several techniques currently under study: time of death analysis, personal identification through skin bacteria, body fluid identification, and microbial biogeography.

The fourth and final part describes the evidentiary framework that will determine the admissibility of testimony based on microbial forensics. Although we now regard DNA typing as the gold standard of forensic science, there were missteps when DNA was rolled out for legal use.[2] To avoid similar missteps in the use of microbial forensics, this part applies the governing admissibility standards to several proposed uses of microbial techniques. The final part offers caveats about some of the uses and identifies areas where further empirical research is arguably necessary. However, the part also points out that research is already under way to fill some of these gaps in the empirical record. Our belief is that several types of testimony based on microbial analysis will soon gain admission in American courts.

I. MICROBIAL ANALYSIS

A. The Microbial Cloud

For centuries, medical science generally ignored the vast majority of microbes—those that are benign to humans, our animals, and our crops—and concentrated instead on those rare microbes that are pathogenic (the ones that cause disease). It is understandable that medical science had such blinders. The impetus for modern microbiology came from sometimes desperate efforts to explain and cure outbreaks of terrible illnesses such as cholera, typhoid, plague, and yellow fever. These diseases killed millions of people and occasionally threatened whole civilizations. Such awful consequences virtually forced the medical community to primarily focus its early research efforts on pathogens.

However, the bulk of microbial life is not pathogenic. Our bodies are home to some two to six pounds of microbial life—organisms that do not share our DNA but replicate and live on our skin, hair, in our colons, in our blood, between our toes, and in our mouths; they are virtually everywhere upon and inside us[3]. There are approximately ten times as many non-human cells in our bodies as there are cells containing our DNA.[4] Some, perhaps most, of those microbes are generally not harmful to us, and some appear to even be beneficial. By way of example, the bacteria in our colons (of which there are trillions upon trillions) are essential to proper digestion.[5] There are also suggestions that our bacterial communities may aid in the development of the immune system, fight off pathogens, regulate our metabolism, and even shape our behavior.[6] Most likely, we have just scratched the surface of the myriad effects that microbes have on human beings.

Although scientists have known about the collection of microbes found in and on people and other organisms for many years, it was not until the late 2000s that the "human microbiome" began to gain significant attention. In the late 2000s, scientists dubbed this symbiotic group of bacteria and other creatures the "microbial cloud."[7] The effects and interactions of the microbial cloud are important enough that some scientists have proposed designating it a new organ of the human body. For instance, bacteria appear to be essential to our bodies' process for digesting fat although in different individuals, different species of bacteria do the work.[8]

A growing appreciation of the human microbiome and what it may do is beginning to have profound effects on medicine. For example, medicine is now expanding the use of fecal transplants (also known as fecal microbiota translation (FMT) or stool transplant) to "infect" a patient with healthy intestinal bacteria to treat microbe-related diseases.[9] Doctors can effectively cure patients suffering from Clostridium difficile with a single infusion of fecal material from a healthy donor. Doctors transplant the fecal material through enemas, a form of "poo tea", or, in the future, perhaps in pill form. Biologists have learned that some species of mammals engage in coprophagia, the consumption of feces, which may be done to obtain bacteria needed to digest various vegetable matter.[10] These species include elephants, koalas, and pandas.

The microbial cloud not only has noteworthy medical applications; the same concept also has important potential forensic applications. For example, the early stages of a post-mortem body's decomposition are largely governed by unchecked growth of the person's anaerobic intestinal flora. Through a study of post-mortem bacteria, an expert might determine a decedent's time of death with much greater accuracy than under current techniques.

The estimation of time of death does not exhaust the forensic possibilities. In our normal activity, we constantly shed microbes. We shed them through breathing, coughing, touching, eating, drinking, and sex. Unlike forensic fingerprints or DNA in blood

[2] Paul C. Giannelli, Edward J. Imwinkelried, Andrea Roth & Jane Campbell Moriarty, Scientific Evidence § 18.05 (5th ed. 2012)[hereinafter 2 Giannelli].

[3] Jonathan Eisen, Ted Talk, available at http://www.ted.com/talks/jonathan_eisen_meet_your_microbes, 2:40 -3:10 (last visited April 11, 2014) [hereinafter Eisen]; see also Jeroen Raes, TedX, http://tedxtalks.ted.com/video/Jeroen-Raes-at-TEDxBrussels.

[4] NIH Human Microbiome Project, available at http://www.nih.gov/news/health/jun2012/nhgri-13.htm (last visited April 11, 2014) [hereinafter NIH Human Microbiome Project].

10 NIH Human Microbiome Project, supra note 9.

[5] NIH Human Microbiome Project, supra note 9.

[6] Eisen, supra note 8, at 7:10-:30.

[7] Eisen, supra note 9, at 14:21.

[8] NIH Human Microbiome Project, supra note 9.

[9] Eisen, supra note 8, at 9:50 to 11:10.

[10] Eisen, supra note 8, at 9:50

droplets, the bacteria we shed are themselves living organisms. Wherever the bacteria land, they continue to grow and spread. These new colonies share a common communal and genetic link with the communities they left behind. We leave a discoverable living trail behind us, wherever we go and whatever we do. The existence of that trail gives rise to the possibility of using microbial analysis as a means of personal identification.

B. Phylogenetic Analysis

The use of microbial analysis for personal identification relies not only on the general notion of the microbial cloud, but also on more specific, new genetic techniques enabling analysts to easily and cheaply identify individual bacteria and viruses. Until the late 20th century, scientists had to isolate microbes from their environment and then grow them in the laboratory (a process known as culturing) to identify them. This process was cumbersome; many microbial species resist culturing and were consequently effectively hidden from discovery. Today, the most potent method of identification is by genetic sequencing in which the strings of chemical components making up the organism's genome are "read." The components come in four chemical forms or "letters" known as bases; these are abbreviated as A, C, T, and G. The order of the bases is called a DNA sequence. Most importantly, scientists can identify microbes by their particular sequences which differ between and even within species. Significantly, scientists can read the sequences of microbes without ever growing the organism in the laboratory—for instance, from swabs collected at a crime scene.

A single bacterium has a genome usually containing anywhere from 1,000,000 to 10,000,000 individual letters comprising its full "genome sequence," while viruses are much smaller, a few thousands to tens of thousands of letters. In theory, one could read the entire genome sequence of multiple organisms in order to perform a forensic analysis. However, in practice that would be overkill. Instead, researchers frequently focus on smaller portions of the genome that serve as diagnostic markers for particular microbes. One such bacterial marker has been a single gene known as the "small submit rRNA (ss-rRRNA) gene." [11] This gene represents only about 1,500 bases out of the genome, but it is a very powerful tool because a version of the gene is found in all living beings.[12] Although the gene is found in all organisms, the sequence of the bases differs between species and is not very prone to mutation.[13] Scientists can then compare the sequences of these ss-rRNA genes between species in much the same way that scientists line up bone structures between different vertebrates, to build a "pedigree" or evolutionary tree of the organisms. Further, researchers have already constructed public databases with billions of ss-rRNA sequences from various organisms and samples to which one can compare data from a new sample.[14] By building evolutionary trees of the ss-rRNA genes, one can then determine what organisms were present in a sample in much the same way as a paleontologist would determine the species found in a lake bed by looking at the bones found in the bed and comparing them to bones of known organisms. One can take a sample such as a crime scene swab and read the sequences of the ss-rRNA genes of the microbes found in the sample relatively easily and cheaply. Since the relative abundance of the bacterial communities are thought to be unique to each person and reasonably constant over time[15], in theory an analysis of the bacterial communities will link a person's identity to bacteria on crime scene debris. This process is conceptually similar to the way that law enforcement has used fingerprints for more than a century. Simply stated, microbial science may become the biggest advance in forensic science since the advent of DNA matching.

In addition to taking a census of the bacterial communities, forensic scientists can trace the genetic lineage of individual cells. They cannot use microbial genetic material in exactly the same way as human DNA. When searching for a DNA match in CODIS, forensic experts may correctly assume that a suspect's DNA will remain stable. Mutations or other changes to DNA generally occur during reproduction; but since police are usually uninterested in the suspect's children or grandchildren, law enforcement has little need to consider any mutations or variation to the DNA they find.

In sharp contrast, when examining microbial genes for forensic purposes, it seems unlikely that one could ever observe a "perfect" match when comparing microbiomes. There is never a direct match. The microbial cloud of a person is composed of more than a trillion individual organisms. These microbes are constantly dying, reproducing, and occasionally undergoing mutations and other types of genetic change. Since bacterial cells can reproduce themselves in as little as fifteen minutes and viruses usually within less than 24 hours, their genomes may exhibit substantial differences in a very short time. A single day in the life of such a bacterial colony can produce as many generations as humans have since the time of the Peloponnesian Wars, 2500 years ago.

With this rapid cycle of microbial reproduction and mutation

[11] A. Fabrice & R. Didier, Exploring Microbial Diversity Using 16S rRNA High-Throughput Methods. 2 Journal of Computer Science and Systems Biology 75 (2009).

[12] Except for viruses.

[13] No living creature or genetic sequence is completely resistant to mutation. However, the ss-RRNA gene is extremely conserved between generations.

[14] Just one of them is the Ribosomal Database Project, available at http://rdp.cme.msu.edu/seqmatch/seqmatch_intro.jsp (last visited April 21, 2014).

[15] Elizabeth Costello et. al., Bacterial Community Variation in Human Body Habitats Across Space and Time, Science 326 (2009) [hereinafter Costello].

occurring with each generation, evolutionary forces work many orders of magnitude faster in microbes than in animals. For bacteria, measurable differences in abundance of particular species can be mapped in terms of a few hours or days, such that taking a "census" of all bacterial species present can give a forensic scientist a picture of who was there at what time. Viruses accumulate genetic changes within weeks, and these differences mean that there will often not be an exact match between all the genes of two viruses even when they originate from the identical source. Nevertheless, viruses sharing a recent common ancestor will have substantial more similarities in their genetic makeup than viruses of two randomly chosen individuals. These similarities can be charted by using phylogenetic analysis.

Phylogenetic analysis is the study of evolutionary relationships between groups of organisms. For instance, biologists have applied phylogenetic analysis to determine the "tree of life", that is, the general relatedness of all living things[16]. While phylogenetic analysis does not require genetic material, DNA and RNA sequences in bacteria and viruses can serve as data permitting statistical analyses to determine the probability of common ancestry and to outline the structure of evolutionary trees. This type of analysis has already allowed scientists to hypothesize that the most recent common ancestor of humans, chimpanzees and, bonobos (our closest relatives in the animal kingdom) lived between five million and seven million years ago. That timeframe represents approximately 250,000 human generations. A similar analysis of human mitochondrial DNA allowed scientists to determine that the most recent female common ancestor of all living humans lived between 100,000 and 200,000 years ago—most likely in East Africa.[17]

Phylogenetic analysis will be essential for any future microbial forensic work. Consider the following situation: A couple has sex, one of them is HIV positive, and after a few weeks the other person tests HIV positive. If we assume that person became infected during that act of sex, his or her virus is genetically related to his or her sexual partner's virus; they have a very recent common ancestor. Several months later, a forensic researcher collects samples from both persons and sequences the virus in their blood. In those few, intervening months, many generations (and thus mutations) have separated both viral strains. In order to link or "match" the two viral strains, a scientist must conduct an evolutionary analysis of the viral genes to find common markers. To a high degree of probability, these markers can indicate a recent common ancestor and provide a statistically probable link between the two. Whether that link is direct or indirect can never be ascertained, since the viruses will never be identical.

Another form of phylogenetics is community analysis. Rather than look for mutations in individual cells, the study here is to determine the bacterial phylotypes in a given sample. For instance, as with the previous example of a hand touching a doorknob, a community of microbes is transferred from the surface of the skin to the doorknob. The palm surface of any two individuals share only 13% of the same bacterial phylotypes.[18] The bacteria on the doorknob are genetically related to the ones remaining on the person's skin. One week later, a forensic researcher collects samples from both the person's fingertips and the doorknob. In that one week interval, our person may have washed her hands many times, killing not only pathologic bacteria but also symbiotic or neutral bacteria comprising the bacterial colonies on her hands. Since the bacteria grow back within hours, several thousands of generations (if not more) can separate the bacteria sampled on the person's hand from those left on the doorknob. In the meantime, the bacteria on the doorknob may have multiplied and mutated as well, though likely not as much because the environment is much less suitable for bacterial growth. By examining the community and taking a census of the phylotypes present, it is possible to match persons by comparing the composition of the bacterial communities. The phylogenetic analysis used here is quantitative and/or weighted phylogenetics.

II. PRIOR COURTROOM USES OF PHYLOGENETIC ANALYSIS: PROOF OF INFECTION TRANSMISSION

Phylogenetic analysis has been accepted as evidence by American and foreign courts in cases of infection.[19] The earliest forensic use of microbial genetic data involved HIV infection by a Florida dentist of his patients, investigated in 1992.[20] However, since that case settled out of court after pre-trial discovery, no court had an opportunity to use evidentiary standards to evaluate the reliability of the phylogenetic analysis. Many other cases have followed, including the first use of HIV forensic evidence in court where the

[16] Carl Woese, A Manifesto for Microbial Genomics, 8 Current Biology Magazine, No, 22, page R781 (Nov. 5, 1998)

[17] Rebecca L. Cann, Mark Stoneking & Allan C. Wilson (1987), Mitochondrial DNA and Human Evolution, 325 Nature 31 (1987); Pedro Soares P et al. (June 2009), Correcting for Purifying Selection: An Improved Human Mitochondrial Molecular Clock, 84 Am. J. Hum. Genet. 84 740 (2009).

[18] Fierer, supra note 5, at 6477.

[19] It should be pointed out that there are significant differences in the science used herein. RNA viruses, like HIV and HCV, have significantly different life cycles, mutation rates and genetic material than bacteria and other microbes. Nonetheless, the techniques used in the HIV and HCV cases that follow will form the legal basis for microbial forensics when they are first asserted in the courtroom; the techniques for studying bacteria and other microbes are similar.

[20] This was the case of Dr. David Akers. See AIDS and a Dentist's Secrets, New York Times, June 6, 1993, p. 3.

defendant was convicted of rape based on the viral genetic analysis.[21] Two other cases are documented here.

A. Case One: The Spanish Prosecution of Dr. Juan Maeso

To date, the most extensive use of phylogenetic analysis in court was in the prosecution of Dr. Juan Maeso, an anesthesiologist practicing in Valencia, Spain.[22] In February 1998, a utility company's doctors noticed a large outbreak of Hepatitis C (HCV, for Hepatitis C Virus) among their patients in Valencia; some 20 people were infected.[23] Twenty diagnosed cases is significant, since HCV is normally asymptomatic and decades can pass before the damage done by the virus becomes noticeable. When HCV is not cleared immediately it establishes a chronic infection that can be quiescent for decades.[24] Chronic infections can in the long run lead to cirrhosis, liver cancer, and liver failure. The disease is transmitted primarily by blood products through unscreened blood transfusions, contaminated needles, traumatic (that is, bleeding) sexual intercourse, and organ transplants. HCV is more common in the developing world where sterile techniques are not rigorously followed. Thus, the discovery of 20 cases appearing simultaneously in a well-developed country such as Spain amounted to an extraordinary outbreak.

The 20 diagnosed cases prompted a much larger epidemiological study. The epidemiological study identified and correlated 322 new cases of HCV to two hospitals in Valencia.[25] While the epidemiologists considered numerous potential common denominators including the surgeon, surgery room, type of surgery and type of anesthesia, the most common factor shared by the victims was their anesthesiologist, Dr. Juan Maeso.[26] Although these facts seemed to incriminate Dr. Maeso, the epidemiological study indicated only correlation with the victims and provided little evidence that Dr. Maeso had caused the infections

To obtain an expert evaluation of the significance of the facts, the Spanish court turned to academic geneticists from Valencia.[27] The court tasked the experts to evaluate the infected individuals and Dr. Maeso to determine which, if any, of the 322 HCV victims had been infected by Dr. Maeso.[28]

HCV, like HIV and influenza, uses an RNA genetic information carrier, which is particularly prone to accumulating mutations.[29] Rates of evolution (accumulated mutations) are measured in nucleotide substitutions per site per year. For human DNA, the rate measures 10-7 substitutions per site per year.[30] In contrast, RNA viruses yield evolutionary rates more than ten thousand times faster.[31] Given the number of victims and the time period involved, the geneticists expected to find an extensive evolutionary tree linking the victims' viruses.

Blood samples were taken from Dr. Maeso and all 322 suspected victims.[32] In order to build a control into the study, the experts sampled an additional 44 HCV-positive patients from the Valencia region with no known association to Dr. Maeso or the two hospitals in question.[33] Under Spanish law, Dr. Maeso was required to provide one blood sample but was allowed to refuse any additional samples.[34] Initially, the experts attempted to determine whether the victim's HCV was more likely to have an ancestral relation to Maeso's HCV (the Presumed Source) than to the control patients' HCV.

The geneticists methodically approached their task given the tools available to them some 15 years ago. First, they sequenced two parts of the HCV virus, known to exhibit different rates of evolution. The geneticists took a 229-nucleotide sample from the NS5B gene, located on a relatively conserved part of the genome that encodes non-structural proteins.[35] In addition, they analyzed a second part of the genome, a 406-nucleotide fragment from the E1-E2 region that encodes structural proteins and is prone to changes.[36]

The NS5B Analysis

The outcome of the NS5B analysis was not very incriminating. In fact, it was difficult to distinguish any of the infected parties,

[21] Albert J, Wahlberg J, Uhlén M., Forensic evidence by DNA sequencing. Nature. 1993 Feb 18;361(6413):595-6.

[22] González-Candelas et al., Molecular Evolution in Court Analysis of a Large Hepatitis C Virus Outbreak from an Evolving Source. BMC Biology 2013 11:76, July 19, 2013, at p. 1 [hereinafter González-Candelas]; see also Vandamme AM, Pybus OG. Viral phylogeny in court: the unusual case of the Valencian anesthetist. BMC Biol. 2013 Jul 19;11:83.

[23] Shaoni Bhattacharya, Nature, Feb. 27, 2014, at p. 425 [hereinafter Bhattacharya].

[24] Centers for Disease Control, Hepatitis C Information for the Public, available at http://www.cdc.gov/hepatitis/C/index.htm (last visited April 21, 2014).

[25] González-Candelas, supra note 27, at 2.

[26] González-Candelas, supra note 27, at 2.

[27] González-Candelas, supra note 27, at 1.

[28] González-Candelas, supra note 27, at 2. The court instructed the experts to address six questions: "1) check whether the suspected source was actually responsible for the outbreak, 2) ascertain which patients had been infected from a common source and could be considered as included in the outbreak and who had been infected from alternative sources, 3) discard alternative sources or the existence of different but simultaneous outbreaks, 4) determine the duration of the outbreak, 5) date the time of infection for each patient involved in the outbreak, and 6) determine the date of infection of the source."

[29] Ana B. Abecasis, A. M. Geretti J. Albert, L. Power, M. Weait & M. Vandamme, Science in Court: The Myth of HIV Fingerprinting, 11 Lancet Infect. Dis. 78 (2011)[hereinafter Vandamme].

[30] Vandamme, supra note 34.

[31] The mutation rate is measured in nucleotide substitutions/site/year.

[32] González-Candelas, supra note 27, at 2.

[33] González-Candelas, supra note 27, at 2.

[34] González-Candelas, supra note 27, at 2.

[35] González-Candelas, supra note 27, at 2.

[36] González-Candelas, supra note 27, at 2.

either from the suspected victims or the control group. Standing alone, the results of the NS5B analysis did not support the hypothesis that the HCV infections of the patients in the outbreak group were more closely related to Dr Maeso than to the HCV infections of the control group who had no connection to Dr. Maeso. Although they used two different computational methods to analyze the data, the results did not separate the outbreak group from the control group.[37]

The E1-E2 Hypervariable Region Analysis

In contrast, the outcome of the E1-E2 analysis was highly incriminating.

The phylogenetic analysis. Ninety seven patients in the outbreak group exhibited virus sequences with close similarities to the Presumed Source.[38] The data indicated that these samples had a relatively recent single common ancestor, while the common ancestor of the control groups and the outbreak patients was much older, they were much more distantly related. Unlike the NS5B analysis, these findings separated the outbreak group from the control group, correlating Dr. Maeso's infection with those of the outbreak victims. At this point, the most that can be inferred with confidence is that the Presumed Source and many of the outbreak victims were infected from a common source.

However, even at this point it could not be inferred that Dr. Maeso was the source. The source could be an unclean surface in an operating room, a non-doctor working in those rooms or one of the other outbreak patients.[39]

The molecular clock. However, the court-appointed experts did not conclude their analysis there. Given enough data, in some cases a molecular clock can be constructed to determine the time of infection. Given the times of infection, a vector sequence (i.e. who was infected earlier than whom) can be inferred. Genomic material mutates at a relatively constant rate over time, though at different rates for different species and subspecies.[40] After determining the number of mutations in patients with known infection dates, an investigator can conduct a regression analysis to establish a formula for computing an approximate date of infection. In Dr. Maeso's case, the geneticists isolated 24 patients who: (a) had only a single contact with the Presumed Source; (b) tested negatively for HCV before that contact, but (c) tested positive for HCV afterwards.46 Using this data, the scientists created a time formula and mapped the results for all those infected.

The molecular clock was very helpful to the prosecution. Dr. Maeso's infection was much older than those of the outbreak victims, and for many of the outbreak patients the approximate estimated timing of the infection did not contradict the hypothesis of infection during surgery. There were far more mutations evident in the viruses derived from Dr. Maeso than in almost any other infected person. These findings were consistent with the hypothesis that he was the source of the outbreak patients' infection (although standing alone those findings did not prove the hypothesis).[41]

The Spanish court accepted both the experts' molecular clock indicating a much earlier infection and their phylogenetic tree pointing to an extremely high likelihood of related infection. Along with other evidence, the court used this information to sentence Dr. Maeso to 1,933 years in prison.[42] Dr. Maeso has since exhausted all his potential appeals. The conclusion of Dr. Maeso's legal proceeding allowed the court-appointed geneticists to publish a full account with raw genetic data in November 2013.

B. Case Two: The American Prosecution of Dr. Richard Schmidt

On the night of August 4, 1994, Dr. Richard Schmidt entered the apartment of his mistress and announced he was going to inject her with a vitamin B12 shot.[43] His mistress, Janice Trahan, was in bed at the time with their three-year-old son. Despite Ms. Trahan's refusal, Dr. Schmidt injected her with a substance that he claimed was vitamin B12.[44]

Four months later, Ms. Trahan was diagnosed with HIV and HCV[45] The prosecution theorized that Dr. Schmidt had taken blood samples from two of his patients, one HIV-positive and one HCV-positive, and inserted those blood samples in his syringe. On

[37] González-Candelas, supra note 27, at 3. "Neighbor-joining and maximum likelihood phylogenetic trees obtained from the NS5B sequences failed to group all the control samples in a monophyletic group. … As a consequence, the phylogenetic signal in this region was too low to reliably separate the local controls, the patients infected from a common source and the patients infected from alternative sources."

[38] González-Candelas, supra note 27, at 3.

[39] Vandamme, supra note 34.

[40] Philippe Lemey & David Posada, Molecular Clock Analysis, The Phylogenetic Handbook 362 (eds. P. Lemey, M. Salemi & Vandamme 2009) [hereinafter Lemey].

[41] At first blush, this result may be counter-intuitive. However, consider languages as an analogy. Languages also mutate through gradual changes over the centuries. There is a similar effect with the spread of English, where a source region (i.e. England) later seeds numerous "infections" (e.g. Australia, New Zealand, Canada, and the United States). Even when those infections grow much bigger than the original host (e.g. the United States), the English spoken near the source (i.e. England) has much greater variety; there is more time for mutation and evolution at the original source. From a statistical standpoint, the variability of language spoken in the UK compared to the US is a clear indication that the United States was "infected" with English later than the U.K. It is therefore not surprising that similar phylogenetic techniques are used for investigation of languages (Remco Bouckaert et al.,Mapping the Origins and Expansion of the Indo-European Language Family, 337 Science 957 (2012); Letters, Corrections and Clarifications, 342 Science 1446 (2013).

[42] Bhattacharya, supra note 28, at 424.

[43] State v. Schmidt, 771 So. 2d 131, 135 (La.App. 2000)[hereinafter Schmidt].

[44] Schmidt, supra note 49, at 135.

[45] Schmidt, supra note49, at 135

that evening while his wife was bathing, Dr. Schmidt quickly left his home, drove to his mistress's apartment, and injected her with the contaminated syringe. He then returned home before his wife finished her bath.[46] To prove its case against Schmidt, the prosecution hired Dr. Michael Metzker to conduct a phylogenetic examination of Ms. Trahan's HIV infection.[47]

Dr. Metzker employed a phylogenetic technique similar to the one used in the Spanish prosecution. Taking 32 HIV positive controls from the city where Dr. Schmidt and Ms. Trahan lived as well as samples from the suspected donors (Schmidt's patients) and the victim (Ms. Trahan), Dr. Metzker constructed a phylogenetic tree.[48] As in the Spanish prosecution, the objective was to compare the HIV (and HCV) infections of the controls, the victim, and the Presumed Source patients. The evolutionary tree showed that the HIV sequence from one of Dr Schmidt's patients and the victim were far more similar to each other than to any member of the control group, the same was found for the HCV sequence. Using three different computational phylogenetics techniques (parsimony, minimum evolution, and Bayesian analysis), Dr. Metzker concluded that there was a strong indication of recent common ancestry between the donor infection and Ms. Trahan's.[49]

In the trial court, Dr. Metzker's methodology was subjected to a Daubert hearing and ruled admissible—the first instance of phylogenetic analysis admitted in the United States.[50] The prosecution presented Dr. Metzker's expert testimony as well as testimony describing a parallel study conducted at the University of Michigan.[51] The defense called its own witness, Dr. Bette Korber. Both the prosecution's and the defense's experts testified that while phylogenetic analysis is capable of excluding suspects, it is generally incapable of proving either a direct infection or the direction of an infection.[52] Further, both side's experts concluded that intervening parties cannot be excluded in a phylogenetic analysis.[53]

The trial jury convicted Dr. Schmidt, and he was sentenced to fifty years at hard labor.[54] Schmidt took an appeal to the Louisiana Third Circuit Court of Appeals, the intermediate appellate court in that state. The court affirmed the trial judge's rulings. In upholding the conviction, the court of appeals pointed to: Ms. Trahan's direct testimony that the defendant had injected her, circumstantial "consciousness of guilt" evidence that Dr. Schmidt attempted to conceal or destroy his office's routine records of the suspected HIV donor's blood draw earlier on the same day as Ms. Trahan's injection, and the phylogenetic analysis. The Louisiana Supreme Court subsequently denied Schmidt's petition for further review of his conviction.[55]

Summary

The Spanish and Louisiana prosecutions were among the first to employ microbial analysis. Both cases were tried around 2000. Since then a number of criminal and civil cases around the world have utilized similar techniques. Although the use of these techniques in court was a novel development, both cases were receptive to evidence of phylogenetic analysis.

The Schmidt case is obviously of importance in the United States because it was the first time phylogenetic analysis had been admitted in an American legal proceeding over a Daubert objection. However, the Spanish case is perhaps the best documented and certainly the case with the largest data collection. In the Spanish case, phylogenetic analysis played a prominent role in the prosecution's case-in-chief. The Spanish case is also notable because it may be the only case to date in which the court admitted a molecular clock analysis.

III. THE BRAVE NEW WORLD OF MICROBIAL FORENSICS: OTHER POTENTIAL USES OF MICROBIAL ANALYSIS

As we saw in Part II, the early legal cases involving microbial analysis focused on the transmission of viral infection. However, in the near future the courts will probably encounter other, more

[46] Schmidt, supra note49, at 142.

[47] Schmidt, supra note 49, at 144-45.

[48] Schmidt, supra note 49, at 152.

[49] Michael L. Metzker et al., Molecular Evidence of HIV-1 Transmission in a Criminal Case, 99 Proceedings of the National Academy of Sciences, no. 22, Oct. 29, 2002, at p.14296 [hereinafter Metzker]. The level of confidence was 95,826 out of 100,000 parsimony bootstrap replicates, 10,000 out of 10,000 replicates with a maximum likelihood distance analysis, and 25,000 out of 25,000 sampled trees for Bayesian analysis.

[50] State v. Schmidt, 699 So. 2d 448 (La.App. 1997).

[51] Schmidt, supra note 49, at 145.

[52] The defense further complained that the control sample should have been restricted to HIV-positive patients with characteristics similar to both the suspected donor and Ms. Trahan, i.e., heterosexual, non-intravenous drug using females and homosexuals males infected in 1994. However, that restriction would have greatly reduced the confidence of the computations in the case, and the court did not agree to the restriction. Schmidt, supra note 49, at 145.

[53] This would be the case if the suspected donor had infected an unknown third party, such as another illicit lover, who then went on to infect Ms. Trahan. As a result of this and other attacks on the evidence, Ms. Trahan was forced under oath to give a detailed sexual history including seven male partners. Schmidt, supra note 49, at146. Five of Ms. Trahan's former sexual partners were subpoenaed to testify at trial. Schmidt, supra note 49, at 138. None of her former partners tested positive for HIV or HCV.

[54] Schmidt, supra note49, at 135. Louisiana sentences for hard labor are usually carried out at the Angola facility, colloquially called "The Farm", where agriculture and manufacturing take place. It is a former slave plantation.

[55] Phylogenetic analysis is at the very core of evolutionary science. Hence, it is somewhat ironic that the first acceptance of phylogenetics in the United States was in Louisiana, which in 1982 mandated the scientific treatment of Creationism in schools.

esoteric applications of microbial analysis. In the final analysis, phylogenetic analysis underpins virtually all microbial forensic work. There will likely never be a direct match between populations of microbes. Consequently, it will always be necessary to build an evolutionary tree through phylogenetic computations to determine how close the relatedness is between different populations. Alternatively, populations of millions of microbes may be compared at one time to find matching communities. In all cases, "matching" will never reveal the samples to be identical but rather so closely related that they can be used to discriminate between hypotheses put forward by court. This is all in the nature of phylogenetics.

Today forensic research laboratories are overflowing with ideas how best to exploit the microbial cloud. This part of the article examines four potential uses other than transmitting infections, that will likely surface in court soon: (1) the estimation of the post-mortem interval (PMI); (2) personal identification based on the analysis of skin microbial communities; (3) the identification of the type of body fluid; and (4) biogeographical analysis to determine the source of a soil sample. This list is not exhaustive. Microbial forensics is an active field of study brimming with new ideas. However, a review of these four uses should give the reader a sense of the general direction and pace of microbial forensics. Depending on future research and the judicial acceptance of the research findings, microbial forensics could well become the most powerful forensic tool since DNA matching.

A. The Estimation of Post-Mortem Interval

The Traditional Techniques

It is often critical to determine when the decedent died. The defendant may have an airtight alibi at a particular time. The current techniques for determining post-mortem intervals (PMI) are imprecise. It is sometimes said that this field of forensic science is still in the Dark Ages.[56] Most of the techniques were developed in the 19th century. Although estimates based on the newest technique, forensic entomology, are subject to a huge number of variables,[57] the courts quickly embraced the technique in part because the courts realize the weaknesses of the other traditional techniques. However, as decomposition is mediated by the body's bacterial flora, microbial research offers a more direct measurement of PMI. Generally, the state of a corpse determines what techniques can be used. Scientists have identified stages of decomposition as: (a) fresh, before decomposition begins; (b) active decay, when bloating and rupture occur; and (c) advanced decay, after decomposition fluids begin to leak.

When a body is fresh, that is, discovered within hours after death, there are three primary means to determine PMI. First, the pathologist can rely on algor mortis, the falling of the body temperature.[58] Estimation based on body temperature (measured from 3.5" inside the rectum or around the mass of the liver) has long been considered the most accurate short-term measure of PMI. Yet, numerous external factors can affect the reading. These factors include the mass of the body (larger bodies have a smaller surface area to volume ratio, retarding a body's ability to cool), the position of the body (a supine body is exposed for 80% of its total surface area and will cool faster than a body in the fetal position with only 60% of its surface exposed), the clothing and other coverings, air movement, humidity, air temperature (bodies do not cool under hot or tropical conditions) and full or partial immersion in water (water is a better conductor of heat than air).[59] In sum, although there are tens of judicial opinions admitting PMI estimates based in part on algor mortis, in the view of Dr. David Carter, "Algor mortis is not [scientifically] acceptable for estimating PMI."[60]

Rigor mortis refers to the stiffening of the corpse. The onset of rigor mortis is another early indicator of PMI, though a much less accurate one than algor mortis. The popular rule-of-thumb is that rigor starts to set in within six hours, becomes fully established in another six hours, remains for 12 hours, and wears off after an additional 12 hours.[61] Numerous factors, including environmental temperature and the degree of muscle activity before death, may affect the course of rigor mortis.[62] Moreover, children and the elderly proceed into rigor faster than adults, as do persons who expire from particular diseases or by particular means (e.g. asphyxiation and carbon monoxide poisoning). Experimental data show that corpses exhibit rigor on a wide bell curve of intervals, from less than two hours to more than 13 hours after death. Again, despite the widespread judicial acceptance of opinions resting in part on rigor mortis, in the words of Dr. Carter, "Rigor mortis is not [scientifically] acceptable for estimating PMI."[63]

Yet, another technique is to examine the corpse's eyes for concentrations of potassium. However, this technique has major limitations. For instance, experts even disagree over the proper

[56] Marshall Houts, Time of Death: Still the Dark Ages of Proof, 10 Trauma 7 (Aug. 1968).

[57] Giannelli, supra note 7, at § 19.8[a], at 283-86.

[58] Derrick Pounder, lecture notes, Head of Department of Forensic Medicine, University of Dundee, available at http://www.dundee.ac.uk/forensicmedicine/notes/timedeath.pdf, page 3 (1995) (last visited April 21, 2014) [hereinafter Pounder].

[59] Pounder, supra note 64, at 5.

[60] David Carter Ph.D., Assistant Dean and Director of Forensic Sciences, Chaminade University of Honolulu, An Overview of Forensic Taphonomy, http://forost.org/seminar/Segundo_seminario/Overview_of_Forensic_Taphonomy.pdf, at 50, (last visited 4/18/14) [hereinafter Carter].

[61] Pounder, supra note 64, at 9.

[62] Pounder, supra note 64, at 7-10.

[63] Carter, supra note 66, at 52.

formula to use to apply this method.[64]

The reality is that once 24 to 48 hours have passed, the body has cooled to ambient temperature, and rigor has passed, the traditional, short-term means of determining PMI are of little use. Forty-eight hours after death, the accuracy of PMI forensic techniques deteriorates further. All bodies decompose, but at different rates. The primary factors impacting the rate of decomposition are: temperature, moisture, pH, and the partial pressure of oxygen.[65]

Given the weaknesses of the traditional short-term indicia of PMI, in the past 20 years pathologists have placed greater focus on environmental changes in and around a corpse. Thus, exerts have turned to blow-fly larvae, "gravesoil" pH, and other external measures to estimate PMI.

In particular, the entomological analysis of flies and beetles is now in widespread use and quickly gained judicial acceptance. When a mammal dies, its body becomes a rich source of nutrients for the creatures and environment around it. Forensic scientists have used blowfly larvae for decades to estimate PMI. Forensic entomology often yields a range of possible PMI dates, and the accuracy of a PMI estimate resting on entomological analysis is subject to changing weather or seasonal conditions, local growth curves of the insect community, and uncertainty about the interval between actual death and the deposition of eggs in the tissue.[66] These variables leave room for substantial inaccuracy in PMI estimates.

Gravesoil pH is another environmental factor that has gained increasing attention in the past 20 years. For corpses in advanced decay, changes to the gravesoil can indicate an approximate PMI. After rupture, fluids from a corpse leak into the soil around the corpse. By measuring the pH and chemical composition of the soil immediately adjacent to the corpse and comparing those measurements to those for surrounding soil, it is possible to venture a rough gauge of PMI. However, such measurements are fraught with potential error. The sources of possible error include the specific body composition, the geometry of corpse and surrounding terrain, the location and extent of the rupture, and environmental factors, notably temperature and precipitation.

Microbial Techniques

Microbial analysis has inherent advantages over methods described above. The non-human bacteria living in and around the human body mediate the process of decay. In microbial PMI studies, researchers consider two stages: pre-rupture and post-rupture. After death, anaerobic organisms living in the body's intestines create byproducts that cause the body to bloat. Once the bloated body ruptures and outside air filters into the corpse, activity shifts from anaerobic bacteria to infiltrating aerobic bacteria. These aerobic bacteria come primarily from the surrounding soil and the skin of the corpse. By studying the sequencing of the communities of microorganisms, an analyst can more accurately estimate the time of death.

In 2013, a team of researchers led by Jessica Metcalf released the results of a study in which bacterial communities of decomposing mouse bodies were studied over 48 days.[67] Forty mice were euthanized and their corpses placed in identical environments. Then, five each were destructively examined for microbial communities at eight different time points: days 0, 3 (fresh period); days 6, 9, 13 (active decay); and days 20, 34, 48 (advanced decay). Their bodies were sampled, both internally and externally, to collect microbial communities that were then sequenced.

Dr. Metcalf's team concluded that samples of the microbial communities on the skin of the head of the dead mice led to the most reliable measurements of PMI. Her team was able to estimate PMI to within an error of +/- 3.30 days, with a standard deviation of 2.52 days. These results are better than those possible with traditional short-term techniques.

Although Metcalf's findings are promising, the research in this field is far from complete. A similar 2013 study of microbial communities using three swine bodies by Jennifer Pechal[68] also found a direct link between time of death and the sequence of bacterial growth. Another limited study, using two human remains, was published in 2014. This study created a catalog of bacterial communities found on human bodies during the first two weeks of decomposition.[69] Much research is ongoing in this field.[70]

Considering the great potential of this technique and the numerous limitations of the traditional methods of estimating PMI,

[64] Pounder, supra note 64, at 17-18. Pounder notes that a sample group of experts differed by as much as 9.5 hours for a body approximately 40 hours dead; for a body approximately 24 hours dead, their disagreement varied up to 12 hours.

[65] Arpad Vass, The Elusive Universal Post-mortem Interval Formula, 204 Forensic Science Int'l 34 (2011). Vass publishes aerobic and anaerobic rule-of-thumb formulae to help police estimate PMI with greater accuracy, but offers no confidence intervals for the formulae.

[66] J.K. Tomberlin, et.al, A Roadmap for Bridging Basic and Applied Research in Forensic Entomology, The Annual Review of Entomology 403 (2011).

[67] Metcalf, supra note 5, at 1.

[68] Jennifer L. Pechal et al., The Potential Uses of Bacterial Community Succession in Forensics as Described by High Throughput Metagenomic Sequencing, 128 Int'l J. Legal Medicine 193-205 (2014).

[69] Embriette Hyde, The Living Dead: Bacterial Community Structure of a Cadaver at the Onset and End of the Bloat State of Decomposition, Oct. 30, 2013, available at http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0077733 (last visited April 21, 2014).

[70] We are indebted to David O. Carter, PhD., Assistant Dean and Director of Forensic Sciences, Chaminade University of Honolulu, for his assistance and insights into forensic taphonomy.

it seems likely that PMI estimation based on microbial analysis will be the first non-infectious use of phylogenetics in court.

B. Personal Identification Through the Analysis of Skin Bacterial Communities

Although PMI analysis addresses the 'when' question, the 'who' question can also loom large in a criminal case. In the long term, the personal identification of individuals through analysis of bacterial communities living on their skin may be microbial forensics' most powerful tool—the next big thing since DNA. Each of us has a community of bacteria living on our skin.[71] As we move through the world and touch items, surfaces and other people, our bacterial cells become dislodged and trail behind us. Our bacterial communities are unique, with only some 13% of the bacterial kinds (technically known as "phylotype" which stands for the phylogenetic type) shared between any two individuals. Furthermore, these communities are relatively stable over time. While washing your hands will eradicate the current members of that community, the bacteria grow back within hours. Further, while there are changes to the community on your body over time, even after months any natural change in the composition of the microbes on an individual is much less than the difference between individuals. In short, albeit constantly changing, your body's bacterial community appears to be unique and relatively stable.

In a 2010 study, a research team headed by Noah Fierer discovered that identifiable bacterial communities are readily transferred from skin to computer keyboards and mice. Further, even when compared against a pool of more than 270 possible matches, these microbial communities were easily associated with the persons who owned or touched them.[72] In testing the permanence and longevity of the transferred communities, researchers left computer keyboards and mice at room temperature for two weeks before sampling. These tests yielded the same results: the phylogenetic signal on the keyboard was much closer to the owner or user than to any of the 270 controls. These results suggest that bacterial colonies might function as a means of personal identification as viable as fingerprints or possibly even DNA.

The people we live with influence our skin bacterial communities and vice versa. A recent study by Se Jin Song and other researchers[73] showed much greater similarity of skin bacterial flora between cohabiting family members than between random individuals in a community. Over time, bacteria are shared between persons. Some of these bacteria take root, grow, and thrive. These bacteria also spread from and to pets living in a household. Hence, search for a person could potentially be narrowed down by using skin bacterial communities from persons or pets living in her household.

As in the case of the PMI studies based on microbial analysis, these are still the early days in adapting phylogenetic techniques to identify persons. The published studies involve rather small samples in controlled situations. Fierer's study compared three computer keyboards and nine computer mice, each with a single owner, to a universe of 270 other people. Song's research studied only 159 persons living in 60 family units. However, if later studies confirm the prior outcomes, the use of microbes to identify individuals may become a very valuable technique. Criminals do not always leave useable fingerprints on the guns that they fire. Likewise, a crime scene may not yield enough human DNA—careful or lucky perpetrators wear gloves and do not bleed, spit, ejaculate, or otherwise leave trace DNA. However, they will still leave behind a trail of their unique microbes. Even without a direct sample of skin bacteria, samples from cohabiting family members or pets could furnish an important investigative clue that could ultimately lead to the culprit's identification.

However, a cautionary note is warranted. The fictional example in this article's introduction hypothesized a CODIS bacterial database in the year 2018. It is not a foregone conclusion that there will be such a database. It may prove infeasible to construct that database by that date—or ever. The ability to build the database will depend on additional research investigating the consistency of our skin bacterial communities over time and the effect of outside influences on the rate of change. Further, to frustrate their later identification, persons may be able to take measures to change or mask their bacterial skin communities. The promise is evident, but so are the problems.

C. The Identification of the Type of Body Fluid

The Traditional Techniques

In the course of an investigation, law enforcement personnel often must attempt to identify blood, fecal particles, semen, vaginal fluid, and saliva left at a crime scene. Sometimes the quantity of fluids is little more than a stain. Worse still, multiple types of fluids can be mixed together. Conventional techniques using antigens or enzymatic activity are problematic, particularly with very small or older samples.[74] In particular, current techniques have great trouble distinguishing vaginal fluid[75]. In a rape case, proof of the presence of vaginal fluid can be essential for the prosecution.

[71] Fierer, supra note 5, at 6477.

[72] Fierer, supra note 5, at 6478.

[73] Song, supra note 5.

[74] Giampaoli et al., Molecular Identification of Vaginal Fluid by Microbial Signature, Forensic Science Int'l: Genetics 559 (2012)[hereinafter Giampaoli].

[75] Alan Gunn & Sarah Pitt, Microbes as Forensic Indicators, Tropical Biomedicine 320 (2012)[hereinafter Gunn].

Microbial Techniques

The human microbiome offers a unique solution. Each type of body fluid harbors a different flora of bacteria. By determining the genus and species of bacteria present in a sample, investigators can pinpoint reliably which type of body fluid is present. Further, genomic extraction and PCR require only very small samples.

In an experiment reported by Giampaoli in 2012, the researchers succeeded in differentiating types of fluid through bacterial identification in a relatively small trial.[76] The researchers used 47 swab samples from volunteers: 24 vaginal swabs, nine oral swabs and four fecal swabs.[77] In addition, some forensic samples were included as well as single species bacterial controls. The researchers were able to distinguish between the various types of fluids by using bacterial identifiers. While traditional methods have difficulty distinguishing between saliva and vaginal fluid, Giampaoli's methodology successfully separated saliva and vaginal fluid in a mixed sample.[78]

D. Soil Mapping

In a given case, the linchpin issue may be tying the defendant to the site of the actus reus. A common trope of crime dramas is the defendant's muddy boot. The investigators often attempt to develop information about the mud to link it and the defendant to a crime scene. However, the muddy boot can present two very different scenarios. In the first, the police already know the crime scene and endeavor to match the mud on the boot to the soil at the scene. Some commentators refer to this situation as a conformational match. In the second scenario, the police face a more difficult challenge; the crime location is unknown, but the police hope to use the mud on the boot to help them identify the crime scene. Commentators sometimes state that in this situation, the police are endeavoring to develop a random match probability for the mud.

Traditional techniques are sometimes adequate to establish linkages in the first scenario where the police have already identified the crime scene. The traditional techniques analyze the soil's physical properties: color, consistency, mineralogy, geophysics, texture, particle size, and color.[79] A microbial typing technique, similar to Fierer's, could also be used to link microbes in two soil samples to each other. That link could help establish the defendant's presence at the site of the actus reus.

However, the use of microbial analysis may be more useful in the second scenario in which, at the outset of the investigation, the crime scene is unknown. Microbial analysis offers a reversible method of matching soil samples through biogeographic mapping. Soil is embedded with living organisms, including bacteria, nematodes, fungi, and protozoa. A grain of soil contains between 106 to 1010 organisms representing over 50,000 different species.[80] Indeed, soil is so rich in microorganisms that a 1gm soil sample contains between 1 μg and 100 μg of DNA.[81] This genetic information can be extracted and mapped geographically.

Scientists employing microbial analysis have already succeeded in preparing biogeographical maps. Most recently, Griffiths drilled more than 1,000 soil cores to develop a map of the United Kingdom.[82] Griffith cataloged 16S rRNA diversity between the samples. With additional resources, it should be feasible for researchers to create a map with much finer detail.

However, the application of microbial techniques to soil analysis is perhaps the most speculative use discussed in this article. Before microbial soil profiling should be accepted by the courts, the expert should be required to validate methodologies to: (a) determine the variability of bacterial communities from location to location; (b) develop analytical approaches combining discrimination power, robustness, and reliability; and (c) create objective statistical measures for the differences or similarities between samples.[83] Until there is adequate validation of all three steps, microbial analysis of the proverbial muddy boot may not generate admissible evidence.

IV. THE GOVERNING EVIDENTIARY STANDARDS AND THEIR APPLICATION TO THE EMERGING TYPES OF FORENSIC MICROBIAL EVIDENCE

A. The Governing Evidentiary Standards

In 1993, the United States Supreme Court rendered its decision in Daubert v. Merrell Dow Pharmaceuticals, Inc.[84] In Daubert, the Court made two formal holdings. One was a ruling that the enactment of the Federal Rules of Evidence in 1975 had impliedly overturned the prevailing, common-law standard for determining the admissibility of scientific evidence. The previous standard had been the general acceptance test, requiring an expert to base her testimony on theories and techniques that were generally accepted

[76] The bacteria studied were Lactobacillus crispatus, Lactobacillus gasseri, Staphylococcus aureus, Streptococcus salivarius, Streptococcus mutans and species of Enterococcus. Giampaoli, supra note 80, at 561. All of these except the Enterococcus species currently have patents pending for fluid identification. See, e.g., patent application 13/682,071, Quantitation and Profiling of Vaginal Microflora, Sergey Balashov.

[77] Giampaoli, supra note 80, at 560.

[78] Giampaoli, supra note 80, at 562

[79] Jennifer Young, Limitations and Recommendations for Successful DNA Extraction from Forensic Soil Samples, Science and Justice, Feb. 18, 2014.

[80] George Sensabaugh, Microbial Community Profiling for the Characterization of Soil Evidence, Criminal and Environmental Soil Forensics 52 (2009)[hereinafter Sensabaugh].

[81] Sensabaugh, supra note86, at 51.

[82] Robert Griffiths et al., The Bacterial Biogeography of British Soils, Environmental Microbiology 1642 (2011).

[83] Sensabaugh, supra at 86, pages 55-56.

[84] 509 U.S. 579 (1993).

within the relevant scientific fields. The Court construed Federal Rule of Evidence 402 as abolishing uncodified exclusionary rules of evidence. The Court professed that it could not find any statutory language that could reasonably bear the interpretation that it incorporated the general acceptance standard.

However, in its next breath the Court issued a second holding establishing a new standard for the admissibility of scientific evidence. The Court looked to the text of Federal Rule of Evidence 702 that refers to "scientific . . . knowledge." The Court adopted an essentially methodological definition of that expression. The Court held that in order to qualify as reliable, admissible "scientific . . . knowledge," a theory or technique must be supported by adequate empirical validation. To give lower court judges with some guidance, Justice Blackmun provided a list of factors that judges may consider: whether the theory is testable, whether it has been tested, whether the theory has been subjected to peer review, whether the technique has a known or ascertainable error rate, whether there are established standards for using the technique, and whether the theory or technique has garnered general acceptance.[85] However, the justice not only stated that the list was non-exhaustive; the justice also emphasized that the trial judge's inquiry is a "flexible" one.

In 1997, the Court handed down its decision in General Electric Co. v. Joiner.[86] There the Court clarified the nature of the judge's inquiry under Daubert. The Court stated that the trial judge must scrutinize the logical connection between the empirical data cited by the expert and the ultimate conclusion the expert proposes testifying to. The Court commented that if the trial judge finds that there is too great an analytical gap between the data and the conclusion, the judge ought to exclude the expert's conclusion.

The Court completed its trilogy of decisions in 1999 in Kumho Tire Co. v. Carmichael.[87] In some cases, the proponents of expert testimony concluded that their evidence could not pass muster under the rigor of the Daubert standard and attempted to develop a stratagem to evade Daubert. They noted that Rule 702 refers to "scientific, technical, or specialized knowledge." Seizing on the alternative wording of the statute, they argued that Daubert's reliability test applies only to purportedly scientific testimony. The thrust of the argument was that if the proffered testimony was "technical" or "specialized" in nature, the expert's theory need not satisfy the reliability test. The Court rebuffed the argument. The Court opined that it would be difficult to clearly differentiate between "scientific," "technical," and "specialized" knowledge. More fundamentally, the Court asserted that in the statute, all three adjectives modify the same noun, "knowledge." The Court reasoned that Congress' choice of that noun is the source of the reliability requirement; the proposed testimony must constitute more than subjective belief or unsubstantiated conjecture. The Court acknowledged that the factors listed in Daubert were most apropos for scientific testimony and conceded that a trial judge might need to consider other factors to evaluate the reliability of non-scientific expertise. However, the Court declared in no uncertain terms that the requirement for a showing of reliability applies across the board to all types of expert testimony.

Since Daubert is a non-constitutional decision based on the statutory interpretation of the Federal Rules Evidence, the states remain free to reject the Daubert standard. Even the states with statutes modeled after Federal Rules of Evidence have the constitutional authority to construe their statutes differently. However, in the two decades since the Daubert decision, over three-fifths of the states have adopted some variation of Daubert's reliability standard.[88]

B. The Application of the Standards to the Emerging Types of Microbial Analysis

The microbial cloud may yield new forensic tools of great power. It is as if scientists have given us a new set of glasses to peer at the world. But Daubert and its progeny preclude courts from facilely endorsing these new methodologies.

PMI Estimates Based on Microbial Analysis

Full population studies are missing from most of the techniques cited here.[89] For instance, this article discussed three studies investigating the use of microbial analysis to estimate post-mortem interval (PMI). All three studies can be characterized as preliminary in nature. Even the largest study examined the remains of only 40 mice; the other two dealt with three pig bodies and two human cadavers respectively. The experiments were limited in their species of choice, the number of subjects, and relevant extrinsic circumstances such as weather, predators, and the external microbial biome. These studies provide a tentative indication that microbial approaches may eventually improve estimates of the post-mortem interval, but the research is still at an early stage.

[85] The Court demoted general acceptance from the status of a litmus test to that of a mere factor. If a theory or technique has been in circulation long enough to have gained general acceptance, presumably other researchers have examined the underlying research and found it to be satisfactory. Thus, the fact of general acceptance serves as circumstantial evidence of the methodological soundness and adequacy of the underlying empirical data and reasoning.

[86] 552 U.S. 136 (1997).

[87] 526 U.S. 137 (1999).

[88] Paul C. Giannelli, Edward J. Imwinkelried, Andrea Roth & Jane Campbell Moriarty, Scientific Evidence §§ 1.14-15 (5th ed. 2012)[hereinafter 1 Giannelli]. As the 2014 supplement to the treatise notes, since 2012 California and Florida, two states that were formerly leading proponents of the traditional general acceptance test, have adopted Daubert. California adopted the Daubert standard by case law while Florida did so by legislation.

[89] An important exception is phylogenetic analysis.

The Use of Microbial Analysis to Distinguish Types of Body Fluid

Body fluid analysis is similarly in the early stages of development. However, a number of patents have already been filed to take advantage of the process.[90] In the primary Giampaoli study, only 47 fluid samples were taken. Subsequent, larger trials might confirm the validity of the technique, or they could expose weaknesses in the method. For instance, there are suggestions that women of different races exhibit different flora in the vaginal fluid.[91] The indication is merely a "suggestion," because again, the research in this field is in its early development.

Personal Identification Based on Microbial Analysis

Tracking persons from their skin bacterial communities is an exciting prospect, but it is too early to pronounce that the technique definitely satisfies Daubert. The Fierer study on computer keyboards and mice was limited to three owners of keyboards and nine owners of mice, compared to 270 random controls. While the successful matching in the study was undeniably impressive, using the technique to identify criminals from a national population is a huge further step. Joiner's caution against unwarranted extrapolation is pertinent. Daubert listed the known or ascertainable error rate in a relevant consideration, and at this point the technique's susceptibility to false positives has not been quantified. Without further studies justifying the extrapolation to larger populations, the technique might not survive a Daubert hearing.

Using microbial analysis as a method of personal identification poses the same distinction between conformational matches and random match probabilities mentioned in the discussion of soil biogeography. Where there is a suspect and the limited purpose of the analysis is to confirm suspicion aroused by other evidence, a court needs only confirmatory evidence. However, that type of confirmational match using phylogenetic analysis of bacterial communities may represent the outer limit of the current state-of-the-art. The technique could be used in this manner in cases in which DNA, fingerprints, or other traditional identification techniques are unavailing. As in the Schmidt case, the phylogenetic signal would play a limited, albeit significant, role as corroborating evidence.

Developing reliable random match probabilities based on the analysis of skin bacteria is a more daunting challenge. Constructing the sort of national or even statewide skin bacteria database hypothesized in the introduction may not prove possible. While the available studies indicate that a person's skin bacterial community is relatively stable over time, those studies have been limited to a three month window.[92] In the case of human nuclear DNA, a person's entire genome is stable throughout life; for bacterial communities, there is a constant shifting of populations. How long can a database retain utility if there are significant shifts in a person's bacterial community every six months? Authorities would have to frequently reconstitute the database. Furthermore, once the technique is established and the public becomes familiar with the technique, criminals may endeavor to develop forensic counter-measures. Might a criminal be able to intentionally change the makeup of her skin bacterial community? Or, a criminal might try to disguise himself by sampling someone else's skin community and cultivating it on himself before committing a crime. At this point, these are nagging, unanswered questions.

There are substantial concerns about phylogenetic analysis. Complex statistical mathematical manipulations are necessary to link genetic data between bacterial or viral samples. Such sophisticated statistical analysis defines the field of computational phylogenetics. The field uses three primary techniques: neighbor-joining, maximum likelihood, and Bayesian inference. Especially when there is relatively little data available for statistical analysis, these methods may yield different results.[93] Further, the statistical techniques are premised on several basic assumptions, inter alia: each sample has had a not too different rate of evolution, genes have not recombined to an extent that it blurs the relationship, there is no bias in sequence content (convergent evolution or homoplasy), and there was no cross-contamination between the bacterial communities after separation. These assumptions are not always present or reasonable.

To take an extreme example underscoring the difference between human lives and the lives of bacteria, consider horizontal gene transfer. When a bacterial cell breaks apart, the genetic molecules become freely available in the environment. Another cell from the same or related species can take those molecules, transport them to its own genetic molecules, and cut-and-paste a piece of the outsider's genome into its own. Scientists are uncertain about how frequently this kind of transfer takes place,[94] but there is little doubt that such transfers are historically responsible for transmitting genes coding for toxins between pathogenic organisms.

[90] Supra note 82.

[91] Gunn, supra note 81, at 321

[92] Costello, supra note 20, at 1694.

[93] The lack of phylogenetic signal within the NS5B region of the Valencia case is one example of this problem. However, with 322 victims, 44 controls and a hypervariable genome, the E1-E2 region gave data that was confirmed by using several techniques. Similarly, with full access to blood samples of the reported donor and victim in the Schmidt case, computations done with multiple techniques showed very high confidence of close prior ancestry between the viruses.

[94] Jonathan Eisen Ph.D., UC Davis Genome Center, personal communication.

V. CONCLUSION

The microbial cloud is rather like gravity in the 17th century: something that everyone experiences and generally ignores, but that has enormous, sometimes surprising, implications. Since the beginning of time, people knew that objects were drawn toward the ground. In the 17th century, Isaac Newton transformed this common, everyday occurrence into a mathematical force that accurately predicted the motions of the sun, moon, and stars.

Similarly, throughout most of the 20th century, scientists and medical professionals ignored the existence of non-pathogenic microbes. Scientists realized that these communities existed, but they rarely explored the significance of the phenomenon. Indeed, as there are some 10^{30} bacteria on Earth, as compared to some 10^{10} humans, it is not a stretch to state that life on Earth is entirely microbial in nature, with some very few exceptions where cells have clumped together for mutual benefit. All humans, trees and jellyfish now living are three examples of those exceptions.

Microbial forensics could become the most powerful forensic tool of the early 21st century. We now appreciate that our world and our bodies are carpeted and infused with layer upon layer of bacterial communities, each with its own genomic blueprint. By studying the communities' patterns of growth and transfer, we may be able to track human movements in the world.

Sooner or later—probably sooner—litigants will offer testimony based on the techniques of microbial analysis. It seems likely that post-mortem interval (PMI) estimates based on microbial analysis will be the first to surface in court. The traditional PMI techniques were developed in the 19th century and still suffer from major limitations. Microbial techniques to identify the type of body fluid may be the next in line for courtroom use. Inventors have already filed patents and manufactured special equipment to separate body fluid evidence based bacterial content. In the long term, the development of phylogenetic matching of skin bacterial communities promises might dramatically expand law enforcement's ability to identify criminals.

Beginning with Schmidt, many American cases have already accepted the use of microbial analysis to trace viral transmission of disease, but it would be a mistake to leap to the conclusion that the courts will rush to approve all the other possible uses of microbial analysis. In jurisdictions still subscribing to the traditional general acceptance standard, testimony based on novel theories of microbial analysis will certainly face an uphill battle. The new theories may also face intense scrutiny in Daubert jurisdictions. While Daubert no longer invariably requires a showing of general acceptance, general acceptance remains a factor in the trial judge's reliability analysis; and some courts ascribe a good deal of weight to that factor. Moreover, in 2000, reflecting on the Daubert line of authority, the Supreme Court generalized that Daubert and its progeny prescribe "exacting standards of reliability."[95] A trial judge who takes Daubert seriously will demand a strong showing of the methodological soundness and adequacy of the empirical data and reasoning underlying the emerging techniques of microbial analysis.

It is exciting to project that the emergence of microbial analysis may be the most significant forensic development since the advent of DNA typing. However, as the introduction noted, there were missteps in the early legal treatment of DNA typing. At first, the courts almost unquestioningly accepted DNA testimony and the accompanying estimates of random match probabilities.[96] However, from the outset the laboratories used the multiplication or product rule to generate the probabilities, and that formula requires that the probabilities multiplied be independent. At the time, most American laboratories were using the single-locus-probe technology—analyzing sites that were often close together on the same chromosome. The 1992 National Research Council report quite correctly pointed out that when the population frequencies are for sites so close together, there is insufficient assurance of the independence of the probabilities.[97] Thus, there was a grave risk that the random match probabilities overstated the rarity of the DNA profiles. American laboratories almost immediately shifted to multi-locus probes, targeting sites on different chromosomes, to correct the problem. We should ensure that like DNA typing, microbial analysis enhances the accuracy of judicial fact-finding without committing mistakes similar to the early errors in the DNA cases. Microbial analysts can learn valuable lessons from the DNA experience.

APPENDIX: QUALIFYING A MICROBIAL FORENSIC EXPERT

Hypo: The defendant is accused of inflicting grievous bodily harm without intent, for having transmitted HIV to an unprotected sexual partner while knowing this would put the partner at risk of infection, and denying to this partner his HIV infection. In addition to other evidence (testimonies, medical records of both complainant and defendant), the prosecution is attempting to introduce evidence of the complainant's and defendant's blood analysis, from which HIV genetic sequences were recovered. This hypo is chosen as it comes closest to current admissibility in court: it uses well-established scientific methods and is intended for

[95] 528 U.S. 440 (2000).

[96] Giannelli, supra note 7, at § 18.04[c].

[97] Giannelli, supra note 7, at § 18.04[c][3], at 104.

circumstantial evidence only. Characters are Prosecutor (P) and Expert (E).

P: Good morning, Mrs. Expert. I'm the Prosecutor and we're here to discuss the methodologies you employed to come to your conclusions. Have you prepared a report for this court that runs to about 30 pages?

E: Good morning, Mr. Prosecutor, yes that is correct.

P: If the witness could be shown her statement, your Honour. Is that the report you prepared?

E: Yes it is.

P: Before we go to your evidence about this case, I want to ask some questions about your qualifications. What is your current position, where do you work at the moment.

… Here follows a lengthy discussion about the employment position of the expert, her field of expertise, her experience with the techniques used, her experience as a forensic expert, which shows that the expert has a long-standing experience in forensic analysis of HIV transmission cases.

P: Now that we have established your expertise as a forensic expert in HIV transmission cases, would you care to explain to us what kind of tests and analyses you performed and how these are relevant to this court. Let us start with the virus, how does HIV spread and what is a transmission?

E: The most common ways for HIV to spread is sexually, but it can also spread through direct blood contact such as needle sharing in injecting drug users, or from mother to child. In this case, we are talking about sexual transmission through virus that has been passed together with body fluids during sex.

P: So we use the word transmission when one person infects another person. And is it correct that after transmission, the virus in the newly infected person is circulating in the blood and can be recovered from the blood?

E: Yes, this is in fact what we did. We were given two samples of blood of both individuals, the two samples on different dates, and we genotyped the virus in all 4 samples.

P: With genotyping, do you mean you determined the genetic information of the virus?

E: Yes, in our laboratory, we extract RNA from the plasma of the blood, transcribe it into cDNA and then, we amplify the DNA through a polymerase chain reaction, or PCR. This gives us enough material to sequence targeted regions of the genome. We targeted the polymerase gene and the envelope gene, for which there are large public databases of HIV sequences. That is how we genotype the virus in the blood.

P: And this is well established in the scientific community? Peer-reviewed?

E: Yes to both of those questions. The genotyping studies have been done since many years and peer-reviewed researchers agree that the regions we sequenced can be used to investigate transmission of HIV.

P: If I may ask some questions which you have detailed in your statement relating to this genotype, is it possible to make an assessment whether the virus in two individuals are related?

E: Yes, to some extent. With these sequences we are constructing phylogenetic trees, and check whether the viral sequence of both complainant and defendant are clustering together.

P: What is a phylogenetic tree?

E: With a phylogenetic tree, you try to estimate how closely related the viruses are. You can consider it a sort of pedigree, a family tree, the closer you cluster in such a tree, the closer your relationship. Each time a virus replicates, it makes mistakes, and these are called mutations. If two people infected each other, the virus passes between them and at that time they have the same virus. But from that time on, the virus replicates in both of them and accumulates mutations. We have software that counts those mutations, and clusters together viruses with the fewest number of mutations between them.

P: Are you saying that such a phylogenetic tree can tell us who infected whom?

E: No, it is more complicated than that. Since these viruses replicate fast, many mutations are happening in a short time, we are unlikely to find a direct genetic match. So even if the viruses cluster together, you don't know whether they infected each other, or were both infected by a hypothetical third person, or there might have been even more third parties in between them.

P: So, what are you looking for then?

E: What we are looking for is whether the relatedness of the virus between both individuals is closer than between any two infected individuals within the same local epidemic. So is their virus closer related than by chance due to the fact that these individuals live in the same area and might have contracted the virus through other means than sex with each other.

P: So, you can never link both individuals directly, how close can you say their viruses are related?

E: If we have the proper control sequences from other locally infected people, we can say whether or not they belong to a transmission cluster. This means we cannot say they infected each other, but their virus is closer related than compared to other locally circulating viruses. If their virus is less closely related, meaning that they are found in different clusters, then we can say they did not infect each other.

P: And how can you see that on such a tree? Is that the tree there as attachment to your statement?

E: Yes, it is, two trees in fact, we analyzed two genes.

P to Her Honour: Does your Honour have the attachments on your copy of the statement? I just want to ask some questions about those attachments.

Her Honour: I have the attachment referred to. You may refer to the attachments. I must say all this is a bit complicated. If you want the jury to consider the trees, you must make sure it is properly understood.

P to Her Honour: That is my intention your Honour.

P to E: Let me guide you through the tree. So you have all kinds of horizontal lines here, and at the right you have labels.

E: Yes, those labels are the names of the samples. The tree is in fact on its side, so it grows from left to right. The lines are called branches, and they split up in two each time someone infects another person. We call that split a node. If you consider that time is from left to right, the more to the right the split is seen, the more recent such an infection occurred. Compare it with a pedigree, that grows from top to bottom, so your grandparents are at the top, your parents below, you and your siblings and cousins below that. So the lines of the pedigree run from top to bottom, not from left to right like in this appendix, but basically it is the same, you could just turn this appendix on its side and then the lines would grow from top to bottom. Now the split between you and your siblings lies at the level of your parents, while the split between you and your cousins lies at the level of your grandparents, that is higher up the tree. You can say you are closer related to your siblings than to your cousins. The lower the split in a pedigree - in our phylogenetic tree this is the more to the right - the closer the relationship. Everything to the right of a split or node in a phylogenetic tree we call a cluster. If the virus in two individuals forms a cluster with no other viruses in it, then they are more closely related to each other than to the other viruses in the tree, provided the cluster is statistically supported.

P: And if other viruses are in the same cluster?

E: Then they also belong to the same transmission cluster, at least if the cluster does not break up into smaller clusters that are also significantly supported.

P: So a transmission cluster can have viruses from more than two individuals?

E: That is correct, and that is also why you can never prove direct transmission with phylogenetic trees, it can only be used as circumstantial evidence. On the other hand, if the viruses from complainant and defendant belong to different statistically supported clusters, then you can say they did not infect each other, at least not recently.

P: What do you mean not recently?

E: Well, say they infected each other 20 years ago, then each of them could have since then infected other people, with which their virus would cluster closer than with the person they infected 20 years ago.

P: Is that why you say that the samples should be taken as closely as possible to the event?

E: That is correct, with time, so many mutations have accumulated, that it become more and more difficult to identify transmission clusters.

P: How well founded is phylogenetics?

E: Phylogenetics is a well-established science. Evolutionary processes, meaning the accumulation of mutations, have been studied through genetics since at least 1974. It has been used not only in microbiology, but also in the evolutionary history of all animals.

P: And there are no dissenting voices about the use of phylogenetics to assess relatedness?

E. In the scientific community, phylogenetics is well understood and supported. The use of phylogenetic techniques to infer relationships is not disputed. There are lots of publications about this.

P: And these studies are peer reviewed?

E: Yes, they are peer reviewed.

P: Is there a known rate of error with these methods?

E: There are numerous factors that can affect phylogenetics, but those are dependent on how one practices the theory. Such practices should be reviewed on a case-by-case basis. In the case before us, using well-known genetic markers in HIV, the statistical support for identifying clusters has been indicated on the tree presented.

P: And how well is it established that such techniques can be used as a forensic technique for transmission investigations of HIV?

E: It is well established that it can be used in court, but there is some dispute on how best to use such techniques, but there are already some guidelines and I followed those guidelines.

P: Thank you, Mr. Expert.

POSTSCRIPT

This article is largely based on Mr. Steussy's research paper in Professor Imwinkelried's Spring 2014 Scientific Evidence seminar.