New technological developments in identification and monitoring of new and emerging infections

Gert Uves van Zyl, Division of Medical Virology, Stellenbosch University and National Health Laboratory Service, Tygerberg Business Unit, Cape Town, South Africa

Introduction

The year 2020 marked awarding the Nobel Prize in Physiology or Medicine for the discovery of hepatitis C virus (HCV). More than 30 years before, in 1989, Michael Houghton and others published the first HCV genome sequence (Choo et al., 1989; Houghton, 2009). This discovery of HCV was enabled by preceding developments and culminated work of more than a decade since the first description in 1978, by Harvey Alter, of non-A non-B hepatitis as an agent transmissible by blood (Alter et al., 1978). However, after the 1989 first identification of a HCV sequence, it took another 7 years, until 1996, for the viral genome to be fully characterized (Kolykhalov et al., 1997). The many years required from clinical description to full genome characterization of HCV, by the end of the 20th century, stand in stark contrast to the full genome characterization of SARS coronavirus-2 (SARS-CoV-2), the virus responsible for coronavirus infectious disease 2019 (COVID-19), within a few weeks after the novel infection was recognized (Zhou et al., 2020; Zhu et al., 2020). Such a rapid discovery is a result of multiple technological developments, reviewed in this article. Moreover, immediate global sharing of the SARS-CoV-2 full genome sequence enabled the development of diagnostic tests and was essential for the fastest ever development and licensing of vaccines.

Emerging infections can be defined as either new infections, newly diagnosed infections, or infections that have increased in their geographic distribution, or which are increasingly found in vulnerable population groups, such as individuals with immunosuppression or intravenous drug users (Morens et al., 2004; Morse, 1995). Moreover, a combination of epidemiological evidence with microbiological identification has often resulted in the discovery of novel disease associations of known infectious agents

Table 1   Factors contributing to the emergence or re-emergence of pathogens.

Factor	Emerging and re-emerging pathogens
Incursion into natural habitats	Exposure to bat-borne paramyxoviruses such as Nipah virus, phyloviruses such as Ebolavirus and Marburg virus
Global travel	Rapid dissemination of 2009 H1N1 Pandemic (H1N1pdm09 virus); SARS-CoV-2 global pandemic
Host immunosuppression: HIV, organ transplant, immune suppressive therapies for cancer or connective tissue diseases	Opportunistic infections
Climate change	Vector borne disease (e.g., arboviruses)
Live oral polio vaccination to achieve herd immunity	Vaccine-derived polioviruses
Vaccine hesitancy	Measles outbreaks
Exotic animal trade	West African rodents: Monkeypox virus

(Morens et al., 2004). The last 30 years saw the discovery of many new or emerging infectious diseases, or infections that re-emerged or increased in incidence after periods of lower occurrence. The discovery and monitoring of novel and emerging pathogens have been made possible by many new developments in diagnostics, of which some of the most significant will be reviewed in this article.

Several factors are contributing to the emergence and re-emergence of pathogens (Table 1). Population growth has increased human incursions into regions where they are at increased risk of zoonotic infections. Climate change has increased the distribution of arthropod vectors and risk of human exposure to these pathogens. Global travel and trade networks have also resulted in an increasingly connected world population with the rapid dissemination of pathogens once they become established in humans. Changes in population vulnerability also contribute. HIV has resulted in the emergence of multiple opportunistic infections, which have recently declined with wider availability of antiretroviral combination therapy. Aging populations in industrialized countries may be increasingly vulnerable to particular pathogens due to immune senescence. Immune suppressive therapy is used in the context of organ transplantation, chemotherapy for cancer, or treatment for connective tissue diseases. These factors provide vulnerable populations to common and opportunistic pathogens (Morens et al., 2004). In addition, socio-economic development, secular trends, and vaccination behavior may either result in overall changes in the risk of infections or in the age of acquisition of particular infections, which influence whether individuals present with symptomatic or asymptomatic infection (Jacobsen and Koopman, 2005). Changes in the age of infection influence whether infections would occur in women of child-bearing-potential and resultant congenital infection, in case of rubella (Panagiotopoulos et al., 1999). Whereas vaccine campaigns have reduced the burden of many infectious diseases, it may have contributed to lesser awareness of the risks of vaccine preventable disease. Paradoxically the success of vaccination may have become the reason why people are less aware of real risks of acquiring infections and more aware of possible risks of vaccination and susceptible to antivaccine propaganda. This has resulted in the re-emergence of measles in settings where it has been eliminated especially among groups with a low vaccine coverage (Coombes, 2017; Lo and Hotez, 2017). A lack of antibiotic stewardship, which involves the use of broad spectrum antibiotics, while not clinically indicated, has contributed to the emergence of highly and broadly drug-resistant bacteria (Baur et al., 2017). It is therefore clear that the threat of emerging and re-emerging infectious diseases is unlikely to diminish in the near future. Fortunately technological development in the diagnosis and monitoring of infectious diseases may enable future rapid identification and response to these emerging health threats.

Identifying agents and proving an etiological role in disease

Koch's postulates, which require the presence of an infectious agent in cases with disease, and its absence in those without, and the isolation of the agent in pure cultures, were initially very important to identify infectious etiologies. However, it had limitations. Many organisms cannot be readily grown in cultures. Moreover, the association between infectious agent and disease is not always a simple single agent deterministic association but could be more complex—infections could be present without causing disease or may not be the only (or necessary cause) of a disease. In addition to infection, it may require other factors such as host risk factors to cause disease. In the latter case the infectious agent is not a sufficient cause (Byrd and Segre, 2016; Falkow, 2004). Disease complications may also be the result of chronic infections with a long time lag between infection and disease (O'Connor et al., 2006) which makes it more difficult to directly associate the infectious agent and disease.

An adaptation of Koch's postulates was therefore proposed based on molecular criteria that relied on identification of pathogen sequences (Falkow, 1988). However, complex etiologies require epidemiological evidence that shows an increased risk in individuals who have particular infections compared to those without. In the case of organisms that result in pathologies such as neoplasms, or other delayed complications of infections a combination of molecular and epidemiological evidence is often required. Cervical cancer is almost always the result of infection with oncogenic types of human papillomaviruses (Nicolás et al., 2019) and hepatitis C virus and hepatitis B virus infections are strongly associated with liver cancer. However, other infections such as human herpes virus 8 (HHV8) and Epstein bar virus (EBV) require cofactors before oncogenesis is likely (Mesri et al., 2014).

Proving the association of chronic *Helicobacter pylori* stomach infection with duodenal ulcers and stomach cancer has required epidemiological and intervention studies, in addition to identifying the infection in people with disease (Choi et al., 2018; Preda et al., 2009). Recent developments in human genetics, the analysis of mRNA expression (transcriptomics), the study of protein structure, function and interaction (proteinomics), and metabolic pathways (metabolomics) have resulted in increasingly rich data, which help to explain the interaction of host genetics, infections, and the environment. In future, novel approaches to analyze such rich data could allow for a better understanding of cellular level pathways of diseases with complex etiologies. This could also provide novel drug targets to treat these diseases (Karczewski and Snyder, 2018).

Traditional approaches to identifying infectious diseases

The identification and diagnosis of infectious diseases have advanced along with technological development and innovations that involve both biochemistry and engineering. Time-tested or traditional diagnostic approaches have their value and limitations, summarized in Table 2 and discussed below. These methods remain valuable in the discovery and characterization of emerging pathogens and are often complementary to more recently developed molecular assays.

Microscopy and culture-based isolation

Since the discovery of microorganisms by light microscopy and the germ hypothesis developed by Pasteur, Koch, and others, culture media have been developed that allowed for the isolation of bacteria (Blevins and Bronze, 2010; Pasteur, 1881). However, viruses, which are responsible for the majority of emerging infections, were initially elusive as they require living cells (either cell culture or live animals) for growth and laboratory isolation. Viruses are also too small to be held back by bacterial filters or to be identified by light microscopy (Lwoff, 1957). The development of cell culture systems greatly aided the identification of viruses from patient samples (Leland and Ginocchio, 2007). Similarly intracellular bacteria, such as mycobacteria, chlamydia, mycoplasmas, and rickettsiae, are difficult to grow and require special media and often long incubation periods or laboratory animals to isolate them (Doern, 2000). Electron microscopy, which has a much higher resolution than light microscopy, allows the characterization of viral structure and, based on the identified morphology, can then direct further investigations to identify the particular species (Curry et al., 2006).

Immune-based diagnostics

Sera from patients with infections or who recovered from infections proved to be very valuable for diagnosing novel agents and the earliest techniques involved immunodiffusion, precipitation, and counterimmunoelectrophoresis reactions that helped to identify the causative agents (Gocke and Howe, 1970; Stanford, 1973). The underlying principle is the hybridization or binding of antibodies from serum to an antigen. The study of antibody and antigen reactions is often called "serology" as antibodies were initially identified from serum samples. The detection of antibodies would signify prior human exposure to an infectious agent, whereas the detection of antigen using antibodies (often produced in laboratory animals or laboratory cultures) would indicate the current presence of the antigen in a sample or culture isolate from a sample. Different technologies were developed to identify antigen-antibody hybridization, which include immunofluorescence, enzyme immunoassays, and chemiluminescence. Although serological methods were some of the first used for diagnosis of infectious agents, they remain relevant for the diagnosis of novel infectious agents especially when used to determine the prevalence of antibodies as evidence of exposure to a particular pathogen. Serological investigations are applied to determine the extent of exposure of the population to new agents and has been used to assess the extent of COVID-19 spread, since many cases may remain asymptomatic or undiagnosed (Larremore et al., 2021; Metcalf et al., 2016). The study of the presence of antibodies among human populations, referred to as seroepidemiology, is discussed in more detail below.

Table 2   Value and limitations of traditional approaches to identify and diagnose infections.

Method	Value in pathogen identification	Limitations
Bacterial growth media (serum broths and other)	Pure cultures of fast growing bacteria	Cannot isolate viruses or fastidious bacteria
Cell culture methods	Isolation of particular viruses	Many viruses do not show cytopathic effect; and others do not grow in available cell culture systems
Polymerase chain reaction (PCR)	Sensitive detection of low concentration agent	Requires some prior knowledge of the sequence to design primers
Serology	Investigate the presence of antibodies in the community (seroprevalence surveys); determine if newly discovered agent had been circulating in the population before	Indirect evidence by detecting antibodies as part of the immune response, which is variable and could wane after infection. Cross reactivity between related agents (e.g., from the same family such as flaviviruses)

Limitations of traditional diagnostic methods

Despite their value, traditional laboratory diagnostic methods have limitations (Table 2). Cell culture-based isolation of viruses and culture of fastidious bacteria are cumbersome and slow and electron microscopy requires a high pathogen concentration and therefore lacks sensitivity (Lipkin and Firth, 2013). Moreover, for many infectious agents, culture isolation systems are not readily available. The identification of these organisms therefore relies on genetic identification, often referred to as molecular diagnostic methods.

The rapid development in molecular methods

Nucleic acid amplification: The polymerase chain reaction and alternatives

The most important breakthrough in the rapid characterization of viruses and fastidious organisms was the development of the polymerase chain reaction (PCR) by Kary Mullis in 1985 (Mullis et al., 1986; Mullis, 1990) for which he was awarded the Nobel prize in 1993. PCR allows the efficient amplification of DNA to detect low concentrations of infectious agents, directly from biological material, without requiring prior isolation in cell culture. An adaptation of this method, first reverse transcribing RNA to DNA, followed by PCR, called reverse transcriptase PCR (RT-PCR), allowed for the identification of RNA viruses.

Alternatives to PCR are isothermal nucleic acid amplification reactions such as nucleic acid sequence-based amplification (NASBA) or transcription-mediated amplification (TMA), which apply similar principles, loop-mediated isothermal amplification (LAMP) and recombinase polymerase amplification (RPA). The benefit of these isothermal methods is that they may require simpler instrumentation as thermocycling is not required (Gill and Ghaemi, 2008; Li and Macdonald, 2015).

When PCR is followed by nucleotide sequencing it allows for the full characterization of an agent and through phylogenetic analysis a description of its relatedness to other agents. However as PCR amplification requires primers to bind to specific DNA sequences in the template, it requires some prior knowledge of the infectious agent's sequence to allow identification. Therefore, identifying agents that are unexpected or distantly related to known agents offers a challenge and requires adaptation of PCR or alternative approaches for molecular diagnosis (Allander et al., 2005; Reyes and Kim, 1991).

Improvements in biochemistry and engineering, which include automation of extraction and amplification in partially or fully automated closed systems, resulted in nucleic acid amplification technology moving from specialized research laboratories to high-throughput diagnostic laboratories. This had been evident in automated viral load monitoring of HIV and HCV to assess treatment success, and multiplex molecular syndromic diagnosis of respiratory disease, meningitis, and sexually transmitted diseases (Greub et al., 2016). In addition, blood banks make use of molecular methods for blood donor screening to ensure the safety of blood products from donors, who may be in the antibody negative window period (Vermeulen et al., 2009). Recently the importance of high-throughput molecular diagnostics has become most evident, when laboratories have been required to rapidly scale up to test thousands of samples for SARS-CoV-2 by RT-PCR. This was necessary for individual diagnostic and epidemic control purposes (Gorzalski et al., 2020; Opota et al., 2020). In contrast to large high-throughput automated laboratory instrumentation, technological development has also resulted in smaller footprint and easy-to-use, near-patient molecular diagnostic instruments that have become increasingly important in the diagnosis of emerging infectious diseases, at the bedside, or point of care (POC). This has played a major role in the diagnosis of Ebolavirus in field hospital settings and rapid diagnosis of SARS-CoV-2, when required for clinical diagnosis or cohort nursing of infected or uninfected individuals (Raftery et al., 2018; Zhen et al., 2020).

Methods that make use of nucleic acid amplification, primarily with PCR, have overcome some of the limitations of traditional methods. Nevertheless, identification of unknown novel infections from clinical samples poses several challenges for which novel solutions were sought (Table 3).

Dealing with human cellular background

Identifying pathogen genetic material among a much higher concentration of normal human cellular nucleic acid (DNA and RNA) is a challenge. One approach to overcome this is representational difference analysis (RDA) developed in 1993. This approach selectively enriches for DNA sequences that are unique to a particular type of diseased tissue (Lisitsyn et al., 1993). An infectious etiology had long been suspected for Kaposi sarcoma (KS), but it was only when RDA was applied to these tumor tissues that it was

Table 3   Challenges to viral discovery from clinical samples.

Challenge	Impact	Solutions
High concentration of host nucleic acid	Competing host material result in time-consuming process of screening	Pre-analytical *Viruses*: ultracentrifugation and DNAse treatment of samples (viral particles are protected) *Subtraction*: representational difference analysis (RDA)
Unknown sequence	PCR amplification requires specific primer binding sites	Sequence independent (unbiased) amplification methods; Random PCR; degenerate PCR

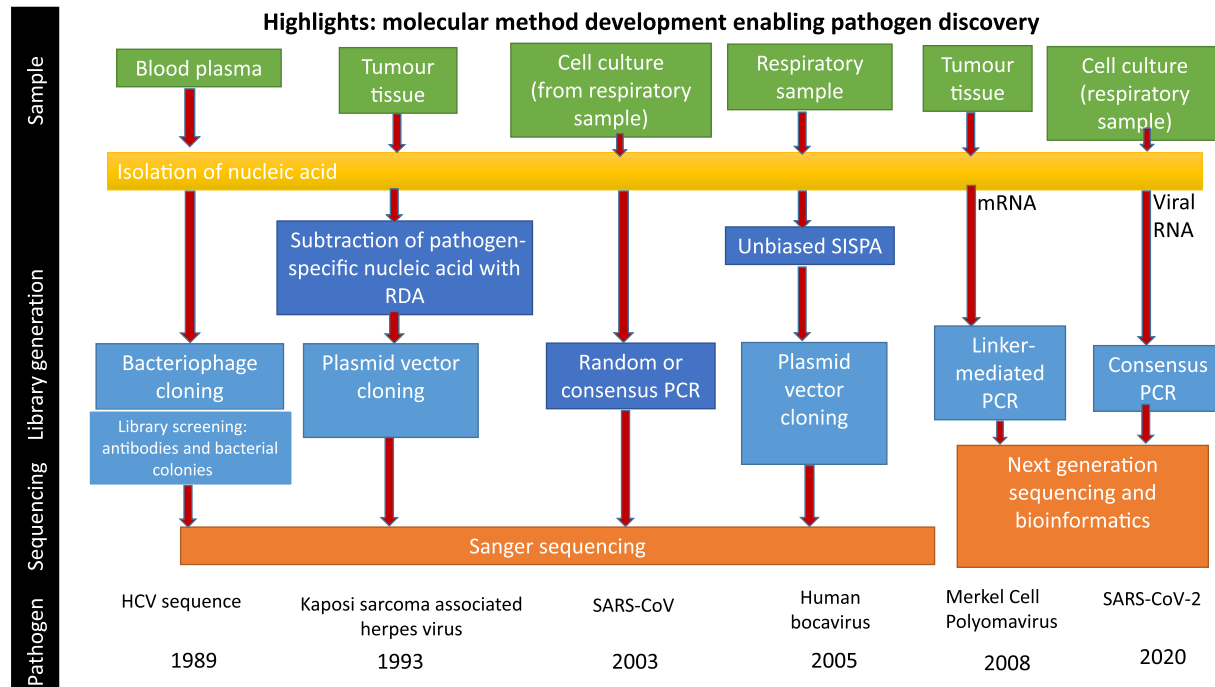**Highlights: molecular method development enabling pathogen discovery**

Fig. 1 The first Hepatitis C virus (HCV) sequence was obtained from a bacteriophage library in 1989. Representational difference analysis (RDA) facilitated the discovery of human herpes virus 8, in the presence of a high concentration of cellular DNA. The use of consensus primers (targeting conserved regions) and random primers facilitated the discovery of Severe acute respiratory syndrome (SARS) coronavirus in 2003. Sequence independent single primer amplification (SISPA) was a subsequent method that allowed for the non-specific amplification in samples where cellular background was reduced by DNAse treatment. Merkel cell associated polyomavirus was identified by linker-mediated PCR, next generation sequencing of mRNA and digital (bioinformatics based) subtraction transcriptome data of normal tissues from that of tumor tissue. Consensus PCR and next generation sequencing of cell-culture supernatant helped to identify SARS-coronavirus 3 (SAR-CoV-2) the cause of Coronavirus disease 2019 (COVID-19).

found that a previously unidentified human herpes virus, called Human Herpesvirus 8 (HHV-8) or Kaposi Sarcoma-Associated Herpesvirus, was present in all KS tissues (Chang et al., 1994). The secret of RDA is that it uses a modified PCR to selectively amplify gene sequences that are present in diseased tissue but absent or at much lower concentration in healthy tissues. It starts with hybridizing a low concentration of amplified products (tester amplicons) from the diseased tissue, ligated to primers, to an excess concentration (driver amplicon) from normal tissues. DNA strands that are unique to the diseased tissue are then preferentially exponentially amplified (Lisitsyn et al., 1993), which facilitates identification of the disease-specific sequences. Subsequently to identifying HHV-8 as the cause of KS, it was found to be associated with Multicentric Castleman's Disease (Soulier et al., 1995) and Primary Effusion Lymphoma (Cesarman et al., 1995). Other viruses identified using RDA are human GB virus (GBV) (Simons et al., 1995), a virus closely related to HCV, but which appears nonpathogenic, and TT virus (TTV) (Nishizawa et al., 1997), from the family Anelloviridae, which is commonly found in blood and is of uncertain significance. Despite the value of RDA, it is very cumbersome, whereas some recently developed biochemical methods are faster and more practicable (Fig. 1).

Unbiased amplification methods

DNase Sequence-Independent Single Primer Amplification (DNase-SISPA) was used to allow nonspecific amplification of viral RNA or DNA from serum, which enabled the identification of human parvovirus 4 (Jones et al., 2005). In short, filtered plasma samples were treated with DNAse to remove human DNA, whereafter viral RNA and DNA were extracted. Thereafter RNA was reverse transcribed and complementary DNA (cDNA) synthesized with random primers, followed by restriction endonuclease digestion of viral DNA or double-stranded DNA, reverse transcribed from RNA, and adaptor ligation to allow nonspecific amplification, cloning, and sequencing. A similar approach: Virus Discovery based on cDNA-AFLP (Amplified Fragment Length Polymorphism) (VIDISCA) was used to discover Coronavirus NL-63 in 2004. This method starts with blood plasma, serum, or culture supernatant. Cellular material and mitochondria are first removed by centrifugation, followed by DNAse treatment to digest mitochondrial or cellular DNA present in cell fragments, whereas nucleic acid in viral particles remains protected. Viral nucleic acid is then isolated and viral RNA is reverse transcribed to cDNA, followed by second strand synthesis. Double-stranded DNA (either directly from viral DNA or from reverse transcribed RNA) is digested with restriction enzymes, anchors are ligated, which contain primer binding sequences. This is followed by PCR amplification and finally by nucleotide sequencing providing nucleotide sequences from an unknown agent, without requiring any prior sequence knowledge (Van Der Hoek et al., 2004).

Improvements in sequencing technology

First-generation sequencing

The first widely adapted sequencing method was the Maxam and Gilbert technique which relied on chemical treatment generating breaks in a small proportion of DNA strands, separately for each of the four nucleotides. Each of these separately cleaved products is then separated into individual fragments with polyacrylamide gel electrophoresis (Maxam and Gilbert, 1977). Fred Sanger developed a method, which later largely replaced the Maxam and Gilbert method. This relies on DNA polymerase incorporating a mix of normal and chain terminating nucleotides during the process of primer extension with complimentary DNA synthesis. As the chain terminating nucleotides are incorporated, fragments of differing lengths are generated, each ending on a respective chain terminating nucleotide (Sanger et al., 1977). Initially the terminating nucleotides were radiolabelled, but once this was replaced by four different fluorescent labels (Prober et al., 1987) it opened the way for automated capillary electrophoresis, which is the version of Sanger sequencing that remains in use today. Each capillary has as single readout, called an electropherogram, which shows the respective nucleotides, in color, incorporated at each position. Scaling up automated Sanger sequencing relied on increasing the number of capillaries run in parallel. This unfortunately increases instrument footprint and despite Sanger sequencing's elegance exposed its limited scalability: in essence each sequencing reaction and capillary, in which electrophoresis takes place, aims to provide one consensus sequence read. The multitude of automated sequencers therefore required for the human genome project highlighted the limitations of automated Sanger sequencing. Making sequencing more scalable therefore required an approach that would allow the simultaneous sequencing of multiple different templates in one reaction. Several different solutions to this have been developed and have been referred to next-generation sequencing (NGS) or massive parallel sequencing (MPS) (Thermes, 2014; von Bubnoff, 2008).

The development of next-generation sequencing

These sequencing methods employ different principles, all to achieve the objective of multiple simultaneous sequencing reactions, which are physically separated and separately detected (Table 4). As NGS evolved, particular methods including PacBio Single Molecule, Real-Time (SMRT) Sequencing, and Oxford Nanopore Technologies (ONT) developed approaches that enabled the sequencing of very long reads of single genomes. Hence, these massive parallel sequencing methods became known as third-generation sequencing or single molecule sequencing (SMS) to distinguish them from shorter read NGS methods.

When used for pathogen discovery, in principle two different approaches can be followed. The one relies on specific enrichment for related pathogens, with a PCR or other amplification method that includes primers that bind to sequences that are conserved across species or genera, which is then followed by sequencing to characterize the pathogen. The second is nonspecific, attempts to detect all possible infectious agents, present in a sample, and relies on nonspecific amplification or direct sequencing of extracted nucleic acid. This is followed by the alignment of these diverse genetic sequences to bioinformatics databases. This latter process is referred to as metagenomics.

Table 4   Overview of next and third generation sequencing methods.

Method	Separation	Chemical principle	Detection
454 sequencing	DNA linked to beads separated on picotiter wells	As the primer extends, pyrophosphate (PPi) is released; ATP is generated and luciferase mediates conversion of luciferin to oxyluciferin	Light is released
Ion Torrent	DNA is separated on beads in micromachined wells	As primer extends, hydrogen (protons) are released.	Semiconductor chip detects change in voltage
Illumina	Oligonucleotides on nanowells capture DNA (bind to adaptors)	Bridge amplification generates clusters. Sequencing by extension of reversible terminating dNTPs	Four dNTPs are tagged with different fluorophores; laser excites fluorophores and fluorescent light is recorded by a digital camera.
PacBio Single Molecular Real-Time (SMRT) sequencing	Flowcell with thousands of transparent bottom picotiter wells: Zero-mode wave guides (ZMW)	Polymerase is immobilized on the bottom of the ZMD and adds fluorescent nucleotides by primer extension	Movie of light signals is detected as sequential bases are added by immobilized polymerase
Oxford Nanopore Technologies	Individual DNA molecules pass through nanopores that bridge a synthetic resistive polymer membrane	Nucleotide bases occupy a nanopore, as they pass through, resulting in a current change	Current changes associated with successive series of bases passing through the nanopore are resolved bioinformatically into a nucleotide sequence.

Applications of next-generation sequencing

The development of NGS and data from the human genome project made a new approach to the discovery of novel agents possible: Merkel cell carcinoma (MCC), a rare but aggressive cause of skin cancer that affects immunocompromised or elderly persons, was identified through digital transcriptome subtraction (DTS) (Feng et al., 2008). In short, RNA libraries from next-generation pyrosequencing were aligned with expected transcriptomes (the transcriptomes representing expected mRNA sequences from transcribed human genes from the National Center for Biotechnology Information (NCBI) databases). Of the remaining (subtracted) mRNA transcripts one was found to match a polyomavirus sequence. Further investigations found this Merkel Cell Polyomavirus (MCPyV) to be integrated in the large majority of MCC tissues (Feng et al., 2008).

NGS has offered a new metagenomic approach to diagnostics and pathogen discovery. Nonspecific DNA amplification involving transposase, used as library preparation, followed by NGS with Illumina, which uses solid-phase bridge amplification, allows the direct sequencing of low concentration DNA libraries (Marine et al., 2011). This unbiased approach was used to identify a novel arenavirus in cluster of three patients who died after receiving solid organ transplantation from the same donor (Palacios et al., 2008). It has also allowed for the unbiased diagnosis of infections in returning travelers (Jerome et al., 2019). These metagenomic approaches have been shown to correlate well with routine diagnostic methods such as multiplex PCR (Graf et al., 2016; Huang et al., 2019) but works best if the viral loads are sufficiently high (Huang et al., 2019; Kustin et al., 2019). Sensitivity may be increased by enriching first for viral sequences with probe capture before next-generation library preparation (O'Flaherty et al., 2018; Wylie et al., 2015, 2018).

Metagenomic NGS nevertheless holds several remaining challenges which hamper its widespread adoption. A challenge of NGS is the computational and bioinformatics needs to assemble sequence reads and match that to existing databases to identify known and possible novel pathogens (Dutilh et al., 2017). Metagenomic NGS also yield information, including viral genome sequences that are of unknown clinical significance, and which could be nonpathogenic, such as anelloviruses. Nevertheless, anellovirus viral loads have been used to monitor the degree of immunosuppression in organ transplant recipients (Wylie et al., 2018).

Another potential use of metagenomic NGS is to identify agents in nonhuman hosts or arthropod vectors before they emerge as human pathogens as part of a One Health Approach: Multiple emerging human pathogens may be of bat origin, such as filoviruses Ebolavirus and Marburg virus, paramyxoviruses (Hendra and Nipah viruses), and coronaviruses (SARS-CoV, MERS, and SAR-CoV-2). Therefore a metagenomics study of bat viruses may be able to identify viruses of possible zoonotic origin before these spread amongst humans (Geldenhuys et al., 2018).

It is clear that improvements in nucleic extraction, amplification, and sequencing have facilitated the discovery and monitoring of new pathogens. Fig. 1 summarizes a few highlights in this development process.

Developments in epidemiological evidence

Molecular epidemiology

Molecular epidemiology is the study of the genetic variation of infectious agents in human populations in terms of its occurrence: in space, by person (age, gender, ethnicity, socio-economic factors and health-related) and occurrence in time. Organisms are sampled to represent human populations and the genetic sequence characterized. Sequences are compared using phylogenetic analysis, which describes the genetic relatedness in phylogenetic trees and ancestral association between variants. This has become an increasingly valuable tool to characterize the distribution of emerging and re-emerging organisms and inform interventions to curb their spread (Holmes, 1998). Improvements in sequencing technology that facilitate rapid sequencing and computing power that allows for the fast and accurate construction of phylogenetic trees have made this increasingly practicable. A few of multiple cases where this has informed public health interventions are discussed below. The sequencing of organisms in different populations, studied in time, by geographic location and according to risk factor exposure has enabled the modeling of the origin of outbreaks, based on the phylogenetic relationships of organisms. The study of how epidemiological processes, immune responses, and evolution shape phylogenetic relationships has been called "phylodynamics." When an infection emerges or is reintroduced, phylodynamic analysis can help to establish the likely geographic origin and whether there are one or multiple origins. The overall diversity and "date stamped" sample collection could be used to calibrate molecular clock models, which can be used to estimate the date of emergence of the most recent common ancestor. This has been used to establish when particular infectious agents entered human populations. From such studies it emerged that HIV-1 has entered human populations likely before 1930 in central Africa (Gryseels et al., 2020; Worobey et al., 2008) and that circulating human coronaviruses infections such OC-43 and NL-63 are likely the result of historic species jumps from bats (Huynh et al., 2012; Vijgen et al., 2005).

Seroepidemiology and sero-archeology

When an organism is newly identified, it is important to establish whether the infection is truly new or just previously undiagnosed. Historic serum banks could help to determine whether an agent is truly new. This helped to determine that human metapneumovirus, an important respiratory infection in young children, although first identified in 2001, had been present in humans for at least 43 years before (Van Den Hoogen et al., 2001). In contrast, at the time when SARS CoV and SARS-CoV-2 were first detected, it was

apparent that there had been no pre-existing population immunity toward these novel agents (Ip et al., 2004; Takahashi et al., 2020). After SARS-CoV-2 had spread widely and had been classified as a pandemic, seroprevalence studies of SARS-CoV-2, when adjusted properly for diagnostic error, proved to be very useful in assessing population immunity, the risk for local epidemic spread, and the role of asymptomatic spread in transmission (Larremore et al., 2021).

Similarly, the strains involved in historic influenza pandemics have been studied by investigating birth cohorts and assessing the influenza strains to which different age groups have antibodies. The most important antibodies toward influenza are directed against surface proteins, hemagglutinin (H) and neuraminidase (N), which are used for influenza A typing. When pandemic influenza A strains share similar antigens to prior strains, the older strains are often displaced by the new pandemic strain. Human exposure to particular strains is assessed by performing hemagglutinin (H) inhibition and neutralizing antibody studies. Using this approach it was discovered that H3 influenza strains circulated before 1891 and were reintroduced during the "Hong Kong" pandemic of 1968 when influenza A H3N2 emerged (Reichert et al., 2012; Simonsen et al., 2004). During seasonal influenza epidemics, mortality is higher at the extremes of age, in infants and elderly individuals. However during some influenza pandemic years, older individuals showed lower excess mortality as they had been protected by prior exposure to previously circulating strains that had antigens in common with the emerging pandemic strain (Reichert et al., 2012).

Case studies: Identification and monitoring of important emerging and re-emerging pathogens

The technological development required for the detection and monitoring of infectious diseases involves genomics and immunology. The development of NGS with increased computer power and more robust bioinformatic algorithms and mathematical models of evolution have had a major role in detecting and monitoring emerging infections. This coupled with faster automated nucleic acid amplification technology and a robust biotechnology industry, which supplies reagents fast and have developed point of care and large high-throughput instruments, have enabled an increasingly rapid response to emerging pathogens. A few examples, where technology played a role in the detection and response to emerging or re-emerging infections, are discussed below, but there are multiple other examples, which cannot be covered in a single article.

SARS and other coronaviruses

Prior to 2002, the known human coronaviruses 229E and OC43 were associated with mild upper respiratory viral disease. The first time when a coronavirus was associated with severe pneumonia in humans was when an atypical pneumonia, later named Severe Acute Respiratory Syndrome (SARS), emerged in Guangdong Province, China in November 2002. This virus, called "severe acute respiratory syndrome coronavirus (SARS-CoV)" was identified using a combination of established and newer methods. Vero cell culture and electron microscopy on the culture supernatant identified viruses that resembled coronaviruses. Following viral cell culture isolation, different groups, respectively, used consensus PCR primers (designed based on an alignment of known coronaviruses) (Drosten et al., 2003) or degenerate/random primers and DNA microarrays (Wang et al., 2003) which enabled them to amplify and characterize the SARS-CoV genome using Sanger sequencing.

The identification of SARS-CoV has led to an increased awareness of the possible role of coronaviruses in respiratory infections. Subsequently the VIDISCA method, described above, was used to identify human coronavirus NL-63 (HCoV-NL-63) from a 7-month-old child with bronchiolitis and conjunctivitis. HCoV-NL-63 was thereafter found to be a common respiratory pathogen especially in children (Bastien et al., 2005; Ebihara et al., 2005; Kaiser et al., 2005). Thereafter another novel coronavirus, human coronavirus HKU1 (HCoV-HKU1), was identified from a nasopharyngeal aspirate of a 71-year-old man with pneumonia by using conserved coronavirus primers and thereafter using a mixture of conserved and degenerate primers to amplify the full genome and identify the sequence by automated Sanger sequencing (Woo et al., 2005). HCoV-HKU1 was later found in children and found to be occurring worldwide (Esper et al., 2006; Lau et al., 2006). In 2012, a novel coronavirus was isolated from a 60-year-old male patient from Saudi Arabia with pneumonia and renal failure. A pan-coronavirus PCR, based on the method used for the discovery of SARS-CoV (Drosten et al., 2003), was used to amplify the viral nucleic acid, followed by next-generation 454 pyrosequencing to identify sequences of this novel coronavirus. This virus was named Middle East Respiratory Syndrome Coronavirus (MERS-CoV) based on the geographical distribution of most cases. Human infections with the virus have a mortality of around 35%. Although the virus likely originated from bats, dromedary camels were found to be the reservoir host, with zoonotic infection of humans and household, and healthcare infections being important means of transmission (Killerby et al., 2020).

The identification and monitoring of SARS-CoV-2 and its impact on medical interventions

SARS-CoV-2, the etiology of COVID19, was identified using a combination of consensus PCR and NGS (Zhou et al., 2020) and separately by cloning and NGS and Oxford Nanopore Sequencing of viral nucleic acid from culture supernatant (Zhu et al., 2020). Phylogenetic analysis revealed that this virus belongs to the same subgenus of Betacoronavirus as SARS-CoV namely Sarbecovirus (Lu et al., 2020; Zhou et al., 2020).

Identification of genetic variants

SARS-CoV-2 NGS and phylogenetic and phylodynamic analysis allowed for the global tracing of SARS-CoV-2 variants and provided insights into its pandemic dissemination. The presence of multiple lineages at particular locations indicated multiple introductions (Alm et al., 2020; Candido et al., 2020; Tegally et al., 2021). Moreover, displacement of lineages indicated that particular variants having the G614 spike protein mutation were likely more infectious than other variants (Korber et al., 2020). Later data from the United Kingdom showed that a variant, B.1.1.7, with multiple distinguishable mutations had about 70% higher transmissibility than historic lineages (Kirby, 2021; Tang et al., 2021). Similarly, other new variants emerged: B.1.351 in South Africa (Makoni, 2021), P.1 in Brazil (Candido et al., 2020), and B.1.617 in India (Thiagarajan, 2021), which displaced previous variants. Sequencing of these variants detected additional spike protein mutations which raised concerns about vaccine escape and which may explain the circulation of some of these variants in regions that had a high level of population immunity prior to the emergence of these variants (Sabino et al., 2021). Subsequently several studies showed a lower ability of vaccine- or natural infection-inducted antibodies to neutralize the B.1.351 and P.1 variants, impacting on the effectiveness of vaccination (Fontanet et al., 2021; Kupferschmidt, 2021). Genomic surveillance has thus become important to understand the transmission dynamics of emerging infections and to inform public health interventions such as vaccination rollout. This has already had impact on the selection of SARS-CoV-2 vaccines that provide better protection against emerging variants. Moreover, vaccine constructs can be adapted to protect against vaccine escape variants. Whereas vaccines elicit a response to several epitopes and immune escape is most likely to be partial, antibody-based biologicals may be most susceptible to such escape mutations. These antibodies have been developed for putative treatment or prevention and act by neutralizing viral infection through binding to the viral receptor-binding domain. Therefore, any changes that impact the antibody-binding epitopes could be adversely affected by such mutations and may render these biologicals ineffective against the emerging variants. Nevertheless, a cocktail of different antibodies targeting separate epitopes may retain effectiveness against emerging variants (Baum et al., 2020).

West Nile virus

West Nile virus, from the genus *Flavivirus*, Japanese Encephalitis serocomplex, had been detected previously in Africa, Eurasia, Europe, and Australia, when it first emerged in New York City in 1999. During the New York outbreak, of 59 cases, who were hospitalized, seven died. West Nile virus is a mosquito-borne virus with birds as the primary hosts (Campbell et al., 2002). Since the New York outbreak, it spread across most of North America (Beasley et al., 2003; LaDeau et al., 2007). Genetic analysis revealed that the strain emerging in New York was most closely related to lineage I, previously detected in Israel (Beasley et al., 2003) but it remains unknown whether it reached the United States via an infected bird or mosquito. Nevertheless the rapid spread of West Nile virus in the Americas shows the risk posed by international trade or travel to introduce arthropod-borne viruses (also referred to as arboviruses) into new territories when susceptible vectors are already present in these regions.

Chikungunya

Chikungunya is a mosquito-borne alphavirus which was first isolated from a Tanzanian patient in 1953. It is associated with fever and arthralgia during the acute presentation and in some individuals arthralgia may persist for months to years. In Africa, the natural hosts are nonhuman primates, whereas human infections are thought to sustain outbreaks in densely populated Asian settings, although monkeys have been found to be infected in Malaysia. There are three viral genotypes: West African, East/Central/South African (ECSA), and Asian (Morrison, 2014). Since 2004 ECSA genotypes increased in geographic range and spread to Indian Ocean Islands, such that by the end of 2005 there were more than a million suspected cases in this region. Therefore, there was a concern that the ECSA genotype could spread to other tropical regions such as the Americas. Then in December 2013, the French National Reference Centre for arboviruses reported that Chikungunya had been detected in the Caribbean Island of Saint Martin. Subsequently it spread to many countries in South and Central America with almost half a million cases by July 2014. Surprisingly molecular epidemiological analysis revealed that the Asian genotype, and not the suspected ECSA genotype, had been responsible for transmission in the Americas (Leparc-Goffart et al., 2014; Morrison, 2014).

Zika virus

Zika virus, a flavivirus, transmitted by *Aedes aegypti* mosquitoes was first described in 1947 in Uganda and was associated with mild human infection in Africa and Asia in the 1960s to 1980s. The first signal of increased global spread was in 2007 when a large outbreak occurred on the island of Yap followed by outbreaks in French Polynesia in 2013 to 2014. Thereafter Zika virus was identified in the Americas with a large outbreak occurring in 2015 in northern Brazil (Grubaugh et al., 2018; Kindhauser et al., 2016). The large outbreaks in the Pacific and Americas provided strong epidemiological evidence, based on geography and timing, that Zika virus infections are associated with congenital microcephaly when women are infected during the first trimester (Brady et al., 2019; Vissoci et al., 2018). Further strong support was provided by a prospective case control study, enrolling 91 infant cases with microcephaly and 173 controls without microcephaly. Zika virus infection is diagnosed by PCR, in the acute phase, or specific neutralizing antibodies after recovery, as IgM antibodies could be cross reactive to other flaviviruses (Landry and St George, 2017). Thirty two (35%) of the cases with microcephaly had Zika virus infection, confirmed with RT-PCR or serology, vs none of the

controls (de Araújo et al., 2018). The rapid spread of Zika virus emphasized the risk of introduction of arthropod borne viruses (called arboviruses) into new geographic settings where their vectors are present, but where there is a lack of population immunity, contributing to rapid spread and a high incidence of human infections. Molecular clock analysis, based on the genetic diversity of the Brazil outbreak, showed that Zika virus had likely been introduced as early as 2013, suggesting that early vector control could have limited the catastrophic spread (Grubaugh et al., 2018). Monitoring of arboviral infections in humans and insect vectors is therefore crucial to allow early detection and timely vector control programs.

Ebola and other filovirus re-emergence: Tracing the source

Filoviruses are associated with systemic illness and viral hemorrhagic fever in humans. The first filovirus to be isolated was Marburg virus in 1967 in laboratory workers who processed blood and tissue samples from African green monkeys (Slenczka and Klenk, 2007). In 1976 two related filoviruses, associated with hemorrhagic fever viruses, were isolated: Ebolavirus (EBOV) in the Congo (then Zaire) and Sudan virus in South Sudan (then Sudan). Before 2013, outbreaks of filoviruses were localized and limited in size, with until then, the largest outbreak being in 2000, involving 425 cases with Sudan virus in Gulu, Uganda. However, from 2013 to 2016 an epidemic spanned Guinea, Liberia, and Sierra Leone, with 28,646 EVD cases and 11,323 deaths. The scale of the epidemic and the duration of sustained human transmission were unprecedented (Spengler et al., 2016). During this outbreak and subsequent monitoring of regional outbreaks, molecular epidemiology came to its own. By December 2013 only 29 Ebolavirus and 65 Marburg full length sequences had been published. Thereafter next-generation sequencing allowed the generation of cumulatively more than 800 complete filovirus genome sequences by 2020 (Di Paola et al., 2020). Molecular epidemiological data provided valuable information about the transmission dynamics and origin of outbreaks and allowed appropriate targeted medical interventions. Targeted EBOV sequencing and unbiased metagenomic sequencing helped to identify natural hosts of viruses from the Ebolavirus and Marburg virus genera. Phylogenetic analysis revealed transmission chains and the presence of super-spreading events, suggesting that targeting these events could dramatically reduce transmission. The persistence of EBOV in semen and the almost identical sequences identified in sexual partners helped to establish sexual transmission as a mechanism by which outbreaks can be reseeded after transmission cycles have apparently been interrupted. Persistence in semen and re-emergence, fueled by sexual transmission, provide a plausible explanation for the epidemic resurgence that had been observed during these outbreaks (Crozier, 2016; Di Paola et al., 2020; Thorson et al., 2021).

Tracing HIV: History and future

A syndrome of immune suppression, with an epidemiology suggestive of a transmissible agent, was first described in men who had sex with men (MSM) from Los Angeles, California and also found in hemophiliacs (Gottlieb et al., 1981, 1983). To distinguish this from primary immunodeficiency, this was named Acquired Immunodeficiency Syndrome (AIDS) (Gottlieb et al., 1983). Then in 1983, the causative virus, later named HIV-1, was isolated by Barré-Sinoussi et al. (1983). Unlike in the United States and Europe, spread in Africa was predominantly through heterosexual exposure and mother to child transmission, with a generalized epidemic and a high incidence, first observed in Central and East Africa (Quinn et al., 1986). Moreover, a serum sample from 1959 from Kinshasa tested positive, indicating a much earlier presence in Africa than the first diagnosis of AIDS (Nahmias et al., 1986). However, it was through the use of molecular clocks, calibrated with early blood samples from Central Africa, that it became clear that HIV likely first crossed the species barrier, and started to spread amongst humans, around the early part of the 20th century. This timing coincided with large colonial population disruptions, mining and urbanization in the then Belgian Congo (Gryseels et al., 2020; Worobey et al., 2008). By studying the relatedness of simian immunodeficiency viruses from chimpanzee subspecies, *Pan troglodytes troglodytes*, it became apparent that they were the source of this cross-species transmission (Gao et al., 1999). Evidence of other chimpanzee viruses in hunters from Central Africa further substantiated that the hunting and slaughtering of chimpanzees could have resulted in a cross-species jump (Calattini et al., 2007).

Whereas HIV-1 caused a global pandemic, the region of high HIV-2 prevalence is restricted to West Africa. Molecular clock analysis revealed that both epidemic HIV-2 subtypes, A and B, were likely transmitted from sooty mangabeys to humans in the first half of the 20th century, with exponential epidemic growth that coincided with the Guinea-Bissau independence war from 1963 to 1974 (Lemey et al., 2003).

Apart from the value of molecular epidemiology to study the history of the HIV-1 pandemic, identifying transmission clusters is an important tool in attempts to end HIV epidemics on country level (Han et al., 2020). Identifying transmission clusters and acting on these by providing additional support services form part of the United States strategy that aims to end HIV transmission by 2030 (Nichols and Kissler, 2020).

Re-emergent poliovirus

Polioviruses belong to Enterovirus C species and are associated with a risk of paralysis of about 1 in 200 to 1 in 1000. There are three poliovirus serotypes: 1, 2 and 3. In 1988 the World Health Assembly resolved to eradicate poliomyelitis (Forty-First World Health Assembly, 1988). Two types of vaccines are used to prevent polio: live oral polio vaccine (OPV) and inactivated polio vaccine (IPV). Although IPV provides effective individual protection against paralytic polio, administering OPV types 1, 2, and 3 (OPV-1, 2, and 3) provide better herd immunity and its use has been essential in limiting community spread and hence OPV use is essential for

poliovirus eradication. However, OPV strains pose a risk of reverting to virulence and could be transmitted onward, classified as circulating vaccine-derived polio virus (cVDPV), which can be differentiated by sequencing from wild-type poliovirus. Therefore the end stage of poliovirus eradication relies on the withdrawal of OPV as the source of cVDPV (Yan et al., 2021). In 2015 and 2019, respectively, wild-type poliovirus 2 and 3 were declared eliminated. Since fewer mutations are required for OPV-2 to revert to virulence than other types, cVDPV had mostly been due to OPV-2. Therefore, to curb the threat of cVDPV, bivalent OPV (bOPV), which lacks OPV-2, replaced trivalent OPV (tOPV) in May 2016. This approach was initially successful, as it reduced the source of cVDPV by the end of 2016. However, from 2017 to 2019, cVDPV started to rebound, as herd immunity toward poliovirus-2 declined. Paradoxically, OPV-2 withdrawal resulted in an increase in cVDPV derived from OPV-2 (cVDPV-2), which had continued to circulate in insufficiently immune populations. This has resulted in the reintroduction of OPV-2 in regions with cVDPV-2 to mop up such outbreaks (Alleman et al., 2020; Chard et al., 2020). Subsequently a new OPV-2 vaccine construct that is more genetically stable and less likely to revert to virulence has been investigated in clinical trials (De Coster et al., 2021). This vaccine may in future help to prevent cVDPV-2.

It is evident that molecular epidemiological investigations of poliovirus have been and would remain essential in targeting medical control measures. An effective eradication effort is reliant on the fast and accurate genotyping of poliovirus from patients and environmental sources and construction of transmission networks by geographical location and time. This is essential in establishing whether strains identified are wild-type or vaccine-derived and to trace the geographic sources of transmission chains. In addition, full genome sequencing can provide valuable information about the role of recombination of vaccine polioviruses with other Enterovirus C genotypes and its role in regaining virulence. Improvements in next-generation sequencing methods are likely to continue to contribute to the success of eradication efforts in providing critical knowledge about the molecular epidemiology of polioviruses (Jorgensen et al., 2020).

Tracing the source of food-borne outbreaks

Listeria monocytogenes (*L. monocytogenes)* is a gram-positive bacterium responsible for listeriosis in humans. It has the ability to survive in food production plants, when disinfection is insufficient (Gandhi and Chikindas, 2007). Whole genome sequencing of *L. monocytogenes* provides a more accurate classification of types than prior methods such as pulse field gel electrophoresis (PFGE). Moreover phylogenetic analysis could identify outbreak clusters and provide accurate molecular data to trace the source of an outbreak (Lüth et al., 2018).

Noroviruses are highly infectious, environmentally stable, and antigenically diverse feco-orally transmitted viruses that belong to the family *Caliciviridae* and are some of the most important causes of large gastroenteritis outbreaks associated with food vendors, hospitals, cruise ships, and large events. Molecular epidemiology has been useful in identifying the source and transmission dynamics of norovirus infections. Moreover monitoring global diversity and identifying conserved antigenic domains and antibodies that are cross-protective, between strains, may be useful for vaccine development efforts (Hasing and Pang, 2021).

Ribosomal RNA and metagenomics

16S rRNA PCR and sequencing have been used for bacterial metagenomics in environmental and biological samples and have enabled the diagnosis of novel bacteria (Woo et al., 2008), and increased the detection of known bacteria (Cummings et al., 2020), when not clinically suspected. It may be particularly useful when working with fastidious or slow growing organisms as it could provide a fast and unbiased diagnosis. This approach has also been used to detect unexpected bacterial species in brain abscesses (Al Masalma et al., 2009). However, cost and optimal workflow are current constraints. The faster result generation with ONT third-generation sequencing, compared to other next-generation sequencing methods, may increase the utility of 16S rRNA PCR and sequencing, especially if the costs of sequencing can be further reduced and if sequencing accuracy of these novel methods improves further in the near future (Heikema et al., 2020).

As ticks are important vectors of bacterial and viral disease, metagenomic microbiome analysis of ticks for potential bacterial pathogens is important in identifying the risks for human infection, even before human outbreaks might occur. PCR amplification and sequencing of 16s rRNA have been used to identify multiple bacteria belonging to the families *Rickettsiaceae*, *Enterobacteriaceae*, and *Pseudomonadaceae* (Lambert et al., 2019; Sperling et al., 2017). A similar approach also allows for the metagenomic identification of tick-borne bacteria in blood from patients (Kingry et al., 2020; Rodino et al., 2021).

Pan-fungal primers selected in the internal transcribed spacer (ITS) of ribosomal RNA are useful for fungal metagenomics in environmental and biological materials. Fungal infections of hospitalized patients have often been shown to have an environmental origin and identifying the source requires ecological sampling. Drug-resistant *Candida auris* has been associated with serious blood infections in hospitalized patients (Prestel et al., 2021), but the likely source has not been known. Recently, ITS primers for fungal metagenomics were used to identify *Candida auris* in coastal wetlands of the Andaman islands (Arora et al., 2021). It is not yet known, however, if these wetlands are the main natural reservoir of *Candida auris*.

Monitoring infectious agents for drug resistance

Next-generation sequencing could be applied to viruses, bacteria, or parasites and be especially useful to detect novel mechanisms of drug resistance. NGS genome sequencing (WGS) of *Plasmodium falciparum* was able to identify candidate genes involved in

artemisinin resistance, which is an essential drug in malaria treatment and prevention (Ariey et al., 2014). Despite some workflow challenges, NGS is also gaining traction for HIV drug resistance genotyping, as it is more sensitive for low frequency drug resistance variants (Ji et al., 2020). Although the prevalence of Mycobacterium tuberculosis (MTB) is low, in many industrialized countries, it is a re-emerging pathogen in impoverished and resource-limited settings and a high incidence has been seen in settings with a high HIV prevalence. The use of genome capture methods to concentrate MTB from sputum samples, which contain a lot of host DNA, improves efficiency of WGS of MTB (Brown et al., 2015). Whole genome sequencing of MTB has several important applications: it can be used to identify drug-resistant lineages (Oppong et al., 2019), understand the molecular epidemiology of MTB to enable targeted preventions, and investigate the molecular mechanisms of new drugs, which have been recently licensed for MTB treatment (Iketleng et al., 2018). Considering the multiple drug resistance mechanisms, WGS will likely become increasingly valuable for the rapid identification of drug resistance against second-line drugs in case of multiple drug-resistant (MDR) and extensively drug-resistant (XDR) MTB strains and to investigate drug resistance to novel MTB drugs (Faksri et al., 2019). The same principles apply to investigating phylogenetic relationships of gram-negative and gram-positive bacteria, associated with hospital associated drug resistant outbreaks, and could help to understand the transmission dynamics for the purpose of targeted interventions that would limit nosocomial spread of highly resistant bacteria (Gröschel et al., 2020; Mahmoud et al., 2020; Quainoo et al., 2017).

False findings: Associations with chronic fatigue, prostate cancer, and XMRV

The high sensitivity of molecular methods such as PCR that is based on amplification of DNA increases the risk of false findings. Investigations that aim to discover novel disease associations of infectious agents should ensure that there are appropriate controls that would detect contamination and that cases and controls are treated the same. A novel agent, Xenotropic MLV-related Virus (XMRV), was identified in 2006 and was soon associated with chronic fatigue syndrome (CFS) (Lombardi et al., 2009) and prostate cancer (Schlaberg et al., 2009). However these findings could not be reproduced by others, even when including the same patients; and there was evidence of contamination with identical MLV sequences (Kearney et al., 2012) likely originating from reagents used prior to PCR (Erlwein et al., 2011; Kearney et al., 2012). Similarly, contamination of cell cultures used in prostate cancer research with murine viruses was the likely source of the artefactual association of XMRV with prostate cancer (Das Gupta et al., 2012; Schlaberg et al., 2014).

Conclusion: Identifying and monitoring emerging infections now and in the future

The development of metagenomics methods that allow for the unbiased discovery of infections in humans, the environment, zoonotic hosts or vectors, has the potential to revolutionize pathogen discovery. Moreover improvements in sequencing technology with next-generation sequencing and more recently long-read third generation sequencing have decreased the time from the first identification to full characterization of genomes of newly discovered infectious agents. This is probably most evident in the rapid full characterization and international sharing of the full SARS-CoV-2 genome a few weeks after its discovery. Without such speed of SARS-CoV-2 characterization the fast development of diagnostic tests and the fastest ever development and approval of vaccines, within a single year, would not have been possible.

It has also became apparent that nonhuman mammal hosts are the source of many emerging infections. Molecular epidemiological studies and molecular clock models have shown that HIV originated from primates and that bats host many viruses of concern that have either historically crossed the species barrier or which in future could result in pandemic spread. It is therefore important that the monitoring for pandemic threats involves a One Health Approach, which combines veterinary and human health and which includes the investigation of arthropod vectors of infectious disease.

Globalization has resulted in a higher level of interconnectedness. Moreover high population density and urbanization have increased the risk of rapid spread of respiratory pathogens. Therefore, an infection that emerges in one continent can rapidly spread globally. Nevertheless, it also offers the opportunity for global collaboration in identifying agents of pandemic potential. Such global collaboration, which involves sharing information and technological developments, would remain essential in the rapid identification and response to current and future emerging infectious diseases and pandemic threats.

References

Al Masalma M, Armougom F, Michael Scheld W, Dufour H, Roche PH, Drancourt M, and Raoult D (2009) The expansion of the microbiological spectrum of brain abscesses With use of multiple 16S ribosomal DNA Sequencing. *Clinical Infectious Diseases* 48: 1169–1178. https://doi.org/10.1086/597578.

Allander T, Tammi MT, Eriksson M, Bjerkner A, Tiveljung-Lindell A, and Andersson B (2005) Cloning of a human parvovirus by molecular screening of respiratory tract samples. *Proceedings of the National Academy of Sciences of the United States of America* 102: 12891–12896. https://doi.org/10.1073/pnas.0504666102.

Alleman MM, Jorba J, Greene SA, Diop OM, Iber J, Tallis G, Goel A, Wiesen E, Wassilak SGF, and Burns CC (2020) Update on Vaccine-Derived Poliovirus Outbreaks — Worldwide, July 2019–February 2020. *Morbidity and Mortality Weekly Report* 69: 489–495. https://doi.org/10.15585/mmwr.mm6916a1.

Alm E, Broberg EK, Connor T, Hodcroft EB, Komissarov AB, Maurer-Stroh S, Melidou A, Neher RA, O'Toole Á, and Pereyaslov D (2020) Geographical and temporal distribution of SARS-CoV-2 clades in the WHO European Region, January to June 2020. *Eurosurveillance* 25: 2001410. https://doi.org/10.2807/1560-7917.ES.2020.25.32.2001410.

Alter HJ, Holland PV, Purcell RH, and Popper H (1978) Transmissible agent in non-A, non-B hepatitis. *Lancet* 311: 459–463. https://doi.org/10.1016/S0140-6736(78)90131-9.

Ariey F, Witkowski B, Amaratunga C, Beghain J, Langlois AC, Khim N, Kim S, Duru V, Bouchier C, Ma L, Lim P, Leang R, Duong S, Sreng S, Suon S, Chuor CM, Bout DM, Ménard S, Rogers WO, Genton B, Fandeur T, Miotto O, Ringwald P, Le Bras J, Berry A, Barale JC, Fairhurst RM, Benoit-Vical F, Mercereau-Puijalon O, and Ménard D (2014) A molecular marker of artemisinin-resistant *Plasmodium falciparum* malaria. *Nature* 505: 50–55. https://doi.org/10.1038/nature12876.

Arora P, Singh P, Wang Y, Yadav A, Pawar K, Singh A, Padmavati G, Xu J, and Chowdhary A (2021) Environmental Isolation of *Candida auris* from the Coastal Wetlands of Andaman Islands, India. *MBio* 12. https://doi.org/10.1128/mBio.03181-20.

Barré-Sinoussi F, Chermann JC, Rey F, Nugeyre MT, Chamaret S, Gruest J, Dauguet C, Axler-Blin C, Vézinet-Brun F, Rouzioux C, Rozenbaum W, and Montagnier L (1983) Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). *Science* 220: 868–871. https://doi.org/10.1126/science.6189183.

Bastien N, Robinson JL, Tse A, Lee BE, Hart L, and Li Y (2005) Human coronavirus NL-63 infections in children: A 1-year study. *Journal of Clinical Microbiology* 43: 4567–4573. https://doi.org/10.1128/JCM.43.9.4567-4573.2005.

Baum A, Fulton BO, Wloga E, Copin R, Pascal KE, Russo V, Giordano S, Lanza K, Negron N, Ni M, Wei Y, Atwal GS, Murphy AJ, Stahl N, Yancopoulos GD, and Kyratsous CA (2020) Antibody cocktail to SARS-CoV-2 spike protein prevents rapid mutational escape seen with individual antibodies. *Science* 369: 1014–1018. https://doi.org/10.1126/science.abd0831.

Baur D, Gladstone BP, Burkert F, Carrara E, Foschi F, Döbele S, and Tacconelli E (2017) Effect of antibiotic stewardship on the incidence of infection and colonisation with antibiotic-resistant bacteria and *Clostridium difficile* infection: A systematic review and meta-analysis. *The Lancet Infectious Diseases* 17: 990–1001. https://doi.org/10.1016/S1473-3099(17)30325-0.

Beasley DWC, Davis CT, Guzman H, Vanlandingham DL, Travassos Da Rosa APA, Parsons RE, Higgs S, Tesh RB, and Barrett ADT (2003) Limited evolution of West Nile virus has occurred during its southwesterly spread in the United States. *Virology* 309: 190–195. https://doi.org/10.1016/S0042-6822(03)00150-8.

Blevins SM and Bronze MS (2010) Robert Koch and the "golden age" of bacteriology. *International Journal of Infectious Diseases*. https://doi.org/10.1016/j.ijid.2009.12.003.

Brady OJ, Osgood-Zimmerman A, Kassebaum NJ, Ray SE, de Araújo VEM, da Nóbrega AA, Frutuoso LCV, Lecca RCR, Stevens A, Zoca de Oliveira B, de Lima JM, Bogoch Il, Mayaud P, Jaenisch T, Mokdad AH, Murray CJL, Hay SI, Reiner RC, and Marinho F (2019) The association between Zika virus infection and microcephaly in Brazil 2015–2017: An observational analysis of over 4 million births. *PLoS Medicine* 16: e1002755. https://doi.org/10.1371/journal.pmed.1002755.

Brown AC, Bryant JM, Einer-Jensen K, Holdstock J, Houniet DT, Chan JZM, Depledge DP, Nikolayevskyy V, Broda A, Stone MJ, Christiansen MT, Williams R, McAndrew MB, Tutill H, Brown J, Melzer M, Rosmarin C, McHugh TD, Shorten RJ, Drobniewski F, Speight G, and Breuer J (2015) Rapid whole-genome sequencing of mycobacterium tuberculosis isolates directly from clinical samples. *Journal of Clinical Microbiology* 53: 2230–2237. https://doi.org/10.1128/JCM.00486-15.

Byrd AL and Segre JA (2016) Adapting Koch's postulates. *Science* 351: 224–226. https://doi.org/10.1126/science.aad6753.

Calattini S, Betsem EBA, Froment A, Mauclère P, Tortevoye P, Schmitt C, Njouom R, Saib A, and Gessain A (2007) Simian foamy virus transmission from apes to humans, rural Cameroon. *Emerging Infectious Diseases* 13: 1314–1320. https://doi.org/10.3201/eid1309.061162.

Campbell GL, Marfin AA, Lanciotti RS, and Gubler DJ (2002) West Nile virus. *The Lancet Infectious Diseases*. https://doi.org/10.1016/S1473-3099(02)00368-7.

Candido DS, Claro IM, de Jesus JG, Souza WM, Moreira FRR, Dellicour S, Mellan TA, du Plessis L, Pereira RHM, Sales FCS, Manuli ER, Thézé J, Almeida L, Menezes MT, Voloch CM, Fumagalli MJ, Coletti TM, da Silva CAM, Ramundo MS, Amorim MR, Hoeltgebaum HH, Mishra S, Gill MS, Carvalho LM, Buss LF, Prete CA, Ashworth J, Nakaya HI, Peixoto PS, Brady OJ, Nicholls SM, Tanuri A, Rossi ÁD, Braga CKV, Gerber AL, de Guimarães APC, Gaburo N, Alencar CS, Ferreira ACS, Lima CX, Levi JE, Granato C, Ferreira GM, Francisco RS, Granja F, Garcia MT, Moretti ML, Perroud MW, Castiñeiras TMPP, Lazari CS, Hill SC, de Souza Santos AA, Simeoni CL, Forato J, Sposito AC, Schreiber AZ, Santos MNN, de Sá CZ, Souza RP, Resende-Moreira LC, Teixeira MM, Hubner J, Leme PAF, Moreira RG, Nogueira ML, Ferguson NM, Costa SF, Proenca-Modena JL, Vasconcelos ATR, Bhatt S, Lemey P, Wu CH, Rambaut A, Loman NJ, Aguiar RS, Pybus OG, Sabino EC, and Faria NR (2020) Evolution and epidemic spread of SARS-CoV-2 in Brazil. *Science* 369: 1255–1260. https://doi.org/10.1126/SCIENCE.ABD2161.

Cesarman E, Chang Y, Moore PS, Said JW, and Knowles DM (1995) Kaposi's sarcoma–associated herpesvirus-like DNA sequences in AIDS-related body-cavity–based lymphomas. *The New England Journal of Medicine* 332: 1186–1191. https://doi.org/10.1056/nejm199505043321802.

Chang Y, Cesarman E, Pessin MS, Lee F, Culpepper J, Knowles DM, and Moore PS (1994) Identification of herpesvirus-like DNA sequences in AIDS-associated Kaposi's sarcoma. *Science* 266: 1865–1869. https://doi.org/10.1126/science.7997879.

Chard AN, Datta SD, Tallis G, Burns CC, Wassilak SGF, Vertefeuille JF, and Zaffran M (2020) Progress Toward Polio Eradication — Worldwide, January 2018–March 2020. *MMWR. Morbidity and Mortality Weekly Report* 69: 784–789. https://doi.org/10.15585/mmwr.mm6925a4.

Choi IJ, Kook M-C, Kim Y-I, Cho S-J, Lee JY, Kim CG, Park B, and Nam B-H (2018) *Helicobacter pylori* therapy for the prevention of metachronous gastric cancer. *The New England Journal of Medicine* 378: 1085–1095. https://doi.org/10.1056/NEJMoa1708423.

Choo QL, Kuo G, Weiner AJ, Overby LR, Bradley DW, and Houghton M (1989) Isolation of a cDNA clone derived from a blood-borne non-A, non-B viral hepatitis genome. *Science* 244: 359–362. https://doi.org/10.1126/science.2523562.

Coombes R (2017) Europe steps up action against vaccine hesitancy as measles outbreaks continue. *BMJ*. https://doi.org/10.1136/bmj.j4803.

Crozier I (2016) Ebola virus RNA in the semen of male survivors of Ebola virus disease: The uncertain gravitas of a privileged persistence. *The Journal of Infectious Diseases* 214: 1467–1469. https://doi.org/10.1093/infdis/jiw079.

Cummings LA, Hoogestraat DR, Rassoulian-Barrett SL, Rosenthal CA, Salipante SJ, Cookson BT, and Hoffman NG (2020) Comprehensive evaluation of complex polymicrobial specimens using next generation sequencing and standard microbiological culture. *Scientific Reports* 10: 1–12. https://doi.org/10.1038/s41598-020-62424-x.

Curry A, Appleton H, and Dowsett B (2006) Application of transmission electron microscopy to the clinical study of viral and bacterial infections: Present and future. *Micron* 37(2): 91–106. https://doi.org/10.1016/j.micron.2005.10.001.

Das Gupta J, Luk K-C, Tang N, Gaughan C, Klein EA, Kandel ES, Hackett J, and Silverman RH (2012) Absence of XMRV and closely related viruses in primary prostate cancer tissues used to derive the XMRV-infected cell line 22Rv1. *PLoS One* 7: e36072. https://doi.org/10.1371/journal.pone.0036072.

de Araújo TVB, Ximenes RADA, Miranda-Filho DDB, Souza WV, Montarroyos UR, de Melo APL, Valongueiro S, de Albuquerque MDFPM, Braga C, Filho SPB, Cordeiro MT, Vazquez E, Cruz DDCS, Henriques CMP, Bezerra LCA, Castanha PMDS, Dhalia R, Marques-Júnior ETA, Martelli CMT, Rodrigues LC, Dhalia C, Santos M, Cortes F, Kleber de Oliveira W, Evelim Coelho G, Cortez-Escalante JJ, Campelo de Albuquerque de Melo CF, Ramon-Pardo P, Aldighieri S, Mendez-Rico J, Espinal M, Torres L, Nassri Hazin A, Van der Linden A, Coentro M, Santiago Dimech G, Siqueira de Assuncao R, Ismael de Carvalho P, and Felix Oliveira V (2018) Association between microcephaly, Zika virus infection, and other risk factors in Brazil: Final report of a case-control study. *The Lancet Infectious Diseases* 18: 328–336. https://doi.org/10.1016/S1473-3099(17)30727-2.

De Coster I, Leroux-Roels I, Bandyopadhyay AS, Gast C, Withanage K, Steenackers K, De Smedt P, Aerssens A, Leroux-Roels G, Oberste MS, Konopka-Anstadt JL, Weldon WC, Fix A, Konz J, Wahid R, Modlin J, Clemens R, Costa Clemens SA, Bachtiar NS, and Van Damme P (2021) Safety and immunogenicity of two novel type 2 oral poliovirus vaccine candidates compared with a monovalent type 2 oral poliovirus vaccine in healthy adults: Two clinical trials. *Lancet* 397: 39–50. https://doi.org/10.1016/S0140-6736(20)32541-1.

Di Paola N, Sanchez-Lockhart M, Zeng X, Kuhn JH, and Palacios G (2020) Viral genomics in Ebola virus research. *Nature Reviews. Microbiology*. https://doi.org/10.1038/s41579-020-0354-7.

Doern GV (2000) Detection of selected fastidious bacteria. *Clinical Infectious Diseases* 30: 166–173. https://doi.org/10.1086/313586.

Drosten C, Günther S, Preiser W, van der Werf S, Brodt H-R, Becker S, Rabenau H, Panning M, Kolesnikova L, Fouchier RAM, Berger A, Burguière A-M, Cinatl J, Eickmann M, Escriou N, Grywna K, Kramme S, Manuguerra J-C, Müller S, Rickerts V, Stürmer M, Vieth S, Klenk H-D, Osterhaus ADME, Schmitz H, and Doerr HW (2003) Identification of a novel coronavirus in patients with severe acute respiratory syndrome. *The New England Journal of Medicine* 348: 1967–1976. https://doi.org/10.1056/nejmoa030747.

Dutilh BE, Reyes A, Hall RJ, and Whiteson KL (2017) Editorial: Virus discovery by metagenomics: The (im)possibilities. *Frontiers in Microbiology* 8: 1710. https://doi.org/10.3389/fmicb.2017.01710.

Ebihara T, Endo R, Ma X, Ishiguro N, and Kikuta H (2005) Detection of human coronavirus NL63 in young children with bronchiolitis. *Journal of Medical Virology* 75: 463–465. https://doi.org/10.1002/jmv.20289.

Erlwein O, Robinson MJ, Dustan S, Weber J, Kaye S, and McClure MO (2011) DNA extraction columns contaminated with murine sequences. *PLoS One* 6: e23484. https://doi.org/10.1371/journal.pone.0023484.

Esper F, Weibel C, Ferguson D, Landry ML, and Kahn JS (2006) Coronavirus HKU1 infection in the United States. *Emerging Infectious Diseases* 12: 775–779. https://doi.org/10.3201/eid1205.051316.

Faksri K, Kaewprasert O, Ong RTH, Suriyaphol P, Prammananan T, Teo YY, Srilohasin P, and Chaiprasert A (2019) Comparisons of whole-genome sequencing and phenotypic drug susceptibility testing for *Mycobacterium tuberculosis* causing MDR-TB and XDR-TB in Thailand. *International Journal of Antimicrobial Agents* 54: 109–116. https://doi.org/10.1016/j.ijantimicag.2019.04.004.

Falkow S (1988) Molecular Koch's postulates applied to microbial pathogenicity. *Reviews of Infectious Diseases* 10(Suppl 2): S274–S276. https://doi.org/10.1093/cid/10.supplement_2.s274.

Falkow S (2004) Molecular Koch's postulates applied to bacterial pathogenicity—A personal recollection 15 years later. *Nature Reviews. Microbiology* 2: 67–72. https://doi.org/10.1038/nrmicro799.

Feng H, Shuda M, Chang Y, and Moore PS (2008) Clonal integration of a polyomavirus in human Merkel cell carcinoma. *Science* 319: 1096–1100. https://doi.org/10.1126/science.1152586.

Fontanet A, Autran B, Lina B, Kieny MP, Karim SSA, and Sridhar D (2021) SARS-CoV-2 variants and ending the COVID-19 pandemic. *Lancet* 397(10278): 952–954. https://doi.org/10.1016/s0140-6736(21)00370-6.

Forty-First World Health Assembly (1988) Global eradication of poliomyelitis by the year 2000.

Gandhi M and Chikindas ML (2007) Listeria: A foodborne pathogen that knows how to survive. *International Journal of Food Microbiology* 113(1): 1–15. https://doi.org/10.1016/j.ijfoodmicro.2006.07.008.

Gao F, Balles E, Robertson DL, Chen Y, Rodenburg CM, Michael SF, Cummins LB, Arthur LO, Peeters M, Shaw GM, Sharp PM, and Hahn BH (1999) Origin of HIV-1 in the chimpanzee Pan troglodytes troglodytes. *Nature* 397: 436–441. https://doi.org/10.1038/17130.

Geldenhuys M, Mortlock M, Weyer J, Bezuidt O, Seamark ECJ, Kearney T, Gleasner C, Erkkila TH, Cui H, and Markotter W (2018) A metagenomic viral discovery approach identifies potential zoonotic and novel mammalian viruses in Neoromicia bats within South Africa. *PLoS One* 13: e0194527. https://doi.org/10.1371/journal.pone.0194527.

Gill P and Ghaemi A (2008) Nucleic acid isothermal amplification technologies: A review. *Nucleosides, Nucleotides & Nucleic Acids* 27: 224–243. https://doi.org/10.1080/15257770701845204.

Gocke DJ and Howe C (1970) Rapid detection of Australia antigen by counterimmunoelectrophoresis. *Journal of Immunology* 104.

Gorzalski AJ, Tian H, Laverdure C, Morzunov S, Verma SC, VanHooser S, and Pandori MW (2020) High-throughput transcription-mediated amplification on the Hologic Panther is a highly sensitive method of detection for SARS-CoV-2. *Journal of Clinical Virology* 129: 104501. https://doi.org/10.1016/j.jcv.2020.104501.

Gottlieb MS, Schanker HM, Fan PT, Saxon A, Weisman JD, Pozalski I, et al. (1981) *Pneumocystis pneumonia*—Los Angeles. *MMWR* 30: 250–252.

Gottlieb MS, Groopman JE, Weinstein WM, Fahey JL, and Detels R (1983) The acquired immunodeficiency syndrome. *Annals of Internal Medicine* 99: 208–220. https://doi.org/10.7326/0003-4819-99-2-208.

Graf EH, Simmon KE, Tardif KD, Hymas W, Flygare S, Eilbeck K, Yandell M, and Schlaberg R (2016) Unbiased detection of respiratory viruses by use of RNA sequencing-based metagenomics: A systematic comparison to a commercial PCR panel. *Journal of Clinical Microbiology* 54: 1000–1007. https://doi.org/10.1128/JCM.03060-15.

Greub G, Sahli R, Brouillet R, and Jaton K (2016) Ten years of R&D and full automation in molecular diagnosis. *Future Microbiology* 11: 403–425. https://doi.org/10.2217/fmb.15.152.

Gröschel MI, Meehan CJ, Barilar I, Diricks M, Gonzaga A, Steglich M, Conchillo-Solé O, Scherer IC, Mamat U, Luz CF, De Bruyne K, Utpatel C, Yero D, Gibert I, Daura X, Kampmeier S, Rahman NA, Kresken M, van der Werf TS, Alio I, Streit WR, Zhou K, Schwartz T, Rossen JWA, Farhat MR, Schaible UE, Nübel U, Rupp J, Steinmann J, Niemann S, and Kohl TA (2020) The phylogenetic landscape and nosocomial spread of the multidrug-resistant opportunist *Stenotrophomonas maltophilia*. *Nature Communications* 11: 28. https://doi.org/10.1038/s41467-020-15123-0.

Grubaugh ND, Faria NR, Andersen KG, and Pybus OG (2018) Genomic insights into Zika virus emergence and spread. *Cell* 172(6): 1160–1162. https://doi.org/10.1016/j.cell.2018.02.027.

Gryseels S, Watts TD, Kabongo Mpolesha J-M, Larsen BB, Lemey P, Muyembe-Tamfum J-J, Teuwen DE, and Worobey M (2020) A near full-length HIV-1 genome from 1966 recovered from formalin-fixed paraffin-embedded tissue. *Proceedings of the National Academy of Sciences* 117: 12222–12229. https://doi.org/10.1073/pnas.1913682117.

Han X, Zhao B, An M, Zhong P, and Shang H (2020) Molecular network-based intervention brings us closer to ending the HIV pandemic. *Frontiers in Medicine* 14(2): 136–148. https://doi.org/10.1007/s11684-020-0756-y.

Hasing ME and Pang XL (2021) Norovirus: Molecular epidemiology, viral culture, immunity, and vaccines. *Clinical Microbiology Newsletter* 43: 33–43. https://doi.org/10.1016/j.clinmicnews.2021.02.002.

Heikema AP, Horst-Kreft D, Boers SA, Jansen R, Hiltemann SD, de Koning W, Kraaij R, de Ridder MAJ, van Houten CB, Bont LJ, Stubbs AP, and Hays JP (2020) Comparison of illumina versus nanopore 16S rRNA gene sequencing of the human nasal microbiota. *Genes (Basel)* 11: 1105. https://doi.org/10.3390/genes11091105.

Holmes EC (1998) Molecular epidemiology and evolution of emerging infectious diseases. *British Medical Bulletin* 54: 533–543. https://doi.org/10.1093/oxfordjournals.bmb.a011708.

Houghton M (2009) Discovery of the hepatitis C virus. *Liver International* 29: 82–88. https://doi.org/10.1111/j.1478-3231.2008.01925.x.

Huang B, Jennsion A, Whiley D, McMahon J, Hewitson G, Graham R, De Jong A, and Warrilow D (2019) Illumina sequencing of clinical samples for virus detection in a public health laboratory. *Scientific Reports* 9: 1–8. https://doi.org/10.1038/s41598-019-41830-w.

Huynh J, Li S, Yount B, Smith A, Sturges L, Olsen JC, Nagel J, Johnson JB, Agnihothram S, Gates JE, Frieman MB, Baric RS, and Donaldson EF (2012) Evidence supporting a zoonotic origin of human coronavirus strain NL63. *Journal of Virology* 86: 12816–12825. https://doi.org/10.1128/jvi.00906-12.

Iketleng T, Lessells R, Dlamini MT, Mogashoa T, Mupfumi L, Moyo S, Gaseitsiwe S, and de Oliveira T (2018) *Mycobacterium tuberculosis* next-generation whole genome sequencing: Opportunities and challenges. *Tuberculosis Research and Treatment* 2018: 1–8. https://doi.org/10.1155/2018/1298542.

Ip M, Chan PKS, Lee N, Wu A, Ng TKC, Chan L, Ng A, Kwan HM, Tsang L, Chu I, Cheung JL, Sung JJY, and Tam JS (2004) Seroprevalence of antibody to severe acute respiratory syndrome (SARS)-associated coronavirus among health care workers in SARS and non-SARS medical wards. *Clinical Infectious Diseases* 38: e116–e118. https://doi.org/10.1086/421019.

Jacobsen K and Koopman J (2005) The effects of socioeconomic development on worldwide hepatitis A virus seroprevalence patterns. *International Journal of Epidemiology* 34: 600–609. https://doi.org/10.1093/ije/dyi062.

Jerome H, Taylor C, Sreenu VB, Klymenko T, Filipe ADS, Jackson C, Davis C, Ashraf S, Wilson-Davies E, Jesudason N, Devine K, Harder L, Aitken C, Gunson R, and Thomson EC (2019) Metagenomic next-generation sequencing aids the diagnosis of viral infections in febrile returning travellers. *The Journal of Infection* 79: 383–388. https://doi.org/10.1016/j.jinf.2019.08.003.

Ji H, Sandstrom P, Paredes R, Harrigan PR, Brumme CJ, Avila Rios S, Noguera-Julian M, Parkin N, and Kantor R (2020) Are we ready for NGS HIV drug resistance testing? The Second "Winnipeg Consensus" Symposium. *Viruses* 12: 586. https://doi.org/10.3390/v12060586.

Jones MS, Kapoor A, Lukashov VV, Simmonds P, Hecht F, and Delwart E (2005) New DNA viruses identified in patients with acute viral infection syndrome. *Journal of Virology* 79: 8230–8236. https://doi.org/10.1128/jvi.79.13.8230-8236.2005.

Jorgensen D, Pons-Salort M, Shaw AG, and Grassly NC (2020) The role of genetic sequencing and analysis in the polio eradication programme. *Virus Evolution* 6. https://doi.org/10.1093/ve/veaa040.

Kaiser L, Regamey N, Roiha H, Deffernez C, and Frey U (2005) Human coronavirus NL63 associated with lower respiratory tract symptoms in early life. *The Pediatric Infectious Disease Journal* 24: 1015–1017. https://doi.org/10.1097/01.inf.0000183773.80217.12.

Karczewski KJ and Snyder MP (2018) Integrative omics for health and disease. *Nature Reviews. Genetics* 19: 299–310. https://doi.org/10.1038/nrg.2018.4.

Kearney MF, Spindler J, Wiegand A, Shao W, Anderson EM, Maldarelli F, Ruscetti FW, Mellors JW, Hughes SH, Le Grice SFJ, and Coffin JM (2012) Multiple sources of contamination in samples from patients reported to have XMRV infection. *PLoS One* 7: e30889. https://doi.org/10.1371/journal.pone.0030889.

Killerby ME, Biggs HM, Midgley CM, Gerber SI, and Watson JT (2020) Middle east respiratory syndrome coronavirus transmission. *Emerging Infectious Diseases* 26: 191–198. https://doi.org/10.3201/eid2602.190697.

Kindhauser MK, Allen T, Frank V, Santhana RS, and Dye C (2016) Zika: The origin and spread of a mosquito-borne virus. *Bulletin of the World Health Organization* 94(9): 675–686C. https://doi.org/10.2471/BLT.16.171082.

Kingry L, Sheldon S, Oatman S, Pritt B, Anacker M, Bjork J, Neitzel D, Strain A, Berry J, Sloan L, Respicio-Kingry L, Dietrich E, Bloch K, Moncayo A, Srinivasamoorthy G, Hu B, Hinckley A, Mead P, Kugeler K, and Petersena J (2020) Targeted metagenomics for clinical detection and discovery of bacterial tick-borne pathogens. *Journal of Clinical Microbiology* 58. https://doi.org/10.1128/JCM.00147-20.

Kirby T (2021) New variant of SARS-CoV-2 in UK causes surge of COVID-19. *The Lancet Respiratory Medicine* 9: e20–e21. https://doi.org/10.1016/s2213-2600(21)00005-9.

Kolykhalov AA, Agapov EV, Blight KJ, Mihalik K, Feinstone SM, and Rice CM (1997) Transmission of hepatitis C by intrahepatic inoculation with transcribed RNA. *Science* 277: 570–574. https://doi.org/10.1126/science.277.5325.570.

Korber B, Fischer WM, Gnanakaran S, Yoon H, Theiler J, Abfalterer W, Hengartner N, Giorgi EE, Bhattacharya T, Foley B, Hastie KM, Parker MD, Partridge DG, Evans CM, Freeman TM, de Silva TI, Angyal A, Brown RL, Carrilero L, Green LR, Groves DC, Johnson KJ, Keeley AJ, Lindsey BB, Parsons PJ, Raza M, Rowland-Jones S, Smith N, Tucker RM, Wang D, Wyles MD, McDanal C, Perez LG, Tang H, Moon-Walker A, Whelan SP, LaBranche CC, Saphire EO, and Montefiori DC (2020) Tracking changes in SARS-CoV-2 spike: Evidence that D614G increases infectivity of the COVID-19 virus. *Cell* 182: 812–827.e19. https://doi.org/10.1016/j.cell.2020.06.043.

Kupferschmidt K (2021) New mutations raise specter of "immune escape": SARS-CoV-2 variants found in Brazil and South Africa may evade human antibodies. *Science* 371: 329–330. https://doi.org/10.1126/science.371.6527.329.

Kustin T, Ling G, Sharabi S, Ram D, Friedman N, Zuckerman N, Bucris ED, Glatman-Freedman A, Stern A, and Mandelboim M (2019) A method to identify respiratory virus infections in clinical samples using next-generation sequencing. *Scientific Reports* 9: 1–8. https://doi.org/10.1038/s41598-018-37483-w.

LaDeau SL, Kilpatrick AM, and Marra PP (2007) West Nile virus emergence and large-scale declines of North American bird populations. *Nature* 447: 710–713. https://doi.org/10.1038/nature05829.

Lambert JS, Cook MJ, Healy JE, Murtagh R, Avramovic G, and Lee SH (2019) Metagenomic 16S rRNA gene sequencing survey of *Borrelia* species in Irish samples of *Ixodes ricinus* ticks. *PLoS One* 14: e0209881. https://doi.org/10.1371/journal.pone.0209881.

Landry ML and St George K (2017) Laboratory diagnosis of zika virus infection. *Archives of Pathology & Laboratory Medicine* 141: 60–67. https://doi.org/10.5858/arpa.2016-0406-SA.

Larremore DB, Fosdick BK, Bubar KM, Zhang S, Kissler SM, Metcalf CJE, Buckee C, and Grad YH (2021) Estimating SARS-CoV-2 seroprevalence and epidemiological parameters with uncertainty from serological surveys. *eLife* 10. https://doi.org/10.7554/eLife.64206.

Lau SKP, Woo PCY, Yip CCY, Tse H, Tsoi HW, Cheng VCC, Lee P, Tang BSF, Cheung CHY, Lee RA, So LY, Lau YL, Chan KH, and Yuen KY (2006) Coronavirus HKU1 and other coronavirus infections in Hong Kong. *Journal of Clinical Microbiology* 44: 2063–2071. https://doi.org/10.1128/JCM.02614-05.

Leland DS and Ginocchio CC (2007) Role of cell culture for virus detection in the age of technology. *Clinical Microbiology Reviews* 20(1): 49–78. https://doi.org/10.1128/CMR.00002-06.

Lemey P, Pybus OG, Bin W, Saksena NK, Salemi M, and Vandamme AM (2003) Tracing the origin and history of the HIV-2 epidemic. *Proceedings of the National Academy of Sciences of the United States of America* 100: 6588–6592. https://doi.org/10.1073/pnas.0936469100.

Leparc-Goffart I, Nougairede A, Cassadou S, Prat C, and De Lamballerie X (2014) Chikungunya in the Americas. *Lancet* 383(9916): 514. https://doi.org/10.1016/S0140-6736(14)60185-9.

Li J and Macdonald J (2015) Advances in isothermal amplification: Novel strategies inspired by biological processes. *Biosensors & Bioelectronics* 64: 196–211. https://doi.org/10.1016/j.bios.2014.08.069.

Lipkin WI and Firth C (2013) Viral surveillance and discovery. *Current Opinion in Virology* 3(2): 199–204. https://doi.org/10.1016/j.coviro.2013.03.010.

Lisitsyn N, Lisitsyn N, and Wigler M (1993) Cloning the differences between two complex genomes. *Science* 259: 946–951. https://doi.org/10.1126/science.8438152.

Lo NC and Hotez PJ (2017) Public health and economic consequences of vaccine hesitancy for measles in the United States. *JAMA Pediatrics* 171: 887–892. https://doi.org/10.1001/jamapediatrics.2017.1695.

Lombardi VC, Ruscetti FW, Das Gupta J, Pfost MA, Hagen KS, Peterson DL, Ruscetti SK, Bagni RK, Petrow-Sadowski C, Gold B, Dean M, Silverman RH, and Mikovits JA (2009) Detection of an infectious retrovirus, XMRV, in blood cells of patients with chronic fatigue syndrome. *Science* 326: 585–589. https://doi.org/10.1126/science.1179052.

Lu R, Zhao X, Li J, Niu P, Yang B, Wu H, Wang W, Song H, Huang B, Zhu N, Bi Y, Ma X, Zhan F, Wang L, Hu T, Zhou H, Hu Z, Zhou W, Zhao L, Chen J, Meng Y, Wang J, Lin Y, Yuan J, Xie Z, Ma J, Liu WJ, Wang D, Xu W, Holmes EC, Gao GF, Wu G, Chen W, Shi W, and Tan W (2020) Genomic characterisation and epidemiology of 2019 novel coronavirus: Implications for virus origins and receptor binding. *Lancet* 395: 565–574. https://doi.org/10.1016/S0140-6736(20)30251-8.

Lüth S, Kleta S, and Al Dahouk S (2018) Whole genome sequencing as a typing tool for foodborne pathogens like Listeria monocytogenes—The way towards global harmonisation and data exchange. *Trends in Food Science and Technology* 73: 67–75. https://doi.org/10.1016/j.tifs.2018.01.008.

Lwoff A (1957) The concept of virus. *Journal of General Microbiology* 17: 239–253. https://doi.org/10.1099/00221287-17-2-239.

Mahmoud AT, Salim MT, Ibrahem RA, Gabr A, and Halby HM (2020) Multiple drug resistance patterns in various phylogenetic groups of hospital-acquired uropathogenic *E. coli* isolated from cancer patients. *Antibiotics* 9: 108. https://doi.org/10.3390/antibiotics9030108.

Makoni M (2021) South Africa responds to new SARS-CoV-2 variant. *Lancet* 397: 267. https://doi.org/10.1016/S0140-6736(21)00144-6.

Marine R, Polson SW, Ravel J, Hatfull G, Russell D, Sullivan M, Syed F, Dumas M, and Wommack KE (2011) Evaluation of a transposase protocol for rapid generation of shotgun high-throughput sequencing libraries from nanogram quantities of DNA. *Applied and Environmental Microbiology* 77: 8071–8079. https://doi.org/10.1128/AEM.05610-11.

Maxam AM and Gilbert W (1977) A new method for sequencing DNA. *Proceedings. National Academy of Sciences. United States of America* 74: 560–564. https://doi.org/10.1073/pnas.74.2.560.

Mesri EA, Feitelson MA, and Munger K (2014) Human viral oncogenesis: A cancer hallmarks analysis. *Cell Host & Microbe* 15(3): 266–282. https://doi.org/10.1016/j.chom.2014.02.011.

Metcalf CJE, Farrar J, Cutts FT, Basta NE, Graham AL, Lessler J, Ferguson NM, Burke DS, and Grenfell BT (2016) Use of serological surveys to generate key insights into the changing global landscape of infectious disease. *Lancet* 388(10045): 728–730. https://doi.org/10.1016/S0140-6736(16)30164-7.

Morens DM, Folkers GK, and Fauci AS (2004) The challenge of emerging and re-emerging infectious diseases. *Nature* 430(6996): 242–249. https://doi.org/10.1038/nature02759.

Morrison TE (2014) Reemergence of Chikungunya Virus. *Journal of Virology* 88: 11644–11647. https://doi.org/10.1128/JVI.01432-14.

Morse SS (1995) Factors in the emergence of infectious diseases. *Emerging Infectious Diseases* 1: 7–15. https://doi.org/10.3201/eid0101.950102.

Mullis KB (1990) The unusual origin of the polymerase chain reaction. *Scientific American* 262: 56–65. https://doi.org/10.1038/scientificamerican0490-56.

Mullis K, Faloona F, Scharf S, Saiki R, Horn G, and Erlich H (1986) Specific enzymatic amplification of DNA in vitro: The polymerase chain reaction. *Cold Spring Harbor Symposia on Quantitative Biology* 51: 263–273. https://doi.org/10.1101/sqb.1986.051.01.032.

Nahmias AJ, Weiss J, Yao X, Lee F, Kodsi R, Schanfield M, Matthews T, Bolognesi D, Durack D, Motulsky A, Kanki P, and Essex M (1986) Evidence for human infection with an HTLV III/LAV-like virus in Central Africa, 1959. *Lancet* 327: 1279–1280. https://doi.org/10.1016/S0140-6736(86)91422-4.

Nichols BE and Kissler SM (2020) Ending the HIV epidemic in the USA. *The Lancet HIV* 7(7): e454–e455. https://doi.org/10.1016/S2352-3018(20)30051-5.

Nicolás I, Marimon L, Barnadas E, Saco A, Rodríguez-Carunchio L, Fusté P, Martí C, Rodriguez-Trujillo A, Torne A, del Pino M, and Ordi J (2019) HPV-negative tumors of the uterine cervix. *Modern Pathology* 32: 1189–1196. https://doi.org/10.1038/s41379-019-0249-1.

Nishizawa T, Okamoto H, Konishi K, Yoshizawa H, Miyakawa Y, and Mayumi M (1997) A novel DNA virus (TTV) associated with elevated transaminase levels in posttransfusion hepatitis of unknown etiology. *Biochemical and Biophysical Research Communications* 241: 92–97. https://doi.org/10.1006/bbrc.1997.7765.

O'Connor SM, Taylor CE, and Hughes JM (2006) Emerging infectious determinants of chronic diseases. *Emerging Infectious Diseases* 12(7): 1051–1057. https://doi.org/10.3201/eid1207.060037.

O'Flaherty BM, Li Y, Tao Y, Paden CR, Queen K, Zhang J, Dinwiddie DL, Gross SM, Schroth GP, and Tong S (2018) Comprehensive viral enrichment enables sensitive respiratory virus genomic identification and analysis by next generation sequencing. *Genome Research* 28: 869–877. https://doi.org/10.1101/gr.226316.117.

Opota O, Brouillet R, Greub G, and Jaton K (2020) Comparison of SARS-CoV-2 RT-PCR on a high-throughput molecular diagnostic platform and the cobas SARS-CoV-2 test for the diagnostic of COVID-19 on various clinical samples. *Pathogens and Diseases* 78: 61. https://doi.org/10.1093/femspd/ftaa061.

Oppong YEA, Phelan J, Perdigão J, MacHado D, Miranda A, Portugal I, Viveiros M, Clark TG, and Hibberd ML (2019) Genome-wide analysis of *Mycobacterium tuberculosis* polymorphisms reveals lineage-specific associations with drug resistance. *BMC Genomics* 20: 252. https://doi.org/10.1186/s12864-019-5615-3.

Palacios G, Druce J, Du L, Tran T, Birch C, Briese T, Conlan S, Quan P-L, Hui J, Marshall J, Simons JF, Egholm M, Paddock CD, Shieh W-J, Goldsmith CS, Zaki SR, Catton M, and Lipkin WI (2008) A new arenavirus in a cluster of fatal transplant-associated diseases. *The New England Journal of Medicine* 358: 991–998. https://doi.org/10.1056/NEJMoa073785.

Panagiotopoulos T, Antoniadou I, Valassi-Adam E, and Berger A (1999) Increase in congenital rubella occurrence after immunisation in Greece: Retrospective survey and systematic review. How does herd immunity work? *BMJ* 319: 1462–1467. https://doi.org/10.1136/bmj.319.7223.1462.

Pasteur L (1881) On the germ theory. *Science* 2: 420–422. https://doi.org/10.1126/science.os-2.63.420.

Preda A, Burada F, Soare C, Birca A, Moraru E, and Cruce M (2009) *Helicobacter pylori* infection and the development of gastric cancer. *Annals of the Romanian society for cell biology* 14: 212–214. https://doi.org/10.1097/00042737-200111000-00023.

Prestel C, Anderson E, Forsberg K, Lyman M, de Perio MA, Kuhar D, Edwards K, Rivera M, Shugart A, Walters M, and Dotson NQ (2021) Candida auris Outbreak in a COVID-19 Specialty Care Unit — Florida, July–August 2020. *MMWR. Morbidity and Mortality Weekly Report* 70: 56–57. https://doi.org/10.15585/mmwr.mm7002e3.

Prober JM, Trainor GL, Dam RJ, Hobbs FW, Robertson CW, Zagursky RJ, Cocuzza AJ, Jensen MA, and Baumeister K (1987) A system for rapid DNA sequencing with fluorescent chain-terminating dideoxynucleotides. *Science* 238: 336–341. https://doi.org/10.1126/science.2443975.

Quainoo S, Coolen JPMM, van Hijum SAFT, Huynen MA, Melchers WJG, Willem van Schaik E, and Wertheim HFL (2017) Whole-genome sequencing of bacterial pathogens: The future of nosocomial. *Clinical Microbiology Reviews* 30: 1015–1064. https://doi.org/10.1128/CMR.00016-17.

Quinn TC, Mann JM, Curran JW, and Piot P (1986) AIDS in Africa: An epidemiologic paradigm. *Science* 234: 955–963. https://doi.org/10.1126/science.3022379.

Raftery P, Condell O, Wasunna C, Kpaka J, Zwizwai R, Nuha M, Fallah M, Freeman M, Harris V, Miller M, Baller A, Massaquoi M, Katawera V, Saindon J, Bemah P, Hamblion E, Castle E, Williams D, Gasasira A, and Nyenswah T (2018) Establishing Ebola Virus Disease (EVD) diagnostics using GeneXpert technology at a mobile laboratory in Liberia: Impact on outbreak response, case management and laboratory systems strengthening. *PLoS Neglected Tropical Diseases* 12: e0006135. https://doi.org/10.1371/journal.pntd.0006135.

Reichert T, Chowell G, and McCullers JA (2012) The age distribution of mortality due to influenza: Pandemic and peri-pandemic. *BMC Medicine* 10: 1–15. https://doi.org/10.1186/1741-7015-10-162.

Reyes GR and Kim JP (1991) Sequence-independent, single-primer amplification (SISPA) of complex DNA populations. *Molecular and Cellular Probes* 5: 473–481. https://doi.org/10.1016/S0890-8508(05)80020-9.

Rodino KG, Wolf MJ, Sheldon S, Kingry LC, Petersen JM, Patel R, and Pritt BS (2021) Detection of tick-borne bacteria from whole blood using 16S ribosomal RNA gene PCR followed by next-generation sequencing. *Journal of Clinical Microbiology* 59. https://doi.org/10.1128/JCM.03129-20.

Sabino EC, Buss LF, Carvalho MPS, Prete CA, Crispim MAE, Fraiji NA, Pereira RHM, Parag KV, da Silva Peixoto P, Kraemer MUG, Oikawa MK, Salomon T, Cucunuba ZM, Castro MC, de Souza Santos AA, Nascimento VH, Pereira HS, Ferguson NM, Pybus OG, Kucharski A, Busch MP, Dye C, and Faria NR (2021) Resurgence of COVID-19 in Manaus, Brazil, despite high seroprevalence. *Lancet* 397(10273): 452–455. https://doi.org/10.1016/S0140-6736(21)00183-5.

Sanger F, Nicklen S, and Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences of the United States of America* 74: 5463–5467. https://doi.org/10.1073/pnas.74.12.5463.

Schlaberg R, Choe DJ, Brown KR, Thaker HM, and Singh IR (2009) XMRV is present in malignant prostatic epithelium and is associated with prostate cancer, especially high-grade tumors. *Proceedings of the National Academy of Sciences of the United States of America* 106: 16351–16356. https://doi.org/10.1073/pnas.0906922106.

Schlaberg R, Choe DJ, Brown KR, Thaker HM, and Singh IR (2014) Erratum: XMRV is present in malignant prostatic epithelium and is associated with prostate cancer, especially high-grade tumors (*Proceedings of the National Academy of Sciences of the United States of America* (2009) 106(38): 16351–16356. DOI: 10.1073/pnas.0906922106). *Proceedings of the National Academy of Sciences of the United States of America*. https://doi.org/10.1073/pnas.1409186111.

Simons JN, Pilot-Matias TJ, Leary TP, Dawson GJ, Desai SM, Schlauder GG, Muerhoff AS, Erker JC, Buijk SL, Chalmers ML, Van Sant CL, and Mushahwar IK (1995) Identification of two flavivirus-like genomes in the GB hepatitis agent. *Proceedings of the National Academy of Sciences of the United States of America* 92: 3401–3405. https://doi.org/10.1073/pnas.92.8.3401.

Simonsen L, Reichert TA, and Miller MA (2004) The virtues of antigenic sin: Consequences of pandemic recycling on influenza-associated mortality. *International Congress Series* 1263: 791–794. https://doi.org/10.1016/j.ics.2004.01.029.

Slenczka W and Klenk HD (2007) Forty years of Marburg virus. *The Journal of Infectious Diseases* 196: S131–S135. https://doi.org/10.1086/520551.

Soulier J, Grollet L, Oksenhendler E, Cacoub P, Cazals-Hatem D, Babinet P, D'Agay MF, Clauvel JP, Raphael M, Degos L, and Sigaux F (1995) Kaposi's sarcoma-associated herpesvirus-like DNA sequences in multicentric Castleman's disease. *Blood* 86: 1276–1280. https://doi.org/10.1182/blood.v86.4.1276.bloodjournal8641276.

Spengler JR, Ervin ED, Towner JS, Rollin PE, and Nichol ST (2016) Perspectives on West Africa Ebola virus disease outbreak, 2013–2016. *Emerging Infectious Diseases* 22: 956–963. https://doi.org/10.3201/eid2206.160021.

Sperling JL, Silva-Brandão KL, Brandão MM, Lloyd VK, Dang S, Davis CS, Sperling FAH, and Magor KE (2017) Comparison of bacterial 16S rRNA variable regions for microbiome surveys of ticks. *Ticks Tick-Borne Diseases* 8: 453–461. https://doi.org/10.1016/j.ttbdis.2017.02.002.

Stanford JL (1973) Immunodiffusion analysis. A rational basis for the taxonomy of mycobacteria. *Annales de la Société Belge de Médecine Tropicale* 53: 321–330.

Takahashi S, Greenhouse B, and Rodríguez-Barraquer I (2020) Are seroprevalence estimates for severe acute respiratory syndrome coronavirus 2 biased? *The Journal of Infectious Diseases*. https://doi.org/10.1093/infdis/jiaa523.

Tang JW, Tambyah PA, and Hui DS (2021) Emergence of a new SARS-CoV-2 variant in the UK. *The Journal of Infection*. https://doi.org/10.1016/j.jinf.2020.12.024.

Tegally H, Wilkinson E, Lessells RJ, Giandhari J, Pillay S, Msomi N, Mlisana K, Bhiman JN, von Gottberg A, Walaza S, Fonseca V, Allam M, Ismail A, Glass AJ, Engelbrecht S, Van Zyl G, Preiser W, Williamson C, Petruccione F, Sigal A, Gazy I, Hardie D, Hsiao N-Y, Martin D, York D, Goedhals D, San EJ, Giovanetti M, Lourenço J, Alcantara LCJ, and de Oliveira T (2021) Sixteen novel lineages of SARS-CoV-2 in South Africa. *Nature Medicine* 1–7. https://doi.org/10.1038/s41591-021-01255-3.

Thermes C (2014) Ten years of next-generation sequencing technology. *Trends in Genetics* 30: 418–426. https://doi.org/10.1016/j.tig.2014.07.001.

Thiagarajan K (2021) Why is India having a covid-19 surge? *BMJ* 373: n1124. https://doi.org/10.1136/bmj.n1124.

Thorson AE, Deen GF, Bernstein KT, Liu WJ, Yamba F, Habib N, Sesay FR, Gaillard P, Massaquoi TA, McDonald SLR, Zhang Y, Durski KN, Singaravelu S, Ervin E, Liu H, Coursier A, Marrinan JE, Ariyarajah A, Carino M, Formenty P, Ströher U, Lamunu M, Wu G, Sahr F, Xu W, Knust B, and Broutet N (2021) Persistence of Ebola virus in semen among Ebola virus disease survivors in Sierra Leone: A cohort study of frequency, duration, and risk factors. *PLoS Medicine* 18: e1003273. https://doi.org/10.1371/journal.pmed.1003273.

Van Den Hoogen BG, De Jong JC, Groen J, Kuiken T, De Groot R, Fouchier RAM, and Osterhaus ADME (2001) A newly discovered human pneumovirus isolated from young children with respiratory tract disease. *Nature Medicine* 7: 719–724. https://doi.org/10.1038/89098.

Van Der Hoek L, Pyrc K, Jebbink MF, Vermeulen-Oost W, Berkhout RJM, Wolthers KC, Wertheim-Van Dillen PME, Kaandorp J, Spaargaren J, and Berkhout B (2004) Identification of a new human coronavirus. *Nature Medicine* 10: 368–373. https://doi.org/10.1038/nm1024.

Vermeulen M, Lelie N, Sykes W, Crookes R, Swanevelder J, Gaggia L, Le Roux M, Kuun E, Gulube S, and Reddy R (2009) Impact of individual-donation nucleic acid testing on risk of human immunodeficiency virus, hepatitis B virus, and hepatitis C virus transmission by blood transfusion in South Africa. *Transfusion* 49: 1115–1125. https://doi.org/10.1111/j.1537-2995.2009.02110.x.

Vijgen L, Keyaerts E, Moës E, Thoelen I, Wollants E, Lemey P, Vandamme A-M, and Van Ranst M (2005) Complete genomic sequence of human Coronavirus OC43: Molecular clock analysis suggests a relatively recent zoonotic coronavirus transmission event. *Journal of Virology* 79: 1595–1604. https://doi.org/10.1128/jvi.79.3.1595-1604.2005.

Vissoci JRN, Rocha TAH, da Silva NC, de Sousa Queiroz RC, Thomaz EBAF, Amaral PVM, Lein A, Branco MDRFC, Aquino J, Rodrigues ZMR, da Silva AAM, and Staton C (2018) Zika virus infection and microcephaly: Evidence regarding geospatial associations. *PLoS Neglected Tropical Diseases* 12: e0006392. https://doi.org/10.1371/journal.pntd.0006392.

von Bubnoff A (2008) Next-generation sequencing: The race is on. *Cell* 132(5): 721–723. https://doi.org/10.1016/j.cell.2008.02.028.

Wang D, Urisman A, Liu Y-T, Springer M, Ksiazek TG, Erdman DD, Mardis ER, Hickenbotham M, Magrini V, Eldred J, Latreille JP, Wilson RK, Ganem D, and DeRisi JL (2003) Viral discovery and sequence recovery using DNA microarrays. *PLoS Biology* 1: e2. https://doi.org/10.1371/journal.pbio.0000002.

Woo PCY, Lau SKP, Chu C, Chan K, Tsoi H, Huang Y, Wong BHL, Poon RWS, Cai JJ, Luk W, Poon LLM, Wong SSY, Guan Y, Peiris JSM, and Yuen K (2005) Characterization and complete genome sequence of a novel coronavirus, coronavirus HKU1, from patients with pneumonia. *Journal of Virology* 79: 884–895. https://doi.org/10.1128/jvi.79.2.884-895.2005.

Woo PCY, Lau SKP, Teng JLL, Tse H, and Yuen KY (2008) Then and now: Use of 16S rDNA gene sequencing for bacterial identification and discovery of novel bacteria in clinical microbiology laboratories. *Clinical Microbiology and Infection* 14(10): 908–934. https://doi.org/10.1111/j.1469-0691.2008.02070.x.

Worobey M, Gemmel M, Teuwen DE, Haselkorn T, Kunstman K, Bunce M, Muyembe JJ, Kabongo JMM, Kalengayi RM, Van Marck E, Gilbert MTP, and Wolinsky SM (2008) Direct evidence of extensive diversity of HIV-1 in Kinshasa by 1960. *Nature* 455: 661–664. https://doi.org/10.1038/nature07390.

Wylie TN, Wylie KM, Herter BN, and Storch GA (2015) Enhanced virome sequencing using targeted sequence capture. *Genome Research* 25: 1910–1920. https://doi.org/10.1101/gr.191049.115.

Wylie KM, Wylie TN, Buller R, Herter B, Cannella MT, and Storch GA (2018) Detection of viruses in clinical samples by use of metagenomic sequencing and targeted sequence capture. *Journal of Clinical Microbiology* 56. https://doi.org/10.1128/JCM.01123-18.

Yan D, Wang D, Zhang Y, Li X, Tang H, Guan J, Song Y, Zhu S, and Xu W (2021) Implication of a high risk for type 2 vaccine-derived poliovirus emergence and transmission after the switch from trivalent to bivalent oral poliovirus vaccine. *The Journal of Infectious Diseases* 223: 113–118. https://doi.org/10.1093/infdis/jiaa386.

Zhen W, Smith E, Manji R, Schron D, and Berry GJ (2020) Clinical evaluation of three sample-to-answer platforms for detection of sars-cov-2. *Journal of Clinical Microbiology* 58. https://doi.org/10.1128/JCM.00783-20.

Zhou P, Yang XL, Wang XG, Hu B, Zhang L, Zhang W, Si HR, Zhu Y, Li B, Huang CL, Chen HD, Chen J, Luo Y, Guo H, Di Jiang R, Liu MQ, Chen Y, Shen XR, Wang X, Zheng XS, Zhao K, Chen QJ, Deng F, Liu LL, Yan B, Zhan FX, Wang YY, Xiao GF, and Shi ZL (2020) A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* 579: 270–273. https://doi.org/10.1038/s41586-020-2012-7.

Zhu N, Zhang D, Wang W, Li X, Yang B, Song J, Zhao X, Huang B, Shi W, Lu R, Niu P, Zhan F, Ma X, Wang D, Xu W, Wu G, Gao GF, and Tan W (2020) A novel coronavirus from patients with pneumonia in China, 2019. *The New England Journal of Medicine* 382: 727–733. https://doi.org/10.1056/nejmoa2001017.