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.

CHAPTER 4

Keeping Track of Viruses

JACK HIETPAS, LAURA K. MCMULLAN, HOLLY L. HANSON, AND CHARLES M. RICE

Laboratory of Virology and Infectious Disease, Center for the Study of Hepatitis C, The Rockefeller University, New York, New York

DAVID P. MINDELL

Department of Ecology and Evolutionary Biology, Museum of Zoology, University of Michigan, Ann Arbor, Michigan

INTRODUCTION

Viruses are ubiquitous in nature. They prey on all forms of life from archaea and eubacteria to fungi, plants, and animals. It has been estimated that viruses that target sea-dwelling organisms, such as algae, turn over half of the world's biomass each day. As such, viruses are an important part of the ecological balance on earth. Of the vast collection of ribonucleic acid (RNA) and deoxyribonucleic acid (DNA) viruses, research has focused on the tiny fraction that cause disease in humans, domestic animals, and crops. Modern virology, coupled with public health initiatives, has seen many triumphs including the development of effective vaccines against yellow fever, poliomyelitis, smallpox, mumps, measles, rubella, hepatitis A and B, and papillomavirus-associated cervical cancer. The global eradication of smallpox and poliovirus as well as the prediction and successful prophylactic vaccination against influenza represent tremendous gains for public health. Antiviral therapy against human immunodeficiency virus (HIV) and some herpesviruses, as well as rapid detection and containment of emerging viruses such as the severe acute respiratory syndrome coronavirus (SARS-CoV) illustrate the successful application of viral research programs as well as the need for ongoing efforts focused on rapid identification and elucidation of viral distribution, kinetics, and pathogenesis. With the growth of knowledge about replication, transmission, and disease, viruses have also taken the stage as possible biowarfare, bioterrorist, and criminal agents.

In our unpredictable world, viral diagnostics and forensics are becoming increasingly important and sophisticated. The purpose of this chapter is to introduce basic concepts in virology and to provide several examples of how current technology has been used to identify and track viral pathogens. The first part of the chapter discusses virus structure, replication strategies, classification, and evolution. The second part highlights the varied modes of transmission, infection, and disease manifestations as they relate to the different diagnostic methods for virus identification. We follow this with examples of how emerging viruses were identified and tracked (Sin Nombre, Nipah, West Nile, SARS coronavirus, and monkeypox). With regard to viral forensics, the Schmidt case is a fascinating example illustrating how the study of RNA virus evolution and relatedness was used in a criminal investigation. Finally, we provide a few snapshot views into new technologies that allow the creation of manmade or "engineered" viruses, some exciting new frontiers in viral diagnostics, and a few thoughts on the future of virology.

WHAT IS A VIRUS? BASIC VIROLOGY

VIRUS ANATOMY

Viruses are extremely simple "life" forms without metabolic capacity, organelles, translational machinery, or autonomous replicative potential; the essential elements of a virus are consequently minimal. Virus particles are either enveloped or nonenveloped. Enveloped viruses contain one or more envelope glycoproteins embedded in a lipid bilayer that is acquired by budding through host cell membranes. Nonenveloped viruses are made up of a tight protein shell. In both kinds of viral particles, the envelope or protein shell serves to protect the viral genome from the hazardous extracellular environment, since without such protection, the fragile nucleic acid genome would be vulnerable to physical, chemical, and/or enzymatic destruction. Surface proteins also harbor functions required for virus entry into host cells. Inside the virus particle, the viral genome is associated with one or more proteins to form a complex usually referred to as the nucleocapsid. In addition to nucleic acid binding proteins, some viruses also carry enzymes required to initiate replication (like polymerases). Although not all virus particles are highly structured, two kinds of symmetry are common. Helical symmetry is reflected by the arrangement of the protein subunits in a "spring-like" or stacked lock-washer fashion, whereas icosahedral symmetry comprises 20 triangular faces that form the surface of a sphere. In the simplest icosahedral capsid, one protein subunit is at each triangular point of the 20 faces, requiring 60 total subunits.

While host organisms use only DNA for their genetic material, a viral genome may be composed of either DNA or RNA. The size of viral genomes varies greatly. The genome of variola major, the DNA virus that causes smallpox, is 190 kb. The RNA genome of the SARS-causing coronavirus is 29.7 kb; those of Ebola and Marburg viruses are each 19 kb, while the HIV and poliovirus are 9.2 kb and 7.4 kb, respectively.

VIRUS LIFECYCLE

VIRUS ATTACHMENT AND ENTRY

Among the enormous variety of host cell types, how does a virus choose which cell to commandeer for its own replication and propagation? The surface of an enveloped virus particle contains virus-specific proteins that mediate viral recognition of host cell targets. The membrane surface of a host cell is riddled with numerous macromolecules including glycoproteins, glycolipids, and carbohydrate residues that have specific host functions serving as transport channels, mediating signaling pathways, and assisting in antigen presentation. While these macromolecules may be essential for host cell survival and function, some of these macromolecules also serve as receptors for virus particles. The presence of the host cell receptor is essential for viral attachment, but the mere presence of the receptor does not confer upon the cell the ability to support virus replication. Some viruses (such as HIV) require a coreceptor for viral entry without which the virus may be able to bind to the host cell membrane but remain unable to enter the cell. If the virus is not enveloped or is a naked particle, the viral capsid undergoes a conformational change upon binding that releases energy used to breach the host membrane, allowing viral entry. Delivery of viral nucleocapsid, and hence the genome RNA or DNA, occurs via fusion of the virus particle with cellular membrane either at the cell surface or after endocytosis. Virus genome replication can occur either in the cytoplasm or the nucleus, and all viruses usurp the host cell translational machinery to produce virus-specific proteins.

REPLICATION STRATEGIES

In 1971, David Baltimore divided viruses into seven groups based on genetic material, polarity, and mRNA synthesis.[1] Insight into the replication strategy of a virus can be gained by understanding how other well-studied viruses in the same group replicate. The following is a brief description of the seven groups; included where possible, are examples of viruses belonging to each

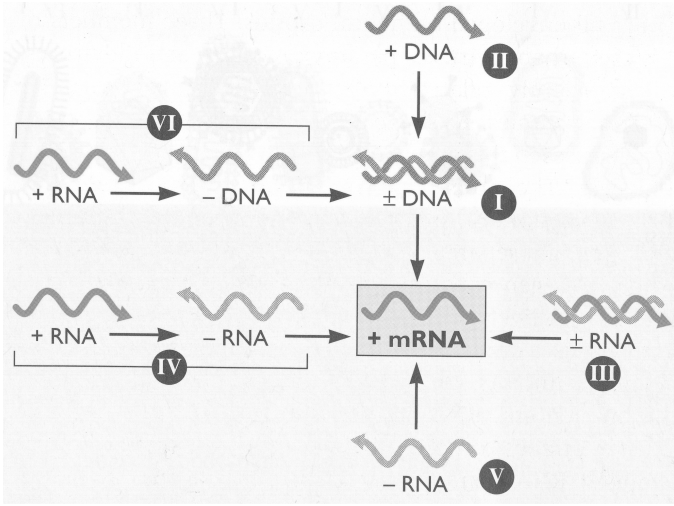


FIGURE 4.1 Schematic illustrating the replication strategies and means of gene expression for the various Baltimore classes. Reprinted with permission from Ref. (56). (See color insert.)

given group. Note that in the following sections nucleic acid genomes will be discussed in terms of strand polarity, where positive, (+), strand nucleic acid has a 5′→3′ polarity identical to mRNA and negative, (–), strand nucleic acid is complementary to mRNA. In cases of (–) strand polarity, the virus must synthesize a complementary (+) strand for gene expression to occur. A simplified schematic is presented in Figure 4.1.

Double-stranded DNA viruses (Class I): Viruses of this class have double-stranded DNA genomes and are subdivided into two groups. The first contains viruses that require replication to take place in the nucleus; the second class contains viruses that replicate in the cytoplasm. Variola major, the cause of smallpox, is an example of an enveloped double-stranded DNA virus that replicates in the cytoplasm; its linear DNA genome is directly transcribed to mRNA, which is then translated by host machinery to produce viral proteins. In contrast, genome replication of herpesviruses occurs in the nucleus.

Single-stranded DNA viruses (Class II): These viruses have single-stranded DNA genomes. Host proteins transcribe mRNAs from the viral genome, which are subsequently translated to viral proteins.

Double-stranded RNA viruses (Class III): Viruses in this class contain double-stranded, segmented RNA genomes. mRNA is synthesized by a virally encoded RNA-dependent RNA polymerase (RdRp) that is, like the RNA genome, contained within the capsid. Most eukaryotic cells do not encode RdRps, so in order for the virus to replicate itself, it must provide this enzyme

for RNA replication. The common etiologic agent causing severe infectious diarrhea in children, rotavirus, has a double-stranded, segmented RNA genome.

Positive-strand RNA viruses (Class IV): These viruses have single-stranded (+)-strand RNA that is directly translated by the host cell to produce viral proteins. The (+)-strand genomic RNA from these viruses can be infectious, in contrast to (−)-strand RNA, which cannot. To generate multiple copies of the genome, these viruses synthesize (−)-strand complementary RNA species that are subsequently transcribed by viral RdRps, to produce more (+) strands. Examples of viruses in this class include poliovirus, West Nile virus, the SARS coronavirus, and hepatitis A virus.

Negative-strand RNA viruses (Class V): In contrast to class IV, viruses of this group contain negative polarity single-stranded RNA molecules as their genome. These viruses are all enveloped and can have genomes that are either segmented or continuous. Some members have ambisense genomes with portions of the genome acting as (+) strands and other portions of the genome having (−) polarity. All members of this class, such as influenza virus, hantavirus, and Ebola virus package both genome RNA and an RdRp into their virion.

Retroviruses (Class VI): This unique class of viruses uses a totally novel scheme for replication and expression. These viruses have two identical copies of single-stranded (+)-polarity RNA molecules as their genome. These RNA molecules are reverse transcribed by the enzyme reverse transcriptase (RT), generating complementary DNA molecules from their RNA templates. Members of this class are called retroviruses, reflecting the fact that their replicative cycle is retrograde (RNA→DNA→mRNA→protein) relative to the central dogma of modern biology, in which the DNA is transcribed to mRNA, which is then translated to protein (DNA→mRNA→protein). A very important virus in this class is the human immunodeficiency virus (HIV), which causes acquired immune deficiency syndrome (AIDS).

Hepadnaviruses (Class VII): Members of this group have partially double-stranded DNA molecules as their genome. These viruses replicate via an RNA intermediate, similar to the retroviruses. RNA is packaged into immature particles where reverse transcriptase uses the RNA template to generate the DNA genome. An example is hepatitis B virus, an important human pathogen that can cause chronic infection and consequent liver damage.

Assembly and Release

Once the genome of a virus has been amplified, it must be packaged into infectious particles. This is a complicated process that is well understood for

only a few viruses. The exact mechanisms that allow selective packaging of viral nucleic acid and particular virus-specific proteins are not completely elucidated. The site of assembly can be in the nucleus, in the cytoplasm in association with membranous organelles (endoplasmic reticulum or Golgi bodies), or at the cell surface. The process of virion release occurs through budding or in the case of naked nonenveloped viruses, through host cell lysis.

VIRUS CLASSIFICATION

Viral nucleic acid, replication mechanisms, and virion release are only some of the characteristics that define a virus. Virologists have devised classification systems to describe and identify viruses based on observed similarities and differences. Initially, the decision of how to classify a virus was rooted in understanding the host and tissue infected by the virus, the resultant disease process, and the method of transmission. As technology advanced, the classification criteria shifted to include physical properties such as virion morphology, stability, filterability, and antigenicity. With the development of negative-staining electron microscopy that allowed scientists to visualize the virus particle, the classification of viruses changed to comprise descriptions of virion size, morphology, and surface characteristics. Finally, technological advances including polymerase chain reaction (PCR), DNA sequencing, monoclonal antibody production, and DNA microchip analysis have allowed researchers to differentiate between single-nucleotide changes present in viral genomes. The complete genomes of numerous viruses have been sequenced, allowing for the identification of species, strains, and mutants that could never before have been possible. When very little information is known about an emerging infection, the classification system and its wealth of descriptive information is useful; predicting possible routes of transmission and pathogenic consequences based on information known about similar well-studied and previously classified viruses may significantly impact public health practices while the new etiologic agent is under study.

The current standardized hierarchical International Committee on the Taxonomy of Viruses (ICTV) system used to classify viruses utilizes sequentially more exclusive levels of order, family, subfamily, genus, and species.[2] This classification system includes viruses that infect vertebrates, invertebrates, bacteria, fungi, plants, algae, and others. A brief description of the classification system follows.

VIRUS ORDER

Virus order is a collection of phylogenetically related virus families that have similar properties. At present there are only two recognized orders, *Monone-*

gavirales and *Nidovirales*. The *Mononegavirales* order is comprised of viruses that have a single-stranded negative-sense RNA molecule as their genome. There are three families in the order *Mononegavirales*: *Filoviridae*, *Rhabdoviridae*, and *Paramyxoviridae*. The *Nidovirales* order comprises the families *Coronaviridae* and *Arteriviridae*. The virus that causes SARS is a member of the *Coronaviridae*. Virus order names are discerned from the other classification level names by ending with the suffix *-virales*.

Virus Family/Subfamily

A virus family is a phylogenetically related collection of genera that share similar properties. Families are given names that end with the suffix *-viridae*. Virus particle morphology, genome characteristics, and/or strategies of replication are used to group viruses into families. Due to the diversity and complexity of the families *Paramyxoviridae*, *Poxviridae*, *Parvoviridae*, and *Herpesviridae*, each is further classified into subfamilies. Subfamilies are given names that end in *-virinae*.

Virus Genus

Names given to genera end with the suffix *-virus*. There is continuing pressure to use increasingly detailed structural, physicochemical, or serological differences to create new genera in many families.[3]

Virus Species

Species is a highly specific classification given to a virus that has met many selective criteria. In 1991, the ICTV defined a species as "a polythetic class of viruses that constitutes a replicating lineage and occupies a particular niche."[3] This definition allows for some plasticity in the properties required for species classification, allowing emphasis in some cases to be on genome properties, in others on structural, physicochemical, or serological properties. The term "quasi-species" refers to a population of closely related viral sequences. This is often used to describe RNA virus isolates where essentially every genome sequence is different due to the high error rate of RNA-dependent RNA replication.

The highly structured and formal classification system set up by the ICTV gives the virologist, clinician, technician, epidemiologist, and first-responder access to valuable information. Once a virus is identified and its classification is defined, information about the virus or similar better-studied viruses such as the kinetic characteristics of virus spread, tissue tropism, disease

pathogenesis, associated clinical symptoms, treatment, and level of required biocontainment can be quickly realized. This system also allows for classification and characterization of a new unknown virus. For example, SARS was a previously unknown virus that was quickly analyzed and placed into the *Coronaviridae* family.

GENERATION OF VIRAL DIVERSITY

Although the classification system facilitates virus identification, the constant emergence of new, diverse viral populations continues to challenge scientists. Three major mechanisms give rise to viral diversity: mutation, reassortment, and recombination. Depending upon the replication strategy of a particular virus, the relative contribution of these pathways differs. Nucleotide misincorporation leading to mutation occurs for all viruses but is dramatically different for viruses with RNA versus DNA intermediates. Errors are made both by DNA and RNA polymerases, but while DNA polymerases have proofreading capacity, RNA polymerases do not. Consequently, errors generated by RNA polymerases are incorporated into the newly generated viral genome at a very high rate. The error rates of DNA polymerases can be as low as one base substitution per every 10^8–10^{11} nucleotides, whereas the RNA polymerases have error rates as high as one base substitution per every 10^3–10^6 nucleotides.[4] In the case of RNA viruses, virus population studies of patient and tissue culture isolates have shown the coexistence of numerous viral genomes with slightly different sequences, or quasi-species.

Homologous or nonhomologous recombination is another mechanism by which new viruses are generated and has played a significant role in virus evolution. Large DNA viruses can not only recombine with each other but also contain genes usurped from the host. RNA tumor viruses have incorporated cellular oncogenes that lead to transformation in some cell types. For RNA viruses, recombination probably occurs via incomplete synthesis and template RNA switching during negative-strand synthesis. This process occurs with high frequency in poliovirus infection, with an estimated 10%–20% of poliovirus genomic RNA recombining in a single growth cycle.[5] In polio eradication efforts, a live virus preparation consisting of three separate attenuated poliovirus strains was used. Despite the safety of each attenuated strain, some vaccinated patients did develop poliomyelitis. Through genetic testing of patient isolates, it became clear that the recovered virus was not the same as the administered vaccine, and that neurovirulence had been restored by recombination among the three attenuated strains.[6–7]

Reassortment can also occur when two related viruses with segmented genomes infect a single cell. For example, reassortment is responsible for gen-

erating new influenza A strains that are antigenically different enough from previous strains to infect and cause disease. Besides the annual flu season, severe pandemics have occurred such as the Spanish flu of 1918, the Asian flu of 1957, the Hong Kong flu of 1968, and the Russian flu of 1977.[8] Influenza virus infects several animal hosts, including humans, pigs, horses, and its natural reservoir, waterfowl. The two major antigenic proteins of this virus, hemagglutinin (HA) and neuraminidase (NP), can be reassorted between different influenza viruses, generating new strains different from the two parental strains (called antigenic shift). Each influenza strain is given a code which identifies its HA and NP proteins. The Spanish flu was caused by the H1N1 strain, the Asian flu was strain H2N2, the Hong Kong flu was strain H3N2, and the Russian flu was strain H1N1.[9]

INTRODUCTION TO VIRAL KINETICS AND OUTCOME

A virus enters the host through any opportunistic opening of the body. Minute breaks in the skin barrier or mucosal linings of the eye, nose, gastrointestinal, and urogenital tracts create portals for the virus to enter host cells and underlying tissues. Replicating virus may establish a localized infection, or the virus, traveling as free virions or associated with immune cells, may spread to other areas of the body. Replication can continue at the initial site of infection or move to secondary organs and tissues. During an active infection, virus particles may be shed through nasal and respiratory secretions, urine, and stool.

Incubation Period

An incubation period is the time between exposure to virus and the manifestation of symptoms of infection. Initial symptoms may be slight, such as a rash or fever, and may not initially prompt patients to seek medical care. Diagnostic samples are taken when people seek medical treatment; therefore, the presence of virus and virus-specific antibodies depends on the incubation period. This observation has important implications for the identification of new viral infections or epidemics. Virus may be detected and isolated for infections with short incubations; however, antibodies may not yet have developed, see Figure 4.2. For example, infections with influenza and adenovirus have very short incubation periods, 1–2 days, before myalgias and fever manifest. Many viruses have invaded the body for over a week before symptoms become apparent. Dengue, measles, and rotavirus are relatively asymptomatic during the first 5–8 days, but may rapidly progress to severe illness soon after. Longer incubation

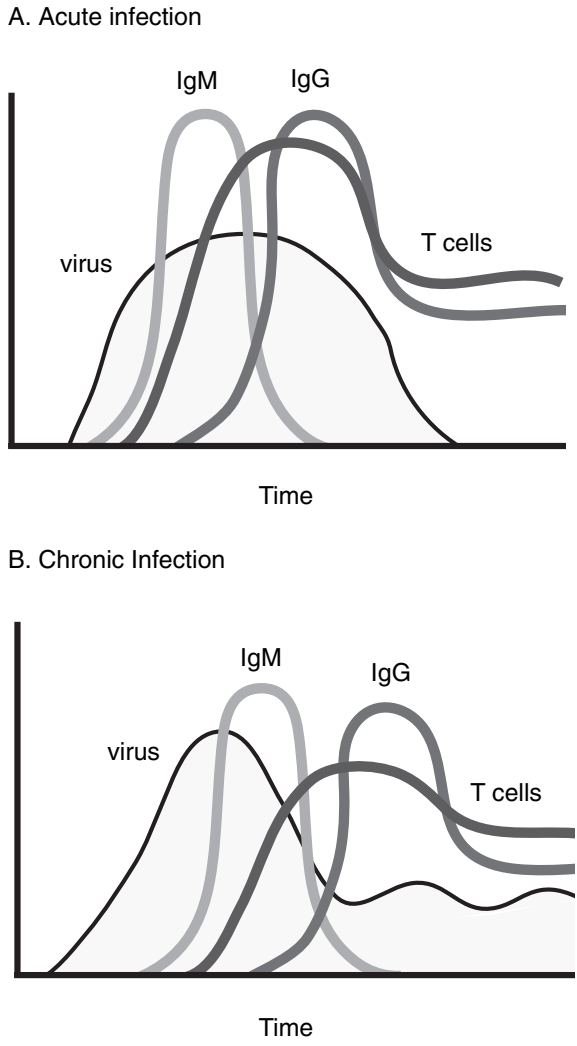


FIGURE 4.2 (A) Adaptive immune response in an acute resolving infection. Low-affinity antibodies of the IgM isotype develop initially, and partially control circulating virus. T cells undergo clonal expansion upon virus antigen stimulation. In response to effector T cells and often higher-affinity IgG antibodies, virus levels decline and the immune system differentiates into memory T cells. Such memory T cells and neutralizing antibody may persist for the life of the patient. (B) Adaptive immune response in a chronic virus infection. Antibodies and T cells develop but are unable to eliminate virus. Virus levels, although typically lower than in the acute stage, are continuously present despite a measurable immune response. Mechanisms leading to chronic infection include generation of viral mutants that escape the adaptive immune response, induction of immune exhaustion or tolerance, and a quiescent state called latency for some DNA and RNA viruses.

periods may be caused by chronic infections. For example, rabies virus may remain silent for 30–100 days after an animal bite before causing sudden disease demanding medical treatment. Other stealth viruses, such as HIV, hepatitis B and C, and papilloma viruses, may bring about a mild ailment initially that may gradually progress to a fatal disease. Patients with chronic infections may seroconvert and have detectable virus even in the presence of antibodies. Others may be persistently infected by a virus, yet never develop symptoms. These carriers may have low levels of virus, which could potentially be transmitted to other individuals despite being below the level of assay detection. An understanding of how the host responds to viral infection aids investigators and health care workers in identifying and treating infection.

The Immune Response

The immune system typically recognizes the invading pathogen at the site of entry, where it launches a generalized innate immune response. Signal molecules recruit inflammatory cells to the area which engulf cell debris and released virions. Virus proteins are presented to the adaptive immune system in local lymph nodes, allowing the host to create a directed defense against the virus. Effector T cells specific for virus antigen stimulate B cells to secrete virus-specific antibodies. There are two predominate types of antibodies produced in response to virus infection, IgM and IgG. The first antibody produced is the IgM isotype and is present in the blood within days and remains in circulation for several months. As the adaptive response develops, B cells switch antibody isotypes to secrete the IgG form. Investigators use the detection of antibodies as indicators of virus exposure, and can utilize knowledge of the presence of specific isotypes to discern acute infection from past or chronic infection. Detection of IgM or rising titers of IgG points to a recent virus exposure. Virus-specific antibodies are also used as tools in numerous serological assays. The specificity of the antibody is useful to distinguish between different virus types and to detect virus protein in clinical samples.

Course of Infection

Most virus infections cause an acute infection with viral particles rapidly cleared from the body. Examples of acute infections are influenza, rhinoviruses, and rotaviruses where infections are brief and virus is eliminated completely from the host. Other viruses are capable of establishing a long-term, persistent infection and may be present for years or throughout the life of the host. During a persistent infection, some viruses are continuously present in the

body with a constant level of viremia, or virus detectable in the blood. For example, lymphocytic choriomeningitis virus is noncytopathic even while actively replicating. Other viruses such as Epstein-Barr virus regulate their replication programs to avoid detection by the immune system, while still others establish infection in certain areas of the body that are immunoprivileged sites, including the central nervous system (CNS) and the eye, where once present, viruses may propagate undetected. Viruses may evade the immune response by infecting the immune cells themselves, such as in HIV infection of $CD4^+$ T cells and dengue virus infection of dendritic cells. Some viruses enter a latent stage where replication temporarily ceases. Reactivation of virus replication can occur weeks to years later, multiple times in the life of the host. Finally, viruses may reduce the rate of virus production so that little or no virus is detectable. As the immune system struggles to control, but not eliminate the infection, virus may be measured in a cyclical pattern with titers alternately rising above and falling below the threshold of detection.

HOW DO YOU IDENTIFY A VIRUS?

An understanding of the established classification system, mechanisms responsible for generating viral diversity, and pathogenesis facilitates the identification of an unknown virus. The sudden emergence of an infectious disease demands rapid methods to identify the source of infection, diagnose patients, and explain routes of transmission. Detection of a novel virus relies on adequate sample acquisition and knowledge gained from previously identified viruses, sequences, and reagents such as antibodies and virus antigen.

SITES FOR VIRUS ISOLATION

The types of samples and the manner of collection depend upon the capabilities of the diagnostic laboratory. The chance of obtaining an active virus sample depends on many factors. The virus concentration varies in the body over the course of infection and may actually not be present when symptoms bring the patient to medical attention. The virus may have been cleared from the area of disease, or the symptoms may reflect damage induced by the immune system rather than direct virus pathology. It is therefore critical to obtain samples early in the infection and from multiple locations in the body. The most common site to look for virus is in the blood. Other clinical samples are generally collected near the site of virus-induced disease or based on known routes of virus transmission. Mosquitoes can deposit West Nile virus and dengue into the skin; poxvirus and papillomavirus replication can create

lesions and warts; the perfuse spread of Ebola allows detection even in the skin. For respiratory infections, samples may include nasal and throat swabs as well as nasopharyngeal aspirates and bronchial lavage fluids. Urine and stool samples are collected for enteric disease, and cerebral spinal fluid is obtained when neurological symptoms are present. Viruses may be very labile, and care must be taken to avoid exposure to harsh treatments such as extreme pH, direct sunlight, and freezing temperatures in order to preserve the potential for identification and study. Generally, samples are transported on ice and evaluated in the lab as quickly as possible. If delays are expected, samples can be frozen, but the recovery of infectious virus may be reduced or eliminated.

Virus Isolation

Historically, viruses were cultivated in the laboratory utilizing human samples to directly inoculate susceptible animals or embryonic chicken eggs. Tissue culture allows virus to be isolated, quantified, and amplified to produce a large stock for further analysis. Primary cell lines derived from animal tissues have a limited lifespan in tissue culture, yet are necessary to initially isolate some viruses. Continuously replicating cell lines are the most convenient, with an infinite capacity to be passaged. Human specimens are clarified and treated with antibiotics before addition to cell culture. The inoculated cells are monitored for any detectable morphologic changes and periodically harvested for other serological assays. There are specific culture conditions for particular viruses. For instance, viruses from the conjunctiva, nasal and respiratory tracts are incubated at cooler temperatures resembling the exterior of the body, and viruses entering cells via fusion proteins requiring cleavage need to have trypsin added to the culture media. Some viruses cannot be cultivated in cell lines but require that clinical samples be injected into suckling mice by specially trained personnel and examined for virus using serological methods. Other culture conditions are being established that involve the use of whole-organ cultivation. Some viruses are even more difficult; hepatitis C virus can only be grown in chimpanzees.

Electron Microscopy for Virus Identification

In the early days of virology, researchers and clinicians could only identify and classify a virus based on its infectivity and associated pathology. Viruses could not be more closely described because of their extremely small size; virion particles can range from tens to several hundred nanometers (nm). The light

microscope, which is a valuable instrument for the identification and study of fungi, bacteria, mold, and spores, does not have the resolving capacity to visualize virus particles. Virologists were limited in their investigations until the 1930's, when the electron microscope (EM) was invented. This instrument uses, instead of visible light, a beam of electrons to form an image of the specimen. The instrument not only produces extremely large magnifications (up to 1,000,000×) but also resolves fine structure at such magnifications.

Virus morphology is as diverse as that seen in other microscopic organisms. Because of this, viruses can be identified based on their appearance under the microscope. The researcher is able to examine both the external and internal structure of the virus. Particle morphology has been observed to appear ribbon-like (rabies), rod-shaped (measles), spherical (poliovirus), and filamentous (Ebola) (Figure 4.3). Many viruses show multiple morphologies under the microscope and are referred to as pleomorphic. An example of a very pleomorphic virus is influenza A, which can appear kidney bean-shaped or filamentous. Some viruses have very characteristic spikes, club or pin-like projections present on the viral envelope. A very interesting example of spikes is seen in the emerging SARS virus.

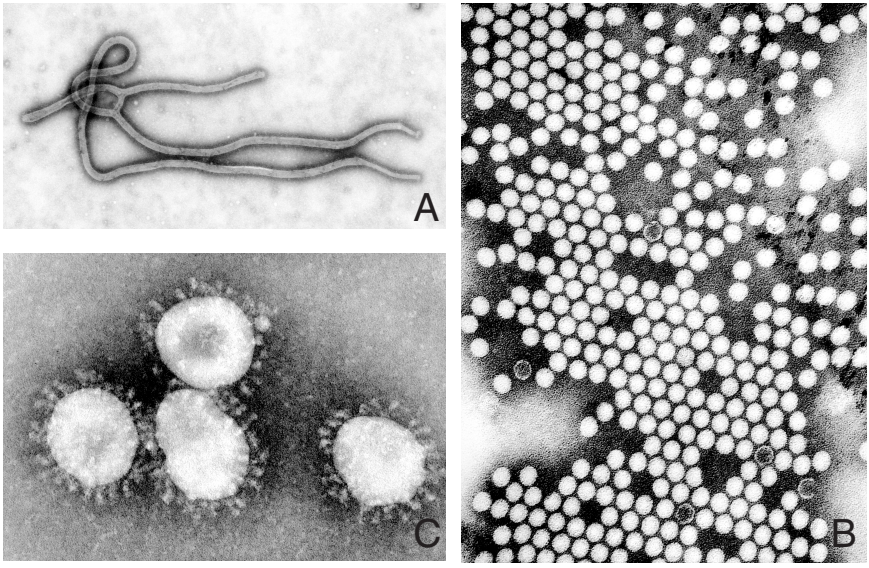


FIGURE 4.3 (A) Ebola virus particles showing filamentous morphology. Courtesy of the CDC/C. Goldsmith. (B) Polio virus particles showing spherical morphology. Courtesy of the CDC/Fred Murphy, Sylvia Whitfield. (C) SARS virus showing indicative "corona-like" morphology. Courtesy of the CDC/Fred Murphy.

There are two main types of visualization for virus identification using EM. The first is negative staining/contrast, which was a revolutionary visualization technique designed in the late 1950s. Due to the low level of electron scatter by viral particles, direct contrast and visualization are difficult. However, using the negative contrast technique, the virus particle is visualized on a black (electron-opaque) background. The degree to which the heavy metal salt stain (electron-rich) penetrates into the virus particle determines the contrast/resolution of virus structure. This procedure provides very useful information on the external structure of the virus, but is somewhat limited in resolving internal structure.

Thin sectioning is a very important sample preparation technique that allows internal structure to be visualized. Tissue samples are embedded with an epoxy-resin, and thin slices are prepared using an ultramicrotome. These slices are then examined using various electron-dense stains and antibodies that can be used to tag specific cellular organelles and virus proteins.

Observance of Cytopathic Effect

A virus may kill an infected cell creating characteristic cytopathic damage or may replicate in cells without any visible effect. The type of changes induced by a virus can be significant for the type of virus and can be observed by the light microscope. The cytopathic effect (CPE) can be focal, diffuse through the cell monolayer, or at the edge of the culture. Cells may appear to be rounded or enlarged, growing in grapelike clusters, indicating adenovirus or herpes simplex virus. Influenza and mumps cause cells to fuse together, creating syncytia that detach from the surface. Cells may fuse and form multinucleated giant cells with granular cytoplasm typical of measles infection. Vaccinia and poxviruses create foci of fused cells, whereas picornaviruses induce proliferation of membranes in the cytoplasm and shrinkage of the nuclei (pyknosis). However, many viruses such as members of *Bunyaviridae*, *Arenaviridae*, and *Retroviridae* fail to produce obvious CPE and can replicate in culture without any noticeable change.

Serology

The specific affinity of antibody-antigen recognition is widely used in virus diagnostics. Polyclonal and monoclonal antibodies can be raised against recombinant virus proteins and virus produced from infected cells. If a virus can be isolated in cell culture, cell lysates and slurries can be used as virus antigens. Infected cell cultures can also be used to test for antiviral reactivity

present in patient serum. Recombinant virus proteins produced in *E. coli* and baculovirus expression systems are also used as antigens to detect antibodies present in human serum, cerebrospinal fluid (CSF), and tissues. Finally, convalescent serum proven to react with infected cells in culture can be used as a source of antibodies for additional assays.

NEUTRALIZATION ASSAY

Antibodies produced during an infection often have the ability to bind the virus and reduce infectivity. These protective, neutralizing antibodies recognize epitopes on the surface of the virus and prevent virus from infecting a cell. Neutralizing antibodies are often used to classify virus into serogroups. Closely related virus families have similar virus coats, and neutralizing antibodies that can bind viruses within a group are said to cross-react. In a neutralization assay, dilutions of neutralizing antibodies are mixed with virus and assayed for remaining infectivity. An unknown virus may cross-react with antibodies from a known serogroup, revealing an antigenic relationship and an initial clue to the genetic identification. More specific and focused assays can then be performed to identify the virus. An extensive bank of serotype-specific antibodies in the diagnostic laboratory increases the chance of discovering a novel virus.

HEMAGGLUTINATION ASSAY

The hemagglutination assay uses the ability of virus proteins to bind and aggregate erythrocytes. Virus is mixed with an erythrocyte suspension in serial dilutions and added to a microtiter plate with V-bottom wells. Unabsorbed red blood cells fall to the bottom point of the well forming a dot, whereas aggregated blood cells uniformly coat the well. The hemagglutination inhibition assay is used to classify virus families that share that ability to bind to erythrocytes. Antibodies specific for a virus family may prevent aggregation of erythrocytes and are assayed by hemagglutination inhibition.

COMPLEMENT FIXATION

Complement fixation uses known virus-specific antigen rather than known anti-virus antibodies. Interactions between virus antigen and antibody cause fixation of complement and result in membrane lysis. Briefly, patient sera are incubated with antigen and a standardized amount of complement. Red blood

cells are coated with anti-red blood cell antibodies that are recognized by complement and added to the sample. If antibodies specific to the virus antigen are present and fix complement the red blood cells will be protected from lysis. Using group-specific virus antigens allows a patient's virus to be identified.

Immunostaining

Antibodies can be used to detect virus antigen in patient tissues and in infected cell cultures. An antibody is linked to a fluorescent dye or enzyme and allowed to bind virus antigen. The complex is visualized under a microscope, allowing location of the virus proteins to be seen inside the cell. Cross-reactive sera can be used to serologically identify an unknown virus as a member of a known family. Polyclonal sera can identify a wide range of viruses, whereas monoclonal antisera can give a specific diagnosis. Additionally, the direct visualization of virus antigen in a cell can provide interesting clues of virus properties beyond the diagnosis.

Enzyme-linked Immunosorbent Assay

The enzyme-linked immunosorbent assay (ELISA) provides a fast method to detect the presence of virus antigen or antibody from a large number of samples. Because of its reproducibility and flexibility, it is the most common serological assay used in the viral diagnostics lab. The general scheme is to capture virus antigens or virus-specific antibodies on a solid surface and to expose the bound complex with a substrate. ELISAs specific for IgM and IgG antibodies can distinguish between recent infection and previous exposure or vaccination. Detection of virus antigen indicates an acute stage of virus infection. The microtiter plate format can test multiple patient samples with different dilutions of antibodies to determine a serological titer, and sequential sampling of patient's serum over the course of an acute infection can allow detection of seroconversion. An increase in titers of IgG antibodies twofold to fourfold over 2 weeks is diagnostic of an active primary infection. ELISAs are frequently used in serosurveys where populations are screened to determine the efficacy of immunization programs or previous exposure to a newly identified virus.

Nucleic Acid-based Methods

The detection of virus nucleic acid has revolutionized diagnostics. The use of molecular techniques has identified viruses that cannot infect cells in culture

or inoculated animals. Most specimens such as blood, tissues, urine, stool, CSF, and respiratory secretions can be treated and the nucleic acid extracted, inactivating the virus and reducing concern about transport or contact in the laboratory. Detection of virus in the blood is diagnostic of an active infection, and the sensitivity of PCR has enhanced our capability to detect very low levels of virus in persistent infections; the universal genetic code allows investigators to detect virus sequence from host insects, animals, bacteria, and fungi where serological reagents may not be available. The specificity of PCR is dependent on the selection of the primer sets. Primers are designed to anneal to conserved regions in a virus genome. Genes that encode proteins essential for the virus life cycle such as polymerases, helicases, and integrases as well as RNA elements in untranslated regions are maintained even in viruses with high mutation rates. There are good PCR amplification targets. To identify a new virus, universal primer sets are designed based on alignments of virus family sequences and selected to minimize base pair mismatches. Primers containing degenerate bases are used to widen the net and find a genetically distant virus. For DNA viruses, PCR can be applied directly. RNA viruses require an initial reverse transcription RT-PCR reaction before amplification. Complex variations on the basic technique of PCR including utilization of multiplex primers, TaqMan probes, and molecular beacons can enhance the ability to identify viral sequences.

PCR products are either sequenced directly or cloned into intermediate vectors for sequencing. Multiple PCR isolates are sequenced in order to verify the authentic virus sequence, as mutations arise during PCR amplification. Sequencing virus directly from clinical samples without amplification in cell culture minimizes cell culture adaptation artifacts. The unknown virus sequences are entered into sequence databases to search for any similarity with known virus genomes. The database should include sequences from clinical isolates, field samples from the natural hosts, and nonpathogenic viruses to assist identification of a novel emerging virus. As with any assay, molecular diagnosis of virus nucleic acid is confirmed with other diagnostic methods such as serology and virus isolation.

MOLECULAR EPIDEMIOLOGY

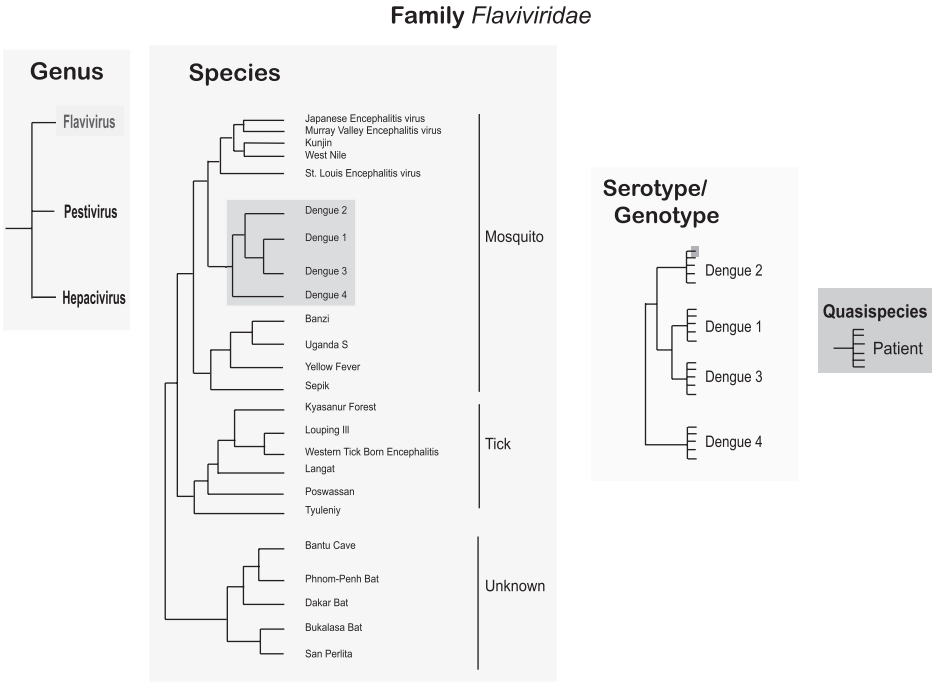
Traditional epidemiology relies on the combination of clinical presentation of disease, identification of the pathogen, and anecdotal circumstances to explain where the infection began and how it spread throughout a population. Common features are established to link the transmission to the source of infection. The addition of molecular tools to the investigative effort allows the

infectious agent to be identified at the genetic level and enhances our understanding of virus origin, emergence, and transmission.

Molecular epidemiology uses phylogenetic methods to reconstruct a path of virus transmission based on heredity. As a virus replicates and moves through a population, mutations accumulate. The genetic variability displayed by the viruses is compared to deduce common ancestors and explain how one virus sequence gave rise to another. Viruses evolve, sometimes very quickly, to adapt to new hosts and environments. The genetic makeup of the pathogen may increase in variability with time and passage. Closely related virus sequences therefore correlate with recent infection and transmission. Using PCR and molecular tools, virus sequences are collected during an outbreak from infected human, animal, or insect reservoirs. These new virus sequences are then compared with a database of known virus sequences. The comparison is not a simple match of nucleotides between the virus genomes; rather, it involves sophisticated algorithms creating clusters of virus sequences sharing a common evolutionary ancestor.

Phylogenetic trees display how a set of virus sequences might have been derived during evolution, and provide guides in the placement and classification of an unknown virus. The trees are a graphical illustration of the evolutionary linkage of newly isolated virus sequences with known virus genomes (Figure 4.4).[10] The outer nodes of the tree display existing virus sequences. The inner branches represent theoretical ancestral virus sequences that gave rise to the recently isolated virus genome. The length of each branch corresponds to the amount of genetic change between the ancestral virus and the currently circulating virus. Additional algorithms may be applied to place a temporal scale with the amount of evolutionary change. A viral sequence that is very distantly related may be used as an outgroup to orient the tree with a direction of evolutionary change. Phylogenetic analyses revealing ancestry enable initial hypotheses about its basic life history, including the virus hosts and transmission patterns, as close relatives tend to be similar in their biology.

The phylogenetic tree can be interpreted as clusters of virus sequences that are used in classification schemes, and often these groupings complement traditional methods. Broad virus families can be defined using sequences from conserved genes such as polymerases and other enzymes required in the virus life cycle. Finer distinctions are noted by using sequences from genes that are more specialized for a particular virus. For example, envelopes and genes encoding structural proteins are often used to define virus subgroups within larger families. These subgroups often correspond to traditional serogroups defined by traditional serological methods that define groups by their ability to cross-react to a particular antibody. Entire genes or portions of genes can be used in the phylogenetic analysis, but complete virus genomes



Adapted from Holmes and Twiddy. Infection, Genetics and Evolution., 3:19-28. 2003

FIGURE 4.4 Phylogenetic tree of the RNA virus family *Flaviviridae*. Originally defined by serological assays, the phylogenetic tree shows inferred genealogical relationships based on maximum likelihood analysis of nucleic acid sequences. The family is divided into separate genera, *Pestivirus*, *Hepacivirus*, and *Flavivirus*. Flaviviruses are further subdivided into species that fall into three groups that can be transmitted via ticks or mosquitoes or without an arthropod vector (black vertical lines). Dengue viruses are separated into four serogroups or genotypes based on serology and envelope protein divergence. Within each Dengue serotype/genotype, representatives responsible for outbreaks of disease have been isolated from around the world. Given that RNA viruses mutate rapidly, a given isolate from an infected individual exists as a nonhomogeneous population variant termed a "quasispecies." Figure adapted from Ref. (10).

are rarely used, as the sequence length is restricted by the amount of computational time.

Epidemiological questions can be resolved with phylogenetic trees. The most fundamental questions in epidemiology are the mode of transmission and the origin of an outbreak. As the virus is passed through a population, genetic differences accumulate. More recent infections generally correspond to more

shared derived genetic changes, and the most closely related will cluster on a tree. Transmission patterns are revealed as virus sequences from isolates are compared with sequences from a database and the clusters of sequences reveal a common ancestor.

An immediate question to solve during a virus outbreak is the mechanism by which the virus spreads. Viruses frequently infect animal or insect vectors that serve to pass the virus to humans. By identifying the virus through sequence analysis, a hypothetical reservoir can be predicted by the placement of the sequence on the phylogenetic tree, since viruses that share a mode of transmission often cluster together. More than one type of vector may be used within the same family of viruses, but individual members depend on a particular vector. Viruses from the family *Bunyaviridae* can be transmitted to humans by such pests as ticks, mosquitoes, flies, and rodents. The particular vector utilized often distinguishes individual subgroups within the family. Phylogenetic trees have been instrumental in proving virus transmission through family members, hospital settings, and susceptible members of a population. By comparing the virus sequence isolated from a patient to virus sequences isolated from individuals in the population transmission routes can be deduced. For example, it can be concluded that a doctor became infected from a patient seen at a hospital rather than from the general community if the virus sequences from doctor and patient are more similar than sequences found in the population.

The geographic location of virus infection can also be predicted by the use of phylogenetic trees. Viruses isolated from reservoirs in different geographic locations can define where people became infected. Arenaviruses are carried by rodents, but domestic rodents are often responsible for causing infections rather than rodents found in the fields or forests. By analyzing virus sequences obtained from infected patients and from house, field, and forest mice, the group of mice serving as a reservoir can be determined, and public health measures to prevent rodents from entering homes can be implemented. The origin of a virus outbreak can involve larger geographic areas. As the SARS outbreak demonstrates, importation of new viruses is an increasing global health concern. By accessing a database of virus genomes from around the world, a new virus outbreak could be rapidly identified as having been imported from another land. As viruses travel, some viruses become endemic, and it is necessary to distinguish between endemic and imported outbreaks to allow appropriate control measures to function. Viruses may also cause seasonal outbreaks. Influenza varies its genome with each year of infection. By understanding the extent to which the virus alters its genome from year to year, phylogenetic analysis assists in predicting the next year's strain and allows scientists to begin production for the upcoming yearly vaccine.

EXAMPLES OF NATURALLY EMERGING VIRUSES

SIN NOMBRE

In May 1993, a clinician reported a cluster of cases presenting with nonspecific illness, fever, headache, and cough that rapidly progressed to respiratory distress. The patients were young, previously healthy, and lived in rural areas in the four-corners region bordering Arizona, Colorado, New Mexico, and Utah. By early June 1993, 24 cases had been reported with a mortality rate approximately 75%.[11,12] Patient blood and autopsy tissues were sent to the Centers for Disease Control (CDC) for investigation. Diagnostic assays failed to identify known causes of respiratory illness; however, patient samples cross-reactived with antibodies to four different hantaviruses. Using hantavirus sera, immunohistochemistry detected hantavirus antigen in autopsy tissue from the lung and other involved organs.[13] Hantaviruses are RNA viruses with a genome in three segments and a member of the family *Bunyaviridae*. Previous outbreaks of hantaviruses involved patients with renal disease leading to hemorrhagic complications, not respiratory distress. Hantaviruses are classified into four antigenic groups, and each strain of hantavirus is tightly associated with a specific rodent host. Although hantaviruses have a worldwide distribution, the dependence on a particular species of rodent host restricts the geographical range. Hantaan (HTN) is associated with the field mouse in Korea, China, and eastern Russia and causes hemorrhagic fever with renal syndrome (HFRS). Seoul (SEO) causes moderate disease in Korea and China and is associated with urbanized rats. A milder form of HFRS is caused by the serogroup Puumula (PUU) in Scandinavia and Europe and is transmitted by the bank vole. At the time of the four-corners outbreak, the only North American hantavirus identified was Prospect Hill (PH), which was detected in two rodent isolates of meadow voles and not associated with any human illness. Patient samples from the U.S. four-corners region cross-reacted with hantavirus antigen and allowed epidemiologists to expand their assays to include RT-PCR and narrow the search for the host to local rodents. The virus was unable to be cultured from patient samples; therefore, molecular identification was crucial for identification. Partial sequences were available for HTN, SEO, PUU, and PH in a database and used to design PCR primers based on conserved areas of the virus glycoprotein. Nucleic acid was extracted from autopsy tissue, and a PCR product was detected in two cases using primer pairs based on PUU and PH. The PCR product was sequenced, revealing a 70% sequence similarity between the new virus and PH and PUU.[14] New primers were designed and could detect products in all current cases. The sequence of the new virus was identified, and the nucleocapsid protein was expressed in bac-

teria and purified and used as a recombinant antigen in other serological assays.[15] The new virus was eventually named Sin Nombre virus (SNV).

Rodents were trapped in areas in and surrounding patient houses. The predominant species was the field mouse, *Peromyscus maniculitis*, and approximately 30% had anti-hantavirus antibodies detected in an IgG ELISA. PCR revealed that 82% had detectable virus in the blood, indicating high viremia even in the presence of circulating antibodies. Sequences from rodent and human cases isolated in the four states were aligned, and identical nucleotide substitutions were present in the rodent and human cases from the same location, indicating that transmission occurred from the field mouse to the human cases.[14] Transmission occurred by inhalation of aerosolized rodent urine. Ecologists reported an increase in the population of field mice during that year due to a surplus food supply. The arid conditions of the southwest and the close proximity of the mice and the patients also contributed to transmission. In this outbreak, there were no reports of human-human transmission.[16]

The discovery of SNV led to a heightened awareness and search for other American hantaviruses. Field investigations and retrospective serological studies identified additional hantaviruses in North America.[17] The North American viruses show an 80%–95% similarity, but are always found only in their specific rodent species.[18] The detection of the new hantaviruses is enhanced by the use of diagnostic reagents developed in this outbreak and refined as new viruses are identified. Additional cases have also been detected in South America. In 1995, Andes virus was isolated in Argentina, the only documented case of human-human transmission.[19] Hantavirus outbreaks have also been identified in Chile, Paraguay, Bolivia, and Brazil, causing both respiratory illness and some renal hemorrhagic disease. South American cases show higher virus titers and also persistently infect a distinct rodent species as a natural reservoir.

After the SNV outbreak, efforts increased to study the relationships between hantaviruses and the rodent hosts worldwide. Phylogenetic analysis of rodent and human virus sequences has shown a close evolutionary relationship, and hantaviruses are often used as models of coevolution. A particular genotype of hantavirus persistently infects one species of rodent even though other rodent hosts may be present in the same environment. This remarkable coexistence has been mapped through changing landscapes. As a rodent species radiates farther away into another habitat, it may evolve into a separate, but genetically similar, species. The hitchhiker hantavirus also evolves with the rodent host, becoming genetically distinct from viruses associated with the ancestral species. This ancient relationship has been mapped using phylogenetic trees, which display common ancestral species as clusters that diverge and separate with geographical and genetic distance. The topology of the rodent tree can be compared and reveals similar patterns with topology of the associated

hantaviruses.[20] Phylogenies have been performed on rodent family, subfamily, genus, and species using mitochondrial DNA. Phylogenetic trees of hantaviruses reveal a similar branching pattern to the rodent tree. Viruses that infect the same subfamily of rodents form identical clades, and closely related hantaviruses are grouped in the same tree pattern as their rodent reservoirs. Although rare, there have been instances of transmission of one virus to another rodent species. This provides another example of molecular evolution where virus segments can mix in the host and form reassorted viruses containing combinations of different virus segments. This can contribute to genetic diversity. Reassortment of the segmented genome has been found in nature and in mixed virus infections in tissue culture, but only between very closely related viruses.[21] Phylogenetic analysis performed on more than one virus segment tests for the chance of reassortment. Even in geographical locations where different rodent species may overlap, virus-host switching and reassortment are very rare events. Each hantavirus is associated with a primary rodent host, and usually a particular rodent species will be the primary reservoir of a single hantavirus. The tight correlation of virus and rodent host in the phylogenetic trees supports the concept of coevolution between virus and host. All subfamilies of the family *Muridae* have a hantavirus-rodent relationship. Fossil records estimate the rodent families *Muridae* to have split into the current subfamilies during the Miocene period, indicating that hantaviruses have been associated with their rodent hosts for 30 million years, an association that began before the division of the rodent family.[22] The coevolutionary relationship and worldwide distribution of hantavirus illustrates that the 1993 SNV outbreak was not due to the emergence of a new virus or virus with an increased virulence, but simply the detection of an ancient virus.

NIPAH

Between the autumn of 1998 and spring of 1999, an outbreak of acute fever with encephalitis occurred in peninsular Malaysia and Singapore, with 265 cases resulting in 105 deaths. The outbreak was associated with a respiratory illness of pigs, and the primary human cases were individuals who had close contact or occupational exposure to pigs.[23] The initial report was released in late September 1998, attributing the cause of illness to mosquito-borne Japanese encephalitis virus (JE). JE is endemic to the region and was believed to involve pigs as amplifiers of virus. Ministries of Health began mosquito control measures and vaccination campaigns to block JE infection.[24] By February 1999, the outbreak continued to spread as pigs were moved from the western-central region in the state of Perak to three other farms located in the

southern states of Selangor and Negri Sembilan, the largest pig-rearing areas in Southeast Asia. Pigs were also exported to two farms in Singapore and resulted in 11 cases and two deaths. Singapore suspended pig importation from Malaysia, and no more cases were reported. Epidemiological analysis discredited JE as the infectious agent of the outbreak, since communities located near the pig farms did not become infected and immunization and vector control programs were having no effect on the increase of cases. As the outbreak continued, new attempts were made to find the pathogen.[25] In March 1999, a virus was isolated by inoculating CSF obtained from fatal cases into Vero cells. The presence of multinucleate syncytia in infected cell culture implicated a paramyxovirus. Thin-section EM detected filamentous nucleocapsids with the hallmark herringbone pattern and also captured virus budding from the plasma membrane, both characteristic of paramyxoviruses.[26] The new virus was named Nipah. A fortuitous event in the identification of the virus was the observed cross-reactivity between sera against Nipah and the paramyxovirus, Hendra virus. No other paramyxovirus sera reacted with Nipah samples.[25]

Hendra virus was identified in 1994 as the cause of respiratory illness in three regions of Queensland, Australia, resulting in the deaths of thoroughbred horses. A horse owner also died of respiratory distress and a stable-hand was infected, but survived. Another individual who assisted with a horse autopsy became ill with a meningitis-like infection and recovered (temporarily), but one year later developed progressive neurological complications and eventually died.[27] Hendra virus was isolated from his CSF. To test for other potential infections, reagents for an IgM and IgG ELISA and antibodies for IHC were developed.

Hendra-specific antibodies and antigen were used to screen initial Nipah cases. Eventually, Nipah-specific reagents were developed for larger serosurveys. In pig tissues, the major source of antigen and pathological damage was seen in the epithelium of the upper and lower airways. Most of the pigs in case farms were IgG-positive for Nipah, yet only a minority displayed any illness and only 5% were fatal.[28] Pigs had a definite respiratory disease comprising a distinctive cough with some developing encephalitis. The primary mode of transmission was suspected to be from respiratory secretions and aerosolized droplets from infected pigs. Domesticated dogs and cats, a few ponies, and bats also tested positive for Nipah. Human autopsy results revealed that the primary pathology was secondary to multiorgan vasculitis due to infected endothelium where multinucleate syncytia with viral antigen were detected. The cerebral cortex and brainstem were also damaged, and neurons were found to have virus inclusion bodies. Nipah antigen was also detected, but to a lesser extent in the lungs, heart, kidney, and spleen, consistent with a systemic infection.[26]

An RT-PCR assay was developed using degenerate paramyxovirus primers annealing to the P gene. Sequencing revealed a 78% similarity between Nipah and Hendra, but less than 50% identity to other members of the *Paramyxovirinae* subfamily.[29] The *Paramyxoviridae* family is divided into two subfamilies, the *Paramyxovirinae* and the *Pneumovirinae*. The *Paramyxovirinae* is separated into three genera; however, phylogenetic analysis did not cluster Nipah and Hendra into any one of the three genera. The *Paramyxovirinae* have a genome size around 15 kb, while both Hendra and Nipah have genomes of 18.2 kb due to a long untranslated region at the 3′ end of the genome.[30] The genome size and phylogenetic differences place Nipah and Hendra into the newly created *Henipah* genus.

The origin of the outbreak was traced to large fruit-eating bats, which are typically named flying foxes. Virus was isolated from urine collected under roosting areas and saliva from dropped fruit by the island flying fox, *Pteropus hypomelanus*. Antibodies have also been detected in the Malaysian flying fox, *Pteropus vampyrus*. Pig farms have fruit trees located near open pig stalls.[31,32] Flying foxes may transmit virus by urine droplets near the pig stalls or food source, or the pigs eat fruit contaminated with flying fox saliva. The outbreak was stopped after culling 1.1 million pigs.

WEST NILE

The West Nile virus outbreak illustrates several important points. First, it illustrates the difficulty in identifying a virus, especially one that had never been seen before in the U.S. Second, it shows that even with an initial misidentification, state and federal agencies were still able to act effectively, knowing how similar viruses could be controlled. And last, this case illustrates that virus distribution can be tracked by monitoring those infected as well as its vector.

In late August 1999, eight concurrent cases of patients having encephalitis and/or profound muscle weakness surfaced in Queens, New York. Geographically, the reported cases were traced to a two by two-mile area of a residential neighborhood, immediately sparking exposure concerns. By the end of the year, 59 patients were hospitalized in New York City, and seven deaths were reported.[33] Cases presented with a mild 3–6-day symptomatic period, including sudden onset of malaise, nausea, vomiting, headache, rash, cough, and sore throat. Less than 1% of those infected developed neurologic disease (i.e., encephalitis);[34] reported deaths were among the young, elderly, or immunocompromised. Also during this initial period it was noticed that several birds throughout the New York area were dying. The birds seemed to be dying from a neurological condition. Samples were sent to the CDC for identification.

Shortly after initial cases were reported, the CDC believed that the infectious agent was likely to be the endemic St. Louis encephalitis (SLE) virus. The CDC based their conclusion on serologic (IgM-capture ELISA) findings, which tested for the presence of patient antibodies to SLE antigens. The laboratory findings of SLE as the infectious agent were consistent with the symptoms of those infected, as well as the fact that SLE was responsible for 4,478 confirmed cases in the U.S. from 1964–1998.[35] SLE is an arbovirus, or a virus transmitted to human via arthropods, such as mosquitoes. Steps were quickly taken to monitor and control mosquito populations.

However, shortly after their initial assessment, the CDC changed their conclusion, and stated that the infectious agent was West Nile virus (WNV), not SLE. They based this on the more discriminating method of sequencing viral RNA isolates via RT-PCR, as well as results from specific monoclonal antibody detection of WNV-specific antigens (envelope glycoproteins). The controversy continued as other independent researchers believed the virus to be Kunjin virus, another closely related arbovirus.[36] It was later confirmed that the agent was indeed WNV through independent genetic testing by the U.S. Army Medical Research Institute for Infectious Diseases (USAMRIID).[33]

The cause for confusion was due to the close relationship between WNV and SLE. Both flavivirieses share significant similarity in their envelope proteins and some antibodies cross-react. Another compounding factor was that WNV had never been seen before in the U.S., although it was responsible for outbreaks in other geographic locations, including Africa, Europe, and the Middle East.

The diagnosis of WNV infection is made by both clinical findings and serologic tests, which detect patient IgM antibodies present in CSF or serum.[34,37] Direct virus isolation has been used, but due to limited sensitivities, is not a primary diagnostic procedure. RT-PCR methods for detecting WNV-specific RNA have been used extensively for detection in tissues and CSF. Blood banks now use RT-PCR to screen the blood supplies, due to the high incidence of asymptomatic WNV-infected donors: 23 persons have been reported to have been infected with WNV after receiving transfusions from 16 WNV-infected blood donors.[38,39]

West Nile virus was also identified as the agent responsible for the deaths of birds observed throughout the New York area. This has become a very important observation, for geographical mapping of the dead birds is valuable for tracking WNV infection as it continues to spread to other states and countries. The CDC stated: “dead-bird-based surveillance has proven to be the most sensitive method for detecting WNV presence in an area.” However the CDC goes on to state “mosquito-based surveillance remains the primary tool for quantifying the intensity of virus transmission in an area, and should remain the mainstay in most surveillance programs for WNV and other arboviruses.”

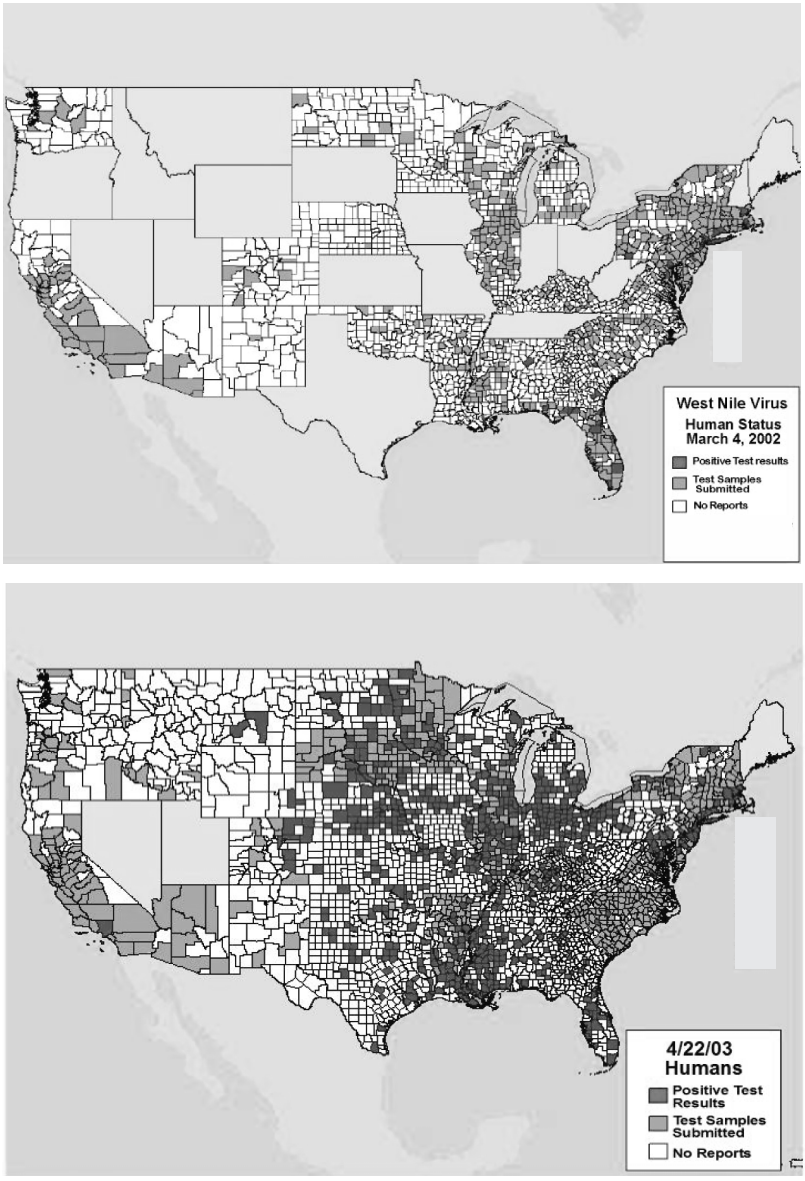


FIGURE 4.5 Spread of West Nile virus. From the initial outbreak in New York in 1999, the maps chart the rapid westward spread of WNV by showing the geographic distribution of human cases in years 2002–2004. Courtesy of the CDC. (See color insert.)

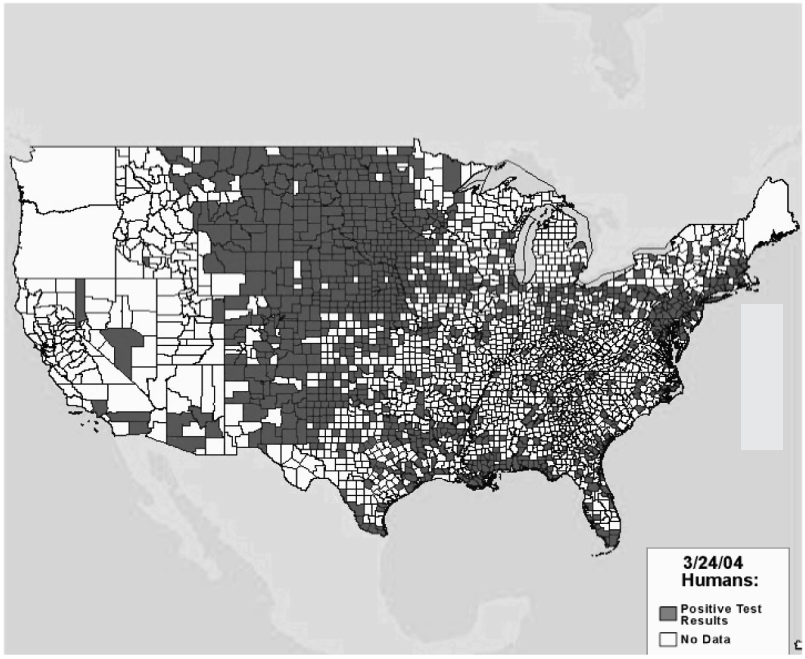


FIGURE 4.5 (*Continued*)

In addition to mosquito and dead-bird surveillance, the CDC as well as state organizations tracked live birds, equine WNV infections, and obvious human infections. These data showed that the virus was continuing to spread westward (Figure 4.5). Computer-assisted tracking services like Arbonet and the U.S. Department of the Interior website,[40] provide up-to-date tracking information. Such data allow for efficient knowledge acquisition regarding geographic distribution and changes in virus intensity, which gives the public the ability to take necessary precautions.

WNV continues to be a problem with over 4,000 cases reported in the U.S. in 2002. Accurate up-to-date tracking of infected humans, animals, and mosquito populations will continue to provide extremely valuable information about virus spread, locations at risk, and the efficiency of controlling arboviruses through mosquito control. Surveillance efforts must be supported to allow effective action against WNV and against future arbovirus epidemics as they occur.

SARS

The worldwide outbreak of SARS demonstrates the importance of rapid diagnosis of an infectious disease to understand how the disease spread and could be contained. The SARS outbreak was confounded by many problems. It is spread by close person-to-person contact allowing localized outbreaks in households and hospitals, and it has an incubation period of 7–10 days, allowing infected individuals enough time to travel anywhere in the world. The symptoms are very general: a high fever and cough or difficulty breathing, making it difficult to trace the spread by clinical presentation.[41] During the early course of the outbreak, investigators did not know the agent causing the infection.

In November 2002, several hundred cases of severe, atypical pneumonia were reported from the Guangdong province in southern China. An infected medical doctor from Guangdong carried the virus to Hong Kong, where he stayed for one night on the 9th floor of the Hotel Metropole. Subsequently, at least 16 guests and visitors to the 9th floor of the hotel became infected and spread the illness to Vietnam, Singapore, and Canada. Hospitals became new epicenters of infection as health care workers and family members became infected. In March 2003, the World Health Organization (WHO) recognized that SARS had become a worldwide health threat, and released a global alert on travel to screen air passengers departing from affected areas. Cases of suspected SARS continued to be reported in new areas including Thailand, Europe, the United Kingdom, and the U.S., but the most heavily affected areas were in China and Hong Kong. In early April, the virus was isolated and identified as a novel coronavirus, SARS-associated coronavirus, or SARS-CoV. The numbers of cases increased at a staggering pace with over 5,000 cases on April 28, 6,000 cases on May 2, and 7,000 cases on May 8 from over 30 countries on six continents.[42] Travel warnings were issued for Toronto, Canada, where an imported case of SARS established an outbreak in a hospital with over 100 cases, which subsequently spilled over into the community. Travel restrictions also applied to Beijing and Guangdong, China and Hong Kong. Although transmission was assumed to be primarily from secreted virus in respiratory fluids, another mode of transmission was found in an outbreak in a Hong Kong apartment building. Improper plumbing in the Amoy Gardens apartment complex probably allowed virus-contaminated sewage to be introduced into the water supply, resulting in almost 300 cases from one building. Epidemiologists began to link a few individuals as sources responsible for infecting a large number of people. These "super spreaders" are individuals who for unknown reasons were capable of transmitting infection to over 100 people. In Singapore, one individual traveling from the Hotel Metropole was credited

for infecting 103 out of 203 cases. Despite these few instances, transmission was primarily due to close contact with infected individuals and remained in the hospital epicenters where import cases had harbored the virus. Rarely, secondary outbreaks in the community occurred. This required containment measures to be enacted. Local authorities began voluntary and mandatory quarantine procedures and eventually, the number of new cases began declining. Officials declared regions to be free of SARS when no new cases were identified after 20 days, twice the time of an incubation period. By early July 2003, there were no new cases of SARS reported. As of September 2003, WHO reported an estimated 8,098 cases and 774 deaths from the SARS-CoV.[43]

The rapid identification of a virus as the etiologic agent of SARS allowed public health officials to address the public on how to kill the virus, any treatment options, and use knowledge of other animal coronaviruses to form hypotheses regarding origin and transmission. The virus was isolated by inoculation of Vero E6 and NCI-H292 cells from an oropharyngeal specimen obtained from a fatally infected SARS patient.[44] The cells showed CPE with cell rounding in focal areas that were refractile to light. The infected cell cultures were treated for thin-section and negative-stain electron microscopy, where the particles were detected with the distinctive halo of the *Coronaviridae* (Figure 4.3). Investigators were extremely fortunate to have an isolate propagating and causing an obvious cytopathic effect in cell culture. Coronaviruses are notorious for being difficult to culture, requiring specific cell lines with supplements, organ culture, or animals. The characteristic particles identified using electron microscopy enabled researchers to swiftly focus on a specific virus group. *Coronaviridae* have three serological groups. Using infected Vero cells as a source of SARS-CoV antigen, antibodies specific for each coronavirus group were tested. Only antibodies from group I recognized SARS-CoV-infected cells. Antibodies from groups I, II, and III all failed to detect virus antigen in autopsy samples despite the presence of syncytial cells and other signs of pathology. Infected Vero cells were again used as antigen to detect the presence of SARS-CoV antibodies from patient samples. Patient sera with high titers of SARS-CoV-specific antibodies were used to develop an ELISA. SARS-CoV antibodies were later raised in animals injected with virus infected cell culture. Random samples from the population did not have detectable IgG antibodies, implying that SARS-CoV had not previously been circulating and instead was a newly emerged virus.[44]

The SARS-CoV was also identified using RT-PCR with a broadly reactive primer set designed based on the polymerase gene. The entire genome was sequenced and phylogenetically compared to other coronaviruses. SARS-CoV was placed equidistant from all other animal and human coronaviruses in all gene regions. This implied that the SARS-CoV did not arise from recombination from any known coronavirus, nor did it help identify the host. To verify

SARS-CoV as the actual agent responsible for disease, two cynomolgus macaques were inoculated with a Vero cell culture infected with serum from a fatal SARS case. SARS-CoV was detected in nasal secretions and fecal material by RT-PCR and immunofluorescence. The virus sequence was identical to the inoculated virus.[45] Histopathology showed cell damage in the lungs, and both macaques seroconverted, proving Koch's postulate for SARS-CoV as the causative agent.[46]

With SARS worldwide, the number of samples requiring diagnostics grew. Designation of SARS-CoV as a biosafety level 3 virus required specialized laboratories, so the WHO coordinated diagnostic labs from nine countries to work together to standardize diagnostic assays, supply similar reagents, ensure proper equipment availability, and analyze results in a comparable manner. Communication was vital with daily teleconferences, secure web sites, and exchange of patient samples and autopsy materials.[43]

The SARS epidemic demonstrates how a delay in reporting and an incorrect diagnosis can allow an epidemic to spread. Initially, the Chinese government misdiagnosed the infectious agent as Chlamydia and underestimated the ease of transmission and potential dispersal throughout the region. Epidemiologists from the WHO were not permitted into the Guangdong province until May 2003, but by then the virus had spread all over the globe. Investigators were distracted by reports of fatalities from a virulent strain of avian influenza, which had resurfaced in Hong Kong in late February 2003. Also, the human metapneumovirus, which also causes respiratory illness, had been identified in cell cultures inoculated with some samples from SARS patients. It was later concluded that the newly identified coronavirus was responsible for SARS, and the metapneumovirus was simply circulating concurrently in some populations.

It is important to find the natural reservoir of the SARS-CoV to prevent its reintroduction into the human population. Retrospective epidemiology of the initial outbreak in Guangdong, China revealed that early cases were local restaurant workers handling wild animals. Investigators tested animals from local marketplaces and performed a serological survey of animal handlers. The SARS-CoV was isolated from the Himalayan palm civet, and antibodies were detected in a raccoon-dog and a Chinese ferret badger as well as some animal handlers. The sequence from the palm civet contained a 29-nucleotide region that was not present in the genomes of virus isolated from the SARS epidemic.[47] It is unknown whether this region contributed to the transmission to humans. It cannot be assumed that the palm civet or other animals in the marketplace are the actual natural reservoir or just infected bystanders.

The future of SARS is uncertain. We still do not understand the natural history of the SARS-CoV. We cannot predict whether and when the virus will reemerge. The virus may cause seasonal outbreaks or establish endemic disease

in rural health care settings. Researchers are assured the virus will appear again and are working to find therapies and a vaccine.

MONKEYPOX

The following case illustrates how trace-back investigations can determine how an African virus infected 81 people across the Midwest in the U.S. On May 20, 2003, in Wisconsin, a young girl was taken to the hospital with a rash and raised fluid-filled bumps on her hand that formed after a bite from her pet (Figure 4.6). The doctor who treated her suspected that the girl may have been infected with either the plague or tularemia. Shortly thereafter, both the mother and father became infected with the unknown agent. On May 30, 2003, upon viewing EM micrographs of the mother's biopsy as well as cell culture, state doctors believed the infectious agent was a member of the orthopoxvirus genus. The EM micrographs (Figure 4.6) were useful in identifying the genus of virus, but were also of limited use for species identification because all members of the orthopoxvirus genus have similar morphology. The agent could have been any member such as cowpox, vaccinia, monkeypox, or even smallpox. On June 6, 2003, local clinicians were able to culture the orthopox virus obtained from the pet prairie dog's lymph nodes. Samples were quickly sent to the CDC in Atlanta, Georgia for identification.

While samples were awaiting identification, several more cases were reported. These were predominantly from Midwestern states, including

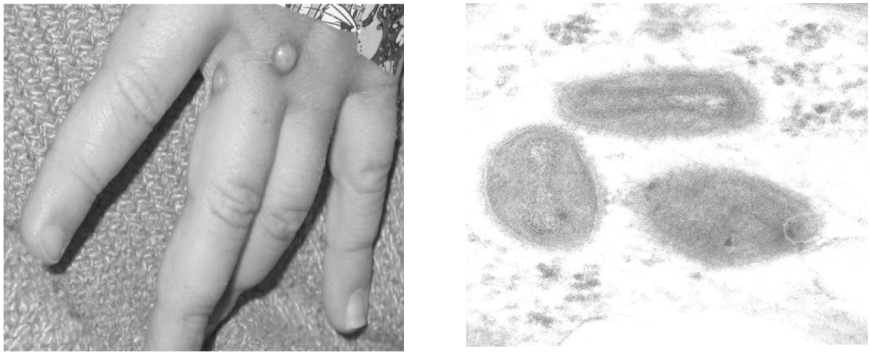


FIGURE 4.6 (A) Monkeypox skin lesion on a young girl's right hand. (B) Electron micrograph of orthopox virus cultured from her mother. Courtesy of the Marshfield Clinic, Marshfield, Wisconsin.

Wisconsin, Illinois, Indiana, Missouri, Ohio, and Kansas. Patients showed signs of rash, fever, cough, and headache, among other symptoms. As the patients were interviewed and information gathered, a common feature became apparent. Most patients had contact with exotic pet prairie dogs. On June 8, 2003, the CDC positively identified the virus as monkeypox.

By July 2, 2003, a final total of 81 monkeypox cases were reported, spanning several states. The CDC confirmed cases using a variety of laboratory techniques including propagation in tissue culture, PCR, immunohistochemical testing, and/or EM.[48] Monkeypox virus got its name when it was first discovered in 1958, after several laboratory monkeys became infected with the virus. There have been few reports of this virus, which is native to rainforest regions of Central and Western Africa. However, in 1996–1997 an outbreak occurred in the Democratic Republic of Congo (DRC), where monkeypox was positively identified in 72 cases spanning thirteen villages.[49] The virus has been found to infect native squirrels, shrews, monkeys, and Gambian rats among other animals representing 14 species.[48] The cause of the epidemic is a result of handling or consuming contaminated bush meat and foodstuffs. The virus was spread most likely through person-to-person contact, due to close-quarters habitation. Of the cases documented, six deaths occurred, accounting for a 3% case-fatality rate.[49]

In the U.S. outbreak several states linked the purchases of the prairie dogs to particular pet stores. The trace-back investigation revealed all of these pet stores purchased prairie dogs from a distributor in Illinois. Through PCR-based genetic testing, the CDC uncovered prairie dogs at the Illinois dealer tested positive for monkeypox. How did these prairie dogs become infected with a virus native to Africa? This distributor, on April 21, 2003, received a shipment of African Gambian rats and dormice. These were then tested by the CDC, and found to be infected with the virus. Investigators then were able to track the rodent shipment to a Texas distributor, which was the initial point of entry for several rodent species imported from Ghana on April 9, 2003. Several of these imported animals were infected with monkeypox. Finally a source for the monkeypox entry into the U.S. was located. A schematic illustration of the trace-back investigation is shown in Figure 4.7.

Very quickly case investigators reconstructed the events of the outbreak. Their trace-back investigation concluded that the method of virus spread to human populations was through contact with monkeypox-infected prairie dogs (animal's body fluids or skin rash); this was in contrast to the outbreak in the DRC where monkeypox spread predominantly via person-to-person contact. The prairie dogs were infected by the monkeypox-carrying rodents at the Illinois animal distributor. Because pet stores from many different states purchased from one distributor, the geographic region that became susceptible drastically expanded.

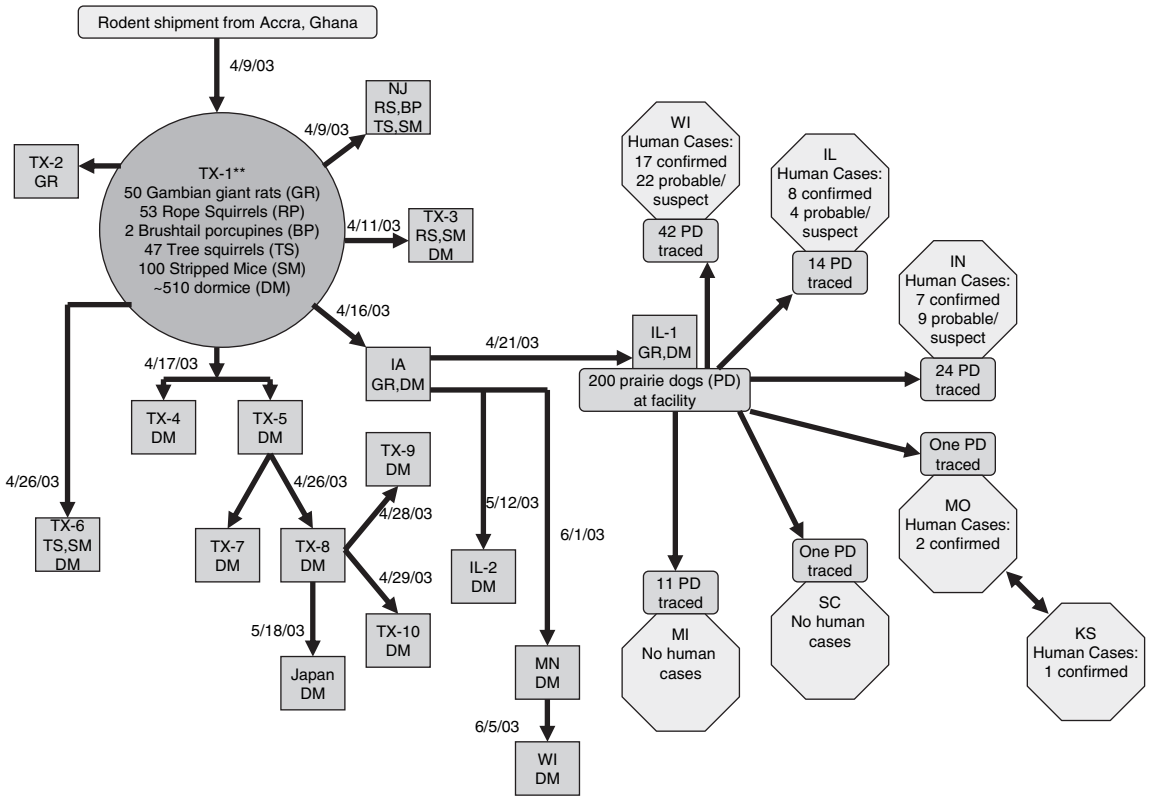


FIGURE 4.7 CDC's trace-back investigation of the monkeypox outbreak. Adapted from the CDC.

VIRAL FORENSICS

THE SCHMIDT CASE

The value of phylogenetic analyses in forensics is well illustrated by the recent and curious criminal case of the State of Louisiana versus Richard J. Schmidt.[50] The uncontested facts in this case are that a gastroenterologist from Lafayette, Louisiana broke into the home of his former mistress and office nurse late at night on August 4, 1994, and that he argued with her and gave her an intramuscular injection. He claimed it was a vitamin B shot. She claimed it was HIV. She had begun feeling ill several months after the injection, and a blood test in January 1995 revealed that she had become infected with HIV. She was a periodic blood donor, and based on tests of those previous blood donations, had a clear record without prior infection. She did not engage in any behaviors placing her at high risk for infection, and her sexual contacts over the previous nine years all tested negative for HIV. He was a community leader and Vietnam War veteran. Schmidt admitted to having a long-term affair with her, but maintained the infection was not his doing.

She went to the District Attorney's office to file charges on learning that she was HIV-positive. Moving quickly, the DA's detectives obtained a search warrant and proceeded to the accused physician's office, where they seized his record books for blood samples drawn from patients and a vial of blood sitting in the refrigerator in a back room in his office. This was unusual; patients' blood samples were sent to the lab soon after being drawn, and none were routinely stored there. The physician claimed that this sample, drawn from one of his HIV-positive patients, was for his own use and research. Was the physician telling the truth? Might this blood sample link the physician to his nurse assistant and former mistress' infection? The next logical step in the investigation would clearly be phylogenetic analyses. Phylogenetic analyses of viral DNAs showing little or no relationship between HIV lineages from the nurse and the alleged source (the blood vial seized from the physicians office) could help demonstrate the physician's innocence, whereas a close, sister relationship among those lineages, in the context of an epidemiological sampling of HIV, would be consistent with the physician's alleged role in transmission.

Are phylogenetic analyses better than other, more routine methods of forensic analyses? In some instances, yes. Analyses based on similarity alone, such as comparison of genetic fingerprinting data in which restriction fragment patterns for hypervariable DNA sequences are compared, are subject to greater bias from similarity due to chance, known as homoplasy, rather than similarity due to common ancestry, known as homology. By contrast, phylogenetic analyses attempt, explicitly, to show the pattern of common descent among

samples analyzed, rather than simple similarity. Phylogenetic analyses have the potential to show homoplasious similarity for what it is, convergence, and not to be misled, when the desired information is evolutionary relationships.

Portions of the HIV env gp120 gene and the RT genes were specifically chosen to be sequenced in an attempt to maximize the phylogenetic information in the dataset. The env gene evolves relatively quickly, being selected upon by hosts' immune systems, and is capable of recovering relationships among recently diverged HIV lineages. The RT gene evolves more slowly, due to greater functional constraints, and can provide insight into relatively older divergences. Using both sequences thus provides a broader range of evolutionary rates than either does alone.

Initial phylogenetic analyses were conducted using maximum parsimony with the software program PAUP,[51] and final analyses were conducted using an explicit model for sequence evolution seeking to account for heterogeneity in rates of change across nucleotide sites and across virus lineages. This was done using the maximum-likelihood optimality criteria in a bayesian context with the program MrBayes.[52] All analyses of the HIV env gp120 sequences from the nurse, the alleged source, and the epidemiological sampling of HIV patients were congruent in showing the HIV sequences from the victim to form a single monophyletic clade, in showing the alleged source to form a single monophyletic clade, and in showing those two clades to be closest relatives (sister taxa) relative to the epidemiological sample. This is consistent with the accusation that the physician used the blood sample from one of his patients to infect the nurse, although this rapidly evolving sequence provides no information regarding the direction of infection.

The more slowly evolving RT sequences also indicate their close evolutionary relationship, but with an additional and valuable piece of information. Based on RT sequences, viruses from the victim arose from within the clade of viruses from the alleged source. That is, the alleged source viruses are paraphyletic (incomplete in this case) unless the victim's viruses are included and nested within that group. This analyses does provide more direct evidence about the direction of infection, with the clear implication that viruses from the alleged source were used to infect the victim. Viral lineages from the alleged source diverged prior to divergences among the victim's viruses. This difference from the tree topologies based on env sequences can be traced to the slower rate of RT sequence change, with longer associated coalescence times for gene lineages, showing an earlier set of divergence events, as expected.

The phylogenetic analyses and rationale above are mainstream methods among evolutionary biologists; however, phylogenetic thinking and explicit use of evolutionary trees to track genealogy and transmission of virus or bacterial strains between individuals is not yet common in the U.S. courts. The

case of the State of Louisiana versus Richard J. Schmidt set a precedent in this area. A pre-trial admissibility hearing was held regarding the proposed use of phylogenetic analyses in the criminal trial accusing Schmidt of attempted murder. Despite the efforts of the defense to block their admissibility, based on arguments that the viruses were evolving too rapidly to allow tracing of their shared ancestry (they were not), the judge ruled that phylogenetic analyses did meet judicial standards of admissibility, being subject to empirical testing, published in peer reviewed sources, and generally accepted within the scientific community.

Though not a panacea, phylogenetic analyses will prove useful in a range of forensic investigations. As with other molecular forensic approaches they can be particularly effective in demonstrating the innocence of accused individuals. They can be also be useful in tracing sources for any transfer of infectious materials whether viral, bacterial, or protozoan, involving accidental contamination or intentional infection in personal crimes or acts of terrorism. However, these applications are potentially limited by rates of sequence change, which must be sufficiently fast to provide a record of phylogenetic relatedness, but slow enough to preserve sufficient phylogenetic signal, prior to its being overwritten with multiple substitutions at individual sites. Application of phylogenetic analyses can also be complicated by the propensity of some viruses to recombine.

The defendant was found guilty of attempted murder and sentenced to 50 years in prison, the maximum allowable under the law. In this particular case, tried by jury, the phylogenetic evidence was consistent with the prosecution's case; however, there was other evidence that the jury may have found even more compelling. This included the physician having hidden the notebooks of his blood sampling and having a history of threats against the victim as she tried to end their affair. On March 4, 2002 the U.S. Supreme Court rejected an appeal of the verdict, thus establishing precedent for use of phylogenetic analyses in U.S. courts of law.

ENGINEERING NOVEL VIRUSES

BEFORE AND AFTER RECOMBINANT DNA

As mentioned earlier, viruses in nature, depending upon their replication strategy and host range, have many ways to evolve through mutation, recombination, reassortment, and selection. These innate properties have been used as tools in virology for decades. Examples include cold-adapted and attenuated live vaccine strains of influenza, cross-species rotavirus reassortants as vaccine candidates, and the attenuated yellow fever virus 17D vaccine strain that was

derived by serial passaging in cell culture. Over the last 30 years, the recombinant DNA revolution has changed virology forever. Besides its impact on our ability to study specific viral nucleic acids and proteins, recombinant DNA technology (and PCR) sped up our ability to determine complete viral genomic sequences and create specific viral variants. The ability of scientists to engineer novel viruses varies greatly. Small positive-strand RNA viruses of bacteria, plants, and animals provided some of the first examples where infectious virus was produced via recombinant DNA. More recently, it has also become possible to engineer the >29-kb RNA genome of coronaviruses and both monopartite and segmented negative-strand RNA viruses. Viruses in the family *Reoviridae*, with double-stranded RNA genomes, have proven more difficult. For herpesviruses, whose genomic DNA is infectious, the same technologies used for cloning large pieces of chromosomal DNA have been applied successfully to propagate overlapping fragments of entire herpesvirus genomes. For the poxviruses, whose genomic DNA is not infectious, recombinant viruses are generated within infected cells by homologous DNA recombination between a plasmid containing the engineered segment and an infecting parental virus.

Synthetic Poliovirus

Typically, recombinant DNA manipulations of viruses begin with viral nucleic acid (RNA or DNA) that is then cloned and amplified in bacteria using plasmid vectors or *in vitro* using PCR. Once the genome sequence of a virus is known, however, it becomes possible to create this sequence artificially using overlapping synthetic oligonucleotides and gene synthesis techniques that have been available for many years. An example of this, which received a great deal of attention from the media and the scientific community, was recently published for an attenuated form of poliovirus.[53] While the poliovirus case came as no surprise to virologists and molecular biologists, it did bring several issues to the forefront. If an infectious virus can be created by synthetic methods, can it ever really be eradicated? In light of potential bioterrorist or biowarfare agents, should there be a restriction on making the genome sequence available in public databases? Should research on these agents be banned, classified, or otherwise regulated? These and other questions will continue to be debated in the years to come.

Mousepox

The ability to engineer viruses raises the concern about modifying existing viruses to make them more virulent. The bulk of examples in virology run

counter to this idea. Viruses are usually highly adapted to a particular niche, and most mutations are deleterious. Propagation of pathogenic viruses in cell culture often leads to adaptation to that environment and attenuation in their animal host. Recombinants or chimeras between even closely related viruses are usually impaired relative to either parent. We do see examples, however, of viruses that are benign in one animal host but highly pathogenic in another species. This is often the case in epizootic emerging viruses. An extreme example of host species-specific pathogenesis is the myxoma poxvirus, which causes a benign cutaneous fibroma in wild rabbits of the Americas but a highly lethal disease in the European rabbit. This virus was used for biological control of feral European rabbits in Europe and Australia. In 2001, an Australian group published a paper describing the construction of recombinant mousepox virus expression interleukin-4 (IL-4) and its pathogenesis in mice.[54] This study provides a striking and sobering example of an engineered virus with enhanced pathogenicity. IL-4 is a cytokine that regulates the immune response at various levels. The recombinant mousepox virus suppressed both innate and adaptive immune responses that normally control infection and was lethal for otherwise genetically resistant mice. Moreover, even animals that had been previously vaccinated and protected from virulent mousepox were susceptible to lethal infection by the mousepox-IL-4 recombinant virus. This study raises obvious concerns about the efficacy of current vaccination against modified versions of the smallpox virus.

DETERMINING THE SOURCE OF AN ENGINEERED VIRUS

Viral replication, mutation, and evolution both help and hinder determination of the origin of a virus. In the case of naturally occurring emerging viruses, sequence comparisons of isolates linked temporally to disease, geography, and species can help pin down the virus origin. Cases involving deliberate dissemination of a naturally occurring isolate, as in the Schmidt case, require not just phylogenetic comparisons but also other forensic evidence. The ability to engineer viruses makes the situation even more complex. While there are numerous laboratory strains of different viruses for which we know the exact sequence, the sequence of an isolate does not prove that it originated from a particular source. Viruses can be transferred and propagated or, as noted above, functional viral genomes can be created synthetically to mimic (or diverge from) a publicly available sequence.

CONCLUSION

In many ways, virology and viral forensics are still in their infancy. We know very little about the vast spectrum of viruses in nature. Although we should continue to pursue studies on known human pathogens, we need to broaden our efforts in virus isolation, viral genomics, and bioinformatics. A comprehensive database of viral sequences is critical for rapid identification of emerging viral pathogens and new diagnostic platforms, such as oligonucleotide arrays.[55] In addition to viral genomics, global proteomic analysis of viral disease processes may uncover molecular signatures that can be used to implicate a particular etiologic agent or class of agents, even in the absence of viral nucleic acid or serology. The global eradication of smallpox ranks as one of humankind's greatest accomplishments. In the wake of this triumph, many viral challenges remain. HIV, hepatitis B and C, influenza, dengue, rotavirus, and many other viral diseases continue to affect tens of millions. It is distressing that in the face of these undeniable challenges we are back to discussing the deliberate use of existing or modified viruses for harm, not good.

ACKNOWLEDGMENTS

Work by D.P.M. was supported by the National Science Foundation (DBI-9974525); J.H., L.K.M, H.L.H., and C.M.R. are supported by grants from the Public Health Service and the Greenberg Medical Research Institute.

REFERENCES

1. Baltimore D. Expression of animal virus genomes. *Bacteriol Rev* 1971;35:235–241.
2. Mayo MA, Pringle CR. Virus taxonomy–1997. *J Gen Virol* 1998;79 (Pt 4):649–657.
3. MHV van Regenmortel, Fauquet CM, Bishop DHL, Carstens EB, Estes MK, Lemon SM, Maniloff J, Mayo MA, McGeoch DJ, Pringle CR, Wickner RB. Virus Taxonomy: The Classification and Nomenclature of Viruses. The Seventh Report of the International Committee on Taxonomy of Viruses. Vol. VIIth report of the ICTV. San Diego: Academic Press, 2000.
4. Domingo E, Holland JJ. RNA virus mutations and fitness for survival. *Annu Rev Microbiol* 1997;51:151–178.
5. Nathanson N, Fine P. Virology. Poliomyelitis eradication—a dangerous endgame. *Science* 2002;296(5566):269–270.
6. Kew OM, Nottay BK, Hatch MH, et al. Multiple genetic changes can occur in the oral poliovaccines upon replication in humans. *J Gen Virol* 1981;56(Pt 2):337–347.
7. Kew O, Morris-Glasgow V, Landaverde M, et al. Outbreak of poliomyelitis in Hispaniola associated with circulating type 1 vaccine-derived poliovirus. *Science* 2002;296(5566):356–359.
8. Morse SS. *Emerging Viruses*. New York: Oxford University Press, 1993.
9. Webster RG, Laver WG, Air GM, Schild GC. Molecular mechanisms of variation in influenza viruses. *Nature* 1982;296(5853):115–121.

10. Holmes EC, Twiddy SS. The origin, emergence and evolutionary genetics of dengue virus. *Infect Genet Evol* 2003;3(1):19–28.
11. Outbreak of Acute Illness-Southwestern United States, 1993. *Morb Mortal Wkly Rep* 1993;42(22):421–424.
12. Khan AS, Ksiazek TG, Peters CJ. Hantavirus pulmonary syndrome. *Lancet* 1996; 347(9003):739–741.
13. Zaki SR, Greer PW, Coffield LM, et al. Hantavirus pulmonary syndrome. Pathogenesis of an emerging infectious disease. *Am J Pathol* 1995;146(3):552–579.
14. Nichol ST, Spiropoulou CF, Morzunov S, et al. Genetic identification of a hantavirus associated with an outbreak of acute respiratory illness. *Science* 1993;262(5135):914–917.
15. Feldmann H, Sanchez A, Morzunov S, et al. Utilization of autopsy RNA for the synthesis of the nucleocapsid antigen of a newly recognized virus associated with hantavirus pulmonary syndrome. *Virus Res* 1993;30(3):351–367.
16. Mills JN, Yates TL, Ksiazek TG, et al. Long-term studies of hantavirus reservoir populations in the southwestern United States: rationale, potential, and methods. *Emerg Infect Dis* 1999;5(1):95–101.
17. Peters CJ. Hantavirus Pulmonary Syndrome in the Americas. *In* Scheld WM, Craig WA, and Hughes JM, eds. *Emerging Infections 2*. Washington, D.C.: ASM Press, 1998:17–64.
18. Morzunov SP, Rowe JE, Ksiazek TG, et al. Genetic analysis of the diversity and origin of hantaviruses in *Peromyscus leucopus* mice in North America. *J Virol* 1998;72(1):57–64.
19. Wells RM, Sosa Estani S, Yadon ZE, et al. An unusual hantavirus outbreak in southern Argentina: Person-to-person transmission? Hantavirus Pulmonary Syndrome Study Group for Patagonia. *Emerg Infect Dis* 1997;3(2):171–174.
20. Plyusnin A, Morzunov SP. Virus evolution and genetic diversity of hantaviruses and their rodent hosts. *Curr Top Microbiol Immunol* 2001;256:47–75.
21. Rodriguez LL, Owens JH, Peters CJ, Nichol ST. Genetic reassortment among viruses causing hantavirus pulmonary syndrome. *Virology* 1998;242(1):99–106.
22. Nichol ST. Genetic analysis of hantavirus and their host relationships. *In* Saluzzo JF, Dodet B, ed. *Emergence and Control of Rodent-borne Viral Diseases*. Annecy, France: Elsevier, 1998: 99–109.
23. Goh KJ, Tan CT, Chew NK, et al. Clinical features of Nipah virus encephalitis among pig farmers in Malaysia. *N Engl J Med* 2000;342(17):1229–1235.
24. Chua KB, Goh KJ, Wong KT, et al. Fatal encephalitis due to Nipah virus among pig-farmers in Malaysia. *Lancet* 1999;354(9186):1257–1259.
25. Chua KB, Bellini WJ, Rota PA, et al. Nipah virus: A recently emergent deadly paramyxovirus. *Science* 2000;288(5470):1432–1435.
26. Goldsmith CS, Whistler T, Rollin PE, et al. Elucidation of Nipah virus morphogenesis and replication using ultrastructural and molecular approaches. *Virus Res* 2003;92(1):89–98.
27. Field H, Young P, Yob JM, et al. The natural history of Hendra and Nipah viruses. *Microbes Infect* 2001;3(4):307–314.
28. Daniels P, Ksiazek T, Eaton BT. Laboratory diagnosis of Nipah and Hendra virus infections. *Microbes Infect* 2001;3(4):289–295.
29. Harcourt BH, Tamin A, Ksiazek TG, et al. Molecular characterization of Nipah virus, a newly emergent paramyxovirus. *Virology* 2000;271(2):334–349.
30. Harcourt BH, Tamin A, Halpin K, et al. Molecular characterization of the polymerase gene and genomic termini of Nipah virus. *Virology* 2001;287(1):192–201.
31. Chua KB, Koh CL, Hooi PS, et al. Isolation of Nipah virus from Malaysian Island flying-foxes. *Microbes Infect* 2002;4(2):145–151.
32. Yob JM, Field H, Rashdi AM, et al. Nipah virus infection in bats (order Chiroptera) in peninsular Malaysia. *Emerg Infect Dis* 2001;7(3):439–441.

33. Outbreak of West Nile-like viral encephalitis–New York, 1999. *MMWR Morb Mortal Wkly Rep* 1999;48(38):845–849.
34. Petersen LR, Marfin AA, Gubler DJ. West Nile virus. *JAMA* 2003;290(4):524–528.
35. http://www.cdc.govncidod/dvbid/arbor/sle_qa.htm
36. Briese T, Jia XY, Huang C, et al. Identification of a Kunjin/West Nile-like flavivirus in brains of patients with New York encephalitis. *Lancet* 1999;354(9186):1261–1262.
37. Campbell GL, Marfin AA, Lanciotti RS, Gubler DJ. West Nile virus. *Lancet Infect Dis* 2002; 2(9):519–529.
38. Pealer LN, Marfin AA, Petersen LR, et al. Transmission of West Nile virus through blood transfusion in the United States in 2002. *N Engl J Med* 2003;349(13):1236–1245.
39. Detection of West Nile virus in blood donations—United States, 2003. *MMWR Morb Mortal Wkly Rep* 2003;52(32):769–772.
40. http://nationalatlas.gov/virusmap.html
41. www.who.int/csr/sars/casedefinition/en
42. www.who.int/csr/sars/country/table2003_09_23/en
43. Severe acute respiratory syndrome (SARS): Status of the outbreak and lessons for the immediate future. Geneva: World Health Organization. Communicable Disease Surveillance and Response, 2003. pp. 1–10.
44. Ksiazek TG, Erdman D, Goldsmith CS, et al. A novel coronavirus associated with severe acute respiratory syndrome. *N Engl J Med* 2003;348(20):1953–1966.
45. Rota PA, Oberste MS, Monroe SS, et al. Characterization of a novel coronavirus associated with severe acute respiratory syndrome. *Science* 2003;300(5624):1394–1399.
46. Fouchier RA, Kuiken T, Schutten M, et al. Aetiology: Koch's postulates fulfilled for SARS virus. *Nature* 2003;423(6937):240.
47. Guan Y, Zheng BJ, He YQ, et al. Isolation and characterization of viruses related to the SARS coronavirus from animals in Southern China. *Science* 2003;302(5643):276–278.
48. Update: multistate outbreak of monkeypox—Illinois, Indiana, Kansas, Missouri, Ohio, and Wisconsin, 2003. *Morb Mortal Wkly Rep* 2003;52(27):642–646.
49. Mwanbal PT, Tshioko KF, Moudi A, et al. Human monkeypox in Kasai Oriental, Zaire (1996–1997). *Euro Surveill* 1997;2(5):33–35.
50. Metzker ML, Mindell DP, Liu XM, et al. Molecular evidence of HIV-1 transmission in a criminal case. *Proc Natl Acad Sci U S A* 2002;99(22):14292–14297.
51. Swofford D. PAUP. Sunderland, Massachusetts: Sinauer Associates, 2001.
52. Huelsenbeck JP, Ronquist F. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* 2001;17(8):754–755.
53. Cello J, Paul AV, Wimmer E. Chemical synthesis of poliovirus cDNA: Generation of infectious virus in the absence of natural template. *Science* 2002;297(5583):1016–1018.
54. Jackson RJ, Ramsay AJ, Christensen CD, et al. Expression of mouse interleukin-4 by a recombinant ectromelia virus suppresses cytolytic lymphocyte responses and overcomes genetic resistance to mousepox. *J Virol* 2001;75(3):1205–1210.
55. Wang D, Coscoy L, Zylberberg M, et al. Microarray-based detection and genotyping of viral pathogens. *Proc Natl Acad Sci* 2002;99(24):15687–15692.
56. Flint SJ, Enquist LW, Racaniello KR, Skalka AM. *Principles of Virology: Molecular Biology, Pathogenesis, and Control of Animal Viruses. 2nd ed.* Washington, D.C.: ASM Press, 2004.