Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Pathogenic viruses: Molecular detection and characterization

I Made Artika\textsuperscript{a,b}, Ageng Wiyatno\textsuperscript{c}, Chairin Nisa Ma'roef\textsuperscript{d}

\textsuperscript{a} Biosafety Level 3 Unit, Eijkman Institute for Molecular Biology, Jalan Diponegoro 69, Jakarta 10430, Indonesia
\textsuperscript{b} Department of Biochemistry, Faculty of Mathematics and Natural Sciences, Bogor Agricultural University, Darmaga Campus, Bogor 16680, Indonesia
\textsuperscript{c} Emerging Viruse Research Unit, Eijkman Institute for Molecular Biology, Jalan Diponegoro 69, Jakarta 10430, Indonesia

\textbf{ARTICLE INFO}

\textbf{Keywords:}
Pathogenic viruses
Virus detection
Polymerase chain reaction
DNA sequencing
Unidirectional workflow

\textbf{ABSTRACT}

Pathogenic viruses are viruses that can infect and replicate within human cells and cause diseases. The continuous emergence and re-emergence of pathogenic viruses has become a major threat to public health. Whenever pathogenic viruses emerge, their rapid detection is critical to enable implementation of specific control measures and the limitation of virus spread. Further molecular characterization to better understand these viruses is required for the development of diagnostic tests and countermeasures. Advances in molecular biology techniques have revolutionized the procedures for detection and characterization of pathogenic viruses. The development of PCR-based techniques together with DNA sequencing technology, have provided highly sensitive and specific methods to determine virus circulation. Pathogenic viruses potentially having global catastrophic consequences may emerge in regions where capacity for their detection and characterization is limited. Development of a local capacity to rapidly identify new viruses is therefore critical. This article reviews the molecular biology of pathogenic viruses and the basic principles of molecular techniques commonly used for their detection and characterization. The principles of good laboratory practices for handling pathogenic viruses are also discussed. This review aims at providing researchers and laboratory personnel with an overview of the molecular biology of pathogenic viruses and the principles of molecular techniques and good laboratory practices commonly implemented for their detection and characterization.

1. Emergence of viral diseases

The emergence of infectious diseases has been a threat to public health and global stability. Historically, emerging infectious diseases have caused the deadliest catastrophic pandemics such as the 1918 influenza pandemic (claiming about 50 million lives), the HIV/AIDS pandemic (claiming about 35 million lives so far), etc. \citep{Morens2013}. Emerging infectious diseases are defined as infections whose incidence in humans has increased within the past two decades or threaten to increase in the years to come. The disease emergence can be caused by the spread of a new pathogen, or by the reappearance (or re-emergence) of a known pathogen after a decline in infection \citep{vanDoorn2014}. Biological, social and environmental factors have been linked to the emergence of infectious diseases. These include changes of the pathogens through evolution, changes in the way human populations interact with each other, and with their environment. In addition, increased susceptibility to infection, increased ease of international travel, climate and weather changes, have also been associated with new diseases emergence \citep{vanDoorn2014}. One of the major agents responsible for causing emerging infectious diseases is the virus. Pathogenic viruses that cause emerging diseases are called emerging viruses \citep{Nii-Trebi2017}.

Cross-species transmission plays an important role in the emergence of viral diseases \citep{Parrish2008}. The majority of the emerging viruses are zoonotic in that, they can be transmitted to humans from animals \citep{Morens2013}. Wildlife, especially mammals and birds, are hosts to a large number of viruses. Most of the mammalian hosts belong to the orders of Chiroptera, Rodentia, Primates and Carnivora. Mammals and birds play an important role as virus reservoirs within which the viruses propagate without causing disease \citep{Mackenzie2013}. Occasionally these viruses can infect other species including humans. Zoonotic viral infections which cause human diseases may come from domestic animals, poultry, livestock and wildlife. When a virus jumps from animal to human, the phenomenon is called zoonotic spillover which potentially has serious human health and economic impacts \citep{Wang2014,Plotwright2015}. The high risks of viral host switching are highlighted by the emergence of SARS coronavirus (CoV), hantaviruses, Ebola and Marburg viruses, Nipah virus, Hendra virus, and human immunodeficiency virus (HIV) \citep{Parrish2008}. Most zoonotic viruses are not readily transmitted...
among humans. In this context, humans can be called dead-end hosts (Causey and Edwards, 2008).

Bats (order Chiroptera) are important reservoir hosts of emerging viruses. As many as 66 different viruses have been isolated or detected in bats and some of them can cause diseases in humans. Some important characteristics of bats have been linked to their roles in the maintenance and transmission of viruses. Bats are unique mammals in that they have the ability to fly. Many species of bats can travel and migrate to extremely long distances which may contribute to viral transmission. Their long life span has been hypothesized to facilitate viral persistence. Many species of bats live in colonies of very high density that has been associated with their ability to efficiently transmit viruses between individuals in close proximity (Calisher et al., 2006). The diversity of the virus in bats is very high. All known types of viral genomic structures and replication strategies have been discovered in bats. Bats are hosts of a range of emerging zoonotic viruses such as Rabies virus, Ebola virus, Marburg virus, Hendra virus, Nipah virus, SARS-CoV, MERS-CoV, etc. which cause severe diseases in humans (Wang and Cramer, 2014; Hayman, 2016). Similar to bats, rodents are also natural reservoirs of a number of life-threatening zoonotic viruses. Comparative quantitative analysis has suggested that bats host more zoonotic viral infections per species than rodents. However, the total number of viruses identified in rodents is higher than that in bats indicating that rodents are important reservoirs of emerging viruses (Luis et al., 2013). Information on the types of virus circulating can also be obtained by detecting the viruses in these animal reservoirs (Mackenzie and Jeggo, 2013).

Birds play important roles in the emergence of zoonotic viruses, either as a reservoir host, or in the spread of novel pathogenic viruses to new areas. Wild birds are known to be reservoirs for several viruses, including the West Nile virus (WNV), Eastern equine encephalitis virus (EEE), St. Louis encephalitis virus (SLE), Western equine encephalitis virus (WEE), influenza A virus, etc. The importance of birds as reservoirs of pathogenic viruses is highlighted by the fact that birds are natural reservoirs of all of the pathogenic avian influenza A virus. Aquatic waterfowl are asymptomatic carriers of all subtypes of influenza A virus (Reed et al., 2003). Influenza A viruses infections outside birds such as in pigs, horses, seals, mink and dogs have also been reported. It is thought that the infection and transmission of avian influenza viruses in these so-called intermediate hosts facilitate the subsequent transfer influenza A virus to humans in a form that can easily be transmitted (Parrish et al., 2015). Birds are also central to the outbreak of West Nile virus because they are the main amplifying host of the virus in nature (Reed et al., 2003; Michel et al., 2018). Furthermore, bird migration provides a mechanism for transmission and spread of pathogenic viruses to new and previously unaffected areas. Some species of birds travel many kilometers each day searching for food or mates, and some species migrate thousands of kilometers each year. The migration pattern of wild birds, for example, has been suggested to serve as transmission pathways of avian influenza virus and may play a role in the spread of the virus among domestic birds (Causey and Edwards, 2008). A better understanding of bird migration patterns and viral diseases of birds would be important to predict future outbreaks of infections caused by emerging zoonotic viruses (Reed et al., 2003). Monitoring studies in wild birds can be used as an early warning system for the incursion of a range of zoonotic viruses (Michel et al., 2018). Different molecular methods, including metagenomics approaches, have been developed and applied for detection and characterization of viruses in birds (Vibin et al., 2018).

A subset of pathogenic viruses requires arthropods, mainly mosquitoes and ticks, for their transmission between vertebrates. These are termed arthropod-borne viruses (arboviruses). The arthropods which facilitate the efficient transmission of viruses between susceptible hosts are termed vectors. The effectiveness of a particular arthropod as a vector depends on its vector competence and vectorial capacity. Vector competence is the inherent ability of a particular arthropod to transmit a given virus. In a competent vector, the virus is acquired during feeding, replicates in the gut and salivary glands, is transferred into salivary secretion, and then ultimately introduced into another host during subsequent feeding. Vectorial capacity is the power of a particular arthropod as a driver of virus emergence which depends on the amount of time between uptake of virus and ability to transmit (Rückert and Ebel, 2018). The most important vectors of arboviruses are mosquitoes. Controlling the spread of arboviruses is challenging. Several arboviruses pathogenic to humans such as Yellow fever virus (YFV), Chikungunya virus (CHIKV) and Zika virus (ZIKV) share the same mosquito vector, Aedes aegypti (Rückert and Ebel, 2018). Furthermore, a virus, for example ZIKV, can be transmitted by mosquitoes of different species such as Ae. aegypti, Ae. albopictus and Culex quinquefasciatus, which are abundant throughout tropical and subtropical regions (Vieira et al., 2019). To be efficiently transmitted by mosquitoes, a virus must successfully infect, replicate, and disseminate in the mosquito’s cells and be able to evade the mosquito antiviral immune responses. Due to increased travel and trade, combined with other factors, the distribution of several mosquito-borne viruses such as Dengue virus (DENV), WNV, CHIKV and ZIKV has expanded, and the viruses are now regarded as the new global pathogens (Rückert and Ebel, 2018). Information on the types of arboviruses circulating in particular area can also be gained by detecting the viruses in the mosquito vectors (Grubaugh et al., 2013; Scheck et al., 2018).

2. Molecular biology of pathogenic viruses

A virus is a small infectious agent that can only multiply inside living organisms by directing the host cell’s machinery to generate more virus. The genetic material of a virus is either RNA or DNA. A virus that has RNA as its genetic material is called a RNA virus, while that having DNA as its genetic material and replicates using a DNA-dependent DNA polymerase is termed a DNA virus (Fig. 1). The term virion refers to the entire virus particle, which typically consists of a molecule of DNA or RNA within a protein coat, and sometimes covered by an envelope (Lodish, 2000). Viruses that can infect and replicate within human cells and cause disease are called pathogenic viruses. Emerging and re-emerging viral diseases have recently attracted worldwide attention. Ebola, zika, H5N1 avian influenza, severe acute respiratory syndrome (SARS), and many other emerging viral diseases have proved fatal and caused worldwide concern (World Health Organization (WHO), 2018; Bloom et al., 2017; Ogden et al., 2017). One of the critical events linked to the emergence of new viral diseases is viral genetic change that allows viruses to become established in a new host species, able to infect humans and cause a local, regional or worldwide outbreak (Pekosz and Glass, 2008). A number of driving factors, such as the increase of the global population, travel, urbanization, land use change, climate change, encroachment of people with the natural environment, extraction industries, etc. have been linked to the capacity for virus emergence (Mackenzie and Jeggo, 2013; Rosenberg et al., 2013; Bloom et al., 2017). In addition, rapid aging of the world population is considered to increase vulnerability to viral diseases (Bloom et al., 2017). Common attributes of emerging viral diseases include unpredictability, high morbidity, and potential for the rapid spread of diseases which may lead to substantial social impacts. The recent Ebola crisis and the outbreak of Zika virus have indicated that the world is unprepared to address emerging viral diseases (Bloom et al., 2017).

Among several high-impact pathogenic viruses which emerged or re-emerged recently include Ebola virus, avian influenza virus, SARS coronavirus and Zika virus. Ebola virus (EBOV) was first detected in 1976. It is one of deadliest re-emerging pathogenic viruses and is the etiologic agent of Ebolavirus disease which is characterized by sudden onset of fever with a case fatality rate which varies from 25% to 90%. The 2014–2016 outbreak in West Africa was the most devastating Ebola outbreak since this virus discovery that resulted in more than 28,000 cases and claimed more than 11,000 lives (Falasca et al., 2015; World Health Organization (WHO), 2018; World Health Organization (WHO), 2018; Bloom et al., 2017; Ogden et al., 2017).
respectively, strong inhibitors of interferon production and signaling for the high level of virulence of Ebola virus. The two proteins are, 2016). The VP35 and VP24 proteins have been indicated to be critical to interfere with the host immune system (Rivera and Messaoudi, 2016). The introduction of the virus into the human population is believed to occur through close contact with the blood, secretions, organs or other bodily fluids of infected fruit bats and other animals such as non-human primates, the forest antelope and porcupines found in the rainforest. Human-to-human transmission is then responsible for spreading the virus. This can occur via direct contact with the organs and bodily fluids of infected people and via direct contact with contaminated surfaces and materials (Rivera and Messaoudi, 2016; Baseler et al., 2017). Molecular methods based on PCR amplification of Ebola genome fragments have been developed and implemented for detection of Ebola virus. The genes targeted in these assays include the L, GP, NP, VP35, VP40, genes (Broadhurst et al., 2016; Ro et al., 2017).

The ultimate source of Ebola viruses in nature and the mechanisms by which they emerge to infect humans have been a mystery that hinders the development of prevention and control measures (Baseler et al., 2017). According to the World Health Organization (WHO), fruit bats of the Pteropodidae family are thought to be the natural hosts of the Ebola virus. The introduction of the virus into the human population is believed to occur through close contact with the blood, secretions, organs or other bodily fluids of infected fruit bats and other animals such as non-human primates, the forest antelope and porcupines found in the rainforest. Human-to-human transmission is then responsible for spreading the virus. This can occur via direct contact with the organs and bodily fluids of infected people and via direct contact with contaminated surfaces and materials (World Health Organization (WHO), 2018). The virus gains entry into the human body through breaks in the skin or through the mucosal surface and then binds to the host’s receptors. The GP protein on the viral envelope plays important roles in mediating receptor binding and fusion. Following GP binding to its receptor, the virions enter the cell by endocytosis. The acidification of endocytic vesicles is followed by a fusion of virus and host membranes and the release of the viral RNA genome into the host cell cytoplasm. The high virulence of EBOV is attributed to the ability of the virus to interfere with the host immune system (Rivera and Messaoudi, 2016). The VP35 and VP24 proteins have been indicated to be critical for the high level of virulence of Ebola virus. The two proteins are, respectively, strong inhibitors of interferon production and signaling (Baseler et al., 2017).

Avian influenza virus (H5N1) is a subtype of the influenza A virus that can cause illness in humans. Up to January 2020, WHO had reported 861 laboratory-confirmed cases of human infection with H5N1, 455 of which were fatal, resulting in a mortality rate of 53% (World Health Organization (WHO), 2020). The genome of this virus consists of 8 gene segments encoding 11 viral proteins. At the molecular level, several genes and gene products have been identified which may be related to the high pathogenicity of H5N1 influenza viruses. The hemagglutinin (HA) protein is a sialidase responsible for cleaving the HA of the progeny virions from receptors on the surface of host cells, promoting the release of the new virions. NA is the target of the neuraminidase inhibitor, oseltamivir. It has been reported that H274Y substitution of the NA protein confers resistance to oseltamivir. The three viral polymerase proteins, PB1, PB2, and PA, are involved in viral RNA synthesis. It was found out that the amino acid at position 627 of PB2 determines the level of virulence of H5N1 human isolates in mice. Glutamic acid at this position confers low pathogenicity, whereas lysine at this position confers high pathogenicity (Korteweg and Gu, 2008). For molecular detection of H5N1 the common target genes include the matrix gene (M), HA and NA genes (World Health Organization (WHO), 2011).

SARS is one of the emerging viral diseases caused by a coronavirus (SARS-CoV) that continues to pose a threat to public health. The SARS epidemic in 2003 resulted in more than 800 deaths. Since the future re-emergence of SARS cannot be excluded, intensive studies to understand the biology of the SARS-CoV were carried out. It was found that the genome of the virus is a positive-sense single-stranded RNA molecule consisting of 29,751 bases. It contains 14 functional open reading frames: (Satija and Lal, 2007; McBride and Fielding, 2012) two of them (the replicase gene, with two open reading frames (ORF) 1a and 1b), encode proteins needed for viral RNA synthesis, 4 open reading frame structural proteins (spike, membrane, nucleocapsid and envelope) and the other 8 encode accessory proteins. Due to extensive translational and enzymatic processing, the viral genome expression within the host
cells forms 4 structural, 8 accessory and 16 nonstructural proteins (Chow et al., 2003; Satija and Lal, 2007). The detailed molecular basis for the very high virulence of SARS-CoV has yet to be elucidated. The spike (S) glycoprotein (Frieman et al., 2011) and the envelope (E) protein (DeDiego et al., 2014) have been indicated to play roles as virulence determinants. The main target organ of SARS-CoV is the lung (Ye et al., 2007) although multiple organ infection has also been indicated (Gu et al., 2005). The virus may enter cells as an endosome in a protease-dependent fashion (Nagata et al., 2010). Apart from this, it has been postulated that the virus may gain entry into the target cells by direct membrane fusion at the target cell surface and by a pH-dependent endocytosis (Satija and Lal, 2007). The SARS-CoV spike protein is thought to play a critical role in this process. Angiotensin-converting enzyme 2 (ACE2) has been identified as the functional receptor for SARS-CoV. It is proposed that the virus infects epithelial cells through interaction of the spike protein and ACE, and then multiplies in these cells to cause intercellular transmission (Ye et al., 2007). The SARS-CoV RNA targeted in the molecular detection of SARS-CoV include the ORF 1b region of the 5′-replicase gene, nucleocapsid (nuc) gene, and the 3′-noncoding region (NCR) (Mahony and Richardson, 2005).

Zika virus (ZIKV) is an emerging mosquito-borne flavivirus that causes Zika viral disease in human. The recent outbreak of the disease and the wide spread of the virus has caused worldwide fear and triggered scientific effort to study this virus in attempt to accelerate the development of intervention strategies aimed to prevent and treat Zika disease (Wang et al., 2017). The ZIKV genome is single-stranded positive-sense RNA harboring a single open reading frame. The open reading frame encodes a polyprotein which is cleaved into 3 structural proteins (capsid, precursor membrane, and envelope proteins) and 7 nonstructural proteins (NS1, NS2A, NS2B, NS3, NS4A, NS4B, and NS5) (Li et al., 2017). Zika virus can cause neurological disorders. The recent Zika virus epidemic in Brazil has been linked to fetal microcephaly. The virus enters host cells through receptor-mediated endocytosis. The virus entry and replication stimulates the host antiviral and proinflammatory responses, apoptosis, and disrupts regulation of glial cell proliferation (Wang et al., 2017). The N-linked glycosylation of the envelope (E) protein and the precursor membrane (prM) protein (Nambala and Su, 2018) have been indicated to be important determinants of the Zika virus virulence and neuroinvasion (Annamalai et al., 2017). Although ZIKV has been known as an arthropod-borne virus with Aedes species mosquitoes as the primary vector, multiple routes of ZIKV transmission have been identified. It has been reported that Zika virus can also be transmitted from human to human through sexual contact, blood transfusion, and vertical transmission from mother to fetus (Gregory et al., 2017; Wang et al., 2017). The Zika virus causes similar symptoms to those caused by Dengue and Chikungunya viruses. Therefore, a reliable assay, such as molecular detection of viral nucleic acid, is required to detect Zika virus in infected people. The genomics domain commonly targeted for molecular detection of this virus include the membrane (M), envelope (E), nonstructural proteins (NS) 1, NS2B, NS3 and NS5 genes (Corman et al., 2016; Perkasa et al., 2016).

Emerging virus outbreaks are difficult to predict as they are the result of complex interaction between host, vector, virus, and environment (de Jong et al., 2018). Southeast Asia is considered to be especially vulnerable to potential outbreaks of pathogenic viruses and a number of pathogenic viruses have been reported to circulate in this region. These include Nipah virus (NIV) (Looi and Chua, 2007), H5N1 (Adisasmito et al., 2013), Dengue virus (Saumono et al., 2015; Dhenni et al., 2018), Rabies virus (Susilawati et al., 2012), West Nile virus (Myint et al., 2014), Chikungunya virus (Kosash et al., 2013; Riswari et al., 2016), Zika virus (Perkasa et al., 2016), Coxackievirus virus (Wiyatno et al., 2016), Measles virus (Hartoyo et al., 2017), Rhinovirus C (Wiyatno et al., 2018), Japanese encephalitis virus (Gao et al., 2013), Seoul virus (Hofmann et al., 2018) etc. The 2019 novel coronavirus (2019-nCoV) which emerged in Wuhan, Hubei, China, has also been identified in this region (Huang et al., 2020). The Indonesian archipelago, in particular, has been regarded as one of the hotspots for the emergence of pathogenic viruses. Indonesia has been diagnosed with the highest number of avian influenza cases in Southeast Asia. Until January 2020, Indonesia had 200 cases of human infection with the avian influenza A(H5N1) virus, 168 of which were fatal, resulting in a case fatality rate of 84% (168/200) which is much higher than the world case fatality rate of 53% (World Health Organization (WHO), 2020). In addition, in this archipelago, West Nile virus was detected for the first time in 2014 (Myint et al., 2014) and Zika virus circulation was confirmed in 2016 (Perkasa et al., 2016).

3. Polymerase chain reaction

Rapid and accurate detection of emerging viruses is essential for rapid response, optimized clinical care and to limit the spread of these viruses. Ideally, a diagnostic test needs to be rapid, cheap, accurate, and applicable in remote settings (Powers and Waterman, 2017; Chertow, 2018). In the past, diagnosis of pathogenic viruses was based on virus isolation and serology. However, these methods have some drawbacks. Virus isolation is expensive and laborious. Serology tests lack the sensitivity and specificity required for detection of viruses at a low level. Dramatic advances in molecular methods have revolutionized the detection and characterization of emerging viruses. Molecular methods are methods which are commonly employed in molecular biology studies and other disciplines of biological sciences dealing with manipulation and analysis of nucleic acids and proteins. Currently, molecular methods find wide applications in the diagnosis and research of pathogenic viruses. One of the molecular methods widely used in the detection and identification of pathogenic viruses is the polymerase chain reaction (PCR).

PCR is an in vitro enzymatic process used to amplify a single, or a few copies, of DNA across several orders of magnitude, resulting in thousands to millions of copies of a specific DNA fragment. It is a powerful technique that has become one of the most widely used techniques in molecular biology because it is quick, inexpensive and simple (Joshi and Deshpande, 2011). Ingredients needed for PCR assay include template DNA, primers, nucleotides, and thermostable DNA polymerase. The DNA polymerase is the key enzyme responsible for linking individual nucleotides together to form the PCR product. The nucleotides constitute four bases, adenine (A), thymine (T), cytosine (C), and guanine (G) that act as the building blocks used by the DNA polymerase to synthesize the PCR product. The primer is a short piece of single-stranded DNA (generally about 18–22 bases) with a defined sequence complementary to the target DNA that is to be detected and amplified. A number of factors need to be considered in designing primers. The size of the primers is optimized to be long enough for adequate specificity and short enough for the primer to bind easily to the template at annealing temperature. During the PCR process, the primers function as starting points for DNA synthesis. They are required because the DNA polymerases can only add new nucleotides to an existing strand of DNA. The pairing of primers in the reaction (forward and reverse primers) specify the exact DNA fragment to be amplified (Garibyan and Avashia, 2013).

Good primer design and optimized PCR conditions are essential for a successful reaction. For example, in a study to characterize human influenza A(H3N2) circulating in Indonesia, to obtain complete coding sequence of HA and NA gene of the virus, specific primer sets had to be developed and further tested using Indonesian H3N2 virus samples (Agustiningish et al., 2016). In addition, in developing PCR-based assays for detection of pathogenic viruses, it is also of very important to take into account the high genetic diversity, especially among the RNA viruses – such as those from the Flaviviridae and Bunyaviridae families – and the potential of cross-reactivity among viruses with a close relative. To address these challenges, a bioinformatics workflow based on two consecutive basic local alignment search tool for nucleotide (BLASTn) steps has been developed and evaluated. This approach
allows rapid selection of highly conserved and specific genomic regions among the complete taxon of viruses investigated to be used as candidate dates for molecular markers from which real-time PCR primer sets are generated (Schneeberger et al., 2017). Similarly, a formulation of universal primer reagent based on a mixture of 10, 18–22-nucleotide PCR primers (5 primer pair groups) was developed for rapid and specific detection of many variants of Dengue viruses with high diagnostic sensitivity. The primer sequences of this so-called primer cocktail were computed such that their probability of mispriming with the host genome background, the human DNA, is extremely low. Initial experimental testing showed that the primers amplified the genome of Dengue viruses and not that of the human. In addition, based on computational testing against 291 strains of other non-dengue flaviviruses, the Dengue virus specificity of the primers were predicted to be very good. Specificity among DENV serotypes was also found to be very good. The primer cocktail was predicted by electronic-PCR to have an ability to detect 95% of the 1688 dengue strains (listed in the Broad Institute Dengue Virus Database) with perfect primer match. With reduced stringency which allows one mismatch and one insertion per primer, it was predicted that the primer cocktail is able to detect 99% of the Dengue virus strains analyzed. This strategy therefore provides a single-PCR diagnostic tool with broad diagnostic sensitivity across Dengue virus strains and serotypes (Gijavanekar et al., 2011). Diagnostic sensitivity, also called true-positive rates, is defined as the percentage of samples with confirmed infection (by the ‘gold standard’ method) which will have positive results. Other parameters commonly used to measure reliability of a diagnostic assay are specificity, accuracy and precision. Specificity (also called the true negative rate) is the percentage of infection-free samples which will have a negative result. This qualitative assessment shows the capability of the test to distinguish target virus from the non-target viruses. Accuracy is defined as the percentage of correct results which describes how close the obtained results are to those obtained with the reference method. Precision refers to the reproducibility of the assay on the same sample in obtaining similar results (Soul, 2016).

The basic principle of PCR is simple. The process involves three major steps: denaturation, annealing, and extension. In the denaturation step, the two strands of the DNA double helix are physically separated at a high temperature (90 °C to 97 °C). The double strands of DNA are held together by hydrogen bonds, and this step is intended to break the hydrogen bonds between all complementary base pairs. In the annealing step, the temperature is lowered in order to provide optimal temperature for the primers to anneal to the DNA template strands to prime the extension process. The two DNA strands function as templates which allow DNA polymerase to selectively amplify the target DNA. In the extension step, the mixture is heated to 72 °C, the optimal temperature for DNA polymerase activity. This step allows extension to occur at the end of the annealed primers to generate a complementary copy strand of DNA. The quantity of the DNA is effectively doubled through the extension steps in each PCR cycle (Joshi and Deshpande, 2011; Garibyan and Avashia, 2013).

For detecting pathogenic viruses using PCR, several steps are involved. Firstly, the genetic material of the viruses is extracted, isolated and purified from the specimens. Types of specimens commonly collected are those which contain the maximum amount of the target virus (Johnson, 1990), for example, whole blood and oral swabs are collected and used for detection of Ebola virus (World Health Organization, 2014); posterior-pharyngeal (throat) swabs or nasal swabs for detection of avian influenza A(H5N1) virus (World Health Organization, 2006); respiratory, stool, and rectal swab (Chan et al., 2004a), sputum (Wang et al., 2005), urine, serum, plasma and peripheral blood leukocytes specimens (Mahony and Richardson, 2005) for detection of SARS-CoV; serum (Perkasa et al., 2016), plasma, urine (Corman et al., 2016), whole blood, amniotic fluid, cerebrospinal fluid (CSF), organ tissues such as placenta, fetal brain, eyes (Landry and George, 2017) for detection of Zika virus; saliva, CSF, tears, skin biopsy, urine for detection of rabies virus (Fooks et al., 2009), serum for detection of Coxackievirus (Wiyatno et al., 2016), West Nile virus (Myint et al., 2014), Dengue virus (Fahri et al., 2013), blood for detection of Chikungunya virus (Riswari et al., 2016). etc.

In addition to specimen type, successful molecular detection of pathogenic viruses is dependent upon the time of sample collection, quality and proper handling of the specimen. The probability of a successful detection is increased when the time interval between specimen collection and processing is not prolonged. It is important to note that RNA viruses in particular are generally unstable and easily inactivated by extreme heat. It is generally recommended that clinical specimens for virus detection and isolation should be immediately placed in a refrigerator with as little delay as possible. To maintain virus integrity and infectivity, specimens should be placed in a good transport medium and protected from thermal inactivation (Johnson, 1990; Killian, 2008). For molecular detection of Ebola virus, for instance, specimens should ideally be taken when a patient exhibits symptoms. Specimens can be stored at room temperature for up to 24 h. If specimens can only be transported and tested within one week, they should be stored between 0 and 5 °C for testing by RT-PCR. For long term storage, specimens should be stored at −70 °C and freeze-thaw cycles should be avoided. For shipment, clinical specimens should be properly packaged, labeled, marked and documented (World Health Organization (WHO), 2014). For avian influenza A (H5N1) molecular detection, a throat swab should be taken within three days of onset of symptoms. The virus is generally detectable in throat swabs from the point of onset of symptoms. Blood serum or plasma should be taken during the first 7 days after the development of symptoms because by that time the patient is most likely to have detectable RNA in the bloodstream. The swab specimens should be put into a viral transport medium, maintained with refrigeration (2–8 °C) and transported to the laboratory as soon as possible. The specimens should be stored at −70 °C or in liquid nitrogen and repeated freezing and thawing of specimens must be avoided to prevent loss of infectivity (World Health Organization (WHO), 2006; Killian, 2008). For detection of SARS-CoV, the collected specimens were refrigerated at about 10 °C until delivery, usually on the same day. Respiratory, stool, and rectal swab specimens were collected in viral transport medium, and urine samples were transported in sterile containers. During delivery, specimens were kept in ice boxes. Molecular analysis is carried out on fresh specimens without prior freezing and thawing (Chan et al., 2004a). Similarly, for molecular detection of Zika virus, specimens should be kept refrigerated at 2–8 °C and tested within 48 h. All types of specimens may be kept frozen at −20 °C for up to one week. For storage longer than one week, specimens should be frozen at −70 °C. Repeated freezing and thawing of specimens should be avoided (World Health Organization (WHO), 2016).

In order to minimize viral exposure which may cause viral infections to laboratory workers, adherence to recommended biosafety guidelines during diagnostic testing of pathogenic viruses is critical (Artika and Maroef, 2017). For example, all work with infectious Ebola viruses should be performed within a BSL4 facility (Ro et al., 2017). For handling specimens suspected of containing avian influenza A virus (H5N1), the WHO recommended the use a BSL2 containment and use BSL3 work practices such as always wearing gloves, goggles, eye protection, etc. Any procedure that may generate aerosols or droplets should be performed in a biological safety cabinet (World Health Organization (WHO), 2005). Avian influenza A virus propagation must be carried out in a BSL3 facility (Artika and Maroef, 2017). For SARS, testing of respiratory or stool specimens must be done at least using BSL2 level precautions in a class II biological safety cabinet. Laboratory workers must wear the required personal protective equipment including long sleeved gloves, goggles, eye protection, and N95 masks. Isolation of SARS-CoV should be carried out within a BSL3 facility (Mahony and Richardson, 2005). Specimens from individuals suspected of having Zika virus infection should be handled in a BSL2 facility. A
risk assessment should be performed to determine if certain procedures or specimens require higher levels of biocontainment. Similarly, potential aerosol generating activities should be performed within a biosafety cabinet (Chosewood and Wilson, 2009; CDC, 2019). A critical step in processing specimens containing pathogenic viruses is viral inactivation which renders specimens noninfectious hence allowing subsequent procedures to be carried out at a lower biosafety level and improves laboratory workflow and throughput. Chemical inactivation using TRIZol LS reagent is one of the common methods for viral inactivation (Broadhurst et al., 2016; Ro et al., 2017).

In the case of DNA virus detection, the isolated viral DNA is directly used as a template for PCR amplification. For detection of RNA viruses, however, the viral RNA molecule must first be reverse transcribed into a complementary DNA (cDNA) because RNA molecule is not an efficient substrate for the Taq DNA polymerase commonly used in PCR. This is carried out using reverse transcriptase, an RNA-dependent DNA polymerase which can synthesize cDNA using RNA as template. The generated single stranded cDNA is then converted by DNA polymerase into a double stranded DNA and then the double stranded DNA is further amplified using standard PCR. This method is termed reverse transcription polymerase chain reaction (RT-PCR) that can be carried out in either a one step or a two step procedure (Fig. 2). In one-step RT-PCR, the RT step is coupled with PCR. All the components including specific primers are put into a single tube as with the PCR reaction. In this process, mRNA undergoes reverse transcription to cDNA and then amplification in a single reaction. In a two-step procedure, as the name implies, the process is carried out in two steps. The first step is a separated reverse transcription reaction to form cDNA. The second step is the amplification of cDNA using PCR (Wacker and Godard, 2005; Singh et al., 2014).

For viral detection, PCR can be utilized to detect viruses at the same level of classification, targeting particular gene or conserved region for viruses from the same family or genus. However, there might be slight gene variations between viruses even in a conserved area that can reduce the template-primer binding affinity. To overcome this problem, some methods and modifications have been developed to improve pathogen detection sensitivity and specificity including variations on PCR steps and cycles, the use of consensus and degenerate primers (primer multiplexing and inosine) (VanDevanter et al., 1996) and the addition of components to the reactions (such as DMSO) have proven to be useful to increase the yield of PCR products (Jensen et al., 2010). In conventional PCR, additional reaction using primers targeting part of sequence inside the PCR products generated from the first amplification can increase sensitivity and specificity of the assay. This method is known as nested/seminested PCR. This modification is very useful for rare templates or PCR with high background by maximizing the total cycles while minimizing non-specific ampiclon (Lam et al., 2007; Abrahao et al., 2010). Availability of a positive control is another major challenge for laboratories or hospital in attempt to detect emerging viruses using PCR. Extracted nucleic acid from PCR-positive samples (Vieira et al., 2019), plasmid containing sequence of particular virus (Watanabe et al., 2010), expressed RNA (Tong et al., 2008) and also in vitro synthesized sequences (Zhai et al., 2007) can be used as alternatives for a PCR positive control. There are two methods which can be used to perform PCR amplification: conventional PCR and real-time PCR. Apart from its quantitative ability, the advantages of real-time PCR include improved rapidity, sensitivity, reproducibility and a reduced risk of carry-over contamination (Mackay et al., 2002). The real-time PCR methods rely upon a fluorescent reporter of the probe used for detection and quantitation. The signal of the reporter increases in direct proportion to the amount of PCR product in a reaction. The three most popular probes used in the detection of pathogenic viruses are the SYBR Green 1, TaqMan and molecular beacons (Parida, 2008). The SYBR Green 1 binds to double-stranded DNA, and upon excitation emits light. It only fluoresces when intercalated into double stranded DNA. The intensity of the fluorescence signal is therefore dependent on the quantity of double stranded DNA present in the reaction. The advantages of SYBR Green 1 are that it is cheap, easy to use, and sensitive. The main disadvantage of this probe is that it is not specific, since the probe binds to all double stranded DNAs formed during the PCR reaction, including the non-specific PCR products, primer-dimers, etc. (Arya et al., 2005; Parida, 2008). The TaqMan Probe is a hydrolysis probe. It has a fluorescent reporter dye attached to its 5′ end and a quencher dye at its 3′ terminus. When the TaqMan probe is intact, the reporter and quencher stay close to each other preventing the emission of any fluorescence. If the viral target sequence is present the TaqMan probe anneals downstream from one of the primer sites. During PCR, when the polymerase replicates a template on which a TaqMan probe is bound, the 5′ exonuclease activity of the polymerase cleaves the probe. This separates the fluorescent and quenching dyes and fluorescence is detected. Fluorescence increases in each cycle in proportion to the rate of probe cleavage (Arya et al., 2005; Parida, 2008). Similarly, the molecular beacons also contain bound fluorescent and quenching dyes at either end of a single-stranded DNA molecule. However, they are also designed to adopt a stem-and-loop structure whilst free in solution to bring the fluorescent dye and the quencher in close proximity and prevent fluorescence emission. The loop portion of the molecule is complementary to the target nucleic acid molecule and the stem is formed by the annealing of complementary arm sequences on the ends of the probe sequence. When the probe sequence in the loop hybridizes to a complementary nucleic acid target sequence during the annealing step, a conformational change occurs that forces the stem apart. This results in separation of the fluorophore from the quencher dye, hence, emission of fluorescence (Arya et al., 2005; Parida, 2008). A fluorescent signal that is detected above the threshold is considered a real signal that can be used to define the threshold cycle (Ct) for a sample. Ct is defined as the fractional PCR cycle number in which the reporter fluorescence is greater than the minimal detection level, the threshold. The presence of more viral nucleic acid template at the start of the reaction leads to a fewer number of cycles reaching the point at which the fluorescent

Fig. 2. Diagram of the one-step and two-step reverse transcription polymerase chain reaction (RT-PCR). In one-step RT-PCR, the reverse transcription step is coupled with PCR. The viral RNA undergoes reverse transcription to cDNA followed by PCR amplification in a single reaction. In the two-step RT-PCR, the process is carried out in two steps. The first step is a reverse transcription reaction to form cDNA. The second step is the amplification of cDNA using PCR (Wacker and Godard, 2005; Singh et al., 2014).
signal is significantly above the background (Arya et al., 2005). Viral quantitation can therefore be carried out by comparing the Ct values of the samples at a specific fluorescence value with similar data collected from a series of standards by a calculation of a standard curve (Mackay et al., 2002).

The PCR-based detection methods have been successfully employed for the detection of both pathogenic DNA and pathogenic RNA viruses. For example, PCR amplification of 242 bp specific portion of hepatitis B genomic DNA, the surface (S) gene, was employed for the detection of hepatitis B virus in patients with liver disease (Riaz et al., 2016). The RT-PCR method was used to detect West Nile virus based on amplification of a 242-bp sequence of the NS5 gene from flavivirus positive-samples (Myint et al., 2014). TaqMan quantitative real-time RT-PCR assays have also been developed and used for the specific detection and quantification of North American eastern equine encephalitis (EEE) and western equine encephalitis (WEE) RNAs based on amplification of 69 bp and 67 bp of the corresponding genome sequences, respectively (Lambert et al., 2003; Vina-Rodriguez et al., 2016). In addition, the PCR-based detection methods have also been successfully employed for the detection of pathogenic DNA viruses such as adenoviruses (Allard et al., 1990), herpes viruses (Shimada et al., 2017), polyomaviruses (Vanchiere et al., 2005), poxviruses (Scaramozzino et al., 2007; Putkuri et al., 2009) and pathogenic RNA viruses such as Ebola (Cherlopilid et al., 2016), Zika (Perkasa et al., 2016), SARS (Chan et al., 2004b), MERS-CoV (Lai et al., 2014), H5N1 (Ng et al., 2005; Karo-karo et al., 2019), Chikungunya (Riswari et al., 2016), coxsackievirus (Wiyatno et al., 2016), rhinovirus C (Wiyatno et al., 2018), etc. Other molecular methods for detecting pathogenic viruses such as direct nucleic acid detection by a biosensor (Saylan et al., 2019) and Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) diagnostics using CRISPR-CRISPR-associated system (CRISPR-Cas) (Chertow, 2018) have also been developed. In addition, loop-mediated isothermal amplification (LAMP), a recent development of a diagnostic method offers advantages in terms of sensitivity, specificity, efficiency and rapidity and can be performed at the point of care for detection of emerging viruses (Notomi et al., 2000; Wong et al., 2017; Lopez-Jimena et al., 2018; Sabalza et al., 2018).

4. DNA sequencing

The ability of pathogenic viruses to adapt to new environments, cross species barriers, develop resistance to antiviral drugs, and to evade existing vaccines is associated with the accumulation of mutations throughout their genomes. Complete genomic information of the pathogenic viruses circulating in the human population is therefore critical for the effective control of human viral disease. DNA sequencing technologies allow determination of the exact nucleotide sequence of each viral genome, in order to better understand viruses especially in terms of genetic diversity, evolution, pathogenesis, ecology, and vaccine design (Leal and Zanotto, 2000; Dijkeng and Spiro, 2009). Sequence data are also important for the identification of viruses in circulation including the discovery of new viruses. Together with PCR screening, DNA sequencing can be employed to detect previously unidentified viruses. In addition, sequence data can be used to track the ancestral relationships between viruses and to infer their possible origins (Leal and Zanotto, 2000; Haagmans et al., 2009). Understanding the molecular mechanisms of virus emergence and transmission is critical. Virus and host genomic data may contribute to rapid identification of critical mutations that enable viruses to spread efficiently, interact with different receptors and cause disease in different hosts. The majority of pathogenic viruses causing recent outbreaks of public health are zoonotic (Haagmans et al., 2009). Efforts toward virus whole genome sequencing is also aimed at facilitating the development of reagents and resources for more accurate molecular diagnosis (Dijkeng and Spiro, 2009).

DNA sequencing is a process to determine the sequence of nucleotides in a piece of DNA. Sequence determination of viral genomes is most commonly performed using the dideoxy-chain-termination technology also known as the “Sanger Method”. This method requires a single-stranded DNA template, a DNA primer, a DNA polymerase, normal deoxynucleosidetriphosphates (dNTPs), and dideoxynucleosidetriphosphates (ddNTPs). The viral DNA sample is divided into four separate sequencing reactions and only one of the four dideoxynucleotides (ddATP, ddGTP, ddCTP, or ddTTP) is added. The method is based on the use of dideoxynucleotides (ddNTPs) to terminate DNA strand elongation. Chemically, the ddNTPs are similar to the dNTPs, but lack a 3’-OH group needed for the formation of a phosphodiester bond between two nucleotides, causing the DNA polymerase to stop extension of DNA when a ddNTP is incorporated. Viral DNA to be sequenced is first denatured into single strands using heat. A primer is then annealed to one of the template strands. This primer is specifically designed so that its 3’ end is in the position next to the DNA sequence of interest. The primer, or one of the nucleotides, should be radioactively or fluorescently labeled in order to allow detection of the final product on an electrophoresis gel. The key to this method, is that all the reactions start from the same nucleotide and end with a specific base. As a result, bands of varying lengths are generated. To accelerate the sequencing process and improve sensitivity, automated sequencing was then developed. In this version, the reactions are carried out in a single tube containing all four ddNTPs, each labeled with a different color dye that will fluoresce at different wavelengths. The DNA sequencers separate strands on the basis of their size by using capillary electrophoresis, and detect and record dye fluorescence. Since the four dyes fluoresce at different wavelengths, the identity of each band is read according to the wavelengths at which it emits fluorescence. The outcomes are then depicted in the form of a chromatogram, which is a diagram of colored peaks that correspond to the nucleotide in that position in the DNA sequence (Obenrader, 2003). For whole genome sequencing of RNA viruses, the most commonly used strategy involves the design of overlapping amplicons that span the entire genome followed by the targeted amplification of genomic regions by reverse-transcription-PCR. Sequence data generated are then assembled (Dijkeng and Spiro, 2009).

Recent advances in high-throughput sequencing (HTS), also referred to as next generation sequencing (NGS) methods have revolutionized the viral genomics field. Unlike Sanger Sequencing, HTS technologies provide massively parallel analyses resulting in a vast amount of sequence data. Massive sequence data is generated by parallel construction of randomly short fragmented DNA, or cDNA synthesized from RNA, called template libraries, on an isolated solid surface or beads. Nucleotide bases are recorded by equipment which can detect luminescence or changes in electrical charges. Enormous digital data generated by this equipment should be analyzed using bioinformatics. Currently, HTS has arrived at the third-generation, which is able to perform a sequence from a single nucleotide, eliminates steps of library preparation and simplify the overall workflow. Albeit, data reliability and accuracy is still being improved, third generation sequencing offers powerful genomic analysis at lower cost in the future (Chiu and Miller, 2016; Kulski, 2016). It has also been applied to sequence genomes of pathogenic viruses (Marston et al., 2013; McGinnis et al., 2016; Yudhaputri et al., 2017).

Apart from PCR-based detection, known and unknown pathogenic viruses in circulation can also be detected by high-throughput sequencing of small RNAs populations followed by bioinformatics analysis (Massart et al., 2019). This is due to the fact that during the infection process, viruses are targeted by the host silencing machinery and, as a consequence, virus derived small RNAs (sRNAs) accumulate in infected host cells. This powerful approach was first developed for identifying and characterizing viruses which infect plants. High-throughput sequencing of sRNAs potentially allows the identification of all types of virus regardless of their genome nature or structure (Massart et al., 2019). The strategy has also been used for detecting human
pathogenic viruses such as human papillomavirus type 18 (HPV-18), hepatitis B virus (HBV), hepatitis C virus (HCV), human immunodeficiency virus type 1 (HIV-1), swamp monkey retrovirus (SMRV), Epstein-Barr virus (EBV) (Wang et al., 2016) and enterovirus-A71 (EV-A71) (Lee et al., 2020). It is noteworthy that computational analysis of a large number of very small size of viral and host sRNA sequences may be challenging especially when the proportion of the of viral sRNA sequences is very low, or when identifying novel viruses which are not represented in the current reference sequence databases. If the host genome sequence is available, a filtering step can be used to separate viral sRNAs from the host sRNAs which may facilitate virus genome reconstruction. Overall the high-throughput sequencing of sRNAs may be used as a reliable tool for detection of known and unknown pathogenic viruses (Massart et al., 2019). Recently, metagenomics and bioinformatics approaches have also been developed for molecular assay of viral and other pathogen infection (Schneeberger et al., 2016). Detecting pathogenic viruses in specimens collected from body compartments such as upper respiratory tract or gastrointestinal tract is challenging as these body compartments are colonized by various pathogens. The high-throughput sequencing technology has also been employed in a metagenomic diagnostics approach proven to be useful for simultaneous detection of DNA viruses (Adenoviridae, bocavirus, cytomegalovirus) together with other pathogens (bacteria, protozoa, helminths) in stool specimens from patients with persistent diarrhea. These microorganisms are suggested to be the causative pathogens of digestive disorders (Schneeberger et al., 2016). Metagenomics is direct genetic analysis of microbial genomes in the environmental samples without a need to propagate the microorganisms. This approach provides detailed information on the presence and diversity of pathogenic organisms in a particular environment. It is proposed that the metagenomic diagnostics strategy can be followed by further techniques such as RNA-based meta-transcriptomics analysis to obtain more information on infections with RNA viruses (Schneeberger et al., 2016). An unbiased protocol for metagenomic detection of known and novel emerging viruses from clinical specimens has also been established. The protocol has been extensively validated by using real-time PCR and next-generation sequencing, and has the ability to increase the amount of detectable virus nucleic acids and improved the detection of viruses about 75,000-fold compared with other tested approaches. The protocol is intended to be used in metagenomic studies to increase the likelihood of detecting viruses from any biological source (Kohl et al., 2015). Various molecular methods which have been used, or developed, for pathogenic virus detection and characterization are summarized in Table 1.

Genomic analysis of pathogenic viruses can be started by generating a multiple sequence alignment of virus sequences that can then be used to construct a phylogenetic tree representing an evolutionary relationship between various viruses (Li et al., 2014; Agustiningisih et al., 2018). As viral genomes evolve by the accumulation of mutations, the amount of differences in nucleotide sequence between a pair of viral genomes indicates how recently those two genomes shared a common ancestor. Two viral genomes that diverged in the recent past are expected to have fewer differences compared to a pair of viral genomes whose common ancestor is more ancient. Thus, by comparing three or more viral genomes it is possible to analyze the evolutionary relationships between them. The phylogenetic method can be used to reconstruct a tree-like pattern that describes the evolutionary relationship between the viruses being studied. The phylogenetic tree reconstruction is based on data obtained from comparisons of virus nucleotide sequence made by aligning virus sequences to score nucleotide differences. Sequence alignment, therefore, is a critical part of the process because it will determine the correctness of the phylogenetic tree being constructed. Virus sequences to be included in the alignment must be homologous in that they are derived from a common ancestral sequence. It is important to note that the phylogenetic analysis must also include at least one outgroup, a homologous sequence known to be closely related to the viral sequences being analyzed, in order to locate the root of the phylogenetic tree and enable identification of the evolutionary route (Brown, 2002).

Phylogenetic analysis has become a useful way to investigate evolutionary genetic relationship among pathogenic viruses. For example, the method has been used to analyze evolutionary relationship of pathogenic viruses recently detected in Indonesia such as the West Nile virus, the Coxsackievirus B3 and the Zika virus. The phylogenetic analysis of a West Nile virus revealed that the virus circulating in Indonesia is most closely related (99% nucleotide identity) to the West Nile virus strain (B956) isolated from Uganda that belongs to the lineage 2. There are two main genetic lineages of West Nile virus, lineage 1 and lineage 2. The phylogenetic analysis has also confirmed the relationship of the Indonesian strain with other lineage 2 West Nile virus sequences. West Nile virus is a zoonotic mosquito-transmitted arbovirus belonging to the Flaviviridae family. The lineage 2 strains have been reported to cause outbreaks in Europe (Myint et al., 2014). The phylogenetic analysis of Coxsackievirus B3 (CVB3) detected for the first time in Indonesia showed that the virus is closely related (97% homology) to the CVB3 from Taiwan. The analysis was based on the nucleotide sequence of the VP1 region which is considered as the most reliable method for determination of enterovirus genotype. CVB3 is a member of the enterovirus B group of the family Picornaviridae, an important human pathogen associated with a number of diseases such as type 1 diabetes mellitus, myopericarditis, aseptic meningitis, hemorrhagia, pancreatitis and foot, and mouth diseases (HFMD). The virus can cause fatal diseases especially when it infects children and newborn infants (Wiyatno et al., 2016). Genomic characterization of Zika virus by whole genome sequencing, followed by phylogenetic analysis showed that the Zika virus isolated from Indonesia is distantly related to the Brazilian Zika virus which has been linked to microcephaly (Yudhaputri et al., 2017). In addition, phylogenetic analysis has also been employed to study the evolution and phylodynamic pattern of human influenza A/H3N2 virus circulating in Indonesian archipelago from 2008 to 2010. The study showed the existence of genetic drift, due to genetic evolution of the virus. The mutation pattern of the virus was not geographically related as viruses isolated from same years tended to cluster together, regardless of geographical origin. Multiple lineages of the H3N2 viruses were found to co-circulate in each year and the existence of a particular lineage was suggested as the result of adaptation or ecological factors (Agustiningisih et al., 2018). DNA sequence analysis was also used to reveal the reassortment events of gene segments among avian influenza A(H5N1) viruses circulating in Indonesia from 2015 to 2016 (Karo-karo et al., 2019).

The ultimate goal of genomic characterization of pathogenic viruses is to provide information about the viruses for the development of better diagnostic tests, improved vaccine design and more effective antiviral drugs. Complete genome sequences of pathogenic viruses can be used to improve development of methods for virus detection and identification. For example, the primary sequences of Ebola viruses have been used to design primers for Ebola molecular detection (Joob and Wiwanitkit, 2015). In addition, considering the relatively small sizes of most viral genomes, and the increased sequencing efficiency, full genome sequences of new pathogenic viruses can become a standard for virus identification (Cotten et al., 2014). Vaccines are the most effective way to eradicate pathogenic viruses. Induction of protective immunity through vaccination can be a critical tool to protect a population at risk against potential viral diseases. Complete genomic information of pathogenic viruses is of importance for the rapid and rational design of vaccines based on identifications of key viral protein antigens that induce immune response rather than on the whole viruses. The genome-derived vaccines are based on the concept that selection and design of antigen is critical for vaccine efficacy. Computational algorithms can be used to select critical antigens using the genomic data of pathogenic viruses (Garcia-Sastre and Mena, 2013; Moise et al., 2016). Virus genome sequences also provide information to identify...
genes that encode proteins which are good vaccine targets (Seib et al., 2009). In addition, advances in viral genomics offer the opportunity to develop more effective antiviral drugs. Genome sequences provide a route to investigate mechanisms of virus pathogenesis and a repertoire of antiviral drug targets from which novel antiviral drugs can be developed (Yudhaputri et al., 2017). Beside their roles in virus detection and characterization, molecular methods are applicable to develop vaccines and therapeutics to combat pathogenic viruses (Artika et al., 2013; García-Sastre and Mena, 2013).

5. Development of laboratory capacity

It is possible that pathogenic viruses of global public health and socioeconomic risk emerge in regions with limited laboratory capacity for their molecular detection and characterization. Hence, to minimize the impact of emerging viral diseases, it is critical to establish local laboratory capacity as this plays a critical role in the surveillance, diagnosis, and monitoring of viral disease, as well as in the understanding of the viral genetic changes. The program might cover improving laboratory facilities, ensuring availability of required equipment, building human resource capacity and ensuring quality laboratory management. It is important to generate reliable and accurate data from a diagnostic laboratory needed to guide control measures. Improved laboratory capacity is essential to enhance national self-reliance in responding to viral disease outbreaks and to support an international agenda in reducing the spread and adverse impact of emerging viral diseases. The development of laboratories with an established capacity for molecular diagnosis and research of emerging viruses may be challenging because investment and long-term commitment is required for sustainability of capacity improvements (World Health Organization (WHO), 2008; Wertheim et al., 2010; Chua and Gubler, 2013). Laboratory management is critical for successful establishment of laboratory capacity. It is important to ensure the availability of qualified personnel, appropriate facilities, equipment, reagents, materials, SOP, training programs, implementation of good laboratory practices, etc. (World Health Organization (WHO), 2009).

Good facilities for preservation of viral sample, especially RNA viruses, are important as viral RNA molecule is unstable and highly prone to degradation. In underdeveloped countries, accurate laboratory diagnosis of RNA viruses is frequently hampered by a lack of sample preservation due to difficulties in the transportation and storage of clinical material collected from suspected outbreaks located far from a laboratory. Maintenance of the cold chain is important to avoid sample degradation which causes difficulties in detecting and characterizing field isolates. Facilities required for cold chain maintenance for sample transport and storage are important to keep samples or material within a certain temperature range during all stages of delivery, processing and storage to ensure the viral preservation. The effective mean to maintain the integrity of RNA viruses for laboratory analysis is by freezing. Hence, the availability of deep freezers for adequate specimen storage and power back-up systems for an uninterrupted power supply is also crucial (Wolking, 2013; Fowler et al., 2014). In order to improve effectiveness and efficiency, it is necessary to maximize the use of existing resources and facilities. Active participation of local scientists and experts in scientific activities and training programs should be encouraged. The establishment of a laboratory network is essential to facilitate scientists within the regions to build sustainable relationships that encourage experience and information exchange (Wertheim et al., 2010).

PCR has become the molecular technique of choice for detection of viral nucleic acids in the laboratory. One of the advantages of PCR is its ability to produce large numbers of copies of a target sequence from a very small sample amount, even single copies of DNA which makes PCR a very sensitive and powerful diagnostic tool. However, the analytical sensitivity of PCR may also bring about problems as it can generate false-positive results due to sample-to-sample contamination. In the laboratory, false-positive results can also be caused by carry-over of DNA from a previous amplification of the same target sequence. This may happen due to aerosolization of PCR products especially during the post-PCR analysis. PCR contamination has been an issue in laboratories

Table 1
Various molecular methods which have been used or developed for detection and characterization of pathogenic viruses.

| Methods                          | Application                                      | Example of targeted virus(s)                        | References                  |
|----------------------------------|--------------------------------------------------|-----------------------------------------------------|-----------------------------|
| Polymerase chain reaction (PCR)  | Detection of DNA viruses                         | Hepatitis B virus                                   | Riaz et al., 2016           |
| Reverse transcription (RT)-PCR   | Detection of RNA viruses                         | West Nile virus, Coxackievirus                      | Myint et al., 2014; Wiyatno et al., 2016 |
| Real time RT-PCR                 | Detection and quantification of RNA viruses      | Dobra (DOB) virus, Hantaan (HTN) virus, Puumala (PUU) virus, Seoul (SEO) virus, 2019 novel coronavirus (2019-nCoV) | Aiello et al., 2005; Huang et al., 2020 |
| Nested PCR                       | Detection of DNA viruses                         | Hepatitis B virus                                   | Le et al., 2015              |
| Nested RT-PCR                    | Detection of RNA viruses                         | Eastern Equine Encephalitis (EEE), Western Equine Encephalitis (WEE), Venezuelan Equine Encephalitis (VEE) | Sanchez-Seco et al., 2001 |
| Duplex nested-RT-PCR             | Detection of RNA viruses                         | Nipah virus                                         | Wacharapluesadee and Hemachudha, 2007 |
| Multiplex nested PCR             | Simultaneous detection of different pathogenic viruses | Influenza A viruses (H1N1, H3N2, H5N1), influenza B virus | Lam et al., 2007             |
| Loop-mediated isothermal amplification (LAMP) | Detection of pathogenic viruses with high sensitivity, specificity, efficiency and rapidity under isothermal conditions | Hepatitis B virus, adenovirus, cytomegalovirus, chikungunya virus, zika virus | Notomi et al., 2000; Wong et al., 2017; Lopez-Jimena et al., 2018; Sabalza et al., 2018 |
| Biosensor                        | Real time detection of viral nucleic acids with high sensitivity, specificity and portability | Human immunodeficiency virus (HIV), hepatitis B virus, ebola virus, zika virus, norovirus, dengue virus | Saylan et al., 2019          |
| Metagenomic diagnostics         | Simultaneous detection of multiple pathogenic viruses | Reovirus, influenza A virus, vaccinia virus, sendai virus, adenoviridae, bacavirus, cytomegalovirus | Kohl et al., 2015; Schneeberger et al., 2016 |
| CRISPR diagnostics              | Rapid, accurate, and portable diagnostic tools for pathogenic viruses | Zika virus (ZIKV), dengue virus (DENV), human papillomavirus (HPV) | Chertow, 2018                |
| High-throughput sequencing of small RNAs | Detection of pathogenic viruses human papillomavirus type 18 (HPV-18), hepatitis B virus (HBV), hepatitis C virus (HCV), human immunodeficiency virus type 1 (HIV-1), squirrel monkey retrovirus (SMRV), Epstein-Barr virus (EBV) and enterovirus-A71 (EV-A71) | Wang et al., 2016; Lee et al., 2020 |                |
| Whole genome sequencing          | Characterization of virus                        | Zika virus, dengue virus                            | Yudhaputri et al., 2017; Dhemni et al., 2018 |
conducing the detection of emerging pathogenic viruses. Therefore, it is very important to develop strategy to avoid cross-contamination of samples and PCR products during PCR process. One of the strategies lies in the layout and operation of the diagnostic laboratory. Physical separation of spaces for reagent preparation, nucleic acid extraction, and amplification and product analysis room (Fig. 3) is recommended. Workflow between these rooms/areas must be unidirectional in that from clean areas to contaminated areas, but not from contaminated areas to clean areas (Wertheim et al., 2010; Viana and Wallis, 2011; Public Health England, 2013).

The reagent preparation room, also commonly called the clean room, is dedicated to reagent storage, aliquoting reagent stocks, and preparation of reaction mixes prior to the addition of sample nucleic acid. It is very important to maintain this room free of nucleic acid extracts, samples, cDNA and PCR products. The nucleic acid extraction room is used for extracting nucleic acid from viral samples. This room is also used for adding extracted DNA or cDNA and the positive control to the PCR reaction mixes. The amplification and product analysis room is used for carrying out PCR and post-PCR analysis such as separation of PCR product using agarose-gel-electrophoresis. The PCR machines are used for carrying out PCR and post-PCR analysis such as separation of PCR product using agarose-gel-electrophoresis. The PCR machines are placed in this room. In an ideal situation, the work station for PCR should be divided into four separated rooms: reagent preparation room, nucleic acid extraction room, PCR amplification room and product analysis room. However, many laboratories only have two rooms available. Therefore, pre-PCR and extraction are carried out within defined areas of a larger room and amplification and product analysis are performed in a second room. For performing real-time PCRs, only three rooms may be needed as post-PCR analysis is not required (Wertheim et al., 2010; Viana and Wallis, 2011; Public Health England, 2013).

It is worthwhile to note that the handling of pathogenic viruses in any diagnostic and research laboratory is dangerous as the viruses can cause accidental infection of laboratory workers. These risks can be minimized through the application of biosafety principles and practices. The activities need to be performed within a biocontainment environment using appropriate equipment and personal protective equipment (Artika and Ma’roef, 2017). A diagnostic virology laboratory must be set up to meet biosafety level 2 (BSL2) standards or above. It is desirable that the laboratory is equipped with biosafety level 3 (BSL3) facilities because most of the pathogenic viruses recently emerged belong to the risk groups which require BSL3 facilities for handling specimens and cultures. However, if resources are constrained, a BSL2 laboratory with negative pressure can be used to handle most of the emerging viruses (World Health Organization (WHO), 2008). To address the possibility that pathogenic viruses may be misused for malicious purposes, implementation of a biosecurity system in any laboratory working with pathogenic viruses is recommended. The laboratory biosecurity system should cover physical security, virus control and accountability, personnel security, transport security, and information security; the implementation of which needs to be integrated with the overall bio-safety plan (Artika and Ma’roef, 2018).

6. Conclusion

Infectious diseases caused by the emergence of pathogenic viruses have become a threat to global public health. The 2014–2016 devastating Ebola crisis in West Africa followed by the spread Zika viruses in Brazil and the continuing circulation of Nipah virus in Asia reminds us that every region in the world needs to be prepared to face emerging viral disease outbreaks. Laboratories play critical roles in viral outbreak investigations, especially to provide diagnostic information on the viral agents required for patient management and as a guide to control virus spread. Molecular methods, in particular PCR and DNA sequencing technologies, have become the methods of choice for the detection and characterization of emerging and re-emerging pathogenic viruses. Although many factors to drive virus emergence have been suggested, predicting new outbreaks of emerging viral diseases is difficult. Deadly viruses of pandemic potential may emerge in the poorer global regions with limited capacity for early detection and identification of emerging viruses. Despite rapid advancements of molecular techniques, their implementation in low-resource environments frequently remains cumbersome due to logistic, financial, and educational constrains. Development of laboratory capacity, especially in the regions predicted to be hotspots for the emergence of pathogenic viruses, is critical. Advances in molecular biology techniques should be contingent with the development of more effective diagnostic tests for emerging viruses. Future efforts should be directed to development of molecular methods which are inexpensive, accurate, and provide results rapidly, and which are applicable to the identification of emerging pathogenic viruses in resource-limited settings. As handling pathogenic viruses pose the risk of infection to laboratory personnel, and also potentially to community and environment, good laboratory biosafety and biosecurity systems need to be established.

Declaration of Competing Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

Acknowledgements

This work was carried out without external funding support. We thank Dr. John Acton for his assistance at the manuscript stage.

References

Abrahao, J.S., Drumond, B.P., Trindade, G.D.S., da Silva-Fernandes, A.T., Ferreira, J.M.S., Alves, P.A., Campos, R.K., Siqueira, L., Bonjardim, C.A., Ferreira, P.C.P., Kroon, E.G., 2010. Rapid detection of orthopoxvirus by semi-nested PCR directly from clinical specimens: a useful alternative for routine laboratories. J. Med. Virol. 82, 692–699. https://doi.org/10.1002/jmv.21617.

Adisasmito, W., Aisyah, D.N., Aditama, Y.T., Kusrastuti, R., Trihono, S.A., Sampurno, O.D., Prasenoshadi, S.N.A., Mamahit, M.J.N., Svenson, A., Dreyer, N.A., Goker, R., 2013. Human influenza A H5N1 in Indonesia: health care service-associated delays in treatment initiation. BMC Public Health 13, 571–576. https://bmcpublichealth.biomedcentral.com/articles/10.1186/1471-2458-13-571.

Agustiningsih, A., Trismananto, H., Setiawary, V., Artika, I.M., Muljono, D.H., 2016. Primer development to obtain complete coding sequence of HA and NA genes of influenza A/H1N1 virus. B.M.C. Res. Notes 9, 1–6. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5004302/.

Agustiningsih, A., Trismananto, H., Restuadi, Artika I.M., Hellard, M., Muljono, D.H., 2018. Evolutionary study and phylodynamic pattern of human influenza A/H1N1 virus in Indonesia from 2008 to 2010. PLoS OnePLOS One. 13, 1–13. https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0204127.

Allard, A., Girones, R., Juto, P., Wadell, G., 1990. Polymerase chain reaction for detection of infectious diseases caused by the emergence of pathogenic viruses. Infection, Genetics and Evolution 81 (2020) 104215
of adenoviruses in stool samples. J. Clin. Microbiol. 28, 2659–2667. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC268252/.

Ammalalui, A.S., Pattinaik, A, Saooh, B.R., Muthukrishnan, E, Natarajan, S.K., Steffen, D., Hu, L.X., Delhon, G, Osorio, F.A., Petro, T.M., Xiang, S.-H., Pattinaik, A.K., 2017. ZIb-2 virus encoding a reverse transcriptase protein is attenuated and defective in neuroinvasion. J. Virol. 91 e01438, 17. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5868755/.

Artika, I.M., Ma'roef, C.N., 2017. Laboratory biosafety for handling emerging viruses. Am. J. Biochem. Biotechnol. 14, 108–116. https://thesciencedirect.com/abstract/doi/10.1016/1878-0859(16)31086-6.

Artika, I.M., Budirahardyja, E, Ekowati, A.L., 2013. Molecular cloning and heterologous expression of human interferon alpha2b gene. Am. J. Biochem. Biotechnol. 9, 429–429. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3422175/.

Arya, M., Sgrellig, I.S., Williamon, M., Gomemmall, L., Arya, N., Patel, H.R.H., 2005. Basic principles of real-time quantitative PCR. Expert. Rev. Mol. Diagn. 5, 209–219. https://www tandfonline.com/doi/full/10.1586/14737914.5.2.209.

Baker, L., Chertov, D., Johans, K.M., Feldmann, H., Moens, D., 2017. The pathogenesis of ebola virus disease. Annu. Rev. Pathol. Mech. Dis. 12, 15.1–15.22. https://wwwannualreviews.org/doi/10.1146/annualrev-050216-100506.

Bloom, D.E., Black, S., Rappuoli, R., 2017. Emerging infectious diseases: a proactive approach. Proc. Natl. Acad. Sci. U. S. A. 114, 4055–4059. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5404224/.

Broadhurst, M.J., Brooks, T.J.G., Pollock, N.R., 2016. Diagnosis of Ebola virus disease: past, present, and future. Clin. Microbiol. Rev. 29, 773–793. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5107479/.

Brown, T.A., 2002. Molecular phylogenetics. In: Genomes, 2nd edition. Garland Science. Manchester, UK. https://www.ncbi.nlm.nih.gov/books/NBK211222/.

Calisher, C.H., Childs, J.E., Field, H.E., Holmes, K.V., Schouszu, T., 2006. Bats: important reservoirs of filoviruses. Clin. Microbiol. Rev. 19, 521–545. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1539106/.

Cauley, D., Edwards, S.V., 2008. Ecology of avian influenza viruses in birds. J. Infect. Dis. 197, S29–S33. https://academic.oup.com/jid/article/197/Supplement_1/S29/ 842172.

CDC, 2019. Laboratory safety when working with zika virus. Centers for Disease Control and Prevention. https://www.cdc.gov/zika/professionals/laboratory/safety/laboratory-zika.html.

Chua, K.B., Gubler, D.J., 2013. Perspectives of public health laboratories in emerging infectious disease. J. Clin. Microbiol. 51, 77–94, 9–9. https://www.sciencedirect.com/science/article/pii/S1386653212007008.

Chow, K.Y.C., Hon, C.C., Hui, R.K.H., Wong, R.T.Y., Yip, C.W., Zeng, F., Leung, F.C.C., 2014. Coronavirus and pandemic potential. J. Clin. Microbiol. 52, 1099–1101. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3644304/.

Gijavanekar, C., Aňez-Lingerfelt, M., Chen Feng, C., Putonti, C., Fox, G.E., Sabo, A., 2016. Detection of enterovirus D68 from human respiratory samples using a highly sensitive primer ‘cocktail’ FEBS J. 278, 1676–1687. https://onlinelibrary.wiley.com/doi/full/10.1111/1742-4658.12982.

Gregory, C.J., Oduyebo, T., Brault, A.C., Brooks, J.T., Chung, K., 2017. Modes of transmission of zika virus. J. Infect. Dis. 216 (Suppl. 10), S875. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5686755/.

Gurubhagavathan, N.D., McMenamy, S.S., McFadden, G., Revell, N.D., Lee, J.S., 2013. Multi-gene detection and quantification of enterovirus D68 using a novel method for detection with a highly sensitive primer ‘cocktail’. J. Med. Virol. 85, 1431–1436. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3777887/.

Hoek, L., 2014. Full genome virus detection in fecal samples using sensitive nucleic acid preparation, deep sequencing, and a novel iterative sequence classification algorithm. PLoS Negl. Trop. Dis. 8, e2878. https://journals.plos.org/plosntds/article?id=10.1371/journal.pntd.0002878.

Huang, C., Yeming, W., Li, X., Ren, L., Zhao, J., Hu, Y., et al., 2020. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. The Lancet. https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(20)30185-X.

Hofmann, J., Weiss, S., Kuhns, M., Zinke, A., Heinsberger, H., Kruger, D.H., 2018. Optimal strategy for detection of emergence of enterovirus D68. J. Med. Virol. 90, 1709–1722. https://onlinelibrary.wiley.com/doi/full/10.1002/jmv.25358.

Huang, C., Yeming, W., Li, X., Ren, L., Zhao, J., Hu, Y., et al., 2020. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. The Lancet. https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(20)30185-X.

I.M. Artika, et al.
Infection, Genetics and Evolution 81 (2020) 104215

Dhenni, R., Hartoyo, E., Wiyatno, A., Antonjaya, U., Ma’roef, C.N., Myint, K.S., Safari, D., 2017. Occurrence of measles genotype D8 during a 2014 outbreak in Banjarmasin, South Kalimantan, Indonesia. Int. J. Infect. Dis. 54, 1–3. https://www.sciencedirect.com/science/article/pii/S1201971216312139.

Dhenni, R., Karyanti, M.R., Putri, N.D., Yohan, B., Safi, E., 2017. Ocorrência de genótipo D8 de sarampo durante uma epidemia de 2014 em Banjarmasin, Sul Kalimantan, Indonésia. Revista Brasileira de Virologia. 49, 1–9. 10.2217/17497612.2017.17497612-477.

Dhenni, R., Hartoyo, E., Wiyatno, A., Antonjaya, U., Ma’roef, C.N., Myint, K.S., Safari, D., 2017. Ocorrência de genótipo D8 de sarampo durante uma epidemia de 2014 em Banjarmasin, South Kalimantan, Indonesia. Int. J. Infect. Dis. 54, 1–3. https://www.sciencedirect.com/science/article/pii/S1201971216312139.

Dhenni, R., Karyanti, M.R., Putri, N.D., Yohan, B., Safi, E., 2017. Ocorrência de genótipo D8 durante uma epidemia de 2014 em Banjarmasin, South Kalimantan, Indonesia. Int. J. Infect. Dis. 54, 1–3. https://www.sciencedirect.com/science/article/pii/S1201971216312139.

Dhenni, R., Hartoyo, E., Wiyatno, A., Antonjaya, U., Ma’roef, C.N., Myint, K.S., Safari, D., 2017. Ocorrência de genótipo D8 de sarampo durante uma epidemia de 2014 em Banjarmasin, South Kalimantan, Indonesia. Int. J. Infect. Dis. 54, 1–3. https://www.sciencedirect.com/science/article/pii/S1201971216312139.
Ye, J., Zhang, B., Xu, J., Chang, Q., McNutt, M.A., Korteweg, C., Gong, E., Gu, J., 2007. Molecular pathology in the lungs of severe acute respiratory syndrome patients. Am. J. Pathol. 170, 538–545. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1851867/.

Yudhaputri, F.A., Trimarsanto, H., Perkasa, A., Yohan, B., Haryanto, S., Wiyatno, A., Soebandrio, A., Myint, K.S., Ledermann, J.P., Rosenberg, R., Powers, A.M., Sasmono, R.T., 2017. Genomic characterization of zika virus isolated from Indonesia. Virology 510 248–2. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5747357/.

Zhai, J., Palacios, G., Towner, J.S., Jabado, O., Kapoor, V., Venter, M., Grolla, A., Briese, T., Paweska, J., Swanepoel, R., Feldmann, H., Nichol, S.T., Lipkin, W.I., 2007. Rapid molecular strategy for filovirus detection and characterization. J. Clin. Microbiol. 45, 224–226. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1828965/.