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 9

Keeping Track of Viruses

Catherine L. Murray, Thomas S. Oh, and Charles M. Rice
Laboratory of Virology and Infectious Disease, Center for the Study of Hepatitis C, The Rockefeller University, New York

INTRODUCTION

Viruses are the most abundant biological entities on earth (1). These obligate parasites infect every form of life, from archaea and eubacteria to fungi, plants, and animals; even viruses can be affected by a coinfecting satellite species (2). Viruses play key roles in global ecology—they form a vast reservoir of genetic diversity, influence the composition and evolution of host populations, and affect the cycling of chemical compounds through the environment (3). While research has focused on the tiny fraction that causes disease in humans, domestic animals, and crops, sequencing surveys have suggested that the majority of viruses are completely unknown (1). The ability of viruses to jump species barriers, move between habitats, and circle the globe rapidly underscores the importance of continued vigilance for naturally emerging or deliberately engineered outbreaks. This chapter reviews methods of isolating, identifying, and tracking viruses with potential applications to microbial forensic investigations.

WHAT IS A VIRUS?

Viruses are extremely simple "life" forms without metabolic capacity, organelles, translational machinery, or autonomous replicative potential (4). Virus particles constitute a minimal set of components, primarily those required to deliver the genome to the target cell and initiate replication. Consequently, virus particles (or virions) are extremely small, most in the range of 20 to 200 nm in diameter. A notable exception is a recently discovered "giant virus," termed mimivirus, for "mimicking microbe," which has a particle diameter of 400 nm, comparable to a small bacterium (5). Virions are not only diverse in size, but also in composition, morphology, and genome

Microbial Forensics. DOI: 10.1016/B978-0-12-382006-8.00009-8

characteristics. Virus particles may be irregular in shape or possess a distinct symmetry, such as helical or icosahedral. Particles may be surrounded by a host-derived membrane(s), termed "enveloped," or a tight protein shell, termed "nonenveloped." Inside the virion, the genome is associated with nucleic acid-binding proteins; some viruses carry additional factors, such as enzymes required to initiate replication. While bacteria, fungi, parasites, plants, and animals use exclusively deoxyribonucleic acid (DNA) as their genetic material, a viral genome may be composed of either DNA or ribonucleic acid (RNA). The genome may be single stranded (ss) or double stranded (ds), circularized or linear, consist of a single nucleic acid strand, or be "segmented" on multiple molecules. Viruses do not share any characteristic sequence that is conserved across all families, as are ribosomal (r)RNAs in cellular organisms. Virus genomes also vary greatly in size. The ssRNA genomes of poliovirus and human immunodeficiency virus (HIV) are 7.4 and 9.2 kb, respectively, whereas the dsDNA genome of mimivirus is approximately 800 kbp.

VIRUS LIFE CYCLE

Virus Attachment and Entry

Viruses must enter a target cell in a way that does not do excessive damage to the host or alert immune defenses (6). This is generally accomplished by hijacking normal cellular processes, including receptor–ligand binding, endocytosis, and nuclear import. The virion attaches by binding to a protein, lipid, and/or carbohydrate displayed on the cell surface. Envelope glycoproteins, or the spikes and indentations of the nonenveloped virus shell, participate in these initial interactions. The specific cellular molecule to which a virus binds is termed its "receptor." Some viruses, such as HIV and hepatitis C virus (HCV), bind to several receptors and coreceptors, which perform distinct roles in complex multistep uptake pathways. Receptor binding initiates internalization of the virus particle, transport to the appropriate cellular compartment, and uncoating of the genome. Enveloped virus glycoproteins are triggered to mediate fusion of the viral and host membranes during uptake. Delivery of a replication-competent viral genome to a permissive intracellular site is the first step in establishing a productive infection.

Replication Strategies

The diversity of viral genomes necessitates a variety of replication strategies. Viruses are divided into seven groups based on genetic material, polarity, and messenger (m)RNA synthesis (7). Polarity refers to the protein-coding capacity of a nucleic acid strand, where positive (+) strand nucleic acid has a 5′→3′ polarity, identical to mRNA, and negative (−) strand nucleic acid has a 3′→5′ polarity, complementary to mRNA (Figure 9.1).

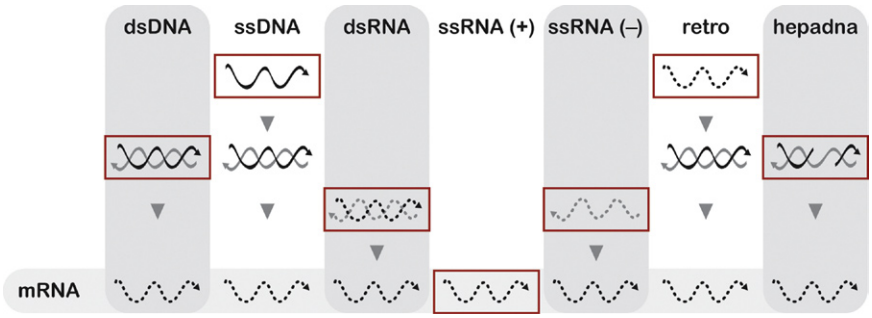

FIGURE 9.1

Replication strategies of viruses. All virus genomes must be used to produce mRNA in order for the viral proteins to be expressed inside the cell. The schematic represents the seven classes of viruses, according to the Baltimore classification, and the intermediates through which mRNA is produced. The nucleic acid character of the viral genome is indicated by a box. Black, positive-strand nucleic acid; gray, negative-strand nucleic acid. Arrows on nucleic strands indicate their directionality, pointing from 5′ to 3′ ends. DNA is indicated as solid lines, RNA as dashed lines. The partially double-stranded nature of the hepadnavirus DNA genome is indicated by a gap.

Double-Stranded DNA Viruses

Viruses with dsDNA genomes may replicate in the nucleus or cytoplasm, with transcription of the genome into mRNA by host or viral RNA polymerases, respectively. Variola major, the causative agent of smallpox, is an example of an enveloped virus with a linear dsDNA genome.

Single-Stranded DNA Viruses

These viruses have an ssDNA genome of (+) polarity. The genome is converted to dsDNA in the nucleus and is subsequently transcribed and translated by host machinery to produce viral proteins. Parvoviruses, which cause rash in children and often-fatal infection in dogs, are members of this group.

Double-Stranded RNA Viruses

Viruses in this class contain segmented dsRNA genomes. mRNA is synthesized by a virally encoded RNA-dependent RNA polymerase (RdRp). Because most eukaryotic cells do no encode this type of enzyme, the virus must import its own RdRp within the incoming virion. Rotavirus, a common etiologic agent of severe infectious diarrhea in children, has a dsRNA genome.

Positive-Strand RNA Viruses

The ssRNA genomes of these viruses are translated directly by host ribosomes in the cytoplasm. The virally encoded RdRp then replicates the genome through a complementary (−) strand intermediate. Examples of viruses in this class include poliovirus, West Nile virus, and HCV.

Negative-Strand RNA Viruses

The ssRNA genomes of these viruses may be either segmented or continuous. Some are ambisense, with portions of the genome acting as (+) strands and others having (−) polarity. All members of this class, which includes influenza and Ebola viruses, import an RdRp that transcribes the viral genome into mRNA.

Retroviruses

Retroviruses package two identical molecules of (+) polarity ssRNA. A virally encoded enzyme termed "reverse transcriptase (RT)" generates dsDNA from the RNA templates. The name "retrovirus" reflects the fact that this replicative cycle is retrograde (RNA→DNA→mRNA→protein) relative to the central dogma of modern biology (DNA→mRNA→protein). HIV, the virus that causes acquired immune deficiency syndrome (AIDS), is a retrovirus.

Hepadnaviruses

Members of this group have a partially double-stranded DNA genome and replicate via an RNA intermediate, similar to retroviruses. mRNA is packaged into immature particles before conversion to DNA by the virally encoded RT. An example of this group is hepatitis B virus, an important etiologic agent of chronic liver disease.

Assembly and Release

Transport of the amplified genome to a new permissive host requires the production of infectious viral particles. This is a complicated process that is well understood for only a few viruses. Similar to replication, virion assembly takes place at defined intracellular locations, such as in the nucleus, at membranous cytoplasmic organelles, or at the cell surface. Virions can be released from the cell by noncytopathic budding or through host cell lysis.

HOW DO YOU IDENTIFY A VIRUS?

Sudden emergence of an infectious disease demands methods to rapidly and accurately identify the infectious agent, diagnose patients, and explain routes of transmission. A "staged" approach is often employed, in which epidemiology, pathology, and serological assays suggest candidate pathogens, which can be confirmed by nucleic acid-based methods. In the absence of suspects, microarray, next-generation sequencing, or subtractive cloning can be informative (8,9).

Culture and Cytopathic Effect

Patient samples may be directly infectious to immortalized cell lines, allowing the pathogen to be isolated, quantified, and amplified. Some viruses do not grow well in cultured cells, but may be coaxed to replicate in primary cells, embryonated eggs, or experimental animals. If culture is successful, the

phenotype of the infected cells can reveal valuable clues about the identity of the virus (4). Infection may kill the cells, creating a characteristic cytopathic effect (CPE). Cells may appear to be rounded or growing in grape-like clusters, indicating adenovirus or herpes simplex virus. Cells may fuse into multinucleated "syncytia," suggesting influenza, mumps, or measles. Poxviruses create foci of fused cells, whereas positive-strand RNA viruses induce proliferation of membranes in the cytoplasm. Many other viruses do not cause noticeable cell damage. Observance of CPE was a critical factor in identifying the causative agent of acute fever with encephalitis among pig farmers on the Malay Peninsula in 1998–1999. Multinucleate syncytia were seen in Vero cells inoculated with cerebral spinal fluid obtained from fatal infections, implicating a paramyxovirus (10). The new pathogen was named "Nipah virus," and the outbreak was stopped after culling over one million pigs.

Electron Microscopy (EM)

Infected cultures or amplified virus can serve as material for visualization by EM. Electron microscopes use a beam of electrons, rather than visible light, to form an image at extremely high magnification (up to 1,000,000×). Staining with an electron-rich "negative stain" and thin sectioning of the specimen can increase contrast and enhance visualization of internal features. Virions can appear ribbon like (rabies), rod shaped (measles), spherical (poliovirus), or filamentous (Ebola) (Figure 9.2). Some viruses show irregular shapes or multiple morphologies, referred to as "pleomorphic." Visualization of particles by EM was a defining step in identification of a novel virus responsible for severe acute respiratory syndrome (SARS), a disease that spread rapidly around the globe in 2003. The virus was isolated by inoculation of cell lines with an oropharyngeal specimen obtained from a fatally infected patient. Cultures were subjected to thin section and negative stain EM, and particles showing the distinctive halo of *Coronaviridae* were detected (Figure 9.2). The characteristic particles enabled researchers to focus swiftly on a specific virus group, and numerous other tests validated and characterized the novel SARS-coronavirus (SARS-CoV) (11).

Serological Assays

The specificity and high affinity of antibody–antigen recognition is widely used in virus diagnostics. Antibodies may be isolated from sera of infected or recovering individuals or be generated experimentally in animals immunized with viral antigens. Serological assays can discover similarities between a novel pathogen and a known virus through antibody cross-reactivity.

Enzyme-Linked Immunosorbent Assay (ELISA)

The ELISA is a rapid and versatile method of detecting antigens or antibodies. Viral proteins or virus-specific antibodies are adsorbed to the surface of a microtiter plate, allowing specific capture of the cognate antibody or antigen

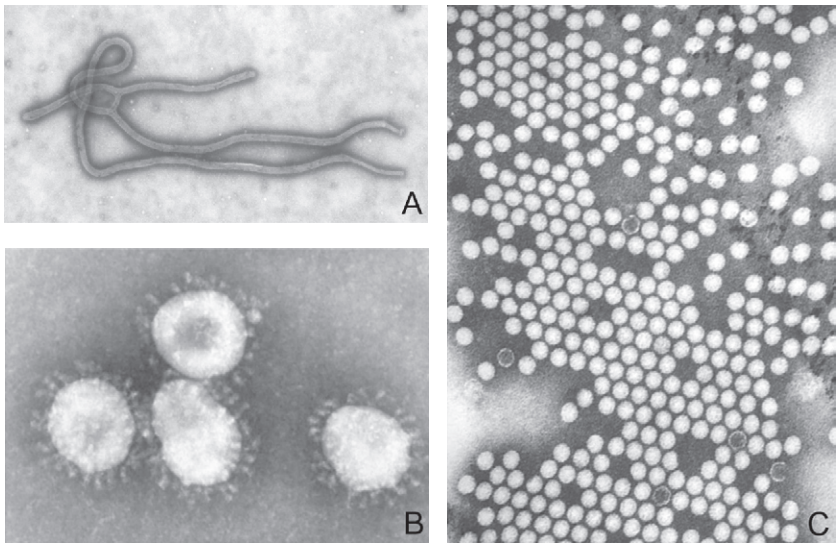

FIGURE 9.2

Electron micrographs of virus particles. (A) Ebola virus particles showing filamentous morphology. Courtesy of the CDC/C.Goldsmith. (B) SARS coronavirus showing characteristic "corona-like" morphology. Courtesy of the CDC/Fred Murphy. (C) Polio virus particles showing spherical morphology; courtesy of the CDC/Fred Murphy and Sylvia Whitfield.

from patient serum. Complexes are detected with labeled secondary antibodies, followed by a colorimetric readout. ELISA can distinguish between different classes of antibodies, such as those indicative of recent infection (IgM) or previous exposure and vaccination (IgG). This assay was used to investigate cases of encephalitis and/or profound muscle weakness in Queens, New York, in 1999. IgM-capture ELISA was used to survey antibodies against common encephalitic viruses. Results implicated St. Louis encephalitis (SLE) virus, a mosquito-borne flavivirus. Sequencing later revealed the agent was not SLE but the related West Nile virus—a pathogen never before detected in the Western hemisphere. The ability of ELISA to detect cross-reactive flavivirus antibodies meant that the appropriate vector control measures could be implemented quickly (12).

Neutralization and Hemagglutination Inhibition Assays

Antibodies produced during infection often have the ability to interfere with the native properties of virus particles. Mixing dilutions of antibodies with a virus sample, followed by measurement of the decrease in virion activity, can be used to identify viruses and to classify them into serotypes. Neutralization assays measure the ability of antibodies to block viral entry. Hemagglutinin inhibition (HI) assays detect antibodies that can block the ability of some viruses to aggregate red blood cells. In May 2009, neutralization and HI assays

were used to demonstrate that previous vaccination against seasonal influenza offered little protection against the novel H1N1 pandemic strain, indicating that the new swine-origin virus was substantially different from those that had circulated in recent years (13).

Immunostaining

Antibodies can be used to detect viral proteins in patient tissues or infected cultures. Specific binding can be visualized by secondary staining using an antibody conjugated to a fluorescent dye (immunofluorescence) or enzyme (immunohistochemistry). Immunohistochemistry was used to investigate a disease cluster in the region bordering Arizona, Colorado, New Mexico, and Utah in spring 1993. Symptoms included fever, headache, and cough that progressed rapidly to respiratory distress and death (14). Patient sera were found to contain antibodies targeting members of the genus *Hantavirus*. Immunohistochemistry of autopsy tissues indicated the presence of hantavirus antigens in endothelial cells of the lung and other involved organs (15). Nucleic acid sequencing confirmed the diagnosis, and the new pathogen was named "Sin Nombre virus."

Polymerase Chain Reaction (PCR)

Development of the PCR in 1987 ushered in a new era of nucleic acid-based pathogen detection systems (16). Amplification of viral genomes allows rapid, specific, and sensitive detection and analysis without the need for *in vitro* culture or quality antibodies. PCR allows a dsDNA target to be amplified exponentially using a pair of oligonucleotide primers designed to flank the region of interest. The PCR product, or amplicon, can be detected by a variety of methods, such as nucleic acid staining. PCR is a sensitive way to confirm the presence of a suspected virus in patients or environmental samples.

Real-Time PCR

Because methods for detecting the final PCR product can be laborious and time-consuming, an alternative strategy is to monitor amplicon synthesis in real time (17). This technique provides a wealth of information, including accurate quantification of the starting template, and is termed "real-time" or "quantitative" PCR. Real-time PCR depends on the emergence of a signal as the amplification reaction proceeds. The simplest form uses reporter molecules that fluoresce when bound nonspecifically to dsDNA. Alternatively, sequence-specific detection can be achieved using an ssDNA "probe" designed to bind within the amplified region. The probe is labeled with a fluorophore in close proximity to a quencher, which dissipates the fluorescence energy until annealing occurs. Real-time PCR is a rapid and effective method for assessing the presence of candidate viruses, distinguishing between genotypes, and measuring viral load.

Multiplex PCR

Considerable savings in time, effort, and sample volume can potentially be achieved by combining multiple PCR reactions in a single tube, termed "multiplex PCR" (17–19). Development of multiplex assays, however, can require significant optimization, and primers must be designed carefully to work well without interference (18). Furthermore, the number of targets that can be distinguished is limited—for example only a few fluorophores are available (17,20). To increase the capacity of multiplex PCR, new methods of amplicon differentiation have been established. One is MassTag PCR, which uses primers labeled with a tag of a known, unique molecular weight. After PCR, the identity of the incorporated tags can be determined by mass spectrometry. This method has been used to multiplex up to 22 respiratory pathogens in a single reaction (21) and for simultaneous detection of viral hemorrhagic fever agents (22). In an alternative method, the precise weight of the amplicon can be measured directly, allowing multiple microbes to be detected in a single complex sample based on amplification of conserved sequences (23–25). These systems, termed "TIGER" or "Ibis T5000" (Ibis Biosciences), identify virus(es) using primers to amplify broadly conserved regions from large groupings of species, followed by electrospray ionization mass spectrometry and analysis of the nucleic acid base composition (i.e., the number of adenosines, cytidines, guanosines, and thymidines in the amplicon) (23). This technology, which is used currently at the Centers for Disease Control (CDC) and the National Bioforensic Analysis Center, has been used to uncover a novel encephalitis virus (26) and detected the second case of novel H1N1 (swine flu) in the United States (26a).

Microarray and Virus Chips

Although PCR is well suited for sensitive detection of a small number of candidate viruses, the technique is inherently biased by primer and/or probe design. Recently, the application of microarray technology has provided a more impartial approach to pathogen discovery. Microarrays, or "chips," consist of short oligonucleotide probes immobilized as spots on a solid support (20,27). Isolated DNA or RNA is labeled with a fluorescent dye and hybridized to the chip; the bound spot(s) indicates the presence of sequences in the target sample. Microarrays have been used extensively since the early 1990s to investigate cell biology (27), but have only recently been adapted for the detection of infectious agents. Initial success was achieved with arrays targeting a few dozen pathogens (28,29). The subsequently developed "virus chip" included probes to the most highly conserved regions of 140 virus genomes (30) and was later expanded to over 1000 different species (31). Species and serotypes not represented explicitly on the chip can also be detected, as they form unique signatures based on hybridization to conserved sequences. The "GreeneChip" is a similar platform that includes probes for bacteria,

fungi, and parasites, as well as viruses (19). The GreeneChip was used in a postmortem diagnosis of a health care worker who had succumbed to fever and liver failure during a Marburg virus outbreak in Angola. Multiplex PCR had failed to detect Marburg or other hemorrhagic fever viruses. Hybridization to the GreeneChip revealed the presence of *Plasmodium*, a parasite that causes malaria (19). Although chip technologies are limited by the need for updates as emerging, mutating, or engineered viruses occur, they are nonetheless an important tool for epidemiologic or microbial forensic investigation.

Next-Generation Sequencing

Unbiased sequencing of the entire microbial population in an environmental sample or diseased tissue has become a real possibility with the advent of next-generation sequencing technologies—termed "high-throughput (HTS)," "deep," "massively parallel," or "Next Gen" sequencing. 454 (Roche), SOLiD (Applied Biosystems), and Illumina (Solexa) represent several of the most widely used platforms (32). HTS uses sheared DNA as a template for millions of parallel amplification reactions. 454 amplification occurs on beads, which are then arrayed in individual wells of a picotiter plate for parallel "pyrosequencing"—using the pyrophosphate released by each nucleotide incorporation to trigger a reporter signal. SOLiD sequencing also begins with amplification on beads, followed by "sequencing by ligation" in which short labeled probes, only two bases in length, bind to and reveal the target sequence over multiple rounds of annealing. Illumina sequencing begins with amplification of DNA clusters on a glass support. Fluorescently labeled nucleotides are then incorporated one at a time, through a series of blocking and unblocking steps, and images of each cycle record the sequence (32). HTS reads are short, but typically sufficient to query a database and discover pathogens with even low similarity to known sequences (9). With new HTS technologies developing at a rate so rapid as to inspire the colloquial term "Now Gen," a major challenge is the need for bioinformatic analysis to keep pace.

High-throughput sequencing was used to unravel the mystery of colony collapse disorder (CCD), a phenomenon that began devastating the honeybee industry between 2004 and 2006. CCD is characterized by the very rapid disappearance of the entire adult bee population of a hive. The observation that irradiated, but not untreated, hives could be repopulated suggested an infectious etiology. In an attempt to identify the agent, total RNA extracted from diseased or healthy colonies was analyzed by pyrosequencing. A number of bacteria, fungi, parasites, and viruses were revealed as candidates, and real-time PCR assays were developed to assess the distribution of each agent. One pathogen correlated most strongly with the occurrence of CCD: Israeli acute paralysis virus, a positive-strand RNA virus that had not been found previously in the United States (33).

Subtractive Cloning

A novel virus may display an unusual or chronic pathology that complicates epidemiology, occurs in conjunction with other microbes, or has a genome that is completely unknown. Subtractive cloning is a classical technique that comprehensively surveys differences between samples and can still be useful in revealing otherwise undetectable pathogens (9). Typically, a cDNA probe derived from infected material is used to identify sequences specific to diseased, but not normal, tissue by sequential rounds of hybridization and amplification. Borna disease virus, a pathogen implicated in a range of behavioral and neuropsychiatric pathologies in animals, and possibly in humans, was discovered by this method (34,35). Disease-specific cDNA clones can also be expressed as proteins and selected by binding to patient, but not healthy, sera. This technique was used to identify HCV, a chronic liver pathogen for which previous detection of antigens, antibodies, nucleic acid, and virus particles had been unsuccessful (36).

VIRAL DIVERSITY AND PHYLOGENETICS

During the investigation of a disease outbreak or suspected biocrime, it may be critical to determine not only the species of virus involved, but how the infection is moving through a population—here the often dramatic diversity generated during viral replication can be highly informative.

Viral Evolution

In most organisms, evolution takes place over very long timescales—much longer than could be observed in a laboratory experiment or in a criminal investigation. Viruses, however, evolve rapidly enough to make the study of genetic change a very relevant and powerful tool for the forensic scientist. What accounts for the remarkable speed of viral evolution? Two important features set viruses apart from other organisms with regard to rates of change: high mutation frequencies and replicative potential. During replication of the genetic sequence of any organism, copying errors are inevitably made, leaving the new sequence with differences from the original. However, whereas such misincorporations occur perhaps once every billion bases in most living cells, in some viruses errors are made as often as once every thousand bases copied. This results from the use of enzymes without proofreading activity (RdRp or RT) rather than higher fidelity DNA polymerases. Not only do mutations occur more frequently, many copies are made and very quickly. While a single cell cycle results in two progeny, a virus might be copied hundreds or even thousands of times in a single life cycle. Because of their inherent simplicity and small size, viruses can be assembled quickly and very large populations can be supported. The number of viruses in a single infected person may be in the billions. The large numbers of genomes and

high error rates result in a diverse and rapidly changing population. In addition to mutation, reassortment and recombination are two other important mechanisms of virus diversity. Reassortment occurs in species with segmented genomes, when two related viruses infect the same cell. As progeny are assembled, each genome segment that is packaged may be derived from either of the two original viruses, producing a virus with a combination of genes. Similarly, genetic material can be recombined from two viruses or even between a virus and its host cell, as events in the process of replicating DNA or RNA result in a new strand that is partially copied from one source and partly from another. In each case, it is the tremendous scale of viral replication that allows these seemingly rare events to have a significant impact on the process of viral evolution.

Phylogenetic Analysis

Generally, the process of genetic variation through mutation, reassortment, and recombination cannot be observed directly; however, resulting viral sequences can. By examining the sequences of many samples, it is often possible to reconstruct a family tree (or phylogeny) of a set of sequences and to infer what series of events occurred, and in what order, to produce that set. This sort of reconstruction is known as phylogenetic analysis. Figure 9.3 shows three nine-base sequences, which differ at two positions. How might these sequences be related? One possibility is that virus A acquired a mutation, giving rise to C, which in turn mutated into B. Alternatively, C may be a common ancestor to both A and B, each differing from C by one mutation. Both scenarios seem quite plausible. A third possibility is that sequence A acquired two mutations

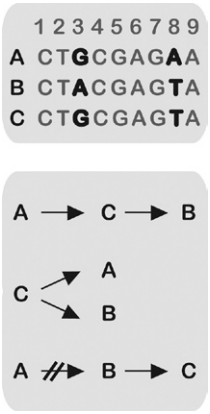

FIGURE 9.3

Phylogenetic analysis. How are these sequences related? (Top) Three nine-base sequences each differing at two positions (bold). (Bottom) Three scenarios that may relate the sequences. The probability that each scenario is correct can be calculated and used to construct a phylogenetic history of the sequence set.

(which may or may not have arisen simultaneously), producing B, and then B mutated once more to become C. This third scenario is certainly possible, but is less likely than the first two because it involves either a double mutation or sequence C arising twice—once in between A and B and again after B. Given a model of how sequences evolve, it is possible to calculate statistically how probable each scenario is relative to the others and then to determine the most likely sequence of events. While the example may seem trivial, as the length and number of samples increase, this type of analysis grows in both complexity and power. The Schmidt case, described below, was decided in large part on the basis of a phylogenetic analysis involving over a hundred viral sequences of nearly 2000 bases in length. Phylogenetic trees are used to identify the origins of new outbreaks, to determine the transmission path from one person to another, and to shed light on the evolutionary history of an unknown virus.

SOLVING A BIOCRIME

The State of Louisiana Versus Richard Schmidt

On August 4, 1994, a physician from Lafayette, Louisiana, gave an intramuscular injection to a former mistress who had recently broken off their affair. He told her he was administering a vitamin B shot, but when she became ill, suspicions mounted (4). A few months later, the victim tested positive for HIV, and Richard Schmidt was accused of using blood from a patient under his care to deliberately infect his former girlfriend—but could it be proven? Multiple viral sequences from the victim, the patient, and infected individuals in the community unrelated to the case were obtained. At first glance, it might be supposed a transmission event could be established by determining whether the patient and the victim carried identical viral sequences; the rapid rate of HIV evolution, however, makes this expectation overly simplistic. Likewise, a measurement of sequence similarity might be enough to establish a relationship between the viruses, but the direction of spread would be unknown. Phylogenetic evidence was needed to unravel the allegation of deliberate virus transmission from patient to victim.

Phylogenetic analysis was conducted on two regions of the HIV genome with different rates of evolution—an important consideration for extracting useful data. If evolution occurs too slowly, all the sequences would be similar, including those from the victim, the patient, and the unrelated controls. If the rate of change is too high, differences between the sequences may be so large that it becomes impossible to determine their relationship. Two portions of the HIV genome were analyzed to help achieve the appropriate range: the envelope (gp120) and RT coding regions. gp120, a structural component of the virion, is relatively plastic and evolves rapidly to escape host immune responses; RT performs conserved enzymatic functions and is less amenable to sequence

change. Analysis of gp120 revealed that all the patient sequences formed one cluster, while all victim sequences formed another cluster; these clusters were related by a common ancestor, which was not shared by any of the other HIV sequences analyzed (Figure 9.4A). This ancestor could be a sequence that existed in the patient before the transmission event and had since disappeared or be a virus from a third individual that infected both the patient and the

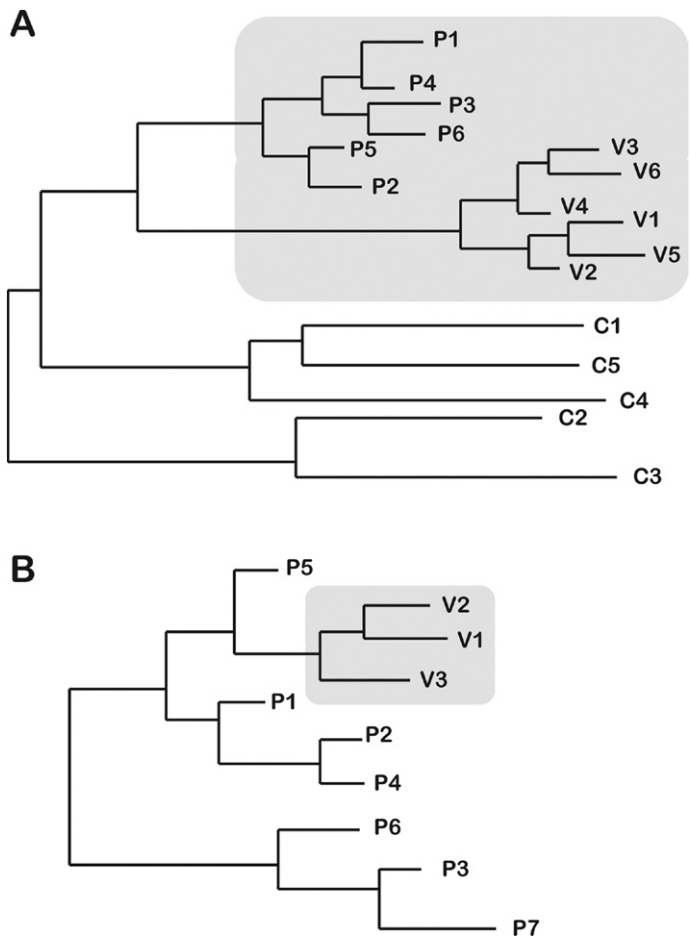

FIGURE 9.4

Schematic examples of phylogenetic trees. Phylogenetic analysis was used to link the source of HIV infection to the victim in the State of Louisiana versus Richard Schmidt case (38). Schematics illustrate the types of sequence clustering that were observed. (A) gp120 sequences from the victim and the patient (grey) shared a common ancestor not shared by unrelated controls. (B) RT sequences from the victim (grey) clustered entirely within the patient sequences. Individual sequences representing those from the patient (P), victim (V), or unrelated controls (C) are indicated with numbers.

victim. The gp120 phylogeny, therefore, shows a relationship but cannot distinguish the direction of transfer. In contrast, the RT coding sequences of the patient and victim clustered together, reflecting the slower evolution of this region (Figure 9.4B). Furthermore, all the RT sequences from the victim shared a single common ancestor contained within the patient sequences. This was a strong indication that transmission took place from the patient to the victim.

Importantly, phylogenetics analysis was used as only one piece of the puzzle, with traditional detective investigation helping to build the case (37). Although the victim worked as a nurse, raising the possibility of occupational exposure, she had a history of unexcluded blood donation and her past sexual partners were HIV negative. Furthermore, a vial of blood from the HIV-infected patient was found in Schmidt's office, a highly unusual occurrence. The State of Louisiana vs Richard J. Schmidt set a legal precedent for the admissibility of phylogenetic evidence in a criminal proceeding, and the defendant is now serving a 50-year sentence for attempted murder (38).

The Case of Kristen D. Parker

Spreading an infectious agent may not be the deliberate intention, but a by-product of criminal activity. Between October 2008 and June 2009, a surgical technician in Denver, Colorado, began stealing pain medication from patients under her care. Kristen D. Parker removed syringes of fentanyl from the operating room and replaced them with her own—often previously used—syringes, filled with saline solution. In the process, the technician exposed hundreds of patients to HCV, a virus she had apparently contracted through her history of injection drug use (39).

When these allegations came to light, the Colorado Department of Public Health and Environment quickly instigated systematic testing of thousands of potential victims for exposure to the virus. To maximize the possibility of effective treatment, it was important to determine rapidly which patients were positive for HCV; in building a case against Ms. Parker, it would also be critical to link the infections epidemiologically and genetically to the suspect. Out of over 5000 patients at two Denver medical centers tested for the virus, almost 70 were positive (40), but how had they been infected? The extensive genetic diversity of HCV provided a quick method to rule out unrelated cases. HCV is classified into six major genotypes, which are further divided into a large number of subtypes showing significant divergence in genome sequences. Patients presenting with viral genotypes other than "1b" were therefore unlikely to have been infected by Ms. Parker. Additional sequencing was conducted on genotype 1b viruses isolated from the suspect and the remaining patients, and the relatedness of the genomes was determined. Overall, more than 20 patients showed strong genetic and/or epidemiological evidence of a transmission link to Ms. Parker. The suspect pled guilty to charges of tampering

with a consumer product and theft of a controlled substance and is now facing at least 20 years behind bars (39).

CHALLENGES THAT REMAIN

Identification and analysis of a toxic agent are critically dependent on sufficient and appropriately collected sample material (41). This may be especially difficult in the case of viruses. Infectious virions can be labile, and harsh treatments such as extremes of pH and temperature should be avoided. Genomic material composed of RNA, unique to viruses, may also be challenging to acquire—RNA is more sensitive to degradation than DNA and must typically be prepared from fresh tissue treated to inhibit ubiquitous nucleases. Finally, the infectious material may have been cleared from the body by the time symptoms become apparent. It is therefore important to collect samples using a variety of different preservation methods, from multiple locations throughout the body, and as early as possible in the infection. Once material has been secured, unbiased identification techniques may be attractive tools for pathogen discovery; however, these methods can implicate hundreds of microorganisms, many of which may have no etiologic relationship to the disease (33). Extensive work may be required to identify a strong candidate for pathogenesis, and formalizing a causal relationship between microbe and disease may not be trivial (8). This diversity, however, may have tremendous value for forensic analysis —it is easy to see that the total microbial composition or phylogeny of multiple species within a sample might facilitate the identification of source material involved in an intentional attack. If the suspected virus is a novel species, appropriate reagents and standards may not be available, and rapid detection or diagnosis may necessitate assays that have not been validated completely (42). Finally, while classical methods, including cell culture, can be simple and highly informative, powerful new techniques require a significant time commitment, as well as specialized equipment and expertise. Because containment of a biological threat, as well as the initiation of a microbial forensic investigation, often requires a rapid point-of-care response, the challenge remains to reduce the time and ease of detection so that an accurate diagnosis can be made by the clinician while maintaining the integrity of the evidence (29,41).

ACKNOWLEDGMENTS

Work in the laboratory of Charles M. Rice is supported by Public Health Service grants from the National Institutes of Health, the Northeast Biodefense Center (U54 AI057158 subcontract 2925-01) and the Greenberg Medical Research Institute. The authors acknowledge Jack Hietpas, Laura K. McMullan, Holly L. Hanson, and David P. Mindell for material derived from the first edition of this chapter.

REFERENCES

[1] R.A. Edwards, F. Rohwer, Viral metagenomics, Nat. Rev. Microbiol. 3 (2005) 504–510.

[2] B. La Scola, C. Desnues, I. Pagnier, C. Robert, L. Barrassi, G. Fournous, et al., The virophage as a unique parasite of the giant mimivirus, Nature 455 (2008) 100–104.

[3] C.A. Suttle, Marine viruses—major players in the global ecosystem, Nat. Rev. Microbiol. 5 (2007) 801–812.

[4] J. Hietpas, L.K. McMullan, H.L. Hanson, C.M. Rice and D.P. Mindell, In R. Breeze, B. Budowle, S.E. Schutzer (Eds.), Microbial Forensics, Elsevier/Academic Press, 2005.

[5] B. La Scola, S. Audic, C. Robert, L. Jungang, X. de Lamballerie, M. Drancourt, et al., A giant virus in amoebae, Science 299 (2003) 2033.

[6] A.E. Smith, A. Helenius, How viruses enter animal cells, Science 304 (2004) 237–242.

[7] D. Baltimore, Expression of animal virus genomes, Bacteriol. Rev. 35 (1971) 235–241.

[8] W.I. Lipkin, Pathogen discovery, PLoS Pathog. 4 (2008) e1000002.

[9] W.I. Lipkin, G. Palacios, T. Briese, Diagnostics and discovery in viral hemorrhagic fevers, Ann. N.Y. Acad. Sci. 1171 (Suppl 1) (2009) E6–11.

[10] K.B. Chua, W.J. Bellini, P.A. Rota, B.H. Harcourt, A. Tamin, S.K. Lam, et al., Nipah virus: A recently emergent deadly paramyxovirus, Science 288 (2000) 1432–1435.

[11] T.G. Ksiazek, D. Erdman, C.S. Goldsmith, S.R. Zaki, T. Peret, S. Emery, et al., A novel coronavirus associated with severe acute respiratory syndrome, N. Engl. J. Med. 348 (2003) 1953–1966.

[12] CDC, Outbreak of West Nile-like viral encephalitis—New York, MMWR Morb. Mortal. Wkly. Rep. 48 (1999) 845–849.

[13] CDC, Serum cross-reactive antibody response to a novel influenza A (H1N1) virus after vaccination with seasonal influenza vaccine, MMWR Morb. Mortal. Wkly. Rep. 58 (2009) 521–524.

[14] CDC, Outbreak of acute illness--southwestern United States, MMWR Morb. Mortal. Wkly. Rep. 42 (1993) 421–424.

[15] S.R. Zaki, P.W. Greer, L.M. Coffield, C.S. Goldsmith, K.B. Nolte, K. Foucar, et al., Hantavirus pulmonary syndrome. Pathogenesis of an emerging infectious disease, Am. J. Pathol. 146 (1995) 552–579.

[16] K.B. Mullis, F.A. Faloona, Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction, Methods Enzymol. 155 (1987) 335–350.

[17] I.M. Mackay, K.E. Arden, A. Nitsche, Real-time PCR in virology, Nucleic Acids Res. 30 (2002) 1292–1305.

[18] E.M. Elnifro, A.M. Ashshi, R.J. Cooper, P.E. Klapper, Multiplex PCR: Optimization and application in diagnostic virology, Clin. Microbiol. Rev. 13 (2000) 559–570.

[19] G. Palacios, P.L. Quan, O.J. Jabado, S. Conlan, D.L. Hirschberg, Y. Liu, et al., Panmicrobial oligonucleotide array for diagnosis of infectious diseases, Emerg. Infect. Dis. 13 (2007) 73–81.

[20] N. Boonham, J. Tomlinson, R. Mumford, Microarrays for rapid identification of plant viruses, Annu. Rev. Phytopathol. 45 (2007) 307–328.

[21] T. Briese, G. Palacios, M. Kokoris, O. Jabado, Z. Liu, N. Renwick, et al., Diagnostic system for rapid and sensitive differential detection of pathogens, Emerg. Infect. Dis. 11 (2005) 310–313.

[22] G. Palacios, T. Briese, V. Kapoor, O. Jabado, Z. Liu, M. Venter, et al., MassTag polymerase chain reaction for differential diagnosis of viral hemorrhagic fever, Emerg. Infect. Dis. 12 (2006) 692–695.

[23] D.J. Ecker, R. Sampath, C. Massire, L.B. Blyn, T.A. Hall, M.W. Eshoo, et al., Ibis T5000: A universal biosensor approach for microbiology, Nat. Rev. Microbiol. 6 (2008) 553–558.

[24] R. Sampath, T.A. Hall, C. Massire, F. Li, L.B. Blyn, M.W. Eshoo, et al., Rapid identification of emerging infectious agents using PCR and electrospray ionization mass spectrometry, Ann. N.Y. Acad. Sci. 1102 (2007) 109–120.

[25] R. Sampath, S.A. Hofstadler, L.B. Blyn, M.W. Eshoo, T.A. Hall, C. Massire, et al., Rapid identification of emerging pathogens: Coronavirus, Emerg. Infect. Dis. 11 (2005) 373–379.

[26] M.W. Eshoo, C.A. Whitehouse, S.T. Zoll, C. Massire, T.T. Pennella, L.B. Blyn, et al., Direct broad-range detection of alphaviruses in mosquito extracts, Virology 368 (2007) 286–295.

[26a] D.J. Faix, S.S. Sherman, S.H. Waterman. Rapid-test sensitivity for novel swine-origin influenza A (H1N1) virus in humans. N Engl J Med 361(7) (2009) 728–9.

[27] M. Dufva, Introduction to microarray technology, Methods Mol. Biol. 529 (2009) 1–22.

[28] W.J. Wilson, C.L. Strout, T.Z. DeSantis, J.L. Stilwell, A.V. Carrano, GL. Andersen, Sequence-specific identification of 18 pathogenic microorganisms using microarray technology, Mol. Cell Probes 16 (2002) 119–127.

[29] C.W. Wong, C.L. Heng, L. Wan Yee, S.W. Soh, C.B. Kartasasmita, E.A. Simoes, et al., Optimization and clinical validation of a pathogen detection microarray, Genome Biol. 8 (2007) R93.

[30] D. Wang, L. Coscoy, M. Zylberberg, P.C. Avila, H.A. Boushey, D. Ganem, et al., Microarray-based detection and genotyping of viral pathogens, Proc. Natl. Acad. Sci. USA 99 (2002) 15687–15692.

[31] D. Wang, A. Urisman, Y.T. Liu, M. Springer, T.G. Ksiazek, D.D. Erdman, et al., Viral discovery and sequence recovery using DNA microarrays, PLoS Biol. 1 (2003) E2.

[32] E.R. Mardis, Next-generation DNA sequencing methods, Annu. Rev. Genomics Hum. Genet. 9 (2008) 387–402.

[33] D.L. Cox-Foster, S. Conlan, E.C. Holmes, G. Palacios, J.D. Evans, N.A. Moran, et al., A metagenomic survey of microbes in honey bee colony collapse disorder, Science 318 (2007) 283–287.

[34] W.I. Lipkin, G.H. Travis, K.M. Carbone, M.C. Wilson, Isolation and characterization of Borna disease agent cDNA clones, Proc. Natl. Acad. Sci. USA 87 (1990) 4184–4188.

[35] S. VandeWoude, J.A. Richt, M.C. Zink, R. Rott, O. Narayan, J.E. Clements, A borna virus cDNA encoding a protein recognized by antibodies in humans with behavioral diseases, Science 250 (1990) 1278–1281.

[36] Q.L. Choo, G. Kuo, A.J. Weiner, L.R. Overby, D.W. Bradley, M. Houghton, Isolation of a cDNA clone derived from a blood-borne non-A, non-B viral hepatitis genome, Science 244 (1989) 359–362.

[37] R. Harmon. In B. Budowle, R.G. Breeze, S.E. Schutzer (Eds.), Microbial Forensics, Academic Press, San Diego, 2005.

[38] M.L. Metzker, D.P. Mindell, X.M. Liu, R.G. Ptak, R.A. Gibbs, DM. Hillis, Molecular evidence of HIV-1 transmission in a criminal case, Proc. Natl. Acad. Sci. USA 99 (2002) 14292–14297.

[39] K. Johnson, Worker in hepatitis case is sentenced to 20 years, N.Y. Times (2009).

[40] CDPHE, Case numbers associated with hepatitis C investigation, 2009, posting date [Online].

[41] B. Budowle, S.E. Schutzer, J.P. Burans, D.J. Beecher, T.A. Cebula, R. Chakraborty, et al., Quality sample collection, handling, and preservation for an effective microbial forensics program, Appl. Environ. Microbiol. 72 (2006) 6431–6438.

[42] S.E. Schutzer, P. Keim, K. Czerwinski, B. Budowle, Use of forensic methods under exigent circumstances without full validation, Sci. Transl. Med. 1 (2009) 8cm7.