FIND
Diagnosis for all

Next generation sequencing for viral haemorrhagic fevers

ACKNOWLEDGEMENTS

This report was developed by PHG Foundation for FIND (the Foundation for Innovative New Diagnostics).

We would like to thank all those who contributed to the development and review of this report. Expert interviewees are acknowledged in Appendix 8.4.

Lead writers:

Chantal Babb de Villiers (PHG Foundation)

Laura Blackburn (PHG Foundation)

Sarah Cook (PHG Foundation)

Joanna Janus (PHG Foundation)

Emma Johnson (PHG Foundation)

Reviewers:

Devy Emperador (FIND)

Jilian Sacks (FIND)

Marva Seifert (FIND/UCSD)

Anita Suresh (FIND)

Swapna Uplekar (FIND)

Publication date: 30 November 2020

Table of Contents

1 INTRODUCTION

In the past few years, pathogen genome sequencing has emerged as a tool to support understanding of the molecular epidemiology of disease outbreaks, complementing and in some cases supplanting more established techniques. Recent advances in sequencing technologies have shown their applicability for research use in outbreak situations, for example, sequencing has been deployed in recent years to support understanding of the epidemiology of Ebola virus disease [1], Zika virus [2] and Lassa fever [3].

However, lack of standardisation in testing and analysis protocols, the choice of reagents and the complexity of current methods can complicate the use of next generation sequencing (NGS) technologies in a range of applications, including as tools for routine surveillance, diagnostics and molecular epidemiology.

FIND (Foundation for Innovative New Diagnostics) aims to provide a comprehensive set of standardised sequencing protocols and workflows in order to support the implementation of NGS technologies and tools for outbreak pathogens in low resource settings. To facilitate this effort, investigation of the scientific landscape of pathogen sequencing as applied to outbreaks is necessary, to identify the most promising applications and to identify any existing workflows and protocols.

This report provides an overview of how sequencing technologies are being used, and could be used in the future, to understand and contribute to efforts to mitigate outbreaks for three viral haemorrhagic fevers of international importance:

- Lassa virus
- Ebola virus
- Yellow fever virus

1.1 Methods

This report summarises current knowledge on technology developments, protocols and global best practice. It is based on desk-based research and analysis informed by official publications, grey literature, peer-reviewed and pre-print literature.

Where appropriate, in-depth interviews (via telephone or video conference) have been conducted with experts and other relevant stakeholders to better understand the enablers and barriers to implementation and adoption of the techniques in appropriate settings. These experts are acknowledged in Appendix 8.4.

Subsequent chapters in this report will cover the following areas:

- Introduction to viral pathogen genomics, including sequencing technologies used to sequence genomes, the applications of these approaches and a summary of diagnostic techniques used on viral haemorrhagic fevers.
- Pathogen specific chapters on Lassa fever, Ebola virus disease and Yellow fever. Each pathogen chapter contains:
    - Disease and pathogen background
    - Research and development needs for disease control
    - Current diagnostics landscape
    - Current uses of NGS, including any protocols available
    - Future needs of NGS, including a gap analysis of where NGS could be applied
- Cross-cutting themes relevant to all three diseases, including where NGS applications intersect, choice of technologies and cost implications
- Discussion and conclusions

A note on terminology

The terms lineage, species and strain are used in the scientific literature to describe different forms of viruses causing disease and these can be used differently and/or interchangeably depending on the pathogen under discussion. For Lassa virus, the term lineage is used to describe the different forms of the virus currently described. The ebolavirus genus is commonly described as having six species with the term strain used to differentiate species sub-types. Therefore, for each pathogen chapter we have used terms as they are most commonly used in the research literature for each virus, in order to maintain consistency.

2 SEQUENCING AND DIAGNOSTIC TECHNOLOGIES

2.1 Sequencing technology platforms and approaches

There are several approaches and technologies that can be deployed for viral genome sequencing – which ones are chosen will be dependent on resource availability and the goals of undertaking sequencing.

There are a number of challenges associated with sequencing viral genomes, including [4]:

- A range of genome types – DNA or RNA, single or double stranded – each of which require different extraction protocols
- High mutation rates and genetic heterogeneity
- Lack of homologous genes between viruses of different phyla, which prevents comparison or surveys of viruses using universal primers
- Viruses are challenging to culture and are often heavily contaminated with host nucleic acid, which can negatively impact sequencing output
- Methods for DNA digestion of host genetic material that rely on methylation patterns are often less effective since DNA viruses in particular often have similar methylation patterns to host DNA.

Despite these challenges, rapid progress is being made in the techniques and technologies being utilised to sequence viral pathogen genomes. For example, metagenomic NGS (mNGS) approaches and long read sequencing are facilitating sequencing directly from patient samples. For each technique, whole or partial genome sequencing can be deployed.

Techniques for viral sequencing

Amplicon sequencing relies on amplification of regions of the viral genome using primers that are complementary to a specific genetic sequence. This allows for deep sequencing of these regions. This approach is often used to detect rare variants that may be present in only a small proportion of the genomes in a sample, and when conducting surveillance or species identification studies. Dependent upon primer design, amplicon-based sequencing can be used for targeted sequencing of selected regions of the genome, or multiple, often overlapping, amplicons can be used to allow for whole genome sequencing. This is particularly useful for viruses with small genomes such as HIV or influenza that require only a few amplicons to cover the whole genome.

Hybrid capture (or target enrichment/bait capture) is typically used to obtain sequencing data for single pathogens from clinical samples. It involves using RNA or DNA probes that are complementary to one or more reference sequences from pathogen(s) of interest. This facilitates the selection and targeting of specific regions of the genome for further amplification and sequencing. This method can be used directly on clinical samples without a need for culture (in some cases) or PCR and can capture the diversity within clinical samples, depending on the probes available. This approach can also be used for metagenomics sequencing. Similar to amplicon sequencing, for small viral genomes hybrid capture methods can be used to sequence either specific regions of the genome, or multiple probes can be used to obtain whole genome coverage.

In addition, for RNA viruses, **RNA sequencing (RNA-Seq)** involves the preparation of a complementary DNA (cDNA) library for sequencing; **direct-RNA sequencing** removes the need for amplification and conversion of RNA into cDNA and can only be performed using nanopore sequencing techniques. Removal of amplification steps means that some bias that would normally be introduced through PCR is avoided, potentially reducing error. It also allows base modifications (such as methylation) to be detected on the unconverted molecule. However, in order to achieve sufficient coverage of the viral genome, a high quality and RNA-rich sample is required.

Metagenomic sequencing – the metagenome includes all nucleic acids present in a sample at one time – this may be restricted to nucleic acids of one type i.e. RNA or DNA. All DNA and/or RNA are extracted from the sample – this will include material from the pathogen of interest as well as the host and any other microorganisms present. Metagenomic sequencing can be targeted or unbiased and detect pathogens direct from a clinical sample (i.e. uncultured) or cultured samples. Targeted sequencing will involve using primers that correspond to one or more genes in multiple organisms. Unbiased sequencing approaches amplify all RNA and/or DNA in a sample. Following library preparation, sequencing can take place by 'shotgun sequencing' whereby all available regions of all nucleic acids are targeted, or by RNA sequencing. Alternatively an additional target enrichment step for high-sensitivity pathogen identification in low-titre infections can be done. Additional host depletion methods can also be used to overcome the sample being dominated by the host's DNA/RNA.

Metagenomic next generation sequencing (mNGS) can be useful when one of a number of illnesses with similar symptoms are suspected, or when co-infection is possible, enabling the identification of multiple pathogens from a single sample. mNGS could also have utility to help ascertain the emergence of a new pathogen, or provide preliminary data on pathogen identity. However, the effectiveness of this approach relies on the public availability of sequence data from a range of pathogens, which provide comparative reference genomes to support identification.

Table 1: Advantages and disadvantages of different viral sequencing approaches (adapted from [4]).

Method	Advantages	Disadvantages
Metagenomic sequencing	-Can sequence novel or poorly characterised genomes -Simple cost-effective sample preparation -Effective in hypothesis free, unbiased approaches to identify a potential underlying pathogen -Lower required number of PCR cycles causes few amplification mutations -Preservation of minor variant frequencies reflects *in vivo* variation -No primer or probe design required, which enables a rapid response to novel pathogens or sequence variants	-High sequencing cost to obtain sufficient data; data analysis also costly -Relatively low sensitivity to target pathogen -Incidental sequencing of human and off-target pathogens raises ethical and diagnostic issues -High proportion of non-pathogen reads increases computational challenges -Coverage is proportional to viral load -Prone to contamination from environmental species and cross-contamination
PCR amplification sequencing	-Tried and trusted well-established methods and trained staff -Highly specific; most sequencing reads will be pathogen-specific, which decreases sequencing costs -Highly sensitive, with good coverage even at low pathogen load -Relatively straightforward design and application of new primers for novel sequences	-Labour intensive and difficult to scale for large genomes -PCR reactions are subject to primer mismatch, particularly in poorly characterised or highly diverse pathogens, or those with novel variants -Iterating standard PCRs across large genomes requires high sample volume -Uneven amplification of different PCR amplicons may influence minor variant and haplotype reconstruction -High number of PCR cycles may introduce amplification mutations -Limited ability to sequence novel pathogens
Target enrichment sequencing	-Single tube sample preparation that is suited to high-throughput automation and the sequencing of large genomes -Higher specificity than metagenomics decreases sequencing costs -Overlapping probes increases tolerance for individual primer mismatches -Fewer PCR cycles (than PCR amplification) limits the introduction of amplification mutations -Preservation of minor variant frequencies reflects *in vivo* variation	-Unable to sequence novel pathogens and requires well-characterised reference genomes for probe design -High cost and technical expertise for sample preparation -Cost and time to generate new probe sets limit a rapid response to emerging and novel viruses -Sensitivity is comparable to PCR, but coverage is proportional to pathogen load; low pathogen load yields low and incomplete coverage

Sequencing systems

There are currently three generations of available sequencing technologies: first generation including Sanger sequencing and associated technologies; second generation comprising next generation sequencing (NGS) short-read technologies; and third generation long-read NGS technologies.

First generation: Sanger and other technologies. Sanger sequencing is not suitable for sequencing at scale, particularly in an outbreak situation, due to the lower throughput of the technique compared to NGS technologies. However, due to the technique's accuracy it is often used in the early stages of outbreaks (with other NGS techniques) to confirm consensus sequences [5].

Next generation sequencing – second generation high-throughput short read technologies

Sequencing by synthesis (Illumina). Sequencing by synthesis represents the basis of the most widely used NGS methods. Illumina systems provide high throughput short-read sequencing.

Sequencing by synthesis nucleotide identification occurs as modified nucleotides are incorporated into newly forming DNA. Fluorescently tagged (modified) bases are detected as they are incorporated. Unlike early chain-termination methods (Sanger sequencing) fluorescently tagged bases do not cause DNA synthesis to stop. Each time a base is incorporated, the attached fluorescent tags are washed away after detection, allowing for more modified bases to be added after this point. The process is repeated until the maximum number of cycles (and therefore sequence length) is reached.

This method requires the prior conversion of RNA molecules into cDNA (performed during library preparation) for bridge amplification, which is a key step in the process.

Advantages of Illumina sequencing

- High accuracy
- One of the more commonly used systems for high resolution genomic analysis - used in around 115 countries – meaning collaborative development of expertise and advancements is easily possible, many genetic or research laboratories may already possess these systems, and bioinformatic pipelines are relatively well-established
- High levels of sample multiplexing are possible, meaning a high number of samples can be run at once

Limitations

- Longer time required for sequencing runs
- Most platforms are large in size, some require specialised infrastructure for safe use
- Relatively short reads may limit accuracy in certain genomic regions

Ion torrent semiconductor sequencing (Thermo Fisher Scientific). Thermo Fisher supplies semi-conductor sequencing through its Ion range of systems. Sequencing utilises a semi-conductor chip and a bead-based system which provide a physical platform for sequencing through DNA synthesis. This system relies on the detection of hydrogen ions released as nucleotides are incorporated.

Advantages of Ion Torrent sequencing

- Comparatively short sequencing runs enable faster return of results
- Low substitution error rate
- Some systems facilitate a highly automated workflow for easy adoption and consistent application of sequencing

Limitations

- Lower throughput in comparison to other NGS technologies
- Shorter reads than are possible with other NGS technologies

DNA nanoball sequencing (Beijing Genomics Institute and MGI Tech). DNA Nanoball sequencing (DNBSEQ) platforms produced by MGI Tech are available through the Beijing Genomics Institute (BGI). DNBSEQ utilises circularised reads which are repeatedly amplified using rolling consensus amplification to create a single long strand of DNA. A barcode and

primers, which also enable circularisation, are attached to the target sequence during library preparation. The sequence is then massively amplified forming what are known as DNA 'nanoballs' (DNBs). The DNBs are then loaded onto a flow cell with embedded wells which facilitate detection of nucleotide integration through light detection in a similar manner to sequencing by synthesis. The repeated nanoball sequence is read to generate high accuracy consensus sequence data.

Advantages of BGI and MGI technologies

- Flexible sequencing including range of run times, reads lengths and output
- High throughput
- Linear amplification reduces error accumulation during amplification

Limitations

- Shorter reads than are possible with other NGS technologies
- Highest throughput systems are very large and will require appropriate laboratory space

Next generation sequencing – third generation long read technologies.

Long read single molecule sequencers use distinct base technologies to read longer contiguous strands of DNA than other NGS sequencing platforms. Reads of 10,000 – 100,000 base pairs (bp) in length are produced, with the potential for molecules of more than 100,000 bp to be read contiguously. As read lengths increase, it is more likely that a read will be distinct from other reads. This allows them to be computationally reassembled with less ambiguity. This is particularly useful for sequencing of highly polymorphic or highly repetitive genomes.

Single molecule real-time sequencing (Pacific Biosciences). Pacific Biosciences' (PacBio) sequencing technologies use single molecule real-time (SMRT) sequencing to produce high accuracy long-read sequence data.

Target nucleic acid molecules are individually immobilised in microscopic pits called zero-mode wave guides (ZMWs) embedded in SMRT flow cells. The many embedded ZMWs contain a fixed DNA polymerase and are open to a pool of free labelled nucleotides. Sequence information is collected through the high-precision detection of miniscule emissions of light which are produced as complementary nucleotides are incorporated into the target molecule.

Advantages of PacBio sequencing

- Capable of high throughput, equivalent to that of Illumina sequencing platforms
- Capable of producing very high accuracy consensus reads – highly accurate long reads (HiFi) sequencing reads around 15,000 bases in length at over 99% accuracy
- Produces long reads
- Errors are random, not systematic, and can therefore be overcome with deeper sequencing
- Sequences read in real time – allowing for termination when user determines enough reads have been generated

Limitations

- Systems are larger in size than alternatives, requiring appropriate laboratory space

Nanopore sequencing (Oxford Nanopore Technologies). ONT's systems are designed to be relatively mobile, generate ultra-long reads and be more accessible to those with less technical experience and expertise.

Extracted nucleic acids are prepared for sequencing by ligation of a motor protein and adapter sequence at the ends of each strand. RNA may either be amplified and converted to

cDNA prior to sequencing as occurs with other sequencing systems or the RNA can be read directly – direct RNA sequencing – without prior amplification or conversion to cDNA.

During sequencing, tagged, single stranded DNA or RNA molecules are fed through a membrane-bound protein pore – a 'nanopore' – by a motor protein. As each DNA or RNA nucleotide is fed through the nanopore, it interrupts the electrical current that exists across the pore, these signals are detected by the sequencing system. The pattern of disruption can be read to determine the base sequence.

Advantages of Oxford Nanopore Technologies sequencing

- Rapid and flexible - particularly useful for sequencing smaller genomes
- Mobile sequencing – the small size and high portability of some systems means that these can be used in the field
- Many reagents do not require cold storage meaning they can be safely stored in environments where refrigeration is not possible or unreliable
- Simpler user interface and analysis platforms
- Simultaneous examination of methylation possible using direct RNA sequencing

Limitations

- Limited barcoding means that it is more challenging to run multiple samples in one run, increasing time taken to run these samples. Currently the mobile sequencing units are not capable of providing the same level of multiplexing as other NGS technologies.
- Raw signal output files are very large – this makes files difficult to store. As software and pipelines for analysis evolve rapidly, it is useful if not essential for these files to be available for subsequent, quality analysis of the data. This also hinders depositing of raw data onto online databases.
- Higher error rate in homopolymeric regions

2.2 Applications of viral pathogen genome sequencing

There are three broad areas, outlined below, where viral pathogen genome sequencing has utility.

Surveillance

Identifying disease origins: both zoonotic origins and of outbreaks within human populations

Transmission dynamics: tracking person-person or person-environmental transmission, understanding in which situations transmission is occurring (community/hospital/environmental) and local/national/international transmission routes

Examining viral population structure: to monitor viral evolution within or between people, geographical regions, or through time.

Tracking disease prevalence: the frequency of infections in a population at a particular point in time

Outbreak investigation: when used in conjunction with more traditional epidemiology techniques such as contact tracing, the higher resolution information provided by sequencing can inform genomic epidemiology, by helping to: unravel complex transmission events, disprove suspected transmission events, identify disease 'super-spreaders'. Sequencing can also support detection of novel pathogens and species identification.

Use of sequencing in these areas can be deployed responsively as new outbreaks emerge but also as part of an established pathogen surveillance service that monitors diseases of importance.

Diagnosis and infection management

Pathogen diagnosis: Genome sequencing can be deployed in cases where conventional methods have not identified a pathogen causing a disease.

Detection of mixed infections/co-infection: particularly those infections with similar clinical presentations, this will support clinical decision making around treatment.

Outbreak management: in healthcare settings, particularly hospitals, use of sequencing can support conventional infection control efforts to determine the sources and spread of infection. This also allows comparison of sequences collected from patients to determine the origin of infection and who is infecting who, to understand transmission events (regional spread, local vs. imported transmission), identification of infection 'hotspots', and to inform mitigation strategies.

Antimicrobial resistance: Sequencing can be used to understand the genetic mechanisms underlying drug resistance or susceptibility, in particular to select appropriate therapies depending on the pathogen's drug susceptibility profile.

Determine vaccine suitability: For those pathogens where a vaccine is available, sequencing data can be used to assess if the vaccine(s) available are still able to protect against the specific viral strain that is currently causing disease.

Research and development

Development of novel therapeutics and vaccines: including identification of novel drug targets, detecting possible drug resistance, understanding virus susceptibility to treatment, informing the design of vaccines.

Informing diagnostic tests: WGS benefits diagnostics by providing up to date information on viral genomes and how they are evolving. This information can be used to ensure that the primer sequences used in PCR-based tests, and other molecular diagnostics that rely on genomic information, are accurate and will enable detection of the pathogen variants in circulation. This is important to support outbreak management and ongoing clinical care of patients. Sequencing is used to identify regions of the genome for use in diagnostic testing, and also to monitor genetic changes in these regions to ensure that diagnostic tests remain effective.

Characterisation of virulence factors and resistance markers: which can support understanding of how the virus might be causing disease in the human population, or potential therapeutics that may or may not be effective.

Improving understanding of virus biology: including viral mutation rates and rare variant detection, to increase understanding of how the virus might change in response to selection pressure and provide information on variants that may start circulating in populations as new infection clusters emerge.

Understanding the link between host genomics and disease severity and susceptibility: which will have an impact on the development of future therapeutic approaches and could help to stratify patients for preventive or treatment measures.

2.3 Databases for pathogen genome sequence data

Genome sequences of the pathogens described in this report are available in a number of publicly available databases. The most widely used genome sequence databases are those covered by the International Nucleotide Sequence Databases (INSD) consisting of: NCBI GenBank; the DNA Databank of Japan (DDBJ); and European Molecular Biological Laboratory (EMBL). Together these provide the principal repositories for DNA sequence data. The Sequence Read Archive (SRA) is also included in the INSD. 95% (705 out of 743)

of nucleic sequence data databases directly link to or download nucleic sequence data from the INSD [6].

The National Institute of Allergy and Infectious Diseases (NIAID) in the US funds the Virus Pathogen Resource (ViPR) which is a freely available resource that supports research into the viral pathogens that present a risk to public health, listed on the NIAID category A-C priority pathogen lists. This list includes pathogens causing (re)emerging infectious diseases, and viral haemorrhagic fevers. ViPR integrates data from external sources (e.g. NCBI GenBank, UniProt – an online resource of information on protein sequence and function) and direct submissions, and also provides curation and analysis pipelines, as well as bioinformatics analysis and visualisation tools [7]. This database has the most extensive and accessible coverage of VHF sequences.

Further information on sequence databases, including an overview of how they are operated and managed, can be found in the recent UN Convention on Biological Diversity report entitled *Combined study on digital sequence information in public and private databases and traceability* [6]. While sharing of sequence data via databases can facilitate research across different specialisms and geographies, there is currently inconsistency in the use of databases between different national and international organisations and between research groups.

The databases listed above have different data sharing policies and are also capable of storing different forms of the data, such as:

- Unprocessed, raw electrical signal files
- Raw reads
- Processed data
- Metadata
- Short reads
- Genome assemblies

Submission of data is voluntary but researchers are highly encouraged to share data. Selection of which database to release data on can be determined by a range of factors, including:

- The regulatory framework
- Sequencing data available
- Preferred form of data to upload
- Familiarity with the database

Each of these factors will have an impact on the ability of researchers and public health authorities in different jurisdictions to optimise use of pathogen sequence data.

The data deposited in public sequence databases, while comprehensive, can contain more errors than curated, but more limited, databases such as the FDA-ARGOS or the FDA Reference Viral database (RVDB) [8]. The genomes available in the FDA-ARGOS database meet the quality metrics defined for providing reference-grade genomes for regulatory use. The availability of high quality reference genomes is important to support alignment of sequence data which facilitates accurate clinical interpretation of results. This is particularly important when considering unbiased metagenomic NGS – if a pathogen(s) does not have a universally accepted reference genome for mNGS it is challenging to compare assay performance between different laboratories. These databases are also continuously being updated, in order to correct misannotations and increase the representation of organisms – both in terms of numbers (and quality) of sequences per species and also in terms of the number of species that have sequences available.

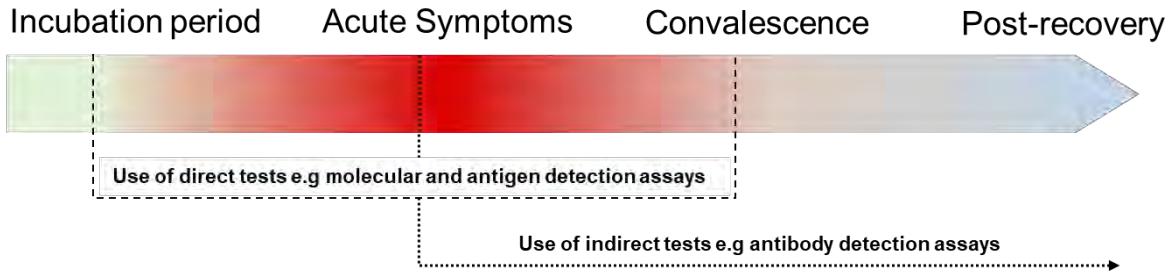
2.4 Summary of commonly used diagnostic methods for viral haemorrhagic fevers

There are several different methods that can be used to diagnose VHF infections, which can be divided into:

Direct methods: These detect virus directly and include viral culture and detection of viral nucleic acids or proteins (including antigens). These tests require a sufficient amount of virus to be present in a sample in order for the chosen assay method to work; their use is therefore determined by the viral load, which can vary between patients and over disease course, as well as the sensitivity of the chosen test method. Typically, the use of direct tests is limited to the early symptomatic stages of infection when viral load tends to be highest. As they directly detect the virus they may be more accurate in diagnosing a current infection than indirect methods.

Indirect methods: The most commonly used are serological tests to measure antibodies produced in response to the virus. These assays are not useful for early stages of infection before antibodies have been generated. However they can be used over a longer time period after the virus itself is no longer detectable, for example to diagnose past infections which is useful for surveillance purposes. A potential limitation is that in some cases a patient may already have antibodies from previous infections or vaccination, leading to false positive results when diagnosing current infections.

Figure 1: General illustration of how direct and indirect test methods are suited for use at different stages of infection. The precise time-points and stages of infection for which different methods are suitable will vary depending on the characteristics of each virus, as well as which tests are available.



Viral culture

Viral culture is traditionally considered the gold standard method to allow virus isolation and characterisation. A patient sample such as tissue or blood is added to specific types of mammalian cells in laboratory culture, and the type of virus is often identified by analysis of the morphological changes in infected cells [9]. Techniques such as immunoassays, western blot, RT-PCR, and sequencing can be used to confirm and refine results. Whilst still used in specific circumstances such as hard-to-diagnose infections and for characterising new strains, the use of more rapid diagnostics means that today viral culture is rarely used as a first line test [10].

Advantages:

- Independent of genetic variation among virus lineages and strains
- Can be used to detect multiple viruses
- Can be used to determine viral load as well as allowing further characterisation and study of virus

Limitations:

- May require higher levels of biocontainment facilities, often biosafety level four (BSL-4) for suspected haemorrhagic fevers, making it unsuitable for lower resource settings
- Due to the skills and materials needed it is often only available in a few reference laboratories
- It can take a week or more for effective viral growth which is impractical in endemic areas or where speed is required.

Immunological tests

Immunological tests either detect antigens or antibodies to diagnose a viral infection:

Antigen detection tests: These tests detect viral antigens and so are a direct method of testing, used in the early stages of infection when the virus is still present. The tests often make use of antibodies to detect the antigens, and can be used on multiple body fluid sample types containing the virus, as well as tissue infected with the virus.

Antibody detection tests: These tests detect IgM and IgG antibodies produced in response to a viral infection and so are an indirect method of testing, used once antibodies have been produced in the later stages of infection, or after the infection has resolved. How long antibodies can be detected following an infection can range from months to several years, depending on the type of virus, the type of antibody, and the individual patient. These tests often make use of viral antigens, which the antibodies bind to. Unlike antigen tests they primarily require blood serum or plasma samples, where the antibodies circulate. Some severe infections can suppress the production of IgG and/or IgM and cause false-negative results.

There are several further advantages and disadvantages inherent to both types of immunological tests:

Advantages:

- Typically highly sensitive in detecting multiple viral strains (via use of different tests)
- Can be highly specific if used in appropriate situations where other viruses which may cross react are not present (see Limitations)
- Detection of antigens or antibodies not susceptible to the contamination issues that can affect PCR assays
- Some types of methods can be designed for rapid detection in low resource settings (table 2).

Limitations:

- Some tests are prone to cross reactivity with several viruses of the same class, as well as vaccine strains, due to the presence of similarly structured antigens and anti-viral antibodies. This can limit test specificity under certain circumstances and lead to false positive results and a requirement for further confirmatory tests.
- For some viruses, the antigens or antibodies produced by different virus strains may be sufficiently different to require the use of separate tests for each strain, requiring knowledge of which strain to test for.
- There is the potential for tests to become ineffective and produce false negative results if a virus mutates enough to alter antigen/antibody structure.
- The proteins used in immunological tests often require refrigerated storage, limiting their use in the field.

Table 2: Common immunological assays used in VHF diagnosis, and their specific advantages and disadvantages.

Assay	Advantages	Limitations
ELISA (Enzyme-linked immunosorbent assay) -Most commonly used as an indirect test to detect viral antibodies, also used to detect antigens -Relies on complementary binding of antigens/antibodies to the ligand of interest, indicated by enzyme catalysed colour change	-Higher throughput, up to 384 well plates -Can quantitate ligand of interest -Multiple sample types can be used; sample inactivation allows testing in BSL-2 -Simple procedure, cheap once developed -Rapid (3-48 h)	-Laboratory preparation of antigens/antibodies and test validation can be complex, labour intensive and expensive -Techniques and reagents required can limit use to reference laboratories -Lack of commercial tests leads to lack of standardisation in laboratory assays
Plaque Reduction Neutralization Tests (PRNT) [11] -An indirect test for the presence of virus neutralising antibodies Used as a confirmatory test when ELISAs inconclusive -Relies on viral culture techniques	-Considered gold standard confirmatory test: higher sensitivity and specificity -Can quantify level of virus neutralising antibodies in sample	-Viral culture requires high BSL laboratories -Time consuming test turnaround (a few days) -Cross reactivity can occur e.g. from previous infections or vaccination
Immunohistochemistry -Typically a direct test, uses labelled antibodies used to bind and visualise viral antigens in sample -Commonly used post-mortem, also to characterise cell cultures	-Antigen detection plus morphological analysis of tissue increases diagnostic accuracy -In some circumstances possible to screen for several viruses at once	Same as for ELISAs
Lateral flow based immunoassays [12] -Similar principles to ELISA, can be used as a direct or indirect test -Use low cost paper or membrane based platforms	-Rapid, low cost, low volume samples needed, easy to perform/interpret -Suitable for point of care -One step assay, little to no sample preparation, simple/no laboratory infrastructure needed -Can be used in array format	-Limited sensitivity, confirmation using other test methods often needed -Issues with samples can affect precision and analysis time of tests e.g. time for sample to flow through the assay unit -Physical manufacturing challenges can affect test performance

Molecular tests

Molecular tests directly detect viral nucleic acids. Most tests that are used as routine diagnostics require knowledge of the genetic sequence of a virus to facilitate their design, after which they can be used as highly specific and often very sensitive tests. When the genetic sequence of the virus is completely unknown, or it is very unclear which virus is causing a patient's symptoms, unbiased genetic sequencing may be used as a diagnostic tool. In addition, sequencing as a research tool is often critical to the design of primers used in non-sequencing molecular tests. However, for tests to be most effectively used frequent monitoring of the genetic sequence of viruses is required, especially in RNA viruses with high mutation rates, to ensure that molecular tests continue to work as expected. Any type of body fluid or tissue yielding sufficient nucleic acid is suitable for molecular testing.

RT-PCR: Reverse transcription PCR (RT-PCR) based tests are the most commonly used molecular diagnostics for RNA viruses. Viral RNA is first reverse transcribed into complementary DNA (cDNA). Primers (short sequences of DNA) are designed to bind to target sequences in the cDNA specific to the virus in question. The target sequence(s) are then amplified many times until the DNA reaches a detectable level, allowing the presence of the virus to be confirmed and the viral load to be quantified. There are several variations of PCR methods, with real time PCR being an increasingly used approach where fluorescent dyes or primers are added to the reaction to allow the amount of PCR product to be quantified in real time.

Advantages:

- High specificity and sensitivity, once technique is established in laboratory it is easy to use and can generate results quickly.
- The process can be fully automated to reduce manual processing and provide results in 90 min or less.
- As a quantitative method it can be used to evaluate viral load, which may be associated with worse disease symptoms in some patients.

Limitations:

- Requires prior knowledge of the genetic sequence in order to allow the appropriate primers to be designed.
- The specificity of primers can be affected if the virus mutates in a primer binding site, which can lead to false negative results. Secondary confirmation of results may be required.
- Primer design may have to be frequently reviewed and new primers designed. This requires both time and expertise in primer design.
- Where multiple strains of a virus exist it may be difficult to know which primers to use, and availability of the correct primers may be lacking. In some cases multiple sets of primers may be required, increasing cost.
- Laboratory and infrastructure requirements include an electrical power supply, the need for specialised equipment and technical expertise, and storage for temperature sensitive reagents. This can complicate deployment in low resource settings and remote areas, and makes the technique less suitable for mobile laboratory settings.
- Adequate biosafety and containment measures are required.
- Contamination can produce false-positive results in PCR-based assays or the presence of PCR inhibitors can also produce false-negative results.
- Varying sensitivities of assays particularly in early disease.

In addition to RT-PCR, a number of molecular diagnostic techniques are in development or in the research phase (table 3).

Table 3: Molecular diagnostic techniques in development or in the research phase.

Test	Advantages	Limitations
Loop-mediated isothermal amplification of DNA (LAMP) [13] -Like PCR, requires primers to target specific regions of sequence -Test reactions carried out at a constant temperature, unlike PCR	-Highly specific and sensitive -Simple equipment, cheap to run and suitable for use in field, lower resource settings -Rapid, results in ~2h -Can be a quantitative test	-Primer design requires pathogen sequence knowledge -Mutation rates can impact primer specificity -Regular primer review needed -Suitable primers might not be available -False positives a concern
CRISPR based diagnostic assays -SHERLOCK (Specific High-sensitivity Enzymatic Reporter unLOCKing) [14] -Utilises CRISPR to target sequences of interest -Uses lateral flow technology	-Designed to be accurate, quick, low cost, low equipment, point of care tests -Rapid design and production allows new tests to be developed quickly	-Test accuracy relies on accurate and up to date viral sequence information -More evidence needed for cost-effectiveness -Further evidence needed to support test effectiveness in clinical contexts
Programmable DNA nanoswitches [15] -Developed in Zika and SARS-CoV-2 viruses -700 bp length DNA (oligo) bound to target RNA of interest switches from linear to looped form, allowing detection by gel electrophoresis -Research use only	-Development of new oligos relatively rapid, as little as two days -Tests could be multiplexed -Current evidence suggests high specificity -Once developed, assays could be deployed in low resource settings -Can be stored at room temperature	-Test accuracy relies on accurate and up to date viral sequence information -Further evidence of effectiveness needed, validation required, not available for clinical use

3 LASSA FEVER

3.1 Lassa fever disease

Lassa fever is a viral haemorrhagic fever caused by the Lassa virus – *Lassa mammarenavirus* (LASV) – a member of the arenavirus *Arenaviridea* family. The Arenaviridae are a family of RNA viruses whose members are generally associated with rodent-transmitted diseases in humans. Arenavirus infections are relatively common in humans in some areas of the world and can cause severe illnesses.

Lassa fever virus

LASV is an enveloped, single stranded RNA virus with a bisegmented, ambisense genome. The viral genome consists of two segments (figure 2):

- Large (L), 7kb in size. This encodes the RNA-dependent RNA polymerase (L) and a small, zinc-binding (Z) protein.
- Small (S), 3.4kb in size. This encodes the viral glycoprotein (GP) precursor and the nucleoprotein (NP). The glycoprotein spike complex (GPC) drives host cell entry [16].

The terminal 19 nucleotides at the 3′ and 5′ ends of the RNA segments are complementary to each other and are highly conserved among all arenaviruses.

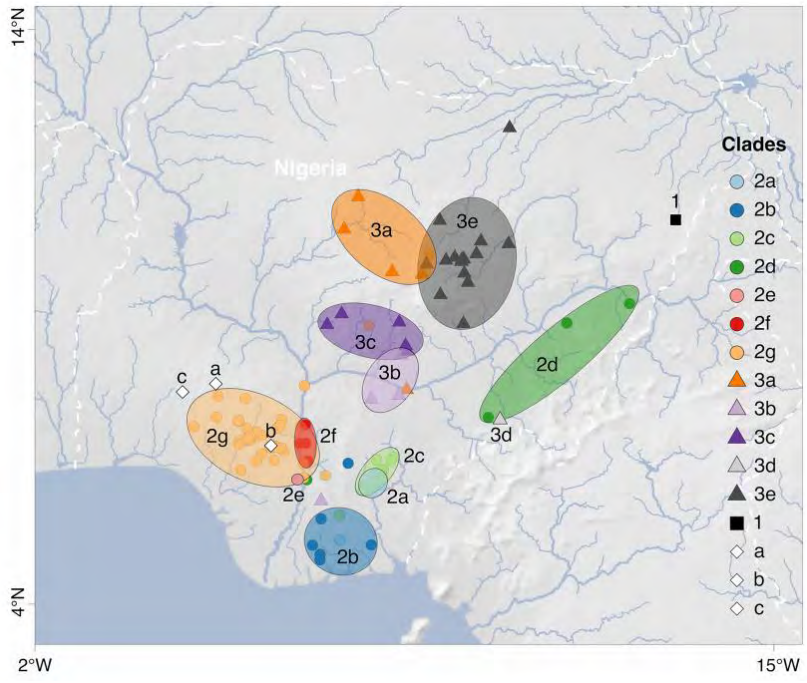
Figure 2: Schematic of the LASV genome, both the Large (L) segment and Small (S) segment. (adapted from [14])

Lassa viruses are classified into six lineages, I–VI, with new evidence indicating the possibility of a seventh, VII.

- Lineages I, II and III are commonly found in Nigeria [17], though lineage I has not been reported in nature for over 40 years
- Lineage IV, represented by the Josiah strain, occurs across West Africa and is transmitted in Liberia, Guinea, and Sierra Leone [18] and is the best studied
- Lineage V is found in Mali and Ivory Coast [19, 20]
- Lineage VI is represented by a cases from Nigeria [21]
- Lineage VII has been proposed and is represented by recently described sequences originating in Togo [22]

Geography plays a role in delineating Lassa virus lineages (figures 3 and 4). In Nigeria, lineages II and III are stably separated along the course of the Niger and Benue rivers (figure 3). Lineage II occurs south of the Benue and Niger Rivers and lineage III to the north. Lineage II has further evolved into lineage IIA mainly occurring south west of the Niger and lineage IIB is confined to south east of the river. The bottleneck that has confined lineages I–III in Nigeria, allowing lineage IV, to travel across West Africa is yet to be determined (figure 4) [23, 24].

Figure 3: Map of LASV lineages and clades in Nigeria demonstrating how the rivers play a role in separating the clades (From [25]).



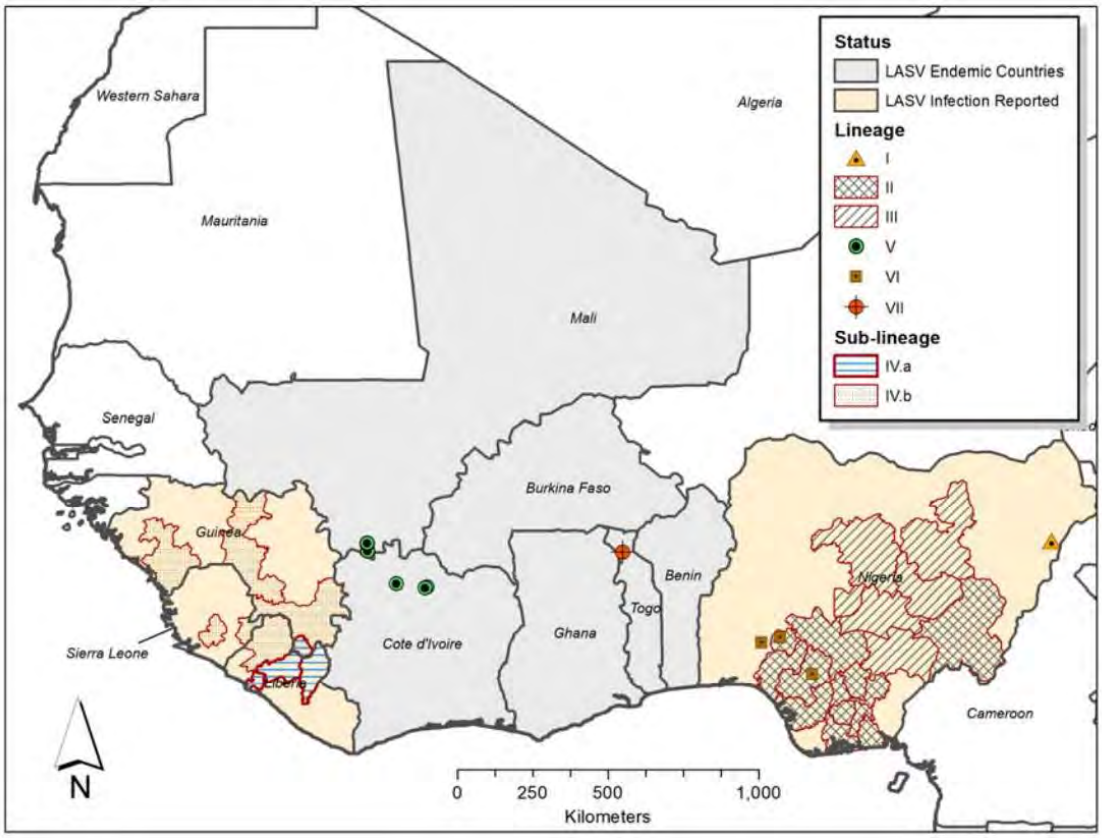
The nucleotide and amino acid divergence among known LASV sequences is up to 27% and 15%, respectively. Substitution rates for both segments of between ~0.8-1.2x10^{-3} substitutions per site per year have been reported [3, 18, 19, 26]. This genetic diversity has a significant impact on the ability to provide one standard molecular test that can be used across all regions where Lassa fever is found or suspected. Taking this heterogeneity into account will be critical for development of assays to detect and identify all LASV lineages – pan-LASV assays – and for development of effective vaccines that will provide protection across multiple lineages [27].

Epidemiology

Lassa fever is a notifiable disease under the WHO's International Health Regulations (IHR) [28]. In Africa, the Integrated Disease Surveillance and Response (IDSR) strategy is in place to help strengthen national surveillance and response from the community up to the national level [29, 30]. Lassa fever requires immediate reporting under this system. There is acknowledgment that these systems need to be strengthened [31, 32], and there have been incidences where the importance of cross-border cooperation and responses was demonstrated [33].

The virus was first described from a case in Lassa, Nigeria, in 1969. The majority of the confirmed cases that have been reported are from the Mano River Union region (Guinea, Liberia, and Sierra Leone) and Nigeria, but there have been a few cases reported, with the isolation of the virus or serological evidence of Lassa fever infection in Benin, Burkina Faso, Ghana, Côte d'Ivoire, Mali, and Togo (figure 4). Cases have also been identified in Germany, the Netherlands, Sweden, the USA, the UK and Japan, largely imported after travel in West Africa [34, 35].

Figure 4: Geographical distribution of Lassa fever lineages in West Africa. The shaded areas and symbols show the distribution of LASV lineages and sub-lineages as observed in West Africa. Shaded areas show administrative regions or states where LASV lineages or sub-lineages have been reported. Symbols show locations where lineages have been observed. From [36].



Epidemiological data are scarce, and the real incidence of Lassa fever is not well known [24]. An estimated 100,000 to 300,000 infections of Lassa fever occur annually, with approximately 5,000 deaths [37]. An estimated 58-60 million people are at risk of the disease [16]. Much of this surveillance information is in part based on serological prevalence studies of the host rodent species, for which seropositivity may have resulted from either LASV infection or possibly infections due to related arenaviruses [27]. In some areas of Sierra Leone and Liberia, it is known that 10-16% of people admitted to hospitals annually have Lassa fever, demonstrating the serious impact the disease has on the region [38]. Cases can occur throughout the year but peak incidence usually occurs in the dry season and peaks around March, when the dry season transitions to the wet season.

The high degree of seroprevalence of LASV-specific antibodies in the general population residing in endemic regions, indicates that most infections are mild (or asymptomatic) and do not result in hospitalisation. In a mild case of LASV infection, symptoms subside and recovery typically commences 8–10 days after disease onset. The overall case fatality ratio (CFR) varies depending on the context. It can be as low as 1% in the general population, approximately 20% amongst hospitalised patients and increases to greater than 50% in high risk groups, including pregnant women and infants. Pregnant women are almost threefold more likely to die as a result of LASV infection than their non-pregnant counterparts [39]. CFR also increases with age [40].

Transmission

LASV is a zoonotic pathogen found in the African rat, also called the natal multimammate rat (*Mastomys natalensis*) which acts as a disease reservoir. It was previously thought to be the only host but recent evidence indicates that LASV is circulating in other rodent species, such as mice [21]. Rodent hosts are chronically infected and asymptomatic. Human infection is incidental to the natural cycle of the virus and occurs when an individual comes into contact with rodent urine or faeces directly – infection can be via broken skin – or with contaminated supplies such as food, or aerosol transmission via contaminated dust.

Person-to-person transmission is associated with direct contact with blood or bodily fluids, containing virus particles, of infected individuals. Contact with objects contaminated with these fluids, such as medical equipment, is also associated with transmission. However, human-to-human transmission only contributes to approximately 19% of all reported cases and is usually observed during nosocomial outbreaks [41]. Hence, the overall transmission pattern is primarily environmental exposure to Lassa fever rather than sustained human-to-human transmission chains. Control measures therefore tend to focus on rodent control and hygiene measures around preventing rodents from entering homes and in particular securely storing food to prevent contamination.

Clinical Symptoms

Lassa fever symptoms develop over an incubation period of 7–10 days, with up to 21 days reported [23, 34, 42]. About 80% of people infected with LASV are asymptomatic or have mild symptoms – these cases will recover but others progress towards more severe illness. The onset of symptoms is gradual, starting with a mild fever, weakness and general malaise. In the remaining 20% of patients, symptoms of more serious disease include fever, weakness, headaches, chest pain, vomiting, and more rarely multiple organ dysfunctions with or without haemorrhage, which is seen in about 20% of these patients [43].

Various degrees of deafness have been shown to occur in 25% of survivors from mild or severe cases. Hearing returns after 1-3 months in only 50% of these patients while the others have permanent hearing loss. Lassa fever is severe during pregnancy resulting in most infected pregnant women losing their foetus.

Early supportive treatment, including correction of fluid and electrolyte imbalances, and symptomatic treatment is imperative and improves survival. The anti-viral ribavirin, if begun within the first six days of symptom onset or exposure to a Lassa fever case, may reduce mortality.

The WHO has a guide on the clinical management of patients with viral haemorrhagic fevers which includes guidelines on the use of PPE, infection management protocols and use of diagnostics [44], as well as technical guidance on many aspects of Lassa fever surveillance and case management [45]. Nigeria also has national guidelines for Lassa fever case management [43].

3.2 Research and development needs for disease control

There have been recent research advances in terms of understanding the structural biology, immunology and genomics of LASV, as well as progress towards an immune-therapy drug and new initiatives to develop a Lassa fever vaccine. The WHO has listed LASV as a high priority pathogen for the development of treatments and prophylactics [46]; key knowledge gaps identified in the WHO Lassa fever Research and Development roadmap include [47]:

- a better understanding of the natural history of Lassa Virus;
- more research on the pathogenesis of infections;
- the determinants of infection and disease severity including host vs. virus determinants, such as virus lineage;

- more accurate epidemiological information on incidence and seroprevalence, due to the limited epidemiological surveillance it is likely that the current incidence and spread of the virus is significantly underestimated.

Lassa fever drugs and therapeutics

A limited number of drugs have been identified and investigated for antiviral efficacy against LASV. Currently, the drug ribavirin is used off-label and is widely accepted by many countries for the treatment of Lassa fever [27] and has been recommended by the WHO in its list of essential medicines for treatment of certain haemorrhagic fevers, including Lassa fever.

Early studies carried out on ribavirin's effectiveness reported that it reduced the risk of mortality to below 5% if administered within the first six days of illness, with benefits greatly decreasing if started later [42]. Oral use is also sometimes recommended as a post exposure prophylaxis for people at high risk of secondary infection [48] although the WHO says there is no evidence to support this [49]. While ribavirin studies have demonstrated a noticeable decrease in mortality for severe cases, fatality rates remain high even when compared to untreated patients, which underscores the need for better therapeutics [34].

Because of the limited number of clinical trials examining the use of ribavirin, and to address concerns around its use, additional data from the original prospective trial by McCormick et al. were released in March 2019 and a reanalysis was performed by independent experts. However, these were inconclusive [50, 51]. More recent investigations studying intravenous administration of ribavirin are currently undergoing clinical trials [52].

Another broad spectrum inhibitor against RNA viruses that has decreased levels of LASV viremia and increased survival in animal models is favipiravir [53]. Animal models have also been used to demonstrate the potential usefulness of small molecules such as benzimidazole-related scaffolds, including ST-193 and its analogue LHF-535 as well as small interfering RNA particles (siRNA) and small molecules that inhibit viral fusion, entry or replication [54]. There is some evidence that treatment with convalescent plasma or with monoclonal antibodies can prevent Lassa fever, however the effectiveness of this approach in humans is currently uncertain [16, 54] as evidence is currently only available from animal models [55, 56].

As such, there is an ongoing need for trials to investigate further therapeutic interventions for Lassa fever. In 2018 the WHO held a workshop on "Efficacy trials of Lassa Therapeutics: endpoints, trial design, site selection"; the next steps identified were [57]:

- Developing an annotated generic protocol for Lassa therapeutic efficacy trials, based on preliminary design consensus
- Finalising a Target Product Profile (TPP) for Lassa therapeutics in consultation with partners (dated 2017)
- Continue efforts towards standardisation and harmonisation of core clinical variables and clinical case management
- Support research capacity and plan for clinical trials

Nigeria has also developed a Lassa Fever National Research Plan [58] which focuses on facilitating and coordinating research efforts within the country.

Vaccines

No vaccine is currently available for Lassa fever but there are a number of vaccine candidates and there is ongoing research [16, 59, 60]. There is support for vaccine development from the WHO and the Coalition for Epidemic Preparedness Innovations (CEPI) and it is anticipated that within the next few years numerous candidates may reach clinical trial stage. Currently there are six active vaccine candidates in the CEPI portfolio: four recombinant viral vector vaccines and two DNA vaccines [24]. Of these, two are in

phase I clinical trials – the Themis measles vector vaccine, and the Inovia DNA vaccine (NCT02483260, NCT00992693).

Glycoprotein complex (GPC) has been the target of LASV antiviral and vaccine research since it is the only antigen displayed on the viral surface and structural analyses of GPC indicate that it is a primary target for neutralising antibody binding. However, the role of nucleoprotein (NP)-specific T-cells in controlling acute infection and mediating immunity, provides some justification for the inclusion of NP in vaccine development [16, 59].

Countries affected by Lassa fever have research sites that are ready to host successful vaccine clinical trials [59]. Vaccine development bottlenecks are thought to include the high cost of biocontainment requirements, the absence of established measures for protection – i.e. signs that a person is immune to infection – and uncertainty around animal models being predictive of vaccine efficacy in humans [16]. Further research is required in all of these areas.

3.3 Current landscape of diagnostics

Diagnosis of Lassa fever

Lassa fever occurs in all age groups, in both sexes and is associated with a broad spectrum of clinical symptoms, which can be non-specific particularly in the early phase of the disease. As such, diagnosis can be challenging based on symptoms alone [61, 62] partly due to the similar symptoms shared with other febrile illnesses prevalent in the geographic regions affected by Lassa fever, such as other viral haemorrhagic fevers, malaria, shigellosis, and typhoid [27]. This makes laboratory diagnosis necessary for proper clinical management, isolation of patients, treatment selection and contact tracing efforts [40, 63]. As a nonspecific febrile illness with a high proportion of asymptomatic cases, infection is often underreported or misdiagnosed, hindering the accurate determination of prevalence. Diagnostic test types differ in complexity, infrastructure requirements and appropriateness for a rapid response.

Current laboratory diagnostics

Detection of viral proteins or LASV-specific IgM or IgG antibodies by enzyme-linked immunosorbent assay (ELISA) or nucleic acids by reverse transcriptase-PCR (RT-PCR) are the most common methods used today. Viral isolation using culture is the gold standard, but is impractical in endemic areas due to the requirement for BSL-4 [23].

The host response to LASV infection is still not completely understood and, combined with the high genetic diversity of the virus, diagnosis of Lassa fever by molecular and immunological methods is currently challenging [23]. There have also been research efforts to develop RT-PCRs capable of detecting the presence of multiple VHFs [64]. Although RT-PCR is more commonly used for viral identification, this may not be adequate as the sole detection method given the high genetic diversity among LASV isolates. Similarly the immune response varies at different stages of the infection and as such may also not be completely reliable all the time. As such a combination of RT-PCR, ELISA and/or tissue culture assays are recommended to improve the reliability of the diagnostic for highly suspected individuals [61, 65]. A number of recently published reviews describe available diagnostics in detail [23, 34, 66-69].

Laboratory diagnostics of LASV infection involve detection of [34]:

- LASV from culture
- LASV RNA using molecular diagnostics, such as RT-PCR
- IgG/IgM antibodies: LASV-specific IgG/IgM antibody response or
- Viral antigens: LASV antigens shed during replication

Virus culture platforms

The main advantage of this method is that it is independent of genetic variation among lineages and can be used to determine viral load as well as allowing further characterisation and study of the virus [68]. Lassa virus can be isolated from an array of patient samples including blood, urine, pleural fluid, throat swab, and histopathological materials (collected during autopsy) including liver, kidney, spleen and heart. Virus isolated from cell culture may be identified by immunoassays, such as an immunofluorescence antibody test (IFA), western blotting, RT-PCR, or sequencing [27].

The main limitation of this method is the requirement of BSL-4 biocontainment and it can take a week or more for effective viral growth in susceptible Vero E6 cells [68]. This means it can be an impractical tool to use in endemic areas and when a rapid result is required [23].

Molecular Diagnostics

Reverse transcription-PCR (RT-PCR) is a molecular tool used for the detection of LASV RNA in blood, tissue, and secretions during the early (acute) stages of infection. While RT-PCR assays to detect LASV RNA can be used to diagnose Lassa fever, the capacity to provide such testing as a routine method in rural healthcare settings in endemic regions of West Africa is limited at the present time [63]. There have been some examples of successful use in mobile laboratory settings [70-73].

The majority of existing assays target the glycoprotein complex (GPC) precursor and the nucleoprotein (NP) on the S segment, but with the availability of more sequence information, assays that target the RNA polymerase (L) and a small, zinc-binding (Z) protein on the L segment have been developed [23].

Standard RT-PCR assays using only amplifying primers are commonly used due to their simplicity. Probe-based real-time RT-PCR assays that use amplifying primers and a specific-labelled probe on a real-time PCR instrument can be deployed to improve specificity. While more technologically complex, this methodology and instrumentation can control temperatures, collect data, analyse results, and therefore has more capacity and faster through-put than a gel-based PCR [23].

There are commercially-available assays from three different vendors (Altona Diagnostics, LifeRiver, LIPSGene/VL diagnostics) that are CE-marked as *in vitro* diagnostics [66]. Further assays are being developed, such as the Aldatus Biosciences PANDAA technology (table 4).

Table 4: Molecular diagnostic tests for LASV (adapted from [34]). Commercial and regulated assays for LASV. Lab developed and in house assays are not included. RUO - Research use only, CE-IVD – CE marked *in vitro* diagnostic

Developer	System	Regulation
Commercial - single assays		
Alton Diagnostics GmbH (Germany)	RealStar® Lassa virus RT-PCT kit 1.0	CE-IVD
Alton Diagnostics GmbH (Germany)	RealStar® Lassa virus RT-PCT kit 2.0	RUO
gensig (UK) Primerdesign Ltd	Primerdesign™ genesig® Kit for Lassa virus Josiah	RUO
LifeRiver (China) Shanghai ZJ Bio-Tech Co. Ltd	Lassa virus real time RT-PCR kit	CE-IVD
LIPSGene/VL diagnostics (Germany)	Lassa virus RT-PCR kit	CE-IVD
Aldatu Biosciences Ltd (USA)	In development PANDAA LASV - pan-lineage detection	-
Commercial - mulitplex assays		
GeneArraytion Biothreat, Febrile kit panels (Taqman or Luminex xMAP)	MultiFLEX: Febrile Agent Panel Mega Febrile panel BioThreat Panel	RUO

RT-PCR tests have been known to produce false-negative results; the amount of the virus in the sample needs to be at a sufficient level for detection, and with the high specificity of primers there may be mismatches due the high degree of genetic diversity with LASV, making RT-PCR an imperfect diagnostic. This highlights a challenge in evaluating the performance of Lassa fever diagnostics [61, 64, 74, 75]. Therefore, due to the genetic diversity of the virus at the nucleotide level, it is recommended that primer sets are designed by clinical laboratories to reflect geographical isolates circulating regionally [17]. There are attempts to have improved primers that target highly conserved regions of the S- and L-segments and that would work on multiple virus lineages [76, 77], however further evaluation of these are needed.

The continuous need to evaluate the appropriateness of existing assays was recently demonstrated by a sequencing study of LASV genomes from Liberian patients. The authors analysed the current molecular diagnostic assays being used in Liberia and found that numerous nucleotide mismatches were present in recently circulating LASV genomes that could impact amplification [18]. This potential of a false-negative PCR result is one of the reasons why it is generally recommended to have a secondary test such as ELISA and/or tissue culture carried out to confirm diagnosis [27, 61].

Molecular Diagnostics in development

A number of novel molecular diagnostics are under development: oligonucleotide array hybridisation coupling, nested-PCR techniques, LAMP, CRISPR, resequencing and transcriptional profiling microarray chips are being examined for bolstering LASV detection methods and developing prognostic indicators [14, 27, 77-82]. Many of these 'in development' approaches are research tools and will still be dependent on robust matching of target sequences to maintain sensitivity and specificity.

Immunoassays

The use of immunoassays based on broadly-reactive monoclonal or polyclonal antibodies can also potentially overcome sensitivity issues related to genetic diversity of RNA viruses, such as LASV. Immunoassays to detect anti-LASV IgM, IgG antibodies, and LASV antigen have been developed and been converted to rapid diagnostic tests that are pan-LASV (can detect most or all lineages), cheap and transportable.

Immunofluorescence assays (IFA) have some use in terms of detecting LASV antibodies, however these are being replaced by ELISA platforms. These are robust enough for use in the field and several research grade ELISA kits exist. Commercially available ELISA kits have been developed for detection of pan-LASV NP (from lineages II, III, and IV) antigen or IgM/IgG. These include the ReLASV® Pan-Lassa IgG/IgM and Pan-Lassa Antigen ELISA test kit from Zalgen Labs [66] (table 5). Performance characteristics, particularly sensitivity and specificity, have not been established for either kit, and field/clinical evaluations are underway [27].

Rapid antigen detection assays

Rapid antigen detection tests are also under development in endemic regions of Africa, such as 'dip-stick' lateral flow immunoassays that work from clinical samples and test for the presence of LASV NP from lineages II, III, and IV. These are used during acute stage LASV infection. One platform (specific for lineage IV only) has been validated and has been CE-marked as an *in vitro* diagnostic (ReLASV® Antigen Rapid Test, Zalgen Labs) and is the only such test that has undergone field evaluations [61]. It has been shown that when combined with an antigen ELISA test sensitivity increases to 90% and specificity to 100% [23]. Commercially available tests are listed in table 5.

Table 5: Serology Tests for LASV (adapted from [34]); commercial and regulated assays for LASV. Laboratory developed and in-house assays are not included.

Developer	System	Regulation
Commercial ELISA and IFT/IFA		
Zalgen Labs /Corgenix (USA)	ReLASV® Pan-Lassa Antigen Elisa test kit	RUO
Zalgen Labs /Corgenix (USA)	ReLASV® Pan-Lassa IgG/IgM ELISA test kit (NP/GP)	RUO
Creative Diagnostics (USA)	Anti-LASV GP polyclocal antibody	RUO
Bernhard Nocht Insitute (Germany)	in development	-
Abbexa (UK)	in development	-
Commercial Rapid Diagnostic Tests		
Zalgen Labs /Corgenix (USA)	ReLASV® Antigen Rapid test	RUO
Zalgen Labs /Corgenix (USA)	ReLASV® Pan-Lassa Antigen rapid test	RUO

Future considerations

A specific knowledge gap identified for Lassa fever diagnostics is that additional field validation data are needed to assess performance characteristics of Lassa fever diagnostic assays against the multiple lineages of LASV that can be found across West Africa.

To guide the development of molecular diagnostics for Lassa fever it is important to take into account where individuals access health care, who will carry out diagnostic testing and report the results, and the response for confirmed Lassa fever cases. Development of more sensitive methods for detection of LASV genomic RNA should also proceed with an emphasis on platforms that are low cost, are quick, produce results that are easy to interpret, and that can be deployed and maintained in difficult environments [61].

Development of a pan-LASV RDT and ELISAs with polyclonal antibodies results in sensitive and specific diagnostic platforms that are more tolerant to genetic variation than PCR assays. However, the expected overall higher sensitivity of PCR-based assays allows the virus to be detected when present at lower levels in the patient, particularly in early (asymptomatic) and post-acute infection stages of Lassa fever. Therefore, a combination of both platforms should be part of a comprehensive diagnostic algorithm in the testing of

suspected Lassa fever patients. Researchers have recommended evaluation of recently developed pan-LASV RT-PCR assays and immunological tests and the development of a more comprehensive testing and/or diagnostic algorithms [61].

3.4 Current use of NGS

Sequencing technologies

A variety of sequencing techniques are used to sequence LASV genomes and a number of methodologies have been adapted to enable sequencing. The biggest challenge to manage is the genetic diversity of LASV lineages, for example careful consideration around which reference sequence to use during bioinformatics analysis is a particular challenge.

The use of sequencing has increased as cost has reduced and sequencing methodologies have improved in terms of their reliability and standardisation. An improved understanding of LASV genetic diversity is also supporting appropriate redesign of probes used for sequencing.

A wide variety of sequencing systems and techniques are used to analyse LASV – there is no one method or technology that dominates over any other. Rather, the sequencing system and method used reflects the availability of technology at the site of the laboratory that carried out the sequencing and that was responding to an outbreak or case presentation.

Sequencing systems used include:

- Illumina systems such as MiSeq, MiniSeq, HiSeq or NovaSeq
- ThermoFisher 3130xl sequencing system or Personalised Genome Machine (PGM)
- Oxford Nanopore Technologies MinION sequencer
- ThermoFisher resequencing pathogen microarray (RPM) which enables detection of 84 pathogens [79]
- Sanger sequencing is frequently used as the method of choice to ensure full genome coverage. The instrumentation used is not often mentioned, however where described the Applied Biosystems 377 fluorescent automated DNA sequencer is used.

Protocols are available in publications, which generally follow manufacturer's instructions. General sequencing protocols for instruments are usually established in most laboratories and adapted as needed based on the pathogen. Organisations such as US CDC or ARTIC also provide sequencing protocols [83, 84].

Sample selection. Samples selected for sequencing tend to be isolated from patients that have laboratory confirmed cases – generally a positive RT-PCR diagnostic test. Sequencing is then carried out on either: the RT-PCR product, culture of a sample, or directly from clinical samples. Clinical samples are primarily serum or plasma and that have been either inactivated or cultured. Other sample types used for sequencing have included breast milk and cerebrospinal fluid [26]. Direct sequencing from these samples is restricted by the viremia content and an amplification step is usually required to improve the success rate of sequencing. When sequencing is not successful reasons can include low viral titres, sample degradation due to lack of sufficient cold chain for storage, or technical sample-handling issues [85]. If NGS is unsuccessful or partial, some laboratories have resorted to Sanger sequencing from clinical samples to obtain a result or complete the sequence.

Lassa virus sequencing studies

The first LASV sequence was produced in the late eighties and is known as the Josiah strain [86, 87]. Currently over 2000 LASV sequences have been uploaded onto genomic sequence databases [88]; the L and S segments are uploaded separately. More than 1000 are samples isolated from humans. Of these, around 500 sequences include the L and S segments from one sample, and can be considered complete genomes. The majority of

sequences are from Nigeria (mainly 2014 and 2018 outbreak), Sierra Leone (mainly 2014 outbreak), and Liberia. The largest studies identified are listed in table 6. Twenty sequences from Benin were produced in 2020. Guinea, Togo, Mali and Ivory Coast also have limited (fewer than 15) sequences represented in databases.

Table 6: The largest (in terms of numbers of samples) LASV sequencing human studies identified.

Reference	Number of sequences	Methodology	Comment
Andersen 2015 [3]	200 from Western Africa Nigeria, Sierra Leonne, Guinea and Liberia: (183 from human, 11 from reservoirs)	Illumina HiSeq2000, HiSeq2500, or the MiSeq (done at the Broad Institute, USA)	On completion this was the largest study examining LASV sequences and provided valuable insight into the viruses' genetic complexity
Siddle 2018 [85]	220 from Nigeria	Illumina MiSeq (Redeemer's University Nigeria), HiSeq 2500 and MiSeq (Broad Institute, USA)	Done at ACEGID, Nigeria, during the 2017-2018 outbreak and contributed to public health understanding of transmission
Ehichioya 2019 [25]	216 S segment and 157 L segment sequences from Nigeria	Illumina MiSeq in samples that successfully cultured and Sanger in those where culture was not successful. Sequencing done in Germany	Phylogenetic analysis of the L and S segments

Sequences published from previously unrepresented areas, have provided valuable insights into LASV diversity. These include sequences of three cases that originated in Togo (repatriated cases that were sequenced in Germany and the USA) [22], and 20 recently confirmed cases from Benin [24]. These are areas that were not considered to be endemic regions but may need closer monitoring and could represent a growing threat of future outbreaks.

The cost of acquiring and maintaining sequencing platforms has become cheaper and more sequencing has become available in Western Africa. Key sequencing studies of LASV that have been carried out at African facilities are:

- Liberia Institute for Biomedical Research [18]; 23 Sequences of West African countries: Liberian, Guinea, Nigeria and Sierra Leone using an Illumina MiSeq
- Irrua Specialist Teaching Hospital (ISTH), Edo State, Nigeria
    - Sanger sequencing resulting in 36 partial sequences for molecular diagnostics development [65]
    - Metagenomics using the Nanopore MinION and validated using Illumina MiSeq, 120 sequences [26] (more detail in case study 2)
- African Center of Excellence for Genomics of Infectious Diseases (ACEGID), Redeemer's University, 220 samples sequenced that originated from ISTH sequenced using Illumina systems [85]

Applications of sequencing

Sequencing has been used for a number of applications in LASV. In addition to the case studies highlighted below, further examples of research in the relevant application areas outlined in chapter two are listed in the table in appendix 8.1.

There is much still to learn about LASV, particularly around virus diversity and disease epidemiology. A number of recent studies have provided information on transmission dynamics that have supported response efforts. For example, two key studies determined that a cluster of disease cases were not nosocomial infections, therefore efforts were directed to rodent control and public health messages to secure food storage areas and exclude rodent contamination [26, 85].

Sequencing projects that aimed to better understand the phylogeny of LASV have contributed to information of viral population structure, viral prevalence, and how the virus is evolving [3, 25]. Information on prevalence can also be obtained from sequencing the virus in the reservoir rat population [20, 89, 90] which further helps to understand disease spread and origins.

In terms of diagnostics, the complete and partial sequencing of samples to determine the best PCR primers to be used in a particular region, has supported the development of the most suitable diagnostics for those regions [65, 78]. Sequencing is also being used to determine which infection is causing disease in patients with febrile disease symptoms, for example through the use of metagenomics [26, 91]. This is then leading to the development of diagnostics to look for multiple infections at a time.

Sequencing has been used to assist with LASV surveillance in a number of ways:

- To detect whether new viral lineages are emerging
- To understand the molecular epidemiology of the virus including identification of transmission chains
- Identifying lineages present in different regions to assist development of diagnostic tests, namely RT-PCR
- To help identify cases of Lassa fever that could lead to an outbreak

Below we outline two case studies to demonstrate use of sequencing in these areas. Whilst these case studies show how sequencing may be used in Lassa fever surveillance, there is not as yet a formalised surveillance system is in place that routinely uses or has recommendations on using LASV sequencing to inform public health measures. Globally, there are moves towards integrating WGS for public health surveillance [92-95].

These examples (and in Appendix 8.1) highlight the variety of sequencing technologies, methods and workflows used, full details of which are available in the papers' methods. All studies reported here were published in journals and/or on publicly available websites such as virological.org, and have deposited sequence data into public databases.

Case study 1: Identifying disease origins

Identifying the origins of an outbreak can assist in disease management due to understanding of where the source is, particularly if originating in an area where it is not usually expected. Once identified, diagnosis, contact tracing and a change in clinical management is possible and can have positive public health and clinical impacts.

Case Study: New Lineage of Lassa Virus, Togo, 2016 [22]

Aim of study: To sequence the Lassa virus from a cluster of imported infections to confirm epidemiological data that the virus had originated and circulated in Togo, a country where the virus was not considered endemic.

Study details: Lassa virus sequencing was performed on samples from three patients who had been diagnosed with Lassa fever and received treatment outside of Africa. The index patient had been a healthcare worker in Togo who was diagnosed with Lassa fever post mortem after receiving treatment in Cologne, Germany. The first secondary case patient had treated the index patient in Togo and was treated in Atlanta, Georgia, USA. The second secondary case patient was native to Germany and contracted the virus after taking part in funerary practices for the index case patient who had died of the disease. Sequencing was performed at a different location for each patient, with each location using different protocols. Phylogenetic analysis was then performed using the sequence data.

Sequencing method: Total RNA sequencing was performed at all three locations using Illumina technology to obtain the majority of the viral genome, with Sanger sequencing used to fill in any gaps when needed. Sequencing was performed either direct from clinical samples (serum or blood), on viral culture, or both. Two locations used the Nextera XT kit (Illumina) for library preparation and one location used the TruSeq total RNA kit (Illumina). In each location NGS was performed using an Illumina MiSeq. One location only sequenced the S segment, whereas the other locations sequenced both the S & L segments.

Advantages of methods chosen:

- All methods were able to capture high coverage of the L and/or S segments, with any gaps filled by use of Sanger sequencing. This provided more information for phylogenetic analysis.
- The use of both viral culture meant that maximum coverage could be obtained whilst any variations detected could be confirmed using the direct from sample sequencing.
- The low scale nature of this study meant Sanger sequencing was a viable technique for generating the material needed.

Information provided: Phylogenetic analysis revealed that the sequences of the three viruses were almost identical, confirming the transmission chains suggested by the epidemiological data. In addition the analysis suggested that the strain in Togo was a new lineage related to lineage II or lineages I/VI, which are all circulating in Nigeria.

How it informed public health: Whilst this study was performed retrospectively, so was not used to inform public health in real-time, it demonstrates the importance of providing genome sequencing to understand the evolution and spread of LASV. These results should inform future surveillance strategies for Lassa fever in Togo, as well as contribute sequence data to enable studies of how LASV emerges and spreads across western Africa which in turn can inform disease prevention and control strategies, including vaccine development.

Source: [22] Whitmer, S. L. M et al. New Lineage of Lassa Virus, Togo, 2016. Emerging infectious diseases. 2018. 24(3): 599-602.

Additional information associated with the study: Contact tracing was done in Togo as a result of these first known cases of Lassa fever in the country [96]. None of the 110 contacts identified and followed for 21 days developed symptoms.

Case study 2: Understanding transmission chains

Tracking person-person or environmental-person transmission, understanding the circumstances under which transmission is occurring (community/hospital/environmental) and local/national/international transmission routes.

Case study: Use of NGS to help understand chains of transmission in the Nigeria 2018 Lassa fever outbreak [26]

Aim of study: To understand the molecular epidemiology of the unfolding outbreak, understand chains of transmission and identify if a new strain was emerging

Study details: Of the 341 LASV cases reported at the Irrua Specialist Teaching Hospital (ISTH) between 1 January and 18 March 2018, 120 samples were selected for sequencing based on PCR analysis of PCR viral load and/or sample origin. Samples were sequenced at ISTH over 7 weeks, with real-time analysis of 36 genomes and subsequent confirmation using all 120 samples sequenced in-country. Phylogenetic analysis was used alongside epidemiological investigations to identify and investigate potential chains of transmission.

Sequencing method: Metagenomics - Random reverse-transcription and amplification by Sequence-Independent Single Primer Amplification (SISPA) using Oxford Nanopore Technology, with libraries sequenced on the MinION. Also performed Illumina sequencing using Nextera XT prepared libraries (Illumina) to validate the LASV results.

Advantages of methods chosen:

- Better able to capture the diversity of LASV strains than other types of sequencing methods that rely on specific primer design, such as amplicon sequencing.
- Small and portable so can be performed on-site, leading to faster results to inform outbreak management
- Able to identify multiple different viruses, allowing the detection of co-infections.

Information provided: Analysis revealed extensive diversity and phylogenetic intermingling with strains from previous years, suggesting independent zoonotic transmission events. This allayed concerns of an emergent strain or extensive human-to-human transmission.

How it informed public health: The researchers communicated their results promptly to the Nigeria CDC forming the basis of their report “Early Results of Lassa Virus Sequencing & Implications for Current Outbreak Response in Nigeria” released on 12 March 2018 [97]. The evidence that there was not extensive human-to-human transmission meant public health resources could focus on community engagement on rodent control, environmental sanitation, and safe food storage.

Source: [26] Kafetzopoulou, L. E., et al. Metagenomic sequencing at the epicenter of the Nigeria 2018 Lassa fever outbreak. Science. 2019. 363(6422): 74-77.

Additional information associated with the study: A perspective by Bhadelia [98] in the same issue of *Science* stated: “[the study] improved understanding of how the disease has been spreading in Nigeria and [has] led to informed and targeted disease-control strategies.”

3.5 Future needs of NGS

WHO roadmap

The WHO Lassa Fever Research and Development roadmap describes potential use of sequencing in the section covering diagnostics [47]. The existing knowledge gap calls for ongoing molecular characterisation (i.e. sequencing) of LASV isolates from both rodent reservoirs and humans to map the geographical distribution, and the monitoring of changes that will enable updating and refinement of diagnostic assays. In addition the roadmap highlights the need for a system to communicate sequencing results to key stakeholders. Key to facilitating this is the creation of networks of laboratories to carry out sequencing and analysis of strains isolated from rodents and humans, over time, in endemic and at-risk areas.

NGS for LASV disease management

There are a number of ways in which LASV sequencing could be used to inform disease management, either real-time or retrospectively:

- To help inform current disease management either prior to or during an outbreak – this can include contributing to diagnosis of new cases of Lassa fever at the start of an outbreak and contributing to outbreak response efforts.
- To generate genetic information to support the maintenance of current diagnostic tests and development of future tests.
- To help inform disease management in future, for example through monitoring which strains are circulating in the population, ensuring that the most appropriate diagnostics are used; once a vaccine is available, to ensure that the vaccine is still effective against the strains in circulation.

LASV is a genetically diverse pathogen and this diversity is not fully understood. A key first step to understanding and mapping this diversity is to sequence a larger number of LASV genomes from a variety of locations and upload the data to public databases (see section 2.3). Knowledge of the variety of lineages and where they are present can provide valuable insights into developing PCR-based tests that are location appropriate, or pan-LASV and are able to detect the variety of the lineages that may be present in a region.

The selection of sequencing method and technology will depend on the purpose – if it is to understand pathogen evolution and diversity, then high quality sequencing that is not error prone will be valuable in furthering knowledge in this area. However if it is to determine the presence of the pathogen in a patient with a fever for diagnostic purposes, time to get a result would be of greater importance. Metagenomic NGS may be appropriate when trying to establish the cause of unknown infection, whilst also ensuring that pathogen genetic diversity is fully captured. However in regions where known diseases are expected, a pathogen-specific amplification may be better suited, particularly if the method is more sensitive, cheaper and quicker to generate a result.

The genetic diversity of LASV poses challenges for the development of vaccines and therapeutics as well as for the implementation of diagnostic tests. The development of diagnostics, therapeutics and vaccines that are suitable for all circulating LASV lineages will require a richer genetic dataset from across the regions where Lassa fever is endemic, including virus circulating in the rodent reservoir. It will be important to incorporate novel, unique genetic signatures into assays that test LASV susceptibility to drug candidates and that these are also incorporated into vaccine clinical trials.

Resources and sequencing efforts will also be required for surveillance both in humans and in the rodent population, in particular to detect spill over events. This balance of needs and how sequencing efforts can be best utilised to meet them is yet to be determined [99].

LASV genetic diversity can also have an impact on genome assembly and interpretation of sequencing. Sequencing analysis still requires specialised bioinformatics and appropriate expertise as well as computational and IT resources. Some pipelines have been established but are continuously evolving and being adapted with updated analysis methodologies. No 'gold-standard' pipeline has been established, with many processes developed in-house and customised based on the needs of the laboratory and study being carried out. Until analysis pipelines become standardised, established scientists and bioinformaticians that are knowledgeable and adaptable to this ever changing environment are invaluable [100]. Appropriate quality management and assurance systems also need to be developed alongside these processes.

Reference based assembly of sequencing data can be very useful, but with LASV sequences, especially those from previously unexplored regions, the reference may not be suitable for the lineage circulating in that region. It may be necessary to use a few representative samples from that region to identify closely related lineages that can then be used for mapping subsequent samples, or developing RT-PCR assays. These initial samples can be assembled through *de novo* assembly, for which common bioinformatics tools also exist [100]. This process can be slower and more computationally intensive but could provide valuable information on novel sequences.

3.6 Lassa fever conclusions

For LASV, sequencing has the potential to improve our understanding of disease epidemiology. This includes metagenomic pan-pathogen sequencing carried out during outbreaks, both in humans and in the rodent reservoir. It is likely that sequencing will be more widely used in the future and as such has the potential to answer questions such as the true incidence of LASV, the transmission bottleneck in LASV infection, and the virus diversity in the reservoir host [23]. This information will also allow for the development of appropriate RT-PCR assays and other diagnostics such as SHERLOCK (table 3).

It is currently unknown whether the divergence of LASV lineages and clades is accompanied by changes in terms of human transmissibility and disease presentation or severity. This issue is not easily addressed for different reasons [101]:

1. Most LASV genomes have been obtained from patients who developed Lassa fever and mild/asymptomatic cases remain mostly unsampled.
2. Most clinical samples derive from patients infected with lineages II and IV, making comparison among all lineages more difficult. Moreover, because different lineages are transmitted in different geographic areas, disparity in time to diagnosis and treatment might influence disease outcome. Broader geographic sampling of LASV, including from the poorly characterised rodent reservoir, is needed to fully understand viral diversity and may be useful for the development of diagnostics, therapeutics, and vaccines [85] as well as predict and prevent future outbreaks.
3. Disease presentation most likely results from the interplay of viral and host factors (e.g. genetic background, age) [40].

Genomic surveys of viruses in reservoirs will most certainly advance our understanding of virus diversity and evolution. However, for mitigation of outbreaks there is a need for proactive, real-time surveillance of human populations. The use of metagenomic sequencing could replace the need for multiple individual pathogen assays [95] when trying to diagnose the cause of generic symptoms in patients with a febrile disease and is likely to become more commonly used in the future.

For LASV, the sequencing technology used is less important than ensuring accessibility to sequencing technology and the generation of good quality complete sequences, containing both L and S segments. Ensuring representation of samples from across the regions at risk of the disease will provide further insight into the epidemiology across these regions.

Generation of these sequences is an important first step in developing the knowledge needed for pan-LASV diagnostics and vaccines. Many of the issues related to achieving this and full utility of sequencing within public health and the control of the disease are relevant across many pathogens and are covered in chapter 6.

4 EBOLA

4.1 Ebola virus disease

Ebola virus disease (EVD) is a severe haemorrhagic fever caused by species in the genus *Ebolavirus*, which is a member of the *Filoviridae* family. Filoviruses are RNA viruses and the family contains four genera, of which two – *Ebolavirus* and *Marburgvirus* – cause viral haemorrhagic fevers in humans and non-human primates.

Ebolavirus is a single-stranded RNA virus with a non-segmented negative-sense genome that has inverse complementary 3' and 5' termini (figure 5). The genome is around 19kb in length and encodes seven functional proteins:

- Nucleoprotein (NP)
- Viral proteins (VP): VP24 (membrane-associated protein), VP30 and VP35 (both polymerase matrix proteins), VP40 (matrix protein)
- RNA polymerase(L)
- Glycoprotein (GP) [102].

Figure 5: Schematic of the ebolavirus genome (adapted from [102]).

Within the genus *Ebolavirus*, six species have been identified, four of which are known to cause disease in humans: *Zaire* (EBOV), *Bundibugyo* (BDBV), *Sudan* (SUDV) and *Taï Forest* (TAFV). Zaire ebolavirus was historically referred to as ZEBOV, however, the accepted abbreviation is now EBOV [103]. Therefore the term ebolavirus is used to discuss all species within the ebolavirus genus, or those that infect humans, depending on the context. EBOV is used to refer to Zaire ebolavirus.

Members of different species in the genus may be distinguished on the basis of glycoprotein gene sequence differences (≥30% amino acid difference), whole genome differences (≥30% nucleotide difference) [104]. EBOV is the most virulent of the viruses in humans and is responsible for the majority of outbreaks as well as the largest number of cases and deaths. The largest known ebolavirus outbreak occurred from 2013 to 2016 in West Africa and was caused by a novel EBOV variant, Makona.

Genetic sequencing has shown that the virus isolated from infected patients in 2014 is 97% similar to the virus that first emerged in 1976 [105]. The estimated rate of mutation for EBOV is ~1.2 x 10^{-3} nucleotide substitutions per site, per year [106], an evolutionary rate that overlaps with other RNA viruses. Greatest variation is seen in the intergenic regions and within specific areas of the *GP*, *NP* and *L* genes, the regions encoding internal viroid proteins are more conserved [107]. Different sections of the *GP* gene sequence have been found to evolve at differing rates, which suggests a selective pressure along the protein sequence that may benefit host immune evasion [107].

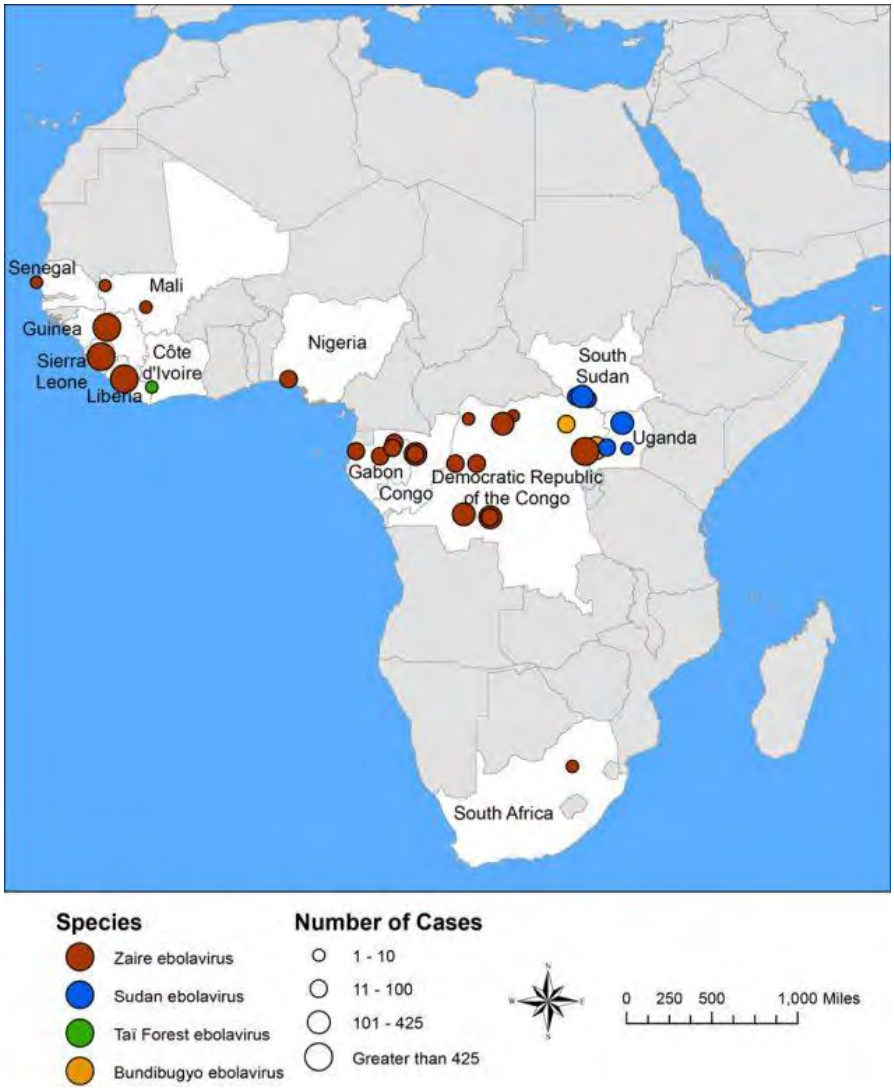
Prevalence and epidemiology

EVD was first identified in 1976 in two simultaneous outbreaks: one in Yambuku, Democratic Republic of Congo (DRC) and the other in Nzara, South Sudan. The disease is endemic in west and equatorial Africa including the DRC. The 2013-2016 West African outbreak was the largest in scale with high incidence and mortality, starting in Guinea then moving across land

borders to Sierra Leone and Liberia. There were 28,616 confirmed or suspected cases and 11,310 deaths with an overall mean case fatality of 62.9% among confirmed cases with clinical outcomes [108]. A further 36 cases and 15 deaths were recorded when the outbreak spread outside these three countries to Nigeria and Mali, some of which received treatment in other countries such as the UK, Germany and US. The most recent outbreak was announced in June 2020 in the DRC and was declared over on 18 November 2020 with a total of 119 confirmed cases and 55 deaths reported in Équateur Province [109].

The majority of ebolavirus outbreaks, including the 2013-2016 outbreak and subsequent outbreaks, are caused by the EBOV species. SUDV is the second most common species to cause outbreaks, with the largest of 425 cases occurring in 2000 in Uganda [110]. Only one outbreak of BDBV has been reported in 2007 in western Uganda, with 56 laboratory confirmed cases and an estimated case fatality rate of 25% [111]. One case of TAFV has been reported in a researcher who performed an autopsy on a dead chimpanzee. She subsequently survived with no haemorrhagic complications [112] (figure 6).

Figure 6: Map indicating location and species of ebolavirus outbreaks 1976-2014 (US Centers for Disease Control and Prevention [113]).



The prevalence of ebolavirus infection in those who have been in contact with EVD patients, including the occurrence of asymptomatic and mildly symptomatic cases, and the risk factors associated with infection, is not fully understood [114]. This is in part due to the results of different estimates of infection rates from studies carried out to address this question. The heterogeneity of the assays used and the populations in which the studies were carried out may also have contributed to the variation in results [115].

How long EVD survivors continue to be infectious, and whether ebolavirus is persistent (continues to replicate) rather than latent (replication has ceased but the virus remains replication competent) in convalescent patients that continue to test positive for infection is another question where research is ongoing. However, it is generally assumed that the infection is persistent rather than latent as studies have found that the virus continues to replicate in those who continue to carry the virus, which is consistent with a persistent infection [116]. The importance of sexual transmission in EVD survivors remains unknown, in particular, many of the studies around persistence and sexual transmission involve studying semen, which means there is relatively little known about viral persistence in women.

Transmission

Though the animal source and mode of transmission of ebolavirus to humans has not been confirmed, it is thought that it was first introduced into the human population by direct contact with wild animals such as fruit bats, porcupines and non-human primates, either by eating bush meat or food contaminated with wild animal faeces [117].

Various factors contribute to the introduction of EVD into the human population and the transmission between individuals including: cultural factors, population density, enhanced mobilisation of people between different regions and countries, health system resources and biological factors associated with the virus. The disease spreads through the human population via direct contact with infected or deceased individuals' blood, secretions, organs or other bodily fluids. Burial ceremonies involving direct contact with the deceased and inadequate infection control precautions for health-care workers are thought to play an important role in transmission amongst the human population. In addition, those who have recovered from EVD may remain infectious if the virus is still present in their blood or bodily fluids. For example, there is evidence from genomic investigations that survivors of EVD with persistent infection can spread the virus either via breast milk or by sexual transmission [1, 118-120].

Clinical symptoms of disease

The symptoms of EVD overlap closely with many other diseases including several VHFs, which can make prompt diagnosis challenging. An estimated 27.1% of EVD infections are asymptomatic [121]. The incubation period of the virus is from two to 21 days. Onset of EVD is typically sudden and includes: fever, fatigue, muscle pain, headache and sore throat. This can then be followed by: vomiting, diarrhoea, a rash, symptoms of impaired kidney and liver function, and sometimes internal and external bleeding. Laboratory analysis can include low white blood cell and platelet counts along with elevated liver enzymes. The severity of EVD is dependent on the host immune system response. Delayed antibody response and expression of pro-inflammatory cytokines results in a more severe disease [112].

During an outbreak, those at highest risk of contracting the disease include: healthcare workers due to their proximity to those infected; family members or others in close contact with infected people; and mourners who have direct contact with bodies during burial rituals as virus levels in the deceased remain high [122]. Pregnant women infected with ebolavirus may experience spontaneous miscarriages and also appear to be at higher risk of more severe disease and death [123], though evidence is limited. There is also some evidence to suggest that the very young (under four years old) and the elderly (over 60) are at higher risk of more severe disease and risk of death [124].

4.2 Research and development needs for disease control

In 2014 the WHO produced an Ebola response roadmap that aims to improve disease response and end transmission of the virus [125]. The WHO also issued two target product profiles for ebolavirus vaccine and diagnostic development [126]. Key needs for robust disease response to Ebola outbreaks identified in the WHO research and development roadmap include:

- The development of rapid, accurate, point-of-care diagnostics for Ebola/Marburg virus infection to inform treatment, outbreak detection, and clinical trials
- Safe and effective treatment and post-exposure prophylaxis to reduce morbidity and mortality from EVD/MVD
- Safe and effective vaccines to prevent EVD/MVD and stop filovirus transmission in human populations.

Drugs and therapeutics

When used early in EVD progression, basic supportive interventions can significantly improve the chances of survival, including: fluids and electrolyte infusions; oxygen therapy; medications to control blood pressure, reduce vomiting and diarrhoea, manage fever and pain; treating coinfections [127].

A clinical trial started during the 2018-2020 outbreak in the DRC was conducted to compare the mortality among patients who received one of three investigational drugs for EVD with a control group who received the monoclonal antibody cocktail, ZMapp [128]. The study found that two antibody-based drugs, mAB114 and REGN-EB3, dramatically reduced the mortality rate of those hospitalised soon after infection. These antibody-based drugs were found to be superior to ZMapp in reducing mortality from EVD. These results led the trial's monitoring board to stop the study and randomise all remaining patients to either mAb114 or REGN-EB3 which are currently being used to treat patients with EVD [128]. Despite these promising outcomes a large proportion of patients in clinical trials still die from the disease, highlighting the potential to improve treatment strategies [128].

In October 2020 the US FDA became the first regulatory body to approve the use of Regeneron's REGN-EB3, now called Inmazeb, for the treatment of EVD [129]. Inmazeb consists of three monoclonal antibodies (atoltivimab, maftivimab and odesivimab-ebgn) that bind to different epitopes of ZEBOV GP, helping to neutralise the virus by blocking its ability to infect cells and/or enhancing the host immune response to target infected cells.

Vaccines

On December 2019 the US FDA approved the first ebolavirus vaccine, rVSV-ZEBOV (tradename "Ervebo"). rVSV-ZEBOV is a recombinant vaccine consisting of a genetically engineered vesicular stomatitis virus expressing the GP of EBOV, which results in a neutralising immune response to infection with EBOV [130]. The vaccine has been found to be safe and protective against EBOV when used in ring vaccination (i.e. vaccinating those close to the infected individual, effectively creating an immunity ring around them). The vaccine is also approved for medical use in the European Union and has been licensed in four African countries, The Democratic Republic of the Congo (DRC), Burundi, Ghana and Zambia [131].

In addition, another vaccine that combines two vaccine components, Ad26.ZEBOV (Zabdeno) and MVA-BN-Filo (Mvabea), was developed and rolled out under research protocol in 2019 to combat the outbreak in the DRC. Ad26.ZEBOV is a recombinant adenovirus vector expressing the GP of EBOV whereas MVA-BN-Filo is a Modified Vaccinia virus Ankara strain containing EBOV, SUDV, Marburg Virus GP and TEBOV NP inserts. This vaccine requires a second "booster" dose 56 days after the initial dose. It was approved for marketing authorisation by the European Commission in July 2020. Several other candidate vaccines have or are undergoing clinical evaluation at different trial phases [132].

The first widespread use of a vaccine occurred in the DRC during the 2018-2020 outbreak, which ended on 25th June 2020. The rVSV-ZEBOV vaccine was given to more than 300,000 people who had been in close contact with infected individuals and their contacts. The vaccine was highly effective resulting in more than 80% of those vaccinated not contracting the disease [133]. Following the approval of the first ebolavirus vaccine late in 2019, Gavi (the global vaccine alliance) initiated the creation of a global emergency stockpile. The stockpile will include rVSV-ZEBOV and other vaccines if and when they successfully complete clinical trials. Gavi-supported low and middle-income countries (LMICs) are able to access the vaccine resource free of charge with additional support for the operational costs of vaccination campaigns as part of the investment funding until 2025 [134].

Success of disease control

The success of EVD control measures depend on the early diagnosis, treatment and quarantining of patients, contacting those who have been in close proximity to those infected and administering vaccinations. Prompt diagnosis is key to abating the spread of the EVD and has an impact not only on patient survival but also the severity of their disease.

Medical countermeasures (MCMs) can face barriers to their effective rollout and maintenance. For example, an unstable political environment interrupted the treatment of patients and vaccination programmes during the 2018-2020 outbreak in the DRC. Logistical challenges, relating to geographical remoteness and accessibility, impacted efforts to manage the most recent outbreak in the DRC's Équateur province, which was declared over on the 18th of November 2020. There can also be challenges around trust and disease stigma, and the willingness of individuals to interact with the health authorities – work is needed at a grassroots level to overcome these.

4.3 Current landscape of diagnostics

Diagnosis of Ebola virus disease

The WHO recommends that all EVD cases (probable and suspected) should be referred to a designated Ebola Treatment Centre or appropriate health care facility where trained medical staff should safely collect the appropriate specimens [135]. The WHO document *Laboratory diagnosis of Ebola virus disease: interim guidance 2014* provides detailed guidance on the type and timing of specimen collection, sample storage, transportation requirements and recommendations that all diagnostic testing should be carried out in BSL-3/-4 facilities unless the pathogen has been inactivated, in which case a BSL-2 laboratory can be used [135]. WHO recommends the detection of viral RNA by RT-PCR and/or antigen detection by specific tests and/or by detection of IgM antibodies. Two negative RT-PCR tests at least 48 hours apart are recommended before patients are discharged from hospital.

It is recommended that the first 25 positive cases and 50 negative specimens (from patients with suspected disease) detected by a country without a recognised national reference Viral Haemorrhagic Fever (VHF) laboratory should be sent to a WHO Collaborating Centre for VHF for secondary confirmation testing [135]. The following WHO Collaborating Centres for VHF have the capacity to confirm EVD:

- National Microbiology Laboratory Public Health Agency of Canada (Winnipeg, Canada)
- Institut Pasteur de Lyon (Lyon, France)
- Centre International de Recherches Médicales de Franceville (Franceville, Gabon)
- Bernhard-Nocht Institute for Tropical Medicine (Hamburg, Germany)
- Kenya Medical Research Institute (Nairobi, Kenya)
- Institut Pasteur de Dakar (Dakar, Senegal)
- National Institute for Communicable Diseases (Johannesburg, South Africa)
- Uganda Virus Research Institute (Entebbe, Uganda)
- Centers for Disease Control and Prevention (Atlanta, United States of America)

Similarly, for countries with a national reference VHF laboratory, the initial positive cases should also be sent to a WHO Collaborating Centre for VHF for confirmation. If results are concordant, laboratory results reported from the national reference laboratory will be accepted by WHO. The requirements for confirmation of positive tests differ between countries, for example the US CDC perform ebolavirus NP and VP40 RT-PCR to confirm test results [136].

During the 2013-2016 outbreak, diagnosis relied on collection of blood samples from symptomatic individuals in a biocontainment laboratory facilities. Limited availability of such facilities caused challenges with specimen collection, data management and resulted in

prolonged turnaround time for results. In response to this, the WHO produced a *Target Product Profile* for EVD diagnostics [137]. This supported efforts to develop a number of novel EVD diagnostics that met the criteria outlined in this document including nucleic acid amplification tests, many of which received WHO and/or FDA emergency use authorisation (EUA) status; a list of approved tests and an outlook on the future diagnostic landscape are reviewed in the literature [138].

Cell culture

The traditional gold standard method to confirm EVD is via cell culture, typically carried out in Vero-E6 cells. This method was used to discover and identify the virus in 1976. The virus can then be directly visualised by electron microscopy or by immunofluorescence microscopy. Whilst these methods are highly definitive, they are slow and require BSL-4 containment, which are typically restricted to research and public health laboratories, reducing the ability to perform the tests in the field.

Immunoassays

There are two main types of immunological assays used to detect and diagnose EVD: ELISAs, which require high biosafety facilities, take longer but are more sensitive and specific, and lateral flow immunoassays, which are much faster and can be used at the point of care but are limited by their low sensitivity and specificity. These tests can be developed to indirectly or directly diagnose infection by the presence of antibodies specific for the target pathogen or the pathogen antigens, respectively.

Enzyme-linked immunosorbent assays

ELISAs for ebolavirus-specific IgM and IgG were first clinically employed by the CDC in 1995 in the DRC and are still in use by the CDC [138]. However, due to the variable onset of humoral response in individuals, detection of IgM and IgG antibodies varies between patients with detection possible between two to 10 days and six to 19 days, respectively, after symptom onset [138]. This means these tests have limited diagnostic utility in time-limited settings. In addition, antibody response is often not detected in infections with fatal outcomes, indicating a lack of humoral response [139]. However, the presence of IgG in survivors is a useful tool for population-level sero-prevalence studies [138].

Antigen detection is therefore the most commonly used approach for ebolavirus. Antigen testing for ebolavirus via ELISA can detect the pathogen from around three days after symptom onset in all patients, regardless of disease severity. ELISA antigen detection tests are employed by some national reference laboratories utilising monoclonal antibodies against NP, VP40, or GP proteins (generated from mice immunised with purified or recombinant ebolavirus proteins) (a list of those used can be found in [140]), however, the use of these tests for clinical diagnosis has not been reported, which is thought to be due to the emergence of faster RT-PCR tests.

Lateral flow immunoassays

A number of commercially-available, point of care rapid antibody mediated antigen capture tests were developed during the 2013-2016 outbreak. This included the first rapid antigen test, ReEBOV Antigen Rapid Test kit (Corgenix, USA), which was assessed by the WHO through Emergency Assessment and Use [141]. Further tests recommended by WHO and/or the FDA include: OraQuick Ebola Radpid Antigen Test (OraSure Technologies) and SD Q Line Ebola Zaire Ag (SD Biosensor). These tests are chromatographic lateral flow immnoassays (LFIs) and are used to detect viral proteins including EBOV VP40, NP and GP in whole blood, plasma or serum.

Typically, LFIs are less sensitive than RT-PCR and work more reliably in patients with higher viral loads. For example, a recent study of their use during the 2013-2016 outbreak found

that none of the four LFIs tested (One step Ebola test (Intec), DEDIATEST EBOLA (Senova), ReEBOV Antigen test Kit (Corgenix) and SD Ebola Zaire Ag (SD Biosensors)) had both high sensitivity and specificity [142]. The researchers found that the sensitivity of the tests ranged from 79.5% to 98.4% and the specificity ranged from 80.2% to 100%. Therefore, the WHO recommends confirmation where possible using molecular testing [141].

Molecular diagnostics

The first RT-PCR tests for EVD, developed by the US CDC during the 1995 EVD outbreak, targeted the L, GP and NP genes. RT-PCR has been found to be more sensitive than antibody or antigen detection ELISAs for EVD diagnosis when evaluated over the complete course of symptomatic infection. For example, assessment of a RT-PCR assay for NP gene detection during the 2000 outbreak in Gulu, Uganda, found that the test was able to detect RNA up to 72 hours before antigen became detectable by ELISA [143]. Sample types suitable for RNA extraction for EVD RT-PCR include blood, saliva and seminal fluid [135].

The detection of ebolavirus RNA by RT-PCR is variable in the first 72 hours of illness so recommendations are that if the test is negative during this period, a repeat test should be done after 72 h of symptoms [135]. How long RNA can be detected in infected patients by RT-PCR varies depending on the type of assay, however, this time period is important as it has implications for the re-entry of EVD survivors into the community.

Real-time amplicon detection using sequence specific probes is faster and more specific than conventional RT-PCR and is currently the most commonly used diagnostic for EVD. Several real-time RT-PCR tests are recommended for emergency use by the WHO and FDA, and six of these are commercially available as kits [144] (table 7). Three other assays are used by the US Department of Defence and the US CDC, however, these are restricted to use by these agencies and are therefore not commercially available.

Real-time RT-PCR has been found to correlate well with quantification of viral loads by plaque assay and the cycle threshold (C_T) values (thus higher viral RNA copy numbers) are indicative of prognosis, survival and whether convalescent patients were likely to carry the infectious virus [138].

Limited data are available regarding the sensitivity and specificity of the various laboratory developed and commercial ebolavirus real-time RT-PCR assays now employed by public health reference laboratories [140]. A 2016 study compared 11 RT-PCR assays (including the six commercially available) and found that although they had up to 100-fold variation in the limit of detection and 1,000-fold variation in the lower limits of quantification the tests were all reliable for EVD diagnosis and viral load estimation [145].

Table 7: Molecular diagnostic tests for ebolavirus (adapted from [138]). Commercial and laboratory developed tests (LDT) for EVD diagnosis that have CE-IVD, WHO emergency use assessment and listing and/or FDA emergency use authorisation status.

Developer	System	Regulation
Altona Diagnostics GmbH	RealStar® Filovirus Screen RT-PCT kit 1.0	CE-IVD, WHO
Altona Diagnostics GmbH	RealStar® Ebolavirus RT-PCT kit 1.0	CE-IVD, FDA
U.S. DoD, LDT	EZI Real-time RT-PCR Assay	FDA
CDC, LDT	CDC Ebola Virus NP real-time RT-PCR assay	FDA
CDC, LDT	CDC Ebola Virus VP40 real-time RT-PCR assay	FDA
TIB MOLBIOL Syntheselabor GmbH	LightMix Ebola Zaire rRT-PCR test	CE-IVD, FDA
Shanghai ZJ BioTech	Liferiver Ebola virus (EBOV) real-time RT-PCR kit	WHO
BioFire Defense	FilmArray NGDS BT-R assay	FDA
BioFire Defense	FilmArray Biothreat-E test	CE-IVD, WHO, FDA
Corgenix	ReEBOV Antigen Rapid Test	WHO, FDA
Cepheid	Xpert Ebola Assay	WHO, FDA

4.4 Current use of NGS

Sequencing technologies

The largest scale sequencing effort undertaken for ebolavirus, including real-time sequencing efforts within affected countries, took place during the 2013-2016 outbreak, with several international groups working collaboratively with local health authorities and national public health institutes to establish sequencing facilities in regional laboratories or in the field, alongside field diagnostic laboratories and treatment centres [1, 146-148]. Prior to the 2013-16 outbreak there were only 29 complete genomes of ebolavirus, however, currently there are 785 complete genomes of ebolavirus on the ViPR database, 687 of those are EBOV [149].

Multiple sequencing approaches have been used, and the choice of technology approach depends on the availability of sequencing platforms, accessible reagents and the proposed application. Multiple research groups have evaluated different aspects of the sequencing workflow to develop protocols that achieve the highest performance for the desired applications. Several commercial sequencer providers have their own Ebola panels for amplicon sequencing such as the Thermo Fisher Scientific Ion AmpliSeq Ebola Research Panel and AmpliSeq for Illumina Ebola Research Panel.

At the beginning of the 2013-16 outbreak researchers at the US CDC evaluated a number of sample and library preparation methods as well as two sequencing technologies – PacBio and Illumina – for their performance, genome coverage and accuracy [150]. Based on their findings they proposed the use of Nextera library construction with Illumina HiSeq2500 sequencing, a method subsequently employed by other researchers [148, 150, 151]. A further 232 samples from the 2013-16 outbreak sequenced at the US CDC were completed using these established methods [152]. This method was optimised for the application of genomic surveillance as well as being sensitive enough to allow for intrahost viral genome variation to be determined.

The ARTIC network has produced sequencing, bioinformatics and phylogenetic analysis protocols for EBOV [153]. The resources include a sequencing protocol using PCR amplification with custom primers to cover the entire genome of EBOV, and bioinformatics

and phylogenetic analysis protocols. The ARTIC protocols are designed to be platform and sample preparation agnostic meaning the protocols can be used across all sequencing technologies and were being used by the INRB for ongoing disease management efforts during most recent outbreaks in the DRC [154-156].

Several methods for genome assembly are employed including *de novo* [148, 150, 157] or reference based assembly from databases such as GenBank using one of several alignment software packages [158-161]. ViPR has an ebolavirus pathogen specific assembly database and analysis tools, however, they still require bioinformatics expertise when used.

The majority of ebolavirus genomes that have been sequenced come from host blood or plasma samples. However, other human biological samples have been reported in the literature such as, semen [119], breast milk and buccal swabs [1] for determining transmission chains and directionality of infection.

Ebolavirus sequencing studies

The location of sequencing facilities and in-country ability to deliver NGS changed dramatically during the 2013-2016 EBOV outbreak. Sanger sequencing was used to determine the emergence of a new EBOV variant, named Makona, in Guinea collected in early 2014, performed by WHO Collaborating Centres for VHF in France and Germany [162]. Samples were also transported to international laboratories such as the US CDC for sequencing throughout the epidemic [150]. However, from September 2014 the Liberian Institute for Biomedical Research (LIBR) established in-country genomic capabilities to monitor the viral evolution of EBOV Makona. The LIBR Center for Genomic Sciences acquired a benchtop sequencer, the Illumina MiSeq System. The operation and development of this sequencer was supported by the US Army Medical Research Institute of Infectious Diseases (USAMRIID) [163]. Similarly in Sierra Leone, the Chinese Mobile Laboratory Testing Team sequenced 175 EBOV genomes using a BGISEQ-100 (Ion Proton) platform between September and November 2014 [164].

Supported by collaborators at the Broad Institute, US, the Kenema Government Hospital (KGH) in Sierra Leone has developed facilities for managing the threat of Lassa fever, including establishing sequencing capabilities. As the EBOV outbreak took hold in 2014, the KGH mobilised their resources to provide diagnostic and surveillance support for the outbreak [165]. From July to November 2014, the Institut Pasteur of Dakar, Senegal, deployed a mobile laboratory in Donka Hospital, Guinea, using an Illumina HiSeq2500 [166].

With financial support from the United States Agency for International Development and Illumina, collaborators from the Broad Institute in the US and Redeemer's University in Nigeria established an Illumina MiSeq System in January 2015 at the African Center of Excellence for Genomics of Infectious Diseases (ACEGID) at Redeemer's University, Nigeria. The purpose was primarily to understand how the mutation rate of EBOV may impact the efficacy of currently available treatments. Later the collaboration obtained an Illumina MiSeq system for the Institut Pasteur de Dakar laboratory in Senegal [167].

The Ebola Outbreak Sequencing Support group was established in July 2015 as a coordinated effort involving the Sierra Leone Ministry of Health, WHO, CDC and collaborators in the UK including the University of Cambridge and the Sanger Institute. A field diagnostic laboratory was equipped to rapidly sequence all new EVD cases in Sierra Leone with the goal of understanding virus phylogeny [1]. The temporary laboratory was established within an Ebola Treatment Centre in Makeni, Sierra Leone for real-time sequencing of EBOV. Once the outbreak had resolved, the Ion Torrent PGM sequencer used in the field was transferred to the University of Makeni, Sierra Leone, to establish the Infectious Disease Research Laboratory where the sequencing facilities now permanently reside.

In parallel, researchers from the University of Birmingham, UK, established sequencing capabilities in Guinea for the real-time genomic surveillance in a field laboratories using a PCR-based amplicon protocol for sequencing on an Oxford Nanopore Technologies (ONT) MinION sequencer [148].

The National Institute of Biomedical Research (INRB) in the DRC has access to Illumina iSeq100 platform [155] and ONT MinION [154] and is being supported by multiple organisations including the University of Nebraska, the USAMRIID and the ARTIC network. Portable sequencing technology (notably, ONT MinION and Illumina ISeq) have been used in field laboratories in the DRC during the most recent outbreaks in the Équateur [155] and Nord- Kivu/Sud- Kivu/Ituri Provinces [155, 156] by the INRB.

Applications of sequencing

NGS is used extensively to understand viral transmission, support surveillance efforts, and collect epidemiological data. NGS data is also used to understand the rate of mutation, biology, characteristics of the disease as well as determining the efficacy of MCMs such as diagnostics, vaccines and treatments.

The extent of genomic diversity observed in viral samples collected at the beginning of an outbreak can be informative about whether there was a single or multiple spill-over events from animals to humans. As outbreaks of EVD have progressed and more genomic sequences have become available, analyses can reveal broader trends such as evolutionary rate of the virus, its geographical migration and factors that contribute to viral transmission, disease outcome and virus-host adaptation. When an outbreak is coming to an end, sporadic cases can be analysed to determine whether the new case is likely from a new spill-over event or transmission from a persistent or undetected infection. Alongside these more imminent uses of the genomic data, NGS can be used to better understand viral and disease biology. A detailed review on the use of NGS for EVD can be found in the literature [168]; a summary table containing examples of research publications highlighting applications of ebolavirus sequencing is available in appendix 8.2. Case studies focusing on particular examples are available below.

Case study 1: identifying disease origin and transmission characteristics

Case Study: Genetic variation of virus genomes from Sierra Leone, 2014.

Aim of study: To uncover the origins of the outbreak, perform phylogenetic analysis to determine the relationship of the variants circulating in 2014 to those that caused previous outbreaks and trace transmission events throughout the human population early in the outbreak.

Study details: Genomic analysis of 99 EBOV genome sequences from 78 patients in Sierra Leone were compared to each other and three published genomes from Guinea, and 20 sequences from previous EBOV outbreaks. Phylogenetic comparison to all 20 genomes from earlier outbreaks suggests that the 2013 West African virus likely spread from central Africa within the past decade.

Sequencing method: Gire et al evaluated two sequencing platforms and several sequencing workflows including sample and library preparations. These included:

- cDNA synthesis with Nextera library construction and Illumina sequencing
- NuGEN Ovation RNA-seq, Nextera library construction and Illumina sequencing
- Pacific Biosciences library construction and sequencing.

Advantages of methods chosen:

- Nextera library construction and Illumina sequencing provided the most complete genome assembly and reliable intrahost single-nucleotide variant (iSNV, frequency >0.5%) identification.
- The authors achieved deep sequence depth of >2000× spanning over 99% of coding regions, essential for identification of iSNVs.

Information provided: The 2013-16 strain in West Africa most likely spread from Central Africa within the past 10 years with comparisons between the 2013-16 viruses and the previous two outbreaks suggesting they shared a common ancestor around 2004. This confirmed that the outbreaks arose from multiple spill-over events from animal to human populations. The genomes from the 2013-16 outbreak were all similar enough to confirm that the outbreak was due to only one spill-over event followed by human to human transmission, which was different from previous outbreaks. This analysis was able to pinpoint the likely spread of EBOV from Guinea to Sierra Leone to 12 Sierra Leonean patients attending a funeral for a Guinean person who died of the virus.

How it informed public health: Genomic analysis may prevent future outbreaks by limiting host-human contacts as well as the initial introduction of filoviruses into the human population. Knowledge that human-to-human transmission was the primary cause of spread helped inform public health officials where to focus efforts to prevent further spread.

Source: Gire, S. K. et al. (2014). Genomic surveillance elucidates Ebola virus origin and transmission during the 2014 outbreak. Science 345: 1369–1372 [150]

Additional information associated with the study: Twelve primer sets from eleven published assays including EBOV-specific and Pan-filovirus primers were screened against a consensus EBOV genome from Sierra Leone. Interrogation of the viral genomes from the Sierra Leonean samples identified mutations that alter protein sequences and other biologically meaningful targets that may be important for the efficacy of MCMs. These findings highlight the importance of regular monitoring the suitability of these countermeasures in light of ongoing viral evolution.

Case study 2: identifying unknown sources and chains of transmission

Prevention of new transmissions from sporadic cases at the end of an outbreak requires rapid and efficient contact tracing. Identifying unknown sources and chains of transmission is therefore key to ending the outbreak. In addition, it is important to differentiate between a persistent infection causing a new chain of infections versus another spill-over event.

Case Study: Ending the 2013-16 EBOV outbreak in Sierra Leone.

Aim of study: To demonstrate that linkage of genomic data with epidemiological data can support contact tracing effort as well as identifying unconventional routes of transmission between individuals.

Study details: 554 EBOV genomes were sequenced from cases across Sierra Leone, including those from Makeni (Bombali district), Port Loko (Port Loko district), Kambia district, Kerrytown (Western Urban district), and Koinadugu district. They obtained genomes from blood, buccal swabs, semen and breast milk samples.

Sequencing method: RT-PCR amplification using the Thermo Fisher Scientific Ion Ampliseq workflow according to the protocol manufacturer with EBOV specific reagents and the Ion Torrent PGM sequencing platform.

Information provided: The genomes sequenced represented 4.5% of EVD cases reported in Sierra Leone during the 2013-16 outbreak and 23.8% of all Sierra Leonean cases during 2015. Evolutionary analysis found there were at least nine viral linages circulating in Sierra Leone.

How it informed public health: The Ebola Outbreak Sequencing Support (EOSS) was established in July 2015 as a coordinated effort from the Sierra Leone Ministry of Health, WHO, CDC and the local sequencing facility to rapidly sequence all new Sierra Leone EVD cases and rapidly place them in phylogenetic context. EOSS processed 21 samples from July-September 2015 and provided an additional level of information to field workers tracing the source of the infection.

For example, genomic analysis helped to determine the sexual transmission of a persistent infection 50 days after the last confirmed case in the area. This was important for public health as it ruled out the possibility of a new spill over event or transmission from another area.

Source: Arias, A. et al. (2016) Rapid outbreak sequencing of Ebola virus in Sierra Leone identifies transmission chains linked to sporadic cases. Virus Evolution 2(1): vew016 [1]

Additional information associated with the study: This study is one of several during the 2013-2016 outbreak that exemplified the ability to perform real-time genomic surveillance in the field. In doing so, the researchers also highlighted the many difficulties and barriers which need to be overcome in order to facilitate prompt, efficient outbreak responses in the future.

4.5 Future needs of NGS

WHO roadmap

The World Health Organisation roadmap for identifying the vision, underpinning strategic goals and prioritisation of research for Ebola/Marburg Research and Development (R&D) [169] is intended to accelerate the collaborative development of MCMs – diagnostics, therapeutics and vaccines – against EVD and Marburg virus disease (MVD), a virus in the same family as ebolavirus. Three areas of the WHO vision have been identified including:

- The development of rapid, accurate, point-of-care diagnostics for Ebola/Marburg virus infection to inform treatment, outbreak detection, and clinical trials;
- Safe and effective treatment and post-exposure prophylaxis to reduce morbidity and mortality from EVD/MVD
- Safe and effective vaccines to prevent EVD/MVD and stop filovirus transmission in human populations

Data produced from NGS will likely play a key role in the successful development of these MCMs as the genomic sequences can be interrogated to understand the biological consequences of viral genomic variation. NGS can also be used as an analysis tool to understand the impact on the virus mutation rate and if resistance to therapeutics is likely to occur.

NGS technologies, particularly portable sequencers, are specifically highlighted in the roadmap for the surveillance and epidemiological analysis of the virus using sequencing platforms. However, other areas of need mentioned in the roadmap that NGS could be useful for include:

- Differentiating EVD and MVD from other diseases with similar symptoms
- Proficiency testing to monitor and evaluate performance of diagnostic assays in the field
- Identification and validation of host biomarkers correlated with patient prognosis and disease progression, such as viral load and transcriptomic signatures.

Future potential of NGS for diagnostics

In addition to monitoring the efficacy of molecular diagnostics, therapeutics and vaccines development, NGS technologies have the potential to be used directly for patient diagnosis. Using NGS technologies could drastically reduce the time between pathogen outbreak and diagnosis of the first case(s) as there is no need for prior knowledge of what pathogen may be causing disease.

Determining which pathogen is causative for a patient's illness can be complicated by the overlapping symptoms presented by VHFs. Targeted amplification panels of primers to screen for multiple pathogens could be used to diagnose febrile patients (e.g. by MassTag PCR [170] and TaqMan qPCR [171]). These methods can be used with or without NGS for diagnosis [172]. Metagenomic analysis can also be useful [151, 173] when investigating an unknown pathogen or determining a diagnosis where PCR testing has yielded no results. Alternatively NGS panels that cover a large variety of known pathogens could also reduce the number of diagnostic tests done to find an answer.

Host genomics is also an active area of research. Early work has begun on investigating the transcriptomic signature of host response to EVD to serve as predictive biomarkers for survival [174]. In addition, investigations into whether transcriptomic biomarkers can be validated that will diagnose EVD before symptom onset are also on the horizon [175]. This is a useful area of research as currently, a major limitation in existing diagnostics is their inability to detect early disease. The discovery of biomarkers early in the infection and those

that may be linked to more severe disease will have benefits for individual patients in terms of clinical management and potentially speed up time to diagnosis and therefore reduce transmission.

Ebolavirus NGS challenges

Many of the opportunities and challenges encountered for EVD sequencing are not unique to this pathogen (see chapter 6). Challenges specific to EVD include ensuring sample representation is adequate to gain the maximum from surveillance efforts. Difficulties in obtaining samples, their storage and preparation during an epidemic can limit the capacity of genetic analysis to uncover directionality of transmission, one of the most valuable applications of NGS for EVD. The larger the number of sequences being produced and from different geographic regions the greater the granularity and accuracy of phylogenetic models to map and predict transmission events. Therefore it is vital that sequences collected are representative of the number of infections, the geographical spread and over time. The metadata associated with samples also needs to be accurate and linked to the sequences when the data are shared. In addition, the number of strains available for sequencing is critical for NGS to be a useful adjunct technology to the epidemiological study of transmission pathways.

The majority of positive samples that get sequenced are from laboratory-confirmed EVD infections within hospitalised patients. Individuals receiving hospital care may differ from those not seeking medical help. The latter may have less severe disease and so sequencing the viral genomes from these patients may give insights into virulence factors, host/virus interactions and viral evolution, though this may be difficult to achieve.

Ebolavirus outbreaks result from multiple spill-over events from the animal reservoir of the virus. There is a need to monitor the potential animal reservoirs for ebolavirus to determine which animals are causal in these spill-over events.

4.6 Ebola conclusions

The use of NGS for ebolavirus sequencing has changed dramatically over the past few years – for example at the beginning of the 2013-16 outbreak samples were sent overseas to international laboratories for sequencing, however the situation evolved to include the development of in-country responses where national institutions were equipped with sequencing facilities and/or deployed mobile laboratories with portable sequencers. Many of these facilities and infrastructures developed during the 2013-16 still exist but there is a varying extent of use between different institutions. Whilst some are actively using the capabilities developed (e.g. INRB in the DRC, ACEGID in Nigeria and the Institut Pasteur in Dakar) it is not clear if or how other institutions are using their sequencing platforms (e.g. the University of Makeni in Sierra Leone and the LIBR in Liberia).

NGS technologies are being used for EVD to confirm new outbreaks, enhance surveillance and epidemiological efforts during an outbreak, and identifying persistent infections and transmission events once the outbreak is ending. NGS is useful for monitoring the evolution of ebolavirus to determine any detrimental impacts to current MCMs such as diagnostics based on the RNA sequence, vaccines and treatments. Efforts should therefore be concentrated on ensuring that samples are representative of the population and over time to support surveillance and monitoring impacts on MCMs during an outbreak.

NGS is also proving invaluable for research into the biology of the virus and disease mechanisms including investigations in novel biomarkers from host response, and characterising the efficacy of interventions. In particular, efforts should be focused on determining the animal reservoirs and monitoring those populations to predict and potentially prevent spill-over events.

A range of approaches and technologies are appropriate for sequencing ebolavirus, and as with other pathogens the choice of technology approach will depend on the availability of technology, the application, as well as the relative advantages and limitations of each.

5 YELLOW FEVER

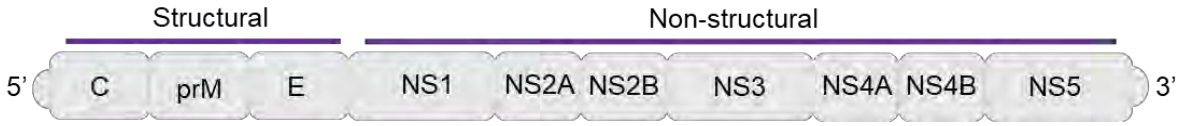
5.1 Yellow fever disease

The yellow fever virus belongs to the *Flaviviridae* family, which includes several other viruses which can cause viral haemorrhagic fevers, including Dengue fever virus. It is an arbovirus (arthropod-borne virus), transmitted amongst vertebrate hosts by arthropod vectors such as mosquitoes and ticks. Like all flaviviruses it is a positive-sense, single-stranded RNA virus, which exploits the presence of host RNA-dependant RNA polymerases in order to replicate.

Yellow fever virus

The yellow fever virus (YFV) genome is approximately 11kb in length, consisting of structural proteins (C, M, and E) that form the viral particle (virion) and non-structural (NS) proteins (NS1, NS2A, NS2B, NS3, NS4A, NS4B, and NS5) that allow the virus to replicate (figure 7) [176]. To date seven genotypes of the virus have been described, comprising five African genotypes (West I, West II, East/Central and East Africa and Angola) and two South American (South America I and II) [177]. There is not a consistent definition of what constitutes a distinct genotype however is has been defined as “distinct lineages which differ by greater than 9% at the nucleotide sequence level” [178]. On average the genetic sequences of the YFV genotypes differ from each other by 20-25% [177]. All genotypes belong to the same serotype so the differing sequences do not appear to affect their cell surface antigens, and the yellow fever vaccine is currently effective against all YFV genotypes.

Figure 7: Schematic of the yellow fever genome (adapted from [176]).



Compared to other RNA viruses the genome of yellow fever evolves relatively slowly, as evidenced by the fact that analysis of the same genotypes sampled decades apart show few sequence alterations [177]. It has been reported that the rate of substitutions ranges from 2 to 5 $x10^{-4}$ substitutions per site per year [178]. There is some evidence that the replicative fitness and virulence differs between the YFV strains belonging to the different genotypes [178]. However there is insufficient data to currently understand why these differences occur, for example if certain genetic variants are responsible for certain physiological traits.

Epidemiology

It is widely accepted that the virus first emerged from Africa between 1500 and 3000 years ago, though the first outbreaks were reported in the Caribbean in 1647 [179]. The virus is currently endemic in tropical and subtropical regions of 33 countries in Africa and 11 countries in South America [180] (figures 8 and 9) which have seen sporadic outbreaks despite the availability of a safe and effective vaccine since the late 1930s.

A recent study carried out as part of the Vaccine Impact Modelling Consortium, estimated the overall burden of yellow fever cases in 2018 to be ~109,000 severe infections and ~51,000 deaths [181]. The majority of cases were in Africa, where there were estimated to be ~100,952 severe infections and ~47,318 deaths. The burden of yellow fever in South America was estimated at ~8,331 severe infections and ~3,883 deaths.

Since the late 1980s there has been a resurgence in yellow fever cases, with an increase from 2016 including in non-endemic regions [182, 183]. This has led yellow fever to be considered a re-emerging disease, with a large number of countries in Africa and South America considered at-risk (figures 8-10).

Figure 8: Areas with risk of yellow fever virus transmission in Africa, produced by the US CDC and updated January 2017 [184].

A large outbreak occurred in Brazil between December 2016 and March 2018 where there were over 2000 confirmed cases in humans including over 500 deaths, as well as 4000 events in non-human primates (epizootics). In recent years, other outbreaks with more than one hundred reported cases have occurred in Uganda (2011), Sudan (2012), Democratic Republic of Congo (DRC) (2014), Angola and DRC (2016), and Nigeria (2017-19).

Aside from outbreaks, 11 cases of yellow fever were imported into China in 2016, resulting in one death. Between 2016-2018 isolated cases have also been imported into several European countries, including Denmark, France, The Netherlands, Switzerland and Great Britain [183].

The introduction of the virus by an infected person into regions with high mosquito density and low vaccination rates presents the highest risk of a large-scale epidemic [185]. Cases are often detected in remote or mobile communities which are hard to reach and so have low levels of immunity [186]. The majority of outbreaks occur when humans come into close proximity with the non-human primate hosts for the disease, the risk of transmission directly between humans is much less common. In order to understand viral epidemiology it is also important to monitor host species, as they can act as an early warning system to signal the likelihood of a new outbreak in an area.

Figure 9: Areas with risk of yellow fever virus transmission in South America, produced by the US CDC and updated May 2018 [187].

Transmission

Different species of mosquito can transmit the virus between non-human primate hosts in endemic countries, maintaining a natural cycle of the virus. Mosquitoes are also responsible for transmitting the virus between non-human primates and humans, and directly between humans. In tropical Africa the mosquito vectors are *Aedes* species whilst in South America mosquitoes are of the *Haemagogus* and *Sabethes* species. The precise species responsible for transmission varies depending on the environment, with some mosquito species residing in forest areas and others predominantly adapted to urban settings. There are three phases of transmission referred to as 'sylvatic', 'intermediate' and 'urban', which are important for understanding yellow fever epidemiology and disease control [185]:

- Sylvatic (or jungle): Transmission occurs mostly between non-human primates in tropical rain forest settings, maintaining a reservoir of the virus. Humans travelling or working in forest settings may be infected by a bite from a mosquito transmitting the virus from a non-human primate.
- Intermediate: This can occur in villages where both humans and non-human primates reside. Semi-domestic mosquitoes which reside both in the wild and around

households transmit the virus between non-human primates and people, often leading to outbreaks across multiple areas. This is the most common type of outbreak in Africa.

- Urban: Transmission by mosquito directly between people, with no need for non-human primate hosts. This can occur if the virus is introduced into heavily populated areas with high mosquito numbers and low vaccination coverage. This form of transmission is rare but has the potential to cause large scale outbreaks.

Yellow fever is a seasonal virus, and whilst outbreaks can occur throughout the year they are more common in the rainy season, when conditions promote mosquito breeding.

In addition to these modes of transmission, yellow fever is occasionally exported out of endemic countries when travellers or those working abroad return home. In this way yellow fever has been detected in the UK, France, the Netherlands and China in recent years. Whilst most of these countries are at very low risk for further yellow fever transmission, some regions of Asia have environments that support non-human primates and mosquitos that could carry the virus and so in theory could be at higher risk.

Clinical symptoms

Upon infection the YFV incubates in the body for a period of 3-6 days. Most patients are asymptomatic or have mild symptoms, which most commonly are fever, muscle pain with prominent backache, headache, loss of appetite, and nausea or vomiting [185]. Most symptoms disappear after three to four days, however up to 25% of patients may then transition into a more serious 'toxic phase' of illness, typically within 24 hours after recovering from initial symptoms. In this phase the patient can suffer high fever alongside kidney and liver damage as well as internal haemorrhage. There are no specific treatments for the yellow fever, and treatment relies on symptom managing such as treating dehydration, fever, liver and kidney failure, which improves outcomes [185]. Approximately 15-50% of patients who enter the acute phase die within 7-10 days [185]. Whilst the disease affects all age groups the highest mortality rates are in infants and the elderly.

5.2 Research and development needs for disease control

The Eliminate Yellow Fever Epidemics (EYE) Strategy has been developed by the WHO, Gavi and UNICEF to support 40 at-risk countries in Africa and the Americas to prevent, detect, and respond to yellow fever suspected cases and outbreaks [188]. The partnership aims to protect at-risk populations, prevent international spread, and contain outbreaks rapidly, primarily through the most effective use of vaccination but also through building surveillance and outbreak response capabilities. The strategy was launched in 2017, and by 2026 it aims to ensure over 1 billion people will be protected against the disease.

Unlike with the WHO roadmaps for Lassa fever and Ebola, the EYE strategy does not lay out specific research and development aims for disease control, but it is intended that a working group will be used to identify these. Beside current priority research areas such as vaccines, vector control, diagnostic and case management issues, the group will identify public health research priorities and activities.

Yellow fever drugs and therapeutics

There are no specific anti-viral treatments available for yellow fever, with medical treatment focusing on easing symptoms such as dehydration, fever and muscle pain. The availability of a successful vaccine (see below) means that disease control strategies are often focused on prevention rather than finding a successful treatment. However the availability of a good antiviral therapy would have obvious benefits in reducing mortality in severe cases, especially whilst vaccine supplies remain limited. As a result new treatments are being actively researched, including a therapeutic anti-YFV antibody which showed early promise in recent phase I clinical trials [189].

Yellow fever vaccination

Yellow fever is primarily controlled through vaccination of at-risk populations, due to the availability of an extremely effective vaccine which is safe and affordable. This vaccine is a live attenuated form of the virus known as the 17D strain. A single dose of yellow fever vaccine is sufficient to confer sustained immunity and life-long protection against yellow fever disease [190]. Several vaccination strategies are used including routine infant immunisation, mass vaccination campaigns designed to increase coverage in countries at risk, and vaccination of travellers going to yellow fever endemic areas. Emergency vaccination campaigns are essential when outbreaks occur in high-risk areas where vaccination coverage is low, which requires a strong surveillance system with rapid detection and reporting of yellow fever cases. Changes in human movement patterns, different risk factors for vulnerable countries and communities, and increased numbers of disease vectors, mean that vaccination approaches will have to be both adapted for individual countries and scaled up to address risk of global disease [188].

Most research and development needs are therefore centred on improving the vaccination strategy, although other forms of disease prevention such as vector control are also important. There are several other areas that require further research and improvement if the vaccination strategy and other efforts to control yellow fever are to be successful. Key areas include:

More rapid and accurate diagnostics

Success of emergency vaccination programmes depends on active surveillance and reporting of suspected cases, in both humans and non-human primate hosts. This requires diagnostic testing, with the WHO recommending that every at-risk country has at least one national laboratory where basic yellow fever blood tests can be performed [185]. Positive results must then often be confirmed by a regional reference laboratory. Significant delays can occur due to the time taken to transport samples to national laboratories, as well as the need to often use several diagnostic tests to overcome the limitations of individual tests. The diagnostic process can take up to a month in some African countries – until 2018 just one regional reference facility in Dakar, Senegal provided yellow fever confirmatory testing for the whole of Africa [191, 192]. EYE hopes to improve testing rates by developing more regional laboratories, as well as introducing more rapid diagnostic testing methods. A new regional reference lab recently opened in Uganda in October 2018, with another planned in Cameroon [191].

Existing molecular tests are in addition frequently inaccurate. This is partly due to disease samples being collected at too late a time point in the disease course for viral RNA to be detected by the test. However inaccuracies can also occur when primers for PCR based tests are designed and optimised to work on certain strains, and so may be less effective on less commonly encountered strains [193]. In addition a lack of standardisation and quality control can result in poor test performance. Research is required to improve the performance of current testing methods, and ensure the correct tests are used in the correct regions. Target product profiles have been developed by WHO partnered with FIND and Gavi to help identify the needs that must be met for three different types of diagnostics: a standardised molecular assay test kit, a standardised immunoassay test kit, and a rapid diagnostic test [191].

Improved surveillance, monitoring and prediction of viral spread

Due to high demand, there is currently a limited global supply of vaccine, which means that vaccinations must be targeted to ensure areas of the highest risk are covered. In addition countries may not be able to afford vaccination of entire populations, and instead it may be a more effective use of resources to target specific regions. This requires improved surveillance and monitoring of the virus including in non-human primate hosts and disease

vectors. This could give earlier notification of areas at particularly high risk and provide information on likely routes of disease transmission, allowing vaccinations and other interventions to be targeted appropriately. It is not clearly understood why some regions in Asia do not have yellow fever, despite having the right conditions for the disease [178]. This needs to be more clearly understood in order to predict future disease risk in these areas.

Increased understanding of vaccine induced immunity

Prior to 2016 the International Health Regulations (IHR) recommended that a 10-year booster dose of the yellow fever vaccine was required to maintain immunity. Reports from the WHO Strategic Advisory Group of Experts (SAGE) and the United States Advisory Committee on Immunization Practices (ACIP) led to the change in guidance that resulted in the current recommendation of only one dose of vaccine being sufficient. However there is still uncertainty over exactly how long immunity lasts, especially among children, and some travellers to high risk areas may still be recommended to have a 10-year booster. Therefore further research into this area to clarify the situation is warranted [194].

Improved understanding of virus biology

YFV research that has been carried out has largely been on the 17D vaccine strain of the virus, resulting in limited understanding of the wild type virus [178]. However the resurgence of yellow fever cases since the late 1980s, with a significant increase from 2016, has highlighted the need to further understand wild type viral biology and genetics, in order to inform and improve disease control [178, 195]. There is limited understanding of virus biology and if the different lineages and strains have different physiological effects, such as causing more or less severe symptoms and outcomes. More genome sequence data is also required to support other disease control methods, for example phylogenetic analysis to track viral evolution and transmission dynamics in humans, primates and mosquito vectors, to better inform epidemiological studies and subsequent viral prevention and control strategies.

5.3 Current landscape of diagnostics

Yellow fever diagnostic tests are required both to inform patient management and to ensure that reported cases are confirmed as soon as possible to allow preventative vaccination strategies to be put in place [191]. Laboratory confirmation is crucial as the early symptoms of yellow fever can overlap with those of other viral haemorrhagic fevers that can occur in the same regions, including dengue, Zika, Ebola virus disease, Lassa fever and Marburg disease. In rare cases the yellow fever vaccine can cause yellow fever symptoms in patients. In these cases, it is important to ascertain whether the symptoms are due to the vaccine or wild type yellow fever, which is not always clear when patient vaccination histories are lacking. Testing is also used on post-mortem cases of suspected yellow fever, to ensure accurate reporting of cases.

Diagnostic process

A number of laboratory diagnostic tests are available for confirmation of suspected cases of yellow fever. Serological tests for IgM antibodies and molecular PCR tests are the most commonly used in routine diagnosis, though further antibody-based tests and occasionally viral culture may be needed to confirm diagnoses, and immunohistochemistry may be used in post mortem cases.

The type of tests that are carried out depend on the guidelines for routine testing procedures in different countries or continents, the availability of different tests and reagents within testing laboratories, and the characteristics of different samples which determines which tests are most suitable to confirm or rule out a case of yellow fever. For example PCR methods are only suitable within a few days of infection (up to 10 days maximum) when the

viral RNA is still detectable, whilst antigen based methods are only suitable once antibodies have developed, from approximately six days post infection [196].

Yellow fever is a BSL-3 pathogen and so requires suitable facilities where the laboratory staff are vaccinated. In some cases where it is unclear if disease symptoms are due to yellow fever or other viral haemorrhagic fever pathogens, a BSL-4 laboratory may be required.

Most tests are developed in-house and not commercially available, or if commercially available are intended for research use only and not fully validated for clinical use [191] (table 8). Therefore testing protocols vary widely between different laboratories, and testing is not widespread but generally performed at a small number of reference laboratories [197]. This results in limited official diagnosis and reporting of YFV as cases often occur in remote and rural areas.

Serological tests can pose particular problems for YFV detection, as they exhibit significant cross-reactivity with antibodies generated in response to other infections circulating in the same areas, including Dengue, Zika or West Nile virus. These tests will also detect vaccine strains of the virus, therefore they are not suitable for use in vaccinated patients and a knowledge of vaccine history is required.

Table 8: Overview of RT-PCR systems for yellow fever virus detection used by 34 laboratories in the Emerging Viral Diseases-Expert Laboratory Network external quality assessment of molecular detection of yellow fever virus, March 2018. 46 tests were used in total, with some laboratories using more than one test. Commercial tests are highlighted in bold. Table adapted from Domingo et al, see original paper for full references [193].

RT-PCR used and specificity	YFV genome target	No. laboratories
In-house YFV specific (wild-type/vaccine strain)		
Domingo et al 2012	5'-UTR	4
Drosten et al 2002	5'-UTR	7
Fischer et al 2017	NS1	2
Weidmann et al 2010	5'-UTR	3
Own design	not specified	7
Own design-adapted from Domingo et al 2012	5'-UTR	1
Fast Track diagnostics Tropic fever Africa (commercial)	unknown	1
Genesig (commercial)	unknown	1
Real Star YFV RT-PCR kit 1.0 (commercial)	unknown	1
In-house YFV vaccine strain specific		
Bae et al., 2003	NS3	2
Fernandes-Monteiro et al. 2015	NS5	1
Mantel et al. 2008	NS5	2
Pan-flavivirus		
Ayers et al. 2006 + seq	NS5	1
Moureau et al. 2007 + seq	NS5	1
Patel et al. 2013 + seq	NS5	3
Scaramozzino et al. 2001 + seq	NS5	4
Own design (pan-flavivirus) + seq	unknown	1
Patel et al. 2013	NS5	1
Scaramozzino et al. 2001	NS5	1
Genekam (commercial pan-flavivirus)	unknown	1
Other, no YFV detection		
Fast Track diagnostics Tropic fever Asia	NA	1
Total number of submitted tests	NA	46

NA: not applicable; YFV: yellow fever virus.

Gaps for current diagnostic tests to overcome:

In 2019 FIND partnered with Gavi and the WHO to publish target product profiles to establish the requirements of three different types of yellow fever diagnostic tests; a standardised molecular assay test kit, a standardized immunoassay test kit and a rapid diagnostic test [191]. During this process several limitations of current diagnostic tests were identified:

- There are currently no commercially available, fully validated diagnostic kits for molecular or serological detection of YFV infections. The use of 'Research Use Only' assays for patient diagnosis means testing, reagents and quality measures are not standardised across laboratories and tests.
- Laboratories with molecular testing capacities all use non-commercial (mainly use laboratory developed tests) and non-regulated PCR assays for both surveillance and patient management.
- Improving case detection through serological testing will require the strengthening and expansion of the current diagnostic testing capacity. It will also rely on the availability of new, easy-to implement diagnostic tests to simplify the serological component of testing algorithms.
- Currently, only RRLs and NRLs are involved in the testing of yellow fever suspected cases. However, hospital based or other laboratories could be involved in the testing of yellow fever suspected cases if new diagnostic methods became available that are easier to use, faster and more accurate.

In addition there are several other features of current tests (aside from the inherent limitations of different methods such as cross reactivity and time period use limitations discussed in chapter 2.4) and testing procedures that present opportunities for improvement. These are in addition to the more systematic issues affecting diagnostic tests in LMICs in particular, such as access to reagents, adequate laboratory facilities or tests designed to work in the field, and cost of diagnostic testing, which are discussed in more detail in Chapter 6.

Uncertain performance of PCR tests across genotypes. Molecular tests are often intended to detect all genotypes or strains, but are frequently only evaluated in one or few genotypes. Therefore it can be unclear how well these tests perform in reality, and could explain the poor performance rates of yellow fever diagnostic labs. In an international external quality assessment study of European laboratories, 13 of 32 produced non-optimal results due to failure to detect one or more YFV strains [193]. If some tests do perform better in certain genotypes than others, it will be necessary to understand which genotype of the virus is likely to result in an outbreak in a particular area, to determine the most appropriate tests to use. Sequencing could be used to identify strains circulating to ensure the most appropriate tests are used, as well as to help develop and validate diagnostic tests that detect multiple strains of the virus.

Lack of understanding of the viral kinetics of yellow fever infection and if different strains affect clinical response. Further research into viral kinetics and how they differ in mild and severe infections could improve the use of different diagnostic tests at different time points of infection, which may vary from patient to patient [177]. Using sequencing to obtain more genetic sequences of different viral strains will allow more research to be carried out to understand if different strains affect the clinical course of infection differently and allow the design and use of appropriate diagnostics.

Poor ability to distinguish between wild type and vaccine strains of yellow fever virus. Current PCR tests may have variable performance, and it is not possible to distinguish between wild type and vaccine strains of YFV using serological methods [198]. More specific tests are needed, with sequencing-based tests potentially a useful option in this scenario.

Use of samples beyond blood plasma and serum. Molecular diagnostic tests most commonly utilise blood plasma or serum samples. However other samples may extend the time period in which tests may be used. E.g. in urine samples viral RNA has been detected just one day after infection and shown to be detectable 21-27 days after the onset of symptoms (or following vaccination) [199, 200]. More research is needed to evaluate the potential of urine and other types of samples such as semen and whole blood.

Use of multiplex testing. In areas endemic for YFV and other flaviviruses it may be useful to perform tests in multiplex for multiple viruses simultaneously, to improve diagnostic yield. However few tests currently exist that facilitate this [197]. Metagenomic sequencing may be suitable in non-outbreak situations to identify which virus is causing symptoms, it may also be useful as a confirmatory test in cases where differential diagnostic testing yields positive results for both yellow fever and other viruses [151, 201].

Tests in development

ELISA YF MAC HD: The WHO and the CDC are currently trialling a new IgM ELISA test called YF MAC HD, which has been developed by the CDC [202]. The test is a complete easy to use kit, providing the necessary reagents and allowing the test to be standardised unlike current laboratory developed tests. In addition, reagents have been designed to be temperature tolerant and the test runs in 3.5 hours compared to the two day average of current ELISAs. The test does not overcome the cross reactivity issues of serology tests, but is still a significant improvement on current ELISAs and will provide a standardised solution to testing. The kit is currently being tested among national laboratories in nine countries – Angola, Cameroon, Democratic Republic of Congo, Senegal, Uganda, Brazil, Colombia, Peru, and Paraguay. It is intended for the test to be fully validated before being used on a larger scale.

NS1 antigen based tests: Non Structural Protein 1 (NS1) is a viral antigen that is highly conserved across flaviviruses. Tests to detect NS1 are currently used for the diagnosis of dengue virus, and tests based on the antigen are being explored for yellow fever, though all are still in the early research phase. Tests explored so far include an NS1-capture ELISA with 80% sensitivity (12/15 cases) and 100% specificity compared to RT-PCR [203], and inclusion of yellow fever and Dengue NS1 in a proof of concept study of a multiplexed silver nanoparticle-based assay, which was more sensitive than the ELISA but not evaluated with clinical samples [204]. More recently a rapid antigen based test kit was described where NS1 was conjugated to colloidal gold particles. The test appeared to be specific for yellow fever and not cross react with other epidemic flaviviruses and alphaviruses, but was not evaluated with clinical samples [205].

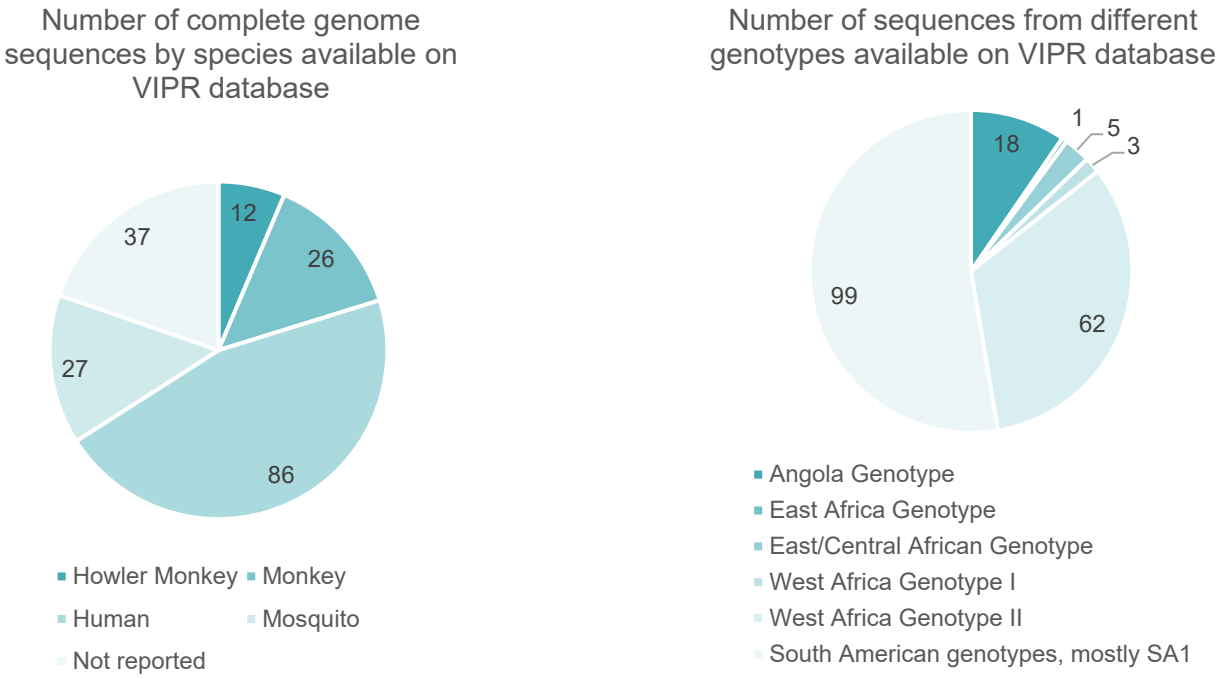
5.4 Current use of NGS

The full genome of YFV was first sequenced in 1985 [206], using sanger sequencing. Analysis of YFV sequencing publications available on PubMed showed that studies making use of pyrosequencing NGS methods first started to be published around 2012, whilst studies utilising more modern platforms such as those available from Illumina, Oxford Nanopore Technologies and Ion Torrent, started to be published relatively recently in 2017.

Despite the long-term use of sequencing, until recently few full YFV genomes had been sequenced. At the start of 2012 it was reported that only 25 full-length YFV sequences were available in the NCBI database, including seven sequences of YFV vaccine and nine sequences from human cases with an adverse event to vaccination [207]. Over the last decade the number of available sequences has greatly increased, likely driven by both the availability of NGS technology as well as multiple studies of recent outbreaks in South America, where much of the sequence data is derived from. As of 30th October 2020, 1031 YFV genomes are available on the ViPR database, of which 232 are complete genomes. Of the complete genomes at least 44 are vaccine strains, potentially slightly more due to

unclear reporting. 86 of the 188 remaining genomes are reported as being from humans, with the others obtained from howler monkeys, other types of monkey, mosquitoes, or were not reported (figure 10).

Figure 10: Schematic of the origin and genotype of yellow fever virus genomes in the ViPR database (accessed Oct 2020).



The current available sequences mostly belong to the South American I genotypes, with some African genotypes much less represented. Whilst this partly reflects the frequency of different genotypes, it is also largely a result of different levels of sequencing in different regions.

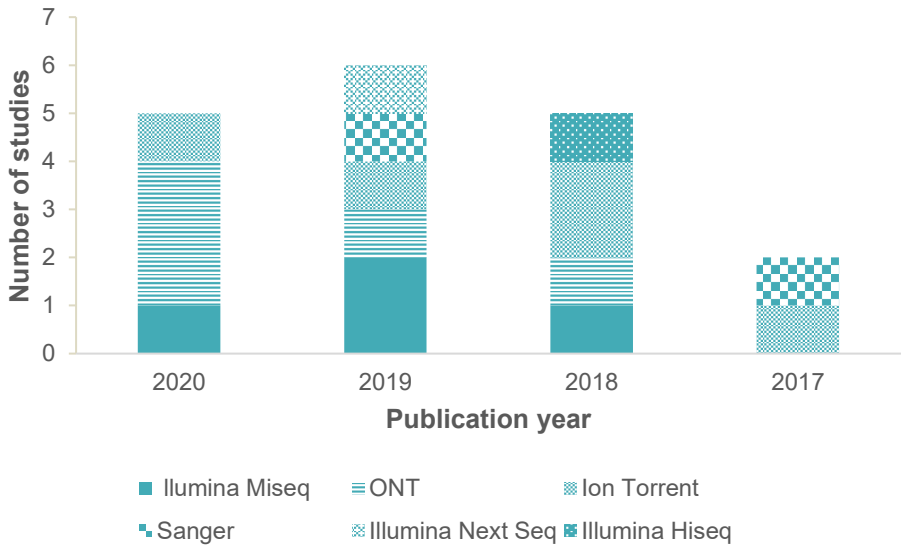
Whilst Sanger sequencing has been used sporadically for decades to identify YFV strains, the more recent use of NGS to produce an increased amount of sequence data and more high quality whole genome sequences has enabled detailed phylogenetic analysis of yellow fever outbreaks, as well as provided more data for research. YFV sequencing has been used worldwide to assist in yellow fever disease management, most commonly in regions of Africa and South America where the virus is endemic, but also in Europe, Asia and North America on occasions when the virus has been imported. Despite this, aside from its use in distinguishing between wild type and vaccine virus strains in the case of potential adverse events, YFV sequencing is currently not routinely used for any public health application. Similarly there is no specific guidance on the use of sequencing to assist in yellow fever surveillance and disease control. Instead sequencing has taken place in the context of research studies, and the extent to which the results have been used to inform public health varies depending on the specific context and collaborations of each study.

Sequencing methods used

As with other VHFs, there is no single method prioritised for viral sequencing and the more general considerations over the type of sequencing data required for the intended application, cost, accessibility of technology and time are the factors that determine which method is chosen. A review of papers indexed by PubMed showed that the majority of studies published between 2017 and 2020 (when modern NGS techniques have been most

widely used) have used an amplicon approach to obtain whole genome data, others have performed RNAseq without the use of amplicons, and one has performed metagenomics. Over the same period Illumina, Oxford Nanopore Technologies (ONT) and Ion Torrent technologies have all been used to a similar extent, though in 2020 the most commonly used technology was ONT (figure 11).

Figure 11: Number of studies listed on NCBI PubMed as being published between 2017-2020 on yellow fever sequencing, with platform used (search carried out 1 November 2020).



The availability of existing publicly available protocols and corresponding tools for data analysis for these technologies will allow for easier data standardisation and comparability between sequencing studies. Several of the most recent studies make use of overlapping amplicon multiplex PCR protocols published in by Quick et al [208] and the protocols and phylogenetic analysis tools made publicly available by Faria et al in 2018 [209]. In addition, although a range of bioinformatics and data analysis strategies are used, publicly available resources such as the KRISP Biotools website help increase accessibility to different forms of analysis [210].

Sample criteria

Sample type: A variety of samples have been used for sequencing studies. In humans blood plasma is commonly used and urine has also been used, but this is less common. In post mortem cases and in non-human primates tissue samples are used, most commonly liver in humans but a variety of tissues may be suitable. Many sequencing studies now sequence directly from the clinical sample, though in some instances viral culture is also used, such as when other research is being carried out to further characterise the sample.

Limited time period for sequencing: As with current molecular diagnostics, sequencing is only useful as long as sufficient viral RNA can be detected in patient samples. In yellow fever, the short period of viraemia means that viral RNA is often only detected in the blood for 3-6 days following infection. In addition viral loads are often low and highly variable. Therefore yellow fever genome analysis requires suitable samples to be collected in this

time period, which may not always be possible if patients only present themselves after a period of symptoms worsening. Urine has been shown to potentially contain YFV RNA for longer time periods which could make it a useful source, but further research into its use is required [200].

Yellow fever virus sequencing applications

One of the most common applications of yellow fever virus sequencing is to assist in surveillance efforts to monitor for new outbreaks of the infection. Similarly, genome sequencing has been used to perform epidemiological investigations in both animal and human populations to monitor viral evolution as well as track routes of disease spread both prior to and during human outbreaks. This can support more accurate targeting of limited vaccine supplies to populations at risk, and also inform other disease control measures, such as vector control. Whilst not routinely used in diagnosis, sequencing can also be used in certain diagnostic scenarios when symptoms are unclear or when there are concerns that symptoms are caused by adverse effects to vaccination. Finally, sequencing has been used in research and development to inform new vaccine design and quality control. In addition to the case studies below, further details of yellow fever sequencing studies are available in Appendix 8.3.

Tracking yellow fever virus transmission during an outbreak

Importantly for yellow fever, using NGS to track transmission during an outbreak is most valuable when performed in both humans and non-human primate hosts. The highly accurate location metadata that can be obtained from non-human primates which do not naturally stray far from their habitats can be useful in tracking an outbreak at higher resolution and placing human sequences in context, as human data can be less easily pinpointed to a specific location and is often not obtained at all. It is also essential to understand if human outbreaks are occurring as a result of sylvatic transmission or have transitioned to urban transmission cycles, which could result in a much larger and harder to contain outbreak. Although it is less commonly performed, sequencing of mosquito vectors and performing epidemiological analysis can also help in understanding chains of transmission. An important consideration is how best to identify and obtain samples from animal hosts. In Central and South America this is relatively straightforward as the virus often kills the animal host, acting as a clear indicator for surveillance and allowing samples to be taken easily. In Africa where the hosts are less susceptible to the virus, identifying epizootics and taking samples for analysis can be more challenging.

Case study 1: Genomic and epidemiological monitoring of yellow fever virus transmission potential

Aim of study: To integrate genomic data with epidemiological, and case distribution data from the 2016-2017 epidemic in Brazil, to estimate patterns of geographic spread, the risks of virus exposure, and the contributions of rural versus urban transmission.

Study details: 62 complete YFV genomes were generated from infected humans (n = 33) and non-human Primates (n = 29) from the most affected Brazilian states.

Sequencing method: A combination of methods were used, with all genomes generated in Brazil. Half were generated in Minas Gerais using a MinION portable YFV sequencing protocol adapted from Quick et al's multiplex PCR method [208]. Other samples were sequenced using an Ion Torrent and an Illumina MiSeq. To place the genomes generated in context 61 publicly available genomes were also used to allow phylogenetic analysis. An automated online phylogenetic tool was developed to identify and classify YFV gene sequences, which is now publicly available.

Advantages of methods chosen: In this case the sequencing technologies appeared to be chosen based on what the participating institutions had available.

Information provided: It was determined that the Brazilian outbreak strains belonged to a single clade within the South America I genotype. Comparison with samples from a previous outbreak showed it was predicted that the outbreak had originated from an endemic area in north-eastern Brazil and then moved South, rather than a lineage re-emerging in Minas Gerais (the centre of the outbreak). Sequencing information was combined with other types of epidemiological analysis to confirm the likely start date of the outbreak, to show that the outbreak was being sustained by sylvatic rather than urban transmission, and to estimate the speed of viral spread which was a relatively rapid 3.3 km/day. The high rate explained the magnitude of the outbreak and suggested the outbreak may have been partly driven by human activity, such as the transport of mosquitoes in vehicles or the trade and hunting of non-human primates.

How it informed public health: Though analysis was performed rapidly, this study was published after the peak of the outbreak had occurred, so was not used to inform public health in real time. However the study provides a demonstration of how genomic surveillance could be used to anticipate risk of yellow fever exposure and the likelihood of sylvatic vs urban transmission. The open access toolkit of protocols and analysis tools developed means that in future this could be more rapidly applied and used to assist in virus control measures. The large number of sequences provided will also be useful for future research studies.

Source: Faria, N. R. et al. Genomic and epidemiological monitoring of yellow fever virus transmission potential. Science. 2018. 361(6405): pp. 894-899 [209].

Monitoring of infected animal hosts

The sylvatic cycle of yellow fever transmission relies on infection of non-human primates, following which the virus is transferred to humans via mosquito vectors. Genomic analysis of the virus during outbreaks in non-human primates (epizootics) can provide valuable information on yellow fever strains present in a region, potentially before any human outbreaks have occurred or been reported. By performing phylogenetic analysis on the sequence data it can be possible to infer the origin of the virus, understand the likeliness of transmission to humans and identify if any major genetic variations have occurred that result in a new strain requiring further characterisation.

An example of where sequencing has been used in this way is in Bahia State in Brazil, which was the only region in North Eastern Brazil to experience no human cases native to the region during the 2016 outbreak, though there were a number of cases of yellow fever in non-human primates. Portable sequencing was used to recover eight complete genomes from confirmed cases of yellow fever in non-human primates, and phylogenetic analysis showed that there were two clusters of viruses, likely from two separate introductions [211]. Subsequent analysis suggested that the infected animals had likely been transported into the area through human activity. The results from the study both revealed that yellow fever may spread into a new area through animal trafficking, which could be used to inform future public health measures, and confirmed that in Bahia the prevention strategies used were effective as despite the high risk of transmission from non-human primates a local outbreak had been prevented.

Identifying the start of a yellow fever outbreak

Countries which are endemic or at high risk of yellow fever often have other viral haemorrhagic fever viruses circulating, for example West Nile Virus, Zika and Dengue Virus in South America and Lassa and Ebola viruses in Africa. As the symptoms of these diseases overlap, particularly in the early stages, it may be difficult to diagnose which virus is responsible, particularly in situations where yellow fever appears to be a much less likely candidate than other more prevalent haemorrhagic fevers. In these situations sequencing has been used in Africa to help diagnose yellow fever when tests for other viruses proved negative, indicating that yellow fever is circulating in the region and allowing measure to manage an outbreak to be put in place [201, 207]. In the future it is possible that sequencing may be useful as an initial test to more rapidly diagnose uncertain cases, which will not only allow better patient management but alert authorities to new outbreaks as quickly as possible.

Case study 2: Real-time metagenomic analysis of undiagnosed fever cases unveils a yellow fever outbreak in Edo state, Nigeria

Aim of study: To diagnose unexplained fever and poor outcomes in 50 patients at Irrua Specialist Teaching Hospital (ISTH) in Edo State, Nigeria between September 2018 to January 2019, after Lassa fever was ruled out. After identifying yellow fever, sequence data was used to understand the outbreak etiology and inform public health responses.

Study details: Of the 50 patient samples, 29 were found positive for YFV by RT-qPCR. Metagenomic sequencing and analysis was then used to confirm the presence of yellow fever and to rule out other viruses. 14 complete virus genomes were then assembled from samples and used for phylogenetic analysis to help understand the origin of the outbreak.

Sequencing method: Metagenomic sequencing was performed on viral RNA extracted from plasma samples, from which cDNA was synthesised using random hexamers. Sequencing libraries were prepared using the Nextera XT kit (Illumina) sequenced on the Illumina MiSeq platform with 101 base pair paired-end reads. Note that for sequence assembly samples with lower Ct values (meaning higher viral loads) were required, so not all samples were suitable.

Advantages of methods chosen:

- Metagenomic sequencing meant that the samples could be used to screen for the presence of several viruses at once, allowing co-infection with other viruses to be ruled out.
- It was possible to assemble genomes from a sufficient number of sequences, providing additional benefit from the use of this technique.
- Multiple samples were analysed, helping to confirm that yellow fever was the cause of the symptoms despite the virus being undetectable in some patients.

Information provided: The presence of yellow fever was confirmed within three days of the samples being received, and the presence of other viruses ruled out. The phylogenetic analysis suggested that the virus clade was not descended directly from previous outbreaks in Nigeria, but instead came from elsewhere in Africa. More detailed analysis was prevented by the limited availability of historical sequence data.

How it informed public health: The information was shared immediately with the referring hospital (ISTH) and the Nigeria Center for Disease Control (NCDC). Based in part on these findings, NCDC and the Nigeria Federal Ministry of Health declared an outbreak in Edo state the following day, prompting more samples to be sent for diagnosis. The sequences are the only complete Nigerian yellow fever genomes from patient samples collected after 1950, and provide a resource to be used to enable future improved phylogenetic studies. The study demonstrates how sequencing can provide an alternative to serology, which according to the authors is only performed in a limited number of laboratories and costs about one thousand US dollars per sample

Source: Ajogbasile, F. V., et al. Real-time Metagenomic Analysis of Undiagnosed Fever Cases Unveils a Yellow Fever Outbreak in Edo State, Nigeria. Scientific Reports. 2020. 10(1): p. 3180. [201]

Identifying imported cases of yellow fever

On the relatively rare occasions when a traveller returns from a yellow fever endemic country and is diagnosed with yellow fever, sequencing has been used in their native country to characterise the viral strain to confirm the origin of the virus, as well as contribute sequence data for use in further studies [199, 212]. Surveillance efforts such as these are particularly important for countries like China that have the right conditions for YFV transmission, in order to confirm if cases are imported or if they might be due to local spread of the virus. Globally, genomic surveillance of yellow fever strains also helps identify how frequently the virus is imported and helps to inform the need for travel vaccinations.

Identifying if symptoms are due to vaccine response or wild type virus

As described earlier, in rare cases yellow fever vaccine strain may cause symptoms that indicate a wild type infection. As vaccination coverage increases and is more widely used, more of these cases are likely to occur. In these situations it can be important to understand if what appears to be a case of yellow fever is due to a wild type strain in an unvaccinated individual or the vaccination itself. There is also the possibility that in the future a strain of yellow fever may develop that the current vaccine is ineffective against, though so far there is no evidence of this. Current PCR tests lack the specificity to distinguish between wild type and vaccine strains of yellow fever, making sequencing an important tool in this regard, which has been used on several occasions. For this application WGS is not necessarily required. For example, sequencing data has been used to inform the development of primers for a Sanger sequencing based virus genotyping tool which analyses only the NS5 region of the genome [198]. This could be used to investigate adverse events following vaccinations without the requirement for the more expensive and complex NGS methods.

Vaccine design and quality control

The WHO recommends that new seed- and working-lots of vaccines should have the viral genome sequenced in order to check vaccine genetic stability [213]. To enable this recommendation to be met one study has developed and standardised a Sanger sequencing protocol for the genetic monitoring of the Brazilian 17D vaccine [213]. In addition, access to genome sequence data can help to inform the design of new vaccines which have fewer side effects than the current 17D vaccine. For example when designing a multi-epitope vaccine one research group used the publicly available genome data of 77 Brazilian strains to ensure that their prediction model was relevant to current strains of the virus, increasing the likelihood of producing a successful vaccine [214].

5.5 Future needs of NGS

EYE strategy and other strategic documents

Use of NGS, either in the form of targeted or WGS sequencing strategies, is not mentioned specifically in the EYE strategy as a mechanism for eliminating yellow fever [188], nor is it included in the 2019 target product profiles for identification of yellow fever infection published by Gavi, FIND and WHO [191]. However this does not mean that sequencing could not be an important tool in helping to achieve the goals outlined in these documents, as demonstrated by the examples of its current uses above and its further potential uses outlined below. It is possible that for yellow fever the potential of sequencing to assist in disease control has not yet been fully appreciated, and will be incorporated into future disease control strategies. In addition, implementation of sequencing infrastructure and protocols for other viruses may lower the barriers to implementation of a YFV sequencing strategy.

Increased number of complete genome sequences

Although there has been an increase in sequencing studies in recent years, more complete genomes are still required. It is not necessary to obtain a full YFV genome for all purposes, for example to identify a strain only a segment of the genome may be required. However for more detailed phylogenetic and spatio-temporal studies full or near full genomes with high coverage of each base position are required to ensure accurate results, as sometimes the genomes of closely related viruses differ by only a few nucleotides [215, 216]. The more sequences available from different periods for phylogenetic analysis, the more accurate and impactful the resulting analysis will be.

Sequencing data that is more evenly representative of different genotypes

Some genotypes are currently more represented in databases than others. In Africa in particular, there have been very few yellow fever sequencing studies, despite multiple recent outbreaks on the continent. The availability of a vaccine that appears effective across all genotypes has likely limited the perceived need for sequencing surveillance and monitoring studies. The lack of sequencing in Africa is likely also driven by a combination of other factors including under reporting of the disease leading to reduced availability of suitable samples, and lack of sequencing infrastructure. In addition, many YFV sequences in South American studies have been obtained from deceased non-human primates, which act as sentinel species for the disease. In Africa non-human primates tolerate YVF better for reasons that are as yet unknown and so mortality is reduced, meaning samples are less readily available for analysis. Future consideration should be given as to how these problems can be overcome to increase the sequencing of YFV genomes from Africa.

Sequencing of YFV in non-human primates and vectors

Sequencing of YFV in non-human hosts alongside sequencing of human cases is required in order to obtain both the most accurate epidemiological data in terms of virus distribution and spread, and allow transmission chains to be characterised as sylvatic, rural or urban [216]. Integrating sequencing alongside standard surveillance measures during epizootics could also provide information about which areas the virus is likely to spread into and at what rate, allowing measures to be taken to prevent human outbreaks in the first place [209, 217]. In addition increased sequencing of mosquito vectors will provide further insight into which species of mosquito are transmitting which genotypes of the virus, and provide further insight into transmission chains [218]. Sequencing in this way requires that deceased non-human primates are available to provide samples, which is more challenging to achieve in Africa than in South America. Therefore this strategy may be less relevant to African outbreaks, unless a solution to humanely and safely sample non-human primates can be found.

More research into viral biology

Little is currently known about the impact of the viral RNA variants on YFV physiology, despite reports that different virus genotypes can influence virulence and replication ability [177]. There is a need for research into this area, in order to better understand virus biology which in turn could lead to an understanding of whether certain strains are more pathogenic than others. In addition a better understanding of the viral biology could help inform future design of vaccines with fewer side effects, as well as new treatments specific to yellow fever.

Informing diagnostic tests

The use of PCR based molecular diagnostic tests for yellow fever is critical for disease management, yet due to the limited amount of sequence data available many of the existing tests may no longer be up to date or validated on specific lineages of the virus. This is applicable despite the slow evolutionary rate of YFV, and may explain the poor performance of many testing laboratories [219]. In addition new types of rapid molecular diagnostic test in

development require accurate and up to date genetic information to inform their design. Therefore more large scale and representative sequencing of the different YFV genotypes is needed to provide the data needed to assess current diagnostics and develop new ones.

5.6 Yellow fever conclusions

There is a global effort to eliminate yellow fever epidemics by 2026, as outlined in the EYE strategy [188]. For this project to achieve its goals it will be essential to build up sufficient immunity via vaccination against yellow fever in the at-risk human population. This in turn requires accurate risk prediction so that the current (and potentially insufficient) vaccine supply can be used as effectively as possible, whilst also reducing the risk of yellow fever by other disease control methods where possible. Use of genome sequencing as a routine surveillance tool could help towards the goals of the strategy and will be most effective when combined with other disease control measures. By collaborating and consulting with others involved in yellow fever control and management, sequencing efforts could be better coordinated and aligned with other goals, whilst methods of best practice can be more rapidly shared.

Currently, aside from its specific use in distinguishing between wild type and vaccine virus strains during potential adverse events, YFV sequencing is performed relatively sporadically and in the context of research studies, rather than from the perspective of answering public health questions. The most immediate need to enable YFV sequencing to be a successful public health tool is to increase awareness of the impact that sequencing can have, so that sequencing is more widely and routinely used by public health authorities. This will require broader challenges to be overcome in terms of the resources and infrastructure needed to allow sequencing to become a routine tool, as well as the need for researchers to more closely communicate with clinicians and public health officials; these are not unique to yellow fever and are discussed further in Chapter 6. There also needs to be a focus on producing whole genome sequences where possible, as these have the most utility for the most applications, from enabling epidemiological analysis to the design of new diagnostic tests.

6 CROSS-CUTTING THEMES

Sequencing, and primarily NGS combined with bioinformatic analysis, has become a powerful tool in the detection, identification and analysis of human pathogens. Given ongoing improvements in sequencing technologies across a range of platforms and ongoing decreases in costs, it is becoming feasible to apply proof-of-principle studies that demonstrate the potential of sequencing on a larger scale. This is occurring in parallel with an ever increasing demand for and improvements in sequencing.

Many challenges exist in ensuring high quality NGS. While some guidelines and principles already exist for implementing NGS for public health [92-94, 220-222] they are not always broadly applicable [100]. The WHO, as part of its research and development blueprint, articulated the valuable role pathogen genetic sequence data has as a source of information in helping to understand and control outbreaks of infectious disease, and the importance of quickly releasing this data into the public domain [223]. Sequencing is powerful and complex, requiring a range of appropriate experience and expertise to manage these complexities and the approaches of the methodologies – including nucleic acid extraction, understanding sequencing platform capabilities and the myriad of bioinformatics tools available. Rapid growth and constant change, particularly in bioinformatics methods, can make sequencing difficult to standardise [100].

However, as use of sequencing is expanding across a range of pathogens and applications, the boundaries between disease-specific silos are diminishing and provision of sequencing is being considered – particularly for viral diseases – as a service that can meet the needs of a range of diseases and applications, rather than for a specific purpose in a particular disease.

The sequencing landscape, particularly on the African continent, is also changing. On 12th October 2020 the African Union Commission and the Africa Centres for Disease Control and Prevention (Africa CDC) announced the launch of the Africa Pathogen Genomics Initiative (PGI). This four-year partnership aims to expand access to next generation genomic sequencing tools and expertise across the continent. This includes strengthening laboratory networks and public health surveillance, including disease prevention and control efforts. Partners in this US$100 million public-private non-profit initiative include the Bill and Melinda Gates Foundation, US CDC, Microsoft, Illumina, and Oxford Nanopore Technologies. The Africa PGI aims to address many of the sequencing needs outlined in this report, in a pathogen-agnostic fashion. As it develops, the initiative will play a pivotal role in the development of flexible and adaptable sequencing facilities and expertise, and their delivery [224].

Key applications where sequencing will have utility in the three VHFs highlighted in this report are:

Surveillance

- Ongoing surveillance strategies to ensure that potential new outbreaks are detected as fast as possible, especially when a patient presents with VHF symptoms but is not easily diagnosed.
- To inform epidemiological investigations to understand local/national/ international virus spread and to put in place appropriate control and public health measures.
- To monitor for the emergence of new strains of viruses that could affect their spread, clinical symptoms – including response to treatment – and ability to evade immunity.

Diagnostics development and evaluation

- To ensure that diagnostic tests being deployed are effective.
- To use the latest genetic information to update these tests as appropriate, for example primers used in PCR tests. This is most relevant to highly diverse and fast evolving viruses, such as LASV and ebolavirus, but is applicable to all VHFs.

Multiple technologies and assays aside from PCR are now being developed that are based on the genome sequence of the pathogen. These include tests such as CRISPR SHERLOCK, where target-specific CRISPR machinery is designed to cut the nucleic acid of the pathogen in specific sequences (table 3). Diagnostics based on viral proteins can also be informed by viral genome data, should genome variants cause changes to proteins that will impact the effectiveness of assays used to detect them.

Identification of unknown infection using metagenomic (m)NGS

Many febrile diseases have similar clinical presentations and as such it can be difficult to determine which single pathogen assay to use for a diagnosis, resulting in sequential testing until a positive result is found. Delays in diagnosis result in delayed or inadequate treatment, prolonged stays, readmissions, and increased mortality and morbidity. With the use of mNGS at presentation it can be possible to determine which infection is causing disease, and any co-infections present. Another advantage of mNGS is that it can provide genomic information necessary for evolutionary tracing, strain identification, virulence determination and prediction of drug resistance. However, it is important to interpret the identification of a microorganism in the clinical context when determining if it is the cause of the illness.

Universal standard references for mNGS as a diagnostic tool have yet to be established, although the FDA-ARGOS database is supporting some research efforts in this area [225]. In addition, test validation, demonstration of reproducibility and quality assurance are required before mNGS can be considered for routine clinical implementation in patient care settings; only a few clinical tests are available for infectious disease [8]. Evidence demonstrating the potential utility of mNGS is building, such as the case studies on examples in Lassa fever [26] and yellow fever [201], which demonstrate proof of concept.

There are a range of technical challenges involved in the implementation of mNGS approaches. These include: optimisation of pathogen titres so the pathogens' genetic signal is not overwhelmed by the host's DNA/RNA; contamination of samples; barcode index switching; computational power required to manage the large volumes of data generated; availability and ongoing updating of bioinformatics pipelines; misaligned sequences; bias in reference database representation; and clinical interpretation of the findings within the clinical context to determine the cause of disease. In certain circumstances it is possible that a pan-pathogen diagnostic test may be better suited and will be available for a fraction of the cost, speed and effort required to carry out mNGS. In particular pan-pathogen PCR tests could cover most of the common pathogens suspected of causing febrile disease in a particular region. mNGS may therefore have most utility in settings where it may be difficult to predict the likely pathogens causing an illness, and/or the pathogens themselves may be very rarely encountered so it may not be feasible to stock the appropriate diagnostic tests. This situation may occur when having to diagnose imported infections in travellers who have visited one or more regions where a range of febrile diseases are endemic [151].

Treatments and vaccines

Genomic information can provide the basis for new drug or vaccine development. In addition, treatments and vaccine efficacy can also be assessed by monitoring changes to the viral genome. Genomic data can also be useful to assess the impact of treatments on viral evolution, for example, whether the pathogen develops mutations that lead to drug resistance in response to treatment.

Impact of pathogen biology and disease transmission

Disease biology will also have an impact on the use of sequencing. The VHFs discussed here vary in their biology, not only in terms of genome mutation rate, but also in terms of how they are transmitted – human to human (ebolavirus) following an initial spill-over event from an animal host, or human to human and multiple animal host to human transmissions (Lassa fever, yellow fever). Therefore, more extensive sequencing efforts may need to be deployed in viruses which have a higher genetic diversity e.g. Lassa fever, in order to keep diagnostics up to date. For diseases with predominantly human to human transmission (e.g. EVD), sequencing can support surveillance and contact tracing efforts. For pathogens transmitted via animal hosts, consideration needs to be given to how much sequencing effort should be deployed to understand the pathogen circulating in these hosts. For example, sequencing could support understanding of spill-over events from wild animal hosts, and how the disease is evolving in these hosts. The unpredictability of EVD emergence in the human population and the uncertainty surrounding the exact animal reservoirs can make implementation of disease control measures based on sequencing data more challenging.

Below we outline general considerations to support the use of sequencing.

6.1 General considerations to support use of NGS

Personnel. Consideration of the skills of the personnel performing the test are required, including a broad skill set in genomic techniques. The laboratory work involving sample inactivation, processing and preparation, library preparation and sequencing, although requiring specialist training, can be relatively easily taught to laboratory workers. The bioinformatic analysis and subsequent interpretation of the results requires more expertise in a rapidly evolving area. The availability of easy to use and/or automated analysis software is limited and still requires the user to have specialist knowledge and understanding of the fundamental aspects of genomic data analysis. A workforce that has an adaptable and responsive skill set will be valuable not only in rapidly changing outbreak situations, but also in routine diagnostic settings. A bottom up approach is required where training opportunities including undergraduate and postgraduate qualifications exist and are funded in areas such as bioinformatics, epidemiology and computational expertise. This will enable countries under recurrent outbreak threat to possess the skills necessary to carry out sequencing and the essential analysis and interpretation, as well as maintaining software and bioinformatics pipelines as required.

Another consideration is that currently there is limited access to engineers that would usually provide maintenance and repair for sequencers, particularly larger devices. Another area of need is in-country bioinformatics and computational infrastructure experts. One key element of the Africa PGI is plans to train in-country experts and technicians in these areas.

Resources and infrastructure. Access to the appropriate reagents is necessary to carry out sequencing effectively. It has been acknowledged that access to sequencing technologies, reagents, diagnostics and vaccines is an issue in LMICs [226]. Boosting local production of these elements, as well as sufficient availability of resources such as reagents that can be imported, is vital [227]. In addition to the computational infrastructure required to support sequencing efforts, access to online or cloud-based systems may be required in order to analyse sequencing data. This will require reliable internet connectivity. Microsoft is one of the partners in the Africa PGI and will be providing computational and bioinformatics support to the initiative via its Azure cloud platform, as well as training of data scientists.

Centres of Excellence. Establishment of regional and national networks with centres of excellence in pathogen genomics, which have wide-ranging laboratory skills in a range of sequencing technologies and applications, will be valuable. These capabilities can be supported by the operation of mobile laboratories that can be dispatched to areas with potential outbreaks, and provide sequencing as required to meet outbreak management needs.

Collaboration. Sequencing is complex and still a specialist field. With the ever-increasing number of platforms, pipelines and methodologies available, collaboration between centres, institutes, consortia, private sector, academia and public health are invaluable in ensuring optimal use of the technology. This includes exchange of knowledge, ideas, protocols, samples as well as aligning goals for carrying out sequencing – which can include surveillance, vaccine development or monitoring of diagnostics. Collaboration, including with public health bodies, is a key element of the Africa PGI and will play a key role in ensuring that sequencing data can be used optimally for public health impact. This will include engagement with public health officials to support their information gathering and decision making around the use of genomics for public health interventions. Pathogen-agnostic research funding to support sequencing investigations will also have a role to play.

Data collection, integration and sharing. Generating genomic sequencing data in isolation has limited use. Genomic data required for epidemiological investigations must be collected alongside the relevant metadata – such as when and where a sample was taken and whether from a human or animal host. This will require guidance on what data should be recorded and how, as well as systems to safely store the data. Data quality standards will be required both for genomic and metadata to ensure data is suitable for use in its intended application.

Most applications of sequencing require access to other sources of high quality sequencing data from online databases and repositories to support integration and interpretation. Whilst there are guidelines for users when uploading sequencing data to these platforms, there is a lack of consistency in reporting the bioinformatic pipelines used. This can make ongoing data analysis challenging. Development of quality-controlled curated reference databases for sequencing pathogens is required to ensure the availability of high quality sequencing data in the public domain. The FDA-ARGOS database is such an example and has produced regulatory grade reference genomes for multiple pathogens endemic to Africa including the three VHFs in this report [225].

To support the effective use of sequencing data, the development of clear guidelines and frameworks on ethical data sharing, as well as the maintenance of good working relationships between different stakeholders such as clinicians, epidemiologists and public health officials, is required. Supporting these efforts across international borders will be particularly important and the Africa PGI will play a key role in this regard.

Sample logistics. Depending on the application, the timely collection, analysis and processing of samples for sequencing, including return of results should be considered. For real-time surveillance in particular, mobile laboratories that make use of portable sequencing devices, positioned close to the source of outbreaks will minimise sample transport distances. To reduce the need for additional biosafety equipment in mobile laboratories, viral inactivation at the site of sample collection can be carried out. However, methods to do this need to be evaluated and measures need to be put in place to ensure that these facilities are available at the source of sample collection. Practical considerations include transport of samples from more remote or inaccessible areas and maintenance of the cold chain to ensure sample integrity prior to processing for sequencing.

Integrated surveillance approaches. The presence of a number of similar diseases, such as VHFs, within a region presents an opportunity to combine surveillance strategies. This could allow more rapid identification of new cases of VHFs and support the diagnosis of patients with uncertain symptoms or where multiple infections are suspected. This will require the establishment of metagenomic or unbiased sequencing methodologies or multiplex/pan-pathogen tests and technologies to facilitate detection of multiple pathogens, along with appropriate bioinformatics tools to support effective data analysis, interpretation and reporting to the relevant health and public health authorities that can respond to the information.

6.2 Choice of sequencing technology

Choice of sequencing platform for the applications outlined in this report will depend on a number of factors including:

- Technical characteristics of the platform:
  - Throughput levels, including minimum samples and pooling / barcoding possibility (sample capacity per run and per kit size)
  - Yield and quality per sample
  - Run time and time for analysis
  - Type of sequencing – chemistry and read length considerations
  - Scalability and flexibility
- Established personnel and facilities for trouble shooting, protocols, and pipelines for analysis
- Infrastructure needed to implement the platform – including available technologies, additional technology needed and personnel requirements
- Costs related to equipment, reagents, other consumables, personnel time, analysis, maintenance
- Proposed sequencing application(s), such as ongoing or real-time surveillance, diagnostics support, vaccine development, research, including the number of samples expected for these applications
- Where sequencing is being deployed – e.g. in a mobile or central laboratory

A combination of these factors can help determine which platform is best suited for a particular purpose. However it should be acknowledged that the majority of sequencing platforms are adaptable and can provide sequencing for almost all situations. Optimising the speed, costs and best utilisation of resources is multi-faceted and each platform will have their own advantages and disadvantages depending on the situation in which sequencing is being deployed. In addition the platform technologies, as well as the reagents, are consistently being updated and refined by manufacturers.

Long-read sequencers, for example instruments manufactured by Pacific Biosciences and Oxford Nanopore Technologies (ONT), can have higher sequencing error rates, but this can be overcome by having higher read depth during sequencing. Therefore, read depth must be taken into account when looking at the sequencing data generated, as not all of it will be of adequate enough depth to ensure low error rates. This makes these instruments more suited for low-throughput sequencing, depending on the application.

ONT instruments use a unique sequencing methodology, and some of their sequencers are more portable than many of the other instruments available. Reagent costs are higher than other platforms, though cost of the instrument itself low. The gigabytes generated per run could be considered high for pathogen sequencing but the benefits of a less technical device, shorter runtimes and real time base calling support the use of the instrument in situations where a rapid answer is required.

Sequencing by synthesis, used by Illumina instruments and some Ion Torrent instruments, has a higher output and is cheaper when compared to the other short read platforms that use sequencing by ligation. Sequencing by synthesis gives very high accuracy and throughput. Of the Illumina instruments, some are more suited for sequencing of larger genomes, such as human genomes. The HiSeq instrument offers a low cost per gigabase, which is more cost effective for larger genomes whereas its use to sequence a genome of a few or less than a million bases in size is not as practical. Whilst the cost per gigabyte of the Illumina MiSeq or MiniSeq may be higher of that of a high throughput instrument, the initial cost of the instrument is also lower and combined with other capabilities of the technology are considered better suited for the purpose of sequencing smaller genomes.

The Illumina iSeq, MiSeq and MiniSeq instruments, all benchtop sized devices, do not require a minimum number of samples and have an output of between 0.2 to 15 gigabytes. The Ion Torrent instruments in this category do offer faster sequencing, but this can lead to a higher error rate, therefore the number of reads per run should also be considered when higher accuracy is required as this can overcome some of the errors.

In addition to broader points discussed above, there are more specific process-related costs to consider. Each step in the preparation of the sample for sequencing has cost and time considerations (table 9). Some manufacturers are reporting on the number of touch points – steps in the process that require personnel hands-on time – and are looking at providing so-called 'turn-key' products that require limited expertise to use, in order to reduce the number of touch points. The Ion Torrent Genexus system is an example of such a system.

Automation of fluid handling is another innovation linked to NGS, however as the set up and consumables required for automation have been designed for a large number of samples for plating, e.g. 96 or 386 well plates, it is better suited in high throughput situations. It can significantly speed up sample preparation time.

Table 9: Turnaround time estimates and cost determinants for routine WGS (table adapted from information in [228] and [4])

Process	Estimated time	Turnaround time: determinants	Cost implications: determinants
Sample to laboratory	10 min - 24 hrs	Distance from patient to laboratory, logistics	Location, transportation logistics, cool and cold chains and metadata collection
Culturing samples (if done)	48 hrs - 7 days	BSL availability	Laboratory and staffing costs
Sample / Extraction	1-2 hrs	Choice of kit/ methodology, automation	Kits versus reagents, technician hands-on time versus automation
DNA quality control	0.5 - 1 hrs	Automation, benchtop equipment availability	Laboratory and staffing costs
DNA library preparation (including enrichment or amplification if done)	4-6 hrs	Method, kit used, automation	Kit choice, automation, reagents
Pooling and loading samples	1-1.5 hrs	Automation	Laboratory and staffing costs
Template generation and sequencing	50 hrs	Platform, chemistry, read length, run protocol	Platform, kits, read lengths, number of samples, number of runs and coverage
Initial analysis - Raw data quality control	1-2 hrs	Number of samples, computing power, available pipelines and software	Commercial versus free software, bioinformatician expertise, computer infrastructure
Specific analysis and report generation	4 hrs	Dependent on data, references, analysis methodology and availability of reports with recommendations	Software, bioinformatician expertise

6.3 Cost implications for choice of sequencing technology

The upfront and ongoing costs of sequencing technologies is an important question, however it is a challenging issue to address for a number of reasons:

- Full economic cost-benefit analyses of applying NGS for pathogen genomics in public health has not been done and any such analysis will be country and region specific
- Estimates of the cost of consumables from general laboratory equipment (e.g. pipettes, centrifuges), plastics (e.g. tips, tubes), reagents and computational requirements (e.g. laptops, software) are variable and will depend on the specific needs of the laboratory and the jurisdiction in which it resides. These consumable costs are often higher in LMICs because of higher costs of shipping, customs formalities, and profit margins for local companies and distributors. They also have to accommodate for fluctuating exchange rates to establish costs. It can be expensive and time consuming for LMICs to import the reagents and other equipment required for sequencing.
- Most sequencing platforms rely on kits that contain proprietary or patented reagents, and it may not be possible to find or develop lower cost alternatives. These kits can be categorised as closed or open systems – closed systems require reagents from a specific producer, typically ensuring high quality reagents but creating dependence on a single supplier.
- The availability and cost of appropriate cool or cold chain shipping and storage may significantly affect the ability to maintain the quality of sequencing reagents. The reagents also have a specific shelf life, sometimes hindering stockpiling.
- The current status of sequencing facilities in the country in question – establishing services from scratch will have different cost implications to augmenting existing facilities.

Tables 9 (above) and 10 (below) contain the available sequencing costs in US dollars, and cost determinants, for a range of technologies and include considerations of the relative advantages and limitations of each.

Table 10: Summary of a selection of sequencing instruments with technical specifications and costs. The costs are based on 2018 figures and given in US dollars. Adapted from [221] and [229]. Since the publication of this information in 2018, a number of sequencing platforms have entered the market, including: Ion Torrent Genexus; Oxford Nanopore Technologies Flongle; BGI DNA Nanoball technology. Abbreviations: bp – base pairs; kb – kilobases; Gb – gigabyte(s); na – not available.

Platform	Read length	Yield (Gb)	Run time	Instrument cost (US$)	Annual contract (US$)	Cost per Gb (US$)	Limitations	Advantages
Illumina MiniSeq	50-150bp	1.6-7.5	7-25 hours	50,000	5,000	200-400	High cost per Gb	Low instrument cost, established technology, low error rate, benchtop size
Illumina MiSeq	75-300bp	0.5-1.5	4-56 hours	99,000	14,000	250-2,000	High cost per Gb	Low instrument cost, established technology, low error rate, scalability, benchtop size
Illumina NextSeq	75-150bp	16-120	15-29 hours	250,000	32,000	33-43	High instrument cost	Low cost per Gb, established technology, low error rate, benchtop size
Illumina HiSeq2500	36-125bp	9-500	7 hours to 11 days	690,000	75,000	30-230	High instrument cost, need for deep multiplexing	Low cost per Gb, established technology, low error rate
Illumina HiSeq2000/4000	50-150bp	105-750	1-3.5 days	740,000-900,000	81,000	22-50	High instrument cost, need for deep multiplexing	Low cost per Gb, established technology, low error rate
Illumina HiSeq X	150bp	800-900 per flow cell	<3 days	1,000,000	93,000	7-10	High instrument cost, need for deep multiplexing, limited compatibility, requires large lab space	Low cost per Gb, established technology, low error rate

Platform	Read length	Yield (Gb)	Run time	Instrument cost (US$)	Annual contract (US$)	Cost per Gb (US$)	Limitations	Advantages
Illumina iSeq	150bp	0.3-1.2	9-17.5 hours	19,900	na	na	Read length	Low initial investment, short run time
Illumina Nova Seq (5000/6000)	150bp	2000-6000	16-44 hours	850,000-950,000	na	na	Long run time	Read accuracy
Ion PGM	200-400bp	0.03-2	3.7-23 hours	49,000	5,000-10,000	400-2,000	High cost per Gb, not able to do paired-end sequencing, poor homopolymer performance	Rapid sequencing run, benchtop
Ion Proton	Up to 200bp	up to 10	2-4 hours	224,000	20,000-30,000	80	Not able to do paired-end sequencing, poor homopolymer performance	Low cost per Gb, rapid sequencing run
Ion GeneStudio S5	200-400bp	0.6-8	2.5-6 hours	65,000	9,000-18,000	80-5,000	High cost per Gb, not able to do paired-end sequencing	Rapid sequencing run
Pacific Biosciences RS II	~20kb	~1	4 hours	695,000	84,000	1,000	13% single pass error rate, very high cost per Gb, high instrument cost	Very long read lengths, can sacrifice length for accuracy, rapid run time
Pacific Biosciences Sequel	~20kb	~5	4 hours	350,000	20,000	1,000	13% single pass error rate, very high cost per Gb, high instrument cost	Very long read lengths, can sacrifice length for accuracy, rapid run time

Platform	Read length	Yield (Gb)	Run time	Instrument cost (US$)	Annual contract (US$)	Cost per Gb (US$)	Limitations	Advantages
Oxford Nanopore MK MinION	Up to 200 kb	up to 10	up to 48 hours	1,000	0	100-400	10% single pass error rate, increased indel errors in repeat regions, high cost per Gb	Very low instrument cost, portability
Oxford Nanopore GridION	100,000+	50-100	30 minutes – 48 hours	2,400	na	na	High error rate	Short run time
Oxford Nanopore PromethION	100,000+	480-960	30 minutes – 48 hours	25,000	na	na	High error rate	Short run time

7 CONCLUSIONS

Sequencing has a broad range of potential uses for the management of viral haemorrhagic fevers (VHFs), including routine surveillance, outbreak responses and genomic epidemiology, diagnostics and vaccine development. For each of the pathogens covered in this report – Lassa fever virus, ebolavirus and yellow fever virus – there is still work to be done to better understand their genomic diversity, epidemiology and treatment.

This genetic diversity, and the populations (both animal and human) in which these pathogens circulate, needs to be taken into account when choosing when and where to obtain samples for sequencing. The decision about what to sequence will not just be impacted by research needs, but also by local health and public health system needs, which will vary depending on whether the system is managing an endemic disease or a new outbreak. In addition to the applications listed above, each of the pathogens described in this report have their own specific needs and challenges where sequencing could be deployed: for example, in Lassa fever it is to help understand lineage diversity; in Ebola for monitoring the variants circulating in the human and animal populations; and in yellow fever it is understanding of viral biology.

However, there are also similarities between these pathogens in terms of where outbreaks occur, disease symptoms, and the biology of RNA infections in the human host. These similarities could be exploited in the development of pathogen-agnostic sequencing approaches, for example by considering how diagnostics and surveillance for different diseases could be combined.

In turn, these issues can inform the countries in which these diseases are endemic regarding what they need in terms of sequencing capabilities. There is currently a lack of studies comparing the cost-effectiveness of different pathogen sequencing approaches in different scenarios. As such, the answer to the question of which sequencing technology and approach to employ will depend on a number of interconnecting factors including set-up and running costs and availability of equipment and infrastructure, personnel availability and expertise, and the application(s) of sequencing.

There is currently no single technology or approach to VHF sequencing that is preferred over any other, and one benefit of many available sequencing platforms is that they can be flexibly deployed for a number of purposes. Sequencing technologies are continually being improved and updated, for example through providing shorter turn-around times, reducing hands on time and providing ‘sample to report’ solutions. Developments of this type will support further real-time uses of sequencing in public health disease management.

An advantage of sequencing is that each pathogen genome sequence has many uses beyond the primary purpose for which it was generated. In order for pathogen sequence data to be used to its full potential, good quality sequences are required that have adequate genome coverage and sequencing depth, as well as associated metadata available to put the sequence in context of when and where a sample was collected, and for what purpose. Uploading of these data into international databases will support data sharing efforts however further work and collaboration is needed to support some of the challenges – technical, legal and regulatory – that can impede sharing of pathogen genome data.

Support for the development and implementation of sequencing should therefore focus on the generation of good quality sequences, rather than focussing on a specific technology or approach used to obtain the sequence. The development of agreed sequencing standards, external quality controls, automated data analysis pipelines, and interpretation and reporting guidelines, will support the production of consistent and comparable sequencing data that can be used for a range of applications.

This application-led and pathogen agnostic approach is being explored by the Africa Pathogen Genomics Initiative (Africa PGI), which among its many goals will be focussing on developing

flexible sequencing facilities and training personnel to develop expertise in the use of a range of sequencing technologies and their applications.

Pathogen genomics using NGS provides an opportunity to transform the management of infectious diseases. Africa PGI will have a very significant role to play in the coming years in how sequencing is further established on the continent and in time is expected to lead the way in the provision of sequencing for infectious disease management, including viral haemorrhagic fevers.

8 APPENDIX

8.1 Summary of NGS studies in Lassa fever

The studies included in this table are examples to illustrate use of sequencing in areas of interest.

Surveillance		Reference	Study purpose	Sequencing methodology
1.1	Identifying disease origins	Case Study: Whitmer [22] – see main chapter on Lassa fever, case study 1		
		Yadouleton [24]	To trace the suspected, probable, and confirmed cases to find the origin and possible nosocomial infections. They also describe genomes of LASV from 20 confirmed patients sampled in 2014 and 2016 in Benin.	1. library preparation using the Nextera XT DNA Library Preparation Kit (Illumina), sequencing on an Illumina MiSeq platform 2. library preparation using the NEBNext® fast DNA fragmentation and Ion Torrent kit (New England BioLabs), sequencing on a Personal Genome Machine (PGM) using an Ion 316v2 chip (Life Technologies)
		Safronetz – small mammals [230]	To determine whether Lassa virus was circulating in southern Mali – included sampling a village visited by a UK patient.	Not specified
1.2	Transmission dynamics	Case Study: Kafetzopoulou [26] – see main chapter on Lassa fever, case study 2		
1.3	Examining viral population structure	Andersen [3]	To generate a catalogue, understand the spread of LASV and understand transmission patterns.	cDNA synthesis and Illumina library construction, adapter-ligated and Nextera XT fragmentation libraries and were pooled. Sequenced on the Illumina HiSeq2000, HiSeq2500, or the MiSeq.
		Ehichioya [25]	Phylogenetic analysis (L and S segment)	Not specified
1.4	Tracking disease prevalence - Note: mainly being done via RT-PCR e.g. in children [231] need to have the appropriate RT-PCR in use	in rodents – Marien [89] / Leski [90] (also Safronetz in 1.1)	To investigate the sequence diversity of strains circulating among small rodents. Leski – peridomestic setting in Sierra Leonne. Marien – over 4 years and 6 rural villages in Guinea.	Leski et al – Resequencing pathogen microarray (RPM) (Affymetrix / ThermoFisher) Marien et al – Sanger sequencing of S segment (partial)

Diagnosis and infection management		Reference	Study purpose	Sequencing methodology
2.1	Diagnosis by pathogen identification	Metagenomic / unbiased sequencing approach: Prof Happi from ACEGID, Nigeria, has discussed this in presentations and a Lancet Editorial [232], however research not published yet. Approach is also being explored for yellow fever		
2.2	Test development for diagnosis	Wiley [18] – expressed concerns in the current molecular diagnostic assays being used in Liberia (see 3.1)		
		Olschläger [77] / Asogun [65] on site test development for molecular diagnostics.	Improved detection of LASV by RT-PCR	Sanger sequencing of S Segment
2.3	Detection of mixed infections/co-infection	There is limited discussion on the topic of mixed LASV infection – full understanding of genetic diversity still needed. Andersen [3] (see 3.3) briefly discusses this: "Did not identify any recombination events within segments, but did find evidence for reassortment between segments in three samples. This could be explained by infections of individual hosts with multiple LASV lineages, followed by shuffling of segments, a process previously observed in vitro with LASV (Lukashevich, 1992) and in vivo with other arenaviruses (Stenglein 2015)"		
		Metsky [91]	A new computational method to improve sensitivity of metagenomics, allowing more genome data to be recovered without enrichment. Designed, synthesised, and validated multiple probe sets, including one that targets the whole genomes of the 356 viral species known to infect humans. Tested during 2018 Lassa Fever outbreak in Nigeria.	Illumina MiSeq or HiSeq 2500 platforms
2.4	Outbreak management	Case Study: Kafetzopoulou [26] – see main chapter on Lassa fever, case study 2		
		Siddle [85]	To assess whether genetic changes in the viral populations circulating in Nigeria might have contributed to the increase in cases. Contributed to public health responses [233]	Library preparation with Nextera XT (Illumina). All sequenced with an Illumina MiSeq, HiSeq 2500 or NovaSeq machine.

Research and development		Reference	Study purpose	Sequencing methodology
3.1	Development of novel diagnostics, therapeutics, and vaccines	All from 1.1 and Vaccine development discussion		
		Wiley [18]	Substantially increase the available pool of data to help foster the generation of targeted diagnostics and therapeutics.	Enrichment based methodology, sequenced on a MiSeq (Illumina) at the Liberian Institute for Biomedical Research. Archival isolates were sequenced using the TruSeq Stranded mRNA kit on an MiniSeq (Illumina) at the US CDC Viral Special Pathogens branch in Atlanta, GA.
		Beitzel [234] – reverse genetics	Facilitate acquisition of hard-to-obtain LASV and enable comprehensive development of medical countermeasures.	Require good sequence information that is representative of the divergent LASV lineages to develop these.
3.2	Characterisation of virulence factors and resistance markers	There is a limited number of relevant studies covering this area published in the literature.		
		Forni [101]	Examined clinical outcomes of cases with lineage IIg, the most frequently represented – 220 cases - that have outcomes recorded in public databases	Data from GenBank
3.3	Improving understanding of virus biology	Andersen [3]	To generate a catalogue of sequences, understand the spread of LASV and understand transmission. *Important paper that pushed the field forward and was the largest of its kind at the time*.	cDNA synthesis and Illumina library construction, adapter-ligated and Nextera XT fragmentation libraries. Sequenced on the Illumina HiSeq2000, HiSeq2500, or the MiSeq.
		As new lineages are still being identified, there is a clear indication that more sequencing across the region is needed to obtain a full understanding of LASV diversity present (both genetic and in terms of virus biology)		
3.4	Understanding the link between host genomics and disease severity and susceptibility	Andersen [235]	Provides further evidence for the hypothesis that LASV may have been a driver of natural selection in West African populations where Lassa haemorrhagic fever is endemic.	Used data from the 1000 Genomes and Haplotype Map consortium projects.

8.2 Summary of NGS studies in Ebola

These studies have many overlapping objectives. Whilst the primary purpose of the majority of studies using genomic sequencing for ebolavirus is for surveillance, many decide other applications for the use of the genomic data. The largest studies (by number of genomes), mostly from the 2014-2016 outbreak are detailed in this table, however, the rest are examples the literature highlighting the different applications listed.

Surveillance		Reference	Study purpose	Sequencing methodology
1.1	Identifying disease origins	Saez [236]	Investigating the zoonotic origin of EBOV.	Fragments of the mitochondrial cytochrome *b* gene from bats and soil samples were sequenced on Illumina MiSeq platform.
		Baize [162]	Phylogenetic analysis of new outbreak to determine origin	Sanger sequencing using L gene-specific primers amplified by PCR. Complete genomes sequencing directly from three patients with high viral load.
		Gire [150]	Genomic surveillance to trace origin and transmission within the human population. They also evaluated multiple methodologies and protocols	Nextera library construction and Illumina HiSeq2500 sequencing provided the most complete genome assembly and reliable intrahost single-nucleotide variant identification.
		Park [152]	Observe both host-to-host transmission and recurrent emergence of intra-host genetic variants	Randomly primed cDNA synthesis Nextera XT library construction on Illumina HiSeq or MiSeq
		Carroll [237]	Epidemiological and evolutionary history of the epidemic from March 2014 to January 2015. Also tracing the origin of the virus in each country and time of transmission.	Epicentre ScriptSeq v2 RNA-Seq Library Preparation Kit with 10-15 cycles of amplification and sequenced on Illumina HiSeq2500.
	Of those studies mentioned in the category above, Baize, Gire, Park and Carroll performed phylogenetic analysis to understand disease transmission dynamics and viral population structure.			

Surveillance		Reference	Study purpose	Sequencing methodology
1.2	Transmission dynamics	Arias [1]	Identifying unknown sources and chains of transmission, particularly at the end of an outbreak	RT-PCR amplification using the Thermo Fisher Scientific Ion Ampliseq workflow according to the protocol manufacturer with EBOV specific reagents and the Ion Torrent PGM sequencing platform.
		Ladner [238]	Tracing the early introductions of EBOV to Liberia and the ongoing spread within the population.	Illumina Nextera XT DNA library preparation kit - unbiased amplification method and sequenced on Illumina MiSeq.
		Quick [148]	Bolstering surveillance efforts with real-time genomic surveillance	Amplification with EBOV specific probes and sequenced on MinION.
		Hoenen [146]	Phylogenetic analysis to determine genetic diversity of viral populations and determine likely introductions to different countries.	Amplification by PCR and sequenced on MinION.
1.3	Examining viral population structure	Kugelman [163]	In-depth view of EBOV diversity in Liberia during September 2014–February 2015	Illumina Nextera XT DNA library preparation kit with Illumina MiSeq.
		Simon-Loriere [166]	Understand how the distinct linages of the strain circulated in Guinea. To define each lineage by genetic variation.	Nextera Library preparation kit and sequenced on HiSeq2500.
		Jiang (2017) [239]	Characterising viral linages.	Amplification by PCR and sequenced on an Ion Proton machine.

Diagnosis and infection management		Reference	Study purpose	Sequencing methodology
2.1	Diagnosis by pathogen identification	Jerome [151]	Using metagenomic sequencing to identify viral pathogens in clinical samples from returning febrile travellers to assess its suitability as a diagnostic tool.	Library preparation was either by the Kapa Library Preparation Kit (KK8232 Kapa Biosystems) or the Nextera XT Kit (Illumina FC-131-1024) in combination with the NEBNext® Multiplex Oligos for Illumina or the Nextera® XT Index Kit. Sequenced on an Illumina MiSeq or NextSeq platform.
2.2	Informing medical countermeasures	Tong [164]	Determine the impact of viral genetic diversity on the currently used MCMs.	PCR amplifications were performed with EBOV-specific primer pairs with overlaps and sequenced on BGISEQ-100 (Ion Proton) platform.
		Mbala-Kingebi [154]	Rapid identification of Zaire ebolavirus outbreak to guarantee the utility of anti-EBOV vaccination.	Genomic DNA Sequencing Kit EXP-NBD103/SQK-LSK208 (Oxford Nanopore Technologies). The VP35 region of the viral genome was sequenced on MinION
2.3	Detection of mixed infections/co-infection	Li [173]	Detect EBOV infection and other co-infections.	Random primer attached to a linker adapter produced libraries amplified cDNA using the Nextera XT DNA library preparation kit. Sequenced on Illumina HiSeq 2500.
2.4	Outbreak management	Mate [119]	Investigating unconventional transmission route via sexual transmission.	Enriched the semen sample for EBOV genomic RNA using the TruSeq RNA Access kit (Illumina) with custom capture probes designed against EBOV, along with other modifications and sequenced on Illumina MiSeq.
		Arias [1]	Investigating unconventional transmission route via sexual transmission and breastfeeding.	See above in section 1.2.
2.5	Antimicrobial resistance	Whitmer [240]	Investigate if treatment for EVD selected for mutations for resistance	Libraries created by NEB- Next Ultra kits (NEB). Regions with low coverage were re-sequenced using Makona-specific primers and RT-PCR. Sequenced on Illumina MiSeq.

Research and development		Reference	Study purpose	Sequencing methodology
3.1	Development of novel diagnostics, therapeutics, and vaccines	Tong [164]	Producing genomic data on the viral genetic diversity to help inform development of MCMs.	See above in section 2.2.
		Guedj [241]	Gain new insights into the antiviral efficacy of favipiravir	PCR amplification of 8 overlapping regions of the genome and sequenced on Ion Torrent PGM platform.
3.2	Characterisation of virulence factors and resistance markers	Ladner [238]	Investigating whether there was incremental adaptation of EBOV to human host.	See above in section 1.2
3.3	Improving understanding of virus biology	Park [152]	Monitoring the evolutionary pressures on nonsynonymous mutations over time.	As above in section 1.1
		Carroll [237]	Understanding viral evolution during an outbreak.	As above in section 1.1
		Piorkowski [242]	Development of a non-human primate model for EBOV infection. Sequencing was done to assess mutation of the virus over the course of disease	PCR amplification of 8 overlapping regions of the genome and sequenced on Ion Torrent PGM platform.
3.4	Understanding the link between host genomics and disease severity and susceptibility	Liu [174]	Identify early stage host factors that are associated with severity of disease	Epicentre ScriptSeq v2 RNA-Seq Library Preparation Kit and sequenced on Illumina HiSeq 2500

8.3 Summary of NGS studies in yellow fever

Surveillance		Reference	Study purpose	Sequencing methodology
1.1	Tracking yellow fever virus transmission during an outbreak	Giovanetti [243]	Aimed to better understand the virus's genetic diversity and dynamics during the recent 2016-2019 outbreak in south eastern Brazil.	Direct from clinical samples- Multiplex PCR method sequenced on MinIon (ONT) using previously published protocol by Quick [208]
		Cunha [215]	Performed molecular epidemiological analysis of fatal cases to explore the origin of the Sao Paulo yellow fever epidemic of 2017-2018.	Total RNA-seq enrichment protocol using Illumina NextSeq platform. Libraries were prepared using the TruSeq Stranded Total RNA HT sample prep kit (Illumina)
		Hughes [244]	The Uganda Virus Research Institute collaborated with the US CDC to determine the molecular epidemiology of the virus causing the 2016 Uganda outbreak.	Direct from sample- RNA seq with Ion Torrent
		Faria [209]	To perform genomic and epidemiological monitoring to understand the transmission potential of the yellow fever virus, identifying its origins, rate of spread, and mechanisms of transmission.	Direct from sample- three different approaches used: Overlapping amplicon sequencing was used for the majority of samples based on method by Quick [208] which were mostly sequenced on a MinION (ONT) with some use of Ion Torrent. RNA sequencing was also performed on some samples using Illumina MiSeq.
1.2	Monitoring of non-human primates and mosquito vectors	Goes de Jesus [211]	Analysis of yellow fever transmission in non-human primates, Bahia, North eastern Brazil, to understand the origins of the virus and the effectiveness of prevention strategies.	Direct from clinical samples- Multiplex PCR method using overlapping amplicons sequenced on Minion (ONT) using previously published protocol by Quick et al [208].
		Hill [216]	Genomic surveillance of yellow fever virus epizootic in São Paulo, Brazil, 2016 – 2018. Used to understand origin of virus, rate of spread and mechanisms of transmission among primates, helping to put human cases into context.	Direct from clinical samples- Multiplex PCR method using overlapping amplicons sequenced on Minion (ONT) using previously published protocol by Quick et al [208]

Surveillance		Reference	Study purpose	Sequencing methodology
		Cunha [218]	Performed NGS to identify and characterise YFV in Aedes scapularis, south eastern Brazil, 2016, to help start to disentangle the role of Aedes scapularis in yellow fever transmission in the Americas.	Direct from clinical samples- Multiplex PCR method using overlapping amplicons sequenced on Minion (ONT), cited Faria et al [209]
		Cunha [217]	To use whole YFV genomes from the 2016 and 2017 epizootic events in São Paulo State, to help determine the virus origin.	Viral culture- RNA seq using Illumina MiSeq. Libraries were prepared using the Nextera XT kit (Illumina)
1.3	Identifying the start of a yellow fever outbreak	Ajogbasile [201]	To diagnose unexplained fever and poor outcomes in 50 patients at Irrua Specialist Teaching Hospital (ISTH) in Edo State, Nigeria, after Lassa fever was ruled out.	Direct from clinical samples- Metagenomic sequencing sequenced on the Illumina MiSeq platform. Libraries were prepared using the Nextera XT kit (Illumina)
		McMullan [207]	To diagnose and provide genotype data on samples which had tested negative for several VHFs including Ebola and Marburg virus.	Direct from clinical sample- Metagenomics approach using pyrosequencing on Genome Sequencer FLX. Sanger sequencing used to fill in gaps in sequence data
1.4	Identifying imported cases of yellow fever	Phan [245]	To construct the genomic sequence of YFV from a Dutch traveller returning from the Gambia-Senegal region, to help provide more information on virus origins.	Direct from clinical sample- RNA seq on Ion Torrent, using standard Ion Torrent library prep

Diagnosis		**Reference**	**Study purpose**	**Sequencing methodology**
2.1	Identifying vaccine and wild type strains of YFV	Rezende [198]	To develop a YFV genotypic tool for investigation of suspected adverse events following yellow fever vaccination.	Direct from sample- NS5 amplicons sequenced by Sanger ABI3130 platform (Applied Biosystems)
Research and development		**Reference**	**Study purpose**	**Sequencing methodology**
3.1	Vaccine design and quality control	Tosta [214]	To design a Multi-epitope based vaccine against YFV, and ensure that the predicted epitopes represented current epitopes.	This study didn't produce sequence data, but analysed publicly available sequences.
		Stock [246]	To better understand the mechanisms of pathology of the yellow fever vaccine seed virus.	PCR amplification of vaccine RNA, sequenced with Sanger
		Pestana [213]	To use Sanger sequencing to develop a robust method for the genetic monitoring of YFV vaccines.	PCR amplification of vaccine RNA, sequenced with Sanger
3.3	Improving understanding of virus biology	Phan [200]	To investigate the shedding mechanisms of YFV from an imported case in the Netherlands after travel to Brazil, to demonstrate the utility of different sample types other than plasma for sequencing.	Direct from clinical samples- overlapping amplicons (primers designed by the group) followed by Ion Torrent sequencing
		Chen [247]	To understand the risk of evolution of YFV within hosts compared to at an epidemic scale, as well as investigate the suitability of different sample types.	Direct from sample-RNA sequencing and amplicon sequencing used, both sequenced using Illumina HiSeq

8.4 Acknowledgements

We thank the following people for contributing their time and expertise to project information gathering:

Professor Ian Goodfellow - Deputy Head of Department (Research) Division of Virology, University of Cambridge

Dr Sarah Hill - Research Associate, Royal Veterinary College, London

Dr Joshua Quick - UKRI Future Leaders Fellow, Institute of Microbiology and Infection, University of Birmingham

Dr Sofonias Tessema - Implementation Science and Pathogen Genomics Expert, Africa CDC Institute for Pathogen Genomics, Ethiopia

9 REFERENCES

URLs correct as of 17 May 2021

1. Arias, A., Watson, S. J., Asogun, D., et al. Rapid outbreak sequencing of Ebola virus in Sierra Leone identifies transmission chains linked to sporadic cases. Virus Evol. 2016. 2(1): p. vew016.
2. Giovanetti, M., Faria, N. R., Lourenco, J., et al. Genomic and Epidemiological Surveillance of Zika Virus in the Amazon Region. Cell Rep. 2020. 30(7): pp. 2275-2283 e7.
3. Andersen, K. G., Shapiro, B. J., Matranga, C. B., et al. Clinical Sequencing Uncovers Origins and Evolution of Lassa Virus. Cell. 2015. 162(4): pp. 738-50.
4. Houldcroft, C. J., Beale, M. A., Breuer, J. Clinical and biological insights from viral genome sequencing. Nat Rev Microbiol. 2017. 15(3): pp. 183-192.
5. Armstrong, G. L., MacCannell, D. R., Taylor, J., et al. Pathogen Genomics in Public Health. N Engl J Med. 2019. 381(26): pp. 2569-2580.
6. Rohden, F., Huang, S., Droge, G., et al. Combined study on Digital Sequence Information (DSI) in public and private databases and traceability. Convention on Biological Diversity. 2020; Available from: https://www.cbd.int/doc/c/1f8f/d793/57cb114ca40cb6468f479584/dsi-ahteg-2020-01-04-en.pdf.
7. ViPR. Virus pathogen resource. 2020; Available from: https://www.viprbrc.org/brc/home.spg?decorator=vipr.
8. Chiu, C. Y., Miller, S. A. Clinical metagenomics. Nat Rev Genet. 2019. 20(6): pp. 341-355.
9. Storch, G. A. Diagnostic Virology. Clinical Infectious Diseases. 2000. 31(3): pp. 739-751.
10. Racsa, L. D., Kraft, C. S., Olinger, G. G., et al. Viral Hemorrhagic Fever Diagnostics. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America. 2016. 62(2): pp. 214-219.
11. Plaque Reduction Neutralization Tests (PRNT). 2020; Available from: https://www.labce.com/spg1050613_plaque_reduction_neutralization_tests_prnt.aspx.
12. Koczula, K. M., Gallotta, A. Lateral flow assays. Essays in biochemistry. 2016. 60(1): pp. 111-120.
13. Dao Thi, V. L., Herbst, K., Boerner, K., et al. A colorimetric RT-LAMP assay and LAMP-sequencing for detecting SARS-CoV-2 RNA in clinical samples. Sci Transl Med. 2020. 12(556).
14. Barnes, K. G., Lachenauer, A. E., Nitido, A., et al. Deployable CRISPR-Cas13a diagnostic tools to detect and report Ebola and Lassa virus cases in real-time. Nat Commun. 2020. 11(1): p. 4131.
15. Zhou, L., Chandrasekaran, A. R., Punnoose, J. A., et al. Programmable low-cost DNA-based platform for viral RNA detection. Science Advances. 2020. p. eabc6246.
16. Purushotham, J., Lambe, T., Gilbert, S. C. Vaccine platforms for the prevention of Lassa fever. Immunol Lett. 2019. 215: pp. 1-11.
17. Bowen, M. D., Rollin, P. E., Ksiazek, T. G., et al. Genetic Diversity amongst Lassa Virus Strains. Journal of Virology. 2000. 74(15): p. 12.
18. Wiley, M. R., Fakoli, L., Letizia, A. G., et al. Lassa virus circulating in Liberia: a retrospective genomic characterisation. The Lancet Infectious Diseases. 2019. 19(12): pp. 1371-1378.
19. Manning, J. T., Forrester, N., Paessler, S. Lassa virus isolates from Mali and the Ivory Coast represent an emerging fifth lineage. Frontiers in microbiology. 2015. 6: p. 1037.
20. Safronetz, D., Sogoba, N., Diawara, S. I., et al. Annual Incidence of Lassa Virus Infection in Southern Mali. The American journal of tropical medicine and hygiene. 2017. 96(4): pp. 944-946.
21. Olayemi, A., Cadar, D., Magassouba, N., et al. New Hosts of The Lassa Virus. Sci Rep. 2016. 6: p. 25280.
22. Whitmer, S. L. M., Strecker, T., Cadar, D., et al. New Lineage of Lassa Virus, Togo, 2016. Emerg Infect Dis. 2018. 24(3): pp. 599-602.
23. Happi, A. N., Happi, C. T., Schoepp, R. J. Lassa fever diagnostics: past, present, and future. Curr Opin Virol. 2019. 37: pp. 132-138.
24. Yadouleton, A., Picard, C., Rieger, T., et al. Lassa fever in Benin: description of the 2014 and 2016 epidemics and genetic characterization of a new Lassa virus. Emerg Microbes Infect. 2020. 9(1): pp. 1761-1770.
25. Ehichioya, D. U., Dellicour, S., Pahlmann, M., et al. Phylogeography of Lassa Virus in Nigeria. J Virol. 2019. 93(21).
26. Kafetzopoulou, L. E., Pullan, S. T., Lemey, P., et al. Metagenomic sequencing at the epicenter of the Nigeria 2018 Lassa fever outbreak. Science. 2019. 363(6422): pp. 74-77.
27. Hallam, H. J., Hallam, S., Rodriguez, S. E., et al. Baseline mapping of Lassa fever virology, epidemiology and vaccine research and development. NPJ Vaccines. 2018. 3: p. 11.

28. International Health Regulations (2005). World Health Organization. 2008; Available from: https://www.who.int/ihr/publications/9789241596664/en/.
29. Fall, I. S., Rajatonirina, S., Yahaya, A. A., et al. Integrated Disease Surveillance and Response (IDSR) strategy: current status, challenges and perspectives for the future in Africa. BMJ Glob Health. 2019. 4(4): p. e001427.
30. Sigfrid, L., Moore, C., Salam, A. P., et al. A rapid research needs appraisal methodology to identify evidence gaps to inform clinical research priorities in response to outbreaks-results from the Lassa fever pilot. BMC Med. 2019. 17(1): p. 107.
31. Isere, E. E., Fatiregun, A. A., Ajayi, I. O. An overview of disease surveillance and notification system in Nigeria and the roles of clinicians in disease outbreak prevention and control. Nigerian Medical Journal. 2015. 56(3): pp. 161-168.
32. Sironi, M., de la Torre, J. C. The chameleonic genetics of Lassa virus. The Lancet Infectious Diseases. 2019. 19(12): pp. 1276-1277.
33. Keïta, M., Kizerbo, G. A., Subissi, L., et al. Investigation of a cross-border case of Lassa fever in West Africa. BMC Infectious Diseases. 2019. 19(1): p. 606.
34. Mazzola, L. T., Kelly-Cirino, C. Diagnostics for Lassa fever virus: a genetically diverse pathogen found in low-resource settings. BMJ Glob Health. 2019. 4(Suppl 2): p. e001116.
35. Wolf, T., Ellwanger, R., Goetsch, U., et al. Fifty years of imported Lassa fever: a systematic review of primary and secondary cases. J Travel Med. 2020. 27(4).
36. Ibukun, F. I. Inter-Lineage Variation of Lassa Virus Glycoprotein Epitopes: A Challenge to Lassa Virus Vaccine Development. Viruses. 2020. 12(4).
37. McCormick, J. B., Fisher-Hoch, S.P. 'Ed.'Lassa Fever. Springer-Verlang Berlin Hamburg; 2002.
38. Lassa Fever. Centers for Disease Control and Prevention. 2020; Available from: https://www.cdc.gov/vhf/lassa/index.html.
39. Kayem, N. D., Benson, C., Aye, C. Y. L., et al. Lassa fever in pregnancy: a systematic review and meta-analysis. Trans R Soc Trop Med Hyg. 2020. 114(5): pp. 385-396.
40. Okokhere, P., Colubri, A., Azubike, C., et al. Clinical and laboratory predictors of Lassa fever outcome in a dedicated treatment facility in Nigeria: a retrospective, observational cohort study. Lancet Infect Dis. 2018. 18(6): pp. 684-695.
41. Lo Iacono, G., Cunningham, A., Fichet-Calvet, E., et al. Using Modelling to Disentangle the Relative Contributions of Zoonotic and Anthroponotic Transmission: The Case of Lassa Fever. PLOS Neglected Tropical Diseases. 2015. 9(1): p. e3398.
42. McCormick, J. B., King, I. J., Webb, P. A., et al. Lassa Fever. New England Journal of Medicine. 1986. 314(1): pp. 20-26.
43. National Guidance on Lassa Fever Case Management Nigeria Centre for Disease Control (NCDC). 2018; Available from: https://ncdc.gov.ng/themes/common/docs/protocols/92_1547068532.pdf.
44. Clinical management of patients with viral haemorrhagic fever. A pocket guide for front-line workers. World Health Organization. 2016; Available from: https://apps.who.int/iris/bitstream/handle/10665/205570/9789241549608_eng.pdf.
45. Technical guidance on Lassa fever. World Health Organization. 2020; Available from: https://www.who.int/emergencies/diseases/lassa-fever/technical-guidance/en/.
46. R&D Blueprint and Lassa Fever. World Health Organization. 2018; Available from: https://www.who.int/teams/health-emergencies-preparedness-and-response/blueprint/lassa-fever.
47. Lassa fever research and development roadmap. World Health Organization. 2018; Available from: https://www.who.int/publications/m/item/lassa-fever-research-and-development-(r-d)-roadmap.
48. Bausch, D. G., Hadi, C. M., Khan, S. H., et al. Review of the Literature and Proposed Guidelines for the Use of Oral Ribavirin as Postexposure Prophylaxis for Lassa Fever. Clinical Infectious Diseases. 2010. 51(12): pp. 1435-1441.
49. Lassa fever - Nigeria. Disease outbreak news. . World Health Organization. 2018; Available from: https://www.who.int/csr/don/01-march-2018-lassa-fever-nigeria/en/.
50. Ludwig, G. V. Letter: FOI request. Final report analysis of a clinical trail, ribavirn and the treatment of Lassa fever. Department of the Army, US Army Medical Research and Materiel Command. 2019; Available from: https://media.tghn.org/medialibrary/2019/03/Dr._Ludwig_memo.pdf.
51. Eberhardt, K. A., Mischlinger, J., Jordan, S., et al. Ribavirin for the treatment of Lassa fever: A systematic review and meta-analysis. Int J Infect Dis. 2019. 87: pp. 15-20.
52. Clinical Trials - Lassa Fever. National Institutes of Health. 2020; Available from: https://clinicaltrials.gov/ct2/results?cond=lassa&term=&cntry=&state=&city=&dist.

53. Raabe, V. N., Kann, G., Ribner, B. S., et al. Favipiravir and Ribavirin Treatment of Epidemiologically Linked Cases of Lassa Fever. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America. 2017. 65(5): pp. 855-859.
54. Warner, B. M., Siragam, V., Stein, D. R. Assessment of antiviral therapeutics in animal models of Lassa fever. Curr Opin Virol. 2019. 37: pp. 84-90.
55. Cross, R. W., Mire, C. E., Branco, L. M., et al. Treatment of Lassa virus infection in outbred guinea pigs with first-in-class human monoclonal antibodies. Antiviral research. 2016. 133: pp. 218-222.
56. Mire, C. E., Cross, R. W., Geisbert, J. B., et al. Human-monoclonal-antibody therapy protects nonhuman primates against advanced Lassa fever. Nature Medicine. 2017. 23(10): pp. 1146-1149.
57. Efficacy trials of Lassa Therapeutics: endpoints, trial design, site selection; WHO Workshop. INSERM, Paris, France. 2018; Available from: https://lassafever.info/sites/default/files/2019-04/LassaTxeval_FinalmeetingReport.pdf.
58. Hall, W., Jamieson, A., Wardle, G. Advancing Epidemics R&D to keep up with a changing world: progress challenges and opportunities. Wellcome Trust. 2019; Available from: https://wellcome.org/sites/default/files/advancing-epidemics-rd-2019.pdf.
59. Salami, K., Gouglas, D., Schmaljohn, C., et al. A review of Lassa fever vaccine candidates. Curr Opin Virol. 2019. 37: pp. 105-111.
60. Lukashevich, I. S., Paessler, S., de la Torre, J. C. Lassa virus diversity and feasibility for universal prophylactic vaccine. F1000Research. 2019. 8: pp. F1000 Faculty Rev-134.
61. Boisen, M. L., Uyigue, E., Aiyepada, J., et al. Field evaluation of a Pan-Lassa rapid diagnostic test during the 2018 Nigerian Lassa fever outbreak. Sci Rep. 2020. 10(1): p. 8724.
62. Shaffer, J. G., Grant, D. S., Schieffelin, J. S., et al. Lassa Fever in Post-Conflict Sierra Leone. PLOS Neglected Tropical Diseases. 2014. 8(3): p. e2748.
63. Boisen, M. L., Hartnett, J. N., Shaffer, J. G., et al. Field validation of recombinant antigen immunoassays for diagnosis of Lassa fever. Sci Rep. 2018. 8(1): p. 5939.
64. Drosten, C., Panning, M., Guenther, S., et al. False-negative results of PCR assay with plasma of patients with severe viral hemorrhagic fever. Journal of clinical microbiology. 2002. 40(11): pp. 4394-4395.
65. Asogun, D. A., Adomeh, D. I., Ehimuan, J., et al. Molecular Diagnostics for Lassa Fever at Irrua Specialist Teaching Hospital, Nigeria: Lessons Learnt from Two Years of Laboratory Operation. PLOS Neglected Tropical Diseases. 2012. 6(9): p. e1839.
66. Emperador, D. M., Yimer, S. A., Mazzola, L. T., et al. Diagnostic applications for Lassa fever in limited-resource settings. BMJ Glob Health. 2019. 4(Suppl 2): p. e001119.
67. Takah, N. F., Brangel, P., Shrestha, P., et al. Sensitivity and specificity of diagnostic tests for Lassa fever: a systematic review. BMC Infect Dis. 2019. 19(1): p. 647.
68. Sharif, A. A., Sharfadi, S. R., Aminu, I., et al. Lassa Virus Diagnostic Platforms: Limitations and Prospects. Asian Journal of Research in Infectious Diseases. 2020. pp. 5-17.
69. Raabe, V., Koehler, J. . Laboratory diagnosis of Lassa fever. Journal of Clinical Microbiology. 2017. 55(6): p. 9.
70. Grove, J. N., Branco, L. M., Boisen, M. L., et al. Capacity building permitting comprehensive monitoring of a severe case of Lassa hemorrhagic fever in Sierra Leone with a positive outcome: case report. Virology journal. 2011. 8: pp. 314-314.
71. Lunkenheimer, K., Hufert, F. T., Schmitz, H. Detection of Lassa virus RNA in specimens from patients with Lassa fever by using the polymerase chain reaction. Journal of clinical microbiology. 1990. 28(12): pp. 2689-2692.
72. Towner, J. S., Sealy, T. K., Ksiazek, T. G., et al. High-throughput molecular detection of hemorrhagic fever virus threats with applications for outbreak settings. J Infect Dis. 2007. 196 Suppl 2: pp. S205-12.
73. Kouadio, L., Nowak, K., Akoua-Koffi, C., et al. Lassa Virus in Multimammate Rats, Côte d'Ivoire, 2013. Emerging infectious diseases. 2015. 21(8): pp. 1481-1483.
74. Demby, A. H., Chamberlain, J., Brown, D. W., et al. Early diagnosis of Lassa fever by reverse transcription-PCR. Journal of clinical microbiology. 1994. 32(12): pp. 2898-2903.
75. Panning, M., Emmerich, P., Ölschläger, S., et al. Laboratory Diagnosis of Lassa Fever, Liberia. Emerging Infectious Disease journal. 2010. 16(6): p. 1041.
76. Vieth, S., Drosten, C., Lenz, O., et al. RT-PCR assay for detection of Lassa virus and related Old World arenaviruses targeting the L gene. Trans R Soc Trop Med Hyg. 2007. 101(12): pp. 1253-64.
77. Olschläger, S., Lelke, M., Emmerich, P., et al. Improved detection of Lassa virus by reverse transcription-PCR targeting the 5' region of S RNA. Journal of Clinical Microbiology. 2010. 48(6): pp. 2009-2013.

78. Olschläger, S., Günther, S. Rapid and specific detection of Lassa virus by reverse transcription-PCR coupled with oligonucleotide array hybridization. Journal of clinical microbiology. 2012. 50(7): pp. 2496-2499.
79. Leski, T. A., Lin, B., Malanoski, A. P., et al. Testing and validation of high density resequencing microarray for broad range biothreat agents detection. PLoS One. 2009. 4(8): p. e6569.
80. Fukuma, A., Kurosaki, Y., Morikawa, Y., et al. Rapid detection of Lassa virus by reverse transcription-loop-mediated isothermal amplification. Microbiology and immunology. 2011. 55(1): pp. 44-50.
81. Pemba, C. M., Kurosaki, Y., Yoshikawa, R., et al. Development of an RT-LAMP assay for the detection of Lassa viruses in southeast and south-central Nigeria. J Virol Methods. 2019. 269: pp. 30-37.
82. Jones, A., Boisen, M., Radtkey, R., et al. Development of a handheld multiplex point of care diagnostic for differentiation of Lassa fever, Dengue fever and Ebola Hemorrhagic Fever. American Association for Clinical Chemistry Annual Meeting and Clinical Lab Expo 2013, Houston,TX, July 28 - August 1, 2013. nanomix. 2013; Available from: https://nano.com/downloads/Nanomix_AACC_poster_Abstract.pdf.
83. Whole genome sequencing protocols. Centers for Disease Control and Prevention. 2016; Available from: https://www.cdc.gov/pulsenet/pathogens/wgs.html.
84. Protocols. Andersen lab, The Scripps Research Institute. 2020; Available from: https://andersen-lab.com/secrets/protocols/.
85. Siddle, K. J., Eromon, P., Barnes, K. G., et al. Genomic Analysis of Lassa Virus during an Increase in Cases in Nigeria in 2018. N Engl J Med. 2018. 379(18): pp. 1745-1753.
86. Auperin, D. D., McCormick, J. B. Nucleotide sequence of the Lassa virus (Josiah strain) S genome RNA and amino acid sequence comparison of the N and GPC proteins to other arenaviruses. Virology. 1989. 168(2): pp. 421-5.
87. Auperin, D. D., Sasso, D. R., McCormick, J. B. Nucleotide sequence of the glycoprotein gene and intergenic region of the Lassa virus S genome RNA. Virology. 1986. 154(1): pp. 155-67.
88. Lassa Fever sequences. Virus Pathogen Resource (ViPR). 2020; Available from: https://www.viprbrc.org/brc/home.spg?decorator=arena_lassa.
89. Marien, J., Lo Iacono, G., Rieger, T., et al. Households as hotspots of Lassa fever? Assessing the spatial distribution of Lassa virus-infected rodents in rural villages of Guinea. Emerg Microbes Infect. 2020. 9(1): pp. 1055-1064.
90. Leski, T. A., Stockelman, M. G., Moses, L. M., et al. Sequence variability and geographic distribution of Lassa virus, Sierra Leone. Emerg Infect Dis. 2015. 21(4): pp. 609-18.
91. Metsky, H. C., Siddle, K. J., Gladden-Young, A., et al. Capturing sequence diversity in metagenomes with comprehensive and scalable probe design. Nat Biotechnol. 2019. 37(2): pp. 160-168.
92. Kozyreva, V. K., Truong, C.-L., Greninger, A. L., et al. Validation and Implementation of Clinical Laboratory Improvements Act-Compliant Whole-Genome Sequencing in the Public Health Microbiology Laboratory. Journal of clinical microbiology. 2017. 55(8): pp. 2502-2520.
93. Chattaway, M. A., Dallman, T. J., Larkin, L., et al. The Transformation of Reference Microbiology Methods and Surveillance for Salmonella With the Use of Whole Genome Sequencing in England and Wales. Frontiers in Public Health. 2019. 7(317).
94. Struelens, M., Albiger, B., Catchpole, M., et al. Expert opinion on whole genome sequencing for public health surveillance. European Centre for Disease Prevention and Control (ECDC). 2016; Available from:
95. Gardy, J. L., Loman, N. J. Towards a genomics-informed, real-time, global pathogen surveillance system. Nature Reviews Genetics. 2018. 19(1): pp. 9-20.
96. Patassi, A. A., Landoh, D. E., Mebiny-Essoh Tchalla, A., et al. Emergence of Lassa Fever Disease in Northern Togo: Report of Two Cases in Oti District in 2016. Case Rep Infect Dis. 2017. 2017: p. 8242313.
97. Ihekweazu, C. Early Results of Lassa Virus Sequencing & Implications for Current Outbreak Response in Nigeria. Nigeria Centre for Disease Control 2018; Available from: https://ncdc.gov.ng/news/121/early-results-of-lassa-virus-sequencing-%26-implications-for-current-outbreak-response-in-nigeria.
98. Bhadelia, N. Understanding Lassa fever. Science. 2019. 363(6422): p. 30.
99. Holmes, E. C., Rambaut, A., Andersen, K. G. Pandemics: spend on surveillance, not prediction. Nature. 2018. 558(7709): pp. 180-182.
100. Maljkovic Berry, I., Melendrez, M. C., Bishop-Lilly, K. A., et al. Next Generation Sequencing and Bioinformatics Methodologies for Infectious Disease Research and Public Health: Approaches,

Applications, and Considerations for Development of Laboratory Capacity. J Infect Dis. 2020. 221(Supplement_3): pp. S292-S307.
101. Forni, D., Sironi, M. Population structure of Lassa Mammarenavirus in West Africa. Viruses. 2020. 12(4).
102. Muhlberger, E. Filovirus replication and transcription. Future Virol. 2007. 2(2): pp. 205-215.
103. Kuhn, J. H., Andersen, K. G., Baize, S., et al. Nomenclature- and database-compatible names for the two Ebola virus variants that emerged in Guinea and the Democratic Republic of the Congo in 2014. Viruses. 2014. 6(11): pp. 4760-99.
104. ICTV 9th Report (2011). International Committee on Taxonomy of Viruses. 2020; Available from: https://talk.ictvonline.org/ictv-reports/ictv_9th_report/negative-sense-rna-viruses-2011/w/negrna_viruses/197/filoviridae.
105. 2014: CDC laboratories produce first genomic sequence of Liberian Ebola. Centers for Disease Control and Prevention. 2014; Available from: https://www.cdc.gov/amd/whats-new/ebola.html.
106. Holmes, E. C., Dudas, G., Rambaut, A., et al. The evolution of Ebola virus: Insights from the 2013-2016 epidemic. Nature. 2016. 538(7624): pp. 193-200.
107. Jun, S. R., Leuze, M. R., Nookaew, I., et al. Ebolavirus comparative genomics. FEMS Microbiol Rev. 2015. 39(5): pp. 764-78.
108. Garske, T., Cori, A., Ariyarajah, A., et al. Heterogeneities in the case fatality ratio in the West African Ebola outbreak 2013-2016. Philos Trans R Soc Lond B Biol Sci. 2017. 372(1721).
109. Ebola virus disease – Democratic Republic of the Congo. World Health Organization. 2020; Available from: https://www.who.int/csr/don/18-november-2020-ebola-drc/en/.
110. Leroy, E. M., Gonzalez, J. P., Baize, S. Ebola and Marburg haemorrhagic fever viruses: major scientific advances, but a relatively minor public health threat for Africa. Clin Microbiol Infect. 2011. 17(7): pp. 964-76.
111. Roddy, P., Howard, N., Van Kerkhove, M. D., et al. Clinical manifestations and case management of Ebola haemorrhagic fever caused by a newly identified virus strain, Bundibugyo, Uganda, 2007-2008. PLoS One. 2012. 7(12): p. e52986.
112. Ebola virus infection: Aetiology. British Medical Journal. 2020; Available from: https://bestpractice.bmj.com/topics/en-gb/1210/aetiology.
113. Ebola Virus Disease Distribution Map: Cases of Ebola Virus Disease in Africa Since 1976. Centers for Disease Control and Prevention. 2019; Available from: https://www.cdc.gov/vhf/ebola/history/distribution-map.html.
114. Bower, H., Glynn, J. R. A systematic review and meta-analysis of seroprevalence surveys of ebolavirus infection. Sci Data. 2017. 4: p. 160133.
115. Diallo, M. S. K., Rabilloud, M., Ayouba, A., et al. Prevalence of infection among asymptomatic and paucisymptomatic contact persons exposed to Ebola virus in Guinea: a retrospective, cross-sectional observational study. Lancet Infect Dis. 2019. 19(3): pp. 308-316.
116. Forrester, J. V. Ebola virus and persistent chronic infection: when does replication cease? Ann Transl Med. 2018. 6(Suppl 1): p. S39.
117. Ebola virus disease. World Health Organization. 2020; Available from: https://www.who.int/news-room/fact-sheets/detail/ebola-virus-disease.
118. Barnes, K. G., Kindrachuk, J., Lin, A. E., et al. Evidence of Ebola Virus Replication and High Concentration in Semen of a Patient During Recovery. Clin Infect Dis. 2017. 65(8): pp. 1400-1403.
119. Mate, S. E., Kugelman, J. R., Nyenswah, T. G., et al. Molecular Evidence of Sexual Transmission of Ebola Virus. N Engl J Med. 2015. 373(25): pp. 2448-54.
120. Diallo, B., Sissoko, D., Loman, N. J., et al. Resurgence of Ebola Virus Disease in Guinea Linked to a Survivor With Virus Persistence in Seminal Fluid for More Than 500 Days. Clin Infect Dis. 2016. 63(10): pp. 1353-1356.
121. Dean, N. E., Halloran, M. E., Yang, Y., et al. Transmissibility and Pathogenicity of Ebola Virus: A Systematic Review and Meta-analysis of Household Secondary Attack Rate and Asymptomatic Infection. Clin Infect Dis. 2016. 62(10): pp. 1277-1286.
122. Ebola virus disease: Fact sheet. World Health Organization. 2021; Available from: https://www.who.int/news-room/fact-sheets/detail/ebola-virus-disease.
123. Mupapa, K., Mukundu, W., Bwaka, M. A., et al. Ebola hemorrhagic fever and pregnancy. J Infect Dis. 1999. 179 Suppl 1: pp. S11-2.
124. Wong, J. Y., Zhang, W., Kargbo, D., et al. Assessment of the severity of Ebola virus disease in Sierra Leone in 2014-2015. Epidemiol Infect. 2016. 144(7): pp. 1473-81.
125. Ebola response roadmap. World Health Organization. 2014; Available from: https://www.who.int/csr/resources/publications/ebola/response-roadmap/en/.

126. Target product profile: Ebola virus vaccine. World Health Organization. 2016; Available from: https://www.who.int/immunization/research/target-product-profile/ebolavaccine/en/
127. Ebola treatment and care. World Health Organization. 2020; Available from: https://www.who.int/csr/disease/ebola/training/treatment/en/.
128. Mulangu, S., Dodd, L. E., Davey, R. T., Jr., et al. A Randomized, Controlled Trial of Ebola Virus Disease Therapeutics. N Engl J Med. 2019. 381(24): pp. 2293-2303.
129. FDA Approves First Treatment for Ebola Virus. US Food and Drug Administration. 2020; Available from: https://www.fda.gov/news-events/press-announcements/fda-approves-first-treatment-ebola-virus.
130. Monath, T. P., Fast, P. E., Modjarrad, K., et al. rVSVDeltaG-ZEBOV-GP (also designated V920) recombinant vesicular stomatitis virus pseudotyped with Ebola Zaire Glycoprotein: Standardized template with key considerations for a risk/benefit assessment. Vaccine X. 2019. 1: p. 100009.
131. Four countries in the African region license vaccine in milestone for Ebola prevention. World Health Organization. 2020; Available from: https://www.who.int/news-room/detail/14-02-2020-four-countries-in-the-african-region-license-vaccine-in-milestone-for-ebola-prevention.
132. Ebola Vaccines. National Institutes of Health. 2020; Available from: https://www.niaid.nih.gov/diseases-conditions/ebola-vaccines.
133. Maxmen, A. World's second-deadliest Ebola outbreak ends in Democratic Republic of the Congo. 2020; Available from: https://www.nature.com/articles/d41586-020-01950-0
134. Gavi Board approves new Ebola vaccine programme. Gavi, the global vaccines alliance. 2019; Available from: https://www.gavi.org/news/media-room/gavi-board-approves-new-ebola-vaccine-programme.
135. Laboratory diagnosis of Ebola virus disease. World Health Organization. 2014; Available from: https://www.who.int/csr/resources/publications/ebola/laboratory-guidance/en/.
136. Guidance for U.S. Laboratories for Managing and Testing Routine Clinical Specimens When There is a Concern about Ebola Virus Disease. Centers for Disease Control and Prevention. 2018; Available from: https://www.cdc.gov/vhf/ebola/laboratory-personnel/safe-specimen-management.html.
137. Target product profile for Zaire Ebolavirus. Rapid, simple test to be used in the control of the Ebola outbreak in West Africa. World Health Organization. 2014; Available from: http://www.who.int/medicines/publications/target-product-profile.pdf.
138. Broadhurst, M. J., Brooks, T. J., Pollock, N. R. Diagnosis of Ebola Virus Disease: Past, Present, and Future. Clin Microbiol Rev. 2016. 29(4): pp. 773-93.
139. Baize, S., Leroy, E. M., Georges-Courbot, M. C., et al. Defective humoral responses and extensive intravascular apoptosis are associated with fatal outcome in Ebola virus-infected patients. Nat Med. 1999. 5(4): pp. 423-6.
140. Reusken, C., Niedrig, M., Pas, S., et al. Identification of essential outstanding questions for an adequate European laboratory response to Ebolavirus Zaire West Africa 2014. J Clin Virol. 2015. 62: pp. 124-34.
141. First Antigen Rapid Test for Ebola through Emergency Assessment and Eligible for Procurement. World Health Organization. 2015; Available from: https://www.who.int/medicines/ebola-treatment/1st_antigen_RT_Ebola/en/.
142. Wonderly, B., Jones, S., Gatton, M. L., et al. Comparative performance of four rapid Ebola antigen-detection lateral flow immunoassays during the 2014-2016 Ebola epidemic in West Africa. PLoS One. 2019. 14(3): p. e0212113.
143. Towner, J. S., Rollin, P. E., Bausch, D. G., et al. Rapid diagnosis of Ebola hemorrhagic fever by reverse transcription-PCR in an outbreak setting and assessment of patient viral load as a predictor of outcome. J Virol. 2004. 78(8): pp. 4330-41.
144. Emergency Use Authorizations for Medical Devices. US Food and Drug Administration. 2020; Available from: https://www.fda.gov/medical-devices/emergency-situations-medical-devices/emergency-use-authorizations-medical-devices#ebola.
145. Cherpillod, P., Schibler, M., Vieille, G., et al. Ebola virus disease diagnosis by real-time RT-PCR: A comparative study of 11 different procedures. J Clin Virol. 2016. 77: pp. 9-14.
146. Hoenen, T., Groseth, A., Rosenke, K., et al. Nanopore Sequencing as a Rapidly Deployable Ebola Outbreak Tool. Emerg Infect Dis. 2016. 22(2): pp. 331-4.
147. Paweska, J. T., Jansen van Vuren, P., Meier, G. H., et al. South African Ebola diagnostic response in Sierra Leone: A modular high biosafety field laboratory. PLoS Negl Trop Dis. 2017. 11(6): p. e0005665.

148. Quick, J., Loman, N. J., Duraffour, S., et al. Real-time, portable genome sequencing for Ebola surveillance. Nature. 2016. 530(7589): pp. 228-232.
149. Ebolavirus. Virus pathogen resource (ViPR). 2020; Available from: https://www.viprbrc.org/brc/home.spg?decorator=filo_ebola.
150. Gire, S. K., Goba, A., Andersen, K. G., et al. Genomic surveillance elucidates Ebola virus origin and transmission during the 2014 outbreak. Science. 2014. 345(6202): pp. 1369-72.
151. Jerome, H., Taylor, C., Sreenu, V. B., et al. Metagenomic next-generation sequencing aids the diagnosis of viral infections in febrile returning travellers. J Infect. 2019. 79(4): pp. 383-388.
152. Park, D. J., Dudas, G., Wohl, S., et al. Ebola Virus Epidemiology, Transmission, and Evolution during Seven Months in Sierra Leone. Cell. 2015. 161(7): pp. 1516-26.
153. Ebola virus. ARTIC Nework. 2020; Available from: https://artic.network/ebov.
154. Mbala-Kingebeni, P., Villabona-Arenas, C. J., Vidal, N., et al. Rapid Confirmation of the Zaire Ebola Virus in the Outbreak of the Equateur Province in the Democratic Republic of Congo: Implications for Public Health Interventions. Clin Infect Dis. 2019. 68(2): pp. 330-333.
155. Mbala-Kingebeni, P., Pratt, C. B., Wiley, M. R., et al. 2018 Ebola virus disease outbreak in Equateur Province, Democratic Republic of the Congo: a retrospective genomic characterisation. Lancet Infect Dis. 2019. 19(6): pp. 641-647.
156. Mbala-Kingebeni, P., Aziza, A., Di Paola, N., et al. Medical countermeasures during the 2018 Ebola virus disease outbreak in the North Kivu and Ituri Provinces of the Democratic Republic of the Congo: a rapid genomic assessment. Lancet Infect Dis. 2019. 19(6): pp. 648-657.
157. Matranga, C. B. , Andersen, K. G. , Winnicki, S., et al. Enhanced methods for unbiased deep sequencing of Lassa and Ebola RNA viruses from clinical and biological samples. Genome Biology. 2014. 15: p. 12.
158. Whitmer, S. L. M., Ladner, J. T., Wiley, M. R., et al. Active Ebola Virus Replication and Heterogeneous Evolutionary Rates in EVD Survivors. Cell Rep. 2018. 22(5): pp. 1159-1168.
159. Wawina-Bokalanga, T., Vanmechelen, B., Marti-Carreras, J., et al. Complete Genome Sequence of a New Ebola Virus Strain Isolated during the 2017 Likati Outbreak in the Democratic Republic of the Congo. Microbiol Resour Announc. 2019. 8(20).
160. Jansen van Vuren, P., Ladner, J. T., Grobbelaar, A. A., et al. Phylodynamic Analysis of Ebola Virus Disease Transmission in Sierra Leone. Viruses. 2019. 11(1).
161. Liu, W. J., Shi, W., Zhu, W., et al. Intra-host Ebola viral adaption during human infection. Biosaf Health. 2019. 1(1): pp. 14-24.
162. Baize, S., Pannetier, D., Oestereich, L., et al. Emergence of Zaire Ebola virus disease in Guinea. N Engl J Med. 2014. 371(15): pp. 1418-25.
163. Kugelman, J. R., Wiley, M. R., Mate, S., et al. Monitoring of Ebola Virus Makona Evolution through Establishment of Advanced Genomic Capability in Liberia. Emerg Infect Dis. 2015. 21(7): pp. 1135-43.
164. Tong, Y. G., Shi, W. F., Liu, D., et al. Genetic diversity and evolutionary dynamics of Ebola virus in Sierra Leone. Nature. 2015. 524(7563): pp. 93-6.
165. Goba, A., Khan, S. H., Fonnie, M., et al. An Outbreak of Ebola Virus Disease in the Lassa Fever Zone. J Infect Dis. 2016. 214(suppl 3): pp. S110-S121.
166. Simon-Loriere, E., Faye, O., Faye, O., et al. Distinct lineages of Ebola virus in Guinea during the 2014 West African epidemic. Nature. 2015. 524(7563): pp. 102-4.
167. Next-Generation Sequencing Aids Researchers in the Fight Against the Ebola Virus Illumina. 2015; Available from: https://emea.illumina.com/content/dam/illumina-marketing/documents/icommunity/sabeti-happi-miseq-ebola-article-1270-2015-001.pdf.
168. Di Paola, N., Sanchez-Lockhart, M., Zeng, X., et al. Viral genomics in Ebola virus research. Nat Rev Microbiol. 2020. 18(7): pp. 365-378.
169. Ebola/Marburg Research and Development (R&D) Roadmap. World Health Organization. 2018; Available from: https://www.who.int/docs/default-source/blue-print/ebola-marburg-research-and-development-(r-d)-roadmap.pdf.
170. Palacios, G., Briese, T., Kapoor, V., et al. MassTag polymerase chain reaction for differential diagnosis of viral hemorrhagic fever. Emerg Infect Dis. 2006. 12(4): pp. 692-5.
171. Pang, Z., Li, A., Li, J., et al. Comprehensive multiplex one-step real-time TaqMan qRT-PCR assays for detection and quantification of hemorrhagic fever viruses. PLoS One. 2014. 9(4): p. e95635.
172. Brinkmann, A., Ergunay, K., Radonic, A., et al. Development and preliminary evaluation of a multiplexed amplification and next generation sequencing method for viral hemorrhagic fever diagnostics. PLoS Negl Trop Dis. 2017. 11(11): p. e0006075.

173. Li, T., Mbala-Kingebeni, P., Naccache, S. N., et al. Metagenomic Next-Generation Sequencing of the 2014 Ebola Virus Disease Outbreak in the Democratic Republic of the Congo. J Clin Microbiol. 2019. 57(9).
174. Liu, X., Speranza, E., Munoz-Fontela, C., et al. Transcriptomic signatures differentiate survival from fatal outcomes in humans infected with Ebola virus. Genome Biol. 2017. 18(1): p. 4.
175. Speranza, E., Caballero, I., Honko, A. N., et al. Previremic Identification of Ebola or Marburg Virus Infection Using Integrated Host-Transcriptome and Viral Genome Detection. mBio. 2020. 11(3).
176. Gardner, C. L., Ryman, K. D. Yellow fever: a reemerging threat. Clinics in laboratory medicine. 2010. 30(1): pp. 237-260.
177. Stock, N. K., Laraway, H., Faye, O., et al. Biological and phylogenetic characteristics of yellow fever virus lineages from West Africa. Journal of virology. 2013. 87(5): pp. 2895-2907.
178. Klitting, R., Fischer, C., Drexler, J. F., et al. What Does the Future Hold for Yellow Fever Virus? (II). Genes. 2018. 9(9): p. 425.
179. Klitting, R., Gould, E. A., Paupy, C., et al. What Does the Future Hold for Yellow Fever Virus? (I). Genes (Basel). 2018. 9(6).
180. Immunization, Vaccines and Biologicals - Yellow Fever. World Health Organization. 2020; Available from: https://www.who.int/immunization/monitoring_surveillance/burden/vpd/surveillance_type/passive/yellow_fever/en/.
181. Gaythorpe, K. A. M., Hamlet, A. T. P., Jean, K., et al. The global burden of yellow fever. medRxiv. 2020. p. 10.1101/2020.10.14.20212472.
182. Vainio, J., Cutts, F. Yellow fever. World Health Organization 1998; Available from: https://apps.who.int/iris/bitstream/handle/10665/64455/WHO_EPI_GEN_98.11.pdf?sequence=1&isAllowed=y.
183. Barrett, A. D. T. The reemergence of yellow fever. Science. 2018. 361(6405): pp. 847-848.
184. Areas with risk of yellow fever transmission in Africa. Centers for Disease Control and Prevention. 2017; Available from: https://www.cdc.gov/yellowfever/maps/africa.html.
185. Yellow fever. World Health Organisation 2019; Available from: https://www.who.int/news-room/fact-sheets/detail/yellow-fever.
186. Weekly epidemiological record - global yellow fever update, 2019. World Health Organization 2020; Available from: https://www.who.int/publications/i/item/10665-333915.
187. Areas with risk of yellow fever virus transmission in South America. Centers for Disease Control and Prevention. 2018; Available from: https://www.cdc.gov/yellowfever/maps/south_america.html.
188. A global strategy to eliminate yellow fever epidemics (EYE) 2017-2026. World Health Organization. 2018; Available from: https://apps.who.int/iris/bitstream/handle/10665/272408/9789241513661-eng.pdf.
189. Low, J. G., Ng, J. H. J., Ong, E. Z., et al. Phase 1 Trial of a Therapeutic Anti-Yellow Fever Virus Human Antibody. N Engl J Med. 2020. 383(5): pp. 452-459.
190. Yellow fever vaccination booster not needed. World Health Organization 2013; Available from: https://www.who.int/mediacentre/news/releases/2013/yellow_fever_20130517/en/.
191. Target product profiles for identification of yellow fever infection. World Health Organization 2019; Available from: https://www.finddx.org/wp-content/uploads/2019/07/YellowFeverVirus-TPPs-2019-07-30.pdf.
192. Berkley, S. Health security's blind spot. Science. 2018. 359(6380): pp. 1075-1075.
193. Domingo, C., Ellerbrok, H., Koopmans, M., et al. Need for additional capacity and improved capability for molecular detection of yellow fever virus in European Expert Laboratories: External Quality Assessment, March 2018. Eurosurveillance. 2018. 23(28): p. 1800341.
194. Staples, J. E., Barrett, A. D. T., Wilder-Smith, A., et al. Review of data and knowledge gaps regarding yellow fever vaccine-induced immunity and duration of protection. npj Vaccines. 2020. 5(1): p. 54.
195. Kleinert, R. D. V., Montoya-Diaz, E., Khera, T., et al. Yellow Fever: Integrating Current Knowledge with Technological Innovations to Identify Strategies for Controlling a Re-Emerging Virus. Viruses. 2019. 11(10).
196. Pan American Health Organization Laboratory Diagnosis of Yellow Fever Infection. World Health Organization. 2018 Available from: https://www.who.int/publications/i/item/laboratory-diagnosis-of-yellow-fever-virus-infection.
197. Waggoner, J. J., Rojas, A., Pinsky, B. A. Yellow Fever Virus: Diagnostics for a Persistent Arboviral Threat. J Clin Microbiol. 2018. 56(10).

198. Maurício de Rezende, I., Alves, P. A., Arruda, M. S., et al. Yellow Fever Virus Genotyping Tool and Investigation of Suspected Adverse Events Following Yellow Fever Vaccination. Vaccines. 2019. 7(4): p. 206.
199. Cui, S., Pan, Y., Lyu, Y., et al. Detection of yellow fever virus genomes from four imported cases in China. Int J Infect Dis. 2017. 60: pp. 93-95.
200. Phan, M. V. T., Mendonca Melo, M., van Nood, E., et al. Shedding of Yellow Fever Virus From an Imported Case in the Netherlands After Travel to Brazil. Open Forum Infectious Diseases. 2020. 7(2).
201. Ajogbasile, F. V., Oguzie, J. U., Oluniyi, P. E., et al. Real-time Metagenomic Analysis of Undiagnosed Fever Cases Unveils a Yellow Fever Outbreak in Edo State, Nigeria. Scientific Reports. 2020. 10(1): p. 3180.
202. CDC and WHO launch trial version of a new ELISA: A kit, simpler and faster to perform. World Health Organization 2018; Available from: https://www.who.int/csr/disease/yellowfev/trial-version-of-a-new-ELISA-test/en/.
203. Ricciardi-Jorge, T., Bordignon, J., Koishi, A., et al. Development of a quantitative NS1-capture enzyme-linked immunosorbent assay for early detection of yellow fever virus infection. Scientific Reports. 2017. 7(1): p. 16229.
204. Yen, C. W., de Puig, H., Tam, J. O., et al. Multicolored silver nanoparticles for multiplexed disease diagnostics: distinguishing dengue, yellow fever, and Ebola viruses. Lab Chip. 2015. 15(7): pp. 1638-41.
205. Kim, Y. H., Kim, T.-Y., Park, J.-S., et al. Development and Clinical Evaluation of a Rapid Diagnostic Test for Yellow Fever Non-Structural Protein 1. The Korean journal of parasitology. 2019. 57(3): pp. 283-290.
206. Rice, C. M., Lenches, E. M., Eddy, S. R., et al. Nucleotide sequence of yellow fever virus: implications for flavivirus gene expression and evolution. Science. 1985. 229(4715): pp. 726-33.
207. McMullan, L. K., Frace, M., Sammons, S. A., et al. Using next generation sequencing to identify yellow fever virus in Uganda. Virology. 2012. 422(1): pp. 1-5.
208. Quick, J., Grubaugh, N. D., Pullan, S. T., et al. Multiplex PCR method for MinION and Illumina sequencing of Zika and other virus genomes directly from clinical samples. Nat Protoc. 2017. 12(6): pp. 1261-1276.
209. Faria, N. R., Kraemer, M. U. G., Hill, S. C., et al. Genomic and epidemiological monitoring of yellow fever virus transmission potential. Science. 2018. 361(6405): pp. 894-899.
210. KRISP biotools introduction. KRISP. 2020; Available from: https://www.krisp.org.za/tools.php.
211. Goes de Jesus, J., Graf, T., Giovanetti, M., et al. Yellow fever transmission in non-human primates, Bahia, Northeastern Brazil. PLoS Negl Trop Dis. 2020. 14(8): p. e0008405.
212. Bae, H. G., Drosten, C., Emmerich, P., et al. Analysis of two imported cases of yellow fever infection from Ivory Coast and The Gambia to Germany and Belgium. J Clin Virol. 2005. 33(4): pp. 274-80.
213. Pestana, C. P., Lawson-Ferreira, R., Lessa-Aquino, C., et al. Sanger-based sequencing technology for yellow fever vaccine genetic quality control. J Virol Methods. 2018. 260: pp. 82-87.
214. Tosta, S. F. de O., Passos, M. S., Kato, R., et al. Multi-epitope based vaccine against yellow fever virus applying immunoinformatics approaches. Journal of Biomolecular Structure and Dynamics. 2019. pp. 1-17.
215. Cunha, M. d. P., Duarte-Neto, A. N., Pour, S. Z., et al. Origin of the São Paulo Yellow Fever epidemic of 2017–2018 revealed through molecular epidemiological analysis of fatal cases. Scientific Reports. 2019. 9(1): p. 20418.
216. Hill, S. C., de Souza, R., Theze, J., et al. Genomic Surveillance of Yellow Fever Virus Epizootic in Sao Paulo, Brazil, 2016 - 2018. PLoS Pathog. 2020. 16(8): p. e1008699.
217. Cunha, M. S., da Costa, A. C., de Azevedo Fernandes, N. C. C., et al. Epizootics due to Yellow Fever Virus in São Paulo State, Brazil: viral dissemination to new areas (2016-2017). Sci Rep. 2019. 9(1): p. 5474.
218. Cunha, M. S., Faria, N. R., Caleiro, G. S., et al. Genomic evidence of yellow fever virus in Aedes scapularis, southeastern Brazil, 2016. Acta Trop. 2020. 205: p. 105390.
219. Domingo, C., Escadafal, C., Rumer, L., et al. First international external quality assessment study on molecular and serological methods for yellow fever diagnosis. PLoS One. 2012. 7(5): p. e36291.
220. Gargis, A. S., Kalman, L., Lubin, I. M. Assuring the Quality of Next-Generation Sequencing in Clinical Microbiology and Public Health Laboratories. J Clin Microbiol. 2016. 54(12): pp. 2857-2865.
221. Technical guide on next-generation sequencing technologies for the detection of mutations associated with drug resistance in *Mycobacterium tuberculosis* complex. World Health Organization.

2018; Available from: https://www.who.int/tb/publications/2018/WHO_technical_guide_nextgen_sequencing/en/.
222. GLASS whole-genome sequencing for surveillance of antimicrobial resistance. World Health Organization. 2020; Available from: https://www.who.int/publications/i/item/9789240011007.
223. WHOs code of conduct for open and timely sharing of pathogen genetic sequence data during outbreaks of infectious disease. World Health Organization. 2019; Available from: https://www.who.int/blueprint/what/norms-standards/gsdsharing/en/.
224. Africa Pathogen Genomics Initiative. Africa CDC. 2018; Available from: https://africacdc.org/institutes/africa-pathogen-genomics-initiative/.
225. Sichtig, H., Minogue, T., Yan, Y., et al. FDA-ARGOS is a database with public quality-controlled reference genomes for diagnostic use and regulatory science. Nat Commun. 2019. 10(1): p. 3313.
226. Nkengasong, J. Let Africa into the market for COVID-19 diagnostics. Nature. 2020. 580(7805): p. 565.
227. Nkengasong, J. N., Tessema, S. K. Africa Needs a New Public Health Order to Tackle Infectious Disease Threats. Cell. 2020. 183(2): pp. 296-300.
228. Rossen, J. W. A., Friedrich, A. W., Moran-Gilad, J., et al. Practical issues in implementing whole-genome-sequencing in routine diagnostic microbiology. Clin Microbiol Infect. 2018. 24(4): pp. 355-360.
229. Whole genome sequencing for foodborne disease surveillance. World Health Organization. 2018; Available from: https://www.who.int/foodsafety/publications/foodborne_disease/wgs_landscape/en/.
230. Safronetz, D., Sogoba, N., Lopez, J. E., et al. Geographic distribution and genetic characterization of Lassa virus in sub-Saharan Mali. PLoS Negl Trop Dis. 2013. 7(12): p. e2582.
231. Akhuemokhan, O. C., Ewah-Odiase, R. O., Akpede, N., et al. Prevalence of Lassa Virus Disease (LVD) in Nigerian children with fever or fever and convulsions in an endemic area. PLOS Neglected Tropical Diseases. 2017. 11(7): p. e0005711.
232. Happi, A. N., Ugwu, C. A., Happi, C. T. Preparing for the next Ebola outbreak: in-country genomic capacity in Africa. The Lancet Infectious Diseases. 2019. 19(6): pp. 569-570.
233. Zusi, K. Rapid genomic sequencing of Lassa virus in Nigeria enabled real-time response to 2018 outbreak. Broad Insitute. 2018; Available from: https://www.broadinstitute.org/news/rapid-genomic-sequencing-lassa-virus-nigeria-enabled-real-time-response-2018-outbreak.
234. Beitzel, B., Hulseberg, C. E., Palacios, G. Reverse genetics systems as tools to overcome the genetic diversity of Lassa virus. Curr Opin Virol. 2019. 37: pp. 91-96.
235. Andersen, K. G., Shylakhter, I., Tabrizi, S., et al. Genome-wide scans provide evidence for positive selection of genes implicated in Lassa fever. Philos Trans R Soc Lond B Biol Sci. 2012. 367(1590): pp. 868-77.
236. Mari Saez, A., Weiss, S., Nowak, K., et al. Investigating the zoonotic origin of the West African Ebola epidemic. EMBO Mol Med. 2015. 7(1): pp. 17-23.
237. Carroll, M. W., Matthews, D. A., Hiscox, J. A., et al. Temporal and spatial analysis of the 2014-2015 Ebola virus outbreak in West Africa. Nature. 2015. 524(7563): pp. 97-101.
238. Ladner, J. T., Wiley, M. R., Mate, S., et al. Evolution and Spread of Ebola Virus in Liberia, 2014-2015. Cell Host Microbe. 2015. 18(6): pp. 659-69.
239. Jiang, T., Jiang, J. F., Deng, Y. Q., et al. Features of Ebola Virus Disease at the Late Outbreak Stage in Sierra Leone: Clinical, Virological, Immunological, and Evolutionary Analyses. J Infect Dis. 2017. 215(7): pp. 1107-1110.
240. Whitmer, S. L., Albarino, C., Shepard, S. S., et al. Preliminary Evaluation of the Effect of Investigational Ebola Virus Disease Treatments on Viral Genome Sequences. J Infect Dis. 2016. 214(suppl 3): pp. S333-S341.
241. Guedj, J., Piorkowski, G., Jacquot, F., et al. Antiviral efficacy of favipiravir against Ebola virus: A translational study in cynomolgus macaques. PLoS Med. 2018. 15(3): p. e1002535.
242. Piorkowski, G., Jacquot, F., Querat, G., et al. Implementation of a non-human primate model of Ebola disease: Infection of Mauritian cynomolgus macaques and analysis of virus populations. Antiviral Res. 2017. 140: pp. 95-105.
243. Giovanetti, M., de Mendonça, M. C. L., Fonseca, V., et al. Yellow Fever Virus Reemergence and Spread in Southeast Brazil, 2016-2019. Journal of virology. 2019. 94(1): pp. e01623-19.
244. Hughes, H. R., Kayiwa, J., Mossel, E. C., et al. Phylogeny of Yellow Fever Virus, Uganda, 2016. Emerg Infect Dis. 2018. 24(8).
245. Phan, M. V., Murad, S. D., van der Eijk, A. A., et al. Genomic sequence of yellow fever virus from a Dutch traveller returning from the Gambia-Senegal region, the Netherlands, November 2018. Euro

surveillance : bulletin Europeen sur les maladies transmissibles = European communicable disease bulletin. 2019. 24(4): p. 1800684.
246. Stock, N. K., Boschetti, N., Herzog, C., et al. The phylogeny of yellow fever virus 17D vaccines. Vaccine. 2012. 30(6): pp. 989-94.
247. Chen, C., Jiang, D., Ni, M., et al. Phylogenomic analysis unravels evolution of yellow fever virus within hosts. PLoS Negl Trop Dis. 2018. 12(9): p. e0006738.