Using a commercial ELISA kit, test and confirm the presence of IgG antibodies to TBEV in the ovine sera provided by collaborators from Norway and Estonia.
2. Use the sera positive for TBEV IgG antibodies to biopan the P8 library, used in the previous experimental chapters, to enrich for phage mimotopes to TBEV IgG.
3. Biopan a previously enriched phage sub-library with binders to LIV IgG antibodies to remove cross reactive antibodies to TBEV IgG. For this experiment, round 2 pooled LIV sub-library from section 4.3.1 was used.

4. The individual output libraries from both experiments will be sequenced on the Ion torrent sequencing platform. Using the machine learning method of training and testing data, a virus specific immunosignature will be identified.

5.2 Materials and methods

5.2.1 Serum samples for TBEV panel.

Serum samples from sheep were provided by collaborators from Norway and Estonia.

5.2.1.1 Ovine serum samples from Norway

Sera from 15 sheep were kindly provided by Katrine Mørk Paulsen (Norwegian University of Life Sciences, Sandnes, Norway) from animals experimentally infected with TBEV as part of a ScandTick Innovation grant (grant number 20200422) supported by the EU Interreg V-A program. The study was conducted over 2017 and 2018. Table 5-1 provides the data known about the study.

Table 5-1: Ovine sera from Norway. Animals were experimentally infected

Animal ID	Day post infection	Year of study	TBE titre	Infected with Anaplasma phagocytophilum
Pre-infection				
1_8	0	2017		
2_36	0	2017		
3_78	0	2017		
4_119	0	2017		
5_264	0	2017		
11_24	0	2017		
12_55	0	2017		
13_108	0	2017		
14_176	0	2017		
15_504	0	2017		
1_26	0	2018		
2_38	0	2018		
3_92	0	2018		
4_143	0	2018		
5_150	0	2018		
Post infection				
1_8	14	2017	240	no
2_36	14	2017	20	no
3_78	14	2017	120	no
4_119	14	2017	160	no
5_264	14	2017	160	no
1_8	31	2017	160	yes
2_36	31	2017	30	yes
3_78	31	2017	80	yes
4_119	31	2017	80	yes
5_264	31	2017	120	yes
11_24	10	2017	480	no
12_55	10	2017	320	no
13_108	10	2017	80	no
14_176	10	2017	160	no
15_504	10	2017	80	no
1_26	18	2018	240	yes
2_38	18	2018	320	yes
3_92	18	2018	40	yes
4_143	18	2018	240	yes
5_150	18	2018	160	yes

The serum samples had been screened for TBEV IgG using a modified version of the TBE IgG: Enzygnost® ELISA (Anti-TBE virus IgG, Siemens Healthcare, GmbH, Marburg, Germany) and results were confirmed by a TBEV-specific serum neutralisation test (SNT) performed by Karin Stiasny at the Center for Virology of the Medical University of Vienna according to their published protocol (Stiasny, Holzmann and Heinz, 2009) and SNT titre is provided in Table 5-1. Some of the animals were co-infected with the obligate intracellular bacterial pathogen *A. phagocytophilum* as part of the study aim.

Serum samples 1_8, 2_36, 3_78, 4_119 and 5_264 were drawn from the same animals at day 0 pre-infection, and days 14 and 31 post infection, respectively. Between bleeds on day 14 and 31, the animals were co-infected with *A. phagocytophilum*.

Serum samples from animals 11_24, 12_55, 13_108, 14_176, 15_504 were drawn on day 10 and these animals were infected with TBEV only. Serum samples from animals 1_26, 2_38, 3_92, 4_143, and 5_150 were drawn on day 18 post infection and these animals were co-infected with TBEV and *A. phagocytophilum*.

5.2.1.2 Ovine serum samples from Estonia

The Estonian ovine sera (n=103) were provided by Arvo Viltrop (Estonian University of Life Sciences). They were field samples taken for serosurveillance. Out of the 103 samples, 56 had been tested for the presence of TBEV antibodies by the collaborator on an unspecified assay and a score of 0, 0.5 or 1 was assigned to the samples (Table 5-2).

Table 5-2: Sample information that was provided for the Estonian sheep sera

Estonian Sample ID	Herd ID	Animal ID	Date of submission	TBEV antibody level
65	20265	1239195	17/03/2012	0.5
66	20265	1473100	17/03/2012	0.5
68	20265	2383781	17/03/2012	0.5
71	20265	2383606	17/03/2012	1
72	20265	1473162	17/03/2012	1
74	20265	809634	17/03/2012	1
75	20265	2383651	17/03/2012	1
76	20265	1905205	17/03/2012	1
77	20265	2687858	17/03/2012	1
78	20265	1671292	17/03/2012	1
143	13582	382465	04/04/2012	0.5
144	13582	2752792	04/04/2012	0.5
145	13582	829748	04/04/2012	0
146	13582	235815	04/04/2012	0
147	13582	2752709	04/04/2012	0
148	13582	1683707	04/04/2012	0
149	13582	1683981	04/04/2012	0
150	13582	829373	04/04/2012	0.5
156		811392	04/04/2012	1
232	12330	961868	14/04/2012	1
233	12330	961158	14/04/2012	1
236	12330	2003382	14/04/2012	1
239	12330	2007502	14/04/2012	1
240	12330	2003412	14/04/2012	1
242	12330	2265162	14/04/2012	0.5
247	26462	1144192	14/04/2012	1
249	26462	505390	14/04/2012	1
253	26462	843898	14/04/2012	1
256	26462	690294	14/04/2012	1
259	25508	1786637	14/04/2012	1
264	25508	2634203	14/04/2012	1
265	25508	3634371	14/04/2012	1
266	25508	2634425	14/04/2012	0
267	25508	2634494	14/04/2012	0
269	25508	2634814	14/04/2012	0
270	25508	2634845	14/04/2012	0
272	25508	2634937	14/04/2012	1
274	11609	980241	14/04/2012	1
275	11609	980265	14/04/2012	1
276	11609	980272	14/04/2012	0.5
278	11609	980302	14/04/2012	0.5

279	11609	1077308	14/04/2012	0.5
280	11609	1077445	14/04/2012	0.5
281	11609	1077483	14/04/2012	0.5
282	11609	1140354	14/04/2012	0.5
284	11609	1236309	14/04/2012	0.5
286	11609	1316209	14/04/2012	1
287	14522	517867	14/04/2012	1
288	14522	708982	14/04/2012	1
289	14522	709354	14/04/2012	1
290	14522	1109030	14/04/2012	1
291	14522	1109085	14/04/2012	0
292	14522	1109184	14/04/2012	0
296	14522	1110500	14/04/2012	0.5
297	14522	1110227	14/04/2012	0.5
300	14522	1352108	14/04/2012	0.5
345	17818	2952161	10/05/2012	nt[1]
350	17818	2952956	10/05/2012	nt
351	17818	2952826	10/05/2012	nt
354	17818	2268934	10/05/2012	nt
408	23254	442541	29/05/2012	nt
410	23254	805537	29/05/2012	nt
412	23254	644075	29/05/2012	nt
413	23254	1295405	29/05/2012	nt
415	23254	1295214	29/05/2012	nt
418	23254	1475944	29/05/2012	nt
419	23254	2046006	29/05/2012	nt
422	26599	714372	30/05/2012	nt
423	26599	589024	30/05/2012	nt
424	26599	2182537	30/05/2012	nt
425	26599	2182391	30/05/2012	nt
426	26599	366953	30/05/2012	nt
428	26599	1078695	30/05/2012	nt
429	26599	2051918	30/05/2012	nt
431	26599	255264	30/05/2012	nt
434	26599	1117332	30/05/2012	nt
435	26599	1739367	30/05/2012	nt

[1] Nt= not tested

477	11297	1346121	14/06/2012	nt
481	11297	1345827	14/06/2012	nt
482	11297	845939	14/06/2012	nt
483	11297	1346084	14/06/2012	nt
484	14331	1114805	18/06/2012	nt
485	14331	1393392	18/06/2012	nt
489	14331	1114676	18/06/2012	nt
490	14331	2258522	18/06/2012	nt
491	14331	1393316	18/06/2012	nt
492	14331	1784718	18/06/2012	nt
493	14331	1393026	18/06/2012	nt
499	14331	843058	18/06/2012	nt
500	14331	843218	18/06/2012	nt
501	14331	843010	18/06/2012	nt
502	14331	843508	18/06/2012	nt
503	14331	2370378	18/06/2012	nt
505	14331	843461	18/06/2012	nt
506	14331	2370392	18/06/2012	nt
507	14331	843294	18/06/2012	nt
509	14331	843140	18/06/2012	nt
510	14331	1114836	18/06/2012	nt
864	11362	2319926	11/10/2012	nt
867	11362	1567953	11/10/2012	nt
873	11362	1404838	11/10/2012	nt
874	11362	754972	11/10/2012	nt
876	11362	1073874	11/10/2012	nt

5.2.2 LIV sera panel

The serum panel containing samples LIV 1–20 is shown in Table 4-1 and were the same sera used to enrich LIV peptide mimotopes in the previous chapter. LIV 21–30 was the set of independent sera used to test the phage clones in the previous chapter (Table 4-15).

Table 5-3: LIV serum samples 21–30

Sample number	Animal Id
LIV 21	14/0570
LIV 22	14/0585
LIV 23	14/0647/2
LIV 24	14/0662
LIV 25	14/0850/1
LIV 26	14/0850/2
LIV 27	14/0938/1
LIV 28	16/0289/3
LIV 29	16/0383/2
LIV 30	17/0663

5.2.3 Immunoassays

ELISA and haemagglutination inhibition assay (HAI) were used to test the sera before they were incorporated into the biopanning panel. The benchmark was a positive result for TBEV IgG and a negative result for LIV IgG.

5.2.3.1 TBEV ELISA

The serum samples were tested for the presence of TBEV IgG using the IMMUNOZYM® FSME (TBE) IgG All Species TBEV ELISA kit (Cat. No. 7701075, Progen), following the manufacturer's instructions. The results from the Progen ELISA are interpreted in Vienna units (VIEU/mL), where <63 is negative, 63-126 is borderline and >126 is positive. The assay kit has 5 QC standards and the OD readings from the QC wells were used to plot a reference curve. The OD values from the test samples were plotted against the reference curve to generate the VIEU/mL values.

5.2.3.2 Haemagglutination inhibition assay for LIV

Dr Mara Rocchi kindly tested the Estonian samples for the presence of LIV IgG using haemagglutination inhibition assay (HAI) at the MRI (Edinburgh). The HAI was carried out on Estonian sera only as the Norwegian sera were from a flock that was infected with TBEV.

5.2.4 Biopanning existing phage sub-library enriched for binders to LIV IgG to reduce cross reactive peptide mimotopes to TBEV IgG.

The aim of this panning experiment was to enrich the round 2 output sub-library enriched for phage peptides to LIV IgG antibodies (chapter 4.3.1) further to make it more specific for LIV IgG binders and remove any phage that could potentially cross react with TBEV IgG. was used in this one round of biopanning.

The biopanning strategy is seen in 'Biopanning round 3' (Figure 5-1). The round 2 phage library was grown, PEG precipitated, and processed to reduce non-specific binding phage to the Protein-G agarose beads and blocked with milk block prior to use in the biopanning experiment.

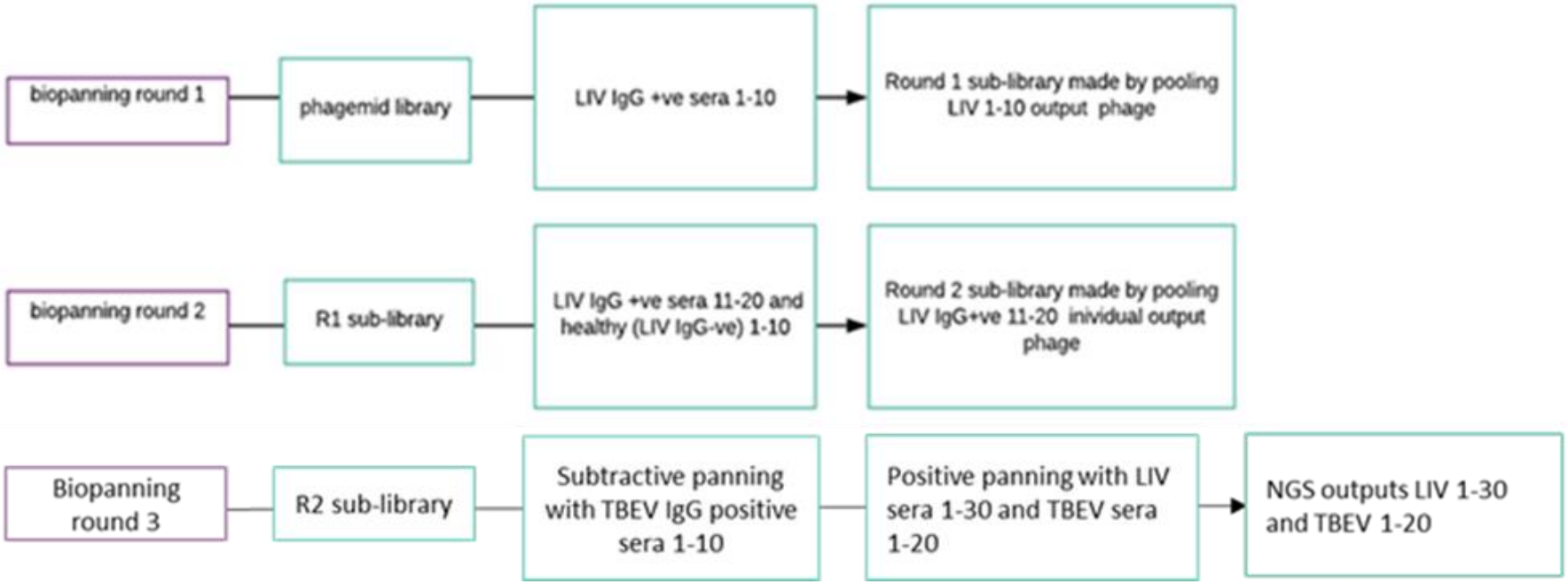


Figure 5-1: Biopanning strategy used in enriching LIV IgG specific phage binders while depleting any that are cross reactive for TBEV IgG

5.2.4.1 Subtractive biopanning

For the subtractive biopanning, 2 µL of serum from ten TBEV IgG positive serum samples (TBEV 1-10) were pooled (20 µL) and incubated with Protein G agarose beads for 30 minutes and then blocked with milk block for further 30 minutes. The agarose beads were added to the phage library and left on a rocking platform for 30 minutes. The library was centrifuged to pellet the beads and the supernatant was carefully removed into a fresh tube without disturbing the bead pellet as described in Chapter 2.1.10.

5.2.4.2 Positive biopanning

The positive panning (screening) round used 30 LIV IG positive sera (LIV 1—30) and 20 TBEV IgG positive sera (TBEV 1—20). For this part, 5 µL of each serum sample was incubated with Protein G agarose beads to capture serum IgG in individual reaction tubes and subsequently blocked with milk block.

The samples were positively biopanned as described in Chapter 2.1.11. The overnight biopanned libraries were rescued as described 2.1.12 and the output library was titrated and stored as described in Chapter 2.1.13. The output libraries were stored in individual and pooled libraries till they were processed for NGS.

5.2.5 Biopanning P8 peptide phage library for binders to ovine serum positive for TBEV IgG

This experiment was conducted to enrich for peptide mimotopes specific to TBEV IgG. In order to achieve this goal, the P8 phage library was initially positively biopanned with a panel of ten TBEV IgG positive sera. In the second round of biopanning, the pooled phage sub-library from the first round was subtractively biopanned with pooled LIV IgG positive sera to deplete any binders that are cross-reactive, and the library was positively biopanned with a separate set of 10 TBEV sera. In the third and final round, the pooled output sub-library from the second round of biopanning was screened against the TBEV sera used in the biopanning, the ten LIV sera used in the subtractive biopanning and a further 20 LIV sera not used in the biopanning process. The biopanning strategy used in this experiment is shown in Figure 5-2.

Glycerol stocks of the output sub-libraries from the third round were used for Ion Torrent sequencing. The TBEV sub-libraries are the target, and the LIV sub-libraries are the control and any peptides that are enriched significantly in the TBEV compared to the LIV sub-libraries are potential mimotopes for TBEV IgG.

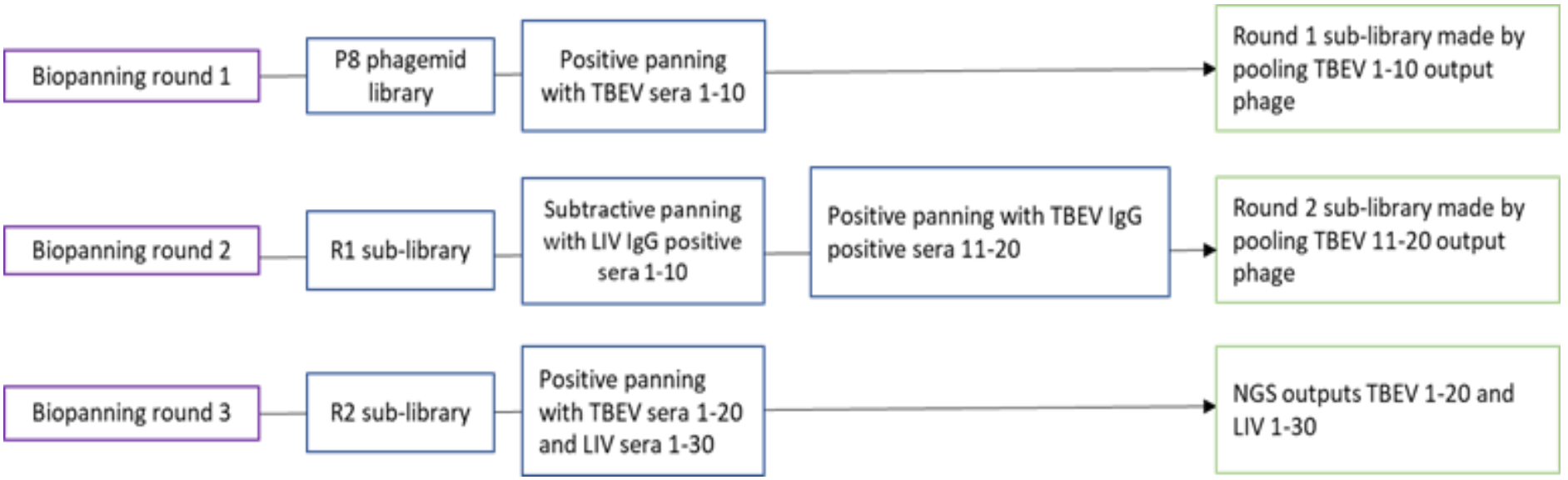


Figure 5-2: Biopanning schematic for TBEV vs LIV experiment, to enrich peptide mimotopes that are specific for TBEV IgG while depleting any that are cross reactive for LIV IgG.

5.2.5.1 Round 1

The first round of biopanning was positive biopanning of the P8 library to 10 TBEV IgG positive serum. The ten output sub libraries from the first round of biopanning were pooled to create the round 1 sub-library.

5.2.5.2 Round 2

The second round of biopanning included subtractive panning to pooled LIV IgG positive sera to reduce phage peptides that are potentially cross reactive to LIV IgG. For the subtractive biopanning, 2 µL of serum from ten LIV IgG positive serum samples (LIV 1-10) were pooled (20 µL).

The library was then positive panning individually with a further ten TBEV IgG positive sera. The output phage sub libraries were stored individually and as the pooled output of that sample group from that round. The pooled sub-library was then used as the input library for the third round of biopanning.

5.2.5.3 Round 3

The third and final round was the screening round where the input sub library was positively panned with all the TBEV samples used in previous rounds (TBEV 1—20); the ten LIV IgG sera used for the subtractive panning and an additional 20 LIV sera (LIV 1—30). There were 50 round 3 output sub-libraries, and these were stored as bacterial glycerol stocks until they were processed for NGS.

5.2.6 Sample preparation for Ion torrent sequencing

The phage sub-library glycerol stocks from the final round of biopanning in sections 5.2.4 and 5.2.5 were processed for Ion Torrent sequencing. Plasmid DNA from all the samples were purified using Qiagen miniprep kit (Qiagen 27106). The concentration of DNA was calculated on the Nanodrop, and 10 ng/µL stock solution was made for each sample. There

were two stages of PCR, the first was to amplify the P8 library specific region which contains the phage peptide. The second stage of PCR added the DNA barcodes which were used to identify the individual sub-libraries when NGS data was analysed as samples from multiple experiments were pooled.

Successful PCR was confirmed by running 10 µL of the PCR products on a 2% agarose gel. The PCR products were cleaned up using Machery-Nagel Nucleospin Gel and PCR clean up kit (Machery-Nagel 740609.250) and the success of the PCR was determined by visualising the bands on a 2% gel following gel electrophoresis of the PCR products. The cleaned product was eluted into 20 µL elution buffer.

The second round of PCR introduced the NGS barcodes and the P1 primer which is necessary for Ion Torrent sequencing. The second round PCR template was 1µL of the cleaned product from round one PCR. The success of the second round was confirmed by gel electrophoresis. The remaining round 2 PCR product was cleaned using the previously mentioned kit. DNA was quantified using the Qubit double stranded BR assay kit.

The barcoded samples were pooled at 10 ng/µL concentration per sample and the pooled DNA was gel purified on a 3% Metaphor gel. The Metaphor gel was run slowly over 3 hours to ensure good separation and a clean band. The band was excised carefully, gel purified, and the concentration of product was quantified using the Qubit.

The pooled DNA was cleaned a further two times using the Ampure XP clean-up kit, and DNA was quantified on the Qubit. A confirmatory 1% agarose gel was run with 50 ng of pooled DNA to confirm that the product was a single band. The DNA was then sent to the Ion Torrent sequencing facility at the University of Pennsylvania in a final volume of 50 µL, at a concentration of 4 ng/µL.

5.2.7 NGS data processing

The NGS data files from Pennsylvania were processed through a data extraction pipeline to uncompress the files; convert the FASTQ files to FASTA format; de-multiplex them into the individual sample files identified by barcodes. The fasta files were translated in all three reading frames taking the amber stop codon into consideration. The 16-mer peptide of interest was identified using the flanking regions 'AEGEF' and 'DPAKA'. The string length of 16-mer was specified to reduce nonsense peptides which have the right flanking region but are not experiment specific.

The peptides and the frequency of reads for each barcode was saved as individual files with the barcode number to trace it back to the original sample. The files were then used for data analysis to identify peptide populations enriched through biopanning.

5.3 Results

5.3.1 Immunoassay results

ELISA and HAI assays were carried out to ascertain the antibody levels in sera before they were chosen for biopanning experiments. The immunoassay results for both tests are shown in Table 5-4.

5.3.1.1 Progen TBEV IgG ELISA

Ninety six samples were tested for the presence of TBEV IgG, 76 of these were from the Estonian batch and 20 from the Norwegian batch (column 3 Table 5-4). The sera were tested as a confirmatory test to the TBEV result from the collaborators.

All the post challenge samples from the Norwegian cohort were tested as they already had a positive result for TBEV. Samples from the Estonian cohort were tested on the Progen ELISA initially if they were assigned a positive TBEV result of 0.5 or 1. Following the LIV HAI results from Moredun, those sera that tested positive for LIV HAI were screened for TBEV IgG on the Progen ELISA kit.

5.3.1.2 LIV HAI

Only Estonian samples were tested for LIV IgG as they were field samples. The TBEV result from Estonia reported that 11 samples had TBEV Ab result of 0, 17 samples had TBEV Ab result of 0.5, 28 h ad TBEV results of 0.1 and 47 were not tested for TBEV. Samples (n=3) that were extremely haemolysed were excluded from the LIV HAI assay.

The HA I assay to determine LIV IgG levels were carried out at the MRI (Edinburgh). One hundred samples were sent for testing, 98 samples were tested, and 64 sera tested positive for LIV IgG, the values ranging from 1:10 to ≥1:640. This data is presented in Appendix 4, Table A5-1. Samples that fit the biopanning criteria are shown in Table 5-4.

Table 5-4: Serum samples from Norwegian and Estonian collaborators tested for TBEV IgG and LIV IgG. Samples that were TBEV positive on the Progen ELISA and had LIV HAI titre ≤1:20 were assigned to the TBEV biopanning panel.

	Source serum Id	Progen ELISA results. Below 65 Vienna units is negative	LIV HAI result reported from MRI	TBEV result reported from source institute
Estonia serum				
1	259	383.55	01:10	1
2	256	277.25	01:10	1
3	272	188.81	01:20	1
4	274	193.33	01:20	1
5	275	232.33	01:20	1
6	286	168.31	01:10	1
7	287	138.78	01:10	1
8	288	165.8	01:10	1
9	289	174.8	01:10	1
10	276	151.65	01:20	0.5
11	874	135.73	01:10	nt
Norway serum				
12	3_78 D14	601.42	nt	120
13	4_119 D14	354.11	nt	160
14	5_264 D14	141.89	nt	160
15	1_8 D31	567.73	nt	160
16	11_24 D10	130.77	nt	480
17	1_26 D18	601.42	nt	240
18	2_38 D18	193.33	nt	320
19	3_92 D18	354.11	nt	40
20	5_150 D18	262.38	nt	160

*nt: not tested.

5.3.2 Biopanning existing phage sub-library to enrich for peptide mimotopes specific to LIV IgG (LIV vs TBEV)

The aim of this biopanning experiment was to enrich a phage sub-library with peptide mimotopes specific for LIV IgG while removing any phage peptides that were cross-reactive to TBEV IgG. Table 5-5 shows the details of the samples used in this experiment: LIV 1–30 and TBEV 1–20 and the rounds of panning each sample was used in.

Table 5-5: Serum samples from LIV and TBEV panels used to enrich for LIV IgG binders with TBEV IgG subtraction. 3* in TBEV panel denotes serum used in subtractive biopanning

Biopanning sample Id	Serum Id	Biopanning round
LIV panel		
LIV 1	14/0647/1	1 and 3
LIV 2	14/0724/2	1 and 3
LIV 3	14/0724/4	1 and 3
LIV 4	15/0441	1 and 3
LIV 5	15/0675	1 and 3
LIV 6	16/0383/1	1 and 3
LIV 7	16/0562	1 and 3
LIV 8	16/0704	1 and 3
LIV 9	17/0433	1 and 3
LIV 10	17/0504/1	1 and 3
LIV 11	14/0859	2 and 3
LIV 12	14/0916	2 and 3
LIV 13	14/0938/2	2 and 3
LIV 14	16/0358	2 and 3
LIV 15	16/0409/2	2 and 3
LIV 16	16/0741/2	2 and 3
LIV 17	16/0797	2 and 3
LIV 18	16/0860	2 and 3
LIV 19	17/0224	2 and 3
LIV 20	17/0319	2 and 3
LIV 21	14/0570	3
LIV 22	14/0585	3
LIV 23	14/0647/2	3
LIV 24	14/0662	3
LIV 25	14/0850/1	3
LIV 26	14/0850/2	3
LIV 27	14/0938/1	3

LIV 28	16/0289/3	3
LIV 29	16/0383/2	3
LIV 30	17/0663	3
TBEV panel		
TBEV 1	3_78 D14	3* and 3
TBEV 2	4_119 D14	3* and 3
TBEV 3	1_8 D31	3* and 3
TBEV 4	1_26 D18	3* and 3
TBEV 5	3_92 D18	3* and 3
TBEV 6	256	3* and 3
TBEV 7	259	3* and 3
TBEV 8	275	3* and 3
TBEV 9	286	3* and 3
TBEV 10	289	3* and 3
TBEV 11	5_264 D14	3
TBEV 12	11_24 D10	3
TBEV 13	2_38 D18	3
TBEV 14	5_150 D18	3
TBEV 15	287	3
TBEV 16	288	3
TBEV 17	874	3
TBEV 18	272	3
TBEV 19	274	3
TBEV 20	276	3

5.3.2.1 Output phage sub-library titre

The titre of the input library used in this experiment, round 2 pooled LIV sub-library, was 3.62 x 10^5 cfu/mL (Table 4-5). The pooled output titre for the LIV and TBEV sub-libraries following screening is shown in Table 5-6. The LIV pooled sub-library was made by pooling LIV 1—30 outputs. The TBEV pooled sub-library was made by pooling TBEV 1—20 output libraries.

Table 5-6: Output titre from one round of biopanning

Biopanning round	Pooled LIV screening	Pooled TBEV screening
Round 3	2.04 x 10^6 cfu/mL	9 x 10^5 cfu/mL

5.3.3 Biopanning P8 library to enrich for peptide mimotopes specific for TBEV IgG (TBEV vs LIV)

This experiment was carried out to enrich phage peptides specific for TBEV IgG while removing any that are cross reactive to LIV IgG. Table 5-7 shows the serum panel used in this biopanning experiment and the samples that were used in the three rounds.

Table 5-7: Serum samples from TBEV and LIV panels used in biopanning to enrich for phage peptides to TBEV IgG while depleting cross reactive peptides to LIV IgG mimotopes with LIV IgG subtraction. 2* in LIV panel denotes serum used in subtractive biopanning

Biopanning sample Id	Serum Id	Panning round used
TBEV panel		
TBEV 1	3_78 D14	1 and 3
TBEV 2	4_119 D14	1 and 3
TBEV 3	1_8 D31	1 and 3
TBEV 4	1_26 D18	1 and 3
TBEV 5	3_92 D18	1 and 3
TBEV 6	256	1 and 3
TBEV 7	259	1 and 3
TBEV 8	275	1 and 3
TBEV 9	286	1 and 3
TBEV 10	289	1 and 3
TBEV 11	5_264 D14	2 and 3

TBEV 12	11_24 D10	2 and 3
TBEV 13	2_38 D18	2 and 3
TBEV 14	5_150 D18	2 and 3
TBEV 15	287	2 and 3
TBEV 16	288	2 and 3
TBEV 17	874	2 and 3
TBEV 18	272	2 and 3
TBEV 19	274	2 and 3
TBEV 20	276	2 and 3
LIV panel		
LIV 1	14/0647/1	2* and 3
LIV 2	14/0724/2	2* and 3
LIV 3	14/0724/4	2* and 3
LIV 4	15/0441	2* and 3
LIV 5	15/0675	2* and 3
LIV 6	16/0383/1	2* and 3
LIV 7	16/0562	2* and 3
LIV 8	16/0704	2* and 3
LIV 9	17/0433	2* and 3
LIV 10	17/0504/1	2* and 3
LIV 11	14/0859	3
LIV 12	14/0916	3
LIV 13	14/0938/2	3
LIV 14	16/0358	3
LIV 15	16/0409/2	3
LIV 16	16/0741/2	3
LIV 17	16/0797	3
LIV 18	16/0860	3
LIV 19	17/0224	3
LIV 20	17/0319	3
LIV 21	14/0570	3
LIV 22	14/0585	3
LIV 23	14/0647/2	3
LIV 24	14/0662	3
LIV 25	14/0850/1	3
LIV 26	14/0850/2	3
LIV 27	14/0938/1	3
LIV 28	16/0289/3	3
LIV 29	16/0383/2	3
LIV 30	17/0663	3

5.3.3.1 Output phage sub-library titre

The phage library used in this experiment was the same P8 random peptide 16-mer phage library used in previous experiments. The starting titre of the library was $1.5x10^{14}$ cfu/mL.

The pooled output sub-library titre from the 3 rounds of panning is shown in Table 5-8. Output phage titre was calculated to quantify the output phage following a round of biopanning. Round 3 pooled phage sub-library for TBEV was made by pooling TBEV 1—20 outputs and the pooled LIV sub-library was made by pooling LIV 1—30 outputs.

Table 5-8: Output phage titre following three rounds of biopanning

Biopanning round	Pooled TBEV library titre (cfu/mL)	Pooled LIV library titre (cfu/mL)
1	$2.6x10^4$	n/a
2	$7x10^5$	n/a
3	$6.1x10^5$	3.7×10^5

5.3.4 Sample preparation for Ion torrent sequencing

Round 3 sub-libraries from both experiments were processed for Ion Torrent sequencing. Appendix 4, Table A-4.1 shows the sample IDs; barcodes, concentration of DNA used in the first round of PCR and concentration of DNA following second round of PCR and the volume of DNA taken into the DNA pool. Figure 5-3 shows gel electrophoresis images A-F which were taken after each round of PCR for the 95 samples which were pooled together into 'Batch 1 NGS 2020', from Table 5-5 and Table 5-7.

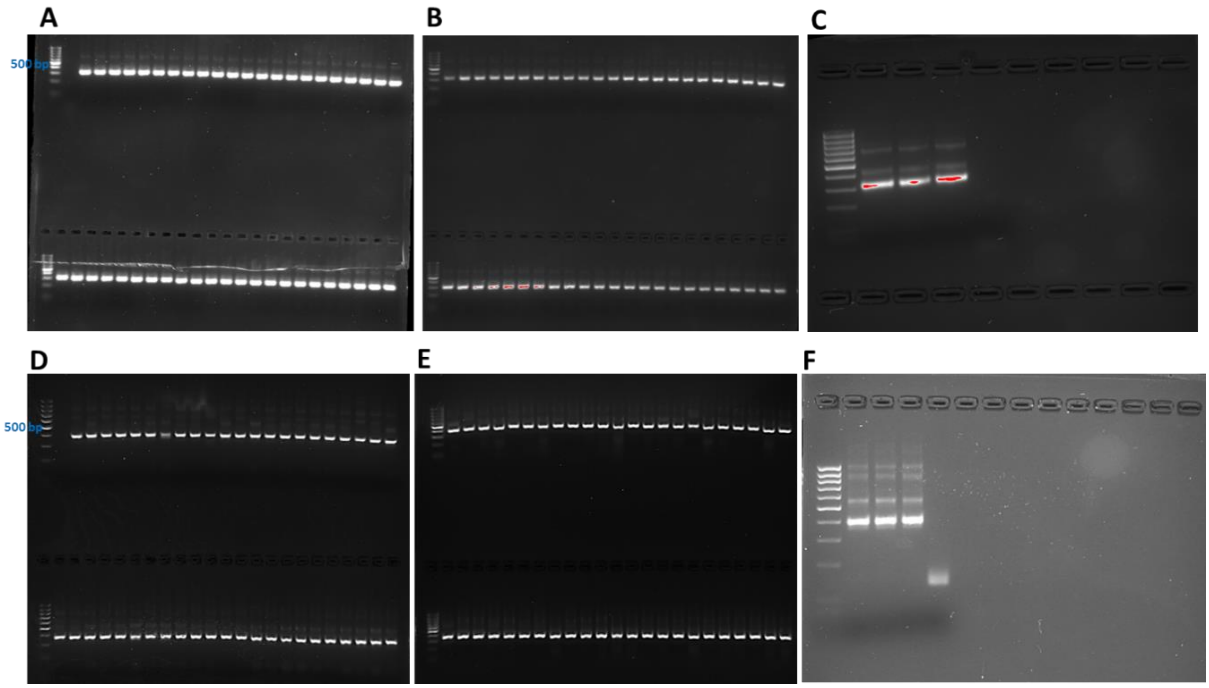


Figure 5-3: Output sub-libraries were processed for Ion Torrent sequencing, the gel images in this figure relate to the panning output libraries (Batch 1 2020 barcodes 1-95) detailed in . The PCR products were run on 2% agarose gels.

Gels A-C shows successful round 1 PCR for all samples. Gel A wells were loaded with samples 1-45 from Round 3 LIV vs TBEV panning. Gel B was loaded with the remaining five samples from LIV vs TBEV and samples 1-42 from R3 of TBEV vs LIV panning. Gel C was loaded with the remaining 3 samples from R3 TBEV vs LIV screening. The negative control was loaded into Gel A well 2. The PCR products are just below 250 bp in size.

Gels D-F are the PCR products from the barcoding PCR, round 2 PCR. The PCR products are ~350 bp in size. Two negative controls were run, Gel D well 2 which was the negative control for this batch and Gel F well 5 which was the negative control from round 1.

5.3.5 Peptide immunosignature specific for LIV IgG

The immunosignature method of data analysis was introduced in section 2.4.4 and the first experiment using this method can be found in section 3.2.6, where a ZIKV specific immunosignature was generated. In this section, a peptide immunosignature for LIV was defined by training and testing the NGS data files of the round 3 sub-libraries derived from the TBEV vs LIV biopanning experiment outlined in section 5.2.4.

Following the third round of biopanning (screening) experiment to enrich for LIV IgG specific mimotopes, the fifty phage sub-libraries from the screening round were processed for Ion Torrent sequencing. NGS barcodes 1–30 were used to barcode the output sub-libraries R3 LIV 1–30 and barcodes 31–51 were used for barcoding the comparator samples TBEV 1–20. The sub-libraries and the associated barcodes are shown in Table 5-9.

Table 5-9: Output libraries from LIV vs TBEV biopanning and the associated NGS barcode

Output library	NGS barcode
LIV 1	BC01
LIV 2	BC02
LIV 3	BC03
LIV 4	BC04
LIV 5	BC05
LIV 6	BC06
LIV 7	BC07
LIV 8	BC08
LIV 9	BC09
LIV 10	BC10
LIV 11	BC11
LIV 12	BC12
LIV 13	BC13
LIV 14	BC14
LIV 15	BC15
LIV 16	BC16
LIV 17	BC17
LIV 18	BC18
LIV 19	BC19

LIV 20	BC20
LIV 21	BC21
LIV 22	BC22
LIV 23	BC23
LIV 24	BC24
LIV 25	BC25
LIV 26	BC26
LIV 27	BC27
LIV 28	BC28
LIV 29	BC29
LIV 30	BC30
TBEV 1	BC31
TBEV 2	BC32
TBEV 3	BC33
TBEV 4	BC34
TBEV 5	BC35
TBEV 6	BC36
TBEV 7	BC37
TBEV 8	BC38
TBEV 9	BC39
TBEV 10	BC40
TBEV 11	BC41
TBEV 12	BC43
TBEV 13	BC44
TBEV 14	BC45
TBEV 15	BC46
TBEV 16	BC47
TBEV 17	BC48
TBEV 18	BC49
TBEV 19	BC50
TBEV 20	BC51

5.3.5.1 Training cohort

Peptide mimotopes specific for LIV IgG +ve sera were trained using the top 250 16-mer peptides by frequency from 20 files: LIV +ve files 1–10 (BC 01–10) and TBEV +ve files 1–10 (BC 31–40). The training files both had 2500 peptides.

The intra group duplicates were removed and this left 806 unique values in the LIV list and 516 unique values in the TBEV list. The peptides

duplicated in both lists were eliminated and this left **405 peptides unique to LIV**, and this list formed the LIV immunosignature.

5.3.5.2 Test cohort

The test cohort was made up with the remaining samples LIV 11–30 (BC11–30) and TBEV 11–20 (BC41–51). The top 400 16-mer peptides from the immunosignature test data set were used for testing the LIV immunosignature. The total 16-mer peptide reads used to normalise the peptide frequencies from each barcode is shown in Table 5-10.

Table 5-10: Total 16-mer peptide read in LIV vs TBEV test barcodes.

Barcode	Total read in file
BC 11	72309
BC 12	74447
BC 13	80457
BC 14	74762
BC 15	83340
BC 16	93408
BC 17	76761
BC 18	7423
BC 19	80862
BC 20	74265
BC 21	78671
BC 22	85307
BC 23	83042
BC 24	78177
BC 25	88794
BC 26	70293
BC 27	74697
BC 28	125722
BC 29	87963
BC 30	87534
BC 41	86279
BC43	70108
BC44	82026
BC45	113506
BC46	91991
BC47	144533
BC48	82505
BC49	100037
BC50	78510
BC51	90188

LIV specific immunosignature

Using the 'VLOOKUP' function in Excel, the 405 LIV specific immunosignature peptides were tested against the test samples and when there was a matched peptide in the training cohort and test cohort, the VLOOKUP function printed the percent frequency value of that specific peptide against the immunosignature panel for each barcode.

5.3.5.3 Heatmap of LIV immunosignature peptides in the test samples

The immunosignature data was exported to GraphPad Prism and plotted as a heat map (Figure 5-4). The heatmap was formatted to show every 25th peptide to avoid overcrowding as there were 405 peptides in the immunosignature. The baseline value for the heatmap was the mean peptide frequency of TBEV test plus (2*STDEV) and this value was 0.51. The highest value on the heatmap was set at 8% and this was the normalised frequency of the peptide with the highest frequency in the immunosignature Test dataset.

The heatmap shows that the biopanning strategy was successful in enriching peptides to LIV IgG as seen by the bands seen in the LIV test data set in the heatmap. There were some bands seen in the comparator TBEV test.

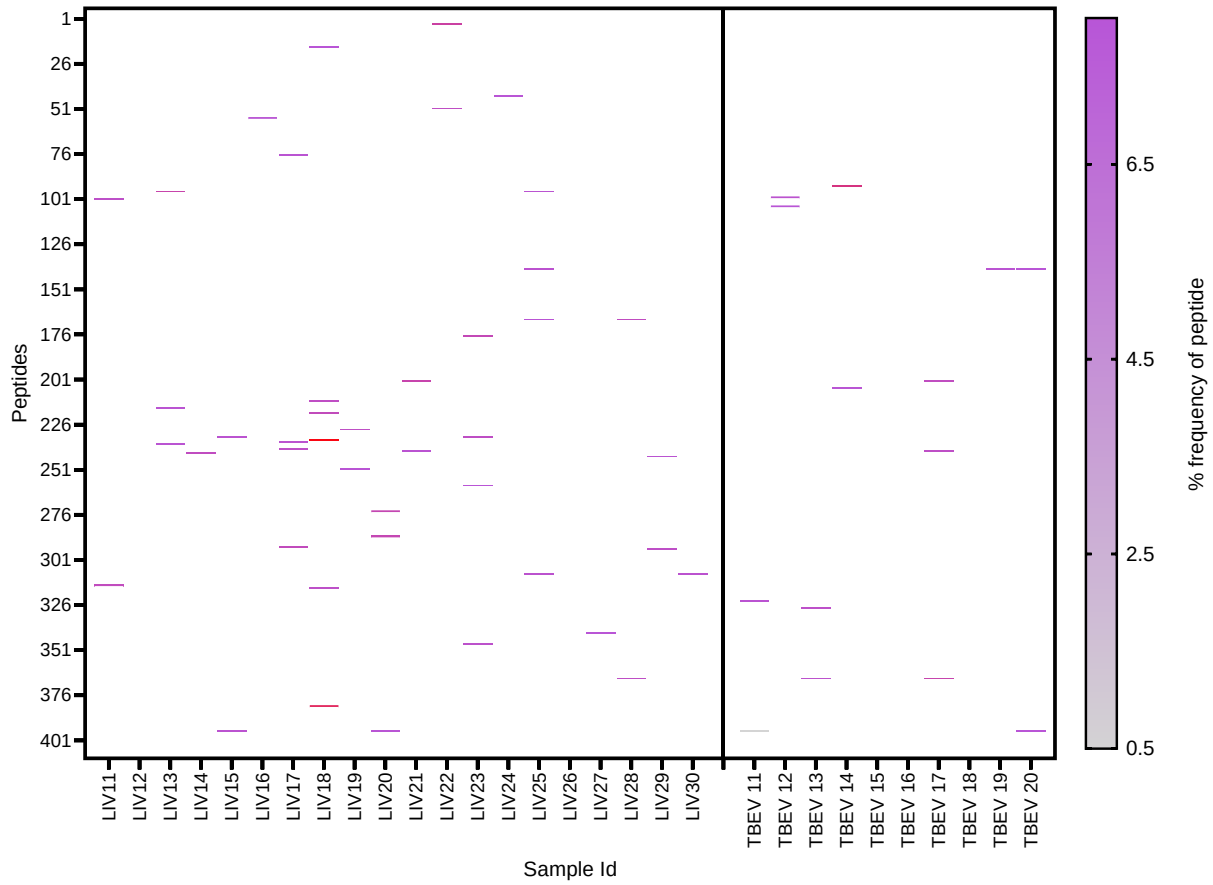


Figure 5-4: Heatmap of LIV immunosignature tested with additional LIV and TBEV samples. The colour of the bands is related to the number of times the specific peptide was enriched in each sample. The baseline value at which the bands register on the heatmap was set to 0.5%.

5.3.5.4 Statistical analysis of the LIV immunosignature *in-silico* assay

The LIV immunosignature *in-silico* assay was analysed using binomial regression, the same test used for assessing the ZIKV immunosignature in Chapter 3 (Section 3.4.3.3).

A Receiver Operating Characteristic (ROC) curve was generated in GraphPad using a subset of the Immunosignature Test set of samples. Samples LIV 11-20 and TEV 11-20 were used to generate the ROC curve. The target samples LIV were assigned the value of 1 and comparator samples TBEV were assigned 0. The ROC plot drawn is shown below in

Figure 5-5 and parameters including the positive/negative cut off for the LIV immunosignature is shown in Table 5-11.

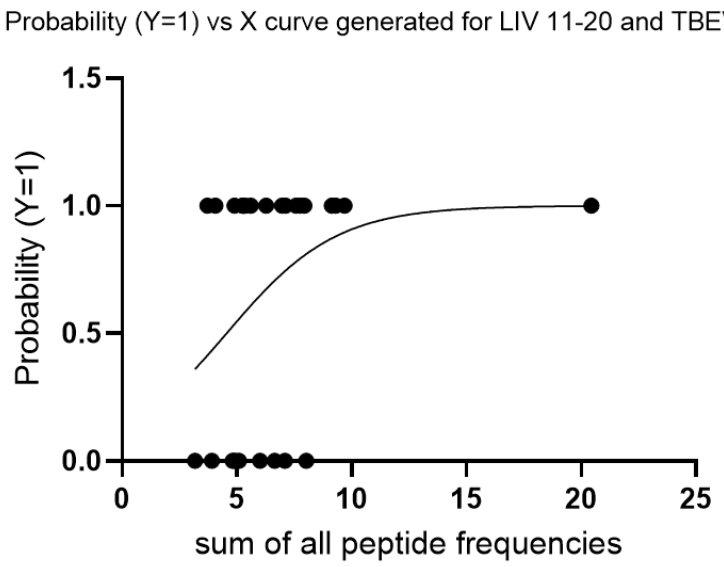


Figure 5-5: ROC curve for the subset of the Immunosignature Test dataset. The target (LIV) samples were assigned value of 1 and the comparator (TBEV) samples were assigned value of 0. The AUC for the curve was 0.780.

Table 5-11: The results from the ROC curve generated for the LIV immunosiganture. The positive/negative cut off value is 6.50.

The ROC analysis parameters calculated for the LIV immunosignature	
Best-fit values	
β0	-3.86
β1	0.594
X at 50%	**6.50**
Std. Error	
β0	2.19
β1	0.337
X at 50%	0.866
Area under the ROC curve	
Area	0.780
Std. Error	0.107
95% confidence interval	0.569 to 0.991
P value	0.0343

In this ROC curve, the X at 50% calculated by GraphPad Prism software, was 6.50 and this is the positive/negative cut off for the LIV immunosignature. Once the ROC curve was generated, the remaining data points (LIV 21-30) were plotted onto the curve. The area under the ROC curve (AUC) for the LIV immunosignature was 0.780 which is classed under 'acceptable' for the diagnostic assay in being able to distinguish between true and false positives. The P value of the ROC plot was 0.034.

The outcome of the binomial regression analysis on all the samples in the LIV immunosignature test set are presented in Table 5-12. Samples where the sum of peptide frequencies was ≥6.5 were considered positive by the model and the rows are highlighted in this table in green shading.

Table 5-12: Binomial regression model results for LIV immunosignature test samples. Samples with predicted probability of being a positive outcome for the assay are shaded in green.

Sample	Sum of peptide frequencies	value of X	Predicted probability
Samples used to generate the binomial regression model			
LIV 11	7.94	1	0.7
LIV 12	4.06	1	0.19
LIV 13	7.12	1	0.59
LIV 14	5.36	1	0.34
LIV 15	9.3	1	0.84
LIV 16	5.6	1	0.37
LIV 17	9.69	1	0.87
LIV 18	20.44	1	1
LIV 19	7.12	1	0.59
LIV 20	7.75	1	0.68
TBEV 11	6	0	0.43
TBEV 12	6.03	0	0.43
TBEV 13	7.09	0	0.59
TBEV 14	8.02	0	0.71
TBEV 15	3.91	0	0.18
TBEV 16	3.17	0	0.12
TBEV 17	6.65	0	0.52
TBEV 18	4.81	0	0.27
TBEV 19	5.09	0	0.3
TBEV 20	4.91	0	0.28
LIV samples tested on the regression model			
LIV 21	5.24		0.32
LIV 22	7.58		0.66
LIV 23	9.13		0.83
LIV 24	6.96		0.57
LIV 25	5.27		0.33
LIV 26	3.71		0.16
LIV 27	6.27		0.47
LIV 28	5.36		0.34
LIV 29	7.53		0.65
LIV 30	4.89		0.28

Setting an *in-silico* positive/negative cut off for the LIV immunosignature using mean plus 2xSD of the comparator.

The sum of percent peptide frequencies seen in column 2 of Table 5-12 were plotted as a bar chart presented in Figure 5-6. The positive/negative cut off was set on the mean+2xSTDEV which was 8.53. At this cut off, 4 out of 20 LIV samples tested positive. All TBEV samples tested negative at the cut off.

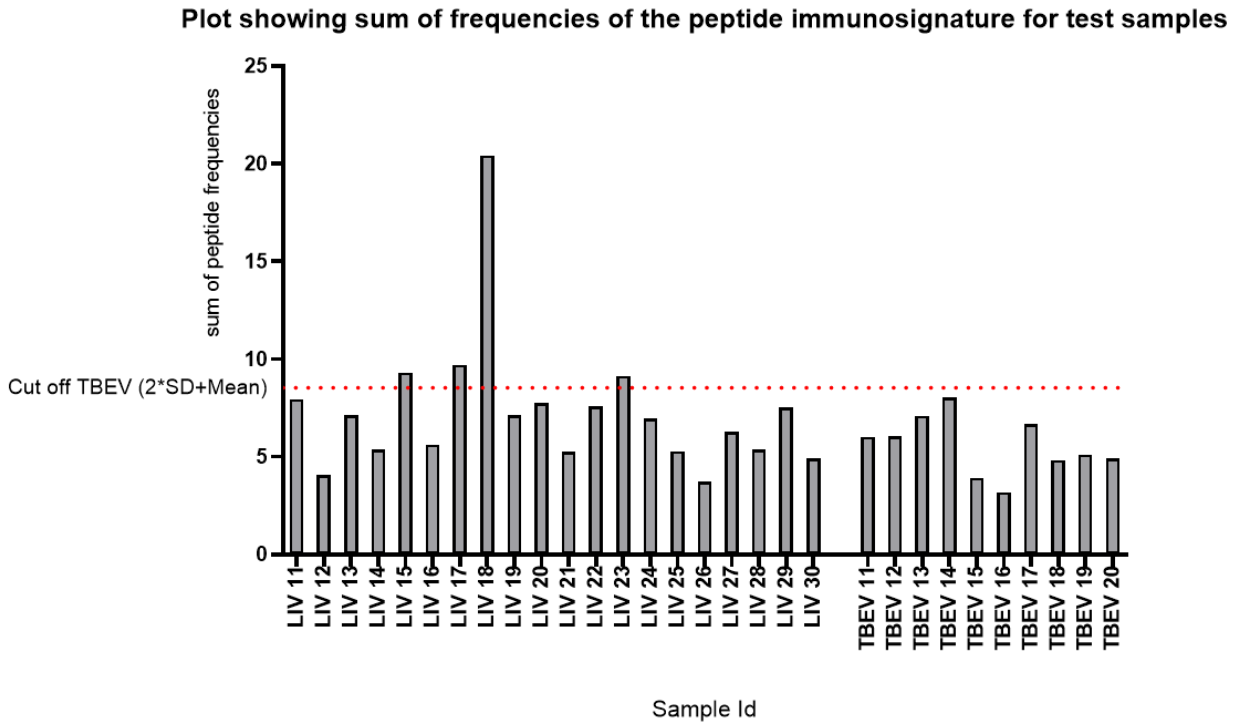


Figure 5-6: Total peptide frequency calculated for each sample with the positive/negative cut off set at 8.53.

5.3.6 Peptide immunosignature for TBEV

Fifty round 3 (R3) sub-libraries from the biopanning were barcoded, pooled, and sequenced on the Ion torrent platform. In this section, a peptide immunosignature for TBEV was defined by training and testing the NGS data files of the round 3 sub-libraries derived from the TBEV vs LIV biopanning experiment outlined in section 5.2.5. The immunosignature analysis method described in 5.3.5 was used in this section.

The samples and their barcodes are shown in Table 5-13. The samples and the barcodes for this experiment were: TBEV 1—20 (Batch 1_BC52–71); and LIV 1—30 (BC72–BC95, Batch 2_BC01–BC06).

Table 5-13: Output libraries from TBEV vs LIV biopanning and the associated NGS barcodes.

TBEV 1	Batch 1 2020 BC52
TBEV 2	Batch 1 2020 BC53
TBEV 3	Batch 1 2020 BC54
TBEV 4	Batch 1 2020 BC55
TBEV 5	Batch 1 2020 BC56
TBEV 6	Batch 1 2020 BC57
TBEV 7	Batch 1 2020 BC58
TBEV 8	Batch 1 2020 BC59
TBEV 9	Batch 1 2020 BC60
TBEV 10	Batch 1 2020 BC61
TBEV 11	Batch 1 2020 BC62
TBEV 12	Batch 1 2020 BC63
TBEV 13	Batch 1 2020 BC64
TBEV 14	Batch 1 2020 BC65
TBEV 15	Batch 1 2020 BC66
TBEV 16	Batch 1 2020 BC67
TBEV 17	Batch 1 2020 BC68
TBEV 18	Batch 1 2020 BC69
TBEV 19	Batch 1 2020 BC70
TBEV 20	Batch 1 2020 BC71
LIV 1	Batch 1 2020 BC72
LIV 2	Batch 1 2020 BC73
LIV 3	Batch 1 2020 BC74
LIV 4	Batch 1 2020 BC75

LIV 5	Batch 1 2020 BC76
LIV 6	Batch 1 2020 BC77
LIV 7	Batch 1 2020 BC78
LIV 8	Batch 1 2020 BC79
LIV 9	Batch 1 2020 BC80
LIV 10	Batch 1 2020 BC81
LIV 11	Batch 1 2020 BC82
LIV 12	Batch 1 2020 BC83
LIV 13	Batch 1 2020 BC84
LIV 14	Batch 1 2020 BC85
LIV 15	Batch 1 2020 BC86
LIV 16	Batch 1 2020 BC87
LIV 17	Batch 1 2020 BC88
LIV 18	Batch 1 2020 BC89
LIV 19	Batch 1 2020 BC90
LIV 20	Batch 1 2020 BC91
LIV 21	Batch 1 2020 BC92
LIV 22	Batch 1 2020 BC93
LIV 23	Batch 1 2020 BC94
LIV 24	Batch 1 2020 BC95
LIV 25	Batch 2 2020 BC01
LIV 26	Batch 2 2020 BC02
LIV 27	Batch 2 2020 BC03
LIV 28	Batch 2 2020 BC04
LIV 29	Batch 2 2020 BC05
LIV 30	Batch 2 2020 BC06

The total number of 16-mer peptide reads in NGS files in the immunosignature test cohort for this experiment is shown in Table 5-14. Peptide counts of those peptides common between the immunosignature and test samples were expressed as percent peptide frequencies using the total 16-mer read for each sample shown in Table 5-14.

Table 5-14: Total peptide reads in the test cohort of TBEV vs LIV panning.

Barcode	Total 16-mer reads
Batch 1 BC 62	71915
Batch 1 BC 63	79092
Batch 1 BC 64	87977
Batch 1 BC 65	83579
Batch 1 BC 66	71728
Batch 1 BC 67	90557
Batch 1 BC 68	72367
Batch 1 BC 69	66518
Batch 1 BC 70	72581
Batch 1 BC 71	75962
Batch 1 BC 82	75478
Batch 1 BC 83	60306
Batch 1 BC 84	77761
Batch 1 BC 85	74141
Batch 1 BC 86	76891
Batch 1 BC 87	90110
Batch 1 BC 88	67785
Batch 1 BC 89	95128
Batch 1 BC 90	74835
Batch 1 BC 91	72980
Batch 1 BC 92	71223
Batch 1 BC 93	67876
Batch 1 BC 94	81936
Batch 1 BC 95	77837
Batch 2 BC 01	51494
Batch 2 BC 02	61167
Batch 2 BC 03	60685
Batch 2 BC 04	52949
Batch 2 BC 05	62368
Batch 2 BC 06	61459

5.3.6.1 Training set

Peptide mimotopes specific for TBEV IgG +ve sera were trained using the top 250 16-mer peptides by frequency from 20 files: TBEV 1—10 (BC 52–61) and LIV 1—10 (BC72–81). Both training files had 2,500 peptides each. The intra group duplicates were removed and this left 1,550 unique values in the TBEV list and 1068 unique values in the TBEV list. The peptides duplicated in both lists were eliminated and this left **1071 peptides unique to TBEV**, which formed the TBEV immunosignature.

5.3.6.2 Test set

The test set were made up of the top 400 peptides from the following sample files TBEV 11—20 (BC62–71) and LIV 11—30 (BC82–BC95, Batch2_BC01–BC06).

TBEV specific immunosignature

The immunosignature peptides were applied to the test samples and using the 'VLOOKUP' function, the peptides that matched with the peptides in the immunosignature list were filtered out along with the percent peptide frequency in that specific sample. The data were tabulated and exported to GraphPad for further analysis.

In GraphPad Prism, the data generated form the immunosignature analysis was plotted on heatmaps and statistical analysis of the sum of the peptide frequencies in the test samples were conducted using binomial logistic regression analysis.

5.3.6.3 Heatmap for TBEV immunosignature

The baseline value for peptides to appear on the heatmap (Figure 5-7) The baseline value for the heatmap was the mean peptide frequency of LIV test plus (2*STDEV) and this value was 0.30. the highest value was 2 and this was the highest peptide frequency common to the TBEV immunosignature.

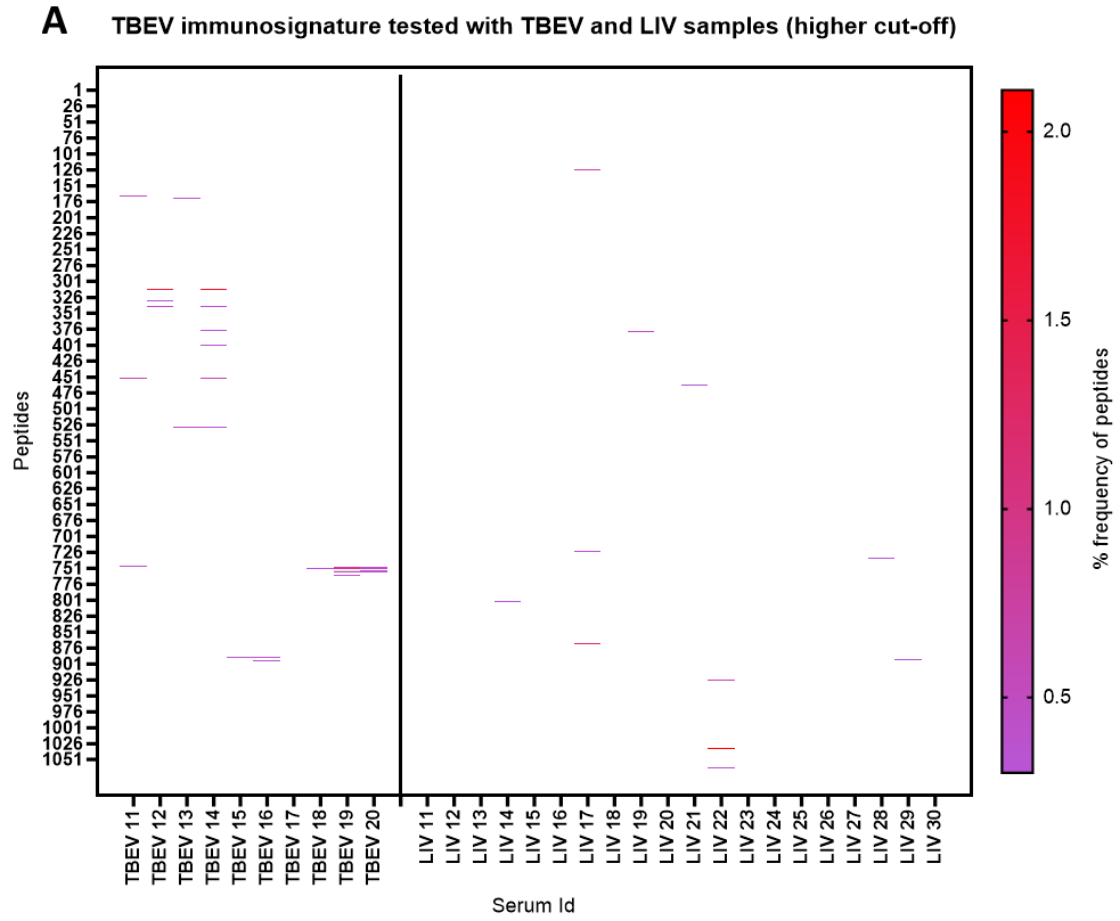


Figure 5-7: Heatmap showing the TBEV immunosignature tested against the TBEV and LIV test samples. The baseline value at which peptides are highlighted on the heatmap was 0.30

5.3.6.4 Binomial simple logistical regression analysis

The sum of peptide frequency values of the peptides common to the immunosignature was calculated for all the test samples in the Immunosignature test cohort. A ROC curve, Figure 5-8, was generated in GraphPad for the immunosignature target samples (TBEV 11–20) and comparator samples (LIV 11-20) The target samples were assigned a value of 1 and the controls were assigned a value of 0. The parameters of the ROC plot are presented in Table 5-15. The remaining LIV samples (LIV 21–30) were fitted to this curve by the software and this generated the predicted probability value seen in column 4 in Table 5-16.

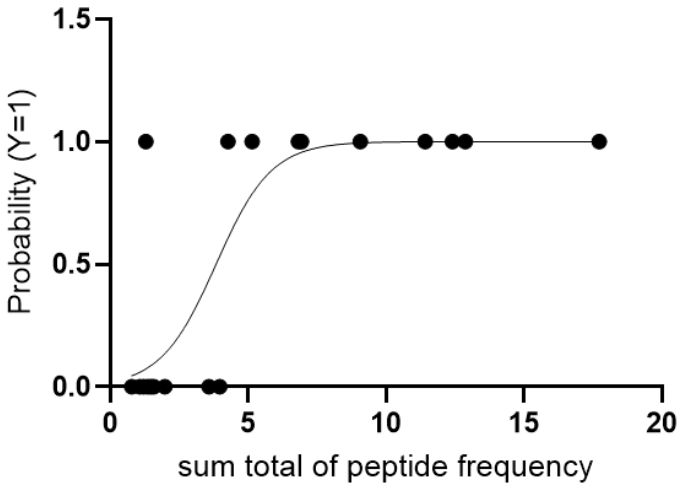


Figure 5-8: ROC curve plotted as part of the binomial logistic regression analysis for the TBEV immunosignature. The curve was plotted using 10 target samples (TBEV 11-20) which were assigned a value of 1 and 10 control samples (LIV 11-20). The AUC for the curve was 0.930.

Table 5-15: Binomial logistic regression results for the TBEV immunosignature.

The ROC analysis parameters calculated for the TBEV immunosignature	
Best-fit values	
β0	-3.82
β1	0.998
X at 50%	**3.83**
Std. Error	
β0	1.69
β1	0.465
X at 50%	0.826
Area under the ROC curve	
Area	0.930
Std. Error	0.0693
95% confidence interval	0.794 to 1.00
P value	0.0012

The positive/negative cut off value generated by the ROC analysis for the TBEV immunosignature was 3.8. The AUC of the ROC curve was 0.930 and this value is considered outstanding for measuring the sensitivity and specificity of diagnostic assays. The P value for the plot was 0.0012.

Samples TBEV 11-TBEV 20 and LIV 11-LIV 20 were used to generate the ROC curve. Table 5-16 shows the outcome of the binomial regression analysis on the remaining samples in the TBEV immunosignature test group. The limited number of samples meant that the *in-silico* assay could only be tested against the control group (LIV 21-30).

Samples which tested positive on the in-silico TBEV immunosignature were shaded in green in this table. There was one false positive in the control set of the TBEV immunosignature test group.

Table 5-16: Sum of the peptide immunosignature of the test samples and predicted probability. Rows with green shading are positive as the cut off value for column 2 is 3.83

Sample	Sum of peptide frequencies	value of X	Predicted probability
Samples used to generate the binomial regression model			
TBEV 11	4.26	1	0.61
TBEV 12	6.82	1	0.95
TBEV 13	6.93	1	0.96
TBEV 14	9.06	1	0.99
TBEV 15	12.41	1	1
TBEV 16	12.87	1	1
TBEV 17	1.28	1	0.07
TBEV 18	5.14	1	0.79
TBEV 19	17.73	1	1
TBEV 20	11.42	1	1
LIV 11	1.35	0	0.08
LIV 12	0.77	0	0.05
LIV 13	1.57	0	0.1
LIV 14	3.96	0	0.53
LIV 15	1.21	0	0.07
LIV 16	1.46	0	0.09
LIV 17	3.57	0	0.44
LIV 18	1.97	0	0.14
LIV 19	1.57	0	0.1
LIV 20	1.06	0	0.06
Samples tested on the regression model			
LIV 21	1.82		0.12
LIV 22	4.14		0.58
LIV 23	2.77		0.26
LIV 24	1.82		0.12
LIV 25	2.13		0.16
LIV 26	2.24		0.17
LIV 27	1.42		0.08
LIV 28	2.07		0.15
LIV 29	2.37		0.19
LIV 30	1.39		0.08

Setting the positive/negative cut off for the TBEV immunosignature *in-silico* assay using mean plus 2SD of the comparator.

The sum of peptide frequencies (Table 5-16 column 2) for the TBEV immunosignature test samples were plotted on a bar chart (Figure 5-9). The positive/negative cut off value was 3.85 and this was derived by calculating the mean of the LIV test set 11-20 plus 2xSTDEV.

Using this cut off, nine of the ten TBEV test samples had a positive result. Seventeen of the LIV samples tested below the cut off for the lower cut off and two tested positive above the higher cut off value.

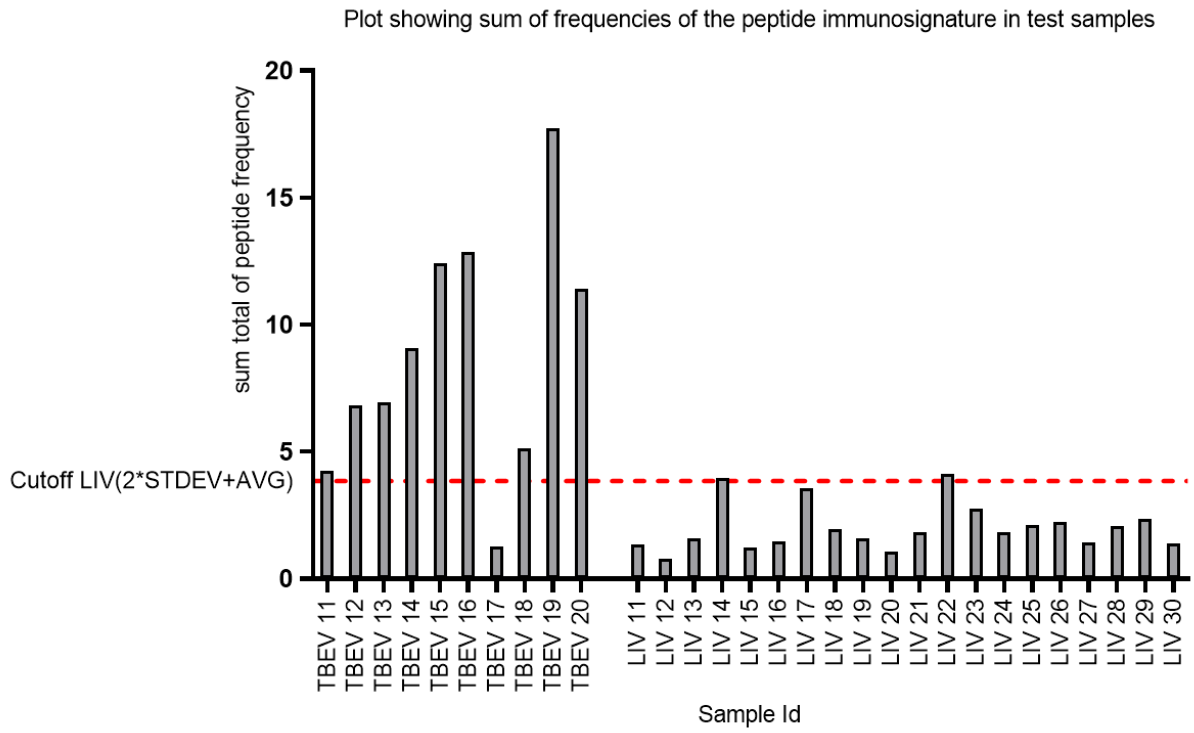


Figure 5-9: Total peptide frequency to calculate success of immunosignature analysis. Cut off values set at 3.85. TBEV 17 was below the cut off value of 3.85 and LIV 14 and LIV 22 were above the cut off.

5.4 Discussion

This chapter showcases the outcomes of two biopanning experiments where NGPD and machine learning were used to identify data patterns known as peptide immunosignatures unique to TBEV and LIV, two very closely related flaviviruses.

In the LIV vs TBEV experiment, a pre-existing phage sub-library enriched with LIV IgG binders was used as the input library for one round of biopanning. This sub-library was subtractively biopanned with pooled TBEV sera and then screened with a panel of 30 LIV IgG sera and 20 TBEV sera.

In the TBEV vs LIV biopanning experiment, the naïve P8 phage library was biopanned over three rounds with the same panel of TBEV and LIV IgG positive sera used in the LIV vs TBEV experiment to enrich for peptide mimotopes to TBEV IgG.

The output libraries from the two experiments were sequenced on the Ion Torrent sequencing platform. A subset of the NGS data was then processed using the machine learning method of data training to define a virus-specific peptide immunosignature. The ability of the peptide immunosignature to act as an *in-silico* assay was tested using the remaining samples. To generate the immunosignature, 10 samples from the target and control were used. To test the immunosignature, the remaining 10 TBEV and 20 LIV samples were used. The TBEV test samples were the control set in the LIV immunosignature analysis and *vice versa*.

Three methods were tested for analysing the outcome of the immunosignature, with the binomial regression analysis being the best out of the three methods tested.

5.4.1 TBEV serum sample selection

The Norwegian sera were not tested for LIV IgG as the risk of LIV exposure was highly unlikely as they were experimentally infected with TBEV as part of a research study (Paulsen *et al.*, 2019). Samples chosen for biopanning were tested on the Progen IgG ELISA to confirm IgG status. Day 14 and Day 31 samples from the same five animals were provided, but only one sample per animal was chosen for the biopanning panel. Biopanning samples TBEV 1, 2, 11 and 12 were from animals inoculated with TBEV only, whereas samples TBEV 3, 4, 5, 13 and 14 were from animals which were also infected with *A. phagocytophlium*. The Norwegian study had reported that there was an increased in TBEV antibody response in animals that were co-infected with TBEV and *A. phagocytophlium.* Information on which study group the sera samples originated from was not provided, but cross-checking with the published data, samples 4, 5, 13 and 14 most likely originated from the co-infected group.

The sheep sera that originated from Estonia were tested for the presence of LIV IgG as they were field samples. There was a correlation noted that those samples which subsequently tested positive for TBEV IgG on the Progen ELISA had cross-reactivity with the LIV HI, with LIV IgG titre ≤1:20. There were 6 samples where LIV IgG titre was ≥1:40. Sera that tested positive on the Progen TBEV IgG ELISA and the LIV HI IgG titre was ≤1:20 (n=11) were picked for the biopanning.

5.4.2 Biopanning strategy

The LIV biopanning experiment used a phage sub-library already enriched for LIV mimotopes. The rationale of picking an already enriched library was that the peptides would be specific for LIV and one round of subtractive panning and screening would be enough to identify the LIV specific immunosignature.

The TBEV biopanning strategy used the naïve P8 peptide library as the starting point. Three rounds of positive panning for TBEV and one round of subtractive biopanning with pooled LIV samples (n=10) in round 2 was carried out.

The strategy used for enriching LIV specific peptide mimotopes were not as successful as the strategy used for TBEV. The subtraction step to remove cross-reactive peptides was in round 2 in the TBEV specific experiments while in the LIV specific experiments this subtraction did not occur till the 3rd round, which was also the final round. By the 3rd round, the diversity of peptides in the phage sub-library would have already been established. This difference of experimental outcome was not anticipated, and the results of the experiments will help in planning better biopanning strategies going forward.

5.4.3 Data analysis of NGS data

The NGS data sets from the biopanning experiments were analysed using the Immunosignaturing analysis introduced in chapter 3. The immunosignature was trained using a subset of the NGS dataset. The immunosignature peptide set was then tested against the remaining target and control samples.

5.4.3.1 Generating the peptide immunosignature

This chapter wanted to test the possibility of generating a pathogen specific peptide immunosignature from NGPD data for use in an *in-silico* assay originally developed from data generated from peptide array experiments (Navalkar *et al.*, 2014; Stafford *et al.*, 2014).

In this chapter the immunosignature training set used the top 250 peptides from each file: LIV (LIV 1—10) and TBEV (TBEV1—10) which then generated the two lists of peptides which were further filtered to make LIV specific or TBEV specific immunosignature peptide list. Here,

the LIV immunosignature was comprised of 405 peptides and the TBEV immunosignature was comprised over 1000 peptides.

An improvement to the pipeline script was written to only allow binning of 16-mer peptides to the barcode files used in the training and testing of data. This allowed for a quicker elimination of nonsense peptides from the files. In this chapter, more peptides per barcode were used for the training (top 250) and testing (top 400) compared to that used in the ZIKV immunosignature training and testing. The numbers were increased to capture as large a sample size as possible.

5.4.3.2 Developing a method to analyse the peptide immunosignature.

Heatmaps:

The outcome of the peptide immunosignature was initially presented as a heatmap with the baseline set on the mean peptide frequency plus 2xSD of the comparator. Although this was a visual representation of the results, it did not provide depth of results and the outcome was not quantitative. This was a useful tool to see whether there were peptides enriched in just the target compared to the comparator and *vice-versa*.

Quantitative *in-silico* immunosignature assay:

The normalised frequency of occurrence of the peptides in the immunosignature test group was used to calculate a positive/negative cut off and to measure the sensitivity and specificity of the immunosignature. The sum of the normalised peptide frequencies of all the peptides common to the peptide immunosignature for each test sample was calculated. Ideally this value will be higher in the target than in the comparator samples.

Binomial regression analysis on GraphPad Prism was the statistical analysis method used to assess the immunosignature. ROC plots are used

to evaluate the performance of diagnostic assays and the AUC provides information on the sensitivity and specificity of the assay (Mandrekar, 2010). GraphPad Prism generates a ROC curve as part of the binomial regression analysis. The software used a subset of the immunosignature test data to generate a ROC plot by assigning a value of one to the target and zero to the control variables. It then curve fitted the remaining unassigned samples to the ROC plot as a 'test' of the immunosignature. This allowed for the outcome of the assay to be truly independent.

In both TBEV and LIV immunosignature *in-silico* assays, only LIV 21-30 can be considered 'true test' as these samples did not contribute to the biopanning process. This meant only specificity could be calculated for the LIV immunosignature and sensitivity for the TBEV immunosignature. The AUC for the LIV immunosignature was 0.780, which is considered acceptable (Mandrekar, 2010). The LIV immunosignature had a sensitivity of 40%, which is poor. This correlates well with the lower AUC value.

The AUC for the TBEV immunosignature was 0.9, a value considered outstanding. The TBEV immunosignature was 90% specific in the small sample set it was tested against. Only 1 of the 10 LIV samples tested positive (LIV 22). This sample was part of the independent sera panel used for the LIV phage ELISA and was the only one that was not detected by any of the phage clones. This sample would be a candidate for testing for TBEV IgG.

The impact of LIV 22 on the biopanning experiments can be considered minimal as subtractive panning was built into the biopanning strategy. The sample was drawn from the animal prior to TBEV being discovered in the UK. If the biopanning experiments were to be repeated, the serum samples in the LIV panel would be tested for the presence of TBEV IgG to consider the current status of TBEV in the UK.

Cut off based on mean of the control plus 2XSD. Another cut off method tested using the sum of peptide frequencies was to set the

positive/negative cut off on the mean of the normalised peptide frequency plus 2*SD of the control. The LIV immunosignature results saw 4 of the 20 LIV samples testing positive and in the case of the TBEV immunosignature, 9/10 TBEV samples tested positive and 2 of the LIV samples tested positive. The results looked similar to the outcome of the binomial regression analysis, but the sensitivity dropped in the TBEV immunosignature *in-silico* assay.

The outcome of the different methods of analysing the immunosignature is that the binomial regression analysis on GraphPad Prism software was the best method of the three methods tested in this chapter. It provided the best outcome for the immunosignatures and that it is generated independent to any manipulation shows that this would be the best method to use.

One of the benefits of the immunosignature analysis is that the data training can be repeated with changes to the number of peptides per barcode or the number of samples used in training the immunosignature. This is the equivalent of selecting the best clones during the LIV phage ELISA development in chapter 4. The list of immunosignature peptides that come through during the training stages might be different, but this is not a disadvantage due to the two stage immunosignature testing process.

The immunosignature(s) will then ideally be tested as an *in-silico* assay, with a set of test samples which has contributed to the biopanning rounds, like samples LIV 11-20 and TBEV 11-20. This part provides the AUC values. The peptide immunosignature that has the best AUC values will have the best sensitivity and specificity value. This peptide immunosignature is then tested against a 'true test' set of samples that have ideally not have contributed to the generation of the sub-libraries in the panning rounds, like samples LIV 21-30. These samples will only have been screened in the final round of biopanning.

5.4.4 Conclusions and future work

In conclusion, the outcomes of this chapter are that the biopanning strategy for enriching peptide mimotopes to TBEV IgG was successful. An immunosignature for TBEV IgG was enriched from NGS sequencing the phage sub-libraries. A method to analyse the immunosignature assay's viability as an *in-silico* assay has been developed and statistical analysis using binomial regression analysis provided an AUC value in the outstanding range. The TBEV peptide immunosignature was tested as an *in-silico* assay against a small number of LIV sub-libraries and was 90% specific.

The LIV immunosignature on the other hand was not as successful, and this is attributed to having used a starting phage library that was already skewed in a certain direction. This impacted the LIV immunosignature and the specificity was 40% with an AUC value at 'acceptable'. Due to the similarity between LIV and TBEV, it can be hypothesised from the results presented here that the immunosignature should have a minimum AUC value of 90%. To improve the LIV immunosignature, the data sets should be re-analysed initially with a smaller peptide depth in the training and testing to see if this improves on the results reported here. Repeating this experiment with a different biopanning strategy might be the other alternative.

The LIV 22 serum sample is a candidate for testing for TBEV IgG due to the result seen in this chapter and its performance in the LIV phage ELISA.

Chapter 6 General discussion and conclusion

This thesis explored the use of traditional phage display and NGPD to develop virus specific diagnostic assays for differentiating between IgG raised against closely related flaviviruses including in co-infection samples. A random peptide phage library expressing 16mer peptides on the major coat protein of the phage was used (Tsoumpeli *et al.*, 2022) and biopanned against immobilised serum IgG. The phage library is introduced to the target attached to a scaffold. A series of washes remove unbound phage, and phage that bind to the target are rescued and in traditional phage display, the sub-library enriched with binders to the target are tested in functional assays such as ELISA. Following the screens, the clones with the best binding ability to the targets were identified using Sanger sequencing. Chapter 4 used this method for the development of a phage ELISA to detect LIV IgG from ovine serum. The LIV IgG sera used in this experiment were clinical samples taken from animals exposed to the virus naturally and the control panel of LIV negative sera came from captive animals with no known exposure to LIV naturally. This experiment was carried out before it was known that TBEV, the closest relative of LIV circulates in the UK (Holding *et al.*, 2019).

Complex experiments such as those outlined in chapters 3 and 5, where virus specific peptide mimotopes were identified from cross-reactive species would not have been easy to identify peptides enriched between the closely related flaviviruses with traditional phage display.

The outputs from this thesis include the development of three immunosignature assays. Firstly, a ZIKV immunosignature was identified which was able to discriminate between ZIKV and DENV in human co-infection samples. A TBEV immunosignature and a LIV immunosignature were also developed following biopanning experiments using ovine sera. Additionally, a novel phage ELISA for detecting LIV IgG positive ovine

sera has also been described in this thesis. Although the assays are still in a developmental phase, the outcomes show that phage display technology has the potential for use as a method to discriminate between closely related flaviviruses.

6.1 Outcomes of the project

The three outcomes of focus in this thesis are, (i) biopanning a random peptide library with flavivirus serum antibodies; (ii) developing a peptide ELISA for detecting LIV IgG from ovine sera and (iii) developing an *in-silico* assay using immunosignaturing analysis to discover virus specific peptide mimotopes using serum positive for flavivirus IgG.

6.1.1 Biopanning methodology for random peptide phage libraries using serum antibodies.

Phage library

The experiments in this thesis used a random peptide library expressing 16-mer peptides on the P8 protein of the bacteriophage (Tsoumpeli *et al.*, 2022). The alternatives to this approach would be to use a targeted library such as a flavivirus specific library (Shriver-Lake *et al.*, 2018; Rajan *et al.*, 2021).

In the first NGPD experiment (Chapter 3), four rounds of biopanning were carried out with the intention to enrich for more ZIKV binders. However, as clones in the phage library are not proportionately represented due to differences in growth and binding kinetics, the clones of interest can be lost (Kuzmicheva *et al.*, 2009). In the ensuing biopanning experiments, the priority was given to maintaining the diversity of clones. This was important especially in experiments outlined in chapter 5, with LIV and TBEV being such close relatives, the peptides specific to the target would already be much smaller than that of the sub-library enriched in chapter 4 (target being LIV IgG positive sera and control being LIV negative sera).

Subtractive biopanning

In the ZIKV biopanning experiment, all four rounds of biopanning incorporated subtractive biopanning to DENV IgG. In chapter 5, only one round of subtractive biopanning with the control sera was carried out for both the LIV and TBEV biopanning experiments.

Due to the time pressure caused by the impact of the Covid-19 pandemic, the biopanning strategy used to enrich for LIV specific peptide mimotopes used a twice enriched LIV sub-library as the starting library instead of the unbiased P8 random peptide library. In this experiment, the subtractive panning to the comparator (TBEV) sera was employed in the final round. The relatively poor performance of the LIV immunosignature compared to that of the ZIKV and TBEV immunosignatures stress the importance of the (subtractive) biopanning strategy.

Serum IgG as biopanning target

The biopanning targets used in this thesis were serum IgG bound to protein G agarose beads. An alternative to this method would be to use a specific protein such as flavivirus NS1 (Lebani *et al.*, 2017). The pros of using serum is the ability to probe multiple antigenic regions compared to just a single antigen if a specific protein were to be used. There are also cons of using serum IgG, these include ensuring that the levels of IgG are detectable and ensuring that the right biopanning strategy is used especially as serum IgG is polyclonal. There is evidence that IgG levels can remain low and undetectable until day 10 (Musso and Desprès, 2020). If using serum IgG as biopanning target, a large panel as possible of sera should be used, so that the output library is not skewed to one sample or a non-specific target, e.g., sample used in the biopanning was positive for a yet-unknown pathogen closely related to the target being biopanned against.

The human sera used in enriching for ZIKV immunosignature in chapter 3 and the Norwegian ovine sera for TBEV were identified molecularly by PCR. However, as IgG is typically produced later in infection after viraemia has passed the window for detecting IgG in PCR positive samples is small. As mentioned above, the biopanning process relies on the availability of the target antibody for the phage to interact with and this interaction leads to the identification of peptide mimotopes to the target antibody. Therefore, an ELISA screening step was introduced prior to NGPD to identify samples that have detectable levels of IgG prior to use in biopanning.

The human sera panel used in the ZIKV vs DENV experiment were collected during the ZIKV outbreak in the Caribbean during 2017. This region is endemic for other mosquito-borne flaviviruses such as WNV, YFV and other mosquito-borne arboviruses such as Chikungunya virus (CHIKV) (Mota *et al.*, 2021). An unknown factor of the sera was the number of times the individual had contracted DENV, and which serotypes they were exposed to as it has been reported that people living in endemic regions can become re-infected multiple times with homotypic DENV (Forshey, Stoddard and Morrison, 2016). Other epitope/mimotope discovery studies on ZIKV/DENV have used a more focused sample set (e.g., pregnant women), conducted longitudinal studies on individuals living in endemic regions or have been conducted with larger sample sets sampled from different geographical locations (Grifoni *et al.*, 2017; Montoya *et al.*, 2018; Kam *et al.*, 2019). An improvement that would be made for the serum selection in Chapter 3 would be a pan-flavi and CHIKV serology screen to understand the polyclonal antibody diversity in the samples. This was not done here due to sample volume, cost and time limitations.

The LIV serum used in chapter 4 and 5 were the same and these samples were submitted to the MRI (Edinburgh) and had tested positive for LIV IgG by haemagglutination. As mentioned earlier, the presence of TBEV in

the UK was not known when the biopanning was done. The LIV positive sera were not tested using a TBEV ELISA and there had been no previous reports of TBEV sero-positivity in ovine sera on the sample collection date.

The TBEV IgG positive sera provided by collaborators in Estonia were tested on LIV HAI and those that had a score of 1:20 or below were used in the biopanning. The cross-reactivity seen in the LIV HAI was used as a confirmatory result for the presence of TBEV and the samples were further tested on the Progen TBEV IgG assay to confirm antibody levels. The Norwegian samples were not tested for LIV IgG levels as these were drawn from animals experimentally infected with TBEV (Paulsen *et al.*, 2019) and therefore were only tested for TBEV IgG levels.

The NGS data analysis that was developed in this thesis is reported in its own section 6.1.3.

6.1.2 Development of a phage ELISA to detect LIV IgG from ovine sera

The current repertoire of serology assays used for detecting LIV does not include an ELISA. Although the current serological assays used for LIV are good assays, some require niche reagents and others are carried out at high containment due to the use of live virus in neutralisation assays. An ELISA using synthetically produced antigen could be carried out at lower containment and allow higher throughput testing of samples. This assay would be useful in studying the seroprevalence of LIV and could also be developed into a diagnostic assay. A proof of concept phage ELISA capable of detecting LIV IgG from ovine sera has been developed in this thesis.

LIV IgG positive and negative serum was biopanned with the P8 peptide phage library for two rounds of biopanning. Eighty phage clones were picked and grown from a dilution plate of the round 2 output phage sub

library. The corresponding peptide sequences for the clones were identified by Sanger sequencing.

In the early developmental ELISA experiments, pooled LIV IgG positive and negative serum used in the biopanning rounds were tested with twenty clones to mark assay parameters. The finalised method was optimised to use neat phage as coating antigen and the incubation steps were carried out at 37°C. The samples were all tested in duplicate to ensure repeatability (intra-assay precision). Four different controls were used provide assurance that the assay was performing adequately and to standardise the results obtained. The positive/negative cut off was calculated on control 1 (pooled LIV negative serum tested with M13KO7) and the positive/negative cut off was set on 3*SD plus mean OD of the control 1.

Once the method was finalised, all the phage clones were tested in duplicate against the pooled serum panels. Twenty one phage clones progressed to the second round of screening where they were tested against the individual members of the serum pools. The positive/negative cut off to progress to the next screening round was relaxed to 2*SD plus mean OD of control 1 of the LIV negative pool. Seven clones progressed to the third round of ELISA screening where each clone was tested against a panel of 40 sera (20 LIV positive and 20 negative sera) not used in the biopanning rounds. The ELISA results were analysed based on each peptide (singleplex) initially and as a combined outcome (multiplex).

The ELISA absorbance values of the singleplex peptides clones tested with sera independent to the biopanning process was analysed for statistical significance using ROC plots. ROC plots provide information on the sensitivity and specificity of diagnostic assays. The area under the curve (AUC) gives a numerical value which signifies the performance of the assay. An assay with very high sensitivity and good specificity would have an AUC value close to 1. The highest AUC value 0.92 was for clone 50

which is 'outstanding' (Mandrekar, 2010) and the next best AUC values were for clones 40 (AUC 0.88), clone 22 (AUC 0.86), clone 61 (AUC0.83) and clone 55 (AUC 0.81). Clones 13 (AUC 0.70) and 16 (AUC 0.71) had the lowest values. The P values generated for the ROC plots of all the clones were <0.05. Clones 22, 40 and 50 had the lowest p values (<0.0001).

The sensitivity and specificity values for the individual clones were calculated independent to the ROC analysis. Some clones had very high sensitivity values but poor specificity values, eg: clones 22 and 40 where the sensitivity was 85% but the specificity values were 60–65%. Other clones had higher specificity value of 90% but sensitivity values of 20 and 55% for clones 13 and 55 respectively. The clone with the best overall sensitivity (90%) and specificity (85%) values was clone 50. The AUC values generated by the ROC plots matched the results of the sensitivity and specificity calculations. The ROC analysis is a useful measure of ELISA performance and should be used in conjunction with the independent calculations of sensitivity/specificity.

The breakdown of ELISA results by individual members of the independent serum panels revealed that by combining multiple peptides, the sensitivity and specificity of the LIV ELISA could be improved. A higher and lower cut off was calculated to predict the best panel of peptides for a multiplex ELISA. One of the LIV negative serum samples (LIV neg 35) tested positive with all seven clones.

At the higher cut off and after masking LIV neg 35, two panels of peptides had a higher sensitivity and specificity values compared to when just the best performing clone by itself (clone 50). For the combination of clones 16, 40, 50, 55 and 61, the sensitivity increased to 95% and the specificity was 78%. For the second combination of clones (22, 40, 50, 55 and 61), the combined specificity was 90% but the specificity increased to 84%.

Both peptide panels have the potential for use as a diagnostic ELISA and as a serosurveillance tool to study the true seroprevalence of LIV (and potentially TBEV) in the UK. Samples that test positive can then be tested on the current gold-standard LIV serology to confirm positivity. The next step in the further development of this ELISA is to produce the peptides synthetically preferably as a polypeptide of the best performing clones in ELISA. As TBEV is also circulating in the UK now, the phage clones should be screened against sera positive for TBEV to understand the extent of cross-reactivity. Ideally this ELISA will be optimised for LIV specificity using TBEV sera.

6.1.3 Development of an *in-silico* assay using pathogen specific immunosignatures

Immunosignaturing assays are a novel diagnostic tools originally developed to map epitopes from diluted blood samples using random peptide arrays (Stafford *et al.*, 2014). The principle behind the immunosignature approach is that by interacting the serum antibody with the random peptide array/library it is possible to interrogate the whole antibody repertoire presented in the sample and identify the weak and strong interactions. A fluorescent labelled secondary antibody is used to probe for antibodies bound to the peptide array slides and downstream processing of the slides yields a pattern of data specific for the target antibody over the other antibodies present in blood. The microarray data generated from the peptide arrays is analysed using machine learning concepts of training and testing data. The data training is run on a subset of the target and comparator files. A unique list of peptides to the target pathology is generated and any that appear in the control data files are removed. The immunosignature is then tested against the remaining target and comparator files and the sensitivity and specificity of the immunosignature is assessed.

This method has been applied to accurately detect biomarkers to multiple cancers (Stafford *et al.*, 2014), conditions such as Alzheimer's disease, and detect between humoral response to vaccine and natural infection for influenza (Legutki *et al.*, 2010). This method has been described to be more sensitive than ELISA as it looks at the weak interactions as well as the strong interactions. in this thesis, the principles of peptide array immunosignature assay was used to analyse the NGS data from the biopanning experiments to identify the peptide mimotopes that are specific to the target.

Immunosignaturing has not been previously published as a method of analysing data generated from NGPD experiments. In this thesis, the data generated from the NGPD experiments in chapter 3 and 5 were analysed using data training and testing to identify peptide mimotopes to the target over the comparator. A previously published data analysis pipeline developed by the Gough group for analysing Ion Torrent data using Z-score analysis was used in the initial processing of the NGS data files.

Unlike the Z-score analysis method of data analysis used by the group, immunosignaturing analysis only uses a subset of peptides in both the training and testing parts of the assay. Advanced machine learning algorithms have been successfully employed for epitope mapping by just using a subpopulation of sequenced data, e.g., the top 50/100/200 peptides arranged by frequency (Luštrek *et al.*, 2013). ROC curves and the AUC values are used to measure the sensitivity and specificity of the immunosignature and following this *in-silico* assay enables a faster probing of a large data set efficiently and the peptides in the immunosignature can be synthesised for use in a physical assay.

Data training and testing

Data training allows for flexibility in the depth of sequences used to develop the immunosignature and in this thesis two different sequence depths were used. In chapter 3, the top 50 peptides from each target and

comparator samples from the data training subset were used to train the ZIKV immunosignature. Once the ZIKV immunosignature peptide list was generated, it was interrogated for sensitivity and specificity against the top 200 peptides from the data testing subset of target and comparator files. In the LIV and TBEV immunosignatures, the training cohort used the top 250 peptides and the testing cohort used top 400 peptides. The use of a larger peptide count did not have an impact on the sensitivity or specificity of the LIV immunosignature. For future experiments it would be beneficial to generate multiple immunosignatures using a range of peptide depths and use ROC analysis to evaluate the best one before testing on an independent serum panel. The sum of peptide frequencies common to the immunosignature was used for the quantitative analysis to assess the immunosignatures.

Heatmaps were used as an early indicator of the performance of the immunosignature. The drawbacks of this method are that it is semi-quantitative and not traditionally used for assessing sensitivity and specificity of a diagnostic assays.

Binomial regression and ROC analysis

Binomial regression is one of the statistical methods of analysing immunosignature data (Brown *et al.*, 2011). This method allows for dichotomous outcome and in the data training, A subset of the samples in the test cohort 10 target and 10 comparator samples were used to generate a ROC curve. ROC analysis is used in measuring performance of serological assays such as ELISA and this was the statistical method used to evaluate the immunosignature assay in this thesis. The area under the curve (AUC) calculation gives a clear indication of the sensitivity and specificity of the assay. The targets were assigned a value of 1 and the comparators were assigned a value of 0. Binomial regression analysis on GraphPad Prism was used to assess the statistical significance of the

results. The software calculates a positive/negative cut off and fits any remaining test samples to the curve it has plotted.

The AUC for the ZIKV immunosignature ROC curve was 0.944 and this shows that the ZIKV immunosignature has very good sensitivity in detecting true positives and specificity in distinguishing them from the DENV samples. The 95% Confidence Interval (CI) values for the ZIKV immunosignature was 0.772-1.0 and the P value for the ROC curve was 0.001. The remaining samples were fitted to the ROC curve and the sensitivity and specificity values for the immunosignature were 82.1% and 69% respectively. Following the ZIKV immunosignature development, improvements were made to the data capture pipeline included a data capture filter which removed peptides below or above the expected peptide size for this library.

In chapter 5, an LIV immunosignature was developed in a similar method and the results from the immunosignature test was analysed in the same way for the ZIKV immunosignature. The AUC for the ROC curve was 0.780, CI values were 0.569-0.991 and the P value for the ROC analysis was 0.03. The TBEV immunosignature had an AUC of 0.93, CI values of 0.794-1.0 and P value of 0.001. The limited number of TBEV samples meant that it was only partially possible to test the immunosignatures against independent sera not used in the biopanning. As a result, the sensitivity and specificity of the immunosignature to the independent sera were not calculated.

This thesis has provided evidence for the possibility of using immunosiganturing and *in-silico* analysis to identify a panel of peptide mimotopes which have high discriminatory power to distinguish between closely related viruses in the case of the TBEV immunosignature. The next steps for the ZIKV and TBEV immunosignatures is to re-train with a range of peptide depths and ascertain the immunosignature with the best AUC

values. At this stage the peptides can be synthesised and tested in immunoassay to assess the biophysical properties.

6.2 Future work and conclusions

Immunosignature assay

The TBEV and ZIKV immunosignatures will be refined further and AUC values will be used to assess the best immunosignature panel for the targets. The replicability of the TBEV immunosignature will be assessed by repeating the screening process and additional convalescent TBEV sera not used in the biopanning will be included as a true test set for the *in-silico* assay. At this point, a decision can be made to synthesise and conduct biophysical assays on the immunosignature peptides from the best *in-silico* assay. The peptide immunosignatures once refined have the potential to be transferred from being *in-silico* to a bench assay in the form of a 'Soluble Phage Array' (SPAr) assay, which can be used for detecting multiple pathogens either for human or veterinary use.

Ion torrent sequencing is not as commonly used as its competitors Illumina and the MinIOn. So, it has become necessary to future proof in case the technology becomes obsolete. The relatively low-cost of the MinION chemistry and its portability as a lab equipment makes it an attractive platform to future proof into. Bridging studies will be conducted to test whether it would be possible to get a similar outcome for the TBEV immunosignature using the MinION platform.

LIV phage ELISA

Further development studies including converting the phage ELISA into a synthetic peptide/multipeptide ELISA and screening with a larger panel of LIV IgG positive sera and also testing for TBEV cross-reactivity. The top peptides enriched in the LIV NGS data sets should also be synthesised and tested for their suitability as an immunoassay candidate. Although

the ELISA was developed for use in ovine sera, it could be adapted for pan-species use and this would be a wonderful resource for serosurveillance studies.

Conclusion

To conclude, the outcomes of this project in identifying virus specific immune responses for a complex problem are a testament to the potential and versatility of the technology. The advent of NGPD and data analysis strategies such as machine learning has changed the landscape of this possibilities associated with this technology.

This thesis also reports the successful application of NGPD to develop *in-silico* assays using immunosignature analysis of the NGS data. Here, two proof of principle *in-silico* immunosignature assays were described, one for ZIKV and the other for TBEV. TBEV specific immunosignature was developed for distinguishing between TBEV and LIV from biopanning with ovine sera. Peptide mimotopes which were ZIKV specific were able to distinguish between ZIKV and DENV IgG from co-infected human serum. The ZIKV immunosignature was tested against an independent set of samples, and this prototype assay demonstrated an acceptable level of sensitivity and specificity.

To put the outcomes of this thesis into the wider context, Coalition for Epidemic Preparedness Innovations (CEPI) has identified the use of *in-silico* methods to study immune responses has been as a key scientific and technological prerequisite in delivering pandemic vaccines in 100 days. This report highlights the use of *in-silico* methods in identifying the early biomarkers of disease and the best immune markers to use or avoid while designing a vaccine.

Appendix 1 Chapter 2 additional information

A.1.1. Phage dilution series

The dilution series used for PEG precipitated phage is shown in table below. The phage dilutions were then mixed with TG1 bacteria to test the titre of the phage.

Table A-1.1: Dilution series used for titrating helper phage

Dilution	PBS buffer (µL)+helper phage (µL)[2]	Final Dilution factor
A	9,990 + 10	1×10^{-4}
B	990 + 10 A	1×10^{-6}
C	990 + 10 B	1×10^{-8}
D	990 + 10 C	1×10^{-10}
E	990 + 10 D	1×10^{-12}
F	990 + 10 E	1×10^{-14}
G	negative	negative

[2]Column 2 shows the serial dilution of helper phage in PBS. In dilution 'A', PEG precipitated helper phage is diluted with 9.990 µL PBS. From each dilution, 20µL was infected into 180µL TG1 and plated on 2YT/*Kan* agar plates

A.1.2. Reagents used for biopanning

Table A-1.2: General reagents used in biopanning

Reagents used in biopanning process		
PEG solution	Dissolve 200g PEG 8000 and 146.1g NaCl in 1L of MilliQ water. Autoclave before use. After autoclaving, the bottle was shaken to ensure the solution was mixed thoroughly to eliminate any crystals being formed. The solution was then stored at 4°C for faster precipitation process.	
25% Glucose solution (500 mL)	125 g D-Glucose Anhydrous was added to 300 mL dH_2O whilst being stirred on a magnetic stirrer and the solution was topped up to 500 mL once the powder dissolved. Solution was filtered through a double compartment vacuum filter flask and stored till use.	Fisher – G/0500/53
1 x PBS	1 tablet in 500 mL water. The solution was autoclaved if sterility was required	Gibco™ 18912014
6x PBS	3 tablets in 250 mL water. Solution was autoclaved if sterility was required	Gibco™ 18912014
18% milk-PBS block	1.8 gm skim milk powder + 10 mL 6x PBS. Solution was mixed thoroughly and made and used on the day	
3% milk-PBS block (M-PBS)	1 mL (18% milk block) + 5 mL 1x PBS. Solution was made and used on the day.	
0.1% PBS-Tween wash buffer	100 mL PBS + 100 µL Tween 20	
20% $MgCl_2$	10 g dissolved in 50 mL dH_2O. Solution was autoclaved before use.	
Freezing media	2YT/*Amp*/Gluc media + glycerol. The glycerol made up 20% of the final volume. Freezing media was made fresh and used on the day.	
Tris buffer (1M, pH 7.5)	6.05 g of Tris base was dissolved in 30 mL of H_2O and pH was adjusted to 7.5 with 5 M HCl. Final volume was adjusted to 50 mL with addition of dH_2O	
Thiamine hydrochloride	100 mg Thiamine HCl dissolved in 10 mL H2O and filter sterilised through 0.45µm filter.	

A.1.3. NGS barcodes

Table A-1.3: Stage 2 PCR barcodes. The common primer is the P1 primer and individual barcodes are assigned to each sample.

Primer name	Sequence (barcode sequence is highlighted as red text and the linker is highlighted in yellow)
P1prim-linker2	CCTCTCTATGGGCAGTCGGTGATCTAGAACATTTCACTTAC
Akey-BC1-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGCTAAGGTAACGTAATCCTTGTGGTATCG
Akey-BC2-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGTAAGGAGAACGTAATCCTTGTGGTATCG
Akey-BC3-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGAAGAGGATTCGTAATCCTTGTGGTATCG
Akey-BC4-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGTACCAAGATCGTAATCCTTGTGGTATCG
Akey-BC5-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGCAGAAGGAACGTAATCCTTGTGGTATCG
Akey-BC6-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGCTGCAAGTTCGTAATCCTTGTGGTATCG
Akey-BC7-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGTTCGTGATTCGTAATCCTTGTGGTATCG
Akey-BC8-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGTTCCGATAACGTAATCCTTGTGGTATCG
Akey-BC9-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGTGAGCGGAACGTAATCCTTGTGGTATCG
Akey-BC10-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGCTGACCGAACGTAATCCTTGTGGTATCG
Akey-BC11-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGTCCTCGAATCGTAATCCTTGTGGTATCG
Akey-BC12-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGTAGGTGGTTCGTAATCCTTGTGGTATCG
Akey-BC13-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGTCTAACGGACGTAATCCTTGTGGTATCG
Akey-BC14-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGTTGGAGTGTCGTAATCCTTGTGGTATCG
Akey-BC15-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGTCTAGAGGTCGTAATCCTTGTGGTATCG
Akey-BC16-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGTCTGGATGACGTAATCCTTGTGGTATCG
Akey-BC17-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGTCTATTCGTCGTAATCCTTGTGGTATCG
Akey-BC18-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGAGGCAATTGCGTAATCCTTGTGGTATCG
Akey-BC19-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGTTAGTCGGACGTAATCCTTGTGGTATCG
Akey-BC20-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGCAGATCCATCGTAATCCTTGTGGTATCG
Akey-BC21-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGTCGCAATTACGTAATCCTTGTGGTATCG
Akey-BC22-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGTTCGAGACGCGTAATCCTTGTGGTATCG
Akey-BC23-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGTGCCACGAACGTAATCCTTGTGGTATCG

Akey-BC24-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGAACCTCATTCGTAATCCTTGTG GTATCG
Akey-BC25-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGCCTGAGATACGTAATCCTTGTG GTATCG
Akey-BC26-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGTTACAACCTCGTAATCCTTGTG GTATCG
Akey-BC27-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGAACCATCCGCGTAATCCTTGTG GTATCG
Akey-BC28-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGATCCGGAATCGTAATCCTTGTG GTATCG
Akey-BC29-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGTCGACCACTCGTAATCCTTGTG GTATCG
Akey-BC30-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGCGAGGTTATCGTAATCCTTGTG GTATCG
Akey-BC31-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGTCCAAGCTGCGTAATCCTTGTG GTATCG
Akey-BC32-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGTCTTACACACGTAATCCTTGTG GTATCG
Akey-BC33-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGTTCTCATTGAACGTAATCCTTG TGGTATCG
Akey-BC34-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGTCGCATCGTTCGTAATCCTTGT GGTATCG
Akey-BC35-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGTAAGCCATTGTCGTAATCCTTG TGGTATCG
Akey-BC36-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGAAGGAATCGTCGTAATCCTTGT GGTATCG
Akey-BC37-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGCTTGAGAATGTCGTAATCCTTG TGGTATCG
Akey-BC38-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGTGGAGGACGGACGTAATCCTT GTGGTATCG
Akey-BC39-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGTAACAATCGGCGTAATCCTTGT GGTATCG
Akey-BC40-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGCTGACATAATCGTAATCCTTGT GGTATCG
Akey-BC41-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGTTCCACTTCGCGTAATCCTTGT GGTATCG
Akey-BC42-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGAGCACGAATCGTAATCCTTGT GGTATCG
Akey-BC43-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGCTTGACACCGCGTAATCCTTGT GGTATCG
Akey-BC44-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGTTGGAGGCCAGCGTAATCCTT GTGGTATCG
Akey-BC45-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGTGGAGCTTCCTCGTAATCCTTG TGGTATCG
Akey-BC46-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGTCAGTCCGAACGTAATCCTTGT GGTATCG
Akey-BC47-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGTAAGGCAACCACGTAATCCTT GTGGTATCG
Akey-BC48-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGTTCTAAGAGACGTAATCCTTGT GGTATCG
Akey-BC49-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGTCCTAACATAACGTAATCCTTG TGGTATCG
Akey-BC50-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGCGGACAATGGCGTAATCCTTG TGGTATCG
Akey-BC51-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGTTGAGCCTATTCGTAATCCTTG TGGTATCG

Akey-BC52-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGCCGCATGGAACGTAATCCTTG TGGTATCG
Akey-BC53-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGCTGGCAATCCTCGTAATCCTTG TGGTATCG
Akey-BC54-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGCCGGAGAATCGCGTAATCCTT GTGGTATCG
Akey-BC55-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGTCCACCTCCTCGTAATCCTTGT GGTATCG
Akey-BC56-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGCAGCATTAATTCGTAATCCTTG TGGTATCG
Akey-BC57-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGTCTGGCAACGGCGTAATCCTT GTGGTATCG
Akey-BC58-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGTCCTAGAACACGTAATCCTTGT GGTATCG
Akey-BC59-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGTCCTTGATGTTCGTAATCCTTG TGGTATCG
Akey-BC60-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGTCTAGCTCTTCGTAATCCTTGT GGTATCG
Akey-BC61-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGTCACTCGGATCGTAATCCTTGT GGTATCG
Akey-BC62-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGTTCCTGCTTCACGTAATCCTTG TGGTATCG
Akey-BC63-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGCCTTAGAGTTCGTAATCCTTGT GGTATCG
Akey-BC64-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGCTGAGTTCCGACGTAATCCTTG TGGTATCG
Akey-BC65-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGTCCTGGCACATCGTAATCCTTG TGGTATCG
Akey-BC66-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGCCGCAATCATCGTAATCCTTGT GGTATCG
Akey-BC67-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGTTCCTACCAGTCGTAATCCTTG TGGTATCG
Akey-BC68-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGTCAAGAAGTTCGTAATCCTTGT GGTATCG
Akey-BC69-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGTTCAATTGGCGTAATCCTTGTG GTATCG
Akey-BC70-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGCCTACTGGTCGTAATCCTTGTG GTATCG
Akey-BC71-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGTGAGGCTCCGACGTAATCCTT GTGGTATCG
Akey-BC72-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGCGAAGGCCACACGTAATCCTT GTGGTATCG
Akey-BC73-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGTCTGCCTGTCGTAATCCTTGTG GTATCG
Akey-BC74-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGCGATCGGTTCGTAATCCTTGTG GTATCG
Akey-BC75-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGTCAGGAATACGTAATCCTTGTG GTATCG
Akey-BC76-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGCGGAAGAACCTCGTAATCCTT GTGGTATCG
Akey-BC77-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGCGAAGCGATTCGTAATCCTTGT GGTATCG
Akey-BC78-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGCAGCCAATTCTCGTAATCCTTG TGGTATCG
Akey-BC79-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGCCTGGTTGTCGTAATCCTTGTG GTATCG

Akey-BC80-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGTCGAAGGCAGGCGTAATCCTTGTGGTATCG
Akey-BC81-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGCCTGCCATTCGCGTAATCCTTGTGGTATCG
Akey-BC82-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGTTGGCATCTCGTAATCCTTGTGGTATCG
Akey-BC83-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGCTAGGACATTCGTAATCCTTGTGGTATCG
Akey-BC84-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGCTTCCATAACGTAATCCTTGTGGTATCG
Akey-BC85-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGCCAGCCTCAACGTAATCCTTGTGGTATCG
Akey-BC86-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGCTTGGTTATTCGTAATCCTTGTGGTATCG
Akey-BC87-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGTTGGCTGGACGTAATCCTTGTGGTATCG
Akey-BC88-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGCCGAACACTTCGTAATCCTTGTGGTATCG
Akey-BC89-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGTCCTGAATCTCGTAATCCTTGTGGTATCG
Akey-BC90-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGCTAACCACGGCGTAATCCTTGTGGTATCG
Akey-BC91-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGCGGAAGGATGCGTAATCCTTGTGGTATCG
Akey-BC92-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGCTAGGAACCGCGTAATCCTTGTGGTATCG
Akey-BC93-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGCTTGTCCAATCGTAATCCTTGTGGTATCG
Akey-BC94-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGTCCGACAAGCGTAATCCTTGTGGTATCG
Akey-BC95-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGCGGACAGATCGTAATCCTTGTGGTATCG
Akey-BC96-lnk1	CCATCTCATCCCTGCGTGTCTCCGACTCAGTTAAGCGGTCGTAATCCTTGTGGTATCG

A.1.4. NGS data analysis

Linux programming uses a command line interface (CLI) which allows the user to type in computing syntax known as 'scripts' to access various programs to process the data. Scripts can be written in a text editor and saved as executable files with the extension .sh; this speeds up data processing if the data 'infile' and 'outfile' locations are specified for each stage of the data processing. Scripting enables the end user to stack several processes which is parsed through by the computer and the output file from one processed is used as input for the next process and this automates the data analysis. Linux terminology used is shown in Table A 1.4

Table A-1.4: List of general Linux terminology used

Terminology	Action taken by software or description of software
Ubuntu	computer operating system powered by Linux
Perl	programming language
Command line interface	commands to the computer program are processed in the form of lines of text.
Terminal	Interface on Ubuntu where user types in
Demultiplex	Finds the barcodes from FASTA files
Translate.pl	Perl script to translate the sequences
<file	This prompt informs the software that this is the input file for the executable command.
>file	This prompt followed by desired output file name saves the results into that file rather than printing the data to the screen. If adding more data to an existing file, >> should be used to prevent overwrite of data
>>file	Create a file if it doesn't exist, otherwise add data to existing file.
.sh	Shell script. This is the Linux equivalent to a .txt or.doc file
Parallel	Processes data from multiple files in parallel
cat (concatemer)	Joins files
grep	Computer program designed to pull string. This allows for identifying the barcode
geany	Text editor for writing long scripts for processing data. Saving the file 'file name.sh' allows the syntax in the file to be run as a shell script
Ls-ltr	To view the results on the screen instead of sending it to an output file
GNU parallel	Data processed faster: eg on an 8processor PC, data is processed 8x faster.

A.1.5. NGS data analysis

The sequenced data file was supplied as zipped FASTQ files with the extension *'gz'*. The file was downloaded, and scripts were used to process the files which included the following processes: conversion of the FASTQ file to FASTA, de-multiplexing the data into individual sample files using the DNA barcodes, translating the DNA files and extracting the peptide sequences.

FASTQ to FASTA conversion

The software 'unpigz' was used to unzip the FASTQ files. The FASTQ files have 4 lines of data and starts with '@'. The FASTQ format contains sequence quality details and requires conversion to FASTA format using script [1], shown below, where **input** is the name of the file that was downloaded from the sequencing facility server and **output** is the name assigned for that batch of files.

```
seqtk seq-a
input.fastq>output.fasta
```

[1]

Demultiplexing files

The FASTA library file was demultiplexed into individual files for each sample using the DNA barcodes as the identifier, an example of the script used is shown below [2]. The script file with all the commands was labelled '*demultiplex.sh* 'and this file read through the *.output.fasta* file, identified the DNA barcode sequence and split it from the rest of the sequence which was saved into the output file named 'split_barcode_BC01.fasta'. The demultiplexed files are named with the barcode number to track each file back to the phage sub-library it originated from.

```
cat file.fasta | grep -E -B 1 '^('.{0,5}DNA barcode' | grep - -v '- -' >
split_barcode BC01.fasta
```
[2]

Once converted to FASTA format (identified by the prefix '>'), each sequencing read was assigned to separate files based on the DNA barcodes used during the library preparation. Standard Ion Torrent barcodes used are listed in table A-1.3.

GNU Parallel for parallel processing of the NGS files

A program named GNU Parallel was used to speed up demultiplexing the files by maximising the processing speed of the computer. It processes multiple individual files through the scripts in *demultiplex.sh* at the same time. GNU Parallel also ensured that the results from each barcode go to their individual output files. The command that ran this part of the processing is shown below [3].

```
Parallel < demultiplex.sh
```
[3]

Translating DNA to peptide sequence

Demultiplexed files were then processed by the next script 'translate.pl', a Perl script that translated the files from DNA format to peptide sequence in all three frames.

The translate.pl script is shown below [4], where each DNA FASTA file is translated in all three reading frames and all three files are concatenated to a single file.

In '***-s opq***', the script allows for identification of the stop codons and incorporates a 'q' when it encounters an amber stop codon (UAG) and '*' for o and p upon encountering them whilst translating the DNA FASTA file.

The '*BC01.all_frames'* file contains 3 copies of sequence identifier, and one will be in the right frame to enable the detection of the 16-mer peptide in the next step.

```
perl translate.pl -i split_barcode_BC01.fasta -s opq -f
1 >BC01.frame1.fasta

perl translate.pl -i split_barcode_BC01.fasta -s opq -f
2 >BC01.frame2.fasta

perl translate.pl -i split_barcode_BC01.fasta -s opq -f
3 >BC01.frame3.fasta

cat BC01.frame1.fasta BC01.frame2.fasta
BC01.frame3.fasta >BC01.all_frames
```

[4]

Identifying the random 16-mer peptide

The 16-mer peptide is presented in the following format 'AEGEF—(X)16—DPAKA', where AEGEF and DPAKA are fixed N and C terminal amino acid sequences from the vector that flanks the randomised peptide '(X)16', which was enriched in the biopanning process. The perl script '*iterate_motifs.pl*' was used for this part. For the perl script to be executed successfully, two specific text files containing the flanking sequences, 'left_motif_list.txt' and 'right_motif_list.txt', need to be in the same directory as the 'BC01.*all_frames*' file which is the input file for this script.

The output peptide files from this step are saved with the extension *.LR.fa* to indicate that the length of the polypeptide included both flanking regions. Aberrant sequences where the left identifier is missing or truncated is saved with extension 'R.fa' and those with missing right identifier is saved with extension 'L.fa'. The files with the extension *.LR.fa* were analysed according to the analysis pipeline below.

The NGS files were named with a prefix of the experiment they originated from and the barcode to track the peptide libraries back to the original phage sub-library it originated from.

Processing multiple libraries

An updated Perl script 'PIPELINE1.05.pl' was applied to later NGS files processed. Using GNUParallel, the process from unzipped FASTQ files to the translated file was sped using script '*parallel<filename_parallel.sh'* [**5**]. Text files containing the barcodes, and the flanking regions and the Perl script were required to be in the same directory for GNU Parallel to process the data.

The script converted the files from FASTQ to FASTA; separated the FASTA files using the barcodes (provided as a text file); processed the barcodes in groups of 12 (provided as text file where the barcodes were batched in groups of 12).

'left AEGEF.txt --right DPAKA.txt' identified the flanking regions in the FASTA file (provided as the text files).

'--minimum 1' filtered the peptide based on the length of peptide expected, in [5] minimum is shown at 1. The peptide length expected from the P8 phage library is a 16-mer.

'--rank --fasta –translate' counted the occurrence of each peptide and generated the FASTA files in all three frames.

```
perl PIPELINE1.05.pl --infile AV_Batch_1.fastq.gz --outfile AV_Batch_1 --barcodes BC01-12.txt --left AEGEF.txt --right DPAKA.txt --minimum 1 --rank --fasta --translate
```
[5]

Z-score analysis scripts

To carry out the comparisons, 'pool' files of all peptides enriched was made for both groups respectively, an example is shown using the scripts [6] shown below.

```
cat target.sample_BC01.frames.fasta.LR.fa
target.sample_BC02.frames.fasta.LR.fa target.sample_
BC03.frames.fasta.LR.....> target.sample_pool.fasta

cat control.sample_BC21.frames.fasta.LR.fa
control.sample_BC22.frames.fasta.LR.fa
control.sample_BC23.frames.fasta.LR.fa... > control.sample_pool.fasta
```

[6]

The individual *target.LR.fa* files were compared against the *control_pool.fasta* files to create a table of peptides seen enriched (higher Z-scores) in the target compared to the control pool.

This allows for the identification of peptides that are favourably enriched in the target cohort of samples. Three scripts [7] were used sequentially for the comparison steps.

The first script compared all the peptides generated from each test sample against all peptides from control samples. The peptides enriched in each target sample compared to the pooled were collated into a table file (*.table*) identified with the barcode associated with the sample.

The second script filtered the data by sorting through the *.table* files and collated another table containing the peptides and their respective Z-scores.

The third script filtered the output data file from the second script and based on the Z-score cut off that was set, in the script shown above the Z-score was set at 5.

```
perl compare.collapsed.sequences
target.sample_BC01.frames.fasta.LR.fa
control.sample_pool_barcodes.fasta > target_sample_BC vs control
pool.table

cat target_sample_BC.vs control pool.table | awk 'NR<2{print
$0;next}{print $0 | "sort -nr -k4"}' > target_sample_BC.vs control
pool.table.sorted

cat target_sample_BC.vs control pool.table.sorted | awk 'NR>1 &&
$4>=5 {print $1, "\t" ,$4}' > target_sample_BC.vs control pool.Z-
score.table
```

[7]

After the Z-score sorting, all the target *.table* files with the chosen Z score were concatenated into a single file which was named '*all_test.samples_BCs_Zscore_vs_comparator_pool*' where all the peptides were ranked by Z-score and the script is shown in [**8**] below.

```
cat ZIKV_01—20_Z5_vs_Denv_pool | sort | uniq -c | sort -nr | sed
's/^[ \t]*//' > ZIKV_01—20_Z5_vs_Denv_pool_ranked_date
```

[8]

Appendix 2 Chapter 3 additional information

The 145 16-mer phage enriched peptides unique to the ZIKV+/DENV- training set of samples. These were the peptides that made up the ZIKV immunosignature

Table A-2.1: 16-mer peptides that formed the ZIKV immunosignature

	Peptide
1	AKVKSLASSRIAESEA
2	ALYTPPGASPNASYNA
3	DGTPISAKQLDLLQRT
4	DKAKQNSSPERL
5	DKHTGLIIAPAQGHLM
6	DLTGSMWKNRPSTPVV
7	DNYQTIFIVNQAQTTS
8	DPYQKPYASQLAVMYQ
9	DRYQIQFTPLQATTLP
10	DTYQLDWAIEPLRTTT
11	ELNSVERDSKNMTSLT
12	ENYIDTILSLPAMHHV
13	ERRKTRPNNTSLELTN
14	FGSTLRYDKTVEENYT
15	FHTTLSYQRQLPPNAC
16	FQTTLRYSNAHELMPT
17	GAIVSDYKNQVMLPKQ
18	GKLIPEFKKSTHTGST
19	GPDNIGKSPWLPSCIP
20	GRPKPPNPKPTLDSCP
21	HALDPASYHNHVFTKQ

22	HGHRKPPNPRPNFPDS
23	HKAEDHTKSGRIQLPL
24	HKLYNPFTQHNQTPSA
25	HKTGSGPQNHTSLLTK
26	HMPPLSAMESTVSWHP
27	HPSGLIDHHVHEVANN
28	HPSGYHVMPTATPYIA
29	HPSGYLGPRKAWISET
30	HPSGYMASAVQIS
31	HPSPIQNSHRWELPLQ
32	HRPPPGASVNSRDAAM
33	HRRKPPNPQILTETIR
34	HTTLTYGKTPVTDQSP
35	IEAQADMQHLDWQRNV
36	ILHSAEKEVKHTPHHS
37	IPEPKPSKVHTLAESS
38	IPIHNLHNPFVTVGEY
39	IPRGARQKPHALPMSA
40	IRHRNAVPSLSDIPRP
41	IWWSSYTEMTWDSMPL
42	KASYTRPHDQGPMQSS
43	KGDRFTSRHTETLQVA
44	KKLNWNTDKLNLAQTP
45	KKPPNPSDPSNRLLQV
46	KPGPITKPIANADNRL
47	KPPRVLKLPDMRSAAT
48	KRDRYSTTVSYPDRTI
49	KTTPSQISDPSPETYP

50	LEIPALTPAEHGGRSQ
51	LKTDGVNFNTPAVPTN
52	LMTKPDPEKVKPWSAG
53	LPAWYQEWLRGMPETN
54	LQHNPFLPLTTIHDVA
55	LQLTKQRPNNASTLPH
56	LRPKEPKPTKDATIDL
57	LRSLDAEHKEHYRSCN
58	MHRKVIRNAPKDTAVA
59	MHSSRSWDNSIAHTPN
60	NKHTGLPRALPIPAQS
61	NKPPNPSPPGPMVIPH
62	NKYVDKSAPNNWNATP
63	NLCSSETEAKSPPYSL
64	NNKTIQYLHTTAVRAE
65	NPRASESRKLVNSATT
66	NPRWIEAGAKPKLPPG
67	NPSGYIPNRLNEPKYS
68	NPTYNFLRKPRSPTSDCSS
69	NRIPTLFTNAGAVHPV
70	NRRPPAGKYSQSEIPP
71	NSLMSLRMATHTATIS
72	NYLPARTAAESGPTAS
73	NYRTKLKEYVHSAPAF
74	PDDIGKTSEITSPGHL
75	PDRGTGLRSAVHHSPA
76	PDYNVLRRSERADPPV
77	PEPSSRASLRDAVKFR

78	PEQGTGRPNIKETLTL
79	PERGTGLAQTHRTVLQ
80	PILYTSTLQYPLPAPE
81	PIPIPNQRPWEPSSPL
82	PLESAELVTKNDQCTS
83	PLIATPKIPTSNYGPE
84	PLPTCIETASCQKNHP
85	PLRSAEKLSKKRFDSS
86	PMLPMLTHAERSTNAI
87	PNNYLALLDAPEEWPI
88	PPGASKNSTASGARTP
89	PPTAMQRTKLNDTATS
90	PRGAKKEATQAPNFLA
91	PRRTPSPHASQILPHY
92	PSLSTSETGAKASSGN
93	PTHFQGPVQIRPCKSY
94	PTKTQPFQMSSPTPQH
95	PTWKAWFRSQSPPPIN
96	PVHKFSYTHGQVAHNP
97	PWHLLYATHPTAQLET
98	PWNAKGLVVAPSTGGM
99	PWSPWPMLEKPSISHS
100	PWTAWPTMQLTVNVPR
101	RGPASRGSAVNATSLP
102	RKQVISTYTLTEPSAR
103	RMKPPNPKQTTYNNGL
104	RNSKLDRLLLPAAVED
105	RPRGASKLDHLPAAAP

106	RSTRQPTVWHNIPAPT
107	RTKPPNPTRPHQIAQI
108	RTPTLFSNTNSAPRWT
109	RTQLKGFDKTVLSDVL
110	RYGPMDNSSLTLTQPV
111	SDDMVQVQYMLTVRDE
112	SKMEYTGNRRTSNFVH
113	SKQSTAHTLNLHAAPS
114	SNSAKSGPNNANSPPV
115	SNTYHSTLRYGTQGKI
116	SPLWWASITREKALPS
117	SPTPMQAPSWAHPTVP
118	SRPSTNPSKTLDTVHA
119	SSHRGSRVSTAAMTPP
120	STTLYYGEDGVPVKLT
121	SWHHRLTNSLPSRTPV
122	TLRDKIQLLLSATAND
123	TMRTWLSAIDRETSAT
124	TNKVQAPQLSSPVYTE
125	TPEPADHAFTILLYSS
126	TPIKRPSSASPIAADI
127	TQVRPPRGARFPGPVP
128	TRRVKGKPLDEPPTGL
129	TRSNSPYAMSLQDEPV
130	TSTLNYGSATPNSTLS
131	TSTSPLPGDSSRLQST
132	TYPPPTSMQSPDPAKI
133	VISPPGRGPAKHPPMG

134	VKDRPNNYGTAKNLQW
135	VKHTLNEVYTMFWSSS
136	VSFDKYRDSTIQAGVS
137	VYKYSLSDTRSSHMTM
138	WNHNPFATLQTTLNYP
139	WQTLGNRPTPRGITNA
140	YPIPLTASERPAVQQL
141	YQPAPMSPGPILTSAT
142	YRKLFRSPPPTPPSET
143	YSTSLQYSTLGTNAKL
144	YSTSLTYGTQYLIGPH
145	YYTATLIHSTNVDSSV

Appendix 3 Chapter 4 additional information

A.3.1. Ovine serum samples provided by MRI

Table A3-1: LIV IgG positive sera provided by MRI. Grey shaded boxes were used in biopanning and in ELISA development study. Yellow shaded boxes were the independent sera used in ELISA for screening with the phage clones.

	Positive Samples	Total Igs	IgG only
1	17/0049	1/2560	1/80
2	17/0224	1/1280	1/160
3	17/0319	1/1280	1/320
4	17/0433	1/2560	1/160
5	17/0504/1	1/5120	1/640
6	17/0663	1/10240	1/320
7	16/0218	1/2560	1/40
8	16/0231	1/1280	1/20
9	16/0289/3	1/1280	1/20
10	16/0358	1/2560	1/320
11	16/0382/1	1/1280	1/80
12	16/0382/2	1/1280	1/160
13	16/0383/1	1/2560	1/160
14	16/0383/2	1/640	1/80
15	16/0409/2	1/2560	1/160
16	16/0472	1/160	1/40
17	16/0562	1/2560	1/640
18	16/0704	1/5120	1/1280
19	16/0741/2	1/640	1/160
20	16/0797	1/320	1/160
21	16/0860	1/2560	1/160
22	15/0347/1	1/1280	1/20

23	15/0356	1/1280	1/80
24	15/0441	1/1280	1/640
25	15/0675	1/5120	1/320
26	14/0561	1/1280	1/320
27	14/0570	1/2560	1/20
28	14/0585	1/640	1/20
29	14/0647/1	1/1280	1/320
30	14/0647/2	1/1280	1/320
31	14/0662	1/640	N (1/10)
32	14/0724/2	1/2560	1/640
33	14/0724/4	1/2560	1/1640
34	14/0829	≥1/10240	1/20
35	14/0850/1	≥1/10240	1/40
36	14/0850/2	≥1/10240	1/30
37	14/0859	1/5120	1/160
38	14/0899	1/1280	1/40
39	14/0916	1/320	1/40
40	14/0936	1/2560	1/80
41	14/0938/1	1/2560	1/160
42	14/0938/2	1/5120	1/640

Table A3-2: LIV IgG negative sera provided by MRI. Grey shaded boxes were used in biopanning and in ELISA development study. Yellow shaded boxes were the independent sera screened with the phage clones.

	LIV IgG -ve samples
1	16/0092/1
2	16/0092/2
3	16/0092/3
4	16/0092/4
5	16/0092/5
6	16/0092/6
7	16/0092/7
8	16/0092/8
9	16/0092/9
10	16/0092/10
11	16/0319/1
12	16/0319/2
13	16/0319/3
14	16/0319/4
15	16/0319/6
16	16/0319/7
17	16/0319/1
18	16/0319/2
19	16/0319/3
20	16/0319/4
21	16/0319/6
22	16/0319/7
23	16/0543/1
24	16/0543/2
25	16/0543/3
26	16/0543/4
27	16/0543/5
28	16/0543/6
29	16/0543/7

30	16/0543/8
31	16/0543/9
32	16/0543/10
33	16/0886/1
34	16/0886/2
35	16/0886/3
36	16/0886/4
37	16/0886/5
38	16/0886/6
39	16/0886/7
40	16/0886/8
41	16/0886/9
42	16/0886/10
43	17/0209/1
44	17/0209/2
45	17/0209/3
46	17/0209/4
47	17/0209/5

A.3.2. Developmental ELISA 4-6

Table A3-3: ELISA absorbance data of PEG precipitated phage clones 1–20 in experiments 4–6. Pooled serum diluted 1:500 was used in all three experiments. Positive/negative cut off was based on control 1 (helper phage tested with LIV negative pool). Plates were read at 450 nm.

phage clone	LIV +ve sera, n°4	LIV-ve sera, n°4	LIV+ve sera, n°5	LIV-ve sera, n°5	LIV+ve sera, n°6	LIV-ve sera, n°6
clone 1	0.826	0.282	2.378	0.706	0.865	0.354
clone 2	0.427	0.553	1.454	1.469	0.79	0.768
clone 3	0.772	0.406	2.56	1.301	1.295	0.779
clone 4	0.804	0.342	2.39	1.026	1.131	0.487
clone 5	1.020	0.524	2.643	1.589	1.533	0.777
clone 6	0.420	0.280	1.387	0.889	0.463	0.351
clone 7	0.521	0.344	1.862	1.138	0.851	0.581
clone 8	0.760	0.432	2.655	1.448	1.224	0.713
clone 9	0.566	0.319	2.119	1.064	0.765	0.509
clone 10	0.740	0.448	1.933	1.291	1.032	0.748
clone 11	0.326	0.334	1.005	0.852	0.375	0.461
clone 12	1.054	0.441	2.464	1.448	1.334	0.563
clone 13	1.366	0.443	2.985	1.358	1.844	0.826
clone 14	0.445	0.457	1.377	1.421	0.542	0.836
clone 15	0.518	0.507	1.734	1.612	0.812	0.681
clone 16	1.015	0.757	2.803	2.238	1.235	0.762
clone 17	0.459	0.367	1.585	1.113	0.522	0.471
clone 18	0.777	0.882	2.08	2.083	0.657	0.921
clone 19	1.656	0.462	2.898	1.204	2.145	0.619
clone 20	0.990	0.398	2.499	1.071	1.289	0.489
helper phage/sera	0.423	0.659	1.398	1.473	0.489	0.719
helper phage/milk block	-0.034	-0.002	-0.022	-0.028	0.489	0.668
PBS/sera	0.136	0.179	0.331	0.307	0.164	0.192
no coating/milk block	-0.020	0.020	-0.031	0.031	-0.018	0.018

A.3.3. Developmental ELISA 7-8

Table A3-4: ELISA absorbance data of PEG precipitated phage clones 21–40 in experiments 7 and 8. Pooled serum diluted 1:500 was used in all three experiments. Plates were read at 450nm within 30 minutes of reaction being stopped. Positive/negative cut off was set to mean+3*STDEV of the OD values for LIV negative serum.

	ELISA 7		ELISA 8	
Phage clone	LIV + pool	LIV - pool	LIV + pool	LIV - pool
Clone 21	2.81	2.71	0.24	0.22
Clone 22	3.59	2.76	0.48	0.17
Clone 23	2.85	2.75	0.36	0.22
Clone 24	2.64	2.23	0.23	0.14
Clone 25	2.47	2.37	0.23	0.22
Clone 26	2.93	2.18	0.35	0.11
Clone 27	2.64	2.43	0.20	0.18
Clone 28	2.55	2.62	0.16	0.20
Clone 29	3.25	2.94	0.26	0.26
Clone 30	3.41	2.99	0.32	0.23
Clone 31	2.88	2.84	0.41	0.32
Clone 32	2.68	2.48	0.54	0.25
Clone 33	2.72	2.44	0.33	0.25
Clone 34	2.41	2.73	0.18	0.23
Clone 35	2.38	2.54	0.14	0.22
Clone 36	2.75	2.85	0.28	0.30
Clone 37	2.83	2.58	0.39	0.24
Clone 38	2.63	2.87	0.21	0.24
Clone 39	2.49	2.69	0.18	0.24
Clone 40	2.65	2.41	0.48	0.18
Cut off		**3.314**		**0.368**

A.3.4. Phage ELISA method

Equipment/materials

- 96 well microtitre plate (Nunc MaxiSorp®)
- Washing buffer: 0.1% Tween 20 in PBS
- Blocking buffer: 3% semi-skimmed milk in 1x PBS
- Stop solution: 1 M H_2SO_4 acid
- Enzyme-conjugated secondary antibody: rabbit anti sheep 1:500 (2uL in 10 mL)
- 3,3',5,5'-Tetramethylbenzidine slow (TMB) substrate (Thermo Scientific 34024)
- 5 ml serological pipettes
- Absorbent tissue
- Discard dish
- Gilson pipettes
- Multichannel pipette
- Spectrophotometer

Method

Coat the 96 well plate with monoclonal or polyclonal phage with 50-100 µL PEG precipitated phage per well in duplicate. Incubate at 4°C overnight

Remove the phage supernatant by flicking the plate over a discard dish lined with 1:100 Distel solution-soaked absorbent tissue.

Wash the plate four times with 300 µL washing buffer per well, flicking the plate to remove the previous wash over the discard dish, between washes. Following the final wash, remove the excess liquid over the discard dish and excess moisture was removed from the plate by thoroughly patting it over a pile of dry tissue.

Block the wells with 200 µL blocking buffer per well for 60 minutes at 37°C.

Wash the plate as previously described.

While the assay plate is being blocked, prepare the test samples by diluting the test serum in block buffer, for pooled serum use 1:500 dilution or for individual sera use 1:200 dilution.

After the incubation step is complete, remove block solution by flicking and pat dry over tissues.

The test serum were added to the plate as per individual experimental plan and left to incubate for 60 minutes at 37°C.

Following the incubation step, flick the plate and wash plate and dry over tissue as described previously.

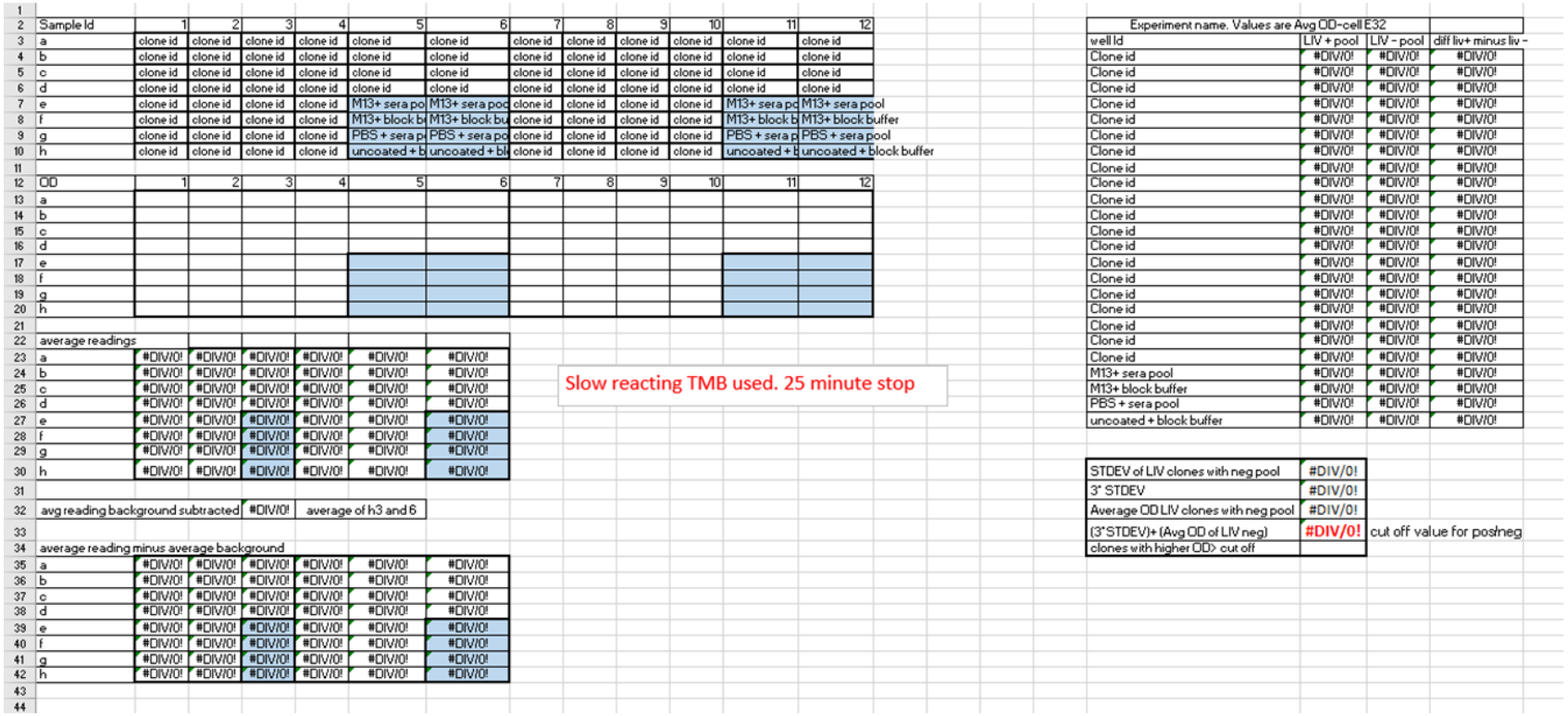
The tagged secondary antibody relevant to the experiment, is added 50-100 µL volume and leave to incubate for 60 minutes at 37°C.

Wash and dry the plate as described previously

The substrate solution is added (50-100 µl) to each well and incubated for 25 minutes at room temperature. If using HRP bound secondary antibodies, use 3,3',5,5'-Tetramethylbenzidine (TMB) substrate (Thermo Scientific 34024).

Once the incubation is complete, stop the reaction with 50-100 µl of Stop solution and read the absorbance of assay plate at 450nm on spectrophotometer.

A.3.5. ELISA Worksheets



Sample Id	1	2	3	4	5	6	7	8	9	10	11	12
a	clone id	clone id	clone id	clone id	clone id	clone id	clone id	clone id	clone id	clone id	clone id	clone id
b	clone id	clone id	clone id	clone id	clone id	clone id	clone id	clone id	clone id	clone id	clone id	clone id
c	clone id	clone id	clone id	clone id	clone id	clone id	clone id	clone id	clone id	clone id	clone id	clone id
d	clone id	clone id	clone id	clone id	clone id	clone id	clone id	clone id	clone id	clone id	clone id	clone id
e	clone id	clone id	clone id	clone id	M13+ sera po	M13+ sera poo	clone id	clone id	clone id	clone id	M13+ sera po	M13+ sera pool
f	clone id	clone id	clone id	clone id	M13+ block b	M13+ block bu	clone id	clone id	clone id	clone id	M13+ block b	M13+ block buffer
g	clone id	clone id	clone id	clone id	PBS + sera p	PBS + sera po	clone id	clone id	clone id	clone id	PBS + sera p	PBS + sera pool
h	clone id	clone id	clone id	clone id	uncoated + b	uncoated + bl	clone id	clone id	clone id	clone id	uncoated + b	uncoated + block buffer

OD	1	2	3	4	5	6	7	8	9	10	11	12
a												
b												
c												
d												
e												
f												
g												
h												

average readings						
a	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!
b	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!
c	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!
d	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!
e	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!
f	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!
g	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!
h	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!

Slow reacting TMB used. 25 minute stop

avg reading background subtracted	#DIV/0!	average of h3 and 6

average reading minus average background						
a	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!
b	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!
c	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!
d	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!
e	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!
f	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!
g	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!
h	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!

Experiment name. Values are Avg OD-cell E32			
well Id	LIV + pool	LIV - pool	diff liv+ minus liv -
Clone id	#DIV/0!	#DIV/0!	#DIV/0!
Clone id	#DIV/0!	#DIV/0!	#DIV/0!
Clone id	#DIV/0!	#DIV/0!	#DIV/0!
Clone id	#DIV/0!	#DIV/0!	#DIV/0!
Clone id	#DIV/0!	#DIV/0!	#DIV/0!
Clone id	#DIV/0!	#DIV/0!	#DIV/0!
Clone id	#DIV/0!	#DIV/0!	#DIV/0!
Clone id	#DIV/0!	#DIV/0!	#DIV/0!
Clone id	#DIV/0!	#DIV/0!	#DIV/0!
Clone id	#DIV/0!	#DIV/0!	#DIV/0!
Clone id	#DIV/0!	#DIV/0!	#DIV/0!
Clone id	#DIV/0!	#DIV/0!	#DIV/0!
Clone id	#DIV/0!	#DIV/0!	#DIV/0!
Clone id	#DIV/0!	#DIV/0!	#DIV/0!
Clone id	#DIV/0!	#DIV/0!	#DIV/0!
Clone id	#DIV/0!	#DIV/0!	#DIV/0!
Clone id	#DIV/0!	#DIV/0!	#DIV/0!
Clone id	#DIV/0!	#DIV/0!	#DIV/0!
Clone id	#DIV/0!	#DIV/0!	#DIV/0!
Clone id	#DIV/0!	#DIV/0!	#DIV/0!
M13+ sera pool	#DIV/0!	#DIV/0!	#DIV/0!
M13+ block buffer	#DIV/0!	#DIV/0!	#DIV/0!
PBS + sera pool	#DIV/0!	#DIV/0!	#DIV/0!
uncoated + block buffer	#DIV/0!	#DIV/0!	#DIV/0!

STDEV of LIV clones with neg pool	#DIV/0!	
3* STDEV	#DIV/0!	
Average OD LIV clones with neg pool	#DIV/0!	
(3*STDEV)+ (Avg OD of LIV neg)	#DIV/0!	cut off value for pos/neg
clones with higher OD> cut off		

Figure A3-1: Worksheet used for data analysis of pooled serum ELISA experiments

A.3.6. Sensitivity/specificity calculations for the phage clones tested with the independent serum panels.

Table A3-5: Sensitivity and specificity calculations for lower cut off.

PHAGE clone		
	LIV pos sera (number)	LIV neg sera (number)
Positive result (Number)	A (true pos)	B (false pos)
Negative result (number)	C (false neg)	D (true neg)
sensitivity	A/(A+C)*100	
specificity	D/(D=B)*100	

PHAGE 13	disease	non disease
positive	2	2
negative	18	18
sensitivity	10	
specificity	90	

PHAGE 16	disease	non disease
positive	17	10
negative	3	10
sensitivity	85	
specificity	50	

PHAGE 22	disease	non disease
positive	15	3
negative	5	17
sensitivity	75	
specificity	85	

PHAGE 40	disease	non disease
positive	17	7
negative	3	13
sensitivity	85	
specificity	65	

PHAGE 50	disease	non disease
positive	15	1
negative	5	19
sensitivity	75	
specificity	95	

PHAGE 55	disease	non disease
positive	11	1
negative	9	19
sensitivity	55	
specificity	95	

PHAGE 61	disease	non disease
positive	6	1
negative	14	19
sensitivity	30	
specificity	95	

Table A3-6: Sensitivity and specificity calculations for higher cut off

PHAGE clone		
	LIV pos sera (number)	LIV neg sera (number)
Positive result (Number)	A (true pos)	B (false pos)
Negative result (number)	C (false neg)	D (true neg)
sensitivity	A/(A+C)*100	
specificity	D/(D+B)*100	

PHAGE 13	disease	non disease
positive	1	2
negative	19	18
sensitivity	5	
specificity	90	

PHAGE 16	disease	non disease
positive	13	5
negative	7	15
sensitivity	65	
specificity	75	

PHAGE 22	disease	non disease
positive	14	2
negative	6	18
sensitivity	70	
specificity	90	

PHAGE 40	disease	non disease
positive	17	4
negative	3	16
sensitivity	85	
specificity	80	

PHAGE 50	disease	non disease
positive	14	1
negative	6	19
sensitivity	70	
specificity	95	

PHAGE 55	disease	non disease
positive	9	1
negative	11	19
sensitivity	45	
specificity	95	

PHAGE 61	disease	non disease
positive	4	1
negative	16	19
sensitivity	20	
specificity	95	

Appendix 4 Chapter 5 additional information

A.4.1. HI results from Moredun Research Institute

HI results for Estonian sheep sera tested at the MRI.

Moredun Research Institute
Surveillance Unit

Moredun Research Institute

Pentland Science Park, Bush Loan
Midlothian, EH26 0PZ
Tel: +44 (0)131 445 5111
Fax: +44 (0)131 445 6111

Email: mara.rocchi@moredun.ac.uk

Consultant Virologist(s)
Francesca Chianini DVM, PhD, MRCVS, Veterinary Pathologist
Mara Rocchi PhD, MRCVS, Head of Viral Surveillance Unit
Mark Dagleish BVM&S, PhD, MRCVS, FRCPath, Head of Pathology

Your Ref No:	Estonian samples	**Our Ref No:**	19/1046
Sender:	Moredun Research Institute	**Date Received:**	
Date Generated:	09-Jan-2020	**Reported By:**	Mara Rocchi PhD, MRCVS

Interim Report:

Final Report:
Antibody to Louping Ill Virus (LIV) detected by HAI test as shown.

Note: antibodies are IgG only.
Two samples (11 and 12) were mixed by accident therefore their titre is unknown.

Report: 12176 Date: 09-Jan-2020/17:19:26 From: Database: LIMS_PROD

A.4.2. Serum samples from Estonia and Norway

Table A4-1: Serum samples from Norwegian and Estonian collaborators tested for TBEV IgG and LIV IgG.

	Source serum Id	Progen ELISA results. Below 65 Vienna units is negative	LIV HAI result reported from MRI	TBEV result reported from source institute
Estonia serum				
1	71	13.42	neg	1
2	74	291.74	0.0972	1
3	150	34.49	neg	0.5
4	156	26.31	neg	1
5	232	23.05	01:10	1
6	236	66.52	01:10	1
7	239	20.79	01:40	1
8	259	383.55	01:10	1
9	240	18.37	neg	1
10	256	277.25	01:10	1
11	265	14.83	neg	1
12	270	17.47	neg	0
13	284	12.21	neg	0.5
14	291	12.45	neg	0
15	292	12.57	neg	0
16	297	56.37	neg	0.5
17	351	24.85	neg	nt
18	354	27.04	neg	nt
19	410	31.76	neg	nt
20	413	23.9	neg	nt
21	415	20.24	neg	nt
22	418	17.13	neg	nt
23	423	407.06	0.1528	nt
24	425	42.95	neg	nt
25	429	548.3	>1:640	nt
26	477	18.03	neg	nt
27	482	19.47	neg	nt
28	483	21.22	neg	nt
29	485	19.58	neg	nt
30	489	17.25	neg	nt
31	490	18.81	neg	nt
32	491	19.36	neg	nt
33	492	19.8	neg	nt
34	493	17.25	neg	nt

35	502	29.52	neg	nt
36	503	19.91	neg	nt
37	505	16.56	neg	nt
38	506	13.42	neg	nt
39	509	15.42	neg	nt
40	510	27.46	neg	nt
41	272	188.81	01:20	1
42	274	193.33	01:20	1
43	275	232.33	01:20	1
44	286	168.31	01:10	1
45	287	138.78	01:10	1
46	288	165.8	01:10	1
47	289	174.8	01:10	1
48	290	Result below Vienna unit threshold	01:10	1
49	65	76.04	01:10	0.5
50	68	68.85	01:10	0.5
51	143	24.62	01:20	0.5
52	144	18.59	01:10	0.5
53	276	151.65	01:20	0.5
54	278	76.84	01:10	0.5
55	279	72.05	01:10	0.5
56	280	72.05	01:10	0.5
57	281	75.24	01:10	0.5
58	282	69.65	01:10	0.5
59	296	62.37	01:10	0.5
60	300	115.84	01:10	0.5
61	350	54.88	01:10	nt
62	408	110.54	01:10	nt
63	412	18.59	01:10	nt
64	419	122.22	01:10	nt
65	422	240.88	01:40	nt
66	424	128.83	01:40	nt
67	428	125.97	01:10	nt
68	431	101.15	01:10	nt
69	434	109.67	01:10	nt
70	435	38.42	01:10	nt
71	481	35.32	01:10	nt
72	484	6.97	01:10	nt
73	499	20.23	01:10	nt
74	500	6.97	01:10	nt
75	873	34.24	01:10	nt
76	874	135.73	01:10	nt
Norway serum				
77	1_8 D14	181.62	nt	240
78	2_36 D14	73.65	nt	20

79	3_78 D14	601.42	nt	120
80	4_119 D14	354.11	nt	160
81	5_264 D14	141.89	nt	160
82	1_8 D31	567.73	nt	160
83	2_36 D31	72.85	nt	30
84	3_78 D31	552.64	nt	80
85	4_119 D31	too low	nt	80
86	5_264 D31	too low	nt	120
87	11_24 D10	130.77	nt	480
88	12_55 D10	56.57	nt	320
89	13_108 D10	76.84	nt	80
90	14_176 D10	123.15	nt	160
91	15_504 D10	116.74	nt	80
92	1_26 D18	601.42	nt	240
93	2_38 D18	193.33	nt	320
94	3_92 D18	354.11	nt	40
95	4_143 D18	too low	nt	240
96	5_150 D18	262.38	nt	160

A.4.3. Ion Torrent Batch 1 NGS samples

Table A4-2: Output phage sub-libraries from Chapter 5 biopanning experiments which were processed for Ion Torrent sequencing.

Biopanning sample Id	Concentration of DNA following miniprep (ng/uL)	NGS barcode	concentration of stock DNA on Qubit (µg/mL)	volume of DNA used to make 100 ng stock (µL)
LIV vs TBEV biopanning output sub-libraries				
LIV 1	162.8	Batch 1 2020 BC01	34.4	2.91
LIV 2	169.1	Batch 1 2020 BC02	39.9	2.51
LIV 3	126	Batch 1 2020 BC03	33.9	2.95
LIV 4	121.4	Batch 1 2020 BC04	33.4	2.99
LIV 5	120.4	Batch 1 2020 BC05	32.8	3.05
LIV 6	82.16	Batch 1 2020 BC06	32.4	3.09
LIV 7	138.3	Batch 1 2020 BC07	19.3	5.18
LIV 8	265.7	Batch 1 2020 BC08	32.4	3.09
LIV 9	143.5	Batch 1 2020 BC09	31.7	3.15
LIV 10	107.7	Batch 1 2020 BC10	22.3	4.48
LIV 11	184.8	Batch 1 2020 BC11	35.2	2.84
LIV 12	139.6	Batch 1 2020 BC12	34.5	2.9
LIV 13	160.5	Batch 1 2020 BC13	39.9	2.51
LIV 14	155	Batch 1 2020 BC14	34.3	2.92
LIV 15	99.34	Batch 1 2020 BC15	35.4	2.82
LIV 16	149	Batch 1 2020 BC16	35.5	2.82
LIV 17	127.8	Batch 1 2020 BC17	40.7	2.46
LIV 18	127.5	Batch 1 2020 BC18	39.4	2.54
LIV 19	188.6	Batch 1 2020 BC19	34.2	2.92
LIV 20	100.4	Batch 1 2020 BC20	30.3	3.3
LIV 21	108.8	Batch 1 2020 BC21	31.4	3.18
LIV 22	102.2	Batch 1 2020 BC22	37.7	2.65
LIV 23	61.57	Batch 1 2020 BC23	31.7	3.15
LIV 24	81.54	Batch 1 2020 BC24	33.3	3
LIV 25	74.78	Batch 1 2020 BC25	36.4	2.75
LIV 26	66.76	Batch 1 2020 BC26	35.3	2.83
LIV 27	95.26	Batch 1 2020 BC27	35.9	2.79
LIV 28	84.14	Batch 1 2020 BC28	28.4	3.52
LIV 29	132.2	Batch 1 2020 BC29	37.2	2.69
LIV 30	110.7	Batch 1 2020 BC30	38.8	2.58
TBEV 1	125.5	Batch 1 2020 BC31	39.6	2.53
TBEV 2	108.4	Batch 1 2020 BC32	40.4	2.48
TBEV 3	101.8	Batch 1 2020 BC33	34.6	2.89

TBEV 4	129.5	Batch 1 2020 BC34	33.1	3.02
TBEV 5	108.6	Batch 1 2020 BC35	33.6	2.98
TBEV 6	105.4	Batch 1 2020 BC36	44.4	2.25
TBEV 7	119.3	Batch 1 2020 BC37	36.7	2.72
TBEV 8	107	Batch 1 2020 BC38	38.8	2.58
TBEV 9	118.5	Batch 1 2020 BC39	38.1	2.62
TBEV 10	96.96	Batch 1 2020 BC40	37.9	2.64
TBEV 11	134.6	Batch 1 2020 BC41	33.5	2.99
TBEV 12	136.2	Batch 1 2020 BC43	39.9	2.51
TBEV 13	130.7	Batch 1 2020 BC44	34.8	2.87
TBEV 14	133.7	Batch 1 2020 BC45	36.2	2.76
TBEV 15	153.6	Batch 1 2020 BC46	33.8	2.96
TBEV 16	114.9	Batch 1 2020 BC47	22.1	4.52
TBEV 17	138.2	Batch 1 2020 BC48	29.1	3.44
TBEV 18	86.34	Batch 1 2020 BC49	29.4	3.4
TBEV 19	129.7	Batch 1 2020 BC50	40.1	2.49
TBEV 20	120	Batch 1 2020 BC51	31.2	3.21
TBEV vs LIV biopanning output sub-libraries				
TBEV 1	421.1	Batch 1 2020 BC52	34.7	2.88
TBEV 2	339.9	Batch 1 2020 BC53	36.5	2.74
TBEV 3	201.8	Batch 1 2020 BC54	34.8	2.87
TBEV 4	482	Batch 1 2020 BC55	31.6	3.16
TBEV 5	276.6	Batch 1 2020 BC56	36.1	2.77
TBEV 6	429.6	Batch 1 2020 BC57	37.2	2.69
TBEV 7	324.4	Batch 1 2020 BC58	40.2	2.49
TBEV 8	324.4	Batch 1 2020 BC59	37.2	2.69
TBEV 9	330.3	Batch 1 2020 BC60	42.5	2.35
TBEV 10	294.3	Batch 1 2020 BC61	39.4	2.54
TBEV 11	335	Batch 1 2020 BC62	38.3	2.61
TBEV 12	276.7	Batch 1 2020 BC63	23.2	4.31
TBEV 13	251.2	Batch 1 2020 BC64	35.9	2.79
TBEV 14	284.8	Batch 1 2020 BC65	37.9	2.64
TBEV 15	255.2	Batch 1 2020 BC66	38.1	2.62
TBEV 16	266.9	Batch 1 2020 BC67	41.9	2.39
TBEV 17	90.98	Batch 1 2020 BC68	40	2.5
TBEV 18	422.9	Batch 1 2020 BC69	38.7	2.58
TBEV 19	231	Batch 1 2020 BC70	21.9	4.57
TBEV 20	262.8	Batch 1 2020 BC71	35.3	2.83
LIV 1	409.9	Batch 1 2020 BC72	33.8	2.96
LIV 2	363.6	Batch 1 2020 BC73	32.1	3.12
LIV 3	474.7	Batch 1 2020 BC74	35.9	2.79
LIV 4	162	Batch 1 2020 BC75	30.3	3.3

LIV 5	367.9	Batch 1 2020 BC76	31	3.23
LIV 6	412	Batch 1 2020 BC77	29.9	3.34
LIV 7	391.4	Batch 1 2020 BC78	32.2	3.11
LIV 8	367.6	Batch 1 2020 BC79	36	2.78
LIV 9	389.6	Batch 1 2020 BC80	35.5	2.82
LIV 10	490.9	Batch 1 2020 BC81	37.4	2.67
LIV 11	505.8	Batch 1 2020 BC82	36.5	2.74
LIV 12	361.4	Batch 1 2020 BC83	33.1	3.02
LIV 13	441.1	Batch 1 2020 BC84	34.1	2.93
LIV 14	448.1	Batch 1 2020 BC85	31.6	3.16
LIV 15	388.9	Batch 1 2020 BC86	37.7	2.65
LIV 16	379.9	Batch 1 2020 BC87	32.7	3.06
LIV 17	413.4	Batch 1 2020 BC88	38.5	2.6
LIV 18	339.1	Batch 1 2020 BC89	34.7	2.88
LIV 19	327.8	Batch 1 2020 BC90	32.2	3.11
LIV 20	409.5	Batch 1 2020 BC91	39	2.56
LIV 21	431.9	Batch 1 2020 BC92	38.7	2.58
LIV 22	241.6	Batch 1 2020 BC93	38.4	2.6
LIV 23	195.3	Batch 1 2020 BC94	35.8	2.79
LIV 24	240.5	Batch 1 2020 BC95	40.2	2.49
LIV 25	468.8	Batch 2 2020 BC01	42.9	2.33
LIV 26	375.4	Batch 2 2020 BC02	38.2	2.62
LIV 27	154.5	Batch 2 2020 BC03	36	2.78
LIV 28	317	Batch 2 2020 BC04	40	2.5
LIV 29	405.9	Batch 2 2020 BC05	32.2	3.11
LIV 30	206.8	Batch 2 2020 BC06	36.2	2.76

A.4.4. Ion Torrent Batch 2 NGS samples

Table A4-3: Output phage sub-libraries from all three experiments were batched together for sequencing submitted in 2020.

Biopanning sample Id	Concentration of DNA following miniprep (ng/μL)	NGS barcode	concentration of stock DNA on Qubit (μg/mL)	volume of DNA used to make 100 ng stock (μL)
TBEV vs P8				
LIV 25	468.8	Batch 2 2020 BC01	42.9	2.3
LIV 26	375.4	Batch 2 2020 BC02	38.2	2.6
LIV 27	154.5	Batch 2 2020 BC03	36	2.8
LIV 28	317	Batch 2 2020 BC04	40	2.5
LIV 29	405.9	Batch 2 2020 BC05	32.2	3.1
LIV 30	206.8	Batch 2 2020 BC06	36.2	2.8
ZIKV/DENV repeat panning and additional sera panning				
DENV r11	147.4	Batch 2 2020 BC07	41.9	2.4
DENV r12	87.19	Batch 2 2020 BC08	23.3	4.3
DENV r13	169.5	Batch 2 2020 BC09	24.7	4
DENV r14	177.7	Batch 2 2020 BC10	35.1	2.8
DENV r15	96.06	Batch 2 2020 BC11	36.1	2.8
DENV r16	90.58	Batch 2 2020 BC12	30.9	3.2
DENV r17	165.7	Batch 2 2020 BC13	41	2.4
DENV r18	147.9	Batch 2 2020 BC14	34.5	2.9
DENV r19	144.4	Batch 2 2020 BC15	33.4	3
DENV 21	87.88	Batch 2 2020 BC16	37.1	2.7
DENV 22	102	Batch 2 2020 BC17	40.5	2.5
DENV 23	159.9	Batch 2 2020 BC18	42.9	2.3
DENV 24	101.2	Batch 2 2020 BC19	37.1	2.7
DENV 25	103.8	Batch 2 2020 BC20	34.1	2.9
DENV 26	97.45	Batch 2 2020 BC21	34.4	2.9
DENV 27	141.2	Batch 2 2020 BC22	39	2.6
DENV 28	132	Batch 2 2020 BC23	42.6	2.3
DENV 29	127.6	Batch 2 2020 BC24	36.4	2.7
DENV 30	106.7	Batch 2 2020 BC25	46	2.2
DENv r21	140.7	Batch 2 2020 BC26	40.7	2.5
DENv r22	160.4	Batch 2 2020 BC27	34.6	2.9
DENv r23	102.8	Batch 2 2020 BC28	42.8	2.3
DENv r24	89.3	Batch 2 2020 BC29	45.8	2.2
DENv r25	164.6	Batch 2 2020 BC30	40.3	2.5
DENv r26	141.4	Batch 2 2020 BC31	42.2	2.4
DENv r27	128.8	Batch 2 2020 BC32	44.9	2.2
DENV r28	127.6	Batch 2 2020 BC33	37.8	2.6

DENv r29	154.8	Batch 2 2020 BC34	43.2	2.3
DENv r30	132.3	Batch 2 2020 BC35	36	2.8
ZIKV r11	172.8	Batch 2 2020 BC36	38.5	2.6
ZIKV r12	166.3	Batch 2 2020 BC37	31.5	3.2
ZIKV r13	171.7	Batch 2 2020 BC38	37.1	2.7
ZIKV r14	176.5	Batch 2 2020 BC39	44.4	2.3
ZIKV r15	126.1	Batch 2 2020 BC40	34	2.9
ZIKV r16	145.4	Batch 2 2020 BC41	29.6	3.4
ZIKV r17	185.9	Batch 2 2020 BC43	33.8	3
ZIKV r18	175.9	Batch 2 2020 BC44	38	2.6
ZIKV r19	83.85	Batch 2 2020 BC45	38.8	2.6
ZIKV r20	150.4	Batch 2 2020 BC46	35.4	2.8
ZIKV 21	127	Batch 2 2020 BC47	37	2.7
ZIKV 22	96.25	Batch 2 2020 BC48	39.5	2.5
ZIKV 23	163.8	Batch 2 2020 BC49	31.3	3.2
ZIKV 24	37.3	Batch 2 2020 BC50	39.1	2.6
ZIKV 25	175.1	Batch 2 2020 BC51	26.1	3.8
ZIKV 26	113.8	Batch 2 2020 BC52	34	2.9
ZIKV 27	102.7	Batch 2 2020 BC53	38.8	2.6
ZIKV 28	120.1	Batch 2 2020 BC54	28.9	3.5
ZIKV 29	81.3	Batch 2 2020 BC55	24.7	4
ZIKV r21	120.5	Batch 2 2020 BC56	35.8	2.8
ZIKV r22	108.2	Batch 2 2020 BC57	37	2.7
ZIKV r23	157.4	Batch 2 2020 BC58	38	2.6
ZIKV r24	124.5	Batch 2 2020 BC59	37	2.7
ZIKV r25	113.6	Batch 2 2020 BC60	49.4	2
ZIKV r26	119.4	Batch 2 2020 BC61	44.7	2.2
ZIKV r27	124.8	Batch 2 2020 BC62	24.3	4.1
ZIKV r28	125.2	Batch 2 2020 BC63	15.6	6.4
ZIKV r29	190.4	Batch 2 2020 BC64	37.2	2.7
LIV vs Healthy				
LIV 1	141.6	Batch 2 2020 BC65	39.9	2.5
LIV 2	182.3	Batch 2 2020 BC66	33.8	3
LIV 3	195.3	Batch 2 2020 BC67	37.5	2.7
LIV 4	159.7	Batch 2 2020 BC68	49	2
LIV 5	131.1	Batch 2 2020 BC69	34.3	2.9
LIV 6	176.1	Batch 2 2020 BC70	14.2	7
LIV 7	200.2	Batch 2 2020 BC71	38.8	2.6
LIV 8	169.4	Batch 2 2020 BC72	33.6	3
LIV 9	195	Batch 2 2020 BC73	33	3
LIV 10	182.5	Batch 2 2020 BC74	30.5	3.3
LIV uninfected 1	202.7	Batch 2 2020 BC75	32.6	3.1
LIV uninfected 2	193.3	Batch 2 2020 BC76	31.9	3.1

LIV uninfected 3	141.3	Batch 2 2020 BC77	40.3	2.5
LIV uninfected 4	195.3	Batch 2 2020 BC78	32.4	3.1
LIV uninfected 5	146.6	Batch 2 2020 BC79	38.6	2.6
LIV uninfected 6	117.9	Batch 2 2020 BC80	46.6	2.1
LIV uninfected 7	172.8	Batch 2 2020 BC81	34.3	2.9
LIV uninfected 8	169.9	Batch 2 2020 BC82	36.7	2.7
LIV uninfected 9	203.6	Batch 2 2020 BC83	36.4	2.7
LIV uninfected 10	131	Batch 2 2020 BC84	38.2	2.6
Cross-reactive library (TBEV vs LIV R2 library)				
TBEV 1	188.5	Batch 2 2020 BC85	36.8	2.7
TBEV 2	168.4	Batch 2 2020 BC86	40.4	2.5
TBEV 3	265.2	Batch 2 2020 BC87	32.4	3.1
TBEV 4	222.3	Batch 2 2020 BC88	36.9	2.7
TBEV 5	298.7	Batch 2 2020 BC89	39.9	2.5
TBEV 6	279.3	Batch 2 2020 BC90	32.7	3.1
TBEV 7	157.2	Batch 2 2020 BC91	43.7	2.3
TBEV 8	161	Batch 2 2020 BC92	41.4	2.4
TBEV 9	170	Batch 2 2020 BC93	39.8	2.5
TBEV 10	283.4	Batch 2 2020 BC94	36.8	2.7

Chapter 7 **References**

Abbott, W.M., Damschroder, M.M. and Lowe, D.C. (2014) 'Current approaches to fine mapping of antigen–antibody interactions', *Immunology*, 142(4), p. 526. Available at: https://doi.org/10.1111/IMM.12284.

Agrawal, S. *et al.* (2016) 'A high affinity phage-displayed peptide as a recognition probe for the detection of Salmonella Typhimurium', *Journal of biotechnology*, 231, pp. 40–45. Available at: https://doi.org/10.1016/J.JBIOTEC.2016.05.027.

Amicizia, D. *et al.* (2013) 'Epidemiology of tick-borne encephalitis (TBE) in Europe and its prevention by available vaccines', *Human Vaccines & Immunotherapeutics*, 9(5), p. 1163. Available at: https://doi.org/10.4161/HV.23802.

Antoniou, E. *et al.* (2020) 'Zika Virus and the Risk of Developing Microcephaly in Infants: A Systematic Review', *International Journal of Environmental Research and Public Health*, 17(11). Available at: https://doi.org/10.3390/IJERPH17113806.

APHA and SRUC (2021) *Sheep Dashboard | Tableau Public*. Available at: https://public.tableau.com/app/profile/siu.apha/viz/SheepDashboard_/Overview (Accessed: 13 September 2021).

de Araújo, T.V.B. *et al.* (2018) 'Association between microcephaly, Zika virus infection, and other risk factors in Brazil: final report of a case-control study.', *The Lancet. Infectious diseases*, 18(3), pp. 328–336. Available at: https://doi.org/10.1016/S1473-3099(17)30727-2.

Ashburn, P.M. and Craig, C.F. (1907) 'Experimental Inverstigations Regarding the Etiology of Dengue Fever', *The Journal of Infectious*

Diseases, 4(3), pp. 440–475. Available at: https://doi.org/10.1093/INFDIS/4.3.440.

Atkinson, B. *et al.* (2016) 'Detection of Zika Virus in Semen', *Emerging Infectious Diseases*, 22(5), pp. 940–940. Available at: https://doi.org/10.3201/eid2205.160107.

Bazan, J., Całkosiński, I. and Gamian, A. (2012) 'Phage display—A powerful technique for immunotherapy', *Human Vaccines & Immunotherapeutics*, 8(12), pp. 1817–1828. Available at: https://doi.org/10.4161/hv.21703.

Bergamaschi, G. *et al.* (2019) 'Computational Analysis of Dengue Virus Envelope Protein (E) Reveals an Epitope with Flavivirus Immunodiagnostic Potential in Peptide Microarrays', *International Journal of Molecular Sciences 2019, Vol. 20, Page 1921*, 20(8), p. 1921. Available at: https://doi.org/10.3390/IJMS20081921.

Bhatt, P. *et al.* (2021) 'Current Understanding of the Pathogenesis of Dengue Virus Infection', *Current Microbiology*, 78(1), p. 17. Available at: https://doi.org/10.1007/S00284-020-02284-W.

Bhatt, S. *et al.* (2013) 'The global distribution and burden of dengue.', *Nature*, 496(7446), pp. 504–7. Available at: https://doi.org/10.1038/nature12060.

Böhm, B. *et al.* (2017) 'Tick-borne encephalitis in a naturally infected sheep.', *BMC veterinary research*, 13(1), p. 267. Available at: https://doi.org/10.1186/s12917-017-1192-3.

Bonheur, A.N. *et al.* (2021) 'A fatal case report of antibody-dependent enhancement of dengue virus type 1 following remote Zika virus infection', *BMC Infectious Diseases 2021 21:1*, 21(1), pp. 1–8. Available at: https://doi.org/10.1186/S12879-021-06482-0.

Breitling, F. *et al.* (1991) 'A surface expression vector for antibody screening', *Gene*, 104(2), pp. 147–153. Available at: https://doi.org/10.1016/0378-1119(91)90244-6.

Bressanelli, S. *et al.* (2004) 'Structure of a flavivirus envelope glycoprotein in its low-pH-induced membrane fusion conformation', *The EMBO journal*, 23(4), pp. 728–738. Available at: https://doi.org/10.1038/SJ.EMBOJ.7600064.

Brockmann, S. *et al.* (2018) 'A cluster of two human cases of tick-borne encephalitis (TBE) transmitted by unpasteurised goat milk and cheese in Germany, May 2016', *Eurosurveillance*, 23(15), pp. 17–00336. Available at: https://doi.org/10.2807/1560-7917.ES.2018.23.15.17-00336.

Brotherston, J.G. and Boyce, J.B. (1970) 'Development of a non-infective protective antigen against louping-ill (arborirus group B): Laboratory experiments', *Journal of Comparative Pathology*, 80(3), pp. 377–388. Available at: https://doi.org/10.1016/0021-9975(70)90068-X.

Brown, J.E. *et al.* (2014) 'HUMAN IMPACTS HAVE SHAPED HISTORICAL AND RECENT EVOLUTION IN AEDES AEGYPTI, THE DENGUE AND YELLOW FEVER MOSQUITO', *Evolution*, 68(2), pp. 514–525. Available at: https://doi.org/10.1111/EVO.12281.

Brown, J.R. *et al.* (2011) 'Statistical methods for analyzing immunosignatures', *BMC Bioinformatics*, 12(1), pp. 1–15. Available at: https://doi.org/10.1186/1471-2105-12-349/FIGURES/3.

Brownlee, A. and Wilson, R. (1932) 'Studies in the Histopathology of Louping-ILL', *Journal of Comparative Pathology and Therapeutics*, 45, pp. 67–92. Available at: https://doi.org/10.1016/S0368-1742(32)80006-8.

Bryan, C.S. *et al.* (1997) 'Featured Report on a Landmark Study Discovery of the Yellow Fever Virus', *International Journal of Infectious Diseases* , 2(1), pp. 52–54.

Buxton, D. and Reid, H.W. (2017) '120 years of Louping-ill Research: an Historical Perspective from the Archive of the Journal of Comparative Pathology', *Journal of Comparative Pathology*, 157(4), pp. 270–275. Available at: https://doi.org/10.1016/j.jcpa.2017.09.002.

Caini, S. *et al.* (2012) 'Tick-borne encephalitis transmitted by unpasteurised cow milk in western Hungary, September to October 2011', *Eurosurveillance*, 17(12), p. 20128. Available at: https://doi.org/10.2807/ESE.17.12.20128-EN.

Calisher, C. *et al.* (1989) 'Antigenic relationships between flaviviruses as determined by cross-neutralization tests with polyclonal antisera', *The Journal of general virology*, 70 (Pt 1)(1), pp. 37–43. Available at: https://doi.org/10.1099/0022-1317-70-1-37.

Calvet, G. *et al.* (2016) 'Detection and sequencing of Zika virus from amniotic fluid of fetuses with microcephaly in Brazil: a case study', *The Lancet Infectious Diseases*, 16(6), pp. 653–660. Available at: https://doi.org/10.1016/S1473-3099(16)00095-5.

Casals, J. (1944) 'Immunological relationships among Central Nervous System Viruses', *Journal of Experimental Medicine*, 79(4), pp. 341–359. Available at: https://doi.org/10.1084/JEM.79.4.341.

Casals, J. and Brown, L. (1954) 'Hemagglutination with arthropod-borne viruses', *The Journal of experimental medicine*, 99(5), pp. 429–449. Available at: https://doi.org/10.1084/JEM.99.5.429.

Casals, J. and Brown, L. v. (1953) 'Hemagglutination with certain arthropod-borne viruses', *Proceedings of the Society for Experimental Biology and Medicine. Society for Experimental Biology and Medicine (New York, N.Y.)*, 83(1), pp. 170–173. Available at: https://doi.org/10.3181/00379727-83-20299.

Chen, L.H. and Wilson, M.E. (2016) 'Update on non-vector transmission of dengue: relevant studies with Zika and other flaviviruses', *Tropical Diseases, Travel Medicine and Vaccines*, 2(1), p. 15. Available at: https://doi.org/10.1186/s40794-016-0032-y.

Chong, M.K. *et al.* (2014) 'Microscopy techniques in flavivirus research', *Micron*, 59, pp. 33–43. Available at: https://doi.org/10.1016/J.MICRON.2013.12.006.

Clapham, H E *et al.* (2021) 'The Journal of Infectious Diseases Using NS1 Flavivirus Protein Microarray to Infer Past Infecting Dengue Virus Serotype and Number of Past Dengue Virus Infections in Vietnamese Individuals', *The Journal of Infectious Diseases* ®, 223, pp. 2053–61. Available at: https://doi.org/10.1093/infdis/jiaa018.

Clements, A.N. and Harbach, R.E. (2017) *History of the discovery of the mode of transmission of yellow fever virus*, *Journal of Vector Ecology*.

Collins, M.H. *et al.* (2017) 'Lack of Durable Cross-Neutralizing Antibodies Against Zika Virus from Dengue Virus Infection', *Emerging infectious diseases*, 23(5), pp. 773–781. Available at: https://doi.org/10.3201/EID2305.161630.

Cooper, P.D. (1962) 'The Plaque Assay of Animal Viruses', *Advances in Virus Research*, 8(C), pp. 319–378. Available at: https://doi.org/10.1016/S0065-3527(08)60689-2.

Corrin, T. *et al.* (2018) 'Powassan virus, a scoping review of the global evidence', *Zoonoses and public health*, 65(6), pp. 595–624. Available at: https://doi.org/10.1111/ZPH.12485.

Cumbie, A.N., Whitlow, A.M. and Eastwood, G. (2022) 'First Evidence of Powassan Virus (Flaviviridae) in Ixodes scapularis in Appalachian Virginia, USA', *The American Journal of Tropical Medicine and Hygiene*, 106(3), pp. 905–908. Available at: https://doi.org/10.4269/AJTMH.21-0825.

Datta, S. and Wattal, C. (2010) 'Dengue NS1 antigen detection: a useful tool in early diagnosis of dengue virus infection.', *Indian journal of medical microbiology*, 28(2), pp. 107–10. Available at: https://doi.org/10.4103/0255-0857.62484.

Davidson, M.M., Williams, H. and Macleod, J.A. (1991) 'Louping ill in man: a forgotten disease.', *The Journal of infection*, 23(3), pp. 241–9. Available at: https://doi.org/10.1016/0163-4453(91)92756-U.

Devadiga, S. *et al.* (2020) 'Dynamics of human B and T cell adaptive immune responses to Kyasanur Forest disease virus infection', *Scientific reports*, 10(1). Available at: https://doi.org/10.1038/S41598-020-72205-1.

Deviatkin, A. *et al.* (2020) 'TBEV Subtyping in Terms of Genetic Distance', *Viruses*, 12(11). Available at: https://doi.org/10.3390/V12111240.

D'Hérelle, F. (1917) 'Sur un microbe invisible antagoniste des bacilles dysentériques', *Comptes Rendus de l'Académie des Sciences—Series D*, 165, pp. 373–375.

Dias-Neto, E. *et al.* (2009) 'Next-Generation Phage Display: Integrating and Comparing Available Molecular Tools to Enable Cost-Effective High-Throughput Analysis', *PLOS ONE*, 4(12), p. e8338. Available at: https://doi.org/10.1371/JOURNAL.PONE.0008338.

Dick, G.W.A., Kitchen, S.F. and Haddow, A.J. (1952) 'Zika Virus (I). Isolations and serological specificity', *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 46(5), pp. 509–520. Available at: https://doi.org/10.1016/0035-9203(52)90042-4.

Doherty, P.C. and Reid, H.W. (1971) 'Experimental louping-ill in sheep and lambs: II. Neuropathology', *Journal of Comparative Pathology*, 81(3), pp. 331–337. Available at: https://doi.org/10.1016/0021-9975(71)90020-X.

D'Ortenzio, E. *et al.* (2016) 'Evidence of Sexual Transmission of Zika Virus', *New England Journal of Medicine*, 374(22), pp. 2195–2198. Available at: https://doi.org/10.1056/NEJMc1604449.

Elshuber, S. *et al.* (2003) 'Cleavage of protein prM is necessary for infection of BHK-21 cells by tick-borne encephalitis virus', *The Journal of general virology*, 84(Pt 1), pp. 183–191. Available at: https://doi.org/10.1099/VIR.0.18723-0.

European Centre for Disease Prevention and Control (ECDC) (2016) *Aedes albopictus - Factsheet for experts*.

Faye, Oumar *et al.* (2014) 'Molecular Evolution of Zika Virus during Its Emergence in the 20th Century', *PLoS Neglected Tropical Diseases*. Edited by B. Bird, 8(1), p. e2636. Available at: https://doi.org/10.1371/journal.pntd.0002636.

Felici, F. *et al.* (1991) 'Selection of antibody ligands from a large library of oligopeptides expressed on a multivalent exposition vector', *Journal of Molecular Biology*, 222(2), pp. 301–310. Available at: https://doi.org/10.1016/0022-2836(91)90213-P.

Fernandez-Garcia, M.-D. *et al.* (2009) 'Pathogenesis of Flavivirus Infections: Using and Abusing the Host Cell', *Cell Host & Microbe*, 5(4), pp. 318–328. Available at: https://doi.org/10.1016/J.CHOM.2009.04.001.

Fleeton, M.N. *et al.* (1999) 'Recombinant Semliki Forest virus particles encoding the prME or NS1 proteins of louping ill virus protect mice from lethal challenge.', *Journal of General Virology*, 80(5), pp. 1189–1198. Available at: https://doi.org/10.1099/0022-1317-80-5-1189.

Földvári, G. (2016) 'Life cycle and ecology of Ixodes ricinus: the roots of public health importance', in M.A.H. Braks et al. (eds) *Ecology and prevention of Lyme borreliosis*. The Netherlands: Wageningen Academic

Publishers (Ecology and Control of Vector-borne Diseases). Available at: https://doi.org/10.3920/978-90-8686-838-4.

Forshey, B.M., Stoddard, S.T. and Morrison, A.C. (2016) 'Dengue Viruses and Lifelong Immunity: Reevaluating the Conventional Wisdom', *The Journal of Infectious Diseases*, 214(7), pp. 979–981. Available at: https://doi.org/10.1093/INFDIS/JIW102.

Foy, B.D. *et al.* (2011) 'Probable Non–Vector-borne Transmission of Zika Virus, Colorado, USA', *Emerging Infectious Diseases*, 17(5), pp. 880–882. Available at: https://doi.org/10.3201/eid1705.101939.

Fuentes, S. *et al.* (2016) 'Antigenic Fingerprinting following Primary RSV Infection in Young Children Identifies Novel Antigenic Sites and Reveals Unlinked Evolution of Human Antibody Repertoires to Fusion and Attachment Glycoproteins', *PLOS Pathogens*, 12(4), p. e1005554. Available at: https://doi.org/10.1371/JOURNAL.PPAT.1005554.

Fumagalli, M.J., Figueiredo, L.T.M. and Aquino, V.H. (2021) 'Linear and Continuous Flavivirus Epitopes From Naturally Infected Humans', *Frontiers in Cellular and Infection Microbiology*, 0, p. 705. Available at: https://doi.org/10.3389/FCIMB.2021.710551.

Game and Wildlife Conservation Trust (2021) *Update on louping ill virus vaccine project - Game and Wildlife Conservation Trust*. Available at: https://www.gwct.org.uk/scotland/research/update-on-louping-ill-virus-vaccine-project/ (Accessed: 13 September 2021).

Gao, G.F. *et al.* (1997) *Molecular variation, evolution and geographical distribution of louping ill virus*, *Acta Virologica*. Available at: http://www.aepress.sk/acta/full/av597c.pdf (Accessed: 27 September 2018).

Gilbert, L. (2016) 'Louping ill virus in the UK: a review of the hosts, transmission and ecological consequences of control', *Experimental and*

Applied Acarology, 68(3), pp. 363–374. Available at: https://doi.org/10.1007/s10493-015-9952-x.

Global Invasive Species Database (2009) *FULL ACCOUNT FOR: Aedes albopictus*. Available at: http://www.iucngisd.org/gisd/species.php?sc=109.

Gordon, W.S. *et al.* (1932) 'Studies In Louping-Ill. (An Encephalomyelitis of Sheep.)', *Journal of Comparative Pathology and Therapeutics*, 45, pp. 106–140. Available at: https://doi.org/10.1016/S0368-1742(32)80008-1.

Gould, E.A. *et al.* (1990) 'Antigenicity of flaviviruses', *Arch Virol*, pp. 137–152.

Gratz, N.G. (2004) 'Critical review of the vector status of Aedes albopictus', *Medical and veterinary entomology*, 18(3), pp. 215–227. Available at: https://doi.org/10.1111/J.0269-283X.2004.00513.X.

Grifoni, A. *et al.* (2017) 'Global Assessment of Dengue Virus-Specific CD4+ T Cell Responses in Dengue-Endemic Areas', *Frontiers in Immunology*, 0(OCT), p. 1309. Available at: https://doi.org/10.3389/FIMMU.2017.01309.

Gubler, D.J. (1998) 'Dengue and dengue hemorrhagic fever.', *Clinical microbiology reviews*, 11(3), pp. 480–96. Available at: http://www.ncbi.nlm.nih.gov/pubmed/9665979 (Accessed: 28 September 2018).

Haby, M.M. *et al.* (2018) 'Prevalence of asymptomatic Zika virus infection: a systematic review', *Bulletin of the World Health Organization*, 96(6), pp. 402-413D. Available at: https://doi.org/10.2471/BLT.17.201541.

Haddow, A.D. *et al.* (2016) 'Genetic Characterization of Spondweni and Zika Viruses and Susceptibility of Geographically Distinct Strains of Aedes

aegypti, Aedes albopictus and Culex quinquefasciatus (Diptera: Culicidae) to Spondweni Virus', *PLoS neglected tropical diseases*, 10(10). Available at: https://doi.org/10.1371/JOURNAL.PNTD.0005083.

Hallauer, C. (1946) '[About virus detection using the Hirst test]', *Schweizerische Zeitschrift fur Pathologie und Bakteriologie. Revue suisse de pathologie et de bacteriologie*, 9(5), p. 553. Available at: https://doi.org/10.1159/000159565.

Henchal, E.A. and Putnak, J.R. (1990) 'The dengue viruses.', *Clinical Microbiology Reviews*, 3(4), p. 376. Available at: https://doi.org/10.1128/CMR.3.4.376.

Henry, T.J. and Pratt, D. (1969) 'The proteins of bacteriophage M13.', *Proceedings of the National Academy of Sciences of the United States of America*, 62(3), pp. 800–7. Available at: http://www.ncbi.nlm.nih.gov/pubmed/5257006 (Accessed: 2 August 2018).

Holbrook, M.R. (2017) 'Historical Perspectives on Flavivirus Research', *Viruses*, 9(5). Available at: https://doi.org/10.3390/V9050097.

Holding, M. *et al.* (2019) 'Detection of new endemic focus of tick-borne encephalitis virus (TBEV), Hampshire/Dorset border, England, September 2019', *Eurosurveillance*, 24(47). Available at: https://doi.org/10.2807/1560-7917.ES.2019.24.47.1900658.

Holding, M. *et al.* (2020) 'Tick-borne encephalitis virus, United Kingdom', *Emerging Infectious Diseases*, 26(1), pp. 90–96. Available at: https://doi.org/10.3201/eid2601.191085.

Hudson, P.J. *et al.* (1995) 'Persistence and transmission of tick-borne viruses: Ixodes ricinus and louping-ill virus in red grouse populations.', *Parasitology*, 111 Suppl, pp. S49-58. Available at:

http://www.ncbi.nlm.nih.gov/pubmed/8632924 (Accessed: 27 September 2018).

Hurwitz, A.M. *et al.* (2017) 'Deep sequencing of phage-displayed peptide libraries reveals sequence motif that detects norovirus', *Protein Engineering Design and Selection*, 30(2), pp. 129–139. Available at: https://doi.org/10.1093/protein/gzw074.

Iannolo, G. *et al.* (1995) 'Modifying Filamentous Phage Capsid: Limits in the Size of the Major Capsid Protein', *Journal of Molecular Biology*, 248(4), pp. 835–844. Available at: https://doi.org/10.1006/jmbi.1995.0264.

Jeffries, C.L. *et al.* (2014) 'Louping ill virus: An endemic tick-borne disease of Great Britain', *Journal of General Virology* [Preprint]. Available at: https://doi.org/10.1099/vir.0.062356-0.

Kam, Y.-W. *et al.* (2019) 'ZIKV-specific NS1 epitopes as serological markers of acute Zika virus infection', *The Journal of Infectious Diseases*, Volume 220(2), pp. 203–212. Available at: https://doi.org/10.1093/infdis/jiz092/5371191.

Karimi, M. *et al.* (2016) 'Bacteriophages and Phage-Inspired Nanocarriers for Targeted Delivery of Therapeutic Cargos', *Advanced drug delivery reviews*, 106(Pt A), p. 45. Available at: https://doi.org/10.1016/J.ADDR.2016.03.003.

Kaufmann, B. and Rossmann, M.G. (2011) 'Molecular mechanisms involved in the early steps of flavivirus cell entry', *Microbes and infection / Institut Pasteur*, 13(1), p. 1. Available at: https://doi.org/10.1016/J.MICINF.2010.09.005.

Kim, Y.H. *et al.* (2018) 'Development of a rapid diagnostic test kit to detect IgG/IgM antibody against zika virus using monoclonal antibodies to the envelope and non-structural protein 1 of the virus', *Korean Journal of*

Parasitology, 56(1), pp. 61–70. Available at: https://doi.org/10.3347/KJP.2018.56.1.61.

Klaus, C. *et al.* (2014) 'Tick-borne encephalitis virus (TBEV) – findings on cross reactivity and longevity of TBEV antibodies in animal sera', *BMC Veterinary Research 2014 10:1*, 10(1), pp. 1–9. Available at: https://doi.org/10.1186/1746-6148-10-78.

Konishi, E. and Kuno, G. (2013) 'In Memoriam: Susumu Hotta (1918–2011) - Volume 19, Number 5—May 2013 - Emerging Infectious Diseases journal - CDC', *Emerging Infectious Diseases*, 19(5). Available at: https://doi.org/10.3201/EID1905.IM0986.

Kraemer, M.U. *et al.* (2015) 'The global distribution of the arbovirus vectors Aedes aegypti and Ae. albopictus', *eLife*, 4. Available at: https://doi.org/10.7554/eLife.08347.001.

Kreusch, T.M. *et al.* (2019) 'A probable case of tick-borne encephalitis (TBE) acquired in England, July 2019', *Eurosurveillance*, 24(47). Available at: https://doi.org/10.2807/1560-7917.ES.2019.24.47.1900679.

Kubinski, M. *et al.* (2020) 'Tick-Borne Encephalitis Virus: A Quest for Better Vaccines against a Virus on the Rise', *Vaccines 2020, Vol. 8, Page 451*, 8(3), p. 451. Available at: https://doi.org/10.3390/VACCINES8030451.

Kucharski, A. and Riley, S. (2016) 'First published report of Zika virus infection in people: Simpson, not MacNamara', *The Lancet Infectious Diseases*, 17, pp. 15–17. Available at: https://doi.org/10.1126/science.

Kuzmicheva, G.A. *et al.* (2009) 'Diversity and censoring of landscape phage libraries', *Protein engineering, design & selection : PEDS*, 22(1), pp. 9–18. Available at: https://doi.org/10.1093/PROTEIN/GZN060.

Labuda, M. *et al.* (1993) 'Efficient transmission of tick-borne encephalitis virus between cofeeding ticks', *Journal of medical entomology*, 30(1), pp. 295–299. Available at: https://doi.org/10.1093/JMEDENT/30.1.295.

Lai, Z. *et al.* (2020) 'Vertical transmission of zika virus in Aedes albopictus', *PLOS Neglected Tropical Diseases*, 14(10), p. e0008776. Available at: https://doi.org/10.1371/JOURNAL.PNTD.0008776.

Lanciotti, R.S. *et al.* (2008) 'Genetic and Serologic Properties of Zika Virus Associated with an Epidemic, Yap State, Micronesia, 2007', *Emerging Infectious Diseases*, 14(8), p. 1232. Available at: https://doi.org/10.3201/EID1408.080287.

Lazear, H.M. and Diamond, M.S. (2016) 'Zika Virus: New Clinical Syndromes and Its Emergence in the Western Hemisphere', *Journal of Virology*, 90(10), pp. 4864–4875. Available at: https://doi.org/10.1128/JVI.00252-16/ASSET/7ADA8AA9-2C82-45D5-9198-1C033AB3777C/ASSETS/GRAPHIC/ZJV9990916350002.JPEG.

Lebani, K. *et al.* (2017) 'Isolation of serotype-specific antibodies against dengue virus non-structural protein 1 using phage display and application in a multiplexed serotyping assay', *PLOS ONE*, 12(7), p. e0180669. Available at: https://doi.org/10.1371/JOURNAL.PONE.0180669.

Legutki, J.B. *et al.* (2010) 'A general method for characterization of humoral immunity induced by a vaccine or infection', *Vaccine*, 28(28), pp. 4529–4537. Available at: https://doi.org/10.1016/J.VACCINE.2010.04.061.

Lessler, J. *et al.* (2016) 'Assessing the Global Threat from Zika Virus', *Science (New York, N.Y.)*, 353(6300), p. aaf8160. Available at: https://doi.org/10.1126/SCIENCE.AAF8160.

Leta, S. *et al.* (2018) 'Global risk mapping for major diseases transmitted by Aedes aegypti and Aedes albopictus'. Available at: https://doi.org/10.1016/j.ijid.2017.11.026.

Levi, J.E. *et al.* (2015) 'Real-time symptomatic case of transfusion-transmitted dengue', *Transfusion*, 55(5), pp. 961–964. Available at: https://doi.org/10.1111/trf.12944.

Lima, M. da R.Q. *et al.* (2019) 'The inability of a dengue NS1 ELISA to detect Zika infections', *Scientific Reports 2019 9:1*, 9(1), pp. 1–7. Available at: https://doi.org/10.1038/s41598-019-55160-4.

Lindenbach, B.D., Thiel, H.-J. and Rice, C.M. (2006) '33 Flaviviridae: The Viruses and Their Replication', in D.M. Knipe and O.M., Eds. Howley (eds). Philadelphia.

Lindenbach, B.D., Thiel, H.J. and Rice, C.M. (2011) 'Flaviviridae: The Viruses and Their Replication', in D.M. Knipe and O.M. Howley (eds) *Fields Virology*. 5th Edition. Philadelphia: Lippincot William & Wilkins, pp. 1101–1151.

Lindqvist, R., Upadhyay, A. and Överby, A.K. (2018) 'Tick-Borne Flaviviruses and the Type I Interferon Response', *Viruses 2018, Vol. 10, Page 340*, 10(7), p. 340. Available at: https://doi.org/10.3390/V10070340.

Lopez-Jimena, B. *et al.* (2018) 'Development and validation of four one-step real-time RT-LAMP assays for specific detection of each dengue virus serotype', *PLOS Neglected Tropical Diseases*, 12(5), p. e0006381. Available at: https://doi.org/10.1371/JOURNAL.PNTD.0006381.

Ludwig, G. v (1993) 'INSECT-TRANSMITTED VERTEBRATE VIRUSES: FLAVIVIRIDAE', *In Vitro Cell. Dev. Biol*, 29, pp. 296–309.

Luštrek, M. *et al.* (2013) 'Epitope Predictions Indicate the Presence of Two Distinct Types of Epitope-Antibody-Reactivities Determined by Epitope Profiling of Intravenous Immunoglobulins', *PLoS ONE*, 8(11), p. 78605. Available at: https://doi.org/10.1371/JOURNAL.PONE.0078605.

Lwande, O.W. *et al.* (2020) 'Globe-Trotting Aedes aegypti and Aedes albopictus: Risk Factors for Arbovirus Pandemics', *https://home.liebertpub.com/vbz*, 20(2), pp. 71–81. Available at: https://doi.org/10.1089/VBZ.2019.2486.

Macleod, J. and Gordon, W.S. (1932) 'Studies in Louping-Ill (An Encephalomyelitis of Sheep): II—Transmission by the Sheep Tick, Ixodes Ricinus L', *Journal of Comparative Pathology and Therapeutics*, 45, pp. 240–252. Available at: https://doi.org/10.1016/S0368-1742(32)80022-6.

Mandrekar, J.N. (2010) 'Receiver Operating Characteristic Curve in Diagnostic Test Assessment', *Journal of Thoracic Oncology*, 5(9), pp. 1315–1316. Available at: https://doi.org/10.1097/JTO.0B013E3181EC173D.

Mansfield, K. *et al.* (2009) 'Tick-borne encephalitis virus - a review of an emerging zoonosis', *The Journal of general virology*, 90(Pt 8), pp. 1781–1794. Available at: https://doi.org/10.1099/VIR.0.011437-0.

Mansfield, K.L. *et al.* (2011) 'Flavivirus-induced antibody cross-reactivity', *Journal of General Virology*, 92(12), pp. 2821–2829. Available at: https://doi.org/10.1099/vir.0.031641-0.

Marchand, E. *et al.* (2013) 'Autochthonous case of dengue in France, October 2013', *Eurosurveillance*, 18(50). Available at: https://doi.org/10.2807/1560-7917.ES2013.18.50.20661.

Marriott, L. *et al.* (2006) 'Detection of Louping ill virus in clinical specimens from mammals and birds using TaqMan RT-PCR', *Journal of*

Virological Methods, 137(1), pp. 21–28. Available at: https://doi.org/10.1016/j.jviromet.2006.05.025.

Matochko, W.L. *et al.* (2012) 'Deep sequencing analysis of phage libraries using Illumina platform', *Methods*, 58(1), pp. 47–55. Available at: https://doi.org/10.1016/j.ymeth.2012.07.006.

Mattingly, P.F. (1957) 'Genetical aspects of the aëdes aegypti problem', *Annals of Tropical Medicine and Parasitology*, 51(4), pp. 392–408. Available at: https://doi.org/10.1080/00034983.1957.11685829.

Michelitsch, A. *et al.* (2019) 'Exploring the Reservoir Hosts of Tick-Borne Encephalitis Virus'. Available at: https://doi.org/10.3390/v11070669.

Montoya, M. *et al.* (2018) 'Longitudinal Analysis of Antibody Cross-neutralization Following Zika Virus and Dengue Virus Infection in Asia and the Americas', *The Journal of infectious diseases*, 218(4), pp. 536–545. Available at: https://doi.org/10.1093/INFDIS/JIY164.

Mota, M.L. *et al.* (2021) 'Serological and molecular epidemiology of the Dengue, Zika and Chikungunya viruses in a risk area in Brazil', *BMC Infectious Diseases 2021 21:1*, 21(1), pp. 1–7. Available at: https://doi.org/10.1186/S12879-021-06401-3.

Müllbacher, A. and Lobigs, M. (1995) 'Up-regulation of MHC class I by flavivirus-induced peptide translocation into the endoplasmic reticulum', *Immunity*, 3(2), pp. 207–214. Available at: https://doi.org/10.1016/1074-7613(95)90090-X.

Musso, D. and Desprès, P. (2020) 'Serological Diagnosis of Flavivirus-Associated Human Infections', *Diagnostics*, 10(5). Available at: https://doi.org/10.3390/DIAGNOSTICS10050302.

Musso, D., Nilles, E.J. and Cao-Lormeau, V.-M. (2014) 'Rapid spread of emerging Zika virus in the Pacific area', *Clinical Microbiology and*

Infection, 20(10), pp. O595–O596. Available at: https://doi.org/10.1111/1469-0691.12707.

Mustafa, M.S. *et al.* (2015) 'Discovery of fifth serotype of dengue virus (DENV-5): A new public health dilemma in dengue control', *Medical journal, Armed Forces India*, 71(1), pp. 67–70. Available at: https://doi.org/10.1016/J.MJAFI.2014.09.011.

Naqid, I.A., Owen, J.P., Maddison, B.C., Spiliotopoulos, A., Emes, R.D., Warry, A., Flynn, R.J., *et al.* (2016) 'Mapping B-cell responses to Salmonella enterica serovars Typhimurium and Enteritidis in chickens for the discrimination of infected from vaccinated animals', *Scientific Reports*, 6. Available at: https://doi.org/10.1038/srep31186.

Naqid, I.A., Owen, J.P., Maddison, B.C., Spiliotopoulos, A., Emes, R.D., Warry, A., Tchórzewska, M.A., *et al.* (2016) 'Mapping polyclonal antibody responses to bacterial infection using next generation phage display', *Scientific Reports 2016 6:1*, 6(1), pp. 1–11. Available at: https://doi.org/10.1038/srep24232.

Navalkar, K.A. *et al.* (2014) 'Application of Immunosignatures for Diagnosis of Valley Fever', *Clinical and Vaccine Immunology : CVI*, 21(8), p. 1169. Available at: https://doi.org/10.1128/CVI.00228-14.

Norrby, E. (2007) 'Yellow fever and Max Theiler: the only Nobel Prize for a virus vaccine', *The Journal of Experimental Medicine*, 204(12), p. 2779. Available at: https://doi.org/10.1084/JEM.20072290.

Nuttall, P.A. *et al.* (1994) 'Adaptations of arboviruses to ticks', *Journal of medical entomology*, 31(1), pp. 1–9. Available at: https://doi.org/10.1093/JMEDENT/31.1.1.

Nuttall, P.A. and Labuda, M. (2003) 'Dynamics of infection in tick vectors and at the tick–host interface', *Advances in Virus Research*, 60, pp. 233–272. Available at: https://doi.org/10.1016/S0065-3527(03)60007-2.

Ortiz García, P.J. (2013) 'Tropical medicine. Surprising new dengue virus throws a spanner in disease control efforts', *Science (New York, N.Y.)*, 342(6157), pp. 76–77. Available at: https://doi.org/10.1126/SCIENCE.342.6157.415.

Pande, J., Szewczyk, M.M. and Grover, A.K. (2010) 'Phage display: Concept, innovations, applications and future', *Biotechnology Advances*, 28(6), pp. 849–858. Available at: https://doi.org/10.1016/J.BIOTECHADV.2010.07.004.

Parmley, S.F. and Smith, G.P. (1988) 'Antibody-selectable filamentous fd phage vectors: affinity purification of target genes', *Gene*, 73(2), pp. 305–318. Available at: https://doi.org/10.1016/0378-1119(88)90495-7.

Paulsen, K. *et al.* (2019) 'Experimental infection of lambs with tick-borne encephalitis virus and co-infection with Anaplasma phagocytophilum', *PloS one*, 14(12). Available at: https://doi.org/10.1371/JOURNAL.PONE.0226836.

Payne, A.F. *et al.* (2006) 'Quantitation of flaviviruses by fluorescent focus assay', *Journal of Virological Methods*, 134(1–2), pp. 183–189. Available at: https://doi.org/10.1016/J.JVIROMET.2006.01.003.

Pierson, T.C. and Diamond, M.S. (2020) 'The continued threat of emerging flaviviruses', *Nature Microbiology 2020 5:6*, 5(6), pp. 796–812. Available at: https://doi.org/10.1038/s41564-020-0714-0.

Pool, W.A., Brownlee, A. and Wilson, D.R. (1930) 'The etiology of "Louping-Ill"', *Journal of Comparative Pathology and Therapeutics*, 43, pp. 253–290. Available at: https://doi.org/10.1016/S0368-1742(30)80026-2.

Porterfield, J.S. (1986) 'The Viruses. The Togaviridae and Flaviviridae', in *The Viruses. The Togaviridae and Flaviviridae*, pp. 1–16. Available at: https://books.google.co.uk/books?id=dQbkBwAAQBAJ&pg=PA7&lpg=PA7&dq=jordi+casals+flavivirus&source=bl&ots=nq0JIF1b34&sig=ACfU3U0C

8s0RKeYOCc3GzePA_Xx2ceNyqA&hl=en&sa=X&ved=2ahUKEwi8m8qQj6jzAhUnREEAHZivAq8Q6AF6BAgMEAM#v=onepage&q=jordi%20casals%20flavivirus&f=false (Accessed: 1 October 2021).

Priyamvada, L. *et al.* (2016) 'Human antibody responses after dengue virus infection are highly cross-reactive to Zika virus', *Proceedings of the National Academy of Sciences of the United States of America*, 113(28), pp. 7852–7857. Available at: https://doi.org/10.1073/PNAS.1607931113.

Priyamvada, L. *et al.* (2017) 'Humoral cross-reactivity between Zika and dengue viruses: Implications for protection and pathology', *Emerging Microbes and Infections*. Nature Publishing Group. Available at: https://doi.org/10.1038/emi.2017.42.

Qi, H. *et al.* (2012) 'Phagemid Vectors for Phage Display: Properties, Characteristics and Construction', *Journal of Molecular Biology*, 417(3), pp. 129–143. Available at: https://doi.org/10.1016/J.JMB.2012.01.038.

Quail, M.A. *et al.* (2012) 'A tale of three next generation sequencing platforms: comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers', *BMC Genomics 2012 13:1*, 13(1), pp. 1–13. Available at: https://doi.org/10.1186/1471-2164-13-341.

Rajan, J. v *et al.* (2021) 'Phage display demonstrates durable differences in serological profile by route of inoculation in primary infections of non-human primates with Dengue Virus 1', *Scientific Reports* |, 11, p. 10823. Available at: https://doi.org/10.1038/s41598-021-90318-z.

Rathore, A.P.S. and st. John, A.L. (2020) 'Cross-Reactive Immunity Among Flaviviruses', *Frontiers in Immunology*, 11, p. 334. Available at: https://doi.org/10.3389/FIMMU.2020.00334/XML/NLM.

Ravichandran, S. *et al.* (2019) 'Differential human antibody repertoires following Zika infection and the implications for serodiagnostics and

disease outcome', *Nature Communications 2019 10:1*, 10(1), pp. 1–14. Available at: https://doi.org/10.1038/s41467-019-09914-3.

Ravn, U. *et al.* (2013) 'Deep sequencing of phage display libraries to support antibody discovery', *Methods*, 60(1), pp. 99–110. Available at: https://doi.org/10.1016/J.YMETH.2013.03.001.

Ray, D.S., Bscheider, H.-P. and Hofschneider, P.H. (1966) 'Replication of the single-stranded DNA of the male-specific bacteriophage M13: Isolation of intracellular forms of phage-specific DNA', *Journal of Molecular Biology*, 21(3), pp. 473–483. Available at: https://doi.org/10.1016/0022-2836(66)90020-9.

Rebollo, I.R. *et al.* (2014) 'Identification of target-binding peptide motifs by high-throughput sequencing of phage-selected peptides', *Nucleic acids research*, 42(22). Available at: https://doi.org/10.1093/NAR/GKU940.

Reed, W.;, Carroll, J.; and Agramonte (2001) *Military Medicine*.

Reid, H. *et al.* (1984) 'Transmission of louping-ill virus in goat milk', *Veterinary Record*, 114(7), pp. 163–165. Available at: https://doi.org/10.1136/vr.114.7.163.

Reid, H.W. (1975) 'Experimental infection of red grouse with louping-ill virus (Flavivirus group): I. The viraemia and antibody response', *Journal of Comparative Pathology*, 85(2), pp. 223–229. Available at: https://doi.org/10.1016/0021-9975(75)90063-8.

Reid, H.W. *et al.* (1978) 'Studies of louping-ill virus (Flavivirus group) in wild red grouse (Lagopus lagopus scoticus).', *The Journal of hygiene*, 81(2), pp. 321–9. Available at: http://www.ncbi.nlm.nih.gov/pubmed/212479 (Accessed: 27 September 2018).

Renato Dulbecco, B. (1952) 'Production of Plaques in Monolayer Tissue Cultures by Single Particles of an Animal Virus', *Proceedings of the National Academy of Sciences*, 38(8), pp. 747–752. Available at: https://doi.org/10.1073/PNAS.38.8.747.

Rey, F. *et al.* (2018) 'The bright and the dark side of human antibody responses to flaviviruses: lessons for vaccine design', *EMBO reports*, 19(2), pp. 206–224. Available at: https://doi.org/10.15252/EMBR.201745302.

Rice, C.M. *et al.* (1985) 'Nucleotide sequence of yellow fever virus: Implications for flavivirus gene expression and evolution', *Science*, 229(4715), pp. 726–733. Available at: https://doi.org/10.1126/SCIENCE.4023707.

Roth, A. *et al.* (2014) 'Concurrent outbreaks of dengue, chikungunya and Zika virus infections – an unprecedented epidemic wave of mosquito-borne viruses in the Pacific 2012–2014', *Eurosurveillance*, 19(41), p. 20929. Available at: https://doi.org/10.2807/1560-7917.ES2014.19.41.20929.

Russell, P.K., Chiewsilp, D. and Brandt, W.E. (1970) 'Immunoprecipitation Analysis of Soluble Complement-Fixing Antigens of Dengue Viruses', *The Journal of Immunology*, 105(4).

Růžek, D., Dobler, G. and Mantke, O.D. (2010) 'Tick-borne encephalitis: Pathogenesis and clinical implications', *Travel Medicine and Infectious Disease*, 8(4), pp. 223–232. Available at: https://doi.org/10.1016/J.TMAID.2010.06.004.

Ryvkin, A. *et al.* (2012) 'Deep Panning: Steps towards Probing the IgOme', *PLOS ONE*, 7(8), p. e41469. Available at: https://doi.org/10.1371/JOURNAL.PONE.0041469.

Sabin, A.B. (1950) 'The Dengue group of viruses and its family relationships', *Bacteriological Reviews*, 14(3), pp. 225–232. Available at: https://doi.org/10.1128/br.14.3.193-258.1950.

Sabin, A.B. (1952) *Research on Dengue during World War II.*, *American Journal of Tropical Medicine and Hygiene* . Available at: https://www.cabdirect.org/cabdirect/abstract/19522901271 (Accessed: 1 October 2021).

Sawyer, W.A. and Frobisher, M. (1929) 'THE FILTRABILITY OF YELLOW FEVER VIRUS AS EXISTING IN THE MOSQUITO', *Journal of Experimental Medicine*, 50(6), pp. 713–718. Available at: https://doi.org/10.1084/JEM.50.6.713.

Shriver-Lake, L.C. *et al.* (2018) 'Selection and Characterization of Anti-Dengue NS1 Single Domain Antibodies', *Scientific Reports 2018 8:1*, 8(1), pp. 1–11. Available at: https://doi.org/10.1038/s41598-018-35923-1.

Simmonds, P. *et al.* (2017) 'ICTV virus taxonomy profile: Flaviviridae', *Journal of General Virology*, 98(1), pp. 2–3. Available at: https://doi.org/10.1099/JGV.0.000672.

Singaravelan, B., Roshini, B.R. and Munavar, M.H. (2010) 'Evidence that the supE44 Mutation of Escherichia coli Is an Amber Suppressor Allele of glnX and that It Also Suppresses Ochre and Opal Nonsense Mutations', *Journal of Bacteriology*, 192(22), p. 6039. Available at: https://doi.org/10.1128/JB.00474-10.

Slon Campos, J.L., Mongkolsapaya, J. and Screaton, G.R. (2018) 'The immune response against flaviviruses', *Nature immunology*, 19(11), pp. 1189–1198. Available at: https://doi.org/10.1038/S41590-018-0210-3.

Smith, G.P. (1985) 'Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface', *Science*, 228(4705), pp. 1315–1317.

Smith, G.P. and Petrenko, V.A. (1997) 'Phage display', *Chemical Reviews*, 97(2), pp. 391–410. Available at: https://doi.org/10.1021/cr960065d.

Smither, S.J. *et al.* (2013) 'Comparison of the plaque assay and 50% tissue culture infectious dose assay as methods for measuring filovirus infectivity', *Journal of Virological Methods*, 193(2), pp. 565–571. Available at: https://doi.org/10.1016/J.JVIROMET.2013.05.015.

Stadler, K. *et al.* (1997) 'Proteolytic activation of tick-borne encephalitis virus by furin', *Journal of virology*, 71(11), pp. 8475–8481. Available at: https://doi.org/10.1128/JVI.71.11.8475-8481.1997.

Stafford, P. *et al.* (2012) 'Physical characterization of the "immunosignaturing effect"', *Molecular and Cellular Proteomics*, 11(4). Available at: https://doi.org/10.1074/mcp.M111.011593.

Stafford, P. *et al.* (2014) 'Immunosignature system for diagnosis of cancer.', *Proceedings of the National Academy of Sciences of the United States of America*, 111(30), pp. E3072-80. Available at: https://doi.org/10.1073/pnas.1409432111.

Stiasny, K., Holzmann, H. and Heinz, F. (2009) 'Characteristics of antibody responses in tick-borne encephalitis vaccination breakthroughs', *Vaccine*, 27(50), pp. 7021–7026. Available at: https://doi.org/10.1016/J.VACCINE.2009.09.069.

Stokes, A., Bauer, J.H. and Hudson, X.P. (1928) 'Experimental Transmission of Yellow Fever to Laboratory Animals.', *American Journal of Tropical Medicine*, 8(2).

Stopar, D. *et al.* (2003) 'Protein-lipid interactions of bacteriophage M13 major coat protein.', *Biochimica et biophysica acta*, 1611(1–2), pp. 5–15. Available at: https://doi.org/10.1016/S0005-2736(03)00047-6.

Tabachnick, W.J. (1991) 'Evolutionary Genetics and Arthropod-borne Disease: The Yellow Fever Mosquito', *American Entomologist*, 37(1), pp. 14–26. Available at: https://doi.org/10.1093/AE/37.1.14.

Tahir, D. *et al.* (2020) 'Interrupted Blood Feeding in Ticks: Causes and Consequences', *Microorganisms*, 8(6), pp. 1–12. Available at: https://doi.org/10.3390/MICROORGANISMS8060910.

Tarr, A.W. *et al.* (2006) 'Characterization of the hepatitis C virus E2 epitope defined by the broadly neutralizing monoclonal antibody AP33', *Hepatology*, 43(3), pp. 592–601. Available at: https://doi.org/10.1002/HEP.21088.

Theiler, M. and Downs, W.G. (1973) 'The arthropod-borne viruses of vertebrates. An account of The Rockefeller Foundation Virus Program, 1951-1970.', *The arthropod-borne viruses of vertebrates. An account of The Rockefeller Foundation Virus Program, 1951-1970.* [Preprint].

Theiler, M. and Smith, H.H. (1937) 'THE EFFECT OF PROLONGED CULTIVATION IN VITRO UPON THE PATHOGENICITY OF YELLOW FEVER VIRUS', *The Journal of Experimental Medicine*, 65(6), p. 767. Available at: https://doi.org/10.1084/JEM.65.6.767.

Tikunova, N. v and Morozova, V. v (2009) 'Phage display on the base of filamentous bacteriophages: application for recombinant antibodies selection.', *Acta naturae*, 1(3), pp. 20–8. Available at: http://www.ncbi.nlm.nih.gov/pubmed/22649612 (Accessed: 30 September 2018).

Tsoumpeli, M.T. *et al.* (2022) 'A Simple Whole-Plasmid PCR Method to Construct High-Diversity Synthetic Phage Display Libraries', *Molecular Biotechnology*, 64, pp. 791–803. Available at: https://doi.org/10.1007/s12033-021-00442-4.

Twort, F.W. (1915) 'AN INVESTIGATION ON THE NATURE OF ULTRA-MICROSCOPIC VIRUSES.', *The Lancet*, 186(4814), pp. 1241–1243. Available at: https://doi.org/10.1016/S0140-6736(01)20383-3.

Tzagoloff, H. and Pratt, D. (1964) 'The initial steps in infection with coliphage M13', *Virology*, 24(3), pp. 372–380. Available at: https://doi.org/10.1016/0042-6822(64)90174-6.

Vasilakis, N. *et al.* (2011) 'Fever from the forest: prospects for the continued emergence of sylvatic dengue virus and its impact on public health', *Nature Reviews. Microbiology*, 9(7), p. 532. Available at: https://doi.org/10.1038/NRMICRO2595.

Wallenhammar, A. *et al.* (2020) 'Revealing new tick-borne encephalitis virus foci by screening antibodies in sheep milk', *Parasites & Vectors 2020 13:1*, 13(1), pp. 1–12. Available at: https://doi.org/10.1186/S13071-020-04030-4.

Walter Reed, B. *et al.* (1900) 'The Etiology of Yellow Fever—A Preliminary Note', *Public Health Papers and Reports*, 26, p. 37. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2329228/ (Accessed: 2 September 2022).

Wee, S. *et al.* (2019) 'Multiplex targeted mass spectrometry assay for one-shot flavivirus diagnosis', *Proceedings of the National Academy of Sciences of the United States of America*, 116(14), pp. 6754–6759. Available at: https://doi.org/10.1073/PNAS.1817867116/SUPPL_FILE/PNAS.1817867116.SD12.TXT.

Westaway, E.G. *et al.* (1985) 'Flaviviridae', *Intervirology*, 24(4), pp. 183–192. Available at: https://doi.org/10.1159/000149642.

WHO (2016) *Situation Report Zika Virus;Microcephaly; Guillan Barre syndrome*. Available at:

http://apps.who.int/iris/bitstream/10665/246091/1/WHO-ZIKV-SRF-16.3-eng.pdf?ua=1&ua=1 (Accessed: 3 October 2018).

Wilder-Smith, A. *et al.* (2014) 'The 2012 dengue outbreak in Madeira: exploring the origins', *Euro surveillance : bulletin Europeen sur les maladies transmissibles = European communicable disease bulletin*, 19(8). Available at: https://doi.org/10.2807/1560-7917.ES2014.19.8.20718.

Wodak, E. *et al.* (2011) 'Detection and molecular analysis of West Nile virus infections in birds of prey in the eastern part of Austria in 2008 and 2009', *Veterinary Microbiology*, 149(3–4), pp. 358–366. Available at: https://doi.org/10.1016/J.VETMIC.2010.12.012.

World Health Organisation (2021) *Dengue and severe dengue*, *WHO Factsheet*. Available at: https://www.who.int/news-room/fact-sheets/detail/dengue-and-severe-dengue (Accessed: 18 November 2021).

World Health Organization (2009) *Dengue guidelines for diagnosis, treatment, prevention and control.*, *WHO guideline*. Available at: www.who.int/tdr (Accessed: 1 October 2021).

Xiao, P. *et al.* (2018) 'Metagenomic Analysis of Flaviviridae in Mosquito Viromes Isolated From Yunnan Province in China Reveals Genes From Dengue and Zika Viruses', *Frontiers in Cellular and Infection Microbiology*, 8, p. 359. Available at: https://doi.org/10.3389/FCIMB.2018.00359/BIBTEX.

Yap, T.L. *et al.* (2021) 'Engineered NS1 for Sensitive, Specific Zika Virus Diagnosis from Patient Serology - Volume 27, Number 5—May 2021 - Emerging Infectious Diseases journal - CDC', *Emerging Infectious Diseases*, 27(5), pp. 1427–1437. Available at: https://doi.org/10.3201/EID2705.190121.

Ytrehus, B. *et al.* (2021) 'LOUPING-ILL VIRUS SEROSURVEY OF WILLOW PTARMIGAN (LAGOPUS LAGOPUS LAGOPUS) IN NORWAY', *Journal of wildlife diseases*, 57(2), pp. 282–291. Available at: https://doi.org/10.7589/JWD-D-20-00068.

Zhang, X. *et al.* (2012) 'Mimotopes selected by biopanning with high-titer HIV-neutralizing antibodies in plasma from Chinese slow progressors', *Brazilian Journal of Infectious Diseases*, 16(6), pp. 510–516. Available at: https://doi.org/10.1016/J.BJID.2012.07.003.