


## Science Needs for Microbial Forensics: Initial International Research Priorities


ISBN  
978-0-309-30245-6

252 pages

PAPERBACK (2014)

Committee on Science Needs for Microbial Forensics: Developing an Initial International Roadmap; Board on Life Sciences; Division on Earth and Life Studies; National Research Council

 Add book to cart

 Find similar titles

 Share this PDF



### Visit the National Academies Press online and register for...

- ✓ Instant access to free PDF downloads of titles from the
  - NATIONAL ACADEMY OF SCIENCES
  - NATIONAL ACADEMY OF ENGINEERING
  - INSTITUTE OF MEDICINE
  - NATIONAL RESEARCH COUNCIL
- ✓ 10% off print titles
- ✓ Custom notification of new releases in your field of interest
- ✓ Special offers and discounts

Distribution, posting, or copying of this PDF is strictly prohibited without written permission of the National Academies Press. Unless otherwise indicated, all materials in this PDF are copyrighted by the National Academy of Sciences. Request reprint permission for this book

# Science Needs for Microbial Forensics

## DEVELOPING INITIAL INTERNATIONAL RESEARCH PRIORITIES

Committee on Science Needs for Microbial Forensics:  
Developing an Initial International Roadmap

Board on Life Sciences

Division on Earth and Life Studies

NATIONAL RESEARCH COUNCIL  
*OF THE NATIONAL ACADEMIES*

In cooperation with  
The Croatian Academy of Sciences and Arts  
The U.K. Royal Society and  
The International Union of Microbiological Societies

THE NATIONAL ACADEMIES PRESS  
Washington, D.C.  
**[www.nap.edu](http://www.nap.edu)**

**THE NATIONAL ACADEMIES PRESS 500 Fifth Street, NW Washington, DC 20001**

NOTICE: The project that is the subject of this report was approved by the Governing Board of the National Research Council, whose members are drawn from the councils of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine. The members of the committee responsible for the report were chosen for their special competences and with regard for appropriate balance.

This study was supported by the Naval Postgraduate School's Project on Advanced Systems and Concepts for Countering Weapons of Mass Destruction (PASCC) via Assistance Grant No. N00244-13-1-0026 awarded by the NAVSUP Fleet Logistics Center San Diego (NAVSUP FLC San Diego) to the National Academy of Sciences, by purchase order SAQMMA13M2120 from the U.S. Department of State to the National Academy of Sciences, and by internal support from the National Academy of Sciences. The views expressed in this publication and/or made by speakers, moderators, and presenters do not necessarily reflect the official policies of the Naval Postgraduate School nor does mention of trade names, commercial practices, or organizations imply endorsement by the U.S. Government. Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the author(s) and do not necessarily reflect the views of the organizations or agencies that provided support for the project.

International Standard Book Number-13: 978-0-309-30245-6

International Standard Book Number-10: 0-309-30245-5

Additional copies of this report are available for sale from The National Academies Press, 500 Fifth Street, N.W., Keck 360, Washington, DC 20001; (800) 624-6242 or (202) 334-3313; Internet, <http://www.nap.edu/>.

Copyright 2014 by the National Academy of Sciences. All rights reserved.

Printed in the United States of America

## THE NATIONAL ACADEMIES

### *Advisers to the Nation on Science, Engineering, and Medicine*

The **National Academy of Sciences** is a private, nonprofit, self-perpetuating society of distinguished scholars engaged in scientific and engineering research, dedicated to the furtherance of science and technology and to their use for the general welfare. Upon the authority of the charter granted to it by the Congress in 1863, the Academy has a mandate that requires it to advise the federal government on scientific and technical matters. Dr. Ralph J. Cicerone is president of the National Academy of Sciences.

The **National Academy of Engineering** was established in 1964, under the charter of the National Academy of Sciences, as a parallel organization of outstanding engineers. It is autonomous in its administration and in the selection of its members, sharing with the National Academy of Sciences the responsibility for advising the federal government. The National Academy of Engineering also sponsors engineering programs aimed at meeting national needs, encourages education and research, and recognizes the superior achievements of engineers. Dr. C. D. Mote, Jr., is president of the National Academy of Engineering.

The **Institute of Medicine** was established in 1970 by the National Academy of Sciences to secure the services of eminent members of appropriate professions in the examination of policy matters pertaining to the health of the public. The Institute acts under the responsibility given to the National Academy of Sciences by its congressional charter to be an adviser to the federal government and, upon its own initiative, to identify issues of medical care, research, and education. Dr. Victor J. Dzau is president of the Institute of Medicine.

The **National Research Council** was organized by the National Academy of Sciences in 1916 to associate the broad community of science and technology with the Academy's purposes of furthering knowledge and advising the federal government. Functioning in accordance with general policies determined by the Academy, the Council has become the principal operating agency of both the National Academy of Sciences and the National Academy of Engineering in providing services to the government, the public, and the scientific and engineering communities. The Council is administered jointly by both Academies and the Institute of Medicine. Dr. Ralph J. Cicerone and Dr. C. D. Mote, Jr., are chair and vice chair, respectively, of the National Research Council.

**[www.national-academies.org](http://www.national-academies.org)**



**COMMITTEE ON SCIENCE NEEDS FOR MICROBIAL FORENSICS:  
DEVELOPING AN INITIAL INTERNATIONAL ROADMAP**

*Members*

- JOHN D. CLEMENTS** (*Chair*), Tulane University, New Orleans, Louisiana  
**MUNIRUL ALAM**, International Center for Diarrhoeal Diseases Research, Dhaka, Bangladesh  
**BRUCE BUDOWLE**, Institute of Applied Genetics, University of North Texas Health Science Center, Fort Worth  
**JONGSIK CHUN**, Seoul National University, Seoul, South Korea  
**RITA R. COLWELL**, University of Maryland, College Park  
**NANCY D. CONNELL**, Rutgers New Jersey Medical School, Rutgers University, Newark  
**PAUL KEIM**, Northern Arizona University, Flagstaff  
**JUNCAI MA**, WFCC-MIRCEN World Data Center of Microorganisms (WDCM), Institute of Microbiology, Chinese Academy of Sciences, Beijing, China  
**ALEMKA MARKOTIĆ**, University Hospital for Infectious Diseases, Zagreb, Croatia  
**GEOFFREY SMITH**, Imperial College London, London, United Kingdom

*Staff*

- FRANCES E. SHARPLES**, Study Director and Director, Board on Life Sciences  
**JO L. HUSBANDS**, Scholar/Senior Project Director, Board on Life Sciences  
**CARL ANDERSON**, Program Associate, Board on Life Sciences  
**JENNA OGILVIE**, Senior Program Assistant, Board on Life Sciences  
**BENJAMIN RUSEK**, Program Officer, Committee on International Security and Arms Control  
**KRISTIN WHITE**, Consultant Writer

## BOARD ON LIFE SCIENCES

### *Members*

**JAMES P. COLLINS** (*Chair*), Arizona State University, Tempe  
**ENRIQUETA C. BOND**, Burroughs Wellcome Fund (retired), Marshall, Virginia  
**ROGER D. CONE**, Vanderbilt University Medical Center, Nashville, Tennessee  
**SEAN EDDY**, HHMI Janelia Farm Research Campus, Ashburn, Virginia  
**SARAH C. R. ELGIN**, Washington University, St. Louis, Missouri  
**DAVID R. FRANZ**, Former Cdr USAMRIID; Consultant, Frederick, Maryland  
**LOUIS J. GROSS**, University of Tennessee, Knoxville  
**ELIZABETH HEITMAN**, Vanderbilt University Medical Center, Nashville, Tennessee  
**JOHN G. HILDEBRAND**, University of Arizona, Tucson  
**RICHARD A. JOHNSON**, Arnold & Porter, LLC, Washington, D.C.  
**JUDITH KIMBLE**, University of Wisconsin, Madison  
**CATO T. LAURENCIN**, University of Connecticut Health Center, Farmington  
**ALAN I. LESHNER**, American Association for the Advancement of Science, Washington, D.C.  
**KAREN E. NELSON**, J. Craig Venter Institute, Rockville, Maryland  
**ROBERT M. NEREM**, Georgia Institute of Technology, Atlanta  
**CAMILLE PARMESAN**, University of Texas, Austin  
**ALISON G. POWER**, Cornell University, Ithaca, New York  
**MARGARET RILEY**, University of Massachusetts, Amherst  
**JANIS C. WEEKS**, University of Oregon, Eugene  
**MARY WOOLLEY**, Research!America, Alexandria, Virginia

### *Staff*

**FRANCES E. SHARPLES**, Director  
**JO L. HUSBANDS**, Scholar/Senior Project Director  
**JAY B. LABOV**, Senior Scientist/Program Director for Biology Education  
**KATHERINE W. BOWMAN**, Senior Program Officer  
**INDIA HOOK-BARNARD**, Senior Program Officer  
**MARILEE K. SHELTON-DAVENPORT**, Senior Program Officer  
**KEEGAN SAWYER**, Program Officer  
**BETHELHEM MEKASHA**, Financial Associate  
**ANGELA KOLESNIKOVA**, Administrative Assistant  
**JENNA OGILVIE**, Senior Project Assistant  
**LAUREN SONI**, Senior Project Assistant

## Preface

In September 2001, shortly after the terrorist attacks on the World Trade Center and the Pentagon, someone placed letters containing *Bacillus anthracis* spores into a mailbox in Trenton, New Jersey. Those letters were addressed to several media outlets, including ABC, NBC, CBS, the *New York Post*, and American Media, Inc. in Boca Raton, Florida. The recipients of one of the letters, Mr. Robert Stevens, was the first person to die in what has become known as the anthrax letters attack and designated as the “Amerithrax” investigation by the FBI. At first, diagnosis of Mr. Stevens’ illness was complicated by the absence of any suspicion that would make anthrax part of the differential diagnosis for a 63-year-old man in Boca Raton. Once a diagnosis of anthrax was made, health care workers and epidemiologists began trying to understand how an office worker could come in contact with *B. anthracis*, the causative agent of the disease anthrax. One possible, though unlikely, scenario that was widely touted at the time was that he had come across an infected animal, perhaps a beaver, while on a recent hiking trip. The receipt of an array of letters filled with white powder soon confirmed that this was a bioterrorist attack and not a natural occurrence. In early October, the perpetrator mailed a second set of letters containing a more highly refined preparation of spores, this time addressed to Senators Tom Daschle and Patrick Leahy. Once it was clear that there was an ongoing threat, extensive efforts were begun to identify the source and characteristics of the material used in the attack. What became immediately clear in the midst of heroic efforts to discern the cause and source of the anthrax mailings



was that we were unprepared to analyze the microbial forensic evidence associated with this attack.

At the time, most diagnostic and epidemiological characterizations of infectious diseases were based on illness, culture, serology, physical characteristics, and metabolic profiles of infectious organisms—processes that take time and require pure cultures of viable microorganisms. There was a nascent field of microbial forensics, which had begun in the United States in the preceding decade with the formation of the Federal Bureau of Investigation's Hazardous Materials Response Unit. The Unit was created in part to support suspected or known bioterrorism investigations by providing investigative leads and supporting prosecutions or exonerations with scientific evidence. The Unit was a law enforcement operation designed to employ forensic science principles and practices to produce evidence that would be admissible in court according to U.S. legal requirements and standards. The Amerithrax investigation accelerated the development of the field of microbial forensics, resulting in remarkable development and applications of new techniques and approaches for using laboratory tools to pinpoint the identity of a microbial agent. Microbial forensics became an essential part of the scientific investigation, which was combined with physicochemical analyses as well as other nonscientific types of evidence to narrow the search for the source of the *B. anthracis* used in the attacks.

Today we find ourselves with a complex infrastructure of government agencies, Select Agent registries, regulated research, environmental monitoring in designated cities, federal and state regulations—all resulting from *one* more or less successful biological attack on the United States. The Amerithrax attack with highly refined material produced by a knowledgeable expert (presumably in a U.S. bioweapons laboratory) resulted in 22 illnesses and 5 deaths. Approximately 4 g of material were used in the Amerithrax attack. At that time, the United States planned and prepared as best it could for an attack involving 50 kg of weaponized anthrax spores released on a city with a population of 500,000, anticipating 125,000 casualties. However, it is unlikely that a nonstate entity would be able to produce that quality or quantity of material undetected. Moreover, aside from *B. anthracis*, there are few (if any) biological agents that can be grown in quantity, viably maintained, stabilized, processed to the appropriate size, and delivered in an aerosolizable form except by a few specialized bioweapons facilities and certainly not by terrorists in a garage or cave. Most exotic microorganisms are just too difficult to grow and keep alive, even in the most sophisticated facility. In addition, the technology involved in weaponizing biological materials is complex, demanding, and requires substantial expertise. The more likely scenario is someone having access to a small amount of unrefined material that he/she uses to make

a few individuals ill (causing perhaps a few deaths), the consequence of which will be a nation paralyzed with fear, not illness.

In that context, microbial forensics becomes more important than ever. How does one differentiate a natural outbreak from an accidental release from a legitimate laboratory, or the use of biological material to commit a crime, bioterrorism, and all-out biological warfare? How can this be accomplished quickly enough to inform law enforcement, the intelligence community, policy makers, and the public in a timely fashion? The traditional clinical laboratory sciences of culture identification, serology, etc. are inadequate for these purposes. It was with this background that the workshop in Zagreb, Croatia, was held in October 2013 with the intent to identify the scientific challenges that must be met to improve the capability of microbial forensics to investigate suspected outbreaks and to provide evidence of sufficient quality to support responses, legal proceedings, and the development of government policies. The workshop also was designed to increase awareness of microbial forensics among the members of the larger international scientific communities and to engage these communities in the development of a plan on how to address scientific challenges.

One of the more important concepts discussed during the workshop was that the techniques of microbial forensics could aid not only law enforcement and policy makers, but also public health workers, in trying to identify the existence and source of natural outbreaks. Indeed, as we saw in the Amerithrax attack, the public health system will likely be the first to encounter and the first line of defense against a biological attack. Most infectious diseases develop gradually, with individual patients seeking medical care through their local health care providers. People vary in their susceptibility to infectious diseases, and subtle clues may signal an attack, such as an increase in frequency of a naturally occurring infectious disease, unusual seasonality, unexpected resistance to antimicrobials, or unusual age distribution. These features are likely to be recognized first in the public health arena, and the more common the tools and techniques are between law enforcement and public health, the more likely that a true attack will be identified early, perhaps in time to administer prophylactic antibiotics or vaccines or prevent a second release. Moreover, the further we get from an actual attack, the less inclined policy makers are to provide financial support for continuing the research necessary to prepare for an attack. As noted in the 2013 President's Report for the Global Partnership Against the Spread of Weapons and Materials of Mass Destruction (U.K. Foreign and Commonwealth Office, 2013):

Many of the capabilities required for detecting and responding to the whole spectrum of natural, intentional, and man-made events are essentially the same. Systems and networks that might be created for rare

events will atrophy through lack of use, whereas systems created for addressing natural, man-made and accidental outbreaks of infectious disease are likely to be used frequently. Relying on tools and systems that are compatible with both rare and common occurrences means that in an instance of a rare event, detection and response will not be delayed by lack of familiarity with the tools or systems of reporting.

On behalf of the entire committee, I wish to extend our sincere gratitude to the excellent staff at the National Academies. It is because of the dedication and extraordinary efforts of Fran Sharples, Director of the Board on Life Sciences at the National Academy of Sciences (NAS), that we were able to complete this ambitious task in so short a time. The committee also wishes to thank Jo Husbands from the NAS staff and our colleagues at the Croatian Academy of Sciences and Arts, Jelena Dukic, the Director of International Cooperation, and Ninja Ivanus from her staff, for their outstanding contributions to the design and organization of the workshop. Our colleagues from the U.K. Royal Society and the International Union of Microbiological Societies provided important support and ideas throughout the process. I also want to thank my fellow committee members for their commitment that made the workshop and writing of this report an enjoyable and rewarding opportunity.

John D. Clements, *Chair*

## Acknowledgments

This report has been reviewed in draft form by individuals chosen for their diverse perspectives and technical expertise, in accordance with procedures approved by the National Academies' Report Review Committee. The purpose of this independent review is to provide candid and critical comments that will assist the institution in making its published report as sound as possible and to ensure that the report meets institutional standards for objectivity, evidence, and responsiveness to the study charge. The review comments and draft manuscript remain confidential to protect the integrity of the process.

We wish to thank the following individuals for their review of this report:

Kenneth Berns, University of Florida  
Aaron Darling, University of Technology, Sydney, Australia  
Sean Eddy, HHMI Janelia Farm Research Campus  
Jens Kuhn, Integrated Research Facility, Fort Detrick, National  
Institutes of Health  
Ian Lipkin, Columbia University  
Stephen Morse, Columbia University  
Karen Nelson, J. Craig Venter Institute  
Tom Slezak, Lawrence Livermore National Laboratory  
Tim Trevan, International Council for the Life Sciences  
David Walt, Tufts University

Although the reviewers listed above have provided many constructive comments and suggestions, they were not asked to endorse the committee's conclusions, nor did they see the final draft of the report before its release. The review of this report was overseen by Ronald S. Brookmeyer, University of California, Los Angeles, and Ronald M. Atlas, University of Louisville. Appointed by the National Academies, they were responsible for making certain that an independent examination of this report was carried out in accordance with institutional procedures and that all review comments were carefully considered. Responsibility for the final content of this report rests entirely with the authoring committee and the institution.

# Contents

<b>BOXES, FIGURES, AND TABLES</b>	<b>xv</b>
<b>ACRONYMS AND INITIALISMS</b>	<b>xix</b>
<b>SUMMARY</b>	<b>1</b>
<b>1 INTRODUCTION: WHAT IS MICROBIAL FORENSICS AND WHY IS IT IMPORTANT?</b>	<b>7</b>
<b>2 MICROBIAL SCIENCE: ECOLOGY, DIVERSITY, AND CHARACTERIZING THE MICROBIAL WORLD</b>	<b>35</b>
<b>3 MICROBIAL FORENSICS AND CLINICAL AND PUBLIC HEALTH CONSIDERATIONS: COMMONALITIES AND DIFFERENCES</b>	<b>59</b>
<b>4 CLINICAL AND FORENSIC APPROACHES TO MICROBIAL IDENTIFICATION</b>	<b>79</b>
<b>5 SAMPLING AND PRESERVATION METHODS</b>	<b>97</b>
<b>6 VALIDATION AND REFERENCE MATERIALS FOR MICROBIAL FORENSICS</b>	<b>117</b>

<b>7</b>	<b>BIOINFORMATICS AND DATA</b>	<b>143</b>
<b>8</b>	<b>FINDINGS AND CONCLUSIONS: INITIAL PRIORITIZED SCIENCE NEEDS FOR MICROBIAL FORENSICS</b>	<b>157</b>
	<b>REFERENCES</b>	<b>179</b>
	<b>APPENDIXES</b>	
<b>A</b>	<b>COMMITTEE BIOGRAPHIES</b>	<b>191</b>
<b>B</b>	<b>CONVENING ORGANIZATIONS</b>	<b>199</b>
<b>C</b>	<b>AGENDA</b>	<b>205</b>
<b>D</b>	<b>LIST OF PARTICIPANTS</b>	<b>211</b>
<b>E</b>	<b>LIST OF PRESENTATIONS</b>	<b>215</b>
<b>F</b>	<b>SPEAKER BIOGRAPHIES</b>	<b>219</b>

## Boxes, Figures, and Tables

### BOXES

- S-1 Priority Categories for Microbial Forensics Needs, 4
  
- 1-1 Statement of Task, 8
- 1-2 Questions Forensics Can Help Answer, 13
- 1-3 Forensic Science Methods and System Elements, 16
- 1-4 Murch's Scenario 1—Unknown Outbreak, No Indigenous Capability, 21
- 1-5 Murch's Scenario 2—Viral Outbreak: Bioattack? Genetic Engineering? Multiple Countries Involved, 22
- 1-6 Murch's Scenario 3—Alleged Strategic Attack, 23
- 1-7 Relevant Articles of the Biological Weapons Convention, 25
  
- 3-1 Epidemiological Clues That May Signal a Biological or Chemical Terrorist Attack, 61
- 3-2 Rapid NGS Diagnostic Applications and Public Health Microbiology, 68
  
- 5-1 Sampling and Handling Goals, 100
- 5-2 Sample Matrices for Evidence Collection, 101
- 5-3 Potential Food and Agriculture Targets, 105
- 5-4 Plants Are Vulnerable and Efficacious Targets for Agroterrorism, 106



- 5-5 Sampling Strategy, 107
- 5-6 Comprehensive Plan for Sampling and Collection, 107
  
- 6-1 Validation for the Analysis Stage, 123
- 6-2 Interpretation of Results, 129
- 6-3 Challenges for Any Target: Source Exclusion, Association, and Attribution, 130
- 6-4 Microbial Forensics Questions: Genetics, 131
- 6-5 A Bioinformatics Toolbox for Microbial Forensics, 133
- 6-6 Inference and Error Issues, 134
- 6-7 Gaps in Standard Reference Materials (SRMs), 136
- 6-8 Databases and Resources, 137
- 6-9 Databases and Resources: The World Data Centre for Microorganisms (WDCM), 140
- 6-10 Proposed Global Catalogue for Microbial Forensics, 142

## FIGURES

- 1-1 Spectrum of risks due to biological agents, 10
- 1-2 The forensic continuum, 15
- 1-3 Time-risk continuum, 33
  
- 2-1 Variants of representative colonies of each of the morphotypes of *B. anthracis* Ames, 42
- 2-2 Worldwide distribution of *B. anthracis* clonal lineages, 47
- 2-3 Three classifications of bacterial pathogens, 49
- 2-4 Distribution of *Y. pestis*, 50
- 2-5 Red lineages represent HIV sequences from the suspected perpetrator, 53
  
- 3-1 The microbial forensic process, 62
- 3-2 Events timeline of enterohemorrhagic *E. coli* O104:H4 outbreak, 65
- 3-3 A ranking of the discriminatory power of microbial typing approaches, 69
- 3-4 Four approaches to using whole-genome sequencing, 71
- 3-5 Differences from the Index Case in the NS5B region, 75
- 3-6 Maximum likelihood tree for cloned sequences in the E1-E2 region, 77
  
- 6-1 Process of implementing a test for diagnostic use, 124
- 6-2 Performance characteristics, error types, and measurement metrics, 126

- 6-3 Sample analysis flowchart, 126
- 7-1 Evolutionary phylogenetic tree of *V. cholerae* based on SNP analysis of 26 genomes, 148

**TABLE**

- 8-1 Priority Categories for Microbial Forensics Needs, 175



## List of Acronyms and Initialisms

AAM	American Academy of Microbiology
ABC	American Broadcasting Company
ABC	Analyzer of Bio-resource Citations
ACD	Advisory Committee to the Director
AMD	advanced molecular detection
ANI	Average Nucleotide Identity
ASM	American Society for Microbiology
ASPR	Assistant Secretary for Preparedness and Response
ATCC	American Type Culture Collection
ATM	atomic force microscopy
AWS	Amazon Web Services
BEAST	Bayesian evolutionary analysis by sampling trees
BfR	Bundesinstitut für Risikobewertung (Federal Institute of Risk Assessment, Germany)
BGI	Beijing Genomics Institute (People's Republic of China)
BLAST	Basic Local Alignment Search Tool
BSL	biosafety level
BWC	Biological Weapons Convention
canSNP	canonical single-nucleotide polymorphism
CAP	College of American Pathologists

CAP	Certified Authorization Professional
CBRND	chemical, biological, radiological, and nuclear defense
CBS	Columbia Broadcasting System (now operating as CBS Broadcasting, Inc.)
CCINFO	World Directory of Culture Collections
CDC	U.S. Centers for Disease Control and Prevention
CEO	chief executive officer
CLIA	Clinical Laboratory Improvement Amendments
CODATA	Committee on Data for Science and Technology
CODIS	Combined DNA Index System
COMCOFs	committees, commissions, and federations
CPU	central processing unit
DGA	French Ministry of Defense
DHHS	U.S. Department of Health and Human Services
DNA	deoxyribonucleic acid
dNTP	deoxynucleotide triphosphates
DOE	U.S. Department of Energy
DTRA	Defense Threat Reduction Agency (part of the U.S. Department of Defense)
EAEC	enteroaggregative <i>Escherichia coli</i>
EBI	European Bioinformatics Institute
ECL	electrochemiluminescence
EDA	Economic Development Administration
EDC	European Centre for Disease Prevention and Control (Sweden)
EHEC	enterohemorrhagic <i>Escherichia coli</i>
ELISA	enzyme-linked immunosorbent assay
EM	electron microscopy
ENCODE	Encyclopedia of DNA Elements
env	HIV retrovirus envelope gene
EPA	U.S. Environmental Protection Agency
ESI	electrospray ionization
EU	European Union
FAO	U.N. Food and Agriculture Organization
FAZD	National Center for Foreign Animal and Zoonotic Disease Defense
FBI	U.S. Federal Bureau of Investigation
FCA	Fellow of the Croatian Academy
FDA	U.S. Food and Drug Administration

FEMA	U.S. Federal Emergency Management Agency
FOI	Swedish Defense Research Agency
GAO	U.S. Government Accountability Office
GCM	Global Catalogue of Microorganisms
GHDP	genomically highly diversified pathogen
GIDP	genomically intermediately diversified pathogen
GISAID	Global Initiative on Sharing Avian Influenza Data
GLEWS	Global Early Warning System for Major Animal Diseases, Including Zoonoses
GMMP	genomically monomorphic pathogen
GOARN	Global Alert and Response Network
GPHIN	Global Public Health Intelligence Network
GPU	graphics processing unit
HAZMAT	hazardous materials and items
HCV	hepatitis C virus
HFRS	hemorrhagic fever with renal syndrome
HHMI	Howard Hughes Medical Institute
HIV-1	human immunodeficiency virus 1
HMM	Hidden Markov Model
HMRU	Hazardous Materials Response Unit
HPA	Health Protection Agency (United Kingdom)
HPS	hantavirus pulmonary syndrome
HSPD	Homeland Security Presidential Directive
HUS	hemolytic uremic syndrome
HVAC	heating, ventilation, and air conditioning
HVR	hypervariable region
IAEA	International Atomic Energy Agency
IBD-BIOM	Inflammatory Bowel Disease Biomarkers Program
iBOL	International Barcode of Life Project
ICDDR	International Center for Diarrhoeal Diseases Research, Bangladesh
ICFMH-WPCM	Working Party on Culture Media of the International Committee on Food Microbiology and Hygiene
ICRC	International Committee for the Red Cross
ICSU	International Council of Science
IDA	Institute for Defense Analyses
IHR	International Health Regulations
IMCAS	Institute of Microbiology, Chinese Academy of Sciences

IMG	Integrated Microbial Genomes
INTERPOL	International Criminal Police Organization
IOM	Institute of Medicine
ISABS	International Society of Applied Biological Sciences
ISO	International Organization for Standardization
IT	information technology
IUBS	International Union of Biological Sciences
IUMS	International Union of Microbiological Societies
IV	intravenous
JAMA	<i>Journal of the American Medical Association</i>
JGI	Joint Genome Institute (U.S. Department of Energy)
LPSN	List of Prokaryotic Names with Standing in Nomenclature
LRN	Laboratory Response Network
LT	Life Technologies Group
MALDI-TOF	matrix-assisted laser desorption/ionization time-of-flight
Mb	megabyte
MDS	minimum datasets
MERS	Middle East Respiratory Syndrome
MGIT	mycobacterial growth indicator tube
MG-RAST	Metagenomics Rapid Annotations using Subsystems Technology
MiSeq	benchtop NGS instrument manufactured by Illumina
MLST	multilocus sequence typing
MLST+	core-genome MLST
MLVA	multilocus VNTR analysis
MPS	massively parallel sequencing
MRCA	most recent common ancestor
MRSA	methicillin-resistant <i>Staphylococcus aureus</i>
NAS	U.S. National Academy of Sciences
NATO	North Atlantic Treaty Organization
NAU	Northern Arizona University
NBACC	National Biodefense Analysis and Countermeasures Center
NBAS	National Biosurveillance Advisory Committee
NBC	National Broadcasting Company

NBFAC	National BioForensic Analysis Center
NBTCC	National Biological Threat Characterization Center
NCBI	National Center for Biotechnology Information
NeCTAR	National eResearch Tools and Resources
NGO	nongovernmental organization
NGS	next-generation gene sequencing
NIAID	National Institute of Allergy and Infectious Diseases, NIH
NIH	National Institutes of Health
NIST	National Institute of Standards and Technology
NRC	National Research Council
NSTC	National Science and Technology Council
OIE	World Organization for Animal Health
OPCW	Organization for the Prohibition of Chemical Weapons
PATH	Program for Appropriate Technology in Health
PC	personal computer
PCR	polymerase chain reaction
PFGE	pulsed-field gel electrophoresis
PGM	Ion Torrent Personal Genome Machine™
PHEIC	Public Health Emergency of International Concern
PI	principal investigator
PIXE	particle-induced X-ray emission
PLOS	Public Library of Science
PNAS	<i>Proceedings of the National Academy of Sciences</i>
Pol	HIV retrovirus polymerase gene
PPE	personal protective equipment
PS	presumed source
PUUV	Puumala virus
QA/QC	quality assurance and quality control
qPCR	quantitative PCR
RAPD	random amplified polymorphic DNA
RDS	recommended datasets
RFLP	restriction fragment length polymorphism
RIPL	Rare and Imported Pathogens Laboratory
RIVM	Dutch National Institute for Public Health and the Environment
RKI	Robert Koch Institute (Germany)
rMLST	ribosomal MLST
RNA	ribonucleic acid



RNJMS	Rutgers New Jersey Medical School
rRNA	ribosomal RNA
R&D	research and development
SARS	Severe Acute Respiratory Syndrome
SEM	scanning electron microscope
SIMS	secondary ion mass spectroscopy
SMRT	single-molecule real time
SNP	single-nucleotide polymorphism
SOAP	Short Oligonucleotide Analysis Package
SOP	standard operating protocols (or procedures)
SRM	Standard Reference Materials
SSD	solid-state disk
S&T	science and technology
TB	terabyte
TEM	transmission electron microscope
TGen	Translational Genomics Research Institute
TIGR	The Institute for Genomic Research
TTP	thrombotic thrombocytopenic purpura
UHID	University Hospital for Infectious Diseases
U.K.	United Kingdom
UKM	University Hospital Muenster (Germany)
U.N.	United Nations
UNESCO	United Nations Educational, Scientific and Cultural Organization
UNICRI	U.N. Interregional Crime and Justice Research Institute
UNIDIR	U.N. Institute for Disarmament Research
U.S.	United States of America
USAMRIID	U.S. Army Medical Research Institute of Infectious Diseases
USDA	U.S. Department of Agriculture
VNTR	variable number tandem repeat
WDCM	World Data Center of Microorganisms
WFCC-MIRCEN	World Federation for Culture Collections— Microbial Resources Centers Network
WGS	whole-genome sequencing
WHO	World Health Organization
WIPO	World Intellectual Property Organization
WMD	weapons of mass destruction

# Summary

One of the fundamental components of any investigation of a suspected accidental release of a dangerous pathogen or the alleged hostile use of biological agents, whether by states or nonstate actors, will be scientific analysis to support efforts for attribution. Science will not offer definitive solutions in such scenarios, but it often plays a special role in supporting other aspects of an investigation. The United States and a number of other governments and international and regional organizations are actively working to identify and support the research that is needed to build those capabilities, which are known generally as “microbial forensics.”

Microbial forensics deals with the genetic and other materials associated with microorganisms. Contrary to the images from popular media, this emerging discipline is still in the early stages of development and faces substantial scientific challenges to provide a robust suite of technologies for identifying the source of a biological threat agent and attributing a bioterror act to a particular person or group. The unlawful use of biological threat agents poses substantial dangers to individuals, public health, the environment, the economies of nations, and global peace. It is likely that scientific, political, and media-based controversy will surround any investigation of the alleged use of a biological agent. For these reasons, building awareness of and capacity in microbial forensics can assist in our understanding of what may have occurred during a bioterror event and facilitate international collaborations that engage the broader scientific and policy-making communities that are likely to strengthen our micro-

bial forensics capabilities. One goal is to create a shared technical understanding of the possibilities—and limitations—of the scientific bases for microbial forensics analyses. Another is to identify the range of scientific needs to continue the field's development.

Microbial forensics has additional specialized needs because of the demand for “evidence” and “proof” in the context of law enforcement or international policy. Many of these relate as much to the quality of the *process* by which material is collected and analyzed as to the science and technology employed. The public health needs highlighted in this report reflect the global reality that in most countries the capabilities and activities relevant to microbial forensics occur only in the context of public health. Particularly the United States has a microbial forensics science community that is truly differentiated from public health, and a limited number of other countries have some basic microbial forensics infrastructure. For most of the rest of the world, microbial forensics is a side activity of public health officials. Since most disease outbreaks will first be recognized through the public health infrastructure, strengthening detection and diagnostic capacities there serves both public health and law enforcement.

With these needs and realities in mind, a group of national and international scientific organizations undertook a collaboration whose centerpiece was a workshop held in fall 2013 in Zagreb, Croatia. As mentioned above, the specific goals of the workshop (and the larger collaboration) were to

- Foster collaboration within the international scientific community to support technical understanding and enhanced research on microbial forensics, and
- Develop the beginnings of an international roadmap for how to do the necessary science, including priorities among potential topics.

The partner organizations were the U.S. National Academy of Sciences (NAS), the Croatian Academy of Sciences and Arts, the International Union of Microbiological Societies, and the U.K.'s Royal Society. The Croatian Academy hosted the workshop at its headquarters in Zagreb. Participants included 59 experts from 21 countries and several international organizations, spanning a range of researchers and clinicians from numerous scientific and technological disciplines related or applicable to microbial forensics as well as technical experts and policy makers with a strong interest in the contributions of science and technology to security. The two-and-a-half-day meeting was the primary evidence-gathering mechanism for an ad hoc committee with substantial international membership

appointed by the NAS's National Research Council (NRC). The NRC committee supplemented the information obtained during the workshop with other evidence to prepare a final consensus report, which offers a series of conclusions for advancing the field. Funding for the project was provided by the U.S. Navy, the U.S. Department of State, and the NAS.

The committee used a generous definition of "science," including research to improve fundamental scientific understanding of microbes; specialized research intended for particular applications in public health, law enforcement, or elsewhere; and an array of technologies and methods that are enabling dramatic advances in both basic and applied research. The committee also identified important procedural and policy needs, such as common understandings and protocols for taking and managing samples within and across nations. Meeting these additional needs is essential to ensure that the scientific and technical developments the committee calls for are to make microbial forensics a more effective tool for investigating suspicious disease outbreaks.

The committee prepared this final consensus report, which draws on the workshop presentations and discussions as well as on additional information obtained from other experts and the literature to reach conclusions for moving forward. The report is not meant to provide a detailed roadmap for the international development of microbial forensics, but rather elucidates the major issues highlighted at the workshop that the committee believes need to be addressed for the global development of the science of microbial forensics. For the purposes of this project, the committee chose breadth over depth.<sup>1</sup> The committee also gave particular attention to those areas, such as increased scientific knowledge about microbial communities and common standards and protocols for analysis, which would benefit from international cooperation and collaboration. The number and variety of issues meant that not all issues could be examined in sufficient detail to develop a true roadmap.

The statement of task for this project calls for the committee to "develop the beginnings of an international roadmap for how to do the necessary science, including priorities among potential topics." The list of needs identified in Box S-1 is long, but the successful development of microbial forensics will require addressing all of them. There are considerable differences in how difficult it will be to achieve the priorities and in whether there are already existing national or international efforts, for example, in basic research or public health or by industry, which

---

<sup>1</sup> Basic information about microbial forensics is available in a number of textbooks (see, e.g., Budowle et al., 2011, and Primorac and Schanfield, 2014). Various national and international advisory bodies have also addressed many of the issues covered in the report (see, e.g., the reports of the U.S. National Biosurveillance Advisory Subcommittee, 2009, 2011).

**BOX S-1**  
**Priority Categories for Microbial Forensics Needs**

**Challenging Tasks and/or Long Lead Times**

- An international collaboration engaging the worldwide scientific community in a systematic effort to identify, monitor, and characterize a far higher proportion of global microbial species to increase knowledge about endemism and background. The effort should begin with known pathogens and then expand to their close relatives as well as emerging pathogens. (Conclusion 1, Basic Science)
- Development of high-confidence methods to distinguish among natural, accidental, and deliberate outbreaks of infectious diseases. (Conclusion 2, Basic Science)
- Increased emphasis on development and validation of processes (sample collection, preservation, handling, storage, packaging, and transportation) and analytical methods for microbial forensics, including establishing standards for most components. (Conclusion 9, Validation and Standards)
- Discussions under the auspices of an international body that has the respect of the international political and scientific communities about how to share microbial forensic data, and for developing and presenting cogent arguments that can be persuasive to political leaders and scientists worldwide. (Conclusion 12, Bioinformatics and Data/Data Sharing)
- An international effort to design and establish more systematic and comprehensive reference collections and databases for pathogens and other microorganisms. This effort could take advantage of existing models, such as the World Data Centre for Microorganisms and the American Type Culture Collection. A model system for a consortium of reference collections and data storage centers could be created and later scaled up to become more inclusive. (Conclusion 13, Bioinformatics and Data/Databases and Reference Collections)

**Ongoing Efforts on Which to Build**

- Increased emphasis on research to determine mechanisms of pathogenicity, including virulence factors and host immune responses. (Conclusion 5, Needs Common to Medicine, Public Health, and Microbial Forensics)

can be drawn upon to help achieve them. Many scientific and technical communities will be able to make contributions. These communities should also be able to take advantage of other initiatives, for example, to advance disease surveillance and diagnostics for public health purposes, particularly if there is a conscious effort to foster communication across the many relevant disciplines and technical fields. There are a number of high-priority needs that are particularly challenging tasks with long lead times to achieving real progress so that efforts should begin or expand soon. The latter needs will require substantial and sustained support

- Priority research to realize the promise of metagenomics and its application to microbial forensics and the development of the forensic value of the other “omics”: proteomics, metabolomics, transcriptomics, glycomics, immunogenomics, etc. (Conclusion 3, Basic Science)
- Greatly improved global disease monitoring and surveillance in humans, animals, and plants to facilitate rapid response and better disease control. (Conclusion 6, Needs Common to Medicine, Public Health, and Microbial Forensics)
- Improved worldwide access to molecular diagnostics (polymerase chain reaction, whole-genome sequencing, etc.), including refinement and distribution of benchtop next generation gene sequencing instruments that are fast and affordable and have simple workflow procedures. (Conclusion 4, Needs Common to Medicine, Public Health, and Microbial Forensics)
- High priority placed on continued research and development to improve physical science applications to microbial forensics. (Conclusion 8, Methods and Technologies)
- Refinement of bioinformatics and statistical methods for evaluating evidence in microbial forensics, including new algorithms that scale to very large or complex datasets. (Conclusion 11, Bioinformatics and Data)

#### **Shorter Lead Times or Industry Incentives**

- Development of more advanced, faster, and cheaper assay and sequencing technologies that can be standardized and made more accessible to benefit both microbial forensics and public health. (Conclusion 7, Methods and Technologies)
- A compilation of all protocols in use (e.g., for sampling, DNA extraction and isolation, sequencing, etc.) and whether and how they have been validated. (Conclusion 10, Validation and Standards)
- Expansion of technically based training to “professionalize” microbial forensics and increase the number of qualified practitioners worldwide by engaging international professional organizations or other entities that have experience providing training in related fields. (Conclusion 14, Training and Education)

from governments as major funders of the research, development, and implementation that will be essential for achieving success.

Box S-1 presents the needs identified in this report organized according to the key features discussed above:

- One set of needs represents tasks, for example, the need to identify and characterize a significantly increased number of microbial species that are particularly challenging and/or require a long lead time to achieve the desired results. Such efforts will require the involvement of governments to provide the research

resources to carry them out over many years and should be given priority by participating institutions.

- The second set represents needs that could take advantage of ongoing efforts to advance the development of microbial forensics, but will require deliberate communication efforts and, in some cases, funding to ensure that microbial forensics applications are actually included and implemented.
- The third set of needs has the advantage of either a relatively short lead time to make substantial progress or the existence of significant markets that will provide incentives for industry to produce what is required. For example, the production of faster and cheaper instruments for diagnostics for medicine and genomic analyses for microbial forensics will probably be conducted by industry, which is always seeking to put improved devices on the market.

The committee recognizes that there is overlap among the categories and that some of the needs would fit within more than one of them. It nevertheless believes that exercises like this can be helpful in thinking about implementation issues and for the development of a more detailed roadmap to guide future efforts.

# 1

## Introduction: What Is Microbial Forensics and Why Is It Important?

Many people have some notion about forensics from their exposure to the many “police procedural” TV shows being broadcast in countries around the world. Popular entertainment has made many familiar with the concept of using human DNA to identify criminals, although the TV version of this practice gives what is probably a less than realistic impression of the complexity of analyses and degree of certainty surrounding DNA evidence. Nevertheless, human DNA typing is a widely accepted means of identifying and convicting perpetrators, exonerating the innocent, and identifying missing persons from such tragedies as mass disasters.

“Microbial forensics” has been defined as “a scientific discipline dedicated to analyzing evidence from a bioterrorism act, biocrime, or inadvertent microorganism/toxin release for attribution purposes” (Budowle et al., 2003). This emerging discipline is still in the early stages of development and faces substantial scientific challenges to provide a robust suite of technologies for identifying the source of a biological threat agent and attributing a biothreat act to a particular person or group. The unlawful use of biological agents poses substantial dangers to individuals, public health, the environment, the economies of nations, and global peace. It also is likely that scientific, political, and media-based controversy will surround any investigation of the alleged use of a biological agent, and can be expected to affect significantly the role that scientific information or evidence can play. For these reasons, building awareness of and capacity in microbial forensics can assist in our understanding of what may



have occurred during a biothreat event, and international collaborations that engage the broader scientific and policy-making communities are likely to strengthen our microbial forensics capabilities. One goal would be to create a shared technical understanding of the possibilities—and limitations—of the scientific bases for microbial forensics analysis.

Toward this end, a group of national and international organizations held a workshop from October 14 to October 16, 2013, in Zagreb, Croatia. The workshop was organized by the U.S. National Academy of Sciences (NAS), the Croatian Academy of Sciences and Arts, the International Union of Microbiological Societies, and the U.K.'s Royal Society. The Croatian Academy of Sciences and Arts hosted the workshop at their headquarters in Zagreb. A planning meeting between NAS and the Royal Society in June 2011 began the process of designing the project with the intent of the Zagreb workshop to (1) identify the scientific challenges that should be met to improve the capability of microbial forensics to differentiate among natural outbreaks, unintentional release, biocrimes, or bioterror attacks; and (2) provide evidence of sufficient quality to support legal proceedings and the development of government policies. The workshop also was designed to increase awareness of microbial forensics among the members of the larger international scientific communities and to engage these communities in the development of a plan on how to address scientific challenges. An international ad hoc committee was appointed by the NAS's National Research Council to organize the meeting; brief committee member biographies can be found in Appendix A. The committee's statement of task is in Box 1-1.

Fifty-nine expert participants from 21 countries took part in the workshop. Participants included researchers and clinicians from numerous scientific and technological disciplines related or applicable to microbial

#### **BOX 1-1** **Statement of Task**

An ad hoc committee with substantial international membership will plan an international symposium and prepare a consensus report of findings and conclusions to address the science needs underlying the development of microbial forensics. The results of the symposium, supported by additional information and data gathering by the committee, are intended to

- Foster collaboration within the international scientific community to support technical understanding and enhanced research on microbial forensics, and
- Develop the beginnings of an international roadmap for how to do the necessary science, including priorities among potential topics.

forensics as well as technical experts and policy makers with a strong interest in the contribution of science and technology (S&T) to security. The two-and-a-half-day meeting included plenary sessions featuring talks by scientific experts and policy makers who reviewed current and developing practices and technologies in, or applicable to, microbial forensics, with an emphasis on defining the challenges confronting the field. Smaller breakout groups enabled more targeted discussion of perceived gaps in microbial forensic capacity, as well as identification and prioritization of possible ways to develop and strengthen the field and its applications.

Information about the convening organizations may be found in Appendix B. Support for the workshop and report was provided by the U.S. Navy through the Naval Postgraduate School, the U.S. Department of State, and the NAS itself. The workshop agenda may be found in Appendix C and a list of participants in Appendix D. A list of all the presentations and biographies for the speakers may be found in Appendixes E and F, respectively. John Clements, the chair of the organizing committee, asked workshop participants to address what tools are needed to enable us to successfully confront a bioterrorism or biocrime event, an unintentional release of a harmful agent, or a public health crisis caused by microorganisms or toxins. In all cases, the goal is the same—to protect the health and well-being of the public. He posed the following questions: What will help move microbial forensics forward in supporting such protection and how should these needs be prioritized?

The committee prepared this final consensus report, which draws on the workshop presentations and discussions as well as on additional information obtained from other experts and the literature to reach conclusions for moving forward. The report is not meant to, nor can it, provide a detailed roadmap for the international development of microbial forensics, but rather elucidates the major issues highlighted at the workshop that the committee believes need to be addressed for the global development of the science of microbial forensics. For the purposes of this project, the committee chose breadth over depth.<sup>1</sup> The committee also gave particular attention to those areas, such as increased scientific knowledge about microbial communities and common standards and protocols for analysis, which would benefit from international cooperation and collaboration. The number and variety of issues meant that not all issues could be examined in sufficient detail to develop a true roadmap.

---

<sup>1</sup> Basic information about microbial forensics is available in a number of textbooks (see, e.g., Budowle et al., 2011, and Primorac and Schanfield, 2014). Various national and international advisory bodies have also addressed many of the issues covered in the report (see, e.g., the reports of the U.S. National Biosurveillance Advisory Subcommittee, 2009, 2011).

### WHAT IS BIOTERRORISM?

Seth Carus (2001) published a now classic compilation of cases from 1900 to 2000 of the illicit use of biological agents by criminals and terrorists. His definition of “bioterrorism” is that it is “assumed to involve the threat or use of biological agents by individuals or groups motivated by political, religious, ecological, or other ideological objectives.” He clearly noted, however, that “most individuals and groups who have used biological agents had traditional criminal motives.” He believes that it is, therefore, essential to separate the clearly criminal perpetrators from those with political agendas, whether the motive is sectarian, religious, or ecological. The available evidence, in fact, suggests that the vast majority of cases involve criminal motives. Such motives include extortion, revenge, a desire to terrorize particular victims to make them worry about their health, and murder.

Both biocrimes and bioterrorism exist on a continuum of risk associated with biological agents (Figure 1-1). The other end of the spectrum deals with natural outbreaks or accidental releases of infectious disease agents. In all cases on the spectrum, public health protection requires that the microorganism first be identified and its source located to stop further cases of exposure. To this degree, medicine, public health, and law

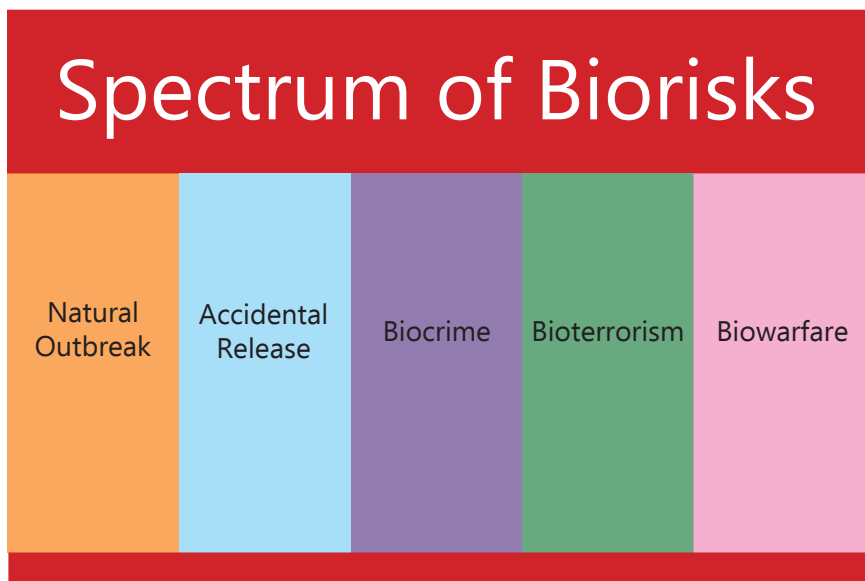


FIGURE 1-1 Spectrum of risks due to biological agents.

enforcement initially have common aims and methods. Microbial forensics, however, has requirements and needs that, in many ways, go beyond those of medicine and public health. Though often applicable to medicine and public health, the methods used in microbial forensics delve deeper into identification of organisms, require standardization and validation, and must meet legal standards for evidence. Hence these topics are the subjects of much of the detailed discussion in this report, while keeping in mind the commonalities with public health that provide opportunities to leverage methods and information across fields.

### THE EMERGING FIELD OF MICROBIAL FORENSICS

Forensic science involves the application of science to the investigation of legal and policy matters. Science may not offer definitive answers in all cases, but often plays a special investigative role. “Science and technology are used to serve as independent ‘witnesses’ in criminal or civil matters, intelligence, and policy” (Murch presentation, 2013). The law enforcement goal is “attribution”—that is, determining who committed the offense. Based on the analysis of biological and other evidence, law enforcement builds a case for attribution to a specific source or sources. The evidence supporting attribution must be robust and suitable for use in legal proceedings and to inform decision making at the highest levels.

Microbial forensics seeks to produce reliable conclusions quickly to protect public health and with sufficient validity and quality to serve law enforcement and policy purposes. In microbial forensics, law enforcement may partner with scientists from microbiology, genetics, public health, agriculture, and many other disciplines to identify and characterize pathogens, or their toxins, implicated in biological events.

Dr. Randall Murch proffered five important questions to frame the background on microbial forensics:

1. Why is there an important need for microbial forensics?
2. What is the current state of the art?
3. How do the forensics used for criminal investigations differ from epidemiological investigations for public health?
4. What are the major research challenges for the field?
5. How can basic science be used to solve the current challenges for microbial forensics and how might this help in other areas, such as public health?

Murch explained that microbial forensics in the United States began in the 1990s with the formation of the Federal Bureau of Investigation’s (FBI’s) Hazardous Materials Response Unit (HMRU). The Unit was cre-

ated to support suspected or known bioterrorism investigations by providing investigative leads and supporting prosecutions or exonerations with scientific evidence. The FBI unit initially drew upon legacy science being developed or performed by scientists at national laboratories and universities across the United States and employed forensic science principles and practices to try to produce evidence that would be admissible in court according to U.S. legal requirements and standards. The FBI also recognized the importance of collaborating with the public health community. At the time of preparations for the 1996 Summer Olympic Games in Atlanta, a collaboration with the U.S. Centers for Disease Control and Prevention (CDC) was established that is still ongoing. The CDC provides, for example, training materials on forensic epidemiology.<sup>2</sup>

Although microbial forensics incorporates basic research to develop techniques or methodologies, the questions asked, processes engaged in, expectations for, and outcomes sought may be different or more demanding than for basic research. The science in microbial forensics, like all science, must (1) be properly validated and accepted by peers and stakeholders (scientific, legal, policy making) before being used, but must also (2) demonstrate that the information generated can answer key investigative and legal questions. Microbial forensics borrows, transitions, and develops science from related disciplines for its own purposes. Meeting these challenges will increase the value of microbial forensics by providing “leaps” in value, confidence, and timeliness for critical decision making and will advance other related disciplines.

In the event of a suspected biological attack, leaders would have questions about the identity and source of the biological threat that intelligence, public health, law enforcement, and forensics must try to answer (see Box 1-2A). Policy leaders want correct answers—and quickly—so they can act appropriately. Forensic science can help answer these questions, and it is essential that the answers be reliable. At the same time, policy leaders have to assess not just whether information is objectively credible and defensible, but also whether it will be accepted by key intended audiences. This can be a major issue at the national level, but becomes a much larger consideration when dealing with international relations, security, and terrorism issues (see next section). Both “elites” and “publics” may reject information because of distrust of the source and other emotive factors. When using the results of science, policy makers (and international negotiators) must take these factors into consideration, too, not just the objective science. It is beyond the scope of this report to examine these political factors in any detail, but they have to be taken into account in developing and applying microbial forensic capabilities.

---

<sup>2</sup> <http://www.cdc.gov/phlp/publications/forensicepidemiology/index.html>.

### BOX 1-2 Questions Forensics Can Help Answer

#### A. A Leader's Perspective: Key and Time-Sensitive Questions

- What is or was it?
- Who did it?
- Did it come from a program or lab we know?
- Will there be more attacks?
- What are we—and nations involved—doing about it?
- What can we know, by when, and with what confidence?

#### B. Investigative, Legal, Intelligence and Policy Perspective Questions

- Did a crime (event of interest) occur?
- What happened?
- How, where, when, and why did it occur?
- Who was involved and responsible?
- What evidence exists? What does it tell us? How strong are the links?
- How reliable and credible is the evidence?
- What alternative explanations are there for the evidence?
- Can we defend our conclusions and actions?



SOURCE: Adapted from Murch (2010).

However, a somewhat varied set of questions could arise from the investigative/legal and intelligence/policy perspectives, as listed in Box 1-2B. These parties will share the same questions as national leaders, but they also can have many additional questions, particularly regarding the accuracy, reliability, validity, credibility, and defensibility of scientific evidence. They may also have information about the likelihood of evidence being accepted and of ways in which different actors might attempt to manipulate information for political reasons. Forensics can contribute to answers, but the evidence is rarely of a clear-cut “smoking gun” type. While legal frameworks exist for evidence, policy frameworks are being developed.

In the United States, evidence in a criminal case must align with the standards of proof of the criminal justice system in order to be admissible in legal proceedings, and standards must be met at every step of the investigation. Investigative leads, for example, must be based on validated, verifiable evidence, in order to enable authorities to gather fur-

ther evidence and/or compel suspected parties to cooperate in a variety of ways. In the United States and many other countries, the “standard of proof in a criminal trial is ‘beyond a reasonable doubt,’ which means the overall evidence must be so strong that there is no reasonable doubt that the defendant committed the crime.”<sup>3</sup> Reasonable doubt, of course, depends on the jurors’ (or the policy makers’) assessment of all the evidence presented during a trial.

Murch defined “scientific attribution” as the assignment of a sample of questioned origin to a source, or sources, of known origin, to the highest possible degree of scientific certainty—while excluding origination from other sources (Murch, 2010). Ideal examples are fingerprint and DNA analyses, which can provide a high degree of scientific certainty that evidence came from one source to the exclusion of all others. However, according to Murch, microbial forensics cannot yet claim that degree of certainty and in many cases may never reach such specificity. In addition, the standard for attribution required in court may be different from the standard required to drive a policy decision.

The process of forensic investigation begins with gathering intelligence and information to lay a foundation for and justify an investigation. This initial information gathering is followed by a time-driven, multidisciplinary, multisource investigation to pursue a rule-in and rule-out process. Crime scene investigation includes evidence identification, collection, preservation, and transport, as well as presumptive testing, which seeks to prove that either (a) the sample probably is a certain substance or (b) the sample is definitely not a certain substance. Laboratory analyses provide deeper characterization and comparison of questioned source and known source samples. Throughout the process, investigators must place the interpretations of analyses and the conclusions drawn into a context, transitioning them to meet the needs of both real-time and end users. Murch, a former FBI official, stated that investigators also should provide alternative interpretations of analyses. Throughout the investigation, forensics is assisting in building and shaping decisions and actions. The process is iterative, aids meeting the burden of proof, is part of the build toward prosecution, and supports exoneration. Finally, all the evidence and information generated are channeled to communication and decision making. Unbiased results, conclusions, and explanations about an event, as well as alternative explanations, must be provided to the legal and policy decision systems and to other stakeholders.

Science plays a role in every phase of a forensic investigation, not only in the establishment of guilt or innocence. For example, science helps to

---

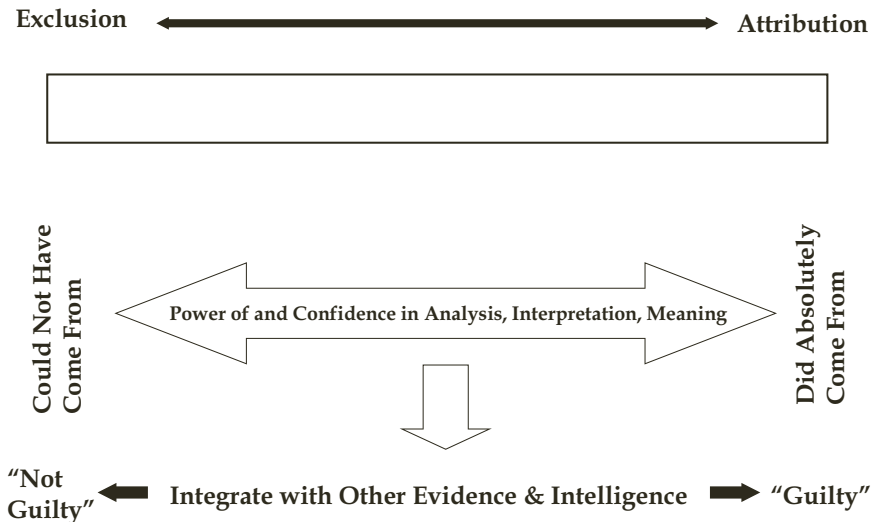
<sup>3</sup> See, for example, the United States Courts website, <http://www.uscourts.gov/FederalCourts/UnderstandingtheFederalCourts/HowCourtsWork/CriminalCases.aspx>.

generate leads, and did so repeatedly during the investigation of the 2001 anthrax letters case. Ultimately investigators will weigh the combined forensic evidence to exclude or attribute to a source (Figure 1-2) to support a finding of guilt or innocence.

Because the information generated by forensic methods is to be used by law enforcement in litigation, the goals of these methods differ from those of traditional research. In addition to the accuracy, reliability, and repeatability demanded in traditional research, forensic methodologies are subject to other rigorous requirements. There are admissibility requirements for “new science” in the U.S. legal process, and similar requirements are being developed for the policy process.

The ideal forensic science methodology would incorporate the goals listed in Box 1-3A. Achieving these goals would help to ensure that samples are collected and handled appropriately to preserve the target evidence as well as possible; analyses and comparisons of known-source

## ***“The Forensic Continuum” Individual Sample Types and Compilation***



**FIGURE 1-2** The forensic continuum represents the evaluation and analysis of a microbial sample to determine its exclusion or attribution value. These results will be integrated with other evidence and information to determine innocence or guilt.

SOURCE: Budowle et al. (2013).



**BOX 1-3**  
**Forensic Science Methods and System Elements**

**A. Forensic Methods**

- Robust collection and preservation of evidence
- Relevant exploitation of samples
- High discrimination
- Enables comparison of “K” and “Q” samples
- Utility across known, encountered sample types
- Accuracy
- Reliability
- Defined and acceptable error rate
- Speed and responsiveness
- Repeatability
- Transferability<sup>a</sup>
- Validity can be independently established
- Probative results are interpretable, explainable, defensible

**B. Forensic Science System**

- Capabilities are matched to submitter and stakeholder needs and requests
- Appropriate deployable assets, and facilities with sufficient resources for operation and maintenance
- Field, transport, and laboratory evidence integrity and security
- Properly credentialed, trained and certified personnel
- Full suite of equipment, fully validated methods, matched to samples received and questions asked
- Comprehensive quality assurance and control program
- Appropriate repositories and database
- Appropriate resources: research, development, test and evaluation, validation and technology transfer

<sup>a</sup> *Transferability means that the method can be performed in other laboratories. It is similar to the concept of reproducibility and important for reanalysis or some form of verification.*

SOURCE: Murch (2008, 2010).

(K) and questioned-source (Q) samples are performed with applicable discrimination and resolution; analyses are reliable and repeatable, with clearly defined error rates or defined limitations; and the process yields interpretable, probative results that can be communicated and supported. The ideal forensic science system would comprise the elements in Box 1-3B and enable the marshaling of appropriate scientific methods, tools, equipment, infrastructure, and personnel to meet the needs of the submitter and stakeholders. Such a system would help ensure that quality control and quality assurance can be maintained through the stages of field assessment and analysis, evidence collection and preservation, lab analysis to characterize an unknown or questioned sample and/or

compare it with a sample from a known source, interpretation and conclusions, and reporting and communication.

The founders of the HMRU recognized that building a microbial forensics discipline would require aggregating a broad range of disciplines (e.g., epidemiology, genomics,<sup>4</sup> metagenomics,<sup>5</sup> and other “omics” disciplines, biostatistics and population genetics,<sup>6</sup> analytical chemistry and biochemistry, microscopy, bacteriology, mycology, virology, clinical medicine for infectious diseases, veterinary medicine, plant pathology, food science, ecology, materials science, process engineering, physical sciences, and bioinformatics and computational science). Murch emphasized that many of these disciplines also are fundamental to public health and medical science. As noted above, the need for a strong union and dynamic collaboration between law enforcement and public health to investigate possible bioattacks from event outset to post-event was recognized early in the development of microbial forensics. Public health, infectious disease medicine, and law enforcement investigations all need to establish whether an event is deliberate, accidental, or natural. Importantly, each can leverage the other’s resources to achieve the same initial objectives. The major difference between the two approaches is that the public health investigation’s goal is to manage the public health response and protect the public’s health and safety, whereas law enforcement’s is to provide safety and security by apprehending and convicting those who committed the attack.

Microbial forensics seeks to answer these questions:

1. *What is the threat agent?* Usually establishing this has not been difficult, although it may not occur in an optimal time frame.
2. *Is it probative or relevant?* Establishing certainty here is more difficult. Scientists may be working with trace quantities, for example, or analysis may require an understanding of the sample background to understand the source.
3. *Can it be linked to a source?* Establishing this demands understanding the power of methods used to discriminate and characterize with acceptable confidence limits.
4. *What are the meaning and weight of the conclusion?*

The goal of the microbial forensics process is to use microbial analy-

---

<sup>4</sup> The science of studying DNA sequences and properties of entire genomes (Alberts et al., 2002).

<sup>5</sup> Metagenomics is the study of a collection of genetic material (genomes) from a mixed community of organisms (<http://ghr.nlm.nih.gov/glossary=metagenomics>).

<sup>6</sup> The study of genetic composition in populations and how natural selection and other factors produce changes in genetic composition.

ses and other evidence to fix a questioned source to a position on a continuum that ranges from “could *not* have originated from” to “*consistent*<sup>7</sup> with having originated from” to “absolutely *did* originate from” a known source. Again, identification is simpler than attribution. Exclusion, association, and attribution are dependent on several key factors, with more value and weight given to attribution derived when more possible sources can be eliminated. Uncertainty and confidence must be stated, either qualitatively or quantitatively.

The magnitude of the microbial forensics workspace is vast. Unlike situations in which a human is a source of the biological evidence (one species, two genders), microbial forensics deals with myriad organisms, including viruses, bacteria, fungi, parasites, and the toxins some of these organisms produce. The vectors by which infectious diseases are spread and the reservoirs in which they reside might also be of importance. Biological agents are not limited to those appearing in the “threat lists” developed by various countries (e.g., the CDC’s “A, B, and C” categories of bioterrorism agents).<sup>8</sup> Nor are humans the only possible target—agricultural plants and animals could be as well. Moreover, threat agents can be bioengineered, and a number of infectious agents could be misused based on the motives, means, resources, and objectives of the perpetrator.

As Murch noted, in light of all this, we should be asking ourselves if the forensic techniques we need to confront this diverse range of potential biological threats are in place today and if they can meet the expectations of all levels of stakeholders?

Much of the work of microbial forensics today is based on studying biodiversity, phylogenetics,<sup>9</sup> phylogeography,<sup>10</sup> genomics, and developing methods with greater sensitivity of detection and level of detail, extraction methodologies, and collection strategies. Science is moving toward faster throughput, cleverer bioinformatics,<sup>11</sup> and other methods. The United States and some other countries have invested heavily in genome

---

<sup>7</sup> Note that although an explanation may be consistent with a set of circumstances, other explanations are not necessarily excluded.

<sup>8</sup> Available at <http://www.bt.cdc.gov/agent/agentlist-category.asp>.

<sup>9</sup> “The study of evolutionary relatedness among various groups of organisms through molecular sequencing data and morphological data matrices” (Herndon, 2012)

<sup>10</sup> “A field of study concerned with the principles and processes governing the geographic distributions of genealogical lineages, especially those within and among closely related species” (Avice, 2000).

<sup>11</sup> “The use of information technology, such as computer programs, to analyze, store, and manage biological data. A common bioinformatics activity is predicting protein products from DNA sequences” (National Institute of Allergy and Infectious Diseases, 2009).

sequencing<sup>12</sup>—pursuing different methods and technologies, including the development of compact benchtop units—and bioinformatics.

Microbial forensics also investigates whether an agent has been genetically manipulated or chemically treated to make it more virulent or dispersible or to mask its characteristics. The agent may have been handled crudely or with great sophistication. Analysis of processing elements, such as methods of growth, separation, washing, drying, grinding, and the use of additives can help further characterize the production process and source of a biological material and are usually determined through physical and chemical analyses that can employ instrumentation such as mass spectrometry.<sup>13</sup>

Traditional forensic evidence (e.g., fingerprints, trace evidence, digital, materials) also is still an important part of attribution that should not be ignored. Microbial forensics requires that scientific investigators safely and properly address the probative classic evidence while also studying the biological agent as evidence.

There are many event scenarios for which we are ill prepared to respond effectively or investigate using microbial forensics. These include introducing a highly aggressive “new strain” of influenza virus during flu season, various scenarios introducing biological threats into agricultural animal populations or crops, and attacks employing pathogens that have been engineered to suggest they originate from a source other than the actual source. If a perpetrator uses nature to his/her advantage, some cases may never be resolved.

In some instances, forensic science and microbial forensics might make only limited contributions for a variety of reasons. The future of microbial forensics lies in bridging those gaps that diminish the capacity. For example, Murch identified needs for better bioinformatics, faster

---

<sup>12</sup>“Determination of the order of nucleotides (base sequences) in a DNA or RNA molecule” (Larkin, 2001). A nucleotide is “a subunit of DNA or RNA consisting of a nitrogenous base (adenine, guanine, thymine, or cytosine in DNA; adenine, guanine, uracil, or cytosine in RNA), a phosphate, and a sugar (deoxyribose in DNA and ribose in RNA). Thousands of nucleotides are linked to form a DNA or RNA molecule” ([http://www.laskerfoundation.com/news/weis/g\\_dictionary.html](http://www.laskerfoundation.com/news/weis/g_dictionary.html)). A base pair comprises “two nitrogenous bases (adenine and thymine or guanine and cytosine) held together by weak bonds. Two strands of DNA are held together in the shape of a double helix by the bonds between base pairs” ([http://www.laskerfoundation.com/news/weis/g\\_dictionary.html](http://www.laskerfoundation.com/news/weis/g_dictionary.html)).

<sup>13</sup> “Method involving specialized instruments for measuring the mass and abundance of molecules in a mixture and identifying mixture components by mass and charge” (U.S. Department of Energy, 2012).

throughput gene sequencing, and an increased focus on biosurveillance, endemism, metagenomics, proteomics,<sup>14</sup> and the other “omics.”

The next section addresses the additional challenges for microbial forensics in an international setting where more than one country is involved.

## INTERNATIONAL DIMENSIONS

The challenges for microbial forensics are difficult enough within the law enforcement context of a single country. Adding elements that span national boundaries poses substantial additional challenges. In his presentation to the workshop, Murch outlined several potential biological-threat scenarios that he has developed to illustrate some of the issues likely to be encountered. Box 1-4 describes a hypothetical scenario of an outbreak of unknown cause in a country with no indigenous investigative microbial forensics capability, and little communication occurs between law enforcement and public health. Although other countries may be able to help, they have limited microbial forensics capabilities to perform the necessary analyses. Meanwhile, the terrorist group that has claimed responsibility leverages confusion by threatening more attacks.

Box 1-5 outlines another hypothetical scenario in which a cruise ship has an outbreak of what initially may be considered to be norovirus, and there is a 25 percent case fatality among affected passengers. Nearby countries have limited expertise and capability in microbial forensics. The World Health Organization (WHO) tentatively identifies the virus. There is evidence the outbreak may be nefarious, but there is uncertainty about conclusions that may be drawn from the analyses (e.g., is it genetically engineered or not?). Media and concerned governments are clamoring. Countries whose citizens are among the affected passengers, including two countries with a strained political relationship, are conducting independent investigations, but no physical evidence other than patient samples has been collected.

For both scenarios, either sharing technologies, protocols, and methodologies a priori or developing Memoranda of Understanding with other countries that possess the required analytical capabilities would allow the affected countries to be better prepared. Sharing of results and a more coordinated approach would provide great benefits and leverage resources.

---

<sup>14</sup> Large-scale analysis of the “proteome” (collection of proteins expressed by a cell at a particular time and under specific conditions) to identify which proteins are expressed by an organism under certain conditions. Proteomics provides insights into protein function, modification, regulation, and interaction (U.S. Department of Energy, 2012).

**BOX 1-4**  
**Murch's Scenario 1—Unknown Outbreak,  
No Indigenous Capability**

- A country reports a suspicious outbreak of an infectious disease in villages populated by an ethnic minority; hundreds are affected.
- Disease with these symptoms has not been reported previously in this country; the public health agency has no experience with it.
- Assistance is requested of the World Health Organization. The causative agent is tentatively identified after 2 weeks.
- A known regional terrorist group claims responsibility for the attacks, but the claim is unverified.
- The country's national police have (1) no expertise or experience with bioterrorism; (2) a limited relationship with the public health agency; and (3) no capability to collect, transport safely, or analyze hazardous evidence.
- No mutual-assistance agreements exist with neighboring countries or countries with necessary capabilities.
- The investigation is at a standstill; it cannot be determined if the outbreak is natural or deliberate.
- Meanwhile, seeing an advantage, the terrorist group claims it is preparing for more attacks.
- Nations experienced in dealing with bioterrorism (including forensic investigation) offer assistance but are not confident they have methods in place to do appropriate analysis for attribution.

SOURCE: Murch presentation, 2013.

The third hypothetical scenario, described in Box 1-6, is the kind of event that continually preoccupies the very highest levels of the U.S. and other governments. Outbreaks of a zoonotic infectious disease occur in the United States and an allied country; a third country has threatened both nations. Microbial forensics attribution efforts appear to rule out other countries with which the United States has tense relationships and to rule in the threatening country, but this attribution cannot be confirmed. The United States accuses the suspect country, and seeks U.N. Security Council support for military action, making clear that the U.S. government is prepared to act unilaterally if necessary to protect its national security and that of its ally. The accused country demands that the United States provide "evidence" to support its charges and provides its own experts to rebut the accusations. Other countries are weighing in on both sides. This scenario raises the question: If attacked with a bioweapon by another country, what quantity and quality of evidence (scientific and other) are

**BOX 1-5**  
**Murch's Scenario 2—Viral Outbreak: Bioattack?**  
**Genetic Engineering? Multiple Countries Involved**

- A popular cruise line reports that a ship that has visited several ports and is now at sea in the eastern Mediterranean is experiencing an aggressive outbreak, perhaps of a foodborne illness. The ship is ordered to anchor off the closest port and await instructions.
- 4 days on, >300 passengers from >15 countries have been affected, with a case fatality of 25 percent. The ship medical staff is overwhelmed but has provided what medical assistance they can.
- Three crewmembers are missing.
- Passengers have been calling relatives and friends to report an attack on the ship with a biological weapon. These people in turn have called the media and their governments, demanding an investigation.
- The cruise line does not have medical resources to move onto the ship; it does not know how to handle the crisis.
- Several national police agencies in the region have no expertise or support for microbial forensics; some have minimal capabilities. None have trained or exercised together, though several public health agencies routinely have collaborated. Two key governments involved have strained political relationships.
- Two weeks on, after a third party gives assistance with capabilities in microbial forensic investigation and public health crisis management, the infectious agent is identified as a probable new strain of norovirus. It appears it may have been engineered; this has not been corroborated.
- Sixty more passengers and crew have died.
- Individual countries with affected citizens are conducting separate investigations; coordination at all levels is inconsistent and uneven. Other than patient samples, no physical evidence has been collected. Gaps in the science are noted. No significant leads or persons of interest have been identified.
- Several countries involved have reached out to the United States and European countries to explore whether additional assistance can be rendered.

SOURCE: Murch presentation, 2013.

needed in order to gain international support for action, how likely is that evidence to be accepted, and how will it affect an ultimate political decision to take action against the perpetrator if that support is not forthcoming? It is essential that such a decision be based on solid evidence whenever possible.

**BOX 1-6**  
**Murch's Scenario 3—Alleged Strategic Attack**

- Tensions are higher than normal in a region owing to a series of bellicose threats and actions by one country against the countries in the region, some of whom are U.S. allies. The U.S. has been threatened also.
- Unusual zoonotic infectious disease outbreaks occur in one of the countries allied with the U.S., as well as in the U.S. itself. Disease spreads rapidly, with high case fatality in infected persons and domesticated animals.
- Public health and agriculture systems are overwhelmed. The public in both countries begins to panic, and politicians voice grave concern.
- Unusual properties of the viral causative agent isolated raise the suspicion of a bioattack.
- Forensic (including microbial) evidence seems to “rule out” as sources other countries with which the U.S. has tense relationships, and to “rule in” the country at issue—although attribution cannot be achieved. U.S. intelligence and other sources “indicate credible links” to the country of concern.
- U.S. leadership publicly accuses the country of concern, and seeks U.N. Security Council support for military action, making clear that is prepared to act unilaterally if necessary to protect its national security and that of its ally.
- The accused country demands the U.S. provide “evidence” to support its charges and provides its own experts to rebut the accusations. Other countries are weighing in on both sides.

SOURCE: Adapted from Murch presentation, 2013.

**INTERNATIONAL LEGAL FRAMEWORKS  
AND INVESTIGATIVE CAPACITIES**

If any of these three scenarios were to occur, global leaders would seek answers from intelligence, public health, law enforcement, and forensics. The answers would then be sought within a particular set of international frameworks and capabilities. The primary international legal frameworks concerned with the use of biological agents as weapons are the 1925 Geneva Protocol, which prohibits the use of biological and chemical weapons, and the 1972 Biological Weapons Convention (BWC), which is the first international disarmament treaty to ban an entire class of weapons.<sup>15</sup> Since these agreements primarily address state-level pro-

<sup>15</sup> The full name of the Geneva Protocol is the Protocol for the Prohibition of the Use in War of Asphyxiating, Poisonous or Other Gases, and of Bacteriological Methods of Warfare; the BWC's full name is the Convention on the Prohibition of the Development, Production and Stockpiling of Bacteriological (Biological) and Toxin Weapons and on Their Destruction. The



grams and actions, in 2004 the UN Security Council adopted Resolution 1540, which puts binding obligations on all UN members to address the risks posed by nonstate actors.<sup>16</sup> None of these agreements has its own investigative or enforcement capacities.

Under the BWC, an allegation of use of biological weapons would be handled through the UN Security Council. The treaty also calls on States Parties to provide assistance to states that have been determined by the Security Council to have suffered an attack (see Box 1-7 for the relevant articles). In the event of an allegation, the U.N. Secretary General has the authority to launch an investigation; the formal title for the process is the Secretary-General's Mechanism for Investigation of Alleged Use of Chemical and Biological Weapons. This is the process that was used to investigate allegations that chemical weapons had been used in the conflict in Syria in 2013 with support from the Organization for the Prohibition of Chemical Weapons (OPCW) and the World Health Organization (WHO). Once Syria joined the Chemical Weapons Convention in October 2013, the subsequent elimination of its chemical weapons program has been verified by the OPCW, which implements the treaty.

Depending on the scenario, efforts to investigate a case of alleged use would face substantial practical challenges. There could be issues of access to sites, availability and conditions of samples, analytical capacity on the ground, concealment of evidence or deliberate efforts to mislead investigators, and concerns for their safety (Casagrande presentation, 2013). Depending on the mandate for an investigation, there would almost certainly be questions related to the sensitivity of information and the willingness and capacity of various actors to share, and at present no international agreement or standard governs what will or will not be shared in a given set of circumstances. The successful use of microbial forensics in such cases involves difficult policy as well as technical issues. Although the BWC does not have a formal investigative capacity, Dr. Piers Millet from the BWC Implementation Support Unit argued that the treaty is a key international platform for addressing challenges in microbial forensics, on both the technical and policy levels. His presentation focused on drivers for enabling and directing advances in microbial forensic capabilities.

---

Chemical Weapons Convention (Convention on the Prohibition of the Development, Production, Stockpiling and Use of Chemical Weapons and on their Destruction) also covers toxins.

<sup>16</sup> U.N. Security Council Resolution 1540 addresses all weapons of mass destruction and obliges U.N. member states "to refrain from supporting by any means non-State actors from developing, acquiring, manufacturing, possessing, transporting, transferring or using nuclear, chemical or biological weapons and their delivery systems" (UN 1540 Committee website, <http://www.un.org/en/sc/1540/#&panel1-16>; accessed December 14, 2013).

**BOX 1-7**  
**Relevant Articles of the Biological Weapons Convention**

**Article V**

The States Parties to this Convention undertake to consult one another and to cooperate in solving any problems which may arise in relation to the objective of, or in the application of the provisions of, the Convention. Consultation and Cooperation pursuant to this article may also be undertaken through appropriate international procedures within the framework of the United Nations and in accordance with its Charter.

**Article VI**

(1) Any State Party to this convention which finds that any other State Party is acting in breach of obligations deriving from the provisions of the Convention may lodge a complaint with the Security Council of the United Nations. Such a complaint should include all possible evidence confirming its validity, as well as a request for its consideration by the Security Council.

(2) Each State Party to this Convention undertakes to cooperate in carrying out any investigation which the Security Council may initiate, in accordance with the provisions of the Charter of the United Nations, on the basis of the complaint received by the Council. The Security Council shall inform the States Parties to the Convention of the results of the investigation.

**Article VII**

Each State Party to this Convention undertakes to provide or support assistance, in accordance with the United Nations Charter, to any Party to the Convention which so requests, if the Security Council decides that such Party has been exposed to danger as a result of violation of the Convention.

SOURCE: Text of the Biological Weapons Convention, signed April 10, 1972.

With 170 participating states, the consensus achieved in the BWC is meaningful. Since 2002 the BWC has held a series of annual intersessional meetings between the treaty review conferences held every 5 years. The two annual meetings of experts and states parties provide both an expert and a diplomatic component, with different issues addressed each year along with some standing topics. The BWC 2004 annual meetings, for example, examined mechanisms in place in the event a biological/toxic weapon was used or an allegation of use was made. In the final report of the meetings, states parties agreed on the value of (1) continuing to develop national capacities to respond to, investigate, and mitigate potential use of bioweapons; and (2) of doing so in cooperation with relevant regional and international organizations (Biological Weapons Conven-

tion, 2004). Although not explicitly defined, relevant organizations could include regional and international scientific bodies. Moreover, the BWC agreed to encourage and assist other states to pursue these activities. Consensus exists that bioweapon defense should be a global enterprise.

The BWC revisited this topic in 2010. Additions to previous agreements were made that stressed the importance of

1. Effective efforts, regardless of whether the outbreak is natural or deliberate. Diseases and toxins that harm humans, animals, plants, and the environment fall within the purview.
2. Recognition that capabilities to quickly and effectively detect, respond to, and recover from an alleged use of a biological/toxic weapon must be in place before they are used (United Nations, 2011).
3. Investigating and mitigating the potential impact of an alleged use of biological/toxic weapons in accordance with national laws and regulations (e.g., for data handling).
4. Coordinating a state-government approach to emergency management since multiple arms of a government would be involved.
5. Addressing the full range of possible implications of meeting these expectations (Biological Weapons Convention, 2010).

Reviewing advances in S&T was a main focus of the 7th BWC review conference in 2011. The work program for 2012 to 2015 that resulted from the meeting comprises three pillars that are addressed each year—one of which is the examination of relevant advances in S&T (Biological Weapons Convention, 2011). Included within its purview are the potential for misuse of S&T (e.g., to make biological weapons) and beneficial applications (e.g., investigating and mitigating bioweapon attacks), as well as interaction between the security and scientific communities.

International interest in ensuring that the S&T reviews focus heavily on potential benefits—including advances in microbial forensics—continues, and the BWC now has an international platform to address such issues. Moreover, mechanisms to respond to the use of bioweapons will be a topic at the next two biannual meetings.<sup>17</sup>

Dr. Dana Perkins from the U.N. 1540 Committee added that microbial forensics is an essential element of a national and international biosecurity infrastructure, both as a deterrent and as a support tool. Similar to nuclear forensics, microbial forensics may be used to detect, prevent, and deter

---

<sup>17</sup> The language in the report is “How to strengthen implementation of Article VII, including consideration of detailed procedures and mechanisms for the provision of assistance and cooperation by States Parties” (Biological Weapons Convention, 2011:21).

acts of bioterrorism and illicit trafficking or use of biological materials. The potential applications of microbial forensics thus may contribute to strengthening biosecurity in the context of Resolution 1540 and to achieving cooperation and synergy among various international security frameworks.

Perkins stated that, in addition to its role in supporting investigations of alleged use, developing and improving microbial forensics methods to detect illicit trafficking and biological materials outside regulatory control, and to prevent and respond to biosecurity events, will strengthen the implementation and enforcement of Resolution 1540. She pointed out for nuclear forensics, the International Atomic Energy Agency (IAEA) has a leading role in facilitating the exchange of information and international collaboration and providing assistance to support law enforcement and assessment of nuclear security vulnerabilities. In contrast, microbial forensics lacks that level of international leadership. She suggested, however, that there is potential for much more widespread and effective cooperation, not only among countries but also among organizations such as WHO, the World Organization for Animal Health (OIE), the Food and Agriculture Organization (FAO) of the United Nations, the Chemical Weapons Convention/Organization for the Prohibition of Chemical Weapons, and the BWC/BWC Implementation Support Unit. (The contributions of WHO, OIE, and FAO are discussed further in the “Biosurveillance” section of Chapter 2.) Another interesting point can be made about the International Atomic Energy Agency, which has a major role in the development and oversight of the peaceful applications of nuclear energy, as well as nuclear weapons security. This suggests that policy mechanisms for mitigating rare events, such as bioterrorism, should be coupled when possible to more routine-use drivers so that the applicable technologies are tried, tested, and familiar.

A sustained effort will be required to (1) build communities of microbial forensics specialists, (2) formulate projects to develop microbial forensics S&T foundations, and (3) raise awareness of possible synergies among its different applications. Perkins believes that these plans should not be confined to investigating the use and alleged use of weapons but should also focus on contributing to prevention and deterrence of weapon proliferation and terrorist threats. Identifying major technical and policy challenges that these plans could support is the subject of the next section and the remainder of this report.

## **AN OVERVIEW OF KEY QUESTIONS AND UNMET NEEDS FOR MICROBIAL FORENSICS**

Significant scientific challenges or needs have long been recognized in microbial forensics, and have persisted with little or no progress toward

resolution. Murch pointed out that raising the bar in science improves our capabilities of innovation and use and also raises the bar against adversaries. In addition, many challenges in microbial forensics are shared by other disciplines,<sup>18</sup> and so, bridging these gaps could provide leaps in “operationally useful” capabilities and knowledge for more than just microbial forensics. One way to meet the needs would be to enable integration of fundamental science, advancing technologies, and forensic applications with those of related fields for multivariate benefits. The benefit would be to push the value of and expectations for microbial forensic investigation/attribution capabilities to the limits of science, technology, and knowledge. This will require recognizing pragmatic priorities—for example, iterative evolutions of technologies, platforms, and methods—that must be designed and implemented to resolve chokepoints and barriers that impede progress and success. In addition, we should identify what the “big leaps” are and invest in them. Advancing national and global biosecurity also requires fostering international, cross-disciplinary collaboration in ways not yet recognized. The benefits would apply to both microbial forensics and other basic and applied science fields.

The needs of microbial forensics share many aspects with medicine and public health, which are discussed elsewhere in this report. Many of the capabilities required for detecting and responding to the whole spectrum of natural, intentional, and man-made events are essentially the same. The advantages of the microbial forensics and public health communities working together are that systems created for rare events—for example, bioterrorism—may suffer because of lack of use, but those created for addressing natural and accidental outbreaks of infectious disease are likely to be used frequently. Availability of tools and systems compatible with both rare and common occurrences suggest that when a rare event does occur, these tools and systems will be ready and detection and response will not be delayed by lack of familiarity with them.

Drs. Randall Murch, Bruce Budowle, and Paul Keim, three of the pioneers of the emerging field of microbial forensics, collaborated on a list of key unmet needs—and questions—in microbial forensics, addressing methodologies, technologies, applications, and practices. This list was presented in the opening plenary session of the Zagreb workshop to encourage the other participants to define their own ideas about needs and questions throughout the workshop. As Murch, Keim, and Budowle see it, the challenges include

---

<sup>18</sup> This is exemplified by a colloquium convened by the American Academy of Microbiology in Washington, DC, on September 27–28, 2006, to discuss problems in microbial taxonomy.

- Being able to discriminate with a high level of confidence among similarly presenting natural, accidental, and deliberate outbreaks, within a matter of hours, anywhere in the world. This capability would aid decision-tree design and inform actions of decision makers at all levels to manage response, recovery, and resolution.
- Establishing the limits of current and near-term microbial forensic characterization methods for identification of priority threat agents at levels more specific than the strain/isolate level. What is the probative value for different methods (e.g., method-specific, agent-specific, single approach, and combined approaches)?
- Being able to rapidly develop and validate<sup>19</sup> new and agile forensic analytical methods as a response to a “surprise” event. Investigators need on-the-shelf capabilities that can be adapted for a wide variety of circumstances.
- Sampling and forensic characterization of any relevant microbial background to provide key context for microbial forensic analyses, interpretation, communication, and resulting decision making. There is insufficient understanding of microbial diversity and endemism to inform assessment of where an attack effort may have been developed or perpetrated, or how perpetrators may have exploited the microbial background. And as technologies become more sensitive, it is more likely that organisms that may not have been known to be endemic may be discovered to reside in a particular region. False positives may increase and methods for assessing the significance of false positives and false negatives are necessary. With higher throughput systems, it is conceivable that background sampling could be performed when an event occurs to attempt to define what may be endemic. Of course such an approach requires access to the geographic location of an event, which may not always be feasible.
- Determining the probative value of a “small signal” (microbe of interest) in a “big noise” (highly cluttered with other material or microorganisms, or “dirty”) sample, with defined confidence. Environmental samples are very dirty samples. What if we find one cell of interest, but do not understand the background—what does that cell really mean? What can the clutter tell us? How

---

<sup>19</sup> “Validation” describes a number of activities in forensics. Scientific validity depends on two major factors: reliability, which is the ability of a technique to produce consistent and objective results with known precision and accuracy; and relevance, which is necessary to make evidence admissible in court. Both data quality and the interpretation of the data must be validated (Velsko, 2011b). See Chapter 6 for a more detailed discussion of all the elements of validation.

should scientific and legal significance be determined and supported when the agent of interest is a minority constituent in a “probative sample”? How much of the threat agent of interest must be contained in a sample to be considered significant?

- Exploiting the “clutter” (microbiota<sup>20</sup> other than the threat agent of interest) in metagenomic samples for forensic value, including potential use in the comparison of samples from known and questioned sources. In metagenomic samples (including highly complex samples), can the “clutter” provide more power to microbial forensic analyses? Can it meet “forensic standards”? What is required to demonstrate viability of this approach, including the limits on analysis conclusions?
- Determining other-than-genomic approaches that can be developed, validated, and integrated for deeper forensic characterization, including other omics (e.g., proteomics), and approaches such as multitarget analysis of culture media.
- Determining the maximal characterization for forensic value that can be achieved for biological toxins. In addition to identification, can analyses be developed to aid rule-in/rule-out determinations?
- Maximally reducing the “discovery-to-decision” timeline, across all threat agents, with maximal probative value and confidence. The answer could be integrated with the decision process.
- Establishing how to best validate low-level analytics (very small quantities of a target analyte) in an operational setting. Precise identification of an analyte (e.g., DNA) might be made to the level of a distinct strain (e.g., canonical single-nucleotide polymorphism [canSNP])<sup>21</sup> or other molecule (e.g., isotope signature)<sup>22</sup> at the few- or even single-molecule(s) level.

---

<sup>20</sup> The microorganisms of a particular site, region, or habitat.

<sup>21</sup> “A single nucleotide polymorphism (SNP) is a DNA sequence variation that occurs when a single nucleotide (A, T, C, or G) in a gene sequence is altered (Orho-Melander, 2006). SNPs are useful for both diagnostic identification and phylogenetic population analysis. There is a model for identifying nucleic acid “signatures” (a set of DNA-related molecular markers) that is based on evolutionary rules and comparative genomic analysis. This model is the foundation for identifying key diagnostic features called “canonical characters,” one of which is a canonical SNP (canSNP). A canonical diagnostic character marks a pivotal evolutionary point in the development of an organism and, therefore, represents multiple evolutionary differences. These “signatures” can be used to discriminate between target and nontarget species. A single canSNP can identify a particular species, subpopulation, and/or isolate (Engelthaler and Balajee, 2011; Keim et al., 2004).

<sup>22</sup> “An isotopic signature (or isotopic fingerprint) is a ratio of stable or unstable isotopes of particular elements found in an investigated material. The atomic mass of different isotopes affects their chemical kinetic behavior, leading to natural isotopic separation process.” (<http://www.springerreference.com/docs/html/chapterdbid/330809.html>).



- International data-sharing forums and quality and nomenclature standards. These are essential. Governmental restrictions on sharing material have limited development of global databases needed to provide confidence in microbial forensic analysis. The political obstacles are likely greater than the S&T obstacles.
- Deep sequencing<sup>23</sup> offers methodological advantages over the Sanger method<sup>24</sup> (throughput, speed, cost) but generates a lot of data. Should databases be established in advance or be generated as events occur?
- Determining how to measure with certainty and report whole-genome–sequencing<sup>25</sup> comparisons performed during forensic analysis (e.g., comparing an evidence sample “profile” with a reference sample that may be considered a direct link or have a common ancestor). Sequencing errors and other factors will likely inflate dissimilarity between samples, creating a degree of uncertainty. Defining and quantifying the error rates associated with each platform and chemistry are critical. How do we accomplish this in a communal way?
- Ensuring that the quality of sequence data and the results of bioinformatics analyses is as high as possible. Factors to consider that affect data interpretation and quality include reliability of standards for genomic data representation; uncertainty about databases (e.g., inferences based on available data, including metadata<sup>26</sup>); sequence errors and uncertainties; criteria for comparisons (match, similar, different, inconclusive); and the rigor of expert reasoning, which should include formulating well-defined hypotheses, and testing methods for assessing the weight of microbial forensics evidence.
- Replicating all the essential details of any particular bioinformatics analysis pipeline by different labs is as much art as it is mathematics and science. Investigators must understand it to use it. It is not easy to transfer; there are analytical complexities. There are also multiple technologies, and different versions

---

<sup>23</sup> “Techniques of nucleotide sequence analysis that increase the range, complexity, sensitivity, and accuracy of results by greatly increasing the scale of operations and thus the number of nucleotides, and the number of copies of each nucleotide sequenced” (<http://www.urmc.rochester.edu/profiles/display/123618>).

<sup>24</sup> A widely used method of determining the order of bases. A base is one of the molecules (adenine, thymine, guanine, or cytosine in DNA; adenine, guanine, uracil, or cytosine in RNA) that forms part of a DNA and RNA nucleotide.

<sup>25</sup> Determining the complete nucleotide sequence of an organism’s DNA.

<sup>26</sup> Data about other data.



of programs and data assessments, and software and hardware change rapidly.

- No bioinformatics software comes as a stand-alone. How much documentation is needed? What baseline truths are needed to ensure that assessments and comparisons of technologies can be made effectively?
- Integrating disparate data to provide a single value (e.g., can genomics analysis be added to physical/chemical analysis to yield a single value or answer; can traditional forensics, intelligence, prior odds, networks, be incorporated?) Is simplicity desirable for conveying information “up the chain” to decision makers? Should disparate data be maintained as separate entities?
- Avoiding the filtering of data on the basis of individual preferences and bias.
- Instituting processes to inform decision makers in a way that ensures that the science is properly understood. Many nonscientists who make decisions based on forensic science make those decisions based in large part on media, such as television.

The overall goal for microbial forensics is to move as far to the left as possible on the time-risk continuum portrayed in Figure 1-3. Microbial forensics results that are very informative, have high confidence, and are rapidly obtained—and perhaps better leveraged with other capabilities—could enable investigators to manage risks so that energy is dedicated to anticipating and preparing for an event rather than reacting to a surprise, scrambling to mitigate consequences, and seeking attribution.

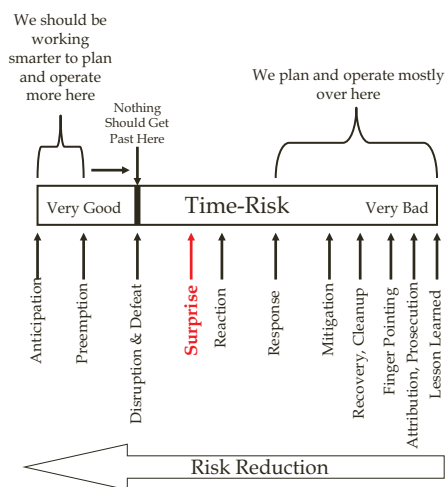
Currently, microbial forensics directs much of its science to operations on the wrong end of this time-risk timeline (Murch presentation, 2013). Dynamic cross-disciplinary international collaboration, however, could serve to improve and develop better microbial forensics capacities, shifting the focus of microbial forensics to the left on the continuum, and at the same time contribute to the strength of other sciences.

## ORGANIZATION OF THE REPORT

As discussed at the beginning of the chapter, the report prepared by the committee draws upon the material presented in the symposium as well as other evidence. This chapter has provided an introduction to the field and to major technical and policy issues. Among its key messages:

- The law enforcement goal of microbial forensics is “attribution”—that is, determining who committed the offense. Based on the

Overall Goal is to Move to the Left as Far As Possible  
(Very Informative, High Confidence and Rapid Microbial Forensics Better  
Leveraged with Other Capabilities Could Help)



**FIGURE 1-3** Time-risk continuum. The goal is to collapse time—to reduce risk by planning and preparing for an event rather than targeting efforts to reacting to an event.

SOURCE: Murch presentation, 2013.

analysis of biological and other evidence, law enforcement builds a case for attribution to a specific source or sources. The evidence supporting attribution must be robust and suitable for use in legal proceedings and to inform decision making at the highest levels.

- Public health, infectious disease medicine, and law enforcement investigations all need to establish whether an event is deliberate, accidental, or natural. Importantly, each can leverage the other's resources to achieve the same initial objectives. The major difference between the two approaches is that the public health investigation's goal is to manage the public health response and protect the public's health and safety, whereas law enforcement's is to provide safety and security by apprehending and convicting those who committed the attack.
- The needs of microbial forensics share many aspects with medicine and public health. Many of the capabilities required for detecting and responding to the whole spectrum of natural, intentional, and man-made events are essentially the same. The advantages of the microbial forensics and public health commu-

nities working together are that systems created for rare events (e.g., bioterrorism) may suffer through lack of use, whereas those created for addressing natural and accidental outbreaks of infectious disease are likely to be used frequently. Developing tools and systems compatible with both rare and common occurrences means that when a rare event does occur, these tools and systems will be ready and detection and response will not be delayed by lack of familiarity with them.

These messages are illustrated and developed in the remainder of the report, with a focus on the major scientific, technological, and policy and process issues that need to be addressed to develop the field of microbial forensics. Chapters 2 through 7 follow the general order of the symposium, while Chapter 8 presents the committee's findings and conclusions, which follow from these chapters and respond to the major messages as well as to the "Key Questions and Unmet Needs" presented in the section above.

## 2

# Microbial Science: Ecology, Diversity, and Characterizing the Microbial World

One might think that the age of great discoveries in biology is past—that humans have been everywhere and seen everything on Earth. But in fact, biology still boasts some formidable unknowns, and many are now being explored in the once invisible world of microbes. In the last 20 years, technological advances have made it possible to explore a microbial world that has proven vastly more extensive, important, and diverse than previously imagined. Analyses of microbial communities in the soil, in the ocean, and even in the human body have shown that previous methods detected only a tiny percentage of the different microbes in these environments. It seems that each technological advance and every new environment sampled reveal even greater diversity in the microbial world. Is there a limit? How can the nature and extent of microbial diversity be satisfactorily characterized?

Source: AAM, 2011a.

Although the world of living things is dominated by microbes, very little is known about the vast majority of them. According to the American Academy of Microbiology, “There are ten-million-fold more bacterial and archaeal cells on our small planet than there are stars in the visible universe, and they may contain as much carbon as all plant and animal life put together” (AAM, 2011b). Much of what is known is based on the very few microorganisms that are culturable in laboratories. Prior to the advent of nucleic acid sequencing, only the phenotypes of culturable microorganisms could be readily studied using tools that had been available for the previous 150 years—growth on selective and differential media,

Gram stains, serotyping with reference antisera, and clinical microbiology techniques—and these are inadequate for the purpose of identifying and attributing the source of the material used in a biological attack because they cannot produce sufficiently fine-grained detail on the nature of the organism (Clements, remarks at Zagreb workshop, 2013).

As noted in Chapter 1, much of the work of microbial forensics today is based on studying biodiversity, phylogenetics,<sup>1</sup> phylogeography,<sup>2</sup> and genomics. The ability to sequence the genomes of microbes has provided an enormous amount of new knowledge about some selected bacteria and viruses. However, even the most basic information remains unknown for most microorganisms and “knowledge of the evolution and ecology of microbial communities lags far behind cellular microbiology” (NRC, 2007:33). Until recently there have been few systematic efforts to collect and describe the microbes living in soil, seawater, freshwater lakes and streams, plants, and even in the guts and on the bodies of humans and other animals. The new techniques of “metagenomics” circumvent the unculturability of the majority of microbes by enabling the study of microbial genes and sequences from entire communities directly from environmental samples. The use of metagenomic techniques, however, is still an area of active development (Darling presentation, 2013). Meanwhile, knowledge of the natural microbial communities residing throughout most of the world is scarce and vastly incomplete. Although the biology of the small fraction of bacteria and viruses that are pathogenic to human beings, livestock and companion animals, and crop or forestry plants is somewhat better known and understood, there is still much to be learned about their evolution, how many strains of pathogen species exist in nature, what their distribution is throughout the world, and how this distribution affects, and is affected by, ecological conditions.

## BIOSURVEILLANCE

Biosurveillance is “the process of active data-gathering with appropriate analysis and interpretation of biosphere data that might relate to disease activity and threats to human or animal health—whether infectious, toxic, metabolic, or otherwise, and regardless of intentional or natural origin—in order to achieve early warning of health threats, early detection of health events, and overall situational awareness of disease activity”

---

<sup>1</sup> The study of evolutionary relatedness among various groups of organisms through molecular sequencing data and morphological data matrices.

<sup>2</sup> A field of study concerned with the principles and processes governing the geographic distributions of genealogical lineages, especially those within and among closely related species.

(Homeland Security Presidential Directive, HSPD-21, 2007).<sup>3</sup> In addition to the scientific benefits that will derive from an improved understanding of the microbial world, the application of that knowledge through improved international biosurveillance capabilities for earlier detection and reporting of new and reemerging infectious diseases underscores the fundamental connections between microbial forensics and public health. As the experience with Severe Acute Respiratory Syndrome (SARS) in 2002 and 2003 demonstrated, the global reach of trade and travel, especially by air, means that diseases with pandemic potential can now reach around the world within hours or days rather than the weeks or months of prior years.<sup>4</sup> Subsequent outbreaks related to various strains of avian influenza (H1N1, H5N1, and H7N9) as well as the Middle East Respiratory Syndrome (MERS) that emerged in 2012 underscored the importance of improving national and international disease surveillance. Many of these diseases originated with animals, emphasizing the need for greater understanding of zoonoses as threats to health:

The WHO is engaging in an ever-increasing number of cross-sectoral activities to address health threats at the human-animal-ecosystem interface. These threats include existing and emerging zoonoses as well as antimicrobial resistance, food-borne zoonoses, and other threats to food safety.<sup>5</sup>

A major breakthrough came in 2005 when the World Health Assembly of the World Health Organization (WHO) adopted revisions to the International Health Regulations (IHR), changing the basis for reporting disease outbreaks from a list of specific diseases to a set of characteristics that constitute a “Public Health Emergency of International Concern.” The IHRs reflect more than a decade of international effort to achieve global consensus on the need to improve disease surveillance capacities and to accept international requirements to do so in a timely manner. Much remains to be done, however; as of the initial July 2012 deadline for reporting progress, fewer than 20 percent of WHO member countries reported that they had been able to meet the IHRs, and a majority requested a 2-year extension.<sup>6</sup> The international assistance being provided to enable countries to achieve the core surveillance and response compe-

---

<sup>3</sup> <http://www.fas.org/irp/offdocs/nspd/hspd-21.htm>.

<sup>4</sup> Between November 2002 and July 2003, the World Health Organization recorded 8,096 probable cases of SARS in 29 countries, with 774 deaths ([http://www.who.int/csr/sars/country/table2004\\_04\\_21/en/](http://www.who.int/csr/sars/country/table2004_04_21/en/); accessed April 6, 2014). One estimate put the economic impact of the outbreak at \$30 billion (<http://www.globalhealth.gov/global-health-topics/global-health-security/index.html>; accessed April 6, 2014).

<sup>5</sup> <http://www.who.int/zoonoses/en/>; accessed April 6, 2014.

<sup>6</sup> See <http://www.who.int/ihr/about/en/> for more information about the IHR.

tencies needed to meet their IHR requirements also serves the development of capacities for microbial forensics.

In addition, national and nongovernmental efforts have already demonstrated the value of increased reporting capabilities. For example, the Global Public Health Intelligence Network (GPHIN), which was originally developed by Public Health Canada in collaboration with WHO, continuously searches media sources for information about infectious disease outbreaks. Pro-Med mail, which in addition to monitoring the media and other sources, relies on reports from its network of 60,000 subscribers in 185 countries, is now a project of the International Society of Infectious Diseases that was begun in the mid-1990s by two U.S. nongovernmental organizations.<sup>7</sup> According to WHO, “more than 60% of the initial outbreak reports come from unofficial informal sources, including sources other than the electronic media, which require verification.”<sup>8</sup>

WHO’s own primary surveillance mechanism is the Global Alert and Response Network (GOARN), “a technical collaboration of existing institutions and networks who pool human and technical resources for the rapid identification, confirmation and response to outbreaks of international importance.”<sup>9</sup> Also relevant to microbial forensics is the Emerging and Dangerous Pathogens Laboratory Networks, another WHO effort to support the development of increased readiness and capacity in high-security public health and veterinary laboratories for detection and response related to both human and animal disease. This network is supported by a number of international working groups, some of which focus on developing common procedures and protocols that could be relevant for microbial forensics as well.<sup>10</sup>

As mentioned above and described in the cases in this chapter and elsewhere in the report, animal, plant, and foodborne diseases can also pose threats to public health and challenges for microbial forensics. This has led to increasing cooperation among WHO, the World Organisation for Animal Health (OIE), and the U.N. Food and Agriculture Organization (FAO). For example, in 2006 the three organizations launched the Global Early Warning System for Major Animal Diseases, Including Zoonoses (GLEWS), which seeks to gain the added value of

---

<sup>7</sup> More information about GPHIN can be found at <http://www.who.int/csr/alertresponse/epidemicintelligence/en/> and [http://www.hc-sc.gc.ca/ahc-asc/pubs/\\_intactiv/gphin-rmisp/index-eng.php](http://www.hc-sc.gc.ca/ahc-asc/pubs/_intactiv/gphin-rmisp/index-eng.php). Information about Pro-Med mail can be found at <http://www.promedmail.org/aboutus/>.

<sup>8</sup> <http://www.who.int/csr/alertresponse/epidemicintelligence/en/>; accessed April 4, 2014.

<sup>9</sup> <http://www.who.int/csr/outbreaknetwork/en/>; accessed April 4, 2014.

<sup>10</sup> Further information can be found at <http://www.who.int/csr/bioriskreduction/laboratorynetwork/en/>; accessed April 4, 2014.

combining and coordinating the alert and disease intelligence mechanisms of OIE, FAO and WHO for the international community and stakeholders to assist in prediction, prevention and control of animal disease threats, including zoonoses, through sharing of information, epidemiological analysis and joint risk assessment.<sup>11</sup>

In the United States, the White House issued the National Strategy for Public Health and Medical Preparedness in response to Homeland Security Presidential Directive-21 (HSPD-21) in 2007. HSPD-21 called for the United States to

establish a biosurveillance capability, with connections to international disease surveillance systems, that can provide “early warning” of a bio-attack or naturally occurring outbreak, and can provide ongoing “near real-time” information about an outbreak as it unfolds. This system should be designed to provide a national “common operating picture” and better situational awareness for state, local, and federal officials, as well as for public and private sector healthcare providers.

In 2008, the Centers for Disease Control and Prevention (CDC) was tasked to lead the establishment of the National Biosurveillance Advisory Subcommittee (NBAS). NBAS provides advice to the Advisory Committee to the Director as well as “leadership and guidance” for implementing the National Biosurveillance Strategy. It issued its first report in 2009 (National Biosurveillance Advisory Subcommittee, 2009), calling for strategic goals and investments in biosurveillance activities and technologies, hiring and retaining trained personnel, and electronic health records and databases. In a second report in 2011, NBAS made further recommendations on governance, information exchange, workforce, and research and development (National Biosurveillance Advisory Subcommittee, 2011). In 2012, the White House issued its National Strategy for Biosurveillance.

More recently, the Global Health Security Agenda, announced in February 2014, is a new partnership created by the United States among WHO, OIE, and FAO along with more than 25 countries to

“... accelerate progress toward a world safe and secure from infectious disease threats and to promote global health security as an international security priority, to

- **Prevent** and reduce the likelihood of outbreaks—natural, accidental, or intentional;
- **Detect** threats early to save lives;

---

<sup>11</sup> <http://www.glews.net/>; accessed April 6, 2014.



- **Respond** rapidly and effectively using multisectorial, international coordination and communication.<sup>12</sup>

Again, success in achieving a number of the Agenda's goals would also mean improved international capabilities for microbial forensics.

The U.S. microbial forensics community has already intensified its efforts to partner with the public health sector, particularly to improve and increase the use of biosurveillance information. Such information would help microbial forensics investigators to better understand the context of an active outbreak and better interpret microbial forensics findings.

The rest of this chapter explores several actual microbial forensics cases to illustrate the techniques that have been developed over the last 10-12 years, some of the relationships between public health and microbial forensics, and the unmet needs that remain in the area of understanding microbial ecology, diversity, and characterization of strains and other genetic variants. Other aspects of public health and microbial forensics, including unmet needs related to technologies and processes, are taken up in subsequent chapters.

### THE AMERITHRAX CASE

Although several attempts to use pathogens as weapons occurred prior to 2001, the most comprehensive major case involving the use of a pathogen as a biological weapon occurred with the anthrax letters (or "Amerithrax") case. In fall 2001, less than a month after the September 11, 2001 attacks on the World Trade Center and the U.S. Pentagon, letters containing spores of the anthrax disease-causing bacterium (*Bacillus anthracis*, or *B. anthracis*) were sent through the U.S. mail. Between October 4 and November 20, 2001, 22 people were sickened by anthrax, with 5 tragic fatalities (Jernigan et al., 2002). In addition, there was extensive environmental contamination caused by the anthrax mailings in U.S. Postal Service buildings, Senate office buildings, and in many other places that processed or received the contaminated letters. The FBI led the effort to characterize the material contained in the letters and identify the individual or individuals responsible for the mailings. This investigation involved extensive scientific study spanning almost 9 years (NRC, 2011). Over the 9 years, the FBI devoted 600,000 investigator hours to the case, which involved 17 Special Agents as well as 10 U.S. Postal Inspectors. During the investigation, 10,000 witnesses were interviewed, 80 searches conducted, 4 million megabytes of computer memory were analyzed,

---

<sup>12</sup> Additional information about the Agenda may be found at <http://www.globalhealth.gov/global-health-topics/global-health-security/ghsagenda.html>; accessed April 6, 2014.

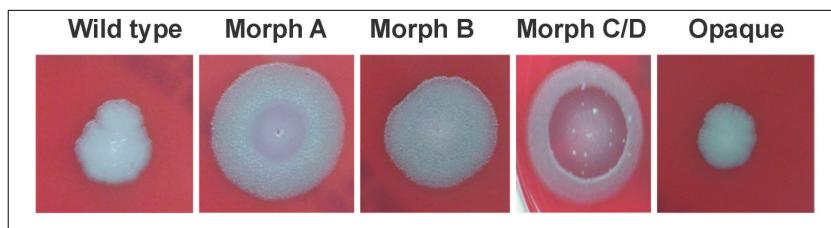
and 5,750 grand jury subpoenas were served. Twenty-nine government, university, and commercial labs assisted with the scientific analyses that were a central aspect of the investigation (U.S. Department of Justice, 2010). Note that, at the time of the anthrax letters mailings, the tools and technologies that were readily available were not adequate and the science of microbial forensics was in its infancy and limited to a few pioneering laboratories (Clements remarks, 2013). The Amerithrax investigation accelerated the development of microbial forensics, resulting in remarkable development and applications of new techniques and approaches for using laboratory tools to pinpoint the genetic identity of a microbial agent (Tucker and Koblentz, 2009). Microbial forensics became an essential part of the scientific investigation, which was combined with physicochemical analyses and other evidence to narrow the search for the source of the *B. anthracis* used in the attacks.

An important point to be made here is that *B. anthracis* is extremely stable genetically. One reason for this stability is that the organism's life cycle includes long periods of dormancy in the form of spores, rendering the genome highly homogeneous (Pilo and Frey, 2011). Paradoxically, reconstruction of the evolutionary history of *B. anthracis* has been challenging because of the same stability that allowed investigators to track its use as a biological weapon (Van Ert et al., 2007). Very early in the investigation the anthrax spores in the letters and the environmental and clinical isolates were identified as the "Ames strain" by Dr. Paul Keim and members of his team at Northern Arizona University (NAU). At the Zagreb workshop, Keim summarized what enabled them to accomplish this identification in the era before whole-genome sequencing (WGS) became commonplace. Research in the mid-1990s to differentiate *B. anthracis* strains had been limited to working with small sections of genomic DNA. The genomes of very few bacteria had, at that point, been fully sequenced. Keim and his colleagues had focused on hypervariable regions in the genetic material of *B. anthracis* called variable number tandem repeat (VNTR) regions. His team was able to identify eight loci in *B. anthracis* that had multiple alleles (i.e., different versions of the same genetic marker). They developed a typing system called multiple-locus VNTR Analysis, or MLVA. In fact, *B. anthracis* was one of the first bacteria upon which this subtyping was performed. In 2001, NAU had a database of about 400 different *B. anthracis* strains and could differentiate these isolates into approximately 90 different genotypes. Both Keim's lab at NAU and the CDC lab independently produced the same result: the MLVA8 genotype was found to be consistent only with the Ames strain genotype that Keim and colleagues had identified earlier (Keim et al., 2000).

The Ames strain was a known laboratory strain used by the U.S. military and other vaccine development teams. The strain was origi-

nally isolated from a dead cow in Texas in 1981 and was shipped by Texas A&M University to the U.S. Army Medical Research Institute of Infectious Diseases (USAMRIID) at Fort Detrick in Frederick, Maryland (NRC, 2011). Over time, it was shared with other labs in the United States and in Canada, Sweden, and the United Kingdom. Thus, although the identification of the Ames strain narrowed the possibilities, making the source likely to be a laboratory that had access to the strain, MLVA was insufficient to identify its source unequivocally. The evidence, therefore, needed to be more painstakingly examined for additional unique and distinguishing genetic and other features that could be compared to samples obtained from laboratories holding the Ames strain to narrow the search for the source and perpetrator(s). Scientists from the Department of Defense had identified several colony morphology variants in the samples from the letters (Figure 2-1). With funding from the National Institutes of Health, the National Science Foundation, and other government agencies, FBI scientists worked with The Institute for Genomic Research (TIGR) in Rockville, Maryland, to identify the genetic basis of the altered appearance of the bacterial colonies (Read et al., 2002). Some of these were determined to be insertions or deletions, while others were associated with single nucleotide polymorphisms (SNPs) (NRC, 2011). FBI investigators then contracted with four laboratories to develop highly specific molecular genetic assays to try to detect some of these variants in the anthrax powder evidence. These assays were used in the examination of the samples the FBI had collected to form its repository of known Ames strain samples.

Unfortunately, the rigor used in collecting and processing samples did not match the rigor of the powerful and validated assays. Validating an evidence-handling stream is very difficult. Achieving high confidence



**FIGURE 2-1** Variants of representative colonies of each of the morphotypes of *B. anthracis* Ames.

SOURCE: Rasko et al. (2011).

would entail processing a large number of blank samples,<sup>13</sup> which is very expensive. In the anthrax letters case, this could not be performed in real time and was not done. A “false positive” result that was uncovered during early evidence handling led to the unnecessary and expensive draining of a Maryland pond. This evidence proved not to be consistent with the Ames strain (NRC, 2011). Trace evidence errors can lead to costly mistakes. Nevertheless, the analysis of the repository samples using the then newly created assays eventually led the FBI to focus attention on a particular flask at USAMRIID containing a spore preparation known as RMR-1029.

In addition, physical science analytical approaches, such as scanning transmission electron microscopy, energy-dispersive X-ray analysis, carbon dating by accelerator mass spectrometry, and inductively coupled plasma-optical emission and mass spectrometry, were used to try to determine the chemical and elemental profiles of the spore powders (NRC, 2011). These kinds of testing were designed to answer such questions as how the material was grown and processed, when the anthrax preparation might have been made, whether there were contaminants or trace elements that would provide a clue to the production location or materials used, and whether there was evidence of an effort to deliberately include additives to improve dispersal of the anthrax spores.

Early in the investigation, elemental analysis of the letter spores detected silicon, leading some to conclude that this material had been deliberately added to aid aerosol dispersal of the anthrax spores. The conclusion that silicon was present was correct, but the interpretation that it had been added to weaponize the spores was not supported. A sensitive and sophisticated analysis performed at the Sandia National Laboratories showed that the silicon signal was inside, not outside the spore coat, which meant the spores had not been coated to decrease clumping (Michael and Kotula, 2009). However, because the erroneous conclusion was leaked to the media, it was never possible to completely mitigate this false conclusion. In fact, controversy concerning the source of the silicon content in the attack material remains active in the literature (Bhattacharjee, 2010; Hugh-Jones et al., 2011).

Two other analytical paths proved to be ineffective. One was stable isotope analysis. Because rainfall can have different oxygen isotope ratios, this analysis was performed to try to identify the region of origin of the

---

<sup>13</sup> “Blanks are analytical quality control samples analyzed in the same manner as site samples. They are used in the measurement of contamination that has been introduced into a sample either (1) in the field while the samples were being collected or transported to the laboratory or (2) in the laboratory during sample preparation or analysis” (U.S. Environmental Protection Agency, 2004; <http://www.epa.gov/oswer/riskassessment/ragasa/pdf/ch5.pdf>).

water used to produce the spores. However, rainfall composition was not sufficiently predictive; in addition, autoclaving versus filtering water for sterilization can change isotope ratios. The second path was predicated on the fact that early nuclear weapons testing had caused a change in atmospheric carbon-14 ratios. This phenomenon suggested a potential way to date the production of the spores and showed that the spores in the letter sent to Senator Leahy were produced between 1998 and 2001. However, this date range did not prove to be useful information.

Keim emphasized, during his presentation at the Zagreb workshop, that one of the most powerful tools in forensics is one that has the ability to differentiate between inclusion and exclusion. The term “inclusion” means that an association or match might exist; that is, it is a failure to exclude some putative relationship. There might be an exact match—as might occur in some cases in human DNA testing—or the sample might be identified as a co-member of the same phylogenetic group. (The term “member” is preferred to “match” because in phylogenetic analysis a “member” may not be identical but would still be part of the same strain.) Conclusions as to whether two things are the same—inclusion—tend to be weak for use for attribution in court, because the clonal nature of microorganisms and unknown history or other factors can be used to discount or reduce the significance of such evidence. In contrast, exclusion can serve as very powerful evidence (Budowle et al., 2008). When it is concluded that something is not the same—not the Ames strain—based on the evolutionary models and assays, these conclusions are very strong. Exclusionary conclusions do not get newspaper headlines, but they are very important as they reduce the potential list of candidates and allow investigators to further focus efforts. The hypothesis testing used in molecular forensics and epidemiology can be summarized as follows:

- “Inclusion” (failure to exclude). Might be an exact match or at least co-members of the same phylogenetic group that are descended from a most recent common ancestor. Conclusions may be weakly supported.
- “Exclusion” (failure to include). Different from the comparator and *very* unlikely to be the same or from the same source. Conclusions may be very strong.

Three examples of the power of exclusion are based on the analyses of the *B. anthracis* strains obtained from (1) the biological-weapons lab accident in Sverdlovsk in the former Soviet Union in 1979 (Meselson et al., 1994), (2) the Aum Shinrikyo terrorist attack in Tokyo in 1993 (Danzig and Hosford, 2012; Keim et al., 2001), and (3) a biological-weapons production facility in Al Hakam, Iraq (analyzed soon after the 2001 U.S. anthrax

letters attack). None of these were consistent with the Ames strain in the anthrax letters. Links to the Japanese terrorist group and the Russian and Iraqi biological weapons programs were therefore ruled out.

Keim named the following as the key challenges associated with the anthrax letters investigation:

1. Initial genetic conclusions about what the anthrax letters strain was and where it originated were based upon limited databases, which investigators recognized. Despite this, the information was shared with policy makers, who proceeded as though these initial conclusions were strong conclusions.
2. Throughout the investigation, genomic technology was undergoing rapid developments from relying on an 8 locus MLVA system to 15- and 60-locus MLVA systems, to SNP analysis, and finally to WGS—which made trying to validate a system in real time very difficult.
3. Trace evidence was analyzed using partially validated systems, leading to costly investigative errors.
4. Scientific and peer-reviewed publication was limited during the entire process. To this day, critical data generated by the investigation remain unpublished.
5. A “kitchen sink” approach to forensics was used. There was strong pressure to “try my method,” including pressure exerted from very high political levels.
6. The public media added to confusion and disruption. There was too much coverage, for example, of the incorrect interpretation of the presence of silicon, yet too little when the false interpretation was corrected.
7. It was a politically charged environment. Certain individuals sought the limelight, and their decision making affected the process of the science.
8. Subject-area experts were also criminal suspects. Those performing analyses were simultaneously being investigated.

Obviously some of these circumstances are unavoidable in today’s environment. However, Keim provided two strong suggestions for improving the forensic approach to microbial attribution. The first is that investigators should define a specific hypothesis or hypotheses. Hypothesis building before obtaining results can reduce bias by averting any tendency to try to fit the results to a desired interpretation. If investigators focus their questions around hypotheses to be tested, it is possible to develop yes/no or inclusion/exclusion criteria in an investigation. The second recommendation is that investigators define a relevant reference



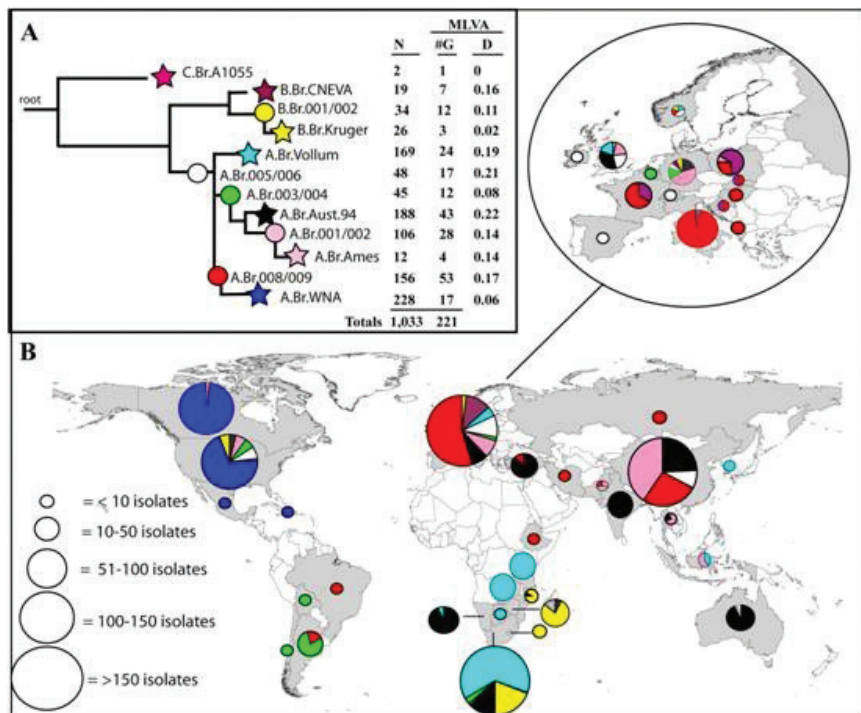
population for genetic analyses. Early in the anthrax investigation, there was inadequate information about both North American and worldwide *B. anthracis* strains.

### THE IMPORTANCE OF PHYLOGENETIC INFORMATION

The Amerithrax case abundantly illustrated the importance of having adequate phylogenetic information about *B. anthracis* Ames, the agent involved in the case. Fortunately, Paul Keim and others had developed a rather extensive, although still incomplete, picture of the evolution of *B. anthracis* and the distribution of major strains throughout the world (Figure 2-2). This knowledge of strains of *B. anthracis* helped to pinpoint quickly the identity of the attack agent as the Ames strain because its genetic characteristics could be compared with other strains in the reference database. It should be noted, however, that the Ames strain is unusual in having a short history and a limited distribution. If the attack strain had been something other than Ames, its identity might not have been as simple to pinpoint.

In a more recent case dealing with anthrax, Dr. Richard Vipond, Operations Manager for the Rare and Imported Pathogens Laboratory at Porton Down in the United Kingdom, reviewed the forensic work employed during a 2009-2010 anthrax outbreak involving contaminated heroin (Price et al., 2012; Vipond presentation, 2013). In the United Kingdom and Scotland, most rare diseases are imported. Porton Down, part of the U.K. public health system, has substantial expertise in this area and is a WHO collaborative center for arbovirus, viral hemorrhagic fevers, and special pathogens, including *B. anthracis*. Historically, anthrax disease in the United Kingdom has been linked to importation and to contamination by animal products of “brownfield” (contaminated land previously used for industrial purposes) sites. Outbreaks in cattle occur in contaminated burial pit sites and imported materials; people renovating buildings may be infected via handling materials such as contaminated horsehair plaster.

Because of their experience with the drumthrax cases, the Porton Down investigators had the tools Keim’s team had developed for the Amerithrax case in place and thus were able to use 13 canonical SNPs (Van Ert et al., 2007) and 8 MLVA loci (Keim et al., 2000) to establish quickly that the heroin-associated strains were all the same, including the one in Germany. The strains matched the Trans-Eurasian group of strains, which is widely distributed, most likely through trade routes (Figure 2-2). The strain had not previously been seen in the United Kingdom, nor is it shown there on the map in Figure 2-2. With real-time PCR methods, resolution down to the single copy is possible, and so detection is not a major issue. There were 53 confirmed cases, and Vipond suspects there



**FIGURE 2-2** Worldwide distribution of *B. anthracis* clonal lineages: phylogenetic and geographic relationships among 1,033 *B. anthracis* isolates. (Panel A) Population structure based upon analysis of data from 12 canonical SNPs. The number of isolates (N) and associated MLVA genotypes (G) within each sublineage are indicated as well as the average Hamming distance (D) as estimated from VNTR data. The major lineages (A, B, C) are labeled, as are the derived sublineages (1-12), which are also color-coded. (Panel B) Frequency and geographic distribution of the *B. anthracis* sublineages. The colors represented in the pie charts correspond to the sublineage color designation in panel A.

SOURCE: Van Ert et al., 2007.

may have been 117 cases in total. Grunow et al. (2013) also reported that in 2012, additional cases occurred in Germany, France, Denmark, and the United Kingdom.

Vipond does not believe that the heroin-related anthrax outbreak represented a biocrime or bioterrorist attack. One argument against the outbreak being due to bioterrorism is that drug users are not an ideal target population for capturing the public's interest. In addition, when the police stepped up their operations in Glasgow in hopes of confirm-



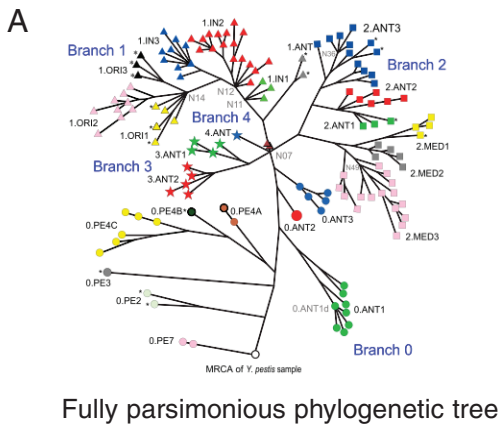
ing the source, they may have contributed to the spread of the disease by driving distributors to sell the heroin farther afield, resulting in cases emerging elsewhere in Scotland and down into England. The German case is thought to have emerged as a result of the normal trafficking route.

Vipond also pointed out that assays are becoming progressively more sensitive and are picking up natural background. A key need is to know what is there already, so one can spot the difference between a deliberate release and a new natural outbreak. One can live with natural background levels of *B. anthracis* in the soil and not become infected. So as detection gets better, one will be able to identify what has been there since time immemorial. The key is being able to distinguish between an anomaly and a deliberate release so no one hits the panic button every time something unusual is seen.

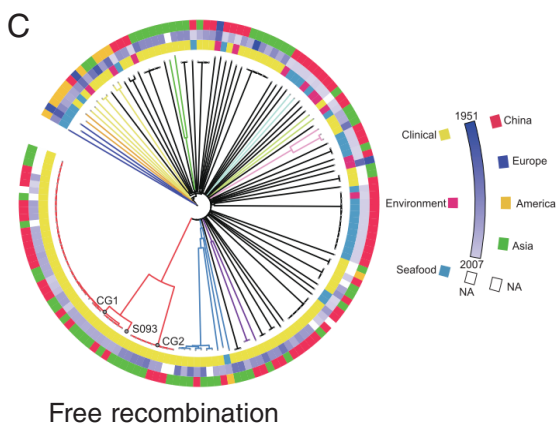
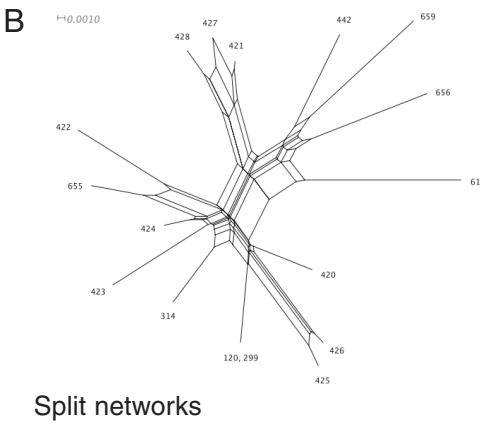
Dr. Ruifu Yang of the Beijing Institute of Microbiology and Epidemiology in China is a recognized expert in, among other things, the phylogeny and evolution of *Yersinia pestis*, the organism that causes plague. At the Zagreb workshop, he reviewed his work with genome sequences of *Y. pestis*, samples of which have been collected from around the world, to show how phylogenetics can aid microbial forensics. He also discussed the effect of environment on a pathogen's ecology and evolution. A pathogen's ecology can strongly determine its genomic diversity. For some microorganisms, ecological diversity is high because the organisms may inhabit ecological niches in soil, the ocean, and even in the human body, in which there are 10 times the number of microbial cells as human cells (Koeppel et al., 2008).

Yang classified bacterial pathogens into three groups based on the rate of a pathogen's genomic recombination, the acquisition of alleles. A genomically monomorphic pathogen, such as *Y. pestis* or *B. anthracis*, rarely exhibits recombination, and thus may be relatively easy to source (see Figure 2-3). A genomically intermediate diversified pathogen (GIDP), such as *Escherichia coli* or *Klebsiella pneumoniae*, undergoes frequent recombination, but this does not necessarily affect the identification of phylogenetic relationships across samples. But a genomically highly diversified pathogen (GHDP), such as *Helicobacter pylori* or *Vibrio parahaemolyticus*, has a high frequency of recombination, and the phylogenetic relationship, especially in the deep branches, is distorted, which makes tracing a source more difficult, although it is still possible to distinguish among populations (e.g., Asian vs. U.S. strains).

An understanding of a pathogen's ecology and genomic diversity will aid in developing microbial forensics strategies to trace a pathogen's source. For example, it would be appropriate to use SNPs to trace all three of the above categories of pathogens; to use VNTRs to trace GIDP and GHDP pathogens; and to use loss or gain of genomic fragments, as well



**FIGURE 2-3** Three classifications of bacterial pathogens. (A) *Y. pestis*: Rare recombination, with parsimonious phylogenetic tree. (B) Intermediate level of recombination. (C) A neighbor-joining tree of 157 *Vibrio parahaemolyticus* genomes. *V. parahaemolyticus* is a highly diversified pathogen in which high-frequency homologous recombination leads to a radiated phylogenetic relationship. SOURCES: A, Cui et al. (2013); B and C, Yang presentation, 2013.



as combined markers, to trace GDP and GHDP pathogens. Databases can be built to accommodate these different target markers.

*Y. pestis* is widely distributed in China. As of 2010, there were 296 counties with 1.437 million km<sup>2</sup> of natural plague foci. In different landscapes, climates, and ecologies, the hosts (e.g., various species of marmot, squirrel, rat) and bacterial strains vary. Dr. Yang and his colleagues are trying to understand the complex interactions between the pathogens, hosts, insect vectors, and environment. They have performed WGS on single strains and employed comparative genomics to understand pathogen diversity, using SNP analysis to understand the transmission of *Y. pestis* through history and the phylogenetic differences in strains. Collaboration with other investigators has enabled them to track the historical transmissions of *Y. pestis* plagues, which have repeatedly emerged from China, to trade routes such as the Silk Road (see Figure 2-4; Morelli et al., 2010). These analyses enabled them to track the evolution of different strains. The Black Plague strain of *Y. pestis* was different from strains before and

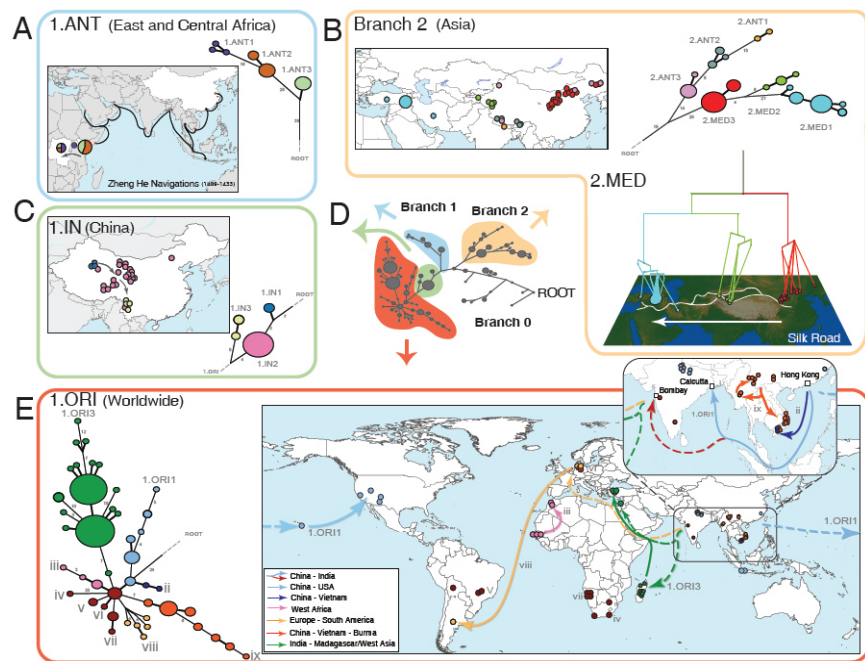


FIGURE 2-4 Distribution of *Y. pestis*.  
SOURCE: Morelli et al. (2010).

after it, and the ecological context in which the strains developed helped shape these differences.

Unfortunately, having information at the level of detail that is available for anthrax and plague is unusual for most microorganisms (AAM, 2009a). As noted in the opening of this chapter, the microbial world has proven vastly more extensive and diverse than previously imagined. Metagenomic techniques now offer the possibility of determining the presence of new microorganismal genes directly from environmental samples or from the bodies of humans, other animals, and plants. To take advantage of this technique, the U.S. National Institutes of Health established a major project on the Human Microbiome and there are comparable European programs in existence. These efforts have produced vast quantities of new information on the bacteria, viruses, and microbial eukaryotes living in or on the human body (e.g., see Human Microbiome Jumpstart Reference Strains Consortium, 2010). One researcher, Rob Knight, has even investigated the possibility that human skin bacteria, which appear to be unique to individuals, can be used as a forensic identification tool. Knight and his team found that residual skin bacteria left on objects could be matched to the skin bacteria of the individual who touched the object. Furthermore, these bacteria can be recovered from objects, such as a computer mouse, for up to 2 weeks (Fierer et al., 2010). The human microbiome is, however, influenced by a wide diversity of factors—for example, health, environment, diet, hygiene—which present many challenges.

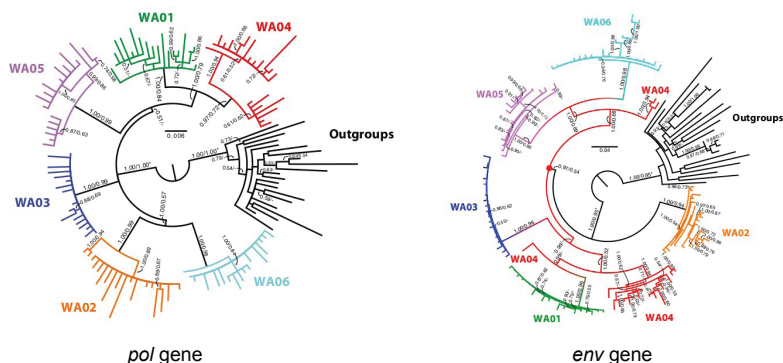
Other groups are seeking to characterize new microorganisms, such as viruses associated with particular animals. Anthony et al. (2013) recently reported on a rigorous and systematic 5-year effort to discover the viruses associated with a single bat, the Indian Flying Fox, which carries Nipah and Hendra viruses, among others. Fifty-five putative viruses (50 of which were new to science) from nine virus families were identified by repeated sampling from this one species. Extrapolating from the numbers derived from the Indian bat to the rest of the roughly 5,500 known mammalian species, the group estimated that there are a minimum of 320,000 undiscovered mammalian viruses waiting to be described, assuming that all mammalian species harbor a similar number of viruses. The importance of this work was characterized as enabling the establishment of preliminary bounds on the potential size of the pool of zoonotic viruses in mammalian wildlife. Other “virus hunters” are working to develop an early warning system to monitor the transmission of infectious diseases from animals to humans. Nathan Wolfe’s work in Africa and Southeast Asia, for example, is designed to develop forecasting capabilities that would identify and mitigate potential new zoonotic disease outbreaks before they become serious pandemics (Wolfe et al., 2007).

Such efforts are, however, barely scratching the surface of the microbial world. Even phylogenetic studies of organisms that are relatively well understood, such as *B. anthracis*, are still incomplete, and there is as yet no worldwide repository of information on pathogen species and strains or any coordinated effort to develop information exchange mechanisms. John Clements pointed out that if public health agencies are not routinely collecting surveillance data (e.g., on influenza cases), data will not exist for law enforcement to use for microbial forensics purposes. Such data could be useful to both sectors.

Dr. Aaron Darling further reviewed microbial ecology and diversity in the context of forensics, focusing on metagenomics, which he defines as the study of DNA from uncultured organisms. He believes metagenomics has enormous potential to help in microbial forensics but that challenges remain for optimization of the technology. A major advantage of using metagenomics is that one can avoid culture bias (some organisms reproduce more quickly than others). There are also, however, disadvantages to metagenomics. One is DNA extraction bias. An organism cannot be sequenced if its DNA cannot be extracted. For example, if the DNA is trapped in spores with hard, uncrackable walls, extraction may not be successful. Therefore, the nucleic acids of an organism that are easily extractable in a sample may suggest that that organism is in greater abundance whereas another organism—perhaps of forensic interest—seems to be of low abundance. A second limitation is that a small number of species greatly dominate most ecosystems. It is difficult to study rare taxa, which comprise the majority of species, and the organism of interest may be naturally present at very low abundance. This situation would require a great amount of sequencing for characterization, higher throughput, or better targeting of informative genetic markers.

Darling stated that most of what is known about the ecology and evolution of pathogens comes from the study of isolates. Like Dr. Yang, he and his colleagues have categorized organisms by their recombination propensity, dividing them into organisms that tend to be monomorphic, those that are intermediately clonal, and those that recombine heavily. Even highly recombinant organisms, however, can reflect obvious phylogenetic structure.

Virologists have led the way in culture-independent forensic analyses, and there have been a number of biocrime cases where phylogenetic evidence was used to support prosecution of individuals for deliberately infecting others. In one such case, an HIV-infected individual was accused of intentionally infecting victims. Analyses of samples of DNA sequences of human immunodeficiency virus 1 (HIV-1) taken from the victims and the alleged suspect were performed in order to reconstruct the phylogenies of the HIV retrovirus polymerase (*pol*) and envelope (*env*)



**FIGURE 2-5** Red lineages represent HIV sequences from the suspected perpetrator. These lineages are ancestral in the *env* genes but not the *pol* genes in victims' samples.

SOURCE: Scaduto et al. (2010).

genes (Scaduto et al., 2010).<sup>14</sup> However, Darling noted with concern that the analysis of the *pol* gene shows a cluster of the victims' *pol* sequences that does *not* include any of the suspect's sequences, whereas analysis of the *env* gene shows the suspect's sequences are ancestral to most of the victims' sequences (see Figure 2-5). The *env* gene was used as evidence to support association, but seemingly discordant results from two genes with different evolutionary rates should be understood and appreciated.

Darling noted the previously mentioned methods focused on SNPs as the basis for inferring phylogenies and performing forensic analyses. He pointed out that WGS, however, has revolutionized the understanding of bacterial evolution. In one study, after sequencing three *E. coli* genomes, researchers compiled a list of all genes in the genomes; only about 40 percent of genes were common to all three strains (Perna et al., 2001; Welch et al., 2002). Although it has been known for some time that bacteria exhibit heterogeneity among strains (e.g., Bergthorsson and Ochman, 1995), it had long been assumed that members of the same species generally have highly homogeneous genomes and are largely identical because they are

<sup>14</sup> The paper by Scaduto et al. presents the molecular evidence used in two U.S. criminal cases: *State of Washington v Anthony Eugene Whitfield*, case number 04-1-0617-5 (Superior Court of the State of Washington, Thurston County, 2004) and *State of Texas v Philippe Padieu*, case numbers 219-82276-07, 219-82277-07, 219-82278-07, 219-82279-07, 219-82280-07, and 219-82705-07 (219th Judicial District Court, Collin County, TX, 2009). The discrepancy noted by Darling comes from the Texas case.



the same species. This belief proved untrue, sparking a controversy over the nature of the definition of a species in bacteria. But it also suggested that one might be able to use gene content, instead of just nucleotide sequences, for forensic characterization because the gene content appears to be evolving almost as fast, if not faster, than the nucleotides.

Darling related information about a joint pilot project by the University of California, Davis and the JGI, that sought to characterize the genetic diversity of microbial organisms worldwide.<sup>15</sup> As noted earlier, if one wishes to build a database of background information about different microbial ecosystems around the globe, an understanding of what exists is needed. The researchers gathered cultures of approximately 100 organisms they deemed the most diverse from culture collections around the world and sequenced their genomes (Wu et al., 2009). They then measured the rate at which they discovered new genes, which was very high. Moreover, novelty in gene content correlates highly with novelty in the 16S rRNA<sup>16</sup> sequence (Wu et al., 2009). Even within a single species, with each additional genome sequenced, discovery of new genes occurs.

As the discovery of additional 16S rRNA diversity continues, it becomes ever clearer that an immense, and probably unachievable, amount of sequencing would be required to characterize all the microbial DNA and RNA on the planet. The amount of 16S rRNA diversity that has been described in cultured samples is dwarfed by the amount of diversity that has been described in uncultured or cloned cultures. Uncultured bacterial diversity greatly exceeds known isolate diversity. How do we gain access to all this uncultured diversity if one cannot grow it? There is a need to be able to build up databases on background microbial communities and use them in forensics.

Although metagenomics has been posed as a solution for this task, Darling indicated that metagenomics in its current form is far from adequate. There are two problematic technical issues. First, samples are complex mixtures. When extracted, the fragmented DNA from the chromosomes from the various microbes in a sample are comingled and they cannot readily be pieced back together. Second, as part of the sample preparation for sequencing, the already fragmented DNA is broken down further to much smaller sizes so the sequencing chemistries can read them. In breaking up the longer molecules, information is lost. While assembly of short read lengths has been possible with homogeneous samples, the assemblies are problematic with complex samples and some genes or anonymous sequences may not be detected. Furthermore, it is

---

<sup>15</sup> For more information see <http://www.jgi.doe.gov/programs/GEBA/index.html>. Accessed November 13, 2013.

<sup>16</sup> The 16S rRNA gene is a section of prokaryotic DNA found in all bacteria and archaea.

difficult to place separate genes, if detected, on the same microbial chromosome even if they originally were syntenic. It becomes unclear which chromosomes are which, which chromosomes were in which cells, and which genes were in which chromosomes. This is a tremendously difficult dataset to analyze, and currently most analyses do not do a good job of reconstructing the cellular origins of the genomic material. To solve this problem, Darling suggests (1) physically “dissecting” microbial communities to preserve information, (2) developing better inference methods with the help of bioinformaticians, and (3) developing alternative sequencing technologies.

Darling has been investigating amplification and sequencing of DNA from single cells obtained directly from environmental samples (i.e., single-cell genomics). He and his colleagues sequenced 201 uncultivated archaeal<sup>17</sup> and bacterial cells from nine different habitats. These organisms belonged to 29 major, but mostly unfamiliar, phylogenetic branches (Rinke et al., 2013). The additional genomic information enabled the researchers to resolve many intra- and interphylum-level relationships and to propose two new superphyla. However, the microbial diversity they were able to capture by this method was only a tiny fraction of organisms in existing genome databases. Therefore this method currently does not scale up for application to microbial forensics.

Darling believes that the most important metagenomic analyses in microbial forensics are phylogenetic analyses. He explained that while the classic, isolate-based genome alignment provides one full-length DNA sequence per isolate, metagenomic alignments provide many DNA fragments, with no information on which fragments originated from the same cell. Current methods for phylogenetic inference assume that complete linkage information is available. In particular, they assume that a full-length DNA sequence for the gene under study is available from each organism being analyzed. However, metagenomic data do not satisfy these requirements because the data frequently contain small nonoverlapping fragments of each organism’s gene sequence. Two inferential approaches that can be used to create phylogenetic trees from metagenomic data include phylogenetic placement, in which the read fragments are “mapped” into a reference phylogeny tree; and co-estimating the read clusters. This process is predicated on the assumption that there is a limited amount of diversity in a metagenomic sample: even if one has 20 million reads, they must combine into a finite number of species. Two ways to implement this approach include Bayesian evolutionary analysis

---

<sup>17</sup> Archaea constitute a domain of single-celled microorganisms. Both the bacteria and archaea domains comprise prokaryotic organisms, but the two domains have different evolutionary histories.



by sampling trees (BEAST), which is a software architecture for analyzing molecular sequences related by an evolutionary tree, and PhyloSift.

Phylogenetics, which provides information on evolutionary relationships based on genetic data, can be used to determine what may be present in a sample. The greater the amount of genetic data, the deeper the resolution of the information on the microbe(s) present. With limited information, such as a single marker, phylum level may be the best that can be determined, which is likely to be of little use for attribution and determining the presence of, for example, a Select Agent. With a whole genome or a number of genes, deeper classification may be feasible. Bayesian hypothesis testing can be applied to any gene or genetic marker phylogeny, including toxin genes of interest. Metagenomic-derived phylogenetic data also can be used to answer the question, "What is the closest match between a sample of interest and a database of related samples?" This can be accomplished using phylogenetic placement data and the "phylogenetic Kantorovich–Rubinstein distance" (Evans and Matsen, 2012).

However, these methods can be applied only to a single gene at a time or to a set of phylogenetically coherent genes. The capability to analyze all genes simultaneously is needed. Given current technology, the information that is contained in whole genomes is not going to be entirely available and some information will not be captured. The analytical challenge of accomplishing this corresponds to reconstructing the histories of coevolution of all genes in a phylogenetic pattern. Evolution is complex and reticulate. Within a species tree, there are phylogenies of individual genes, which include gene duplication, gene conversion, lateral gene transfer, and gene death. The target areas that define a species or strain and distinguish it from near neighbors must be well defined. One or more sites may be required depending on phylogenetic resolution. Efforts are under way using informatics and phylogenetic reconstruction to meet the challenges (Bérard et al., 2012; Boussau et al., 2013). But scaling these methods presents technical, cost, and training challenges and is likely to remain difficult.

## THE ISSUE OF DATA SHARING

Even if there is a major effort to characterize microbial diversity comprehensively through nucleic acid sequencing, unless the data that are collected are shared in such a way that scientists around the world have access to them, they will not provide optimal benefit for either public health or microbial forensics. During the discussion following the workshop session on microbial ecology and diversity, various workshop participants pointed out a variety of problems involved with data sharing.

Paul Keim pointed out that language (e.g., English vs. Chinese) and export-control laws can pose problems for sharing data, particularly for organisms that are considered potential bioweapons. He noted that export control laws can sometimes be circumvented by publication of the data, but Dr. Yang stated that the decision to share is not always left up to the scientists, who are not always permitted to publish everything. Yang said that the worldwide scientific community could press governments to address the need for data sharing, given the bioweapons context, citing the needs of public health and microbial forensics. Yang's database of *Y. pestis*, for example, could aid a traceback in the event of a *Y. pestis* outbreak anywhere in the world. A major theme among the workshop participants was that solutions to data sharing are severely needed. Ideally, countries would have their own microorganism collections and genomic databases and share them. But standards for the technology used would be needed so that all collections would have the same level of testing, accuracy, and quality. In addition, large developed countries have an advantage over small developing countries when it comes to data-generating capacity. It was suggested that an international framework is needed both to encourage and to reward data sharing.



### 3

## Microbial Forensics and Clinical and Public Health Considerations: Commonalities and Differences

This chapter explores the commonalities and differences between the needs for microbial forensics on the one hand and for clinical diagnostics and public health protection on the other in terms of purposes, procedures, technologies, and best practices. The answers to these questions touch on key goals and objectives, sample type, level of characterization required, interpretation and reporting of differences, and differences in governance and associated requirements.

Dr. Stephen Morse of the U.S. CDC outlined what public health needs to know in the event of an infectious disease outbreak due to either natural causes or a biological attack:

- The identity of the agent,
- Whether it has antimicrobial susceptibilities,
- The method of dispersal or dissemination,
- The type of preparation (was it weaponized?) and potential for reaerosolization,
- Who was potentially exposed,
- The area involved and the extent of contamination,
- Whether to advise those in the affected area to evacuate or shelter in place, and
- Whether other factors (e.g., safety of food and water supplies) need to be addressed.

Professor Alemka Markotić, University Hospital for Infectious Diseases, Zagreb, an associate member of the Croatian Academy of Sciences and Arts, also emphasized that the distinction between natural and deliberate infectious disease outbreaks is very important. The Biological Weapons Convention (BWC) does not apply to natural outbreaks of disease. Therefore, it is essential to determine as quickly as possible if the origins of an outbreak are natural or deliberate. Any State Party to the BWC may request assistance when a suspicious outbreak of disease has occurred (see Chapter 1). Various evaluation criteria have been developed to assist in assessing whether an outbreak is suspicious (see, e.g., Box 3-1). An interesting paper by Grunow and Finke (2002) applied a model using many of these criteria to an epidemic of tularemia in Kosovo in 1999 and 2000. Grunow and Finke determined that the outbreak was the result of natural causes. But the procedure used was retrospective and would not have been in time for an intervention to prevent further spread of the disease if the cause had been due to a deliberate act.

Dr. Dana Kadavy of Signature Science, LLC, noted some general differences between the procedures used for clinical diagnostics and public health versus microbial forensics:

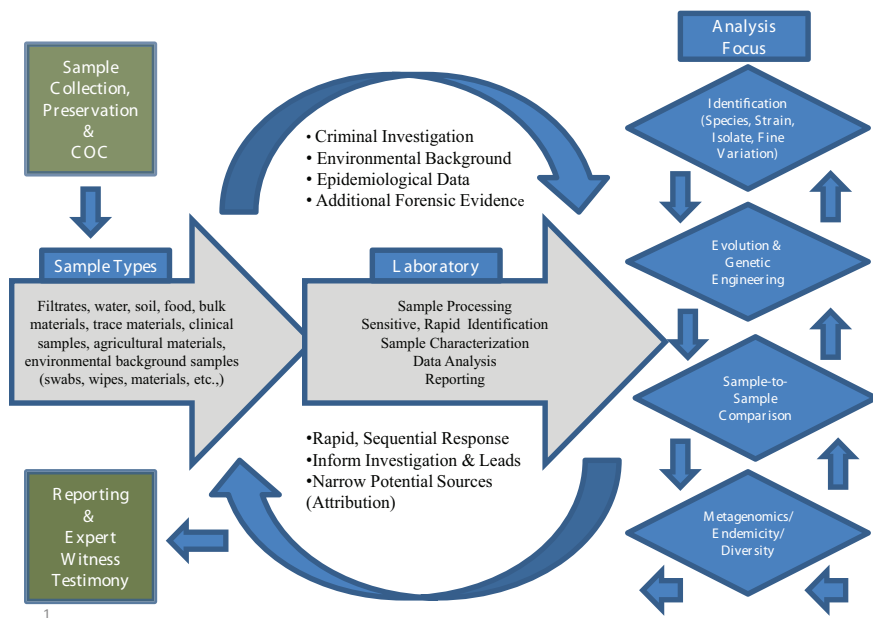
- Clinical diagnostic analyses are designed to identify rapidly the species and possibly the strain of an infectious organism to inform effective treatment strategies. Microbial forensics laboratories generate more fine-grained analyses of evidentiary samples, looking not only at species and strains but also perhaps at isolates of strains that possess unique genetic markers. Such in-depth analyses are likely to take much more time and effort (Figure 3-1).
- For microbial forensics, data must hold up to the scrutiny of judges and juries in a court of law as well as political figures, the media, and the public. Standards are defined for human DNA analysis, and diagnostic data are generated in conformity with Certified Authorization Professional (CAP), Clinical Laboratory Improvement Amendments (CLIA), or other certification standards. However, such guidance is less well defined for microbial forensics.
- In the clinical and public health sectors in the United States, governance lies with the American Medical Association and the Food and Drug Administration. For microbial forensics, governance lies in law enforcement, the judicial system, and perhaps the U.S. Department of Defense.

Microbial forensics may encompass evidence from environmental background, epidemiological data, and other forensic findings. Collec-

**BOX 3-1**  
**Epidemiological Clues That May Signal a  
Biological or Chemical Terrorist Attack**

1. Single case of disease caused by an uncommon agent (e.g., glanders, small-pox, viral hemorrhagic fever, inhalation, or cutaneous anthrax) without adequate epidemiological explanation;
2. Unusual, atypical, genetically engineered, or antiquated strain of agent (or antibiotic resistance pattern);
3. Higher morbidity and mortality in association with a common disease or syndrome or failure of such patients to respond to usual therapy;
4. Unusual disease presentation (e.g., inhalation anthrax or pneumonic plague);
5. Disease with an unusual geographic or seasonal distribution (e.g., plague in a nonendemic area, influenza in the summer);
6. Stable endemic disease with an unexplained increase in incidence (e.g., tularemia, plague);
7. Atypical disease transmission through aerosols, food, or water in a mode suggesting sabotage (i.e., no other possible physical explanation);
8. No illness in persons who are not exposed to common ventilation systems (have separate closed ventilation systems) when illness is seen in persons in close proximity who have a common ventilation system;
9. Several unusual or unexplained diseases coexisting in the same patient without any other explanation;
10. Unusual illness that affects a large, disparate population (e.g., respiratory disease in a large heterogeneous population may suggest exposure to an inhaled pathogen or chemical agent);
11. Illness that is unusual (or atypical) for a given population or age group (e.g., outbreak of measles-like rash in adults);
12. Unusual pattern of death or illness among animals (which may be unexplained or attributed to an agent of bioterrorism) that precedes or accompanies illness or death in humans;
13. Unusual pattern of death or illness in humans that precedes or accompanies illness or death in animals (which may be unexplained or attributed to an agent of bioterrorism);
14. Significant number of ill persons who seek treatment at about the same time (point source with compressed epidemic curve);
15. Similar genetic type among agents isolated from temporally or spatially distinct sources;
16. Simultaneous clusters of similar illness in noncontiguous areas, domestic or foreign;
17. Large numbers of cases or unexplained diseases or deaths;

SOURCE: Reprinted from Khan and Pesik (2011). Copyright 2011, with permission from Elsevier.



**FIGURE 3-1** The microbial forensic process.

SOURCE: Budowle et al., 2013.

tively these aspects serve to help frame the forensics work required, and provide understanding of, and goals for, the types of laboratory analyses to be performed. The range of potential sample types that may be encountered in microbial forensics is vast. Clinical samples also can be highly complex, but the sample types generally are known and their matrices are well characterized. In microbial forensics, samples may come not only from clinical sources, but also environmental, food, and other matrices or be a completely unexpected target, as in a hoax. Sample types range from the simple to the complex, and across a clinical, clinical/forensic, and forensic continuum. They may comprise pure isolates, a prominent isolate in a minor mixture, a trace isolate in a mix of hundreds of organisms, and every other permutation. Samples also may contain copious amounts of a target organism or just a trace of the sought-for (or unknown) organism, which represents an analytical target against a background of sample “noise.” These challenges necessitate alternative sample-processing strategies.

Laboratory analyses for microbial forensics are intended to (1) serve as rapid screening mechanisms, (2) provide information for investigative leads, and (3) narrow the potential sources for a particular microorganism

in order to link it to a particular origin or activity for attribution. Microbial forensic methods need the capability to drill deeply and with very fine variances to fully assign attribution. Analysis extends to strain, subtype, and/or isolate; type and abundance of organism present in simple to complex samples; presence of antibiotic resistance or virulence genes; evidence of genetic engineering and/or isolate evolution (is it endemic, wild type, or a cultured strain repeatedly passed around labs?); and in-depth, sample-to-sample comparison that may be informed by SNPs and other genetic markers. Some of these are also major concerns for public health, for example, the presence of antibiotic resistance or virulence genes, which may dictate differences in treatment options. Reporting requirements for characterization also are specific for the microbial forensics space, and some of these are shared with clinical diagnostic practices and food analysis systems. In such cases, the sectors should leverage one another. The data generated may inform investigative leads and requests for additional samples, and will be used in reporting and for expert testimony. Documentation and reporting must be in place at each step.

The following case studies will illustrate in practical terms how some analytical approaches are shared between clinical medicine or public health and the use of microbial forensics to deal with biocrime.

### **THE RAJNEESHEE CULT IN OREGON, *E. COLI* O104, AND *KLEBSIELLA PNEUMONIAE*: THREE CASE STUDIES OF RESPONSES TO INFECTIOUS DISEASE OUTBREAKS**

In 1984, the Rajneeshee cult deliberately contaminated restaurant salad bars in The Dalles, Oregon, with *Salmonella enterica*. The motivation for this biocrime was to influence the outcome of a local election by incapacitating the local populace to ensure that the cult's candidates would win the county elections. More than 750 people were sickened, but fortunately there were no fatalities. But the deliberate nature of the attack was not confirmed until nearly a year later.

In this instance, the agent and context of the attack mimicked a naturally occurring event that plays out several times a year across the United States, with large, multifocal outbreaks such as the one that plagued the tomato and pepper supplies in 2008-2009, causing disease in more than 1400 people. (Skowronski and Lipkin, 2011:174)

This example of a biocrime (a more frequent category of intentional use of biological agents than bioterrorism; see Carus, 2001) highlights several important points. First, it can be a major challenge to differentiate among natural, accidental, or deliberately caused disease outbreaks. Second, such an outbreak requires similar responses to protect public

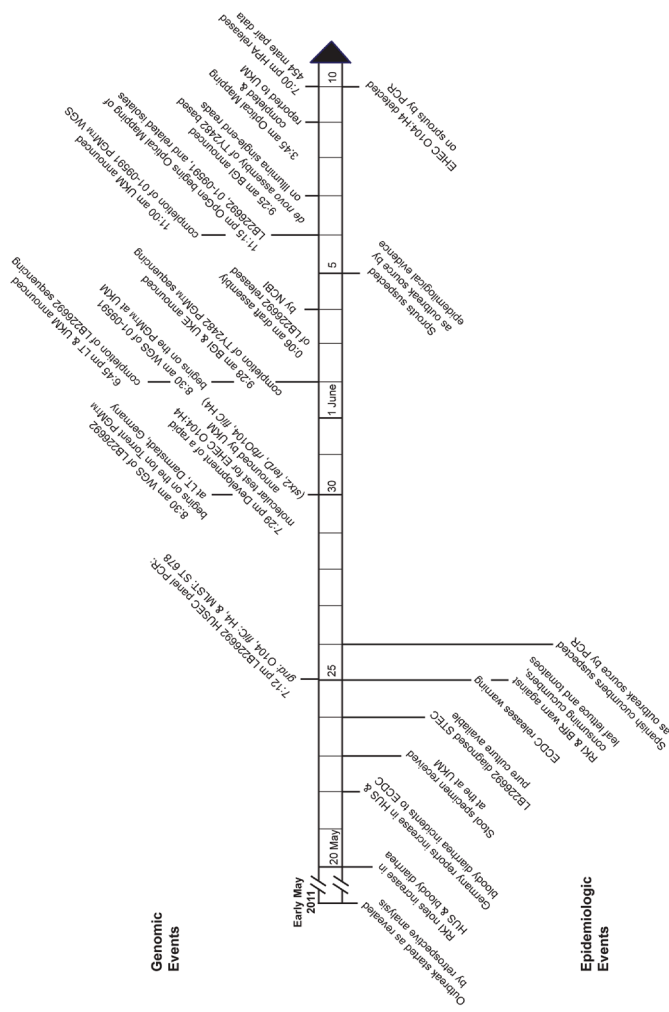


health regardless of the cause. Finally, modern technology makes it possible to determine the identity and characteristics of the organism being dealt with much more rapidly, as the next two cases dealing with natural outbreaks illustrate. The first case is one that also originated in the consumption of food from a salad bar, this time in Germany in 2011.

*Escherichia coli*, or *E. coli*, is usually a harmless bacterium that grows naturally in the intestinal tracts of humans and other animals. However, there are many different strains of *E. coli*, just as there are different strains of *B. anthracis* and *Y. pestis*. Some of these strains are highly pathogenic, such as the infamous O157:H7 and O104:H4 strains. The letter O in the name refers to a marker on the bacterium's surface, which presents in hundreds of different immunogenic forms, in the latter instance, form number 104. The H describes another immunologic marker found on the bacterium's flagellum. Both of these markers are easily detected by specific antibodies and can be used to identify strains. Pathogenic *E. coli* must be able to attach to or invade the cells of the intestinal lining, where they disrupt the normal function of the intestine by producing toxins, such as the Shiga-like toxin produced by the O157:H7 strain (AAM, 2011b). This toxin is very similar to the toxin produced by *Shigella dysenteriae* and is able to kill host cells. Both *S. dysenteriae* and *E. coli* that produce Shiga-like toxin are able to cause hemolytic uremic syndrome (HUS) in infected individuals. Some pathogenic *E. coli* may have extra genes not found in the normal intestinal strains that confer these abilities.

Dr. Dag Harmsen of University Hospital, Munster, Germany, described the recent *E. coli* O104:H4 outbreak in Germany and a *Klebsiella pneumoniae* outbreak in the Netherlands to illustrate the usefulness of next-generation gene sequencing (NGS) in public health responses. NGS is becoming both an essential component of microbial forensic analyses and the new preferred standard in epidemiology. For example, the *E. coli* O104:H4 outbreak has become a textbook case on technology development and microbial surveillance (Lipkin, 2013). It ushered in a new kind of epidemiological investigation and was directly instrumental in the development and evaluation of several new bioinformatics tools. It also provided proof of principle for the capability of smaller sequencing machines, which are reasonably affordable and, perhaps more importantly, allow rapid turnaround times.

As can be seen in Figure 3-2, which provides the timeline of the case, the O104 *E. coli* outbreak began in early May 2011 in the Hamburg area of Germany. In contrast with previous outbreaks, the incubation period was long, approximately 10 days. The outbreak was not declared in Germany until May 20th. Two days later, the European Centre for Disease Prevention and Control issued a warning for Europe. Labs were instructed to send isolates to the University of Münster, and the first samples arrived



**FIGURE 3-2** Events timeline of enterohemorrhagic *E. coli* O104:H4 outbreak. BfR, Bundesinstitut für Risikobewertung (Federal Institute of Risk Assessment, Germany); BGI, Beijing Genomics Institute (People's Republic of China); EDC, European Centre for Disease Prevention and Control (Sweden); HPA, Health Protection Agency (United Kingdom); HUS, hemolytic uremic syndrome; LT, Life Technologies Group; PGM™, Ion Torrent Personal Genome Machine™; RKI, Robert Koch Institute (Germany); UKM, University Hospital Muenster (Germany). SOURCE: Mellmann et al. (2011).

on May 22. After growing a pure culture, the organism was characterized and a screening test was created. This screening test, based on Sanger sequencing, was published on the Internet on May 30, and 15 days later in *Lancet Infectious Disease*. Eventually several other groups sequenced the *E. coli* so that its genome was eventually produced on all available NGS platforms. Several peer-reviewed papers published the sequence very quickly.

Phylogenetic analysis based on Sanger multilocus sequence typing (MLST) showed that the closest related genome was enteroaggregative *E. coli*, which had previously been sequenced by a French group (Mellmann et al., 2011). Harmsen pointed out that routine surveillance is very important. The HUSECO41 *E. coli* strain was identified in a 2008 publication (see Mellmann et al., 2008), yet little notice was taken, perhaps because this organism rarely causes infections. Using the sequence data, the MLST technique showed convincingly that this previous isolate is indeed most closely related to *E. coli* O104:H4. Although Harmsen's group and others have proposed differing hypotheses about *E. coli* O104:H4's evolutionary descent, there is as yet no resolution of this question.

Although there had already been an outbreak in Japan related to bean sprouts, it was not until June 5 that sprouts were suspected as the source. Unfortunately, most people remember eating a salad but will not remember all of its ingredients. This deficiency provides a very strong argument for changing the way epidemiological investigations are conducted, because the goal is to elucidate and contain the outbreak. It takes considerable time to identify a single contaminated component of a mixed salad and there are many false leads because microorganisms and/or toxins can be transferred from one component to another. There are numerous examples where one component of a salad was alleged to be the culprit only to find out later that something else was responsible. In the current case, two clusters of affected people had eaten at the same restaurants, which proved to be the epidemiological breakthrough. Based on this information, a very detailed food supply chain analysis was performed, and that is how the sprouts were finally implicated. Although no isolate had been recovered from sprouts in the German case as of October 2013, the epidemiological investigation confirmed convincingly that sprouts were involved. There was also a smaller outbreak in Marseilles, France, linked to sprouts. In the German case, once the sprouts were identified, the outbreak was essentially over.

Hospitals need fast, sensitive, and specific screening tests. In hospital surveillance terms, the ability to exclude pathogens is very important owing to response and cost. If a hospital can exclude the presence of an outbreak, it can avoid closing wards and putting patients into isolation, which are very costly measures.

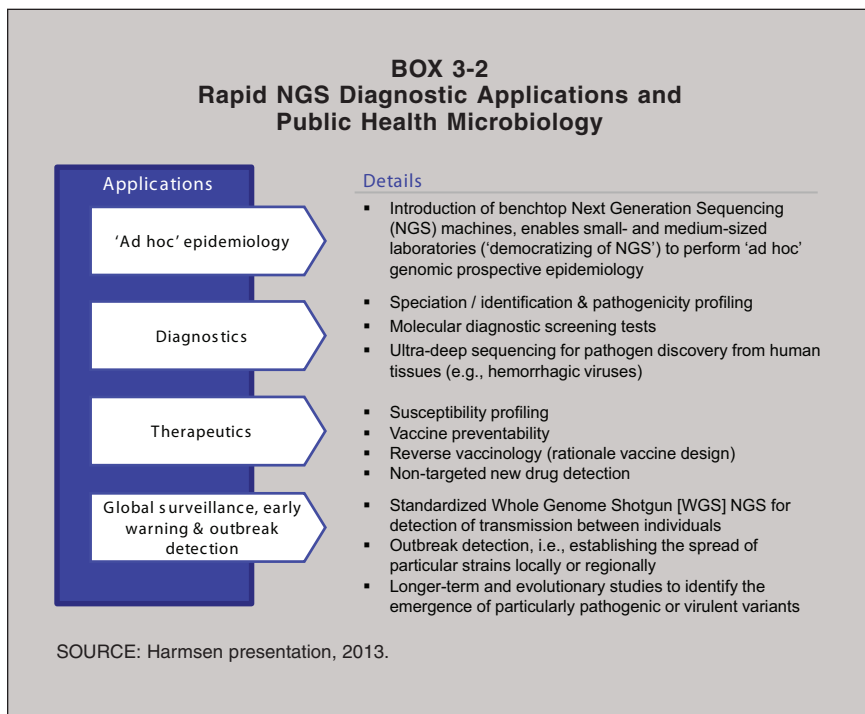
According to Harmsen, while the laboratories and treatment provid-

ers were well prepared to confront the outbreak, communication with the public was not handled as well. For example, an accusation made early in the outbreak implicating Spanish cucumbers came from a local food agency in Hamburg. This agency used screening tests that were not specific for the O104 strain and the Spanish cucumbers were wrongly implicated as the source. Spain sustained substantial economic losses and sought compensation. Communication among the national, state, and local authorities was not always consistent, nor was there agreement between public health and food communications groups. Intense media attention is common in such situations and is not always helpful. Harmsen noted that he himself was misquoted during interviews. He suggests that there should be centralization of information for these kinds of communication.

The *E. coli* case demonstrated that rapid NGS can be used almost in real time during outbreaks, and Harmsen described a second example that proved the usefulness of NGS for diagnostics during an outbreak. Only a month after the German *E. coli* case in June 2011, there was an outbreak of *K. pneumoniae* at Dutch Maastad Hospital in Rotterdam. The strain was multidrug-resistant *K. pneumoniae* OXA-48. The outbreak did not receive the global attention that the German outbreak had, but it was an enormous public health issue in the Netherlands.

The Dutch National Institute for Public Health and the Environment (RIVM) sent *K. pneumoniae* strains to the University of Münster to sequence. A draft genome of the *K. pneumoniae* OXA-48 outbreak strain was developed and compared with other publicly available *Klebsiella* genomes. Scientists identified 36 candidate regions to use in developing a strain-specific multiplex PCR test. They enlisted the help of the Wellcome Trust Sanger Institute in Cambridge in the United Kingdom, which was conducting a global surveillance of *Klebsiella*. By comparing the candidate signature sequences against Sanger's additional 200 *Klebsiella* genomes, they identified two candidate regions that were specific for the Dutch outbreak (Netherlands National Institute for Public Health and the Environment, 2013). This information was given to RIVM, and a multiplex molecular diagnostic test assay that targeted one of the two signatures as well as antibiotic resistance genes was developed. This test assay was supplied to every Dutch hospital and is still used today for screening patients, mainly for exclusion purposes. The case is an excellent example of the important role of genomics in diagnostics as well as microbial forensics.

As detailed in Box 3-2, clinical microbiologists see many potential applications for NGS that can be broadly organized under these categories: (1) ad hoc epidemiology; (2) diagnostics; (3) therapeutics; and (4) global surveillance, early warning, and outbreak detection. Benchtop NGS is a democratizing force, enabling small- and medium-sized labora-



tories to acquire these powerful new diagnostics, and some believe this technology can “leapfrog” into developing countries. NGS can be used for diagnostic and screening tests, and in therapeutics it has been used for susceptibility profiling. Some of the very early genome sequencing was performed for reverse vaccinology for the design of new vaccines, and NGS is now being used for drug development. It seems likely that many of the problems with sequencing will eventually be overcome, and its use could become fairly routine in clinical diagnostics. If this is the case, stored clinical data can also be analyzed for microbial forensics purposes if mechanisms for forensic analysts to access the clinical data can be devised.

Clinical microbiologists are particularly interested in global surveillance for use in detecting early outbreaks. To maximize surveillance data, a kind of molecular “Rosetta stone” is needed—a nomenclature that will enable global comparisons of data.

Harmsen also suggested that another ultimate goal should be to develop plain-language reporting that is understandable to physicians and epidemiologists. Translating “there is a SNP at position 4,000,000” to

## Whole Genome Sequence Typing

Pairwise Genome Alignment	<ul style="list-style-type: none"> <li>- Impossible with draft genomes</li> <li>- Problematic with rearrangement / recombination events</li> <li>- Not additive expandable</li> </ul>
K-mer	<ul style="list-style-type: none"> <li>+ No assembly needed, quickly find closest matching genome</li> <li>- Whole genome reduced to a single number of similarity</li> <li>- Not additive expandable</li> </ul>
Genome-wide SNP	<ul style="list-style-type: none"> <li>+ Works well for monomorphic organisms and 'ad hoc' analysis</li> <li>- Problematic with rearrangement / recombination events</li> <li>- Not additive expandable</li> </ul>
Genome-wide gene by gene allele typing	<ul style="list-style-type: none"> <li>+ Works on gene level, i.e. the element of evolution</li> <li>+ Rearrangement / recombination treated as single event</li> <li>+ Additive expandable when using a 'core genome'</li> </ul>

FIGURE 3-3 Four approaches to using whole-genome sequencing. Pros (green), cons (red).

SOURCE: Adapted from Harmsen presentation, 2013.

“this confers resistance to tetracycline” greatly improves usefulness. For certain bacteria, plain-language reporting could be easily achieved, and several groups are already working to accomplish this.

Figure 3-3 summarizes Harmsen’s thoughts on four approaches for using whole-genome nomenclature. Genome-wide SNPs work especially well for monomorphic organisms, but Harmsen does not think it is the best choice. He agrees with Maiden et al. (2013) in advocating genome-wide, gene-by-gene allele typing. Although MLST is not popular for use with monomorphic organisms owing to its limited discriminatory power, it is still useful for other organisms and is influential from an intellectual point of view. Harmsen suggests using core-genome<sup>1</sup> MLST, or using MLST+, which enables analysis of the seven “housekeeping” genes<sup>2</sup> as

<sup>1</sup> The core genome is the set of genes that are present in all members of a species, suggesting that they are required for essential cellular functions.

<sup>2</sup> Housekeeping genes are constitutive genes that are required for the maintenance of basic cellular function.

well as hundreds or thousands of other genes, which will provide the discriminatory power needed for outbreak investigation for most organisms (Jolley et al., 2012). There are also the benefits of it being additive, expandable, and nomenclature friendly.

NGS is a technology for analyzing all bacteria, whether performing evolutionary or phylogenetic analysis. It could enable standardized hierarchical microbial typing, surveillance and outbreak investigation, evolutionary analysis, and resistome/toxome<sup>3</sup> analysis. Harmsen ranked the discriminatory power of microbial typing approaches in Figure 3-4. He believes that MLST will endure for backwards compatibility; it has a legacy of 15 years of publication and data generation and it can be extracted easily from NGS data. Canonical SNPs offer the same range, if not better, of discriminatory powers. Martin Maiden recently proposed a system called ribosomal MLST (rMLST) (Jolley et al., 2012) to implement a combined taxonomic and typing approach for the whole domain of bacteria. Harmsen believes that MLST+ ranks high for discriminatory power and that it can be standardized. Using SNPs and alleles would offer even more discriminatory power (Köser et al., 2012), but standardizing the method will be challenging.

Harmsen also sees a need for quality assurance and quality control guidelines for microbiology and microbial forensics. Published guidelines are primarily for use in human genetics. There is an initiative called Global Microbial Identifiers<sup>4</sup> that is trying to address this and harmonize, but it is not well funded. He believes that a One Health approach,<sup>5</sup> which links the communities concerned with food safety, veterinary medicine, clinical medicine, and microbial forensics, should be implemented, both within and among countries. Many nations may assume such collaborations are already in place, but during an actual event, difficulties will arise.

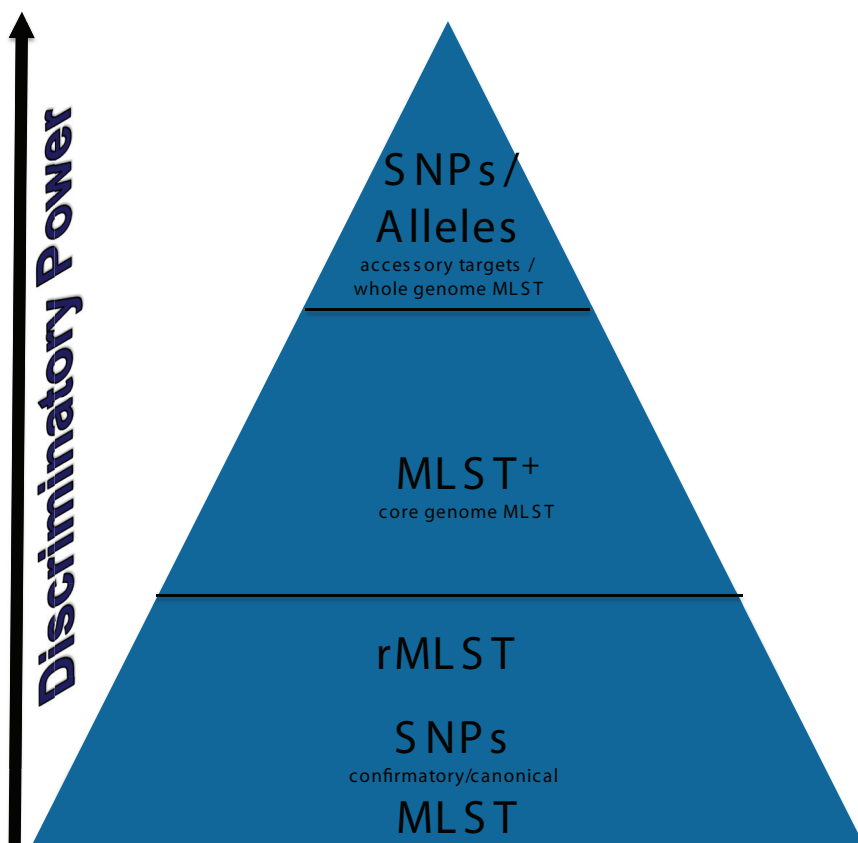
From Harmsen's perspective, the future of NGS looks bright. Phenotyping based on genotyping should become a reality. Organisms such as tuberculosis and methicillin-resistant *Staphylococcus aureus* (MRSA) would be particularly appropriate for this analysis. These capabilities should be coupled with early-warning and geographic information systems.

---

<sup>3</sup> The "resistome" is the collection of antibiotic resistance genes and their precursors in bacteria (Wright, 2007). The "toxome" is the collection of toxicity pathways, most of which are only partly known (Hartung, 2011).

<sup>4</sup> For more information, see <http://www.globalmicrobialidentifier.org/>; accessed October 24, 2013.

<sup>5</sup> The One Health Initiative is a movement to forge all inclusive collaborations between physicians, veterinarians, nurses, and other scientific-health and environmentally related disciplines. More information is available at <http://www.onehealthinitiative.com/>.



**FIGURE 3-4** A ranking of the discriminatory power of microbial typing approaches, from bottom to top, with increasing discriminatory power. MLST, multilocus sequence typing; MLST<sup>+</sup>, core-genome MLST; rMLST, ribosomal MLST. SOURCE: Harmsen presentation, 2013.

### HEPATITIS C IN SPAIN: A PUBLIC HEALTH PROBLEM THAT BECAME A LAW ENFORCEMENT ISSUE

Dr. Fernando González-Candelas, University of Valencia, Spain, described how molecular technology and phylogenetics were used successfully in court to convict an anesthetist of a biocrime involving the infection of multiple patients with hepatitis C virus (HCV). González-Candelas described how phylogenetic inference and coalescent theory



were employed to establish association between the presumed source and those that were infected.

González-Candelas explained that by early 1998, the Spanish public health authorities had noted a steady rise in HCV cases in Valencia since 1994. The cases appeared to be unrelated. Because intravenous (IV) drug use was on the rise and viruses are transmitted easily by blood, the increase was at first attributed to drug use. But in early 1998, a physician reported that four of his patients who had been sent to a private hospital for minor surgery had tested positive for HCV a few weeks afterward. The hospital stay seemed to be the only common denominator. Public health officials launched an epidemiological investigation, analyzing 66,000 surgical patients in two hospitals. They concluded that an anesthetist working at the private hospital was the likely common source. A search for other potentially infected patients revealed a sizeable outbreak of HCV that appeared to be clearly linked to this one medical professional. The anesthetist was ultimately convicted for being responsible for the infection of 275 of patients with HCV.

Both the hospital and the anesthetist were sued. Public health authorities and the judge heading the respective epidemiological and judicial investigations requested that González-Candelas and his colleagues use their expertise in evolutionary biology to specifically ascertain

- Was there an outbreak?
- What was the source?
- Can other sources be excluded?
- Which patients were included in the outbreak and which were not?
- When did it start and how long did it last?
- When was each patient infected?
- When was the index case infected? (González-Candelas et al., 2013)

The question “when was each patient infected?” was highly relevant to the judge and multiple insurance companies who had provided malpractice insurance coverage to the anesthetist during his working years; the time of transmission could determine liability. In addition, the anesthetist contended that he was a simply another victim.

Many people who are infected with HCV remain without symptoms and are unaware of their infection yet still can transmit the virus. Of those infected, 15-30 percent experience spontaneous clearance of the virus, and 70-85 percent become chronic carriers, of whom 25 percent do not experience a progression to disease. But HCV can be silent for many years. Of those in whom illness progresses, a process that can take 10-30 years, out-

comes can include cirrhosis, end-stage liver disease, and hepatocellular carcinoma (Bowen and Walker, 2005).

The HCV genome is organized as a positive sense *single-stranded* RNA virus and has a high mutation rate, similar to that of HIV. It evolves quickly—one million times faster than human DNA—but variability is not equally distributed across the genome, and some areas evolve much faster than others. The rapid evolution is due to high mutation rates, high virus production (approximately  $10^{10}$  viral particles/day in an infected patient), and a short generation time. Cell-infection cycles occur within a few hours (Penin et al., 2004).

The background prevalence of HCV in Spain is about 2.5 percent, which is relatively high, and many of those infected are unaware that they are. Evolution can occur quickly within an infected individual, and over years of infection, the number of viruses that accumulate in that individual can be substantial. In addition, there are reports of compartmentalization of HCV in individuals; various organs and tissues in the same individual can be infected by slightly different and divergent populations of the virus (Di Liberto et al., 2006). For this reason, one cannot expect to find the same virus at the same time throughout the host. In addition, although transmission between infected patients is horizontal (i.e., from person A to person B), the viruses are related through vertical processes. Whenever there is an infection, there is a strong “bottlenecking” of the infections passed on, such that patients infected by the same source may receive slightly different samples of the virus. Therefore one cannot expect to see exactly the same viral genome in the source and all patients infected by the source.

González-Candelas and his team sequenced about 134 clones for each patient who agreed to supply a sample. This limited sampling is why consideration of compartmentalization is so important. He gave an example of patients co-infected with two HCV genotypes, who, depending on the day of sampling, showed only one of the two genotypes, alternating from one day to another. The dynamics of viruses within an individual’s blood are largely unknown, but are more complicated than one simple representation of a viral population. There are many populations, and the longer a person is infected the more opportunity exists for minor differences. Most differences appearing in the widely divergent clade shared by the patients in the Valencia HCV outbreak probably represented sampling from the same source at different times over 10 years. The source would transmit a slightly different virus, but one still recognizable as a sample of the source’s viral diversity; it was still possible to determine the common ancestor. González-Candelas proposed that if one can find the common ancestor of a sample, compared with the common ancestor of

the reference population, one is in a good position to make all additional inferences.

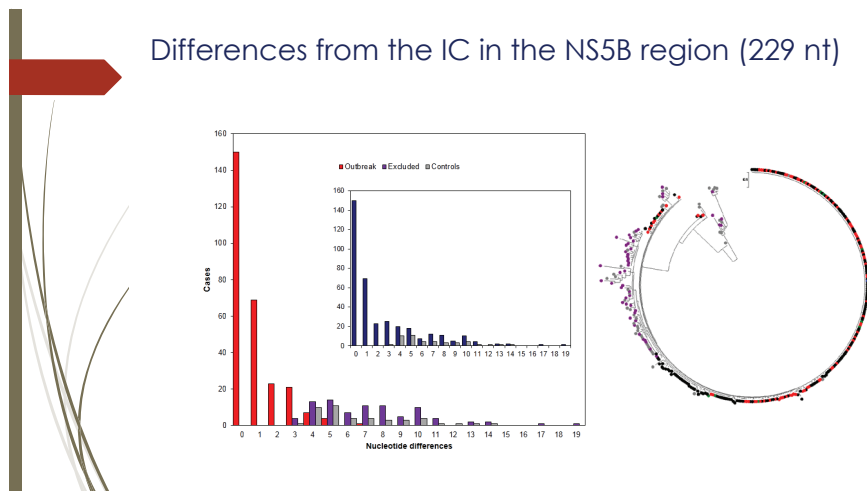
For the court case, phylogenetic inference and coalescent theory<sup>6</sup> were used to analyze the outbreak. González-Candelas believes that these approaches should be incorporated into microbial forensics.

In 1998, González-Candelas and colleagues were using gel-based Sanger sequencing. They decided to use two approaches to analyze the clinical viral samples: (1) direct sequencing of PCR products generated with primers directed to a relatively slow-evolving polymerase region of the HCV genome (NS5B), and (2) sequencing of PCR products derived from the faster-evolving regions (E1 and E2). For the latter, they analyzed several clones from multiple victims to estimate viral diversity in each individual (González-Candelas et al., 2013). They also cloned sample sequences from local population controls, which were not related to the case based on epidemiological determinations. In analyzing 300+ samples, it was also possible that some were positive for HCV but were not related to the source despite the fact that they had either been anesthetized by the alleged source or there were other circumstances that could have made him the presumed source of their infection.

Analyzing a single gene from the virus samples, however, did not provide all the necessary information. The graph in Figure 3-5 shows the distribution of differences when compared with the sequence in the NS5B region (229 nucleotide [nt]) of each patient isolate with that from the presumed source. There were many identical sequences, which appeared to be the “smoking gun” that linked about 150 patients to the anesthetist. The scientists, however, wanted to know how they should characterize possible victims who show a single nucleotide difference between their

---

<sup>6</sup> Coalescent theory and phylogenetic inference: Phylogenetic inference uses the genetic variance between members of a population to infer evolutionary relationships; at the level of specific genes, alleles are used to create trees to represent how members split into separate branches, converge, or become extinct. Coalescent theory, on the other hand, seeks to analyze all known alleles of a specific gene to identify the most recent common ancestor (MRCA). In its simplest form, coalescent theory does not consider evolutionary pressure, recombination, or other interactions with the environment. If one follows the ancestry of two haploid organisms, eventually a single organism (the MRCA) will be identified that gave rise to the two lineages and the two organisms will have coalesced. These analyses can be used to predict the time of appearance of the alleles, as a sort of molecular clock, by converting the degree of change in a sequence to a specific interval of time. Gonzalez-Candelas and his colleagues were the first to use these two approaches in understanding origin and relatedness in the context of a criminal trial. Problems remain with juror perception of the validity of these kinds of data, in view of the complexity of the concepts and the tendency to accept all DNA-based evidence as incontrovertible; validation of the approach for use in criminal court remains to be performed. Nevertheless, the Valencia case highlights the value of using phylogenetics in conjunction with traditional epidemiological data when building a case (Vandamme and Pybus, 2013).



**FIGURE 3-5** Differences from the Index Case in the NS5B region (229 nt). The bar graph (left) shows the distribution of differences in the NS5B region. Sequences were compared with a 229-nt fragment of the NS5B gene derived from the presumed source. The graph shows the distribution of nucleotide differences (Hamming's distance) for sequences derived from patients included in the outbreak (red bars), from patients excluded (dark purple) from the outbreak, and from local controls (gray). The bar graph inset shows the same distribution for putative outbreak samples (dark blue) and local controls (gray) before the former were divided into included and excluded from the outbreak. The neighbor-joining tree (right) was obtained with the NS5B-region sequences of hepatitis C virus (HCV)-1a samples analyzed. Color codes: outbreak sequences are in black, red, and green, excluded from the outbreak are in dark purple, and local unrelated controls are in gray. The presumed source (PS) sequence is shown in blue. No clade was found with bootstrap support higher than 70 percent. SOURCE: González-Candelas et al. (2013).

virus and that of the presumed source—a circumstance that occurs commonly in virology. The court decided that those showing one difference could also be considered victims, which increased the number of potential victims to 200+. But, the scientists asked, if one difference is accepted, why not two differences? They had controls with as few as three differences from the source. Being concerned about false negatives, they asked whether they should use this number as a threshold defining those with three or more differences as not part of the outbreak.

The second problem with NS5B phylogeny is illustrated by the

neighbor-joining tree analysis in Figure 3.5 (right), which failed to group all of the control samples in a monophyletic group. Furthermore, none of the nodes in this tree receive bootstrap support higher than 70 percent by either method. Analysis with the slowly evolving NS5B region was unable to discriminate among the three epidemiological groups: local controls, patients infected by a common source, and patients infected by alternative sources. Therefore the emphasis was shifted to the more rapidly evolving E1-E2 regions.

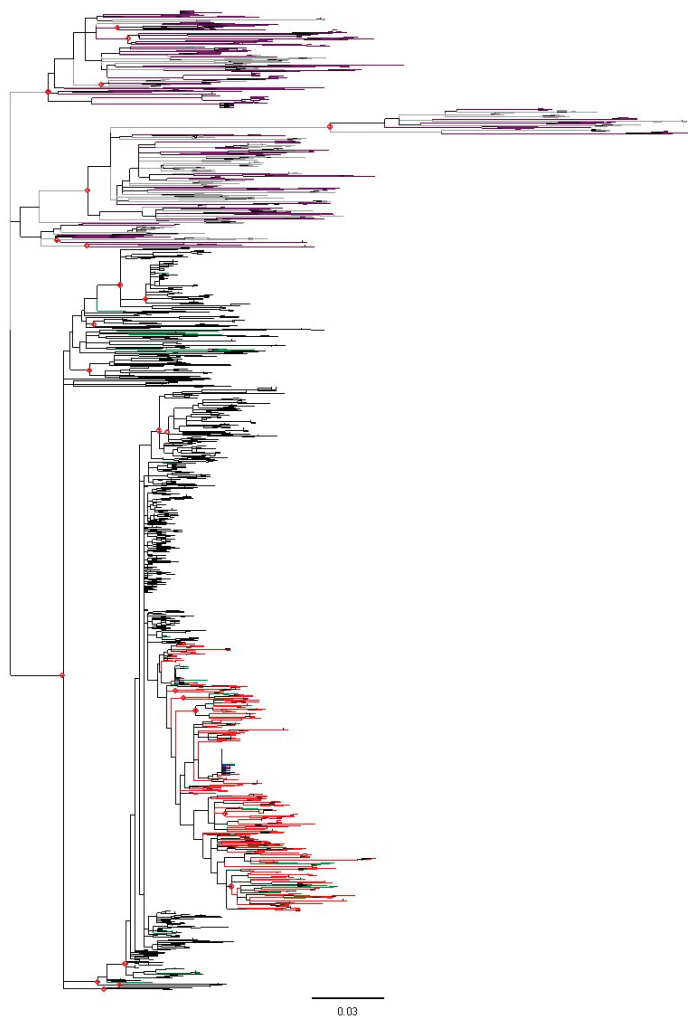
Sequencing of the E1-E2 region was carried out for a total of 4,184 cloned PCR fragments, representing 134 clones sequenced from the presumed source, 321 samples initially considered part of the outbreak, and 42 local controls; the average number of clones from each sample (excluding the presumed source<sup>7</sup>) was 10.77. Multiple alignments were performed with the 4,184 sequences to derive neighbor-joining and maximum likelihood phylogenetic trees for the 4,184 cloned sequences, shown in Figure 3-6. The conclusion of this approach was that sequences from 274 patients were grouped with the sequences from the presumed source, while the second group included all the sequences derived from the local controls and sequences from 47 patients initially considered to belong to the outbreak. Thus the separation between the two groups became clear.

However, while they could show the global analysis of the phylogeny, the data could not represent a forensic conclusion, because forensics normally requires an individual analysis. This requirement had not previously been applied to an epidemiological analysis. They adapted a statistical framework for forensic analysis to molecular epidemiology (González-Candelas et al., 2013). They used Evett and Weir's (1998) method for applying genetic analyses to forensic settings, using a Bayesian framework. Evett and Weir propose that scientific experts limit their contributions to forensic analyses of genetic information to the evaluation of available genetic data in light of two competing hypotheses, those of the prosecution and the defense. The forensic expert should present the likelihood ratio: what is the likelihood of the genetic information seen in light of two mutually exclusive hypotheses combined with other types of evidence? The police or other investigators would provide additional independent evidence.

It is up to the judge or jury to evaluate and integrate this information with the information provided by other types of evidence, ideally in a numerical way. Molecular phylogenetics provides a logical way to provide likelihood ratios or probabilities. In this case, there were two large groups—the outbreak and the non-outbreak. Based on what was

---

<sup>7</sup> Under Spanish law, only a single sample is required to be provided by a suspect.



**FIGURE 3-6** Maximum likelihood tree for cloned sequences in the E1-E2 region. The tree includes 4,184 sequences from a 406-nt fragment of the E1-E2 region including hypervariable region (HVR)1 and HVR2. Sequences were obtained from patients included in the outbreak (274 patients, 3,038 sequences), patients excluded from the outbreak (47 and 559, dark purple), local controls (42 and 453, gray), and the presumed source (PS, 134 sequences, dark blue). Sequences and branches in the monophyletic clade defined by all the cloned sequences from the PS are labeled in red. Sequences from polyphyletic samples with some representatives in the clade delimited by the PS sequences and others outside it are labeled in green. Relevant nodes with bootstrap support larger than 90 percent are indicated by red dots.

SOURCE: González-Candelas et al. (2013).

being tested, the results were either tests of the defense or the prosecution hypothesis.

Eventually, González-Candelas and his team answered all of the questions posed by the court. There was an outbreak, and the source was a practicing anesthetist. Other potential sources could be discarded; molecular surveillance of HCV has never turned up another case with samples from the clade of the outbreak samples. No other common source of infection was found. They identified 275 individuals in the outbreak and excluded 47. The outbreak began at the end of 1988, and ran until 1998, intensifying from 1996 onward. The presumed source had been infected for 10 years, and this infection was prior to the time of his patients' infection. Two-thirds of the estimates for the time of infection coincided with those offered by the prosecution. The anesthetist was convicted of professional malpractice and was sentenced to a lengthy jail term.

For moving forward, González-Candelas' conclusions are

1. Molecular phylogenetics and coalescent theory are essential for microbial forensics.
2. A meaningful statistical treatment (maximum likelihood, Bayesian inference) is mandatory.
3. In particular, this case shows that recent developments in evolutionary biology can be used to estimate dates and places of relevant events.
4. A good sample of a reference population is absolutely essential to draw any conclusion about the origin of a set of samples. This may change from one case to another.
5. The forensic expert is one among others: his/her conclusions have to be evaluated in the appropriate context.
6. Although current sequencing methodologies allow us to work with complete genomic information, a lot of work remains to be done to standardize and control laboratory and analytical procedures.

## 4

# Clinical and Forensic Approaches to Microbial Identification

### TECHNOLOGIES FOR MICROBIAL FORENSICS

In this chapter we review both the biological and physical science techniques and methods that are commonly used or have potential for further development for use in microbial forensics. The law enforcement and legal context of microbial forensics and the challenges imposed by this context are emphasized and contrasted with clinical diagnostics in medicine.

Dr. Dana Kadavy, Senior Microbiologist at Signature Science LLC, reviewed the microbial forensic process from systems and technical perspectives, described challenges in microbial forensics compared with the needs of clinical diagnostics and public health, and reviewed technologies applicable to identification of organisms for microbial forensics. These technologies can be roughly grouped into four classes depending on their targets for detection: protein (antibody and toxin) signatures, nucleic acid signatures other than sequencing (e.g., PCR), gene sequencing, and mass spectrometry (IOM/NRC, 2014).

As previously noted, microbial forensic analyses are conducted in the context of a criminal investigation, which shapes the differences between its technologies and/or best practices and those used for clinical diagnostics and public health. The drivers for the analytical differences between microbial forensics and other sectors are legal requirements it must meet regarding (1) sample type, (2) level of characterization, and (3) interpretation and reporting. Sampling and sample types are discussed elsewhere in this report.



In addition, the level of characterization and reporting requirements are specific for the microbial forensics “space.” Some are shared with clinical diagnostic practices and food analysis systems, and Dr. Kadavy suggested that the two sectors should leverage one another. At the same time, microbial forensics must drill very deep for pathogen identification. Analysis extends to strain, subtype, or isolate; type and/or abundance of organism present in simple to complex samples; the presence of antibiotic resistance and virulence genes; evidence of genetic engineering and/or isolate evolution (is it endemic wild type or a cultured strain repeatedly passed around labs?); and in-depth, sample-to-sample comparison that may be informed by SNPs.

When reporting to a court,<sup>1</sup> investigators must have confidence in their results and understand the power and limitations of the procedures they have used, which can only be gained through accumulated experience in testing and validation. Investigators also must consider reasonable alternative explanations. Moreover, they must provide known error rates and detection limits when possible for a judicial audience that is unlikely to include scientists. These factors inform the choice of analytical technologies to be used. Traditional microbiology and culture are the current gold standard in clinical diagnostics, but it is not always possible or practical to culture microbes in a forensic investigation.

A capability needed in microbial forensics is sample processing to achieve desired sensitivity/accuracy in complex sample matrices, such as soil samples from forensic exhumations of graves linked to the Bosnian–Serbian war. Sensitive nucleic acid–based and antibody-based detection technologies also are needed. Although these diagnostic assays may not provide in-depth information, they potentially offer rapid, inexpensive, high-throughput screening to rule in or rule out various samples. Investigators and responders need to know quickly what kind of threat they face; additional in-depth analysis can be performed later.

Traditional microbiology methods and features do have forensic value. These methods include culture, phage sensitivity, staining, and microscopy (all used in the U.S. anthrax letters case), as well as fatty acid analysis and serotyping. It is not always possible or practical, however, to use these on complex, degraded, and/or inactivated samples, or for isolates or mixtures that cannot be cultured.

During the 1980s and 1990s, antibody-based techniques dominated biological agent detection. The “backbone” of antibody-based identification relies on the ability of the immune system to recognize “non-self” components of pathogens, especially antigenic proteins and polysac-

---

<sup>1</sup> Before trial, experts must prepare reports summarizing their analysis and conclusions, and share the reports with all other parties.

charides (Schaudies, 2014). Most antibody-based identification methods (such as enzyme-linked immunosorbent assay, or ELISA) involve immobilizing antibodies for particular antigenic targets on solid substrates. These “capture” antibodies then bind to the antigens produced by the microorganismal cells. This binding is then detected with a second antibody to the same antigen that contains a “reporter molecule” that produces a detectable signal based on optical absorbance or fluorescence. Examples of antibody-based platforms include Luminex, PathSensors, Inc., and TacBioHawk. (See Schaudies, 2014, for a more in-depth review.) The most notable advantage of antibody- or protein-based detection systems is their speed—many of them can provide answers in 3 to 10 minutes. Some of the systems can reach sensitivities approaching that of nucleic acid amplification techniques, such as PCR-based assays. However, for microbial forensic purposes, they often lack sufficient specificity to discriminate beyond the level of species. Platforms that perform protein and antibody/antigen detection, such as electrochemiluminescence (ECL), are also needed to detect protein toxins, such as ricin and botulinum. Schaudies (2014:175) states that antibody-based systems are “more effective and consistent than nucleic-acid-based systems for the detection and identification of toxins.” This statement is based not on efficacy of an assay, but more on the degree of toxin purification; in some toxin preparations, the DNA that codes for production of the toxin may not be present or nucleic acid concentrations may be reduced below limits of detection. Toxin treatment also may degrade DNA from the sample. As with PCR, these platforms are only as good as the assays developed for the intended target. Knowledge of the target being pursued is necessary as are thoughtfully designed and validated assays to use ECL effectively.

Nucleic acid-based detection and identification technologies provide the ability to examine the genetic as well as structural information associated with a pathogen. Among the early tools developed to access an organism’s genetic information were nucleic acid amplification techniques, such as PCR. PCR allows specific fragments of genomic DNA to be isolated and their copy number amplified. Over time, variations of the basic PCR technique have developed, such as “multiplex PCR,” which allows several DNA genomic regions to be amplified in the same reaction set, and quantitative or “real-time” PCR (qPCR), which can measure the quantity of a target sequence in real time. Multiplex PCR technologies do have limitations. Most allow up to five different fluorophores to be detected simultaneously. Other platforms have attempted to increase this to as many as 20 fluorophores simultaneously, which would be a big advance in multiplex capability. Real-time qPCR is a rapid and sensitive nucleic acid signature detection technology that has proved effective in both clinical and microbial forensic applications. Some platforms

have a closed system, which may be fine for some situations. However, a closed system may not be optimal if an open-architecture system is needed to port novel assays to determine genetic elements more fully or to screen quickly for elements that are most likely of diagnostic value in isolate characterization. The power of qPCR lies with assay selection, some multiplex capability, optimization, and validation of the PCR assays themselves.

There is a wealth of available PCR assays that are validated to certain standards, although they must be validated again in each laboratory that implements the assay(s). Many are Food and Drug Administration (FDA) approved and enable qualitative detection of hundreds of pathogens and hospital-acquired infections to the level of species and sometimes strain. Qualitative tests can provide solid evidence for determining the next step of an analysis as well as generating a quick answer. Some detect antimicrobial resistance genes. There are also assays in the food safety arena to draw upon. Matrices in the food industry are complex, and food is a common target of intentional and accidental contamination. There are many food industry investigations about such events, and microbial forensics should learn from these, understand the assays, and have access to the assays should a food-related microbial forensic investigation arise. There are fewer assays for animal and plant pathogens, yet they are important, and the microbial forensics community can communicate their importance.

The limitation of PCR is that the list of what can be detected is contingent on the validated assays in one's toolbox. Even if one has many assays, it is still limiting and restricts one's multiplexed capability. The MassTag PCR technology marries PCR with traditional chemistry techniques (mass spectrometry) to provide the resolution quality of mass spectrometry coupled with the sensitivity and specificity of PCR. It has, for example, revolutionized the ability to identify respiratory pathogens and hemorrhagic fever viruses.

Mass spectrometry also is useful in identifying many non-nucleic acid chemical species that may provide clues to microbial identity, origins, and production processes. Proteins, peptides, lipids, carbohydrates, inorganic metals, and organic metabolites may provide information about an organism's source environment, how it was produced, and the level of sophistication of the preparation (Wahl et al., 2011). For example, microorganisms respond to the conditions in their environment by altering which of their genes are transcribed and then translated into proteins. For pathogens isolated from samples other than patients, the profile of proteins an organism produces can sometimes provide valuable information about the environment in which the organism originated. A mass spectrometer generally consists of an inlet for introducing the sample,

an ionization source to turn the analyte into charged gas-phase particles (ions), and an analyzer that detects and separates the charged particles based on their mass-to-charge ratio. Two ionization methods are used for biological mass spectrometry: electrospray ionization and matrix-assisted laser desorption/ionization (MALDI). Wahl and colleagues (2011:456-457) point out that although nucleic acid sequencing is invaluable in microbial forensics, proteins “may offer improved stability over DNA markers . . . and have fewer or different inhibitors to analysis methods.” It is also useful to have additional detection modalities. More in-depth reviews of mass spectrometry and its role in microbial forensics are available in Wahl et al. (2011) and Snyder and Jabbour (2014).

Dr. Jongsik Chun of Seoul National University pointed out that in Korea, use of MALDI—specifically the MALDI-time of flight (MALDI-TOF)<sup>2</sup>—is becoming very popular in routine clinical laboratory analyses. The potential of this technique exceeds most techniques discussed by Kadavy except for sequencing. Kadavy agreed that the MALDI-TOF technology is a very good option although it does not have capacity equal to sequencing. The CDC is using MALDI-TOF to identify the protein toxins of *B. anthracis* and *Clostridium botulinum* (Boyer et al., 2011; Kalb et al., 2011). The CDC’s tests are used both for assays for the presence of the toxins and for identifying the correct botulinum neurotoxin subtype for administration of the most effective therapeutic immunoglobulins to affected patients.

Microarray platforms provide similar capabilities but in a more multiplexed format. They have the ability to detect both nucleic acid and protein signatures. Many panels have 1 million assays on a single array chip; some have up to 5 million assays. There are custom and standard microarray panels. Microarray panels are available for microbial detection, SNP detection, and genome-wide association studies, gene expression, and protein presence and abundance. Laboratories that use a microarray reader for other purposes, such as oncology or infectious disease screening, could be leveraged by the development of custom chips that are directed to answering microbial forensic questions. Most assays are available; it is a matter of customizing a single chip to combine them. Microarrays will not achieve the level of information provided by sequencing, but they are quite impressive. They may be a rapid and cost-effective alternative or supplement to sequencing.

For microbial forensics purposes, NGS is a powerful set of technolo-

---

<sup>2</sup> Matrix-assisted laser desorption ionization-time of flight (MALDI-TOF) mass spectrometry can establish the identity, purity, or other aspects of protein samples, genetic materials, and microorganisms. MALDI-TOF uses the time for a target particle to traverse a specific distance after being dislodged from a surface by a precise amount of energy. This allows a determination of the molecular weight of the target. [http://www.ndif.org/public/terms/11543-MALDITOF\\_mass\\_spectrometry](http://www.ndif.org/public/terms/11543-MALDITOF_mass_spectrometry).

gies that offers a state-of-the-art approach. However, NGS also poses a new set of issues for implementation. For example, the same data run on two different systems—the benchtop MPS (massively parallel sequencing) instruments, Personal Genome Machine™ or PGM (or Ion Torrent, Life Technologies) and the MiSeq (Illumina)—can produce somewhat different answers. These issues will have to be resolved or at least understood for the microbial forensics context.

Detter and Resnick (2014) provide a brief review of NGS. In general, sequencing consists of four steps: preparation of the sample by nucleic acid (either DNA or RNA) extraction, library preparation, sequencing, and data analysis. Following extraction of the material from the sample as described above and in Chapter 5, the DNA molecules to be sequenced must be converted into sequencing “libraries.” This preparation entails fragmenting the DNA and then adding adapter molecules to the ends of the fragments to facilitate sequencing. Each DNA fragment in a library is then clonally amplified before sequencing. The precise process and how automated it is depends on the NGS platform being used (Detter and Resnick, 2014).

NGS produces huge amounts of data as “reads,” which are strings of DNA that correspond to the sequence of the original DNA or RNA being investigated. Reads have three important features: length, number (or “coverage”), and quality. All of these differ according to which platform and sequencing kit are used as well as the quality of the sample and library preparation. Kadavy discussed three platforms that were described in a 2012 article that compared the three, head-to-head, using real data (Quail et al., 2012). However, new improved instruments come out frequently and displace those already on the market. The new high-throughput sequencing methods have greatly improved the overall process. Some novel approaches beyond these platforms incorporate single-molecule real time, or “smart cell,” and also preparation adaption-ligation strategies; alternative chemistries, and so, less RNA/DNA input is required; and longer read lengths. These decrease the time to results. The approximate sequencing capacity of thousands of bases per day of 1995 has been increased to billions per instrument per day in 2013.

All sequencing-platform workflows share general themes—fragmenting DNA, various adaptor-ligation strategies, and novel chemistries—and efforts are under way to reduce preparation time and sample input requirements and improve the quality of results. The CDC, the College of American Pathologists, and others have developed guidelines for its use with human diagnostics (Ellard et al., 2012; Gargis et al., 2012; Pont-Kingdon et al., 2012; Rehm et al., 2013). But these guidelines do not address the range of possibilities encountered in microbial forensics, that is, from the homogeneous sample to the very complex metagenomic

or microbiome samples. Sequencing remains a huge investment in time, materials, and money and also demands a huge investment in data analysis because none of the data are meaningful without bioinformatics, and bioinformatics and software issues must be factored in (see Chapter 7). Bioinformatics pipelines that can handle a continuum of samples from pure isolates to complicated metagenomic samples that may contain only trace amounts of the threat agent are needed. Kadavy stated that it is necessary to be prepared for all possibilities, so building on the current technologies to meet future needs must continue.

In terms of bioinformatics, Kadavy stressed that accurate and updated genomic databases are critical. The dialogue has only just begun on how to share such data. Bioinformatics tools must be usable and practical. They must be applicable to the questions posed and validated for forensic application. A challenge is that in a field that is evolving so quickly, how can methods that are themselves continually evolving be validated in a timely fashion and in an economically feasible manner?

Kadavy suggested that investments in technology should be prioritized according to what is best for the common good. It is possible to leverage advancement that helps microbial forensics and public health simultaneously. These fields should not compete for resources, but should share them to meet their respective goals.

The issue of proprietary restrictions that make some technologies inaccessible was raised by Dr. Raymond Lin. For example, there are good publications on MassTag spectrometry, but many of the technologies are proprietary. Consequently investigators have had to develop their own panels, and results have not been good enough for application. Published microarrays are often proprietary, so other groups have had to develop their own chips. One group in Singapore has licensed arrays for about 70,000 pathogens, but they tend to be undersensitive, and do not perform as well on respiratory pathogens. Lin believes that the future may be in using one of the NGS instruments that is less biased as far as target analysis and target enriched with certain non-eukaryotic sequences.

Kadavy agreed that the proprietary issues are an important consideration. It is also important to know what technologies people are using. How does the MALDI-TOF operate in one's own hands? To what do other countries have access? What technologies should one be considering?

She suggested the workshop consider how to link up experts—in microbiology, clinical medicine, public health, research and diagnostics, and food science—in order to leverage and build on knowledge and for reach-back purposes. She noted that small laboratories and research groups can leverage existing infrastructure and expertise in public health systems. How can information, including sequencing data, be shared in



the international microbial forensics community? This will require policy decisions.

Dr. Munirul Alam raised the problem of sharing data in areas where limited resources restrict technologies. Bangladesh is part of PulseNet Asia Pacific (part of PulseNet International<sup>3</sup>), which enables access to a database for threat pathogens. He asserted that labs in developing countries cannot afford everything, so it is important to collaborate as much as possible.

Dr. Ruifu Yang agreed that sharing is key for developing microbial forensics. He has collaborated with others to standardize methods, using the same data but analyzing them in different laboratories, and then publishing the results. If scientists work together, they can share more than just published data.

Piers Millet of the U.N. Office for Disarmament Affairs pointed out the potential of the technology and international collaborative arrangements that enable sharing sequences and related data via the Internet. He cited the example of the rapid response to the posting of the viral sequences of the H7N9 avian influenza by the Global Initiative on Sharing Avian Influenza Data (GISAID) by the Chinese CDC in March 2013.<sup>4</sup> This process enabled Novartis, in partnership with a unit of the J. Craig Venter Institute and support from the U.S. CDC, to synthesize the genes of the new virus and begin initial work on developing a vaccine using its cell-culture production process, which is itself much faster than traditional egg-based flu vaccine development (Brennan, 2013). The synthesis was completed and shared with CDC before physical samples of the virus arrived by mail.

Dr. Gilles Vergnaud, Institut de Génétique et Microbiologie, France, works for the French Ministry of Defense, Defense Procurement (DGA). In 2001, he headed the department of dangerous biological agents in the DGA. Following the U.S. anthrax event, many countries experienced thousands of threat hoaxes. In France alone, there have been 3,000; the defense establishment dealt with most, and his division undertook about 1,500. They needed an easy way to determine which incidents were hoaxes and which, if any, were real threats. A hoax is massively disruptive; many involved hospitals, post offices, and other public spaces. In every case, evidence was processed as quickly as possible to differentiate hoaxes

---

<sup>3</sup> PulseNet International is a network of national and regional laboratories that track foodborne infections worldwide. Each laboratory uses standardized genotyping methods and shares information in real time. More information available at <http://www.pulsenetinternational.org/>. See also Global Food Safety, <http://www.slideshare.net/Adrienna/global-food-safety2013>.

<sup>4</sup> Further information about GISAID can be found at <http://platform.gisaid.org/epi3/frontend#283e41>; accessed March 31, 2014.

from real threat agents, although even this simple question requires a confident answer. He went on to say that he has long had an interest in database issues, but there are many political problems. He believes it will never be possible to share all forensic-quality data, but some level of slightly degraded information may be shared, and he believes that even that would be a great achievement.

Dr. Rocco Casagrande of Gryphon Scientific added a cautionary note by suggesting that a guiding principle when developing new technology should be to consider the consequences of its day-to-day use. Better surveillance capabilities and better technologies for identifying unusual microorganisms can result in “events” that otherwise would not be interpreted as such. Had it been possible, for example, to detect Aum Shinrikyo’s release of anthrax spores in Japan in 1993, there might have been an “event” even though there were no public health consequences. Similarly, in the United States, when grapes imported from Chile were allegedly contaminated with cyanide (Rushing, 1994), the cyanide was insufficient to cause human health concerns, but the detection caused significant economic harm to the exporter. John Clements further noted that it is important to distinguish between meaningful events and background “noise.” For instance, there are approximately seven cases of plague due to *Y. pestis* diagnosed in the United States every year. Although a sharp increase in number or distribution may be important, the mere existence of the disease is “background.” Also, in the Aum Shinrikyo anthrax case, cult members failed to spread the disease because they used the nonvirulent vaccine strain of the organism. Use of an insufficiently discriminatory assay would have picked up *B. anthracis* and may have started a full-scale panic even though the public was at no great risk. There are consequences for public safety, the economy, and individuals for making declarative statements about things that are merely suggestive.

### INVESTIGATING INFECTIOUS DISEASES IN CLINICAL MEDICINE

Professor Alemka Markotić shared her perspective as a scientist and a clinician who daily faces a variety of infectious diseases in patients and has used elements of microbial forensics to deal with public health outbreaks. She offered lessons learned through experience with zoonotic diseases—diseases transmitted between animal species and from animals to humans—which she believes could be useful to public health experts, clinicians, and microbial forensic investigators. She pointed out natural vulnerabilities that terrorists could exploit, and proposed areas where research is needed.

Zoonotics can cause severe disease both in animals and humans and



therefore constitute a serious public health problem. Animals and their vectors play an essential role in maintaining zoonotic infections in nature. Zoonoses may be due to bacterial, viral, or parasitic organisms or may involve unconventional agents. In addition to being a public health problem, many of the major zoonotic diseases (e.g., brucellosis, salmonellosis, listeriosis, trichinellosis, campylobacteriosis, and hepatitis A or E) prevent the efficient production of food of animal origin and create obstacles to the international trade of animal products (World Health Organization, 2014). Markotić noted that many of the agents regarded as potential bio-weapons are zoonotic.

Preparing for a biological threat event requires knowledge of pathogenic agents; a clinical picture of disease incubation, transmission mode, and treatment; knowledge of which molecular diagnostics and technologies to apply to which pathogen; and characterization of a usual versus an unusual outbreak (see Box 3-1). But Markotić stressed that in real life, the picture is not always clear.

Hantaviruses are transmitted to humans through contact with hantavirus-infected rodents or their urine and droppings. Infection with Old World hantaviruses can progress to hemorrhagic fever with renal syndrome (HFRS), and infection with New World viruses to hantavirus pulmonary syndrome (HPS). The clinical picture for hantavirus ranges from mild illness to a severe form with fulminant hemorrhagic fever. Mortality can be as high as 20 percent with Old World hantaviruses and 60 percent with New World hantaviruses.

Croatia is a natural center for many rodentborne zoonoses (e.g., leptospirosis, babesiosis, and HFRS) because of its diverse forest ecology and abundance of small rodents (Markotić et al., 2002a; Tadin et al., 2012). In January through April 2012, 33 patients were hospitalized in Zagreb, experiencing something that resembled flu but did not exhibit its typical symptoms (Tadin et al., 2014). This outbreak occurred against the backdrop of a normal flu season, including cases of influenza A subtype H1N1. However, it was subsequently determined that all patients had either attended the “Snow Queen Trophy” skiing competition at Medvednica Mountain, a popular ski resort, or lived near or had links to the mountain. Acting on their suspicions, Markotić and her colleagues performed point-of-care tests and found that all patients were serologically positive for specific antibodies to Puumala virus, a hantavirus (Markotić et al., 2002b).

There are three lessons to be learned from this event. First, a simple clinical test can quickly distinguish between influenza and HFRS. Results were received within hours of patient admission. Such tests can be a powerful tool for providing orientation in the field at the onset of an unusual outbreak.

This outbreak comprised an extremely high number of HFRS cases

within a small geographical area. Moreover, the timing was unusual because HFRS cases typically peak during the summer months. The patients' clinical symptoms were first compared with those of previously published cases and there were no differences. To confirm the origin of the infection, both human and rodent samples were analyzed. Through collaboration with the forestry and veterinary medicine sectors, rodent samples from different altitudes of the mountain were collected.

Markotić and colleagues discovered that almost 80 percent of rodents at the altitudes of the mountain's ski track were infected with hantaviruses—a rate never reported previously in Croatia or in the literature. Molecular and phylogenetic analyses of viral nucleic acid sequences obtained from human and rodent samples confirmed an almost 100 percent similarity between the samples, which helped rule out the virus having been imported or deliberately created.

The second major lesson learned was that molecular analysis enabled linking the human outbreak with a pathogen in a natural reservoir, rodents, in a specific geographic location. A major contributor to the dispersal of hantaviruses in the Medvednica Mountain outbreak was that, because of warm winter temperatures, the ski resort had distributed aerosolized artificial snow to prepare the snow track for the competition. Moreover, high winds blew the snow off the track.

A third major lesson came from an unusual outbreak of HFRS cases that began in a community of former drug users whose rehabilitation program included gardening near a forest (Medved et al., 2002). Many community members became infected with hantaviruses and developed HFRS, and one died. Fifty-two percent of all community members were infected with hepatitis B or C, or both. Of those with HFRS, 82 percent were co-infected with one or both types of hepatitis virus (Medved et al., 2002). It appeared that the multiple infections were associated with altered immune responses. Markotić suggested that there is a need to consider whether multiple infections pose an investigative problem, both in terms of public health and microbial forensics.

Markotić posed the questions, "What if mother nature can do it better? And what if terrorists copy it?" She and her colleagues, using an existing bank of rodent samples, investigated what pathogens rodents in northern Croatia carry. Their research revealed that the rodents typically carry one to four zoonotic infections (Tadin et al., 2012). In light of this fact, one must consider the public health implications of multiple infections. When physicians diagnose a patient with one microorganism, they tend to be satisfied with that diagnosis, and treat the patient accordingly. Even if we record unusual immunopathogenic responses (e.g., strong inflammatory response or immune suppression) in patients, the possibility of multiple infections is not typically considered. Co-infection and multiple infections

should be the target of public health research. The literature shows that more severe disease occurs in the presence of co-infection (e.g., influenza with pneumonia), as well as higher mortality (Chertow and Memoli, 2013). Bacterial co-infection has occurred in influenza pandemics throughout the world, including those with H1N1 influenza (Wang et al., 2011). In the event of co-infection or multiple infections, there exists a different immunopathogenic basis for disease (Guidi et al., 2011).

Immune response is altered in the presence of more than one infectious disease. Markotić's research shows the level of proinflammatory cytokines to be several times higher in patients infected with both hantaviruses and leptospira than in patients with a single infection (Markotić et al., 2002a). A study examining the interacting roles of cytokines and viral load in Crimean-Congo hemorrhagic fever showed that viral load and immune response parameters can be useful as biomarkers to predict disease severity and outcomes (Saksida et al., 2010). Therefore, in addition to using multiplex technology to detect co-infection or multiple infections, the capability to measure immune response in unusual outbreaks or disease presentations should be developed.

An important "omics" research field that holds great promise for both clinical medicine and microbial forensics is immunogenomics, an emerging field that focuses on the intersection between genomics and immunology. It takes into consideration the host and pathogen proteomes, and combines bioinformatics technology, genomics, proteomics, immunology, and clinical medicine (Gupta et al., 2009). Professor Indrani Karunasagar of the Karnataka Veterinary, Animal & Fisheries Sciences University of India, commented that a syndrome may be precipitated by a certain factor, but the condition is exacerbated by the presence of bacteria. Perhaps addressing changes in climate or other environmental factors might be used for prevention. She agreed that immunogenomic bioinformatics is very important. It might answer, for example, why when two people are exposed to the same situation, one becomes infected and the other does not. The genetic makeup of an individual could aid in addressing possible contributing environmental factors.

Markotić emphasized that there are very powerful simple techniques. Flow cytometry, a well-known cell sorting tool, can sometimes quickly orient one to whether a pathogen is a virus or a bacterium when the clinical picture and diagnostic results are not clearly distinctive. At the Zagreb hospital where Markotić practices, they have used this technology to differentiate HFRS from leptospirosis (Markotić et al., 2011).

To improve clinical and forensic approaches to microbial identification, Markotić believes interdisciplinary approaches and advanced technology are needed, specifically:

1. Molecular epidemiology data and field research.
2. More extensive clinical databases, for example, descriptions of “usual” and “unusual” cases and outbreaks.
3. A translational medicine approach to link basic and applied research to clinical data and diagnostic tool development and a One Health medicine approach to zoonotic diseases.
4. Development of robust, powerful, multiplex molecular diagnostic tools and bioinformatic analysis tools.
5. Development of rapid and inexpensive diagnostic tests.
6. Powerful immune-response measurement and bioinformatics analysis tools.
7. The ability to understand and synthesize a huge amount of data, and react promptly and properly.
8. The ability to quickly communicate (e.g., via telemedicine) with world experts knowledgeable about the clinical manifestation and detection of dangerous pathogens.

In a final note on zoonotics, it was pointed out that the National Center for Foreign Animal and Zoonotic Disease Defense (FAZD, <http://fazd.tamu.edu/>) at Texas A&M University in the United States has a program in which they are trying to integrate data and information from clinical and field observations provided by any stakeholders—from pet owners to production/processing facilities. Participants can electronically enter observations (e.g., presentation, diagnostics) into the system. Normally, there is a problem with unwillingness of people to share such information because any indication of a serious zoonotic situation can have significant financial ramifications. FAZD provides a shield for reporting this information and they are assimilating the data at a high level, looking for multiple loci or anything that will give them information about natural outbreaks, deliberate attacks, or other sources.

#### **CULTURE-INDEPENDENT TESTS: CLINICAL AND PUBLIC HEALTH PERSPECTIVES**

Dr. Stephen Morse of the U.S. CDC spoke about how rapid developments in technology and clinical laboratory test methodologies are leading to culture-independent testing. He discussed the advantages and limitations of culture-independent tests from the clinical and public health perspectives, noting that his comments should not be assumed to reflect CDC policies. In the case of a large outbreak of an infectious disease, particularly an enteric infection, careful epidemiological investigation will likely be needed to differentiate between an intentional and a naturally occurring event.

Currently public health surveillance is isolate-based, but this approach is changing. For example, PulseNet, which was founded in 1996, is the molecular subtyping laboratory network for conducting foodborne disease surveillance. It relies on the availability of isolates for testing using pulse field gel electrophoresis. This culture-based diagnostic and typing method is, however, likely to be replaced by more rapid PCR-based tests (CDC, 2011). A move toward rapid nonculture testing is expected to improve the speed and cost-effectiveness of diagnostics. However, it also likely will result in a decrease in the availability of isolates for PulseNet subtyping and susceptibility testing. The issue of how to adapt to rapidly changing technology is a major one for the U.S. CDC. CDC and its partners in the State Public Health Laboratory network are working to develop higher resolution diagnostics, new rapid and standardized data collection methods, new analytical software, and improved environmental assessments (CDC, 2011). President Obama's FY 2014 budget proposed the establishment of the advanced molecular detection (AMD) program to begin to replace CDC's decades-old methods of detecting microbes, which will soon be obsolete, as well as increasing understanding of their lethality to humans.<sup>5</sup>

In the United States, public health surveillance has a number of functions (Cronquist et al., 2012). Surveillance assists in individual case management and investigation at the local level. It is used to assess disease burden and trends in order to prioritize and assess the impact of population-based control methods. It is relied upon for outbreak detection to protect the population, as well as to identify gaps in control measures. Isolates collected for public health purposes undergo characterization to improve understanding of pathogens and their virulence mechanisms and to assess infection epidemiology.

The needs for testing in patient management differ from those in public health. For patients, an ideal test is fast, accurate and, when necessary, can establish antimicrobial susceptibility. In the case of most infections, a patient will be treated with an antibiotic before the test results arrive, and subtyping or virulence testing is seldom used. As we understand more about the genetics of virulence factors, however, such testing may play a greater role.

In contrast, population-based testing values specificity and sensitivity over speed; the goal is to ensure that outbreak-associated cases are identified unambiguously. The public health sector requires more detailed information on isolates, and subtyping becomes important because it enables detection of outbreaks due to particular strains. In this context,

---

<sup>5</sup> More information on AMD can be found at <http://www.cdc.gov/amd/>; accessed April 7, 2014.

monitoring antimicrobial susceptibility is particularly important for two reasons. First, antibiotic-resistant organisms are associated with more severe disease. Second, livestock and poultry are commonly treated with antibiotics to enhance their growth, which selects for antibiotic-resistant organisms. Monitoring can aid in attribution, particularly in foodborne outbreaks in which antibiotic-resistant organisms are associated with meat and poultry.

From a clinical perspective, the advantages of culture-independent testing to patients include more rapid diagnoses, the potential for easier specimen collection and decreased costs, and the detection of a wider range of pathogens (e.g., non-O157 Shiga toxin-producing *E. coli*) as well as better sensitivity for some pathogens, such as those that are difficult to culture. Challenges include false-positive findings that result in unnecessary treatment or incorrect diagnosis, and an inability to test for antimicrobial susceptibility (Cronquist et al., 2012). Specifically, the presence of a genetic sequence coding for a factor associated with antimicrobial resistance (e.g., an enzyme) does not mean that the gene is expressed or that the microbe is resistant, only that the associated genetic sequences are present.

From a public health/population-level perspective, the advantages of culture-independent testing include rapid case detection; a potential for increased testing, leading to greater case capture; detection of a wider range of pathogens; and again, better sensitivity for some pathogens (Cronquist et al., 2012). Limitations can include resources and time wasted following up noncases, incorrect or unstable estimates of the number of illnesses, and disruption of trend monitoring. Other challenges are the loss of subtyping for outbreak detection, the possible investigation of pseudo-outbreaks, and a decreased ability to monitor trends in subtypes,<sup>6</sup> such as *Salmonella enteritidis* (Cronquist et al., 2012).

### PERSPECTIVE OF A CLINICAL MICROBIOLOGIST IN THE PUBLIC HEALTH SECTOR

Dr. Raymond Lin of the National University of Singapore offered a perspective on the challenges and issues clinical microbiologists who work in the public health sector face in terms of technology, policy, and legal issues.

---

<sup>6</sup> Infectious disease outbreaks are currently monitored by culturing organisms from patients or contacts, identifying those organisms by serology, biochemistry, fermentation patterns, and antimicrobial susceptibility. Changing to a nonculture system would not provide sufficient continuity to allow us to continue monitoring infectious disease trends (down to the subtype level of the microorganisms) that have been established over the last 100 years using the aforementioned technologies.



New technologies, such as the MALDI-TOF mass spectrometry technique, have brought great change into clinical microbiology and public health labs because they enable earlier pathogen identification. However, as with any technology, limitations exist. Currently MALDI-TOF cannot always provide needed differentiation because some important public health pathogens are genomically very similar to nonpathogenic species. For example, *B. anthracis* may not be easily differentiated from *B. cereus* or *B. thuringiensis*, so extra testing may be required.

Dr. Lin pointed out that public communications about the meaning of data generated by the new technologies can also present a major problem. The media and the public typically do not understand the science underpinning a report of “a new pathogen,” “a new chimera,” or “a new mutation.” They assume the worst and fear there is a serious danger to public health, whether there is or not. Better communication could prevent unnecessary disruptions and delays.

In public health, the needs for gene typing are relatively simple. Most investigations focus on outbreaks, and seek clustering over a short period of time to provide leads for further epidemiological investigation. Most commonly, action needs to be taken to address the most likely cause of an outbreak based on the epidemiology, with or without laboratory results. In his experience with foodborne outbreaks, one or two people with diarrhea meeting the case definition may prove to have an unrelated pathogen or unrelated strain. Usually it is not difficult to piece the story together, but a capacity that is very much needed is a database of the baseline types of the worst pathogens—even something as common as *Salmonella*. Certain species are so common that adequate baseline databases are not available to public health authorities to help identify how different a certain isolate may be from what is known to be in the community.

Lin also believes that a problem with typing/sequencing capability is that it may encourage litigation. He asks if, in foodborne outbreaks, whether action should be taken against the food vendor? He tries to avoid being drawn into cases alleging, for example, that one person transmitted HIV to another person. The courts ask many questions, and Lin believes that public health lacks the manpower to spend time in court for such cases. Public health investigators are obligated, however, to be involved when commercial products are suspect in pathogen transmission. In Singapore, there are many imported disease cases, and these must be resolved. Were a “local” transmission of cholera to rise above a certain threshold, for example, the country would lose its cholera-free status. Similarly, people who come from malaria-endemic countries may be infected by more than one strain of *Plasmodium*, be fairly asymptomatic, but still transmit the infection(s), possibly with antigenic variations.

Lin would like to see studies using benchtop NGS instruments. Much

of the literature has focused on the large expensive machines. The PGM (Ion Torrent, Life Technologies) and MiSeq (Illumina), both benchtop machines, are not used much in public health in Singapore.

He agrees that there is a need to build databases and to standardize how to report the data internationally. How does one communicate in plain language so that the implications of an analysis with MLST+, for example, are understandable? Lin agreed with Morse in his concern that viral isolation has become almost a thing of the past in diagnostic labs. He believes that public health labs must make an effort to collect specimens of bacteria and viruses. Even something as common as investigating foodborne outbreaks is still done by culture, but Lin thinks that with an increasing move toward multiplexed PCR, public health will have to take over the responsibility of collecting isolates so that long-term bacterial and viral archives can be created and maintained.





## 5

# Sampling and Preservation Methods

### THE BROAD PERSPECTIVE

Smith (2011) references a Department of Homeland security meeting in 2005 from which a report on sample collection, handling, and preservation was issued. The authors of this report stated that the collection and preservation of vital microbial forensic evidence is a critical element of successful investigation and ultimate attribution subsequent to a biological event.

A primary goal of collection is to obtain sufficient biological agent to support both species/strain or toxin identification for critical public health decisions and complete signature characterization for valuable lead information. Also, the collection of other relevant traditional forensic evidence must not be overlooked. Trace evidence, fingerprints, and other traditional evidence should be collected and preserved in order to support the attribution mission. (Smith, 2011:380)

This chapter explores the principles and practices that guide the collection, handling, and preservation of microbial forensic evidence.

Mr. Adam Hamilton, President and CEO of Signature Science LLC, gave an overview of sampling and preservation, reviewing potential applications of microbial forensics both inside and outside of law enforcement. He also discussed the challenges of collecting and preserving immensely diverse evidence samples for multiple purposes, and made suggestions for specific areas of opportunity for international collaboration to advance the discipline.

In all cases, the bottom line in sampling is that samples must be collected and preserved in a manner that prevents or minimizes degradation or contamination. This requirement makes sampling and preservation as important to the microbial forensic process as is scientific analysis. The key messages of his presentation were that microbial forensics is a multidisciplinary field developed to serve law enforcement needs and provide criminal attribution. However, although it was initially motivated by biocrimes and bioterrorism, there are many new applications for microbial forensics that create both opportunities and challenges. Expanding roles for microbial forensics also provide opportunities for leadership on the development of international guidelines for sampling and handling practices. “Standardized flexibility” is required to accommodate myriad applications. In addition, while quality control is a necessity, so is transparency. Anyone who is potentially affected by the outcome of microbial forensic analyses should be provided insight into what transpired and how conclusions were derived.

The expansion in potential applications of microbial forensics is being driven by the availability and accessibility of new technologies, such as NGS and bioinformatics. As the cost of using these technologies drops, more communities will use them to address the problems they face, as has already been well illustrated by the public health cases discussed in a previous chapter. Security and surveillance are possible new areas of application for microbial forensics. Microbial forensics could be used to identify strategic locations for collecting monitoring samples and also could inform approaches for monitoring systems, such as those needed in agriculture. In the clinical and public health realm, microbial forensics could play a major role in pathway identification of health care–associated infections. The combination of an increase in multidrug-resistant pathogens, shrinking therapeutic pipelines, and enhanced access to health care has the potential to produce a health crisis. Phylogenetics determined through sequencing and other technologies can complement traditional epidemiology. Microbial forensics also could aid the development of diagnostics and interventions for transboundary diseases, particularly animal and zoonotic diseases.

As with other forensics, the collection and preservation of microbial evidence are critical for efficient and successful investigation and attribution. Moreover, a relatively small number of samples may be the basis for highly significant strategic and policy-level decisions. Evidence sampling approaches must include (1) planning and design; (2) protocols for quality assurance and quality control; (3) logistics and preparation; (4) collection, which is a fairly small part of the overall process; and (5) documentation, which is particularly important for maintaining the chain of custody. Evi-

dence handling must take into account storage conditions, packaging and labeling, shipment/transportation, and maintaining the chain of custody.

Best practices need to be developed for sampling and handling systems to create a viable microbial forensics system. This is well recognized, as exemplified by the 2008 recommendation made to the U.S. Department of Defense that “[t]he broad group of government agencies . . . should develop and promote best practices for microbial sample collection that best preserve the genetic and epigenetic signatures of interest” (JASON Advisory Panel, 2009). Because the need for microbial forensics is international, there should be internationally accepted practices. Efforts to evaluate and improve collection methods and standards are under way at the U.S. Department of Homeland Security’s National BioForensic Analysis Center<sup>1</sup> (NBFAC) and National Biodefense Analysis and Countermeasures Center<sup>2</sup> (NBACC), but similar efforts are needed globally in order to advance the science of microbial forensics as a whole.

The goals of sampling and handling are shown in Box 5-1. Hamilton stated that the first goal is to ensure the health and safety of the evidence collectors, any resident populations near the location of an outbreak event, and those who may have already been exposed in a real attack. The design for evidence collection must sufficiently demonstrate the hypothesis or hypotheses that investigators propose. The capabilities of sample analyses must be evaluated with the knowledge that results will be fed not only into a microbial forensics system, but also into other forensic systems. And flexibility must be incorporated throughout, owing to the variety of matrixes and samples that can be collected.

To ensure the broad applicability of preserving and protecting the integrity of samples, the methods that are efficient and low cost are most

---

<sup>1</sup> NBFAC is part of the National Biodefense Analysis and Countermeasures Center (NBACC), which applies science to challenges critical to defending the nation against bioterrorism. The Department of Homeland Security’s Directorate for Science and Technology established the NBACC to be a national resource to understand the scientific basis of the risks posed by biological threats and to attribute their use in bioterrorism or biocrime events. NBFAC’s mission is not primarily research. Rather, NBFAC conducts bioforensic analysis of evidence from a biocrime or terrorist attack to attain a “biological fingerprint” to help investigators identify perpetrators and determine the origin and method of attack. NBFAC is designated by Presidential Directive to be the lead federal facility to conduct and facilitate the technical forensic analysis and interpretation of materials recovered following a biological attack in support of the appropriate lead federal agency (<http://www.dhs.gov/national-biodefense-analysis-and-countermeasures-center>).

<sup>2</sup> NBACC’s National Biological Threat Characterization Center (NBTC) conducts studies and laboratory experiments to fill in information gaps to better understand current and future biological threats; to assess vulnerabilities and conduct risk assessments; and to determine potential impacts to guide the development of countermeasures such as detectors, drugs, vaccines, and decontamination technologies. Neither NBFAC nor NBTC has as its primary responsibility to conduct research on microbial forensics.

**BOX 5-1**  
**Sampling and Handling Goals**

- Health and safety
- Sufficiency of design
- Analytical compatibility
- Preservation and integrity
- Low cost and high efficiency
- Documentation and training
- Quality and transparency
- Flexibility!

SOURCE: Hamilton presentation, 2013.

desirable. Spending \$10 per collection kit for every sample is unrealistic; ideally it should cost pennies. Although compared to the cost of sequencing, \$10 per collection kit may not seem unreasonable. However, it is possible that many samples will be collected on a routine basis, and it would be desirable to be more cost-effective by reducing the cost to a few cents. Lower cost collection devices would allow more samples to be collected and preserved for subsequent analyses, if warranted.

Documentation and training in how to conduct sampling missions are needed so that collection efforts can be delegated to people with other responsibilities or who may not be involved in the microbial forensic field at all (e.g., first responders). Cerys Rees of Porton Down in the United Kingdom noted that because first responders are often police or military personnel, results of sampling are much improved if the microbial forensic experts can advise the samplers before they begin to collect evidence. When the scientists are asked in advance for their recommendations for the sampling process, analytical results tend to be better. Also, when samplers are told what the scientists plan to do with the samples in the laboratory, they are better equipped to sample appropriately. Matts Forsman of the Swedish Defense Research Agency also related that in Sweden, first responders may include police, fire, and rescue squad personnel, but a specially trained bomb squad is then called in to sample, pack, and transport samples to the Swedish National Laboratory for Forensic Science.

Some generic sample matrices for evidence collection appear in Box 5-2. Microbial forensics incorporates both clinical and environmental samples, which in turn vary in size, shape, matrix, procedure, and method used to collect them. Portions or the entire sample or item may be collected. For example, the entire house of the Unabomber, Ted Kaczynski,

**BOX 5-2**  
**Sample Matrices for Evidence Collection**

**Environmental**

- Air and liquid filtrates
- Water
- Soil
- Vegetation
- Wipes and swabs
- Aliquots
- Filtrates
- Food
- Vectors
- Fomites
- Bulk items
- Etc.

**Clinical**

- Swabs
- Feces
- Urine
- Blood and components
- Sputum
- Fluids
- Washes
- Tissues
- Hair
- Etc.

SOURCE: Hamilton presentation, 2013.

was moved to a controlled area for evidence analysis. Standardization is needed for handling bulk as well as minute samples. As pointed out by Dana Kadavy, because microbial forensic investigators deal with such a broad range of sample matrices, they should maximize cleanup procedures to enhance typing success.<sup>3</sup> A laboratory may need more than one sample processing method, particularly if it receives a variety of samples that reside in different matrices. One technology may be particularly effective for processing clinical samples, for example, but not soil, effluents, or plant material. Automated technologies can reduce labor, minimize hands-on time, and reduce human error in extraction and preparing PCR plates and sequencing libraries. However, an automated technology is not necessarily a superior methodology. Validation studies at each laboratory must determine whether the technology reduces error and/or contamination, affects yields, and/or improves the analysis.

Bruce Budowle also touched upon sampling objectives, which may address elements such as real-time monitoring, screening, random sampling, targeted sampling, collecting bulk material and suspicious items, and the conditions of the crime scene. Each element is accompanied by

<sup>3</sup> This is required to remove most of the interfering substances from the matrix and to promote the concentration of the analyte.

a set of criteria that must be considered. Each of the different sampling approaches will require incorporating different considerations when planning the collection process. (For further detail, see Chapter 6.)

Ideally, ready-for-prep samples arrive at the microbial forensics laboratory, are fed into the system, and research and analysis are performed. But in operational terms, collecting samples is complicated. First, the investigator must determine how to optimize the collection of the clinical or environmental samples from the various sources and how to preserve them in an appropriate manner. At the same time, the collection procedures should not disturb other forensic processes. In fact, the goal is to leverage the microbial samples for use for other forensic applications. Finally, the evidence must be preserved so it is available for future analysis, be it for a secondary analytical or validating process, or for analysis one day using a not-yet-invented technique.

Randall Murch added that sampling from a facility or location identified as worthy of investigation but from which law enforcement may wish to gather evidence unobtrusively while building a case can be a challenge. Because they do not wish to draw attention to themselves, investigators may be constrained in how long they are on site, what they carry, including protective gear and sampling equipment, and how many samples they collect. This constraint will affect strategies, design features of sampling kits all the way through the analytics, and conclusions drawn from the effort. How conclusions should be caveated, given the uncertainty that may accompany collection under these circumstances, is a topic that needs further consideration.

Hamilton summed by stating that there is a global need for microbial forensics. Areas of opportunity for international collaboration to advance the discipline include the following:

- Standardized terminology for the discipline;
- Novel solutions (e.g., for sampling from multiple sources and to enable multiple forensic finishes);
- Best practices;
- Validation expectations;
- Documentation standards;
- Quality assurance and quality control (e.g., blanks, proficiency tests, QC samples, spikes)—At what point is each introduced, and in what standardized manner?
- Experimental design—judgmental (targeted) and statistical;
- Information management (e.g., an expert should be able to easily access and analyze a file remotely), capabilities for which should be in place before an event;
- Accreditation;

- Training and certification—collaboration should be multidisciplinary (e.g., public health, veterinary, food safety, environmental science) as well as multisectoral (academia, industry, law enforcement, and defense); and
- Data records that should accompany a sample (e.g., temporal, spatial, custodial, storage) to aid epidemiological and forensic analysis.

Hamilton noted that there are a number of potential starting points for collaboration. These include the public health, veterinary, agricultural, food safety, environmental, defense, law enforcement, industry, and academic communities; international organizations with accrediting agencies and governing bodies; and sovereign states. He noted that two important questions are

- What standards or guidelines should be used as the starting point for microbial forensics?
- What sample characteristics are most critical to preserve for current and anticipated analytical methods?

## SPECIFIC ISSUES

### Food and Agriculture

The National Research Council report *Countering Agricultural Bioterrorism* summarized the historical evidence of planned or actual use of biological agents against livestock and crop plants as follows:

Historical evidence suggests that a number of nations have considered using or have even used biological agents against plants and animals. During World War I, German agents infected horses with bacteria that cause glanders, a fatal equine and human disease. The Soviet Union military also used glanders in the early 1980s during war in Afghanistan. In World War II, the United States, Great Britain, and others all had offensive biological-warfare programs directed against plants or animals, and these continued after the war until three years before the adoption of the Biological and Toxic [sic] Weapons Convention of 1972 (BWC). Some countries continued their offensive programs after 1972. Iraqi bioweapon development in the late 1980s and early 1990s included anticrop efforts based on wheat smut, a disease caused by a fungus in the genus *Tilletia*. There is also evidence that the Soviet Union had an extensive agricultural-terrorism program aimed at animals and plants. (NRC, 2003:18)

Fletcher et al. (2006) also pointed out that agriculture makes an attractive target for terrorism using plant pathogens because it (1) is a



\$1 trillion/year sector of the U.S. economy, (2) employs 17 percent of the U.S. workforce, and (3) is not subject to regular surveillance so that there could be long lag times between the deliberate introduction of a plant pathogen and its detection. Given the economic and health-related importance of agriculture, Fletcher et al. advocated for “the integration of the traditional discipline of plant pathology and the specialized field of forensic science. . . .”

With these thoughts in mind, Dr. Bruce Budowle discussed sampling and preservation in the food and agriculture context. His purpose was to provide a sense of the range of evidence sampling and collection possibilities that a microbial forensic investigation might confront and the outlines of a comprehensive plan for approaching such varied circumstances and challenges. A review of the major issues dealing with plant pathogen forensics can be found in Fletcher et al. (2006), which builds a case for the integration of plant pathology and forensic science. It describes the potential threats to plants and addresses the development of a program in microbial forensics and criminal attribution that addresses crops and other plant targets.

The challenges in microbial forensics might be greater than simply working with an isolate from an individual. The evidence may be a degraded sample on a laboratory floor or in soil or effluent from a clandestine laboratory. In cases where evidence is highly degraded, NGS technology may not work well whereas a real-time PCR assay may work better.

Natural outbreaks in food make good models for microbial forensics. During the passage from “farm to fork” there are many potential targets for investigation and evidence collection. The list in Box 5-3 gives an idea of the range of access for a potential attack. To investigate contaminated sprouts, for example, one may need to go back to the farm or to any of the many places the food has traveled along the way. There have been many examples of deliberate contamination of food: salad-bar tampering in The Dalles, Oregon; cyanide-tainted grapes; *Shigella*-spiked muffins. Each event comes with its own set of circumstances and questions. In the cyanide-tainted grapes case, for example, only two tainted grapes were ever identified. What happened to the others, or were there others?

Contamination can occur unintentionally through bad planning. A country may use mycoherbicides for biocontrol—to halt cocaine production, for example—but inadvertently cause more far-reaching toxicity. In Australia, a virus was introduced to reduce the population of introduced and invasive rabbits. Terrorists may be interested in such biocontrol agents or how they were effectively administered.

Evidence-gathering tasks can also be diverse. In 2003, the USDA Animal and Plant Health Inspection Service had to determine whether an out-

**BOX 5-3**  
**Potential Food and Agriculture Targets**

- Crops
- Grain elevators
- Water supplies
- Food in grocery stores
- Food and agriculture transportation systems
- Farm workers
- Animals
- Livestock producers
- Food processors
- Food handlers
- Processing facilities
- Trucks, railroads, ships
- Restaurants, *and more*

SOURCE: Budowle presentation, 2013.

break of mad cow disease in Canada originated with U.S. cows. Kinship tests were performed on cattle to trace ancestry, and evidence collection included tracking down the hides of parent cattle for DNA trace-back.

Plants tend to make good vehicles for an attack because the perpetrator is unlikely to be at risk when preparing the agent, and plants are poorly protected. Picture the huge areas of unguarded farmland around the world on which it is nearly impossible to impose security. This lack of security (which is impossible to impose) represents both a safety and a forensic challenge. Box 5-4 details some reasons why plants make such an excellent target for agroterrorism.

Various factors must be considered depending on the circumstances when collecting and preserving evidence. For example, in the case of a wheat streak mosaic virus, the mite vector may be implicated. A farm's wheat may be tainted, but it may not be recognized until 8 months later, after the harvest has been cleared. Does the virus exist underground in roots in the winter? Is it in other reservoirs that are in the vicinity, such as grasses? Sample collection under these circumstances is very difficult, and the difficulties do not end with evidence recovery. If samples are maintained at room temperature, for example, other microbes may take over and mask the one of interest. Under certain conditions, the agent may actually be destroyed.

A general approach to sampling appears in Box 5-5A, and a more focused approach—in this case, for wheat streak mosaic virus—appears in Box 5-5B. When considering approaches to evidence collection and

**BOX 5-4**  
**Plants Are Vulnerable and Efficacious**  
**Targets for Agroterrorism**

- Many agents are readily available in nature, from low-security laboratories, even from commercial sources, that require little effort or risk to obtain.
- Most agents pose no risk to human health.
- There is less risk in handling and dispersing the pathogen.
- Once released, an agroterrorism event may go unnoticed for days to weeks or longer.
- By that point, it may be nearly impossible to determine if the event was deliberate or occurred naturally.
- Tracking the perpetrator is more difficult.
- Vulnerable targets have low security.
- Moral barrier that perpetrator must cross is lower.
- Maximum effect may not require many cases.
- Mimicking natural introduction can be effective.
- Multiple point-source outbreaks can be initiated by contaminating imported feed or fertilizer, without even entering the country.
- Agronomic practices reduce the genetic variability and create conditions (large, dense populations) that facilitate disease spread.

SOURCE: Partially adopted from Burden (2010); Budowle presentation, 2013.

sampling, the technology that is available and appropriate must be incorporated into the logic of the process. A sampling and collection strategy should be structured, yet the structure cannot be overly rigid. One may decide, for example, to systematically collect a sample every 3 meters in a field. But if a tainted sample appears half a meter away, rigid rules would prohibit collecting the sample. A comprehensive plan should guide efforts (see Box 5-6), but owing to myriad possibilities that investigators may encounter, there is not a single sample collection and preservation strategy that is suitable in all situations. A key step in the comprehensive plan is to identify experts in advance so that when an event occurs, investigators can quickly develop consultation plans before they begin collection. Again, the chain of custody must be maintained throughout the process.

### **Environmental Contamination: The Public Health View**

Dr. Stephen Morse reviewed the needs of the public health sector in the event of environmental contamination, and provided a sense of how the site, what is known or not known about the extent of contamination, and the nature of the contaminant will direct the manner of sample collec-

### BOX 5-5 Sampling Strategy

#### A. General

- Logical and systematic
  - Air movement
  - Cross-contamination
- Work toward or away from source
- Scheduled
- Risk-based
- Targeted
- Statistical
- Known sources

#### B. Wheat Streak Mosaic Virus

- Develop plan with consultation
- Use experience of practiced collectors for how to collect (pattern and cutting and bagging)
- Maintain chain of custody
- Use targeted and statistical sampling plans—symptomatic/asymptomatic
- Collect mites (rolled up leaves)
- Collect samples from multiple plant parts, multiple plants
- Collect samples from other nearby plants
- Preserve and transport sample material (“breathing” of bags)
- On ice, 4°C

SOURCE: Budowle presentation, 2013.

### BOX 5-6 Comprehensive Plan for Sampling and Collection

- Develop a mechanism for quickly formulating a “consensus” analytical plan when a new sample (or set of samples) arises.
- Keep and update a set of standard operating procedures and validation data for analyzing case samples.
- Maintain a set of documented guidelines, requirements, and procedures for sample preparation for each analytical procedure.
- Maintain approved procedures for handling and storing samples.
- Develop standardized methods for data analysis, reporting, and presentation.
- Maintain reliable channels for sending and receiving samples.
- Maintain secure conduits for data, information, and discussion.
- Develop a mechanism for formulating an on-the-fly validation plan for a new procedure.

SOURCE: Budowle presentation, 2013.

tion and preservation. He views validating collection methods as a major challenge because the number of potential surfaces from which samples can be taken is unlimited.

From the point of view of high-level decision makers, in the event of an environmental threat, sampling and analysis will enable them to (1) determine who has potentially been exposed, (2) characterize the extent of contamination, (3) remediate indoor sites of contamination, and (4) clear the facility for reoccupation or use. A public health agency wants to establish certain facts, some of which also are sought by law enforcement. In fact, the FBI and public health officers often investigate a site simultaneously, collect specimens together, and share specimens and results.

The type and method of the release or event will determine the impact of sampling. If it is a covert event, and the first indication is sick people appearing in an emergency room, collecting environmental samples 1 to 3 weeks after the fact may not accomplish much. On the other hand, if the release point is determined, trace evidence and dispersal items may still be accessible. If it is an overt release, however, and the site of release can be identified, sampling and modeling can be used to determine the extent of contamination in a particular area. Contamination via food or water presents a further level of challenges.

Whatever the circumstances, identification of the crime scene and sampling are generally the first steps. The entire process—from sample collection, to transport, to analysis—must be validated. Investigators who understand the performance parameters of a process can effect superior analyses. After the Amerithrax letters event, the CDC and other responders received criticism from the U.S. Government Accounting Office about the extent of their ability to interpret performance of sampling methods. For example, does a negative result mean an area is free of contamination? Or could a sampling method simply be inadequate, and contamination still be present? Are the limits of the detection method clearly understood? This becomes very important when making a decision to reopen, for example, the Hart Senate office building after the anthrax attack. What degree of assurance can one give to leaders that a building is actually free of contamination? These questions are necessary for remediation but do not contribute realistically to a microbial forensics investigation.

Methods for sampling can be difficult to validate. The *B. anthracis* sampling methods have been evaluated, but these methods represented a low hurdle because the organism is stable in the environment. Less is known about sampling methods for other organisms, particularly Gram-negative organisms and viruses that do not survive for long on environmental surfaces.

Swab and wipe methods have been validated for nonporous surfaces,

and validation is under way for a vacuum method for porous surfaces. To further complicate validation, the sampling tool used is validated only for a specific area of surface. A swab, for example, is validated for a 2 × 2-inch square, and a wipe for a 10 × 10-inch square. Samples cannot be taken from a wall using a single swab because the swab may not be adequate for the sampling or at a minimum no validation was performed to guide the collection under best practices.

In the laboratory, one might encounter various types of surfaces, both porous and nonporous. Nonporous surfaces, for example, might be painted wood or wallboard or various types of wood, plastic, glass, and metal. But each of these nonporous surfaces has individual characteristics. A surface may have an electrical charge or may have micropores, which will affect the ability to remove microorganisms. Similarly, porous surfaces, such as fabric and carpeting, have their own characteristics. Although the number of potential surfaces is unlimited, it is only possible now to validate sampling methods for a small number and then extrapolate the experience as necessary.

Environmental conditions, such as humidity, affect the ability to recover organisms from surfaces. Although one can control such factors and validate methods for the laboratory environment, conditions out in the field are uncontrolled and will vary. Another variable is the person collecting the sample. The individual may or may not be well trained. Even if personnel are trained, they may not perform in the same manner every time, especially if sampling guidelines are not available. Moreover, the person is likely wearing personal protective equipment (PPE), which is hot and uncomfortable. Getting the sample quickly becomes the primary goal.

Transportation methods must be carefully selected. *B. anthracis* is a hardy organism, but some agents require a specific transport medium. Also, after decontamination, particularly via chlorine dioxide or vapor-phase hydrogen peroxide methods, a neutralizing agent must be added to the transport medium to remove the decontaminant, or anything that survives will die before it reaches the laboratory and will not be culturable. Recovery efficiency once in the laboratory is another factor, and again, it will vary from agent to agent.

Sampling approaches are chosen based on the situation. The easiest approach is judgmental, which is typically used in the early characterization phases of a response to establish whether an agent or toxin is present. In a meeting room, for example, an investigator would likely first sample from heating, ventilation, and air-conditioning vents in the wall, and from computer screens, which are great attractants for airborne organisms. When there is a large quantity of organisms or toxin, this sampling method works well. One of the simpler scenarios to handle is a white-

powder incident, where the agent is visible; standard methods exist for collecting white powder for analysis and forensic use. If, however, there is only trace evidence in a room, a random, probability-based method—sampling  $X$  number of places in the room in a random fashion, to provide a 95 percent chance or greater of picking up material—is appropriate. In general, the larger the room, the greater the number of samples required.

After decontamination, a combined judgmental and random sampling would be appropriate to ensure that the area is free of contamination. The U.S. Environmental Protection Agency (EPA) uses composite sampling. They collect samples, put them in the same tube, and analyze the contents. They simply want to know if there is contamination present, regardless of its location. This approach could apply to forensic sampling to reduce labor and cost. But it will reduce the ability to reconstruct the event with that/those sample(s).

In the event of a wide-area release, for example, if someone releases *B. anthracis* over a city, there is a limited time to remediate the area. According to statements made at several exercises, substantial delays in remediating the affected area would reduce the likelihood that people would return. Leaders will want to be able to tell the public when or whether it is safe for them to return, and that decision will also be based on sampling. However, the public could lose confidence in their leadership if any residual infections occur after they move back in.

Morse was asked to comment, in terms of validating collection methods, how much confidence he has that there are no spores when he gets a zero response from sampling a room, and how he establishes probability or statistical confidence limits. Morse responded that a negative result only means that there were no viable spores *detected*, but one does not know if there are spores below the limit of detection; current sampling methods will detect less than 1 spore per  $\text{cm}^2$  on a surface. The CDC feels comfortable saying that a negative response would be equal to less than 1 spore per  $\text{cm}^2$ . Although the analytical method may be able to detect a target at very low levels, sampling may or may not be efficient. Therefore, lack of detection may be related to collection method, sampling efficiency, and limit of detection of the analytical assay.

In the United States, the CDC has a network of analytical laboratories, the Laboratory Response Network, comprising 160 laboratories throughout the country. Similarly, the Food Emergency Response Network has protocols for analyzing food for the presence of threat agents, the Animal Health Network has them for agricultural animals, the EPA has an Emergency Response Laboratory Network, and there is a plant network. All use basically the same methods. The labs can process environmental samples using the exact same method; the results in one lab are equivalent to those in another lab. This is important because in the case of a wide-



area threat release, the nearest lab may be contaminated and unavailable. Samples collected using chain of custody are shipped to the closest lab for analysis. Morse pointed out that one problem with giving confidence limits is how to equate that to risk. Because the lowest dose that will infect a human is unknown for many, if not most, pathogens, it is hard to say there is no risk, which is never zero. Also, detecting something on a surface does not necessarily equate to an inhalational risk; it has to be able to detach from the surface and get into the air, to be breathed in. So it is difficult to say, "I am really confident that this room is safe to be in."

Budowle noted that calculating false-negative rates and the like is a problem because the approach is borrowed from human clinical genetics, in which a true negative can be defined. Minimum criteria should be established for the output of the selected analytical method, such as depth and uniformity of coverage. Defining output thresholds for metagenomic samples may be difficult given the immense quantity of data and microbial diversity; therefore, single-source samples and defined mixtures might be used as a guide. These limitations may be necessary in defining false negatives and false positives. Clearly, there will be ambiguous calls due to sequencing noise and novel genome composition. The specific parameters and settings used to establish thresholds, false-positive, and false-negative rates should be detailed thoroughly to enable sound interpretation and accurate comparison to alternative methods and protocols.

Keim said that the situation he encountered was detecting something above zero, but only in 1 in 1,000 samples. The question was whether that finding was informative or significant given that the sampling had never been validated to that level and the processing labs had never validated the sampling. They concluded it was contamination by the processing system. So one can have a positive result and still not be confident. A single colony of *B. anthracis* on a plate is still just one, and may or may not be meaningful.

### **Perspective from a U.K. Government Laboratory**

Dr. Cerys Rees works with the U.K. Ministry of Defense's Defense Science and Technology Laboratory based at Porton Down. This facility provides chemical and biological analysis capability on behalf of defense and security customers, including the military and law enforcement. The Laboratory's law enforcement experience includes the analysis that supported the first U.K. prosecution under the Chemical Weapons Act, which concerned the production of ricin. White supremacists in northeast England were arrested and convicted in 2009 for producing ricin in their home. Porton Down analyzed samples taken from the house, and by



combining several different methods, produced evidence that supported a conviction.

Proficiency testing is particularly lacking in the biological area. Dr. Rees's colleagues in chemical labs undergo stringent Organization for the Prohibition of Chemical Weapons (OPCW) chemical testing annually to maintain their designated laboratory status. Rees would like to see similar training in the biological area, perhaps starting with confidence-building tests, and then moving on to proficiency testing.

In the United Kingdom, a forensic science regulator who has overall responsibility for activities that support the U.K. criminal justice system oversees evidence sampling. The regulator may soon be granted statutory powers, which would enable him to enforce International Organization for Standardization (ISO) 17025 accreditation on all laboratories that support the criminal justice system. He also is very motivated to require ISO 17020 accreditation of crime-scene examiners and personnel who collect samples. It is the international standard of competence for inspection bodies, which may not fit perfectly with crime-scene examiners, but it is likely the only standard that is potentially applicable.

Rees noted that the Laboratory has frequently worked with defense and security customers to address the requirements for chemical and biological analysis and attribution capability that support the different communities' needs, which appear to be very similar. All require high-confidence information that enables defense customers to make strategic decisions and security customers to support the criminal justice system.

In their laboratories, validated tests are required to assess staff competence and to ensure that there is standardization in how tasks are performed (e.g., chain of custody) so they can demonstrate that the laboratory has high confidence in the information they produce.

Rees and her colleagues would like to pursue attribution based on all materials, not just the agent alone. They would like to develop methods that enable them to exploit the matrix and anything else associated with the sample, including traditional forensic materials, such as fingerprints, hair, and fibers. Her lab is working with police and forensic colleagues to develop these methods. Ideally they would provide information that, coupled with intelligence and situational information, could lead to a higher degree of attribution of an attack and perhaps identify the perpetrators. Dr. Franca Jones, U.S. Department of Defense, also said that there needed to be more discussion about other signatures in microbial samples that might be used for microbial forensics, such as media components. She encouraged that these nongenomic methods under development receive more consideration.

### Weapons Inspections in Iraq

Dr. Rocco Casagrande, founder and managing director of Gryphon Scientific, reviewed the technological challenges he experienced while supporting biological disarmament missions in Iraq for the United Nations. His experience provided an example of the kinds of problems those doing work in the field can encounter, especially in sampling. He emphasized that although his unit was not equipped as well as it could have been, most problems lay in how to apply good technology in an austere environment.

From November 2002 to March 2003 he served as a U.N. biological weapons inspector in Iraq. As chief of the biological analysis laboratory, he collected and tested most of the samples collected by his unit. At peak staffing, the number of inspectors did not exceed 25, and most were field or bioprocess specialists, so he had little assistance in the lab. The lab was small and rudimentary, containing a hood, a real-time PCR machine, and equipment to perform DNA extraction. Real-time PCR was the assay most frequently used owing to its simplicity, low false-positive rate, and excellent sensitivity. The process was labor-intensive because DNA extractions were performed manually. The inspectors were limited to the primer sets provided, and had no input in the choice of reagents. During the first month in Iraq, the equipment was inoperable because they had not been supplied with a transformer adaptable to Iraqi electricity.

Their immunological methods originally included hand-held lateral flow immunoassay devices, but their unacceptable false-positive rates for environmental samples prohibited their use. For toxin detection, they used conventional ELISA kits with antibodies prebound to 96-well plates. In addition to running the lab, Casagrande's duties included inspections, writing site reports, and interviewing Iraqi personnel. Only 4 hours a day could be dedicated to laboratory work; the remaining time was spent in the field.

In early 2003, equipment was found from a facility that had been destroyed by U.S. bombardment in 1991. There were allegations that the facility had been used for the manufacture of biological warfare agents, but the Iraqis claimed it was a baby-formula factory. After the facility was destroyed, the damaged equipment was moved to another site, and it lay outside for 10 years. Inspectors disassembled the damaged equipment, and samples taken from protected joints showed a borderline detection of *Brucella* species. The heterogeneous melting temperature of amplified fragments suggested multiple species were present, which would be consistent with natural contamination. A shortfall in inspectors' equipment, Casagrande noted, was that they were not permitted to pick positive controls, and they had very few. The sample from the damaged equipment proved similar to the *Brucella* species profile in a control Casagrande

created from the U.N. cafeteria milk. Therefore, the PCR technology was effective in detecting the trace contamination of dual-use items that had been blown up, moved, and then lay unprotected in the weather for a decade.

A great problem the inspectors faced, as in forensics in general, was where to apply the good technology and practices they had. Selecting sampling sites can be problematic; some sites are very large. An earthen berm, for example, had been built to encircle the Iraqi facility at Tuwaitha to shelter it from Iranian bombardment in the earlier Iran-Iraq war. From which sites of the berm should samples be taken? From which of the facility's 150 buildings, and from which equipment within those buildings, should samples be taken? The assay technology is effective, but what is needed is better technology to determine the reliability of the intelligence sources who direct inspectors to sampling sites. Better technology is needed to identify activities at a suspect site to determine which of the 150 buildings in Tuwaitha, for example, were of interest to the bioteam. Reliable background information is essential to direct inspectors' investigations.

In 2002, for guidance the inspectors used a mid-1990s Iraqi declaration of agents and activities putatively in their biological weapons program. This declaration, plus traditional threat agent lists, was used to direct procurement of reagents, such as PCR primers, specific for certain microbes. Inspectors were able to detect *B. anthracis* and orthopoxvirus, but not any pathogens not contained in the lists, such as Machupo virus and *Chlamydia psittaci*. Casagrande suggested that usable technology that could detect the presence of threat agents regardless of identity would be revolutionary for this purpose. The technology exists, but in its present form it would be difficult to use in a one-person lab with limited working hours and a single source of electricity. In one case, for example, there were allegations that the BBs within a larger weapon projectile had been coated with *Clostridium perfringens*, which causes gas gangrene, but inspectors lacked the technology to test for this bacterium, so the entire weapon was sent away for analysis. Casagrande suggested a need to search for dangerous traits of pathogens, or developing a more sensitive "zoo" chip.

The ability to identify modifications in an agent would also be extremely useful. It might yield clues about the intended target or use of a strain, determine if particular defenses are no longer effective, and indicate a degree of sophistication regarding the perpetrator. An example might be antibiotic resistance genes. For example, there are roughly 36 genes that encode tetracycline resistance; however, even if inspectors look only for that phenotype, the PCR analysis would need to be a highly multiplexed assay. Moreover, frequently only nonviable samples can be

obtained from a site. They may have been sterilized in place, bombarded, left outside, or subject to efforts to “clean” the sites. Samples often are contaminated with—and outnumbered by—environmental microbes, making analyses more difficult. Therefore amplification for genomic or phenotypic analysis is complex.

Inspectors faced this kind of challenge when collecting samples from R400 aerial biological bombs that had been filled with *B. anthracis* spores, botulinum toxin, or aflatoxin. It was difficult to get a culturable sample even from inside the weapon because the Iraqis had attempted to decontaminate the weapons with potassium permanganate. A challenge also lay in analyzing viruses and toxins. Many threat viruses have RNA as their genetic material, and RNA, even double-stranded RNA, is generally less stable than DNA, and thus is harder to detect. Because it requires an additional step, reverse transcriptase might also reduce detection limits. Another challenge lay in detecting protein and small-molecule toxins, such as botulinum toxin and aflatoxin, which were both on the declaration list. Inspectors used PCR to detect trace contamination of agent DNA because of its superior sensitivity and specificity over the ELISA assays. However, for a number of toxins of interest, such as ricin, inspectors hope to find DNA as well to indicate the presence of inactivated toxins. Although PCR detects DNA or RNA, not protein toxins, analyzing a sample by PCR might detect genetic sequences from the castor bean, even if the ricin had been inactivated or rendered undetectable. The Iraqis had legitimate castor oil plants, for example, and inspectors needed to determine if the ricin being produced was still active, if it was heat-inactivated, or if efforts were being made to extract it. A PCR assay cannot provide that information.

Finally, Casagrande added that one thing that was very important for the weapons inspection/biological disarmament mission was proficiency testing for people collecting samples from dual-use equipment. They needed to understand the equipment, know how to disassemble it and how to sample from the proper places inside the equipment. When inspectors confused very similar looking equipment or sampled from locations that had already been adequately decontaminated, they ran the risk of drawing falsely negative conclusions. Standardization, protocols that are widely disseminated and understood, and testing for the ability to take samples from dual-use equipment would be desirable.



## 6

## Validation and Reference Materials for Microbial Forensics

In the introduction to their chapter on validation of microbial forensics, Murch and Bahr (2011:649) state the following:

The use of effective, robust, and properly validated methods for the collection, preservation, transport, analysis, interpretation, and communication of probative evidence is a linchpin of reliability and confidence-building measures that contribute to the acceptance, use, and understanding of science by investigators, judges, attorneys, juries, the media, and lay public. Stakeholders expect that forensic methods, protocols, and techniques have been validated properly. All science proposed and admitted to court is subject to discovery and scrutiny under U.S. case law and prescribed legal procedures. In recent years, as courts and the media have become more aware of the value, power, risks, and uncertainties of forensic science, whether or not methods have been validated properly is receiving increasing attention.

Microbial forensics is a relatively new discipline and the science and technologies it uses are evolving rapidly. Because of the demands to produce and apply reliable and robust capabilities, validation measures, requirements, and protocols are essential. Dr. Bruce Budowle, one of the early developers of microbial forensics, and his colleagues Steven Schutzer, Roger Breeze, Paul Keim, and Stephen Morse produced a *Microbial Forensics* textbook, now in its second edition (Budowle et al., 2011), which is used as an informative resource regarding many aspects of microbial forensics. In an authoritative presentation at the Zagreb workshop, Dr. Budowle reviewed established and recommended measures and

approaches for validating components of each stage of microbial forensics processes. He provided insight into gaps in validation coverage and the challenges of designing protocols in a system encountering substantial and often undefined diversity and sample variation.

Dr. Budowle began by stressing that a microbial forensics investigation does not have to produce a “smoking gun” solution to be informative. It needs only to provide a conclusion to the level necessary for producing a useful piece of a puzzle. The goal is attribution, which is

- 1.) To establish the exclusion or possible inclusion of the source of a sample. The level of attribution can be determining the species or strain of a microbe to the individualization of a particular sample, that is, coming from a particular isolate or test tube. The latter is not readily achievable with currently validated genetic typing capabilities; and
- 2.) To integrate the microbial forensic findings with other forensic and investigative evidence to address the ultimate question of guilt or lack of guilt.

He outlined four undesirable scenarios for the application of a methodology:

1. *Applying a bad method and doing it poorly.* Operators are often unaware their methods and performance are poor. The information generated lacks a firm basis and is frequently incorrect.
2. *Applying a bad method well.* Operators believe the data are good because they followed standard operating protocols carefully, but the data may be meaningless or, worse, erroneous.
3. *Applying a good method poorly.* Operators produce content for which the criteria or limitations for use are undefined or compromised or, worse, erroneous.
4. *Applying a good method well, but the science community or stakeholders do not accept the data.* The operator may, for example, have validated the method without documenting the validation, so there is no confidence in the system, or there have been documented problems with the particular laboratory and there is a lack of confidence even though the procedure(s) were carried out properly.

The goal, of course, is to perform a good method well. Budowle noted that a great deal of focus is placed on technology. Empowerment by using sophisticated technology can be deceptive if its use is not validated. Moreover, while validation is necessary, perceptions of what constitutes validation vary. In practice, the basis for validation may be as tenuous as “because I’ve used it for so long it must be the right thing,” or “based on my experience of doing this three times” or “it is a community validated method so my use is valid.” Because there are no absolute criteria, and

each application may have unique requirements, validation is a challenge. The goals for validation in microbial forensics are to

- Promote development of a program that is scientifically *valid* and *rigorous*.
- Define criteria for development and *validation* of methods that will support *attribution* for criminal investigations.
- Establish *national/international* working guidelines for *quality* assurance and quality control.

There are two general types of validation for microbial forensics: developmental and internal. Both essentially have the same goal: to define limitations of use and/or interpretation.

1. Developmental validation refers to the acquisition of test data and the determination of conditions and limitations by the developers of the method. Measures or determinants of developmental validation would include specificity, sensitivity, reproducibility, bias, precision, false positives, false negatives, determining appropriate controls, and the choice and/or quality of reference databases (Murch and Bahr, 2011).
2. Internal validation refers to accumulation of test data within the laboratory that intends to use the method to demonstrate that established methods perform as expected. Methods must be tested in one's own laboratory to ensure that they perform as expected and that limitations are understood, so that interpretations do not cross the boundaries of the abilities of the methodology.

Because microbial forensics may be applied in exigent circumstances, "preliminary validation" may be considered. For example, during the U.S. anthrax letters case when a research tool had to be applied to an investigation, Paul Keim's multiple-locus VNTR analysis (MLVA) was used to determine that the *B. anthracis* strain was Ames. In the midst of such an event, a result was needed in a short time to protect human health, and it was unreasonable to wait until the laboratory can validate a procedure to the degree normally desired. Response time does not allow for such a lengthy process. The criteria recommended for considering "preliminary validation" are (1) exigent circumstances, (2) an inability to wait for some lengthy time period to completely validate the method, and (3) the acquisition of limited test data that will enable an evaluation of a method that provides support to investigate a biocrime or bioterrorism event.

One caution raised was being so tied to a protocol that the scientist does not think about the assay and its result and operates more or less



as an automaton. Standard operating protocols (SOPs) are for routine work. A danger of SOPs is that they can restrict analytical thinking and disregard both inculpatory and exculpatory evidence. The ideal balance to strike is one that enables thinking “out of the box” while still applying boundaries on use.

The minimal criteria that should be addressed for validation are

- Specificity,
- Reproducibility,
- Precision,
- Accuracy,
- Robustness,
- Analysis of specified samples commensurate with the intended application of the assay (e.g., reference panels and mock or non-probative materials), and
- Limit of detection.

In his experience, Budowle has encountered scientists who contend that there is no need for a quantitative element in validation when the test result “is a yes/no answer.” He stressed that this view is absolutely incorrect because one must quantify the limits of detection in order to draw inferences about the meaning of what one is trying to assess. An assay will require some level of quantification.

Validation measures should be applied from the very first stage of the investigation—sample collection—and through all subsequent stages of the process, including shipping and storage; extraction, which greatly impacts results; analysis; and finally, interpretation. It is critical that validation be applied to the interpretation stage. Interpretation is dependent on, for example, which databases and procedures are employed for drawing inferences; a value can change based on the database or inference framework used.

If evidence is not collected, it cannot be tested. If it is not collected correctly, results may be suspect or destroyed. The “proper” collection of evidence is not a new concept; clearly, mixing evidence items in a bag is unacceptable. Storing evidence improperly also will be problematic as crucial targets may be destroyed.

Microbial forensics relies not only on microbial evidence but also on additional forensic evidence, and the collection of microbial evidence can sometimes be incompatible with other types of forensic evidence. Given a choice between a fingerprint and a spore, the level of individualization offered by the fingerprint could make it the more desirable piece of evidence to preserve. Swabbing may obliterate the fingerprint and

destroy valuable evidence. Triaging strategies must be incorporated into the sampling process.

A challenge is to develop investigative protocols that optimize the choice and priority of methods. Among factors to consider are the amount of sample available, evidence preservation/conservation, and trade-offs between speed and accuracy and/or precision. Given the limitations of wearing protective equipment speed may become a requisite over accuracy and chain of custody. Making priority decisions given collection constraints should be determined ahead of time and methods developed to minimize the loss of sample and less than desired collection and documentation. Simple strategies, such as pre-labeling tubes before beginning sampling due to difficult manipulation of labels when wearing protective gear, is an example of a simple consideration.

Collection validation must address

- Recovery,
- Stability,
- Integrity,
- Target—organism or analyte (e.g., DNA, RNA, protein, toxin, agar), and
- Influence of sample matrix.

As an illustration, Budowle pointed out that the ideal swab is the one that adsorbs the sample well. But the worst swab for extraction of the target is the one that adsorbs well. A swab that adsorbs well may be better for recovery from surfaces at a crime scene, but the swab matrix holds the target analyte such that it is not efficiently removed from the swab during extraction. The result can be low yield of a sample even though sufficient material was collected at the crime scene. Selection of collection tools often requires that a balance be struck. Samplers must work out these details and ensure that the application aligns with final objectives. Collecting an isolate from a victim is quite different from going into an air duct to collect trace samples. Good procedures exist for collecting, preserving, and shipping blood, but they do not exist for handling a flask of liquid that no one wants to open during collection. It may be necessary to make an educated guess about the target in the sample or the nature of the sample may dictate the best option for preserving the sample. Options to preserve the contents of a flask can vary from maintaining it as a liquid or frozen. There may be presumptive testing that can be performed at the scene or intelligence to guide this determination—rightly or wrongly. If there is a substantial amount of material, it may be desirable to store it multiple ways to increase the odds of preserving the target of interest. Procedures must continue to be developed.

Extraction efficiency depends on the collection medium. Different extraction procedures can be productive or counterproductive in obtaining the particular target medium. When choosing an extraction procedure, analysts must consider

- Specific target—virus, bacteria, fungal, toxin;
- Spore vs. vegetative cell;
- Active vs. inactive (culture plan, amplification plan);
- Analyte that will be assayed—DNA, RNA, protein, lipid, stabilizers, media, fatty acid, etc.;
- Stability of analyte;
- Matrix effect—substrate, other co-extracted analytes, materials such as soil; and
- Downstream assay impact.

Budowle emphasized that it is essential to consider what effect extraction methods will have on downstream analysis. There is no point in extracting material with a particular method if co-purifying compounds impedes the function of an assay downstream.

The basic criteria for validation during the analytical stage appear in Box 6-1A, and additional criteria may be warranted. The amount of target that can be analyzed by an assay should be defined. It is not enough to implement an assay; it is essential to define the acceptable range. A low amount may be acceptable, but associated stochastic effects may affect interpretation. The critical reagents should be defined. Often assays are developed to perform in optimum ranges and performance criteria. Optimization can be misleading as it implies maximum performance. However, a window of performance criterion is established for robust methodologies. It is usually undesirable to design an assay that is at the edge of a window of performance because any small perturbation in some part of the process could cause it to fail. Windows can be set for criteria, for example, as simple as the temperature for annealing primers during PCR. If the highest temperature for successful primer annealing is 63°C it would not be best for a robust assay to set the temperature at 63°C. Given that thermocyclers have some variation, it may be desirable to set the temperature slightly lower so that the temperature in the tube is favorable for primer annealing. It is best to find the range for the performance criterion, and use a condition somewhere within the performance range of the analysis but not at the edge—so it will perform robustly.

Another point is to consider a real-time PCR assay with a linear range between 20 and 30 cycles, and the user observes a positive result somewhere between 40 to 50 cycles. Since the result is no longer in the linear range, one cannot predict quantity of the target, but there is a posi-

**BOX 6-1**  
**Validation for the Analysis Stage**

**A. Basic Criteria**

- Standard operating protocol
- Control—positive, negative, and inhibition
- Specificity
- Dynamic range
- Reproducibility
- Reliability
- Precision
- Accuracy
- Predictive value
- Critical reagents
- Sensitivity
- Limit of detection
- Window of performance for operational steps of assay
- Critical equipment and calibration
- Interpretation criteria for results
- Resolve conflicting results

**B. Broader Criteria**

- Resolution
- Robustness
- Specified examples
- Purity
- Input values
- Quantitation
- Dynamic range
- Basic error rate
- Databases

SOURCE: Budowle presentation, 2013.

tive result. The challenge becomes in determining the meaning of a positive result that may be due to a single molecule. Highly sensitive assays may detect a signal but the reliability or confidence in the result may be suspect.

An important issue that Budowle believes receives too little attention is the process for handling conflicting results or considering alternative hypotheses. Conflicting results will occur. Good scientists should always question results. But in the face of conflicting results, the action to follow could be to verify with orthogonal testing, if available. Alternatively, reports should contain language about results that may be questionable or that alternative hypotheses may be supported.

Budowle and colleagues developed additional criteria (Box 6-1B) for the analysis-stage validation to supplement basic criteria. Their intent was to provide test/data recipients, funders, or stakeholders with a more comprehensive list of criteria. Being more informed, one could ask if appropriate validation criteria were addressed with the development of a method. It is the responsibility of the developers and users to demonstrate why they chose the criteria they did to evaluate and validate the method and, just as importantly, explain why they rejected what they thought unim-

portant. Thus, more accountability and documentation will be associated with “validated” methods. Not all criteria listed would apply in every situation. While testing, for example, specificity across a range of species may be necessary to validate a hand-held assay; once implemented the assay would not require a comprehensive database to effect interpretation. It may, however, apply to determining whether or not the Ames strain is common or rare in a particular circumstance.

With a new technology or method, the results often are compared with the results of an existing technology—for example, comparing MLVA and canonical SNPs to determine if they correlate and where they do not. One would expect that at some point, one technology may resolve better than the other and that observation should not be considered discordant data in itself. If assessing an entirely new technology for assaying a target and generating results with no comparable existing data, one must define the criteria by which to test it, develop the test, and validate the performance. Two approaches to evaluating new technology are portrayed in Figure 6-1.

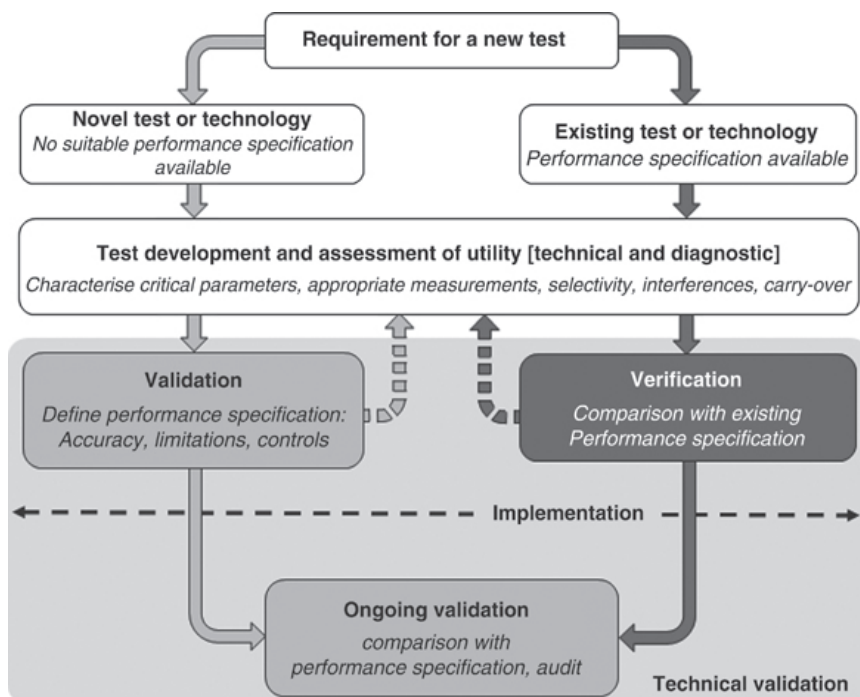


FIGURE 6-1 Process of implementing a test for diagnostic use.

SOURCE: Mattocks et al. (2010).

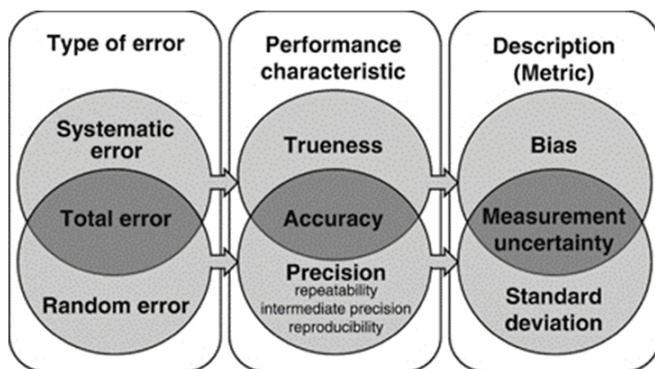
SOPs must be created, and once an entire system is created, another validation must be performed. The SOPs should contain sufficient detail, including, where appropriate,

1. All steps in the procedure,
2. Proper controls,
3. All reagents and preparations,
4. Calibration,
5. Criteria for analysis of results, and
6. Interpretation of results.

Analysis and interpretation are two different things. Analysis could be “I have a positive result.” Interpretation would be, “It is ten times more likely to observe this result if X compared with if Y.” Budowle emphasized that there are feedback validation mechanisms in both approaches.

Validation is an ongoing process; it does not stop once the method is up and running. To illustrate its importance, he described an incident in which a lot of commercially available human DNA identification kits were sold to customers, and subsequently the manufacturer discovered that the deoxynucleotide triphosphates (dNTPs) could degrade while kits sat on the shelf. The company sent letters to crime laboratories apprising them of this possibility, but some laboratories continued using the kits. The signal of results was dropping (i.e., a loss of sensitivity of detection) and in actuality could have presented negative results. This outcome is a very serious problem when working with small amounts of DNA. The positive control supplied with the kit also was diminishing in signal intensity. Yet some users ignored it or were not cognizant of the signal loss. Instead, as long as they had a qualitative call that was consistent with the profile of the positive control, the users were satisfied that quality control was acceptable. Both qualitative and quantitative signals should have been monitored. Had the users done so, it would have been evident early on that the system was not performing appropriately and the assay could have been halted rather than consuming valuable evidence.

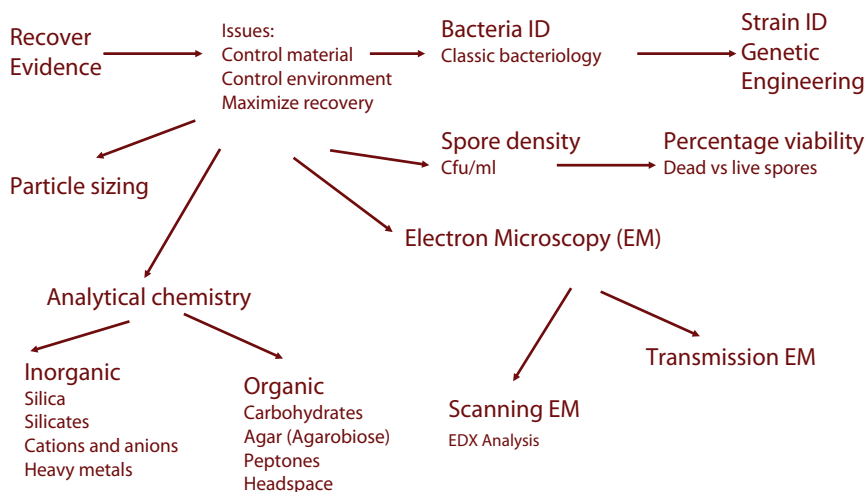
Error is an important consideration because error is ubiquitous. Different types of error are shown in Figure 6-2. There is, for example, systematic error, which includes measurement error. There also is bias, both in methods and in the individuals interpreting the data. It is human nature to want to accommodate data to scenarios that make sense to us, regardless of alternative hypotheses or explanations. This human nature affects how we perceive false positives and false negatives. There is no absolute point where the line can be drawn between the two false categories for acceptable performance. As the line shifts in one direction or the other, more false positives will be obtained at the expense of false nega-



**FIGURE 6-2** Performance characteristics, error types, and measurement metrics. SOURCE: Adapted from Mattocks et al. (2010).

tives, or vice versa (see Mattocks et al., 2010). The reality is that a decision is made based on an evaluation of the necessary risks related to the application at hand with available resources, and these must be defined.

The sample analysis flowchart in Figure 6-3 provides an idea of the variety of microbial analyses one might consider. Investigators might perform classic taxonomic analysis, for example, and analyze for weaponization engineering. They also might examine spore density. If a sample



**FIGURE 6-3** Sample analysis flowchart. SOURCE: FBI Quantico laboratory.



is collected from a host, this analysis might not be very helpful, but if it is collected from a package found in a train station, it might be useful.

The range of possible analyses in the flowchart conveys the challenge encountered with microbial forensic evidence. The field comprises a number of subdisciplines, each requiring substantial expertise. Moreover, it illustrates why next-generation sequencing (NGS, also known as massively parallel sequencing), is so appealing. Many other techniques have been designed to identify a specific target or targets; but they require knowledge about the target to be able to design the assay. But with NGS, the technology can target all bacterial and viral select agents using the same basic approach. NGS also provides high throughput, which can translate into a high sensitivity of detection and resolution. Higher throughput can result in greater depth of coverage, potentially allowing for detection of low-abundance or trace-level targets. Therefore, NGS will likely become one of the primary tools for microbial forensics. To use the technology, or for that matter any of the microbe identification methodologies, a panel of isolates is required to validate the assays. Much thought should go into the appropriate panel to test/validate the tool. The panel should be broader than two samples, but narrower than all possible samples, because the latter would be impractical. The challenges will be to define the panel breadth, how to obtain the materials, and who will make them available.

Many procedures for collection exist but have not been compiled, some have never been disseminated, and some have not been validated. These protocols, however, should be leveraged. Budowle believes that the literature, particularly older literature, should not be ignored. Access to older literature is needed because some perpetrators may not have the same resources that advanced laboratories have for routine microbiological research. It may be delusional thinking to assume that someone will go into a BSL-3 facility to produce a weapon when that someone may actually be in a garage using a washing machine as a centrifuge. If we are not familiar with the old procedures, we may miss telltale signs that a weapon is being developed, miss invaluable leads, and may not collect informative evidence. Digital archives could be created to preserve older literature. Workshop participant Jens Kuhn of the National Institutes of Health (NIH) agreed that access to older literature is very important. Moreover, he believes that researchers who rely solely on PubMed for research and database creation may be unaware that there is a bias in this resource. The journals of many countries and even the literature on specific diseases are underrepresented in PubMed. PubMed is, for example, of little value if you seek case representations on hantaviruses, whereas some countries produce much literature on the subject. Another problem is that younger scientists may mistakenly believe that older literature has little value. It



is important that libraries receive funding, and that libraries retain their old relevant journals, which even the NIH library finds difficult to do.

Budowle believes that interpretation is perhaps the most critical step requiring validation because there is a tendency to become unthinkingly trusting once a protocol is running. For example, in a real-time PCR assay, if a signal is obtained at 48 cycles, it may not have much value, or it may have some lead value. With enough amplification cycles, it is likely everything will become positive. There is a need to have an interpretation of “inconclusive.” He pointed to a situation in 2005 in which a partial signal was indicative of the presence of *Francisella tularensis* in the area of the National Mall in Washington, DC. Four positive amplicons were considered sufficient for identification of *F. tularensis*; however, only two were positive, and there was no clear interpretation for a partial profile and what actions to take.

Budowle believes the concept of “missing data” is also an important factor to be considered when interpreting results and should be part of the validation equation. As an example, he recounted an anecdote about the process the Allies in World War II followed to protect bombers from being shot down during bombing runs in Germany. This anecdote may or may not be true, but either way, it is instructive. Returning planes were examined, the location of bullet holes mapped, and vulnerable areas were reinforced with shielding based on the location of the bullet holes. Much later, it was pointed out that the planes that returned were not necessarily informative. The planes that did *not* return should have been the focus. The investigators were missing key data. Budowle sees this as an ongoing challenge in microbial forensics because a lot of data is missing. Interpretations must be made using limited data, and biology is mutable. The sample an investigator tests today may be linked to something ten passages back, or may have traveled through a host and mutated. We must accept that we lack data, yet build around this to make the process work as well possible. While we can never have absolutes, we must still produce conclusions, but they should be tempered appropriately.

Considerations to help guide this interpretation process are listed in Box 6-2. Putting a confidence limit on a conclusion can be very difficult. In some cases, a qualitative statement may be all that is possible and may be appropriate, as it was in the Amerithrax case in which there was no accompanying statement of a percent confidence that the microbe was the Ames strain. Limitations are a reality. An answer with limitations is not a wrong answer; it simply reflects what we know and can do. Even with a qualitative statement, some verbal constraints are necessary on the strength of the evidence or extant data support. In addition, saying, “I am 92 percent confident of X” may not convey the right information for the circumstances. Instead of quantitative or qualitative statements, we may

**BOX 6-2**  
**Interpretation of Results**

- Statements—qualitative, quantitative, semiquantitative
- Database—type, relevance, representative, quality
- Background data—normal values, reference range, endemicity
- Does a result require follow-up or further analysis—temporal/spatial analysis, effect of passage
- Limits of interpretation
- Statistical approach—match, similarity, most recent common ancestor, identical
- Thresholds
- Software
- Alternative explanations

SOURCE: Budowle presentation, 2013.

need qualitative statements with extra information to guide the end user or other scientists who may review the analysis.

Databases also have limitations. Our databases comprise samples of organisms that infected something. The organisms that did not infect anything may be just as prevalent but have lacked the opportunity or ability to infect. This reality should be considered during the investigation process. We do not know what is endemic, and this can be an issue in the overall interpretation process. If we identify a “match” or a “similarity” or the most common ancestor, the significance of these associations must be supported by something defensible.

Thresholds are another challenge. Thresholds are critical not just for the analytical phase but also for the bioinformatics phase of NGS. Budowle stressed that there always can be alternative explanations and that good scientists should confront them, position them appropriately, and through good hypothesis testing determine whether they have low probability or high probability of support. Especially in forensics it is crucial to follow the scientific method, which is to attempt to disprove one’s hypothesis rather than to prove it.

Some of the major questions that might be asked during a source exclusion, association, and attribution analysis appear in Box 6-3. Perhaps instead of asking whether samples are the same or different, we should be asking “how different are they?” What conclusions can be made, and how do they associate in an evolutionary context? In contrast, one might argue that the question to ask about the evidence in the Amerithax case—the

**BOX 6-3**  
**Challenges for Any Target:**  
**Source Exclusion, Association, and Attribution**

- Are these isolates the same or different? What are the discrimination criteria?
- How dependent is the conclusion on samples taken, method used, and interpretive approach?
- What would an exclusion, association, or “match” really mean? How precise can or should one be?
- Do the methods and interpretation account for variation, evolutionary change (genome dynamics), and influences imparted by environment and ecology?
- How should significance to a comparison be assigned? What are the confidence limits? Can a probability or likelihood be assigned?

SOURCE: Budowle presentation, 2013.

flask marked RMR 1029—might not have been an evolutionary question since the flask’s contents comprised a mixture of samples. A more appropriate question may have been “How probable would it be to observe these variants if they arose from a single source versus from mixtures?” A great quantity of epidemiological data exists that can be consulted to help us understand associations.

Forensic questions include “What is it, and was the release intentional or natural?” How the threat agent was made can be a critical piece of evidence for an investigation, which in turn can feed back to help answer other questions and also may provide a lead to the perpetrator’s identity.

Forensic questions based on genetics-based analyses appear in Box 6-4. Investigators want to know, “Is the information provided with genetic markers probative? Is it meaningful?” For example, in a foodborne pathogen investigation, one might ask if there is an expectation that *E. coli* will be in a room besides what we carry in our bodies? The probability of *E. coli* being in a room occupied by humans (and animals) is high. What the appropriate level is may be unknown but it could affect the confidence of our conclusions.

Richard Vipond of Public Health England, Porton Down, offered an observation based on his experience of trying to take PGM (Ion Torrent Personal Genome Machine™) through validation for metagenomics work. His lab wanted to “freeze” the technology in time to assess performance and result quality. Because the manufacturers are in competition with each other and the technologies are evolving rapidly, every couple of months there is a release of a new version—for example, a chemistry

**BOX 6-4**  
**Microbial Forensics Questions: Genetics**

- What might be deduced concerning the *nature and source* of the evidentiary sample?
- Is the pathogen detected of *endemic* origin or introduced?
- Do the genetic markers provide a significant amount of *probative* information?
- Does the choice of markers allow effective *comparison* of samples from known and questioned sources?
- If such a comparison can be made, how definitively and confidently can a *conclusion* be reached?
- Are the genetic *differences* too few to conclude that the samples are not from different sources (or lineages)?
- Are these differences *sufficiently robust* to consider that the samples are from different sources?
- Is it possible that the two samples have a recent *common ancestor* or how long ago was there a common ancestor?
- Can any samples be *excluded* as contaminants or be recent sources of the isolate?
- Are there *alternative explanations* for the results that were obtained?

SOURCE: Reprinted from Hunt et al. (2009) with permission from Elsevier.

or server upgrade—and this can dramatically change performance. The Porton Down laboratory has seen error rate, critical to any decision it makes, decrease with quality enhancements, and then increase with the introduction of a novel step. It is quite difficult to freeze assessment. The lab cannot even go back and rerun tests using the same reagents and software because these materials will not be available 6 months later. However, the rapidly evolving technology environment must be embraced or there will be a risk of reducing capabilities.

Deep sequencing generates substantial data, and sequencing capabilities are enabling scientists to develop elaborate databases tailored specifically to the demands of a case. This capacity demands novel ways to handle and analyze data that are now routinely obtained in terabyte amounts. Given the immense quantity of data, bioinformatics is an absolute necessity. It is unlikely (at least in the United Kingdom), however, that bioinformaticians will be on staff in application-oriented laboratories, so investigators will need to validate the bioinformatics pipelines and there will be a demand that the pipelines be sufficiently robust.

Factors to consider that affect data interpretation and quality include

- Quality metrics of sequence data;
- Sequence errors and uncertainties;
- Reliable standards for genomic data representation;
- Uncertainty with databases used;
- Inferences based on available data, including metadata;
- Formulation of well-defined hypothesis/hypotheses: testing methods for assessing the weight of microbial forensics evidence;
- Criteria for comparisons: match, similar, different, inconclusive;
- Rigor of reasoning by the expert; and
- Software.

Quality metrics are another important issue. What criteria should be used? Budowle pointed out that while some analysts use Q20 (a quality metric score associated with a base call) for base calling, others use Q30 (a higher quality score by an order of magnitude than Q20). He suggested that neither by itself may be sufficient, and validation is necessary to guide the user on how best to use quality metrics. For example, if the user were to encounter 500 reads with a Q9 score for one base in a homopolymeric stretch in all these reads, despite the low Q score, there may be high confidence associated in this single base with complete representation because the low score might be the result of a chemistry artifact. The PGM system is more refractory than the MiSeq with sequencing of homopolymers but with so many reads (or fragments) that the data may still be reliable because of the quantitative representation. Such considerations will be necessary to validate so that data can be used as effectively as possible. It is very difficult for one lab to replicate all essential details of any bioinformatics pipeline used by another lab because there are so many factors that affect outcomes. Attention to documentation will be requisite.

The extraction method one uses can cause variation of the result, as can the amplification method, including enrichment, PCR, and primer selection. The primers may seem to be an obvious consideration, but the enrichment process or capture method may not.

There are also differences in sequencing techniques. One technique will show a gap in a sequence whereas another will show a base. We must consider how to resolve such differences. Budowle's lab runs analyses on both MiSeq and PGM so he can exploit orthogonal chemistries to help determine what is reliable and valid. Neither system is immune to problems that need to be addressed. This testing enables one to improve both systems. Standards and controls must be created to better assess such concerns.

Box 6-5 lists the components of a "bioinformatics genetic toolbox" for microbial forensics. Budowle noted that there are processes in which alignment is used, and processes in which it is not; the processes are

**BOX 6-5**  
**A Bioinformatics Toolbox for Microbial Forensics**

- Assemblers;
- Aligners;
- Phylogenetic algorithm(s) for clonal and sexually inherited markers, recombination, gene conversion, and horizontal gene transfer;
- Capability to identify informative markers and their power to address specific forensic issues;
- Better understanding of mutation rates and the effects of environment and host on these rates;
- Discrimination and match criteria to quantitatively interpret results with confidence bounds;
- Endemicity;
- Capability to relate diversity to function;
- Capability for comparative and functional genomics;
- Ability to contain or access curated (genetic marker) databases on pathogens and near neighbors and their background occurrence with epidemiological history, when available; and
- Data management with the capability to access and process large amounts of diverse genetic data and to communicate data rapidly with stringent informational security (i.e., fully functioning information interoperability).

SOURCE: Budowle et al. (2005).

different and must be validated. Assembly may be sought for simple samples, but currently would be extremely challenging for metagenomic samples. Phylogenetic algorithms differ and also must be validated. Validation also will be an issue in data management. An enormous amount of data is being generated, and there is a recommendation or trend in the greater science community toward deleting raw data and simply saving the data at the fast level. This approach may be practical and economically appealing, but concordance testing may require that the biological material initially used for sequencing be archived. Otherwise, there may not be proper comparisons made as modifications or novel approaches are developed. Note that the large amounts of data common today require new bioinformatics algorithms because old tools (e.g., BLAST) will not scale to handle large data needs.

There are inference and error validation concerns, as well (Box 6-6). The base error rate of the particular sequencing protocol used must be defined. Moreover, there are different kinds of error. Sequencing errors, which vary from site to site, may occur as a result of chemistry and soft-

**BOX 6-6**  
**Inference and Error Issues**

- Forensic analysis of whole-genome sequence data often will compare two or more sequences, for example, an evidence sample profile with that of a reference sample that may be considered a direct link or have a common ancestor.
- Sequencing error and other factors will most likely inflate the dissimilarity between samples, creating a degree of “uncertainty” to some extent.
- Defining and quantifying the error rates associated with each platform/chemistry is critically important and includes extraction, amplification, library preparation, software.

SOURCE: Budowle presentation, 2013.

ware. Alignments can cause a great deal of noise, and there is other noise that cannot be identified solely by simulation, for example, with a metagenomic sample analysis in which chimeras have been created.

When scientists perform 16S rRNA diversity studies, everything that differs from what has been previously seen is considered “new”—another organism—yet Budowle notes that this might simply be junk. The diversity of what has been identified may be overstated or understated because of stochastic effects in the process.

Validation of materials is essential to a successful microbial forensics program, both for cross-comparing data and for running routine tests. Controls—reference samples, panels, and reagents—are required and must be accessible. The magnitude of the problem of developing material standards is illustrated by the wide variety of targets, which include but are not limited to genes, proteins, morphology, physiology, and biochemistry. Reference materials are needed for all developers and analysts. Moreover, standards will change. The Ames strain has become a sort of standard for *B. anthracis* only because it was used as a weapon. But how we define or select a standard for each species should be made more judiciously. A single strain may never suffice as a standard, but a standard must be established at some level because using 50 standards in a run with one sample may not be practical and currently would be costly. Establishing standards becomes a process in itself.

The number of technologies is wide and expanding. Among them are NGS, other DNA/RNA assays, traditional morphological and biochemical identification approaches, SEM/TEM, micro-Raman spectrometry, atomic force microscopy, isotopes, secondary ion mass spectrometry, particle-induced X-ray emission, and field immunoassays. A single method cannot

address all targets and all evidentiary materials that will be encountered. Nor can one validation method or standard address all technologies; each technology must be validated individually.

There are two types of standards: performance standards and material standards.<sup>1</sup> Most often, users focus on material standards, but performance is equally important. Traceability, a documentation of measurements to some standard, is needed. In the United States, standards developed by the National Institute of Standards and Technology (NIST) are typically used, but the available standards do not meet all needs. We need to decide who will make the standards to bridge the gaps in what is available. Some of the gaps in standard reference materials appear in Box 6-7. A major concern is how to identify a good standard versus “just a standard.” For example, surrogates may or may not suffice; they are often, but not always, suitable (Anderson et al., 2005). Near neighbors will be required to describe the sensitivity and specificity of an assay. Budowle also posed the question of who will prepare these standard reference materials and who will maintain them? Is it government’s, industry’s, or an individual laboratory’s responsibility? Likely the responsibility will vary based on the need.

Whole-genome sequencing of a metagenomics sample will generate sequences that span the genomes of the microorganisms that reside within the sample. In metagenomic samples, species and/or strains can be represented in widely varying abundance. The limited depth of coverage and amplification-bias stochastic effects on portions of individual genomes might affect representation of critical sequences that define species, causing them to be missed. Part of a genome will likely be represented, but phylogenetically and genetically the parts that are detected may not be able to resolve at the species level (most importantly near neighbors) or even at higher taxonomic levels. An analytical assay may perform perfectly well but may be uninformative on the presence of a target even if it is truly in the sample. So, we have to consider this process of resolving at the near-neighbor level and understanding which sequence reads are informative for taxonomic resolution or classification. One should rightly pose the question, whether only one, two, three, or more reads are sufficient to render an interpretation of identification.

Primer development and the quality of primer synthesis require validation as well. Primer design programs do not validate primers, and the need for such validation is underappreciated. Similarly, bar coding or

---

<sup>1</sup> A performance standard specifies what is to be accomplished but does not dictate the particular method or material to be used as long as the desired end is achieved. A materials standard specifically directs that a certain material or method be used to accomplish the desired end.



**BOX 6-7**  
**Gaps in Standard Reference Materials (SRMs)**

- Traceability of controls. At this point, most reference materials lack SOPs for routine typing.  
Preanalytical, assay, and interpretation: reaction mode
- Sensitivity, specificity, contamination, technical issues: framework for investigators.
- How to identify a good standard?
- Supporting materials, for example,
  - nucleic acid
    - extraction method,
    - quantification method, and
    - integrity, purity.
- NIST-traceable model.
- Criteria to qualify a reference sample: What is the reference?
- Standards for preparation of SRM:
  - process
  - appropriate analyte of interest
- Validation—also acquire samples other than reference: range (and the logic behind that)
- Can surrogates suffice?
- Are near neighbors required? Or should this be an individual researcher responsibility?

SOURCE: Budowle presentation, 2013.

indexing for NGS may require reliability testing. Some reads may have uninformative bar codes,<sup>2</sup> likely because the index is low quality, because there may be synthesis errors in the generation of the barcodes. Perhaps adequate differentiation between/among indices should be implemented so that an error does not place sequencing content in the wrong sample. Computationally it should be feasible to define barcodes, but depending on how they are used may limit the numbers of samples that can be indexed.

New reference materials must be generated, but old ones must be maintained, or there may be no good way to compare back with existing or former methods going forward. It will be necessary to select the best methods for generating databases. Concerns that should be addressed

<sup>2</sup> Barcodes or indexes are short unique sequence tags added to every fragment of a sample during library preparation to “tag” the fragments unique to a sample. Thereby, different samples can be pooled (or multiplexed) and data separated (i.e., demultiplexed) bioinformatically after sequencing, based on the unique tagged sequences.

appear in Box 6-8. For years, there has been discussion about the concept of (1) a centralized database or (2) a decentralized virtual centralized database. There are challenges to creating these databases. One is how to convince and incentivize people to share. Potential solutions might be to make funding dependent on sharing, or to provide database access in return for sharing. There will be patent and intellectual property issues. There should be, at a minimum, some centralized knowledge database of what is held by governments and the private sector.

Any centralized entity or group of entities that produces reference materials should have access to a sufficiently comprehensive collection, understand the rationale for isolates, and have sufficient long-term support. Such an entity would need to maintain the highest possible QA/QC, using accepted, standardized methods. It should be responsive to research and development (R&D) and support and incorporate R&D that updates extant knowledge of diversity. Finally, it should be governed as a community resource, while balancing this access against security concerns.

Budowle does not foresee development of an “ideal” reference resource because the characteristics of such a resource will continually change. Instead he foresees the creation of multiple reference resources, some of which may be set up in real time during events. Another limitation to achieving an ideal reference resource is the inability to capture the full diversity of the microbial world and all the permutations neces-

#### **BOX 6-8** **Databases and Resources**

- The Select Agent Registry system has created a database, but it has limitations—no uniformity, and decentralized holdings hamper the speed, accuracy, efficiency, and reliability.
- Databases could refer to physical materials and/or related data; microbes, toxins, nucleic acids; and metadata; genomic, proteomic, transcriptomic, and metabolomic data.
- A centralized, comprehensive physical archive of reference materials would facilitate
  - Implementation of a standardized characterization system,
  - Uniform QA/QC,
  - Development of standard typing techniques,
  - Standardization of new techniques and analytical methods,
  - Reference samples for high-resolution genomic comparisons, and
  - Duplication.

SOURCE: Budowle presentation, 2013.

sary for the multitude of analytical methods. For archival purposes, it is reasonable to maintain representatives of pathogens that may be used in biocrimes or bioterrorism. Centralized collections are likely to be inadequate for most investigations, but suitable for basic research. Culturing strains from an archive will be done on a limited basis owing to the cost of maintaining a comprehensive set that is continually cultured and tested over time to ensure viability.

Microbial forensics investigators need increased access to reference resources, yet there is a real conflict: One position may demand that access be restricted for security reasons while another position is that tests and countermeasures cannot be developed without access to the materials. To make any progress, we must consider and understand both positions.

Budowle's suggestions for immediately moving forward are to

- Identify the experts now. There should be a global consortium of experts who have agreed to be available for, at a minimum, consultation should an event occur.
- Review the current (and past) state-of-the-art technologies.
- Establish scientific working groups and guidelines.
- Establish standards and standardization.
- Better define—and encourage—validation and peer review of the science.
- Share information and capabilities within the law enforcement and intelligence communities.
- Foster partnerships.
- Develop ways for greater access to genomes or microorganisms to facilitate validation.

Budowle emphasized that we must begin an interactive and committed process to address issues *now*. Scrambling to prepare as an event occurs is not a desirable scenario.

Cindi Corbett of Canada's National Microbiology Laboratory agreed that the idea of microbial working groups is a good one. Large problems exist, but they can be attacked one step at a time by the typing experts, database experts, and so forth. She would like to see Round Robins because investigators do things differently. Given the same set of data, would we all come up with an acceptable answer? Initially such an exercise could use just data, instead of an actual sample. Many opportunities exist for collaboration, interacting with the community, and for employing experts in various areas.

## REFERENCE COLLECTIONS AND DATABASES

As noted in numerous places in this report, there is widespread agreement that more reference collections and databases that are properly curated and maintained are needed. Reference collections house actual organisms while databases comprise only genomic or other information about microorganisms.

The American Type Culture Collection (ATCC) is one of the premier sources for microbial reference strains. It contains the world's largest collection of bacteria, viruses, yeast, fungi, protozoa, nucleic acids, and molecular tools. The ATCC Bacteriology Collection contains more than 18,000 strains in over 750 genera as well as more than 3,600 type cultures of validly described species and nearly 500 bacteriophages. For viruses, ATCC houses a wide assortment of cultures for use in research related to pathogenesis, epidemiology, molecular assay development, and vaccine discovery and production. ATCC offers an extensive array of infectious disease organisms intended to promote research leading to novel methods of detecting, minimizing, and treating infectious diseases. It is an entity that could be studied as a model for a reference collection devoted to microbial forensics.

Dr. Juncai Ma, a member of the organizing committee for the Zagreb workshop, is the Director of WFCC-MIRCEN World Data Center for Microorganisms (WDCM) of the Institute of Microbiology at the Chinese Academy of Sciences and an executive of the World Federation of Culture Collections. He discussed the WDCM,<sup>3</sup> which is a user-friendly international database resource for the compilation of data on the location and function of culture collections of microorganisms, cultured cell lines, and genetic elements. The WDCM also provides access to the data, serving as an online gateway to international databases on microbial diversity, culture collection catalogs, services, and molecular data relating to microorganisms.

The WDCM was developed under the auspices of the World Federation of Culture Collections and UNESCO's Microbial Resources Centres (WFCC-MIRCEN). The WFCC is a Multidisciplinary Commission of the International Union of Biological Sciences (IUBS) and a federation within the International Union of Microbiological Societies (IUMS). The WDCM is now located at the Institute of Microbiology, Chinese Academy of Sciences (IMCAS) in Beijing. Brief descriptions of the WDCM databases and resources appear in Box 6-9.

The Information Center of IMCAS is collaborating with the interna-

---

<sup>3</sup> More information is available at <http://www.wdcm.org/>; accessed November 23, 2013.

**BOX 6-9**  
**Databases and Resources:**  
**The World Data Centre for Microorganisms (WDCM)**

1. *World Directory of Culture Collections (CCINFO)*.  
 Directory of all registered culture collections: 652 collections from 70 countries/regions.

Microorganisms: 2,304,840

Viruses: 36,807

Bacteria: 966,143

Cell lines: 31,178

Fungi: 689,946

Can be browsed by country, region, or acronym. Can be searched by collection, strains, or keyword. Drill-down information and statistics on culture collections—including key collection personnel, subject coverage (e.g., agriculture, biotechnology), preservation methods, culture availability criteria, services, and entry and update information.

2. *CCINFO Strains*.  
 List of holdings of registered culture collections.
3. *Reference Strain Catalogue*.  
 Provides access to the reference strains listed by the ISO TC 34 SC 9 Joint Working Group 5 and by the Working Party on Culture Media of the International Committee on Food Microbiology and Hygiene (ICFMH-WPCM) from its *Handbook of Culture Media for Food and Water Microbiology*.<sup>1</sup>
4. *Analyzer of Bio-resource Citations (ABC)*.  
 A platform to support researchers in checking literature citations. Data-mines 3,005 journals; years 1953-2012. Can search by paper, patent, and strain number. Offers drill-down statistics on most-referenced strains. Users can upload their own papers.

tional Barcode of Life (iBOL) project,<sup>4</sup> and with the China Central DNA Barcode of Life program. WDCM collaborates with multiple collections throughout the world to advance the field and to organize educational workshops and symposia. WDCM also cooperates with the International Organization for Standardization (ISO) to develop the WDCM Reference Strain Database for all microbial resources in conformance with ISO standards, and all of this information is available to the public.

Ma believes that the WDCM can make contributions to microbial forensics and perhaps serve as another model for information on culture

<sup>4</sup> More information is available at <http://ibol.org/>; accessed November 23, 2013.

5. *Global Catalogue of Microorganisms (GCM).*

A free information-service platform to help culture collections to manage, disseminate, and share information related to their holdings. Fifty-two different collections from 25 countries. By year end 2014, there will be 100 collections. Small collections will be offered support in creating their own linking homepages. Includes detailed strain information (e.g., patents, sequences, bioinformatics analysis), related citations, isolation sources, geographic origin, phylogenetic analysis, species identification, and access to online exchange of data. Drill-down information available for countries. Searches can be refined by collection, temperature, organism type, and isolation origin, and can be displayed in multiple formats. Android version available; iPad version soon to be released

6. *Statistics on Organism Patents.*

Via collaboration with World Intellectual Property Organization, information and statistics on patents associated with culture collections.

7. *Resources:*

Nomenclature—National Center for Biotechnology Information, Species 2000, List of Prokaryotic Names with Standing in Nomenclature.

Metagenome—Joint Genome Institute (JGI, U.S. Department of Energy). Metagenome portal, European Bioinformatics Institute Metagenomics, metagenome submission guide.

8. *The World Directory of Culture Collections.*

Book. Sixth version will be released in 2014; 191 collections updated their information in 2013.

<sup>1</sup> WDCM Reference Strain Catalogue, available at <http://refs.wdcm.org/home.htm>.

SOURCE: Ma presentation, 2013.

collections and databases. Issues that should be addressed by all culture collections before sharing would include

- Data standards: What do we need? What can we share? Minimum datasets (MDS) and recommended datasets (RDS). Range of microorganisms and software development.
- Data policy, software availability for data sharing, data club (should it be member only?). The data in WDCM's Global Catalogue are currently open to all.

Ma proposed development of a Global Catalogue for Microbial Forensics, as outlined in Box 6-10. He noted that a Global Microbial Forensics Catalogue and its knowledge base would make the discipline more vis-

**BOX 6-10**  
**Proposed Global Catalogue for Microbial Forensics**

**Proposed Content:**

- Definitions of microorganisms groups
- Identify strain providers (collection)
- Documents and material transfer agreements for provision of strains
- Prices and contact information
- Data could include strain data, characterization, genetic fingerprint, whole genomics, protein structure (could be assembled in part based on international literature)
- Papers
- Patents

**Proposed Functions:**

- Provide quick and accurate identification
- Analyze isolation and storage of strains
- Provide experts list for dealing with specific microorganisms
- Consolidate and offer as much information as possible for strains
- Make bioinformatics tools available
- Link experts through online communication tools

SOURCE: Ma presentation, 2013.

ible, and pointed out that it is much easier to begin cooperative efforts among stakeholders when working in the information field. Working groups could be organized to address information issues.

## 7

## Bioinformatics and Data

### THE ROLE OF BIOINFORMATICS IN MICROBIAL FORENSICS

In a 2009 report, the American Academy of Microbiology stated that the application of computer analysis to molecular biology, otherwise known as “bioinformatics,” is “a fundamental corollary to biodefense research” (AAM, 2009b). The report also pointed out a theme very much consistent with the presentations at the Zagreb workshop—that advancements in biodefense and bioinformatics research, although meant to lead to national security improvements, will also reverberate across almost all biological disciplines. Specific fields that will benefit include

- Understanding genetic diversity,
- Epidemiology,
- Vaccinology,
- Global health,
- Metabolic reconstruction,
- Systems biology, and
- Personalized medicine.

As most of the workshop speakers noted, WGS has made possible giant steps forward for epidemiology and microbial forensics. The ability to examine and compare the genetic sequences of bacteria, archaea, viruses, and microbial eukaryotes has revolutionized microbiology and provided researchers with insights “into the processes microbes carry out, their pathogenic traits, and new ways to use microorganisms in medi-



cine and manufacturing” (AAM, 2009b). Gene sequencing, WGS, and the techniques these make possible are moving microbiology forward more rapidly than ever before. They also have given rise to new tools—genomics, proteomics, metabolomics, transcriptomics, molecular phylogeny, and others—that now are available to scientists studying the relationships among microbes, their pathogenicity mechanisms, and their metabolic potential (AAM, 2009b). The AAM report went on, however, to note that a “major effort in functional annotation to understand the DNA sequence we already have” is needed (2009b:16). “Annotation is a critical part of making genome sequences into resources, but it represents a huge bottleneck.” It also called for the implementation of a set of standard annotation platforms and a central annotation resource “with defined methods and standards, and guidelines for people to use the resource” (2009b:16).

### THE INTERPLAY BETWEEN SEQUENCING AND BIOINFORMATICS

Dr. Jongsik Chun of Seoul National University in South Korea believes that WGS provides the ultimate information for both epidemiology and microbial forensics. As an illustration of the power of WGS, he pointed to its use in tracking the transmission of carbapenem-resistant *K. pneumoniae* among individuals during the NIH Clinical Center’s 2011 outbreak (Snitkin et al., 2012). WGS is superior to all other identification methods if cost, time, and bioinformatics are not issues. However, using genomic information is the area of microbial forensics in which the application of bioinformatics is most needed. Chun sought to provide an understanding of the scope of microbial knowledge that is lacking and the uses to which the data that are available could be applied with better tools. He reviewed the challenges faced in assembling and curating databases, as well as in developing bioinformatics to efficiently and effectively exploit the data.

Chun, who is a trained taxonomist, reminded the workshop participants that the traditional concept of species is “groups of actually or potentially interbreeding natural populations that are reproductively isolated from other such groups” (Mayr, 1942). But in the bacterial world, this definition falls short because there is no “sex” or “breeding,” per se. Instead, bacteria may exchange genetic information through lateral transmission both within species and among different species, allowing for high genetic diversity. For bacteria the fundamental concept of a species is actually based on how similarity between strains is measured. Bacte-

rial classification and identification should be based on the comparison between type strain<sup>1</sup> and isolate.

Chun believes that the indirect DNA-DNA hybridization method—the previous “gold standard” for determining the genetic distance between two species—is not a reliable methodology. The 16S rRNA sequencing is revolutionizing taxonomy and now WGS is available. A new index has been developed called Average Nucleotide Identity (ANI). In DNA-DNA hybridization, the DNA would be fragmented and hybridized in solution. The same general analysis now can be done digitally. The DNA of one strain is fragmented, the sequences are compared using BLASTN or MUMmer programs, the similarity of each fragment is established, and the findings are averaged. The method has been demonstrated to be very reliable (Goris et al., 2007).

NGS and ANI have provided for the first time an accurate, objective, and reproducible method for bacterial classification and identification. Sequencing may be expensive but capacity is not a problem. High-throughput NGS sequencers have a capacity equal to tens of thousands of conventional Sanger sequencers. However, Chun noted that big gaps exist between the ability to sequence and the cost of accessing the necessary hardware and bioinformatics. Genomics in microbial forensics presents a typical big data problem: high-volume, high-velocity, and high-variety, coupled with a need for quick answers.

A very conservative estimate puts the number of prokaryotic species at 1 million, of which only approximately 11,000 species have valid names. The routine use of 16S rRNA sequencing is accelerating the description of new species, and in 2012, 721 new species were described. However, Chun pointed out that there is little or no taxonomy funding in many countries, except perhaps China. Chinese microbiologists are now responsible for the description of over 30 percent of new species of bacteria and archaea.

There has been a huge jump in genome releases since the onset of NGS. Of the roughly 11,000 bacteria and archaea species with valid names, 99 percent have 16S rRNA gene sequences in public databases. There are genome sequences, however, for only 24 percent of these prokaryotic species and only 16 percent of prokaryotic species have been sequenced for type strain. Thus about 85 percent of known species have no genome sequences for their type strains. Applying genome sequencing to species-level identification is limited because reference genomes do not exist. Accurate identification based on genome data is possible with ANI, but more type strains should be sequenced.

---

<sup>1</sup> Usually the first strain of a new species of bacteria to be described is known as the “type strain.” A bacterial species is a collection of strains that share many common features (Black, 2002).

Chun pointed out that sampling for genomics is biased toward clinically important strains, such as *E. coli*, *Staphylococcus aureus*, *S. enterica*, and *Enterococcus faecalis*. There are big gaps in genome-sequenced taxa. If there is bacterial contamination in soil, for example, how much background can be identified? Only about 5 percent of detectable soil bacteria can be identified as species with valid names, while 95 percent cannot. This is because most soil bacteria are not cultivable. Perhaps single-cell genomics can solve this problem, but the currently available technology is not suitable for large-scale surveys.

Creating a database is easy, but maintaining a large and rapidly growing database is very difficult. If someone adds 2,000 *E. coli* strains to a database, it can take a year to curate. Many databases have been retired or destroyed, likely owing to the inability to cope with the quantity of new data and the lack of the funding to maintain existing databases. Databases Chun considers to be up to date and stably funded include the following:

- NIH's National Center for Biotechnology Information (NCBI) maintains several databases: GenBank, RefSeq, Microbial Genomes Resources.
- Integrated Microbial Genomes and Metagenomes (IMG) supports the annotation, analysis, and distribution of microbial genome and metagenome datasets sequenced at DOE's JGI.
- The European Bioinformatics Institute's (EBI) Ensembl Genomes.
- EzGenome.<sup>2</sup>

Chun explained that his lab created EzGenome because it wished to provide a prokaryotic genome database that was manually curated, taxonomically correct, and useful for microbial systematics as well as ecology and microbial forensics. In the GenBank database, for example, genomes are identified to the species level in some cases. EzGenome contains about 12,000 prokaryotic genomes. All genome projects have complete hierarchical taxonomic information (from phylum to species) according to the EzTaxon taxonomic system. Some genome projects are identified at the genus level owing to a lack of 16S rRNA sequence and adequate reference genome sequences to which to compare them. ANI-based dendrograms are produced for all genera and families.

Chun agreed that GenBank is a good primary database, but its entries may be incorrectly labeled and include contaminants, likely due to contributor's error, which is a common occurrence in databases. A survey by

---

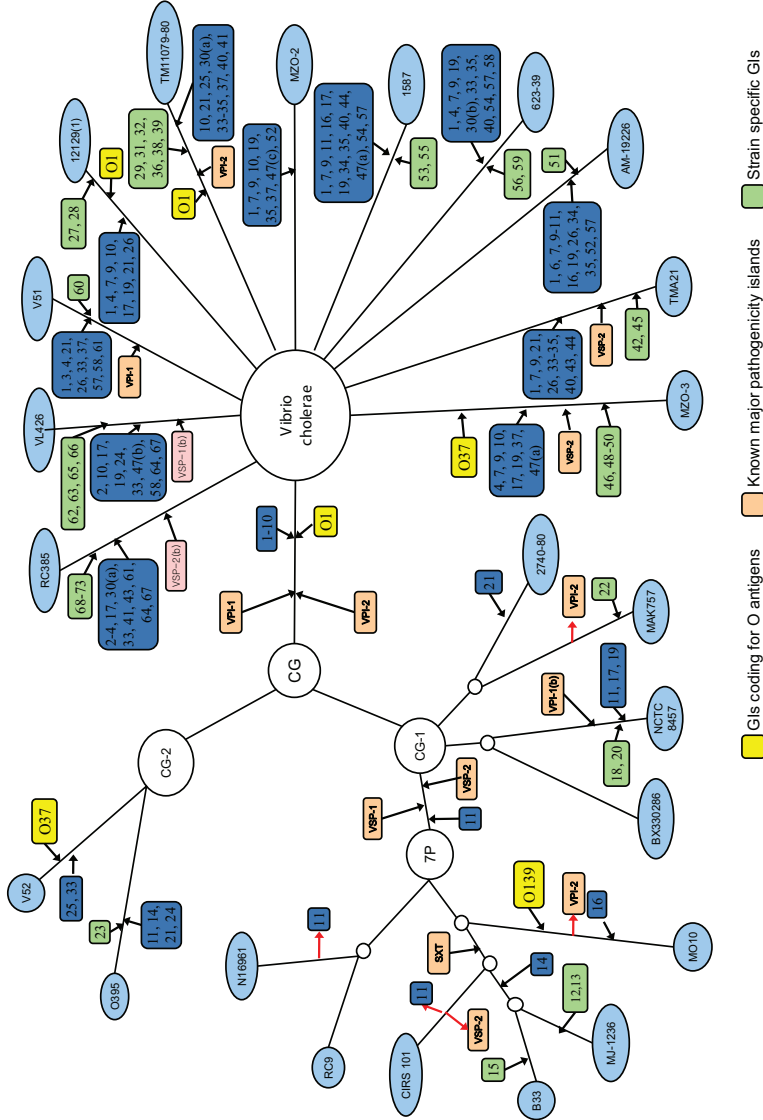
<sup>2</sup> Information is available at <http://ezgenome.ezbiocloud.net/>; accessed November 23, 2013.

Chun's lab of GenBank revealed many errors, and such errors can lead to problems—for example, when taxon-specific primers are designed.

Chun used work his laboratory has done with pandemic strains of *Vibrio cholerae* as an example of the power, and limitations, of current technology and bioinformatics. Cholera researchers continually encounter newly emerging and reemerging pathogenic clones with diverse combinations of phenotypic and genotypic properties, and this hinders control of the disease (Chun et al., 2009). The seven pandemic *V. cholerae* strains belong to O1 or O139 serotypes and have almost identical genome backbones, including only about 100 SNPs. However, they show highly variable gene content in some regions of the genome (Cho et al., 2010). Short-term evolution is caused by a combination of SNPs and lateral gene transfer, but it is possible to trace the gene transfer events. Chun's laboratory performed SNP analysis on a variety of *V. cholerae* strains, mapped these findings, and created the phylogenetic tree in Figure 7-1. It portrays the genomic islands in closely related strains. They discovered that sequence variations among orthologs were more common than SNP variations and that SNP variations appear to be rare among highly conserved genes. They also determined that the degree of genome variation appears to be species specific; for example, there is more variation within *V. cholerae* than within *Streptococcus agalactiae*. The tree appearing in Figure 7-1, however, was based on only 26 genomes and took 6 months to create. How long would it take to perform analysis on 1,000 *E. coli*? Better bioinformatics are needed.

Examples of two targets in microbial forensics are the pure culture sample or a single microorganism or virus, and the metagenomic sample, which can comprise a mixture of organisms (e.g., human, bacteria, viruses). Metagenomic sample analysis involves detecting target-specific information against a large background by using single-gene-based community analysis (e.g., 16S rRNA-based microbiome) or shotgun metagenomic DNA sequencing. Issues, methods, and bioinformatic analyses differ for pure cultures versus metagenomic samples, but Chun emphasized that properly curated genome databases are necessary for both.

Pure culture sample analysis involves the assembly of genomes, taxonomic identification, identification of variation, such as SNPs, gene content, tracing of gene transfer, and matching against a database. Metagenomic sample analysis involves assembly of a metagenome, taxonomic analysis of community composition, and matching against a database. Chun stressed that the technical difficulty of a metagenomic sample analysis is 1,000 times greater than that of a pure culture sample. Clinicians continually ask him, "Can you do quantification using NGS sequencing?" He believes the answer is no, and a method for quantification needs to



**FIGURE 7-1** Evolutionary phylogenetic tree of *V. cholerae* based on SNP analysis of 26 genomes. SOURCE: Chum et al. (2009).

be developed. The components of a bioinformatics program for microbial forensics should include

- Curated, regularly updated databases;
- Data mining tools for large-scale databases;
- Real-time analytical capability;
- Adequate reporting functions; and
- Effective data visualization—it is crucial to import user experience into this process because people (e.g., judges, decision makers, first responders) must understand the outcome of an analysis.

Challenges that Chun sees in microbial forensic bioinformatics include

- Integrating data generated by different methods, such as WGS, multilocus sequence typing (MLST), and pulsed-field gel electrophoresis. The answer lies in bioinformatics and manual curation.
- Integrated real-time sharing of NGS data. Labs A, B, C, and D should be able to share in real time. We can even upload raw data to the cloud servers while sequencing. This is doable (e.g., via BaseSpace genomics cloud computing),<sup>3</sup> but poses a policy problem.
- Storage capacity is not an issue, but processing secondary data for microbes is.

For efficient sharing, standardization must be developed for

- Metadata (data about data) and database format/schema,
- Sequencing instruments,
- Protocols for sequencing-library preparation,
- Minimum sequencing depths (for each species),
- Assembly software and/or pipeline,
- Annotation software and/or pipeline, and
- Other bioinformatics tools.

There are challenges to achieving analysis repeatability. A bioinformatics pipeline typically involves many software tools, parameters, and hidden “know-how.” For this reason it is difficult to reproduce most of the large-scale genomics papers (e.g., genome assembly, metagenome comparison). Two solutions he suggests are

---

<sup>3</sup> A cloud platform directly integrated into the industry’s leading sequencing platforms. Built by Illumina on Amazon’s AWS cloud infrastructure. For more information, see <https://basespace.illumina.com/home/sequence>; accessed November 24, 2013.

- Virtualization. A model could be the ENCODE Virtual Machine and Cloud Resource for human DNA.<sup>4</sup> Complex bioinformatics pipelines and parameters are packaged as an image and can be plugged into any computer.
- Cloud computing. If a lab does not have adequate information technology (IT) infrastructure, software can be applied using cloud computing. The cloud service comes with IT.

There is a need for bioinformatics and databases for the purposes of detecting antibiotic resistance and genetic engineering. A retrospective study revealed the potential for rapid WGS to reduce the time it takes to diagnose a patient with multidrug-resistant tuberculosis from weeks to days (Köser et al., 2013). In this study, the patient's first sputum sample became culture-positive after 3 days in the mycobacterial growth indicator tube (MGIT) culture system. The DNA was extracted directly from the MGIT tube and sequenced using the Illumina MiSeq platform. Antibiotic susceptibility was determined in less than 2 days. For this purpose, a database of SNPs responsible for antibiotic resistance would be enormously helpful. Similarly, the ability to genetically engineer microbes would be greatly empowered by assembling databases of sequence data to compare. This could be particularly useful for analyzing viruses. A great number of kits exist—for example, for PCR and microarrays—and if new virus sequences were to be added to the dataset, one could mine the database to update primers, probes, and other tools.

Chun believes that bioinformatics for metagenomics faces serious hurdles. The computational challenges are (1) the databases are ever-growing, (2) software and hardware are needed to improve searching/matching, and (3) improved data visualization capabilities are needed.

Cloud computing offers a way to share metadata, and to send out streams of data that can enable accurate identification, provide global epidemiology and treatment information, and inform genetic engineering. Overall, Chun believes that cloud computing technology and bioinformatics will find a way. But he sees real barriers in policy. Two key questions are

- Who owns the data? Only 23 *B. anthracis* sequences are in public databases (versus more than 1,000 *E. coli*). So it is likely that many *B. anthracis* genome sequences have not been shared.
- Who is responsible if something goes wrong? If there were to be a problem with a certain software that was used in a court case,

---

<sup>4</sup> For more information, see <http://www.nature.com/encode>; accessed November 24, 2013.



for example, would the developers be liable? This can be a problem, especially because most popular bioinformatics software is developed in academic institutions.

One possible solution to the information-sharing dilemma, since the majority of genome data for biothreats may not be publically available, is defining and establishing two data tracks, one that can be shared and one that cannot. Ideally, academia and governments can at least share standards for data, if not the data itself. In this way, the data can be easily integrated when possible.

In a discussion following Dr. Chun's presentation, Dr. Paul Keim noted that technology development today differs from that of a decade ago. Technologies then were not forward compatible, which limited their power. But WGS is forward compatible. If one investigator has a genome sequence using Life Technologies' PGM sequencer and another using an Illumina MiSeq sequencer, the data can be easily compared, put in a database, and will be forward compatible with any technology that comes along. He agrees that a committee will never decide anything, but the genome is the genome and once we have it, we can use it.

Dr. Munirul Alam of the International Center for Diarrhoeal Diseases Research in Bangladesh suggested that individual committees be formed to guide what typing tools, for example, should be used for each pathogen. Scientists are sporadically doing everything, but not in a collaborative way leading to consensus. The findings could be used in microbial forensics. Dr. Dag Harmsen agreed it would be desirable to have an international committee to agree on typing methods, but thinks that such a thing is unrealistic—committees cannot even agree on names. In addition, needs differ between developing and developed countries, so there will never be one assay method that fits all needs. It is ultimately the users who decide what is going to be used and will declare a *de facto* standard. Also, technology continues to evolve, and novel and even more powerful analytical tools likely will arise in the near future.

## MANAGING LARGE DATASETS

Dr. Aaron Darling reviewed changes that have already occurred and are still under way in the development of hardware, software, and systems for collaborative computing. Enormous quantities of data are being generated, and there are challenges to handling, storing, and analyzing them in reproducible ways that can be validated. He discussed the pros and cons of cloud computing and "software as a service" in a scientific arena. He touched on issues regarding the funding of and responsibility



for microbial forensics dataset management. He offered possible solutions for some of the obstacles he sees.

Looking backward, he noted that when there were few genomes to analyze, researchers performed analyses and then disagreed about the meaning of the analyses because there is a great deal of observational bias when data are limited. Now that a great deal of data exists, however, the taxonomic biases within databases should be considered, in particular, the lack of commensals.

In addition to the uncertainties regarding which organisms to sequence, organisms are naturally and constantly changing. The sequences added to databases represent a moving target because they come from living systems that evolve. Is the population evolving, or is one observing “standing” variation?<sup>5</sup> Moreover, the databases used for microbial forensics are likely to be based on public databases. The standards for contributions to public databases are unlikely to meet the rigor one would expect for admissibility in court. Can these data be trusted?

Ten years ago, a major challenge was how to sequence a genome. Today a major challenge is how to deal with the immense amount of sequencing data being generated. This issue is becoming a major problem, both in isolate genomics and metagenomics. Two years ago, the sequence read archive at NCBI reached 100 terabases, 11 percent of which was metagenomic (Kodama et al., 2012). The MG-RAST<sup>6</sup> resource offers 38 terabases of metagenomic data for download.

Sending large amounts of data over the Internet is not a problem because the aggregate bandwidth of the network is enormous. Moving data is possible using available, well-engineered, peer-to-peer (e.g., BitTorrent and Coral Content Distribution Network) and multicast solutions. However, although the aggregate bandwidth of the network is substantial, the endpoint bandwidth (e.g., at the source and the destination) may be limited, which could be of concern for truly large data volumes. The nature of how data arrives and how analysis is performed is also changing. The trend is away from data arriving in batches and toward delivery of continuous data streams. Three examples of initiatives to release data streams as they are being generated by public health labs and other sources are (1) the Global Microbial Identifier (a multiorganizational international effort), (2) the 100K Foodborne Pathogen Genome Project (University of California, Davis, School of Veterinary Medicine), and (3) the Genome Trakr Network (FDA). The data-stream approach is changing the nature of how analyses are performed. The previous paradigm was that data were collected from various sources, computational analysis was

---

<sup>5</sup> Standing variation is genetic variation for fitness traits in natural populations.

<sup>6</sup> For more information, see <http://metagenomics.anl.gov/>; accessed November 24, 2013.

performed, and the results were published. This approach does not adapt well to streaming data sources and may not be viable going forward. Unfortunately, most of the existing algorithms were engineered with the assumption that all data would be available from the start. Bioinformaticians need to start developing algorithms that scale to arbitrarily large datasets.

Algorithms in bioinformatics already exist that fit those two criteria. They include multiple sequence alignment with profile-hidden Markov models, phylogenetic placement on reference trees, and bloom filters. The nature of the data has changed, and more methods like these must be developed. Once it is possible to analyze all the data, researchers will have some basis from which to make a decision about whether data are useful or can be deleted from the data that are archived for future analysis.

The computing infrastructure used to analyze data also has changed greatly. The approach is shifting from individuals doing analyses on PCs or within a computing cluster to people performing analyses within cloud systems. The Amazon Web Services Elastic Computing Cloud is the best-known of these systems, and there are others, such as OpenStack, a cloud infrastructure that virtually everyone can download and set up on his or her own computer hardware. In Australia, the government set up a national research cloud named NeCTAR. Those at Australian research institutions are issued free allocations for use; in instances when large computing resources and many computing hours are needed, users can make special requests.

Other paradigms for using cloud computing have emerged, including the HTCondor project, which essentially leverages idle computing time; it is a cycle-scavenging system. The Condor can launch virtual machine images on idle computers. If an institution has a large number of computers that sit idle after everyone goes home at 5:00 p.m., that overnight time becomes a large computing resource.

The cloud-computing model is attractive because it enables one to rent time on a large, commercial, professionally managed computing facility, and to pay only for what is used. It enables a researcher to quickly scale computing resources up or down. One does not have to worry about how to dispose of all of the outdated computers in 5 years because the cloud provider will manage that. A challenge, however, is that data motion and storage in the cloud can become very expensive and will increase the funding researchers need from their sponsors. Researchers who wish to perform large dataset analyses in clouds must factor in how to most efficiently move the data around and store it. One possible solution is to use third-party providers for data storage.

Hardware architectures have changed also, largely owing to the fact that central processing unit (CPU) clock speed reached its limits about

10 years ago. A CPU cannot be pushed much further than 2.5 gigahertz without overheating. Speedups are largely attributable to doing more work in the same clock cycles. In particular, the graphics processing paradigm has emerged; these systems are constructed to enable thousands of arithmetic operations in parallel during each clock cycle. Today there are a number of graphics processor-enabled bioinformatics applications. For sequence alignment, there is the Short Oligonucleotide Analysis Package (SOAP); for sequence indexing and generic sequence analysis algorithms, the SeqAn library; and for phylogenetic analysis, graphics processing unit-enabled versions of MrBayes and Bayesian evolutionary analysis by sampling trees, which use the BEAGLE Library (beagle-lib).

Another innovation in computing hardware that is very relevant for large-data processing is solid-state disk (SSD) architectures. SSDs are similar to “thumb drives” but are scaled-up higher performance versions. They use essentially the same technology except that these are used as a normal disk drive in a computer; their random read/write performance is much higher than the classic hard-platter disk drive. When coupled with the classic disk in a unified storage system, these provide both high capacity and the ability to do a very fast random-access and reading and writing of data, which is a very useful feature when performing large genome analyses.

Computing costs have begun to exceed sequencing costs (Loman et al., 2012; Sboner et al., 2011). The major costs are the development of novel algorithms and software and training and maintaining skilled personnel to perform the analyses. When computation results in a public good, one must determine who will pay for it. Darling believes that costs can be divided into three steps:

1. Development of novel algorithms and software (likely by academia);
2. Integration of new software with existing software for usable systems (engineering); and
3. Maintenance, support, and operation of computing.

Creating software is by nature very different from developing a physical technology. Once software is developed, it can be easily copied and disseminated. The cost for distributing it widely is negligible. The prevailing model in bioinformatics, and perhaps scientific software in general, is that through its funding institutions and scientific organizations, the public pays for step 1 and perhaps a portion of step 2. In Darling’s opinion, industry generally supports steps 2 and 3. This model appears to be reasonably workable because if a software system breaks, someone should be accountable, and private industry appears to be in a better posi-

tion to handle this responsibility. A complication, however, is intellectual property considerations. Much of the intellectual property is generated during the first step, and recognition of this must be conveyed to the parties who operate and maintain the software. One of the most common approaches is to make the software open source. The open sourcing of software is actually advantageous for industry because it can avoid investing in the development step. It is attractive for the bioinformatics industry to pick up these open-source tools and operationalize and validate them for users.

In microbial forensics analysis, absolute reproducibility may be an unachievable goal. So a challenge becomes, how reproducible should it be? In theory, computers should make repeatability and reproducibility very easy. Many scientists who practice data analysis, however, lack training in fundamental aspects of computing and software engineering that would enable them to undertake reproducible computing. A recent trend toward educating this community can be seen in the series of workshops called Software Carpentry.<sup>7</sup> These workshops offer training in basic computing skills, such as version control, as well as literate programming, so scientists can generate workflows that are reproducible on different systems and engineered according to standards generally accepted in software engineering. In addition, there are a number of efforts to generate point-and-click visual systems that will enable users to generate reproducible workflows. These include the Galaxy, Taverna, Knime, and Kepler systems. Users can access a large number of small bioinformatics components, which they can connect and reconnect for arbitrarily different and unique workflows, and which they can then share with others. When users perform data analysis, the systems will track and record every step applied to data, and the users can share analysis metadata.

In terms of the operational aspects of repeatability and reproducibility, Darling believes that the ability to examine the software unit should be taken into account. With closed-source software, the best one can achieve is repeatability and copying the closed-source software to another computer; one cannot know how or why it works, nor can the computation being performed by it be independently reproduced. With open-source software, it is possible to examine the nuts and bolts of why it does what it does. Often when software is developed, and a manuscript is published about it, there are inconsistencies, and frequently large discrepancies, between what the software is actually doing and what is described in the manuscript. This can be attributed to the fact that when software is engineered and implemented, the developer must make approximations

---

<sup>7</sup> More information is available at <http://software-carpentry.org/>; accessed November 24, 2013.

to the model described in the manuscript for basic engineering issues because computers have finite resources. The trend toward cloud computing has largely been advantageous, but an important disadvantage exists. In the old model of computing, software developers sold copies of their software and users installed it on their computers. The problem was that the software was difficult to maintain, and it was very difficult for software developers to distribute bug fixes. The new model is that the user does not even receive a copy of the software. Instead, the software is purchased as a Web-based service. The software runs on a server that is managed by the entity that developed it, and the user pays for access. It simplifies software deployment and maintenance; updates are managed centrally from the software developer servers; and piracy issues are eliminated. The disadvantage, however, especially when systems must be validated, is that the user has no control over software versioning. Therefore, reproducible research is not possible. The ever-present fear is that when the service provider decides to fix a bug in its software system, it could have unintended consequences for the analysis a researcher has already completed. Similarly, doing the same analysis on different days may yield very different answers.

Darling thinks that there is probably a way to generate reproducible results and still obtain careful version control of analyses using cloud systems, but it should be something that the service provider builds into the system at the very basic level. This needs to be thought about and discussed with the providers who develop the necessary software.

Chun suggested that there currently are too many software choices, and what is needed is a trusted body to evaluate this software. He added that bioinformatics is engineering and he believes all bioinformatics issues are solvable. Darling agreed, noting that software reviewers don't always run the software or confirm that the source code is correct. A peer-review avenue for software systems is a good idea. The AAM (2009b) report also noted that there is a lack of statistical rigor for bioinformatics algorithms that needs to be addressed.

Some typing methods offer more sensitivity with less variability, and others offer more variability with less sensitivity. What can be relied on in the meantime? Darling said that he cannot comment on what should be relied on, but that genome-based phylogenies and their applicability are based on the type of organism one examines. Different lineages evolve in very different ways. Some recombine heavily, and if one builds phylogenies on these, a different result will emerge than if one builds a genome phylogeny on *B. anthracis*. For *E. coli*, he estimates it would be necessary to examine at least 200–400 kilobases of the genome in a concatenate alignment before getting a consistent answer about the phylogeny (Didelot et al., 2012).

## 8

## Findings and Conclusions: Initial Prioritized Science Needs for Microbial Forensics

The speakers and participants at the Zagreb workshop represent a collection of expertise from many disciplines, such as microbial forensics, epidemiology, public health, clinical medicine, genetics and genomics, phylogenetic analysis, and bioinformatics. Despite the great diversity of backgrounds, there were clear themes that emerged from the presentations and discussions. Most of these themes are consistent with the current literature in the fields represented and with the views of the committee members. In this chapter, the committee synthesizes its views on the needs that emerged and the questions that require answers.

### BASIC SCIENCE

As discussed in Chapter 2, although the world of living things is dominated by microorganisms, very little is known about the vast majority of microbes. Until recently, much of what is known was based on the very few microorganisms that are culturable and that were studied in laboratories using growth characteristics, Gram stains, serology, and other traditional techniques. Today, however, the ability to sequence the genomes of microbes has provided a great deal of knowledge about nutrient cycling, gene regulation, and reproduction for some bacteria and viruses. These processes, however, remain unknown for most microorganisms and “knowledge of the evolution and ecology of microbial communities lags far behind cellular microbiology” (NRC, 2007). In addition, until recently there have been few systematic efforts to collect and describe the

microbes living in soil, seawater, freshwater lakes and streams, on plants, and even commensally in the guts or other surfaces of humans and other animals. The new techniques of metagenomics circumvent the inability to culture the majority of microbes by directly assaying microbial genes, sequences, and genomes from entire communities, such as those available from environmental samples and the human microbiome. The use of metagenomic techniques, however, is still an area of active development. Meanwhile, knowledge of the natural microbial communities residing throughout the world is still highly incomplete. Although the biology of the small fraction of bacteria and viruses that are pathogenic to human beings, livestock and companion animals, and crop or forestry plants is somewhat better known and understood, there is still much to be learned about their phylogeny and evolution, how many strains of pathogenic species exist in nature, what their distribution is throughout the world, and how this distribution interacts with ecological conditions.

**Finding:** For microbial forensics purposes, the dearth of information about microbial diversity, ecology, population genetics, evolution and phylogeny, and worldwide distribution represents a major scientific knowledge gap. Determining an organism's source in the event of a biothreat will be greatly hampered by the absence of baseline information on the natural abundance and distribution of pathogens. If the normal distribution of a pathogen in background settings is unknown, it would be difficult, for example, to discern whether there is an unusually high abundance of that pathogen in an area experiencing an infectious disease outbreak. Background knowledge is also important in determining whether the presence of that pathogen is natural or the result of a deliberate or inadvertent release. Understanding the endemic microbial background is necessary to provide proper context for microbial forensics analyses, interpretations, communication, and resulting decision making. Currently there is insufficient understanding of microbial diversity and endemism to facilitate determining whether a pathogen has been intentionally released, where and how a biothreat agent was developed, and whether the presence of naturally occurring or endemic organisms could be exploited.

**Conclusion 1:** Although efforts to characterize the diversity of bacteria and viruses existing in nature are under way, there is no comprehensive effort ongoing to describe microorganisms. **An international collaboration engaging the worldwide scientific community in a systematic effort to identify, monitor, and characterize a far higher proportion of global microbial species to increase knowledge about endemism and background is needed. The effort should begin with known pathogens and then expand to their close relatives as well as emerging pathogens.**



To mount such an effort, formal international scientific collaborations will need to be created to ensure that technological resources are accessible to all nations, including developing countries that currently lack such resources, and that funding can be leveraged better. Other questions will need to be addressed: for example, what kinds of sampling and forensic characterization programs would be required to provide the most useful information? In addition, should there be an international committee to establish standards for typing and characterization of strains?

**Finding:** Skowronski and Lipkin (2011:173) state that

Perhaps the most significant challenge in microbial forensics is the ability to differentiate the natural from the intentional biological event. Mother Nature continues her relentless assault on human, animal, and plant populations with an impressive array and diversity of microbes, and situational awareness of unnatural malicious intent can be hard to come by.

Although there are criteria (see Box 3-1) for considering whether a disease outbreak is unusual, determining whether such an event is due to natural, accidental, or deliberate causes, and collecting the information to work through these criteria at the time of an event is likely to be time-consuming with slow response time. As noted in the *Biological Response and Recovery Science and Technology Roadmap* recently issued by the National Science and Technology Council (NSTC, 2013), a biological event demands “a quick and effective response in order to minimize loss of life and other adverse consequences and, in the case of suspected criminal activity or terrorism, to thwart ongoing activity and prevent follow-on attacks.” Timelines for making this crucial determination are currently unacceptable from both public health and law enforcement standpoints, which require development and management of appropriate and rapid responses, recovery measures, and resolution.

**Conclusion 2: There is a strong need for the development of high-confidence methodologies to distinguish among natural, accidental, and deliberate outbreaks of infectious disease.** Although the research and development agendas of the U.S. federal departments and agencies are now being coordinated under the NSTC *Roadmap* mentioned above (NSTC, 2013), such coordination on an international basis is justified by the potential for rapid spread of disease through global transportation networks. This is a high-priority need for the research and funding agendas both inside and outside the United States that requires a coordinated effort on an international scale.

Metagenomic techniques offer the possibility of sequencing the genomes of unculturable microbes, which represent the vast majority of



microorganisms, particularly from environmental samples (e.g., soil and water). In epidemiology and public health, the need to move from culture-dependent to culture-independent methods to identify pathogens is also being recognized as an important diagnostic capability. Metagenomics would avoid culture bias, but it is difficult to study rare taxa, which comprise many of the organisms in such samples. The probative value of the “small signal” of the rare microbe of interest in the “big noise” of a cluttered metagenomic sample is undetermined, and validation of the processes to detect rare organisms is necessary but problematic, particularly because of artifacts, stochastic effects, abundance effects, and contamination. Questions that arise include: What is required to demonstrate that this is a viable approach and what are the limits to what could be stated from such analyses? Can these problems be overcome to meet forensic standards? How should the scientific and legal significance be determined and supported when the agent of interest is a minority constituent in a “probative sample”? How much of the threat agent of interest must be in a sample to be significant or meaningful in context?

**Finding:** Metagenomics has great promise for microbial forensics if it can be adapted to improve and speed up the ability to characterize microbial species and communities derived from environmental samples. But it is as yet unclear whether the “clutter” in metagenomics samples can be exploited for forensic value.

**Finding:** The other “omics” (e.g., proteomics, metabolomics, transcriptomics, and glycomics) may also provide applications of value to microbial forensics. Wahl et al. (2011) state that “Molecular variations in DNA sequence are widely used for organism and strain identification, but many other molecules and chemical species may be useful in determining microbial identity and origins; these include proteins, peptides, lipids, carbohydrates, inorganic ions, and organic metabolites.” For example, proteomics may be of forensic value because protein expression profiles can potentially provide information that can be traced back to the environment in which the organism was cultured. Given that microorganisms respond to environmental conditions by changing patterns of gene expression, the types of proteins expressed by an organism can reveal information for forensics about aspects of growth conditions. The chemical species produced by microbes must, of course, be correctly detected and identified and the results appropriately interpreted using computational algorithms and rigorous statistics (Wahl et al., 2011).

**Conclusion 3: Priority research is needed to realize the promise of metagenomics and its application to microbial forensics and the devel-**

**opment of the forensic value of the other “omics”: proteomics, metabolomics, transcriptomics, glycomics, immunogenomics, etc.**

### NEEDS COMMON TO MEDICINE, PUBLIC HEALTH, AND MICROBIAL FORENSICS

Human and veterinary medicine, public health, agriculture, and microbial forensics clearly share many techniques and approaches and all are similarly motivated by the need to protect the health of humans, animals, and crops. With any infectious disease outbreak, the starting point is determining what the agent is, whether for epidemiological or microbial forensic purposes. The use of molecular techniques to identify pathogen species and strains has, over the last 10 years, become fundamental to microbial forensics and is rapidly being adopted in clinical medicine. Although hospitals do not yet routinely use sequencing strategies for diagnostics, recent examples (e.g., Snitkin et al., 2012) have demonstrated that hospitals can deal more quickly and effectively with nosocomial infections if they have the ability to integrate genomic and epidemiological data. The importance of being able to access rapid molecular diagnostic capabilities in developing countries was described in the report *Biosecurity Challenges of the Global Expansion of High-Containment Biological Laboratories* (NRC, 2012).

**Finding:** Access to rapid and accurate molecular diagnostic equipment and techniques is important for controlling and responding to disease outbreaks as well as for microbial forensics.

**Conclusion 4: Improved worldwide access to molecular diagnostics (polymerase chain reaction [PCR], whole-genome sequencing [WGS], etc.), including refinement and distribution of bench-top next-generation sequencing (NGS) instruments that are fast and affordable and have simple workflow procedures, is a critical need.** At a minimum, PCR technology should be universally available. WGS also could be more broadly available with the advent of small benchtop instruments that offer speed and accuracy of analysis and have become less costly. These tools should be in place as soon as is feasible, to allow more rapid detection, higher sample throughput, and swift response for both public health and microbial forensics.

**Finding:** In addition to increasing basic knowledge of the diversity, ecology, phylogeny, and other aspects of basic microbes, there is a need to improve our knowledge of mechanisms of pathogenicity, virulence factors, antibiotic resistance traits, mutation rates, prevalence and impacts of

lateral gene transfer (AAM, 2009b), and host immune responses to pathogens. Some participants at the Zagreb meeting expressed the need for the development of databases on clinical infectious disease agents, which is not currently common practice in medicine and would be a daunting task. Skowronski and Lipkin (2011) pointed out that between 1998 and 2006 there were 40 million hospital admissions for infectious diseases in the United States alone. They stated that

it is literally impossible for the etiological agent to be identified, much less to a standard that would be considered a validated fact generated by a tertiary testing facility (CDC/Laboratory Response network/State Public Health Laboratories) and that would withstand legal scrutiny. . . . In many cases, presumptive diagnoses and minimal screening tests are often used when clinical signs are sufficient to treat. In practice, this has been sufficient for medical care but is not adequate to address the needs of microbial forensics. (Skowronski and Lipkin, 2011:174)

Nonetheless, insights into pathogenesis and the development of drugs and vaccines require live organisms and clinical samples from individuals who have been exposed to live organisms. Skowronski and Lipkin also noted that sequencing, microarrays, high-throughput NGS, and bioinformatics represent new sensitive and exacting techniques for identifying and characterizing pathogens that will eventually have broad applicability to public health, biosurveillance, and forensics and attribution. It also is worth noting that in November 2013, the U.S. Food and Drug Administration (FDA) approved Illumina's MiSeq Dx DNA sequencing system for diagnostics, which until then had only been used for research (Dubay, 2013).

**Conclusion 5: An increased emphasis on research to determine mechanisms of pathogenicity, including virulence factors and host immune responses, is needed.**

**Finding:** A number of the participants in the Zagreb workshop called for better global disease surveillance. Although there are a number of programs already in existence for this purpose (see discussion of biosurveillance in Chapter 2), there is considerable evidence that these programs have yet to fully bear fruit. For example, as noted previously, only 16 percent of countries were able to meet the 2012 deadline for implementing the International Health Regulations, and the U.S. government has recently announced a new Global Health Security Agenda to intensify worldwide efforts to achieve global health security.

**Conclusion 6: Greatly improved global disease monitoring and surveillance in humans, animals, and plants is needed to facilitate rapid response and better disease control.** It also would increase our understanding of background levels for aiding in the detection of unusual outbreaks. Although a daunting task, clinical databases of infectious disease cases in hospitals need to be developed to improve our understanding of pathogen distribution and characterization. An early-warning geographic information system could also be valuable for aiding these purposes.

## METHODS AND TECHNOLOGIES USED IN MICROBIAL FORENSICS

Methods of great value in microbial forensics for identification of microbial agents and analysis of their source include both molecular genetic and nongenetic technologies. PCR, sequencing of 16S ribosomal RNA, multilocus sequence typing, and WGS are molecular genetic technologies used to characterize microbial agents. Other physical science techniques, such as mass spectrometry and electron beam-based methods, can be used to analyze the physical properties of microbial forensic evidence, for example, the presence of additives for stabilization and/or dispersability and physical signatures from the locale where the material was produced (Michael et al., 2011). Mass spectrometry is also useful for analyzing biological toxins, such as ricin and botulinum toxin, which may not contain co-purifying nucleic acid signatures (Johnson et al., 2011). In addition, proteins and other biochemical products may provide clues about an organism's place or conditions of origin.

### Nucleic Acid-Based Technologies

There are a number of approaches for interrogating nucleic acids for attribution. As described in Chapter 4, real-time PCR (qPCR) is a rapid and sensitive detection technology that has proven effective in both clinical and microbial forensic assay applications. The power of qPCR lies in simplicity, assay selection, multiplex capability, optimization, and already-validated assays. There are many available assays that are validated to certain standards, although they must be validated internally in each laboratory that uses the assay. Many methods enable qualitative detection of hundreds of pathogens and hospital-acquired infections to the species and sometimes the strain level and a number of assays are FDA approved. There also are assays in the food safety arena on which microbial forensics can draw. Matrices in the food industry are complex, and food is a potential target of intentional and accidental contamination.

There are many food industry microbial and epidemiological investigations, and microbial forensics can learn from these events and have access to the assays.

Mass spectrometry exploits the different masses of atoms or molecules in a sample material. MassTag PCR combines PCR with mass spectrometry to provide high resolution, sensitivity, and specificity. This technology already has revolutionized the ability to identify respiratory pathogens and hemorrhagic fever viruses.

Microarray platforms accomplish a similar goal of organism identification via detecting nucleic acid or protein signatures in a higher-multiplexed format. Many panels have 1 million hybridization or binding assays on a single array chip, and some have up to 5 million assays. Microarray panels are available for microbial detection, single-nucleotide polymorphism (SNP) detection, genome-wide association studies, gene expression analyses, and protein presence and abundance. Microarray readers for other purposes, such as oncology or infectious disease screening, could be leveraged by the development of custom chips that are directed to addressing microbial forensic issues. Many assays are available, but will need to be customized on a single chip to be more comprehensive for microbial forensic needs. Typically an array is dedicated to a specific organism or small group of organisms, although arrays suited to larger numbers of organisms do exist. Microarrays do not provide as much information as sequencing, but they approach the discriminatory power afforded by sequencing. These platforms may be a rapid and cost-effective alternative or a supplement to sequencing.

Until recently, it was laborious and expensive to sequence an entire microbial genome, and researchers and funding agencies were obliged to choose carefully, organism by organism, what would be sequenced (AAM, 2009b). The Sanger chain termination method dominated for about 30 years before a new generation of technologies significantly changed how sequencing is performed. Today, NGS technologies are allowing rapid sequencing of thousands of microbial isolates as well as complex microbial communities at far less cost and effort (Fricke et al., 2011). A review of various early high-throughput machines coming onto the scene can be found in Parla et al. (2011). There have been numerous comparisons of current NGS devices, among them Quail et al. (2012), Loman et al. (2012), and Jünemann et al. (2013), that evaluate the advantages and disadvantages of the many devices already on the market or expected soon. NGS technology will continue to improve rapidly. Any reference in this report to a related methodology or specific technology for a case study is used for illustration and not to represent the state of the art.

Nevertheless, as a consequence of these recent advances, sequencing and genomics are revolutionizing our understanding of bacterial and viral

evolution and function. It is now possible to sequence the genomes of an unprecedented variety of microorganisms. However, there has not been a strong open effort to prioritize which organisms to sequence for microbial forensics purposes. In 2009, the White House's NSTC (2009) recommended that a survey of all full genome sequences available at the time be conducted to assess whether the number and diversity of sequences are adequate for forensic purposes. On the basis of results of the survey, additional strains for WGS were to be recommended. However, this effort appears to be restricted to the United States and it is unclear what the outcome of the effort has been. The NSTC strategy document in which this recommendation appeared did not specifically address whether the effort was to be international in scope and the role of other countries in the task, if any, is not clear.

In addition, the enormous increase in sequencing has brought with it other needs that must be addressed. For example, the competing technology platforms for sequencing lack standardization, and there are no standards for how sequencing data should be reported. One consequence of this is that it is unclear whether differences in sequence data generated using different devices are a consequence of error or have some other significance. Sequencing error may inflate the similarity or dissimilarity between reference and evidence samples, increasing uncertainty. Defining and quantifying error rates associated with each sequencing platform are critically important, but it is unclear how to go about this task. "Deep" sequencing routinely produces vast (terabyte) quantities of data that demand novel methods for analysis, handling, storage, etc. These data require expert annotation and curation, but there are no standardized annotation software systems. (This topic is covered in more detail under Bioinformatics and Data.)

**Finding:** The development, validation, and standardization of rapid new analytical methods, including sequencing and non-nucleic acid assay technologies, will continue to be important. It is not likely that prolonged research, development, testing, evaluation, validation, and technology transfer programs such as those developed for the anthrax letter attacks case, which took 8 years, will be acceptable each time a biothreat event occurs, particularly if a novel agent is used or if exigent circumstances arise. To be as prepared as possible for the next outbreak event, this work needs to be carried out now, whether it is in response to a natural, accidental, or deliberate event involving a biological agent.

**Conclusion 7: Development of more advanced, faster, and cheaper assay and sequencing technologies that can be standardized and made more accessible to benefit both microbial forensics and public health is nec-**

**essary.** Although industry is likely to pursue technology improvements on its own, increased accessibility of small and affordable benchtop NGS machines, such as those used in the German *Escherichia coli* O104:H4 outbreak, would “democratize” use and enable many countries that cannot afford sequencing to be able to implement such capabilities for routine diagnostic purposes. Further development of NGS would also enable detection and identification of organisms without a specific a priori assay design. Standards for reporting sequence data also need to be developed. Indeed, small handheld devices to digitize DNA results and send sequences over smart phones for real-time detection in the field might be a reasonable technology hope in the future. Of course, if handheld devices come to fruition, they will have to be applicable to, and meet the requirements of, microbial forensic analyses.

### Applications of Physical Science

**Finding:** As noted, a variety of nonsequencing techniques are available for identifying organisms and their biochemical properties. Although nucleic acid-based technologies are very sensitive, they will not identify all threats that may be encountered, such as toxins. Toxins can be purified to a degree that it is impossible to identify even trace levels of DNA or treatment may degrade DNA in the sample. Platforms that perform protein and antibody/antigen detection are being exploited to address this problem. One system is electrochemiluminescence (ECL), which enables antibody-based assays, as well as nucleic acid-based assays. Knowledge of the target being pursued is necessary, as are thoughtfully designed and validated assays, to use ECL effectively.

The textbook by Budowle et al. (2011) has several chapters devoted to non-nucleic acid chemical and physical analysis technologies, such as mass spectrometry and electron beam-based methods. These technologies also appear to be useful, particularly for helping to determine the manufacturing method for a bioterror agent by analyzing the materials associated with the sample, such as agar and silicon in scenarios where residual material, a dispersal device, or other materials are discovered.

**Conclusion 8: High priority needs to be placed on continued research and development to improve physical science applications to microbial forensics.**

### SAMPLING

**Finding:** Chapter 5 contains a great deal of detailed information on sampling and preservation of microbial forensic evidence. Ideally, sampling



should be conducted according to a plan or design appropriate for the context and must include provisions for clear quality assurance and quality control (QA/QC) measures, documentation, handling, storage, packaging, and transportation. It was stressed that results are more reliable if specialists in microbial forensics can advise those who collect samples (e.g., first responders) before they begin work on the crime scene. For microbial forensics, guidelines are needed that will support the criminal justice system and intelligence community and comply with standards for crime scene examiners. Type of release influences the type of sampling that is appropriate. A covert release may result in a delayed response, whereas an overt release allows a more rapid, and possibly more focused, response. It also needs to be recognized that evidence collection for microbial forensics is sometimes incompatible with other forensic techniques. This concern must be taken into account, and triage strategies are needed for sample collection so that critical evidence is not destroyed. Investigative protocols that optimize the choice and priority of methods are needed and should consider sample availability, preservation and conservation of evidence, trade-offs between speed and accuracy and/or precision, as well as other factors. First and foremost, consideration of the health and safety of the people collecting and handling samples is required. Collection protocols also should be based on analytical capability, preservation and integrity of evidence, low cost and high efficiency, documentation, training, quality and transparency, and flexibility, because matrices in which evidence might be collected could be highly diverse and may be difficult to handle or preserve.

**Conclusion 9: The development (and validation) of processes (sample collection, preservation, handling, storage, packaging, and transportation) and analytical methods for microbial forensics, including establishing standards for most components, require a much higher priority.** Existing processes need to be standardized, compiled, and shared worldwide, while new, more efficient ones need to be sought.

## VALIDATION AND STANDARDS

**Finding:** All components of a microbial forensics investigation, from methods for collection and sampling, preservation and handling, to identification of the agent, etc., need to be validated.

Although validation is recognized as necessary and essential, it is ill defined. There is a challenge in translating the requirements because there is no standard way to validate an assay or method. This includes developmental validation of one's own method as well as internal validation by other laboratories implementing the method. Minimum validation criteria



include sensitivity, specificity, reproducibility, precision, accuracy, robustness, analyses consistent with the samples and the intended application of an assay, for example, reference panels and mock or nonprobative materials. Other criteria that tend to apply to many assays but may or may not apply in every situation are resolution, purity, critical equipment calibration, critical reagents, and databases.

Validation presents many other challenges. For example, validating sampling methods is very difficult because there are many types of surfaces and environmental matrices. Another variable is the individual who performs sample collection. Quality management needs for performing microbial forensic work include scientific rigor and validity; methods that will support attribution for criminal investigations; and establishing national and international working guidelines for quality assurance and quality control as applied to microbial forensics. There is a serious need for methods for characterizing uncertainty and resolving conflicting results. All analyses are subject to bias and error, so the limits of a method need to be understood. Interpretations may need to be made with incomplete and mutable data due to the fact that the diversity and variability of microorganisms are unknown and organisms undergo genetic change in response to environmental and other factors. Consequently, methods to accommodate this functional knowledge gap must be devised. Although many analysts develop new tests, there is also a need to look at old, less-used, or abandoned techniques, since nonstate actors may not have access to the latest sophisticated technology. As noted above, although the precise identification of an analyte (DNA, etc.) can now be accomplished to the level of a distinct strain using, for example, canonical SNPs at the few- or even the single-molecule level, the validation of the processes to detect rare and low-level signatures is problematic.

The limits of interpretation also need to be clearly understood, and statistical approaches should be rigorously applied in determining whether two microbial samples “match” or are identical, their degree of similarity, their association, and/or their most recent common ancestor. Given what current science and informatics can provide, what is the best that can be expected, accepted, and communicated regarding the probative characteristics of a microorganism and determining its source with acceptable confidence? It is important to understand what a negative result means, particularly for metagenomic analyses. Finally, analysis is not equal to interpretation, and there can be alternative explanations, so there is a need to test hypotheses and try to disprove them, not just attempt to prove them. Velsko (2011a; see also Conclusion 10) emphasized that we must begin an interactive and committed process to address these issues *now*. Scrambling to prepare as an event occurs is an undesirable scenario.

**Conclusion 10:** There is a great need for establishing criteria and requirements for validation. In addition, rapid and accurate development of additional validated analytical methods for microbial forensics must continue. These efforts should include establishment of standards for most components of microbial forensic analyses. **A compilation of all protocols in use (e.g., for sampling, DNA extraction and isolation, and sequencing) and whether and how they have been validated is essential.**

## BIOINFORMATICS AND DATA

Bioinformatics is an interdisciplinary field that develops and improves on methods for storing, retrieving, organizing, and analyzing biological data. A major activity in bioinformatics is to develop software tools to generate useful biological knowledge. To do so, bioinformatics uses many areas of computer science, mathematics, and engineering to process biological data. Complex machines are used to generate biological data at a much faster rate than before. Databases and information systems are used to store and organize biological data. Analyzing biological data may involve algorithms in artificial intelligence, soft computing,<sup>1</sup> data mining, image processing, and simulation. The algorithms in turn depend on theoretical foundations such as control theory, system theory, information theory, and statistics.

The quality of sequence data and the results of bioinformatics analyses must be as high as possible. Factors to consider that impact data interpretation and quality include the quality metrics of sequence data; sequence errors and uncertainties; reliable standards for genomic data representation; uncertainty with databases used; inferences based on available data, including meta-data; formulation of well-defined hypotheses; and testing methods for assessing the weight of microbial forensics evidence.

Data quality and interpretation both require reliable standards for genomic data representation. First, there must be an understanding of sequence quality, sequence errors, and uncertainties about the output data. Second, for interpretation with respect to attribution, criteria are needed for comparisons: match, similar, different, inconclusive; rigor of reasoning by the expert; and well-defined hypotheses and testing methods for assessing the weight of microbial forensic evidence.

Understanding what various bioinformatics results mean may be difficult because one laboratory cannot precisely replicate all essential details of any particular bioinformatics analysis pipeline used by another laboratory. Analyses are complex, different versions of programs exist, and

---

<sup>1</sup> Use of inexact or approximate solutions for computationally difficult tasks, often when the computing power needs exceed the task.

software and hardware change rapidly. Software libraries are updated frequently, and investigators often find that the software they were using is obsolete before their project is completed or validated. No bioinformatics software comes as a stand-alone. All programs depend on multiple libraries, each of which has its own versions, “bug” fixes, and potential parameter settings. It is sometimes uncertain how changes in any of those library modules might affect answers coming from any particular program. For example, changes in a random number generator seed or in a floating-point math library can cause many programs to exhibit different behavior if they incorporate Monte Carlo calculations. At an even more fundamental level, the version and patch level of the underlying operating system and the CPU chip architecture and release level could affect outputs. Although some differences may be insignificant, others may be quite important and these need to be understood. In addition, there are important issues regarding how much documentation is required. The baseline assumptions that are needed to ensure assessments and effective comparisons of technologies need to be defined. Validating bioinformatics pipelines requires data acquisition at all phases of the process, from extraction to data analysis to interpretation.

**Finding:** The new “deep-sequencing” methods generate data in the range of tens to hundreds of gigabases. Although the capability to detect polymorphisms for attribution at a level that was not possible only a few years ago is a great advance, it also presents new challenges. It requires novel ways to handle and analyze terabytes of data. Given this new demand, bioinformatics is both a necessity and a challenge.

**Conclusion 11: Refinement of bioinformatics and statistical methods for evaluating evidence in microbial forensics is needed, including new algorithms that scale to very large and complex datasets.** Bioinformatics also needs to be made understandable and user-friendly to laboratory users, first responders, the public, and the policy makers. Clear requirements are critical in order to attract industry to develop needed hardware and software for bioinformatics.

### Data Sharing

**Finding:** One highly consistent theme heard at the workshop was that data sharing is critical and desperately needed, encompassing biological information on gene sequences, software, and protocols, and standard operating procedures for all microbial forensic methods. Sharing of such data has the potential to promote international collaboration and cooperation among scientists and, more importantly, to inspire innovation.

It might improve security efforts in the sense of promoting more rapid responses to infectious disease outbreaks, which could be crucial in this age of rapid global transportation. Although most participants agreed on the importance of data and information sharing, many expressed concerns about its feasibility and acceptability to governments and policy makers. While cloud computing offers solutions to data sharing as well as storage issues, these possibilities generate security concerns for microbial forensics. In addition, the quality of what is shared is crucial, so standardization of characterization assays was deemed to be key for data sharing.

**Conclusion 12: Discussions are needed under the auspices of an international body that has the respect of the international political and scientific communities about how to share microbial forensic data, and for developing and presenting cogent arguments that can be persuasive to political leaders and scientists worldwide.**

### Databases and Reference Collections

**Finding:** The importance of having a comprehensive archive, or set of archives, of reference materials was emphasized by numerous participants at the workshop, although it was not clear to what extent an archive should contain organisms, nucleic acids only, or just sequence information, the latter of which is technically not a reference collection but a database. An institution could facilitate the development of standardized nomenclature, typing techniques, and characterization systems; uniform QA/QC requirements; standardization of new techniques and analytical evaluations; and reference samples for high-resolution genomic comparisons. It is important that any centralized reference entity maintain a reasonably comprehensive collection; have sufficient long-term support; maintain the highest possible QA/QC standards; use accepted and standardized methods; strengthen standards for curation and material preparation; be responsive to the needs of research and development and support R&D that updates our understanding of diversity; be governed as a community resource; and maintain procedures that balance access and security concerns. Ideally, a reference archive would meet the needs of both criminal investigators and researchers. However, there is skepticism about the feasibility of capturing the full diversity of the microbial world and all the permutations necessary for the multitude of analytical methods in a single archive. However, if highest priority were given to archiving the pathogens and near neighbors that may be used in biocrimes or bioterrorism, this would have considerable value for microbial forensics. Even if less comprehensive collections are likely to be inadequate for most microbial forensic investigations, they would

still be suitable for basic research and provide at least a nominal degree of centralization, which is required to move forward on an effective system for dynamic response and accountability.

**Conclusion 13: An international effort to design and establish more systematic and comprehensive reference collections and databases for pathogens and other microorganisms could take advantage of existing models, such as the World Data Centre for Microorganisms and the American Type Culture Collection. A model system for a consortium of reference collections and data storage centers could be created and later scaled up to become more inclusive.**

### TRAINING AND EDUCATION

Velsko (2011a:522) provides an informative concluding paragraph:

If there is any lesson to be drawn from past experience, it is that microbial forensic collection and analysis are not very effective when they are conducted as ad hoc activities, by non-specialists, using improvised methods and on-the-fly attempts at validation, without prior review by a knowledgeable community. The utility of microbial forensic analysis rises in proportion to the extent that it is anticipatory, well planned, driven by a cadre of qualified experts, and resourced adequately.

**Finding:** The committee agrees with this statement. Microbial forensics is still a relatively new discipline. Training is needed for a number of purposes, including

- Increasing the availability of trained microbial forensics practitioners. Given that microbial forensics encompasses a number of complex technical areas, any such training would necessarily require multidisciplinary, or even transdisciplinary, approaches. A core microbial forensics training program that is available worldwide is needed, and a determination of who would be responsible for developing and implementing such a program is required as soon as possible. Training, and perhaps even certifications, in microbial forensics disciplines would help make microbial forensics a more widely accepted basis for law enforcement actions for use in both domestic and international settings.
- Increasing the awareness and preparedness of first responders, which is essential for both safety purposes (to prevent accidental exposures of responder personnel to hazardous pathogens) and

for law enforcement needs to ensure that evidence samples are not compromised, crime scenes are not contaminated, etc.

- Improving policy-maker and public understanding of what microbial forensics is and what it can accomplish. Currently, there is scant literature available for either the general public or policy makers that explains either the nature or significance of microbial forensics or the fact that moving microbial forensics to a fully mature and accepted set of law enforcement tools requires substantial long-term scientific effort as well as new policies to facilitate that effort and its implementation.

**Conclusion 14: An expansion of technically based training is needed to “professionalize” microbial forensics and increase the number of qualified practitioners worldwide by engaging international professional organizations or other entities that have experience providing training in related fields.**

### SUMMING UP

In its efforts to identify what is required to develop microbial forensics further, the committee used a generous definition of “science,” including research to improve fundamental scientific understanding of microbes, specialized research intended for particular applications in public health, law enforcement, or elsewhere, and an array of technologies and methods that support both basic and applied research. The committee chose breadth over depth and gave particular attention to those science and technology areas that would benefit from international cooperation and collaboration. The result was an extensive list of needs, as outlined in the preceding section. The committee also identified a variety of procedural and policy needs, such as common understandings and protocols for taking and managing samples within and between nations. These additional needs must be addressed if the scientific and technical advances the committee calls for are to make microbial forensics a more effective tool for responding to natural, accidental, or deliberate disease outbreaks. This report is not meant to, nor can it, provide a detailed roadmap for the international development of microbial forensics, but rather elucidates the major issues that the committee believes need to be addressed for the global development of the science of microbial forensics.

## SETTING PRIORITIES

The list of needs identified by the committee is long, but the successful development of microbial forensics will require addressing all of them. The committee emphasizes, however, that there are considerable differences in how difficult it will be to address the needs and in whether there are already existing national or international efforts, for example, in basic research, public health, or industry, that can be drawn upon to help achieve the desired results. There are also a number of high-priority needs that are particularly challenging tasks with long lead times to achieving real progress so that efforts should begin or expand soon. The latter needs will also require substantial and sustained support from governments as major funders of the research, development, and implementation that will be essential for achieving success.

Table 8-1 presents the needs identified in this chapter organized according to the key features discussed above:

- One set of needs represents tasks, for example, the need to identify and characterize a significantly increased number of microbial species that are particularly challenging and/or require a long lead time to achieve the desired results. Such efforts will require the involvement of governments to provide the research resources to carry them out over many years and should be given priority by participating institutions.
- The second set represents needs that could take advantage of ongoing efforts to advance the development of microbial forensics, but will require deliberate communication efforts and in some cases funding to ensure that microbial forensics applications are actually included and implemented.
- The third set of needs has the advantage of either a relatively short lead time to make substantial progress or the existence of significant markets that will provide incentives for industry to produce what is required. For example, the production of faster and cheaper instruments for diagnostics for medicine and genomic analyses for microbial forensics will probably be conducted by industry, which is always seeking to put improved devices on the market.

The committee recognizes that there is overlap among the categories and that some of the needs would fit within more than one of them. It nevertheless believes that exercises like this can be helpful in thinking about implementation issues and for the eventual development of a more detailed roadmap to guide future efforts.



TABLE 8-1 Priority Categories for Microbial Forensics Needs

**Challenging Tasks and/or Long Lead Times**

- An international collaboration engaging the worldwide scientific community in a systematic effort to identify, monitor, and characterize a far higher proportion of global microbial species to increase knowledge about endemism and background. The effort should begin with known pathogens and then expand to their close relatives as well as emerging pathogens. (**Conclusion 1, Basic Science**)
- Development of high-confidence methods to distinguish among natural, accidental, and deliberate outbreaks of infectious disease. (**Conclusion 2, Basic Science**)
- Increased emphasis on development and validation of processes (sample collection, preservation, handling, storage, packaging, and transportation) and analytical methods for microbial forensics, including establishing standards for most components. (**Conclusion 9, Validation and Standards**)
- Discussions under the auspices of an international body that has the respect of the international political and scientific communities about how to share microbial forensic data, and for developing and presenting cogent arguments that can be persuasive to political leaders and scientists worldwide. (**Conclusion 12, Bioinformatics and Data/Data Sharing**)
- An international effort to design and establish more systematic and comprehensive reference collections and databases for pathogens and other microorganisms. This effort could take advantage of existing models, such as the World Data Centre for Microorganisms and the American Type Culture Collection. A model system for a consortium of reference collections and data storage centers could be created and later scaled up to become more inclusive. (**Conclusion 13, Bioinformatics and Data/databases and Reference Collections**)

**Ongoing Efforts on Which to Build**

- Increased emphasis on research to determine mechanisms of pathogenicity, including virulence factors and host immune responses. (**Conclusion 5, Needs Common to Medicine, Public Health, and Microbial Forensics**)
- Priority research is needed to realize the promise of metagenomics and its application to microbial forensics and the development of the forensic value of the other “omics”: proteomics, metabolomics, transcriptomics, glycomics, immunogenomics, etc. (**Conclusion 3, Basic Science**)



TABLE 8-1 Continued

- 
- Greatly improved global disease monitoring and surveillance in humans, animals, and plants to facilitate rapid response and better disease control. (**Conclusion 6, Needs Common to Medicine, Public Health, and Microbial Forensics**)
  - Improved worldwide access to molecular diagnostics (PCR, WGS, etc.), including refinement and distribution of benchtop NGS instruments that are fast and affordable and have simple workflow procedures. (**Conclusion 4, Needs Common to Medicine, Public Health, and Microbial Forensics**)
  - High priority placed on continued research and development to improve physical science applications to microbial forensics. (**Conclusion 8, Methods and Technologies**)
  - Refinement of bioinformatics and statistical methods for evaluating evidence in microbial forensics, including new algorithms that scale to very large or complex datasets. (**Conclusion 11, Bioinformatics and Data**)

#### **Shorter Lead Times or Industry Incentives**

- Development of more advanced, faster, and cheaper assay and sequencing technologies that can be standardized and made more accessible to benefit both microbial forensics and public health. (**Conclusion 7, Methods and Technologies**)
  - A compilation of all protocols in use (e.g., for sampling, DNA extraction and isolation, and sequencing) and whether and how they have been validated. (**Conclusion 10, Validation and Standards**)
  - Expansion of technically based training to “professionalize” microbial forensics and increase the number of qualified practitioners worldwide by engaging international professional organizations or other entities that have experience providing training in related fields. (**Conclusion 14, Training and Education**)
- 

Microbial forensics has specialized needs because of the demand for “evidence” and “proof” in the context of law enforcement or international policy. Many of these relate as much to the quality of the process by which material is collected and analyzed as to the science and technology employed. The strong presence of public health needs in this report reflects the global reality that in most countries the only capabilities and activities relevant to microbial forensics occur in the context of public health. The United States is largely the exception in having a microbial

forensics science community that is differentiated from public health, reflecting its experiences in the early 2000s with the anthrax letters. Some militaries in developed countries are creating microbial forensic capabilities, and nations such as the United Kingdom, Sweden, Germany, China, and France have basic microbial forensics infrastructures. But for most of the rest of the world, microbial forensics is a side activity of public health officials. Since most disease outbreaks will first be recognized through the public health infrastructure, strengthening detection and diagnostic capacities there serves both public health and microbial forensics. In this regard, the new Global Health Security Agenda as an alliance among over 25 countries, the European Union, and three international organizations (World Health Organization [WHO], World Organization for Animal Health (OIE), and Food and Agriculture Organization (FAO) illustrates the potential for initiatives to address global disease threats of any origin to contribute to building fundamental microbial forensic capabilities. As one of the documents announcing the new collaboration framed the challenge:

An interconnected world is increasing the opportunities for human, animal and zoonotic diseases to emerge and spread globally. Today's health security threats arise from at least 5 sources: the emergence and spread of new microbes; the globalization of travel and food supply; the rise of drug-resistant pathogens; the acceleration of biological science capabilities and the risk that these capabilities may cause the inadvertent or intentional release of pathogens; and continued concerns about terrorist acquisition, development, and use of biological agents. The recent emergence of the H7N9 influenza virus and Middle East Respiratory Syndrome Coronavirus underscore infectious disease as a serious global threat. Since the emergence of Severe Acute Respiratory Syndrome in 2003, the world has made great progress in strengthening local, regional, and international capacity to prevent, detect and respond to emerging infectious disease threats. Yet, despite important accomplishments, much remains to be done to achieve our shared global health security vision. Only 16% of countries reported reaching full compliance with the core IHR [International Health Regulations] competencies by the June 2012 deadline set by the WHO. Vulnerabilities include geographic areas with limited disease surveillance systems, reluctance to share outbreak information or biological samples, emergence of new pathogens and development of drug-resistance, and the specter of intentional or accidental release of biological agents. Multi-sectoral collaboration and the combined resources and expertise of the health and security sectors will be required to efficiently match resources to needs, avoid redundant efforts, and identify gaps. In 2013, for the first time, the G20 called upon countries to strengthen compliance with the WHO IHR—the standard by which the world measures its preparedness for emerging disease threats, as well as bioterrorist events. (White House, 2014)

The committee commends the Global Health Security Agenda and hopes that its implementation will include a conscious effort to develop new capacities and connections among institutions, including those in the scientific community, that have a role in responding to this range of threats. Although the new agenda promotes a welcome multisectoral collaboration between the health and security sectors, it will be important to consider how the needs of law enforcement and legal standards for “evidence” and “proof” can be related to the new Agenda. It is hoped that the Agenda will eventually accommodate clear considerations of the needs of microbial forensics.

## References

- AAM (American Academy of Microbiology). 2009a. *Bioinformatics and Biodefense: Keys to Understanding Natural and Altered Pathogens*. Washington, DC: AAM.
- AAM. 2009b. *Large Scale Sequencing: The Future of Genomic Sciences?* Washington, DC: AAM.
- AAM. 2011a. *E. coli: Good, Bad, and Deadly*. Washington, DC: AAM.
- AAM. 2011b. *The Rare Biosphere*. Washington, DC: AAM.
- Alberts, B., A. Johnson, J. Lewis, M. Raff, K. Roberts, and P. Walter. 2002. *Molecular Biology of the Cell*, 4th Ed. New York: Garland Science.
- Anderson, I., A. Sorokin, V. Kapatral, G. Reznik, A. Bhattacharya, N. Mikhailova, H. Burd, V. Joukov, D. Kaznadzey, T. Walunas, M. D'Souza, N. Larsen, G. Pusch, K. Liolios, Y. Grechkin, A. Lapidus, E. Goltsman, L. Chu, M. Fonstein, S. D. Ehrlich, R. Overbeek, N. Kyrpides, and N. Ivanova. 2005. Comparative genome analysis of *Bacillus cereus* group genomes with *Bacillus subtilis*. *FEMS Microbiology Letters* 250(2):175-184.
- Anthony, S. J., J. H. Epstein, K. A. Murray, I. Navarrete-Maclas, C. M. Zambrana-Torello, A. Solonyov, R. Ojeda-Flores, N. C. Arrigo, A. Islam, S. A. Kahn, P. Hosseini, T. L. Bogich, K. J. Olival, M. D. Sanchez-Leon, W. B. Karesh, T. Goldstein, S. P. Luby, S. S. Morse, J. A. K. Mazet, P. Daszak, and W. I. Lipkin. 2013. A strategy to estimate unknown viral diversity in mammals. *mBio* 4(5):e00598-13; doi:10.1128/mBio.00598-13.
- Avise, J. 2000. *Phylogeography: The History and Formation of Species*. Cambridge, MA: Harvard University Press.
- Bérard S., C. Gallien, B. Boussau, G. J. Szöllösi, V. Daubin, and E. Tannier. 2012. Evolution of gene neighborhoods within reconciled phylogenies. *Bioinformatics* 28(18):i382-i388.
- Bergthorsson, U., and H. Ochman. 1995. Heterogeneity of genome sizes among natural isolates of *Escherichia coli*. *Journal of Bacteriology* 177:5784-5789. Available at <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC177399/pdf/1775784.pdf>.
- Bhattacharjee, Y. 2010. Anthrax investigation: Silicon mystery endures in solved anthrax case. *Science* 327(5972):1435.
- Biological Weapons Convention. 2004. Report of the Meeting of States Parties. BWC/MSP/2004/3. Available at <http://daccess-dds-ny.un.org/doc/UNDOC/GEN/G04/643/00/PDF/G0464300.pdf?OpenElement>. Accessed December 17, 2013.

- Biological Weapons Convention. 2010. Report of the Meeting of States Parties. BWC/MSP/2010/6. Available at <http://daccess-dds-ny.un.org/doc/UNDOC/GEN/G11/600/18/PDF/G1160018.pdf?OpenElement>. Accessed December 17, 2013.
- Biological Weapons Convention. 2011. Final Document of the Seventh Review Conference. BWC/CONF.VII/7. Available at <http://daccess-dds-ny.un.org/doc/UNDOC/GEN/G12/600/60/PDF/G1260060.pdf?OpenElement>. Accessed December 17, 2013.
- Black, J. G. 2002. *Microbiology: Principles and Explorations*, 5th Ed. New York: John Wiley and Sons, Inc.
- Boussau, B., G. J. Szollosi, L. Duret, M. Guoy, E. Tannier, and V. Daubin. 2013. Genome-scale coestimation of species and gene trees. *Genome Research* 23:323-330. Available at <http://genome.cshlp.org/content/23/2/323.short>.
- Bowen, D. G., and C. M. Walker. 2005. Adaptive immune responses in acute and chronic hepatitis C virus infection. *Nature* 436:946-952.
- Boyer, A. E., M. Gallegos-Candela, R. C. Lins, Z. Kuklenyik, H. M. Woolfitt, S. Kalb, C. P. Quinn, and J. R. Barr. 2011. Quantitative mass spectrometry for bacterial protein toxins—a sensitive, specific, high-throughput tool for detection and diagnosis. *Molecules* 16:2391-2413.
- Brennan, Z. 2013. Novartis, Venter Institute collaborate on new H7N9 vaccine with CDC. in-Pharmatechnologist.com. April 9. Available at <http://www.in-pharmatechnologist.com/Regulatory-Safety/Novartis-Venter-Institute-Collaborate-on-New-H7N9-Vaccine-with-CDC>. Accessed March 31, 2014.
- Budowle, B., S. E. Schutzer, A. Einseln, L. C. Kelley, A. C. Walsh, J. A. L. Smith, B. L. Marrone, J. Robertson, and J. Campos. 2003. Building microbial forensics as a response to bioterrorism. *Science* 301(5641):1852-1853.
- Budowle, B., R. Murch, and R. Chakraborty. 2005. Microbial forensics: The next forensic challenge. *International Journal of Legal Medicine* 119(6):317-330.
- Budowle, B., S. E. Schutzer, S. A. Morse, K. F. Martinez, R. Chakraborty, B. L. Marrone, S. L. Messenger, R. S. Murch, P. J. Jackson, P. Williamson, R. Harmon, and S. P. Velsko. 2008. Criteria for validation of methods in microbial forensics. *Applied and Environmental Microbiology* 74(18):5599-5607.
- Budowle, B., S. E. Schutzer, R. G. Breeze, P. S. Keim, and S. A. Morse, eds. 2011. *Microbial Forensics*, 2nd Ed. Burlington, MA: Academic Press.
- Budowle, B., S. Schmedes, and R. S. Murch. 2013. The microbial forensics pathway for use of massively-parallel sequencing technologies. In E. R. Choffnes, L. Olsen, and T. Wizemann, rapporteurs. *The Science and Applications of Microbial Genomics*. Washington, DC: The National Academies Press.
- Burden, R. W. 2010. Extension Agent Critical Infrastructure Threat Awareness Training, Introduction to Food Supply System Threats. Available at: <http://www.uteasternregion.org/EMERGENCY%20PREPAREDNESS/DOWNLOAD%20%20EMERGENCY%20PREPAREDNESS/Extension%20Agents%20&%20Critical%20Infrastructure%20Protection.pdf>. Accessed April 10, 2014.
- Carus, W. S. 2001. *Bioterrorism and Biocrimes: The Illicit Use of Biological Agents Since 1900*, 8th rev. Washington, DC: Center for Counterproliferation Research, National Defense University.
- CDC (U.S. Centers for Disease Control and Prevention). 2011. Minutes of the Meeting of the Board of Scientific Counselors, Office of Disease Control and Prevention, November 9, 2011. Atlanta, GA.
- Chertow, D. S., and M. J. Memoli. 2013. Bacterial coinfection in influenza: A grand rounds review. *Journal of the American Medical Association* 309(3):275-282.
- Cho, Y. J., H. Yi, J. H. Lee, D. W. Kim, and J. Chun. 2010. Genomic evolution of *Vibrio cholerae*. *Current Opinion in Microbiology* 13(5):646-51.

- Chun, J., C. J. Grim, N. A. Hasan, J. H. Lee, S. Y. Choi, B. J. Haley, E. Taviani, Y. S. Jeon, D. W. Kim, J. H. Lee, T. S. Brettin, D. C. Bruce, J. F. Challacombe, J. C. Detter, C. S. Han, A. C. Munk, O. Chertkov, L. Meincke, E. Saunders, R. A. Walters, A. Huq, G. B. Nair, and R. R. Colwell. 2009. Comparative genomics reveals mechanism for short-term and long-term clonal transitions in pandemic *Vibrio cholera*. *Proceedings of the National Academy of Sciences of the United States of America* 106(36):15442-15447.
- Cronquist, A. B., R. K. Mody, R. Atkinson, J. Besser, M. Tobin D'Angelo, S. Hurd, T. Robinson, C. Nicholson, and B. E. Mahon. 2012. Impacts of culture-independent diagnostic practices on public health surveillance for bacterial enteric pathogens. *Clinical Infectious Disease* 54(Suppl. 5):S432-S439.
- Cui, Y., C. Yu, Y. Yan, D. Li, Y. Li, T. Jombart, L. A. Weinert, Z. Wang, Z. Guo, L. Xu, Y. Zhang, H. Zheng, N. Qin, X. Xiao, M. Wu, X. Wang, D. Zhou, Z. Qi, Z. Du, H. Wu, X. Yang, H. Cao, H. Wang, J. Wang, S. Yao, A. Rakin, Y. Li, D. Falush, F. Balloux, M. Achtman, Y. Song, J. Wang, and R. Yang. 2013. Historical variations in mutation rate in an epidemic pathogen, *Yersinia pestis*. *Proceedings of the National Academy of Sciences of the United States of America* 110(2):577-582.
- Danzig, R., and Z. Hosford. 2012. *Aum Shinrikyo: Insights into How Terrorists Develop Biological and Chemical Weapons*, 2nd Ed. Washington, DC: Center for a New American Security.
- Detter, J. C., and I. G. Resnick. 2014. State of the art for autonomous detection systems using genomic sequencing. Pp. 197-213 in Institute of Medicine and National Research Council, *Technologies to Enable Autonomous Detection for BioWatch: Ensuring Timely and Accurate Information for Public Health Officials: Workshop Summary*. Washington, DC: The National Academies Press.
- Di Liberto, G., A. M. Roque-Afonso, R. Kara, D. Ducoulombier, G. Fallot, D. Samuel, and C. Feray. 2006. Clinical and therapeutic implications of hepatitis C virus compartmentalization. *Gastroenterology* 131(1):76-84.
- Didelot, X., G. Méric, D. Falush, and A. E. Darling. 2012. Impact of homologous and non-homologous recombination in the genomic evolution of *Escherichia coli*. *BMC Genomics* 13:256. Available at <http://www.biomedcentral.com/1471-2164/13/256>. Accessed April 10, 2014.
- Dubay, L. 2013. Next-gen DNA sequencing system receives FDA approval for clinical use. *BioOptics World*. Available at: <http://www.bioopticsworld.com/articles/2013/11/next-gen-dna-sequencing-system-receives-fda-approval-for-clinical-use.html>. Accessed April 10, 2014.
- Ellard, S., H. Lindsay, N. Camm, C. Watson, S. Abbs, G. R. Taylor, and R. Charlton. 2012. Practice Guidelines for Targeted Next Generation Sequencing Analysis and Interpretation. Clinical Molecular Genetics Society. Available at [http://www.acgs.uk.com/media/815227/bpg\\_for\\_targeted\\_next\\_generation\\_sequencing\\_2011134.pdf](http://www.acgs.uk.com/media/815227/bpg_for_targeted_next_generation_sequencing_2011134.pdf). Accessed April 16, 2014.
- Engelthaler, D. M., and S. A. Balajee. 2011. Forensics and epidemiology of fungal pathogens. Chapter 18 in *Microbial Forensics*, 2nd Ed., B. Budowle, S. E. Schutzer, R. G. Breeze, P. S. Keim, and S. A. Morse, eds. Burlington, MA: Academic Press.
- Evans, S. N., and F. A. Matsen. 2012. The phylogenetic Kantorovich-Rubinstein metric for environmental sequence samples. *Journal of the Royal Statistical Society, Series B: Statistical Methodology* 74(3):569-592.
- Evetts, I. W., and B. S. Weir. 1998. *Interpreting DNA Evidence: Statistical Genetics for Forensic Scientists*. Sunderland, MA: Sinauer Associates.
- Fierer, N., C. L. Lauber, N. Zhou, D. McDonald, E. K. Costello, and R. Knight. 2010. Forensic identification using skin bacterial communities. *Proceedings of the National Academy of Sciences of the United States of America* 107(14):6477-6481.

- Fletcher, J., C. Bender, B. Budowle, W. T. Cobb, S. E. Gold, C. A. Ishimaru, D. Luster, U. Melcher, R. Murch, H. Scherm, R. C. Seem, J. L. Sherwood, B. W. Sobral, and S. A. Tolin. 2006. Plant pathogen forensics: Capabilities, needs, and recommendations. *Microbiology and Molecular Biology Reviews* 70:450-471.
- Fricke, W. F., T. A. Cebula, and J. Ravel. 2010. Genomics. Chapter 28 in *Microbial Forensics*, 2nd Ed., B. Budowle, S. E. Schutzer, R. G. Breeze, P. S. Keim, and S. A. Morse, eds. Burlington, MA: Academic Press.
- Gargis, A. S., L. Kalman, M. W. Berry, D. P. Bick, D. P. Dimmock, T. Hambuch, F. Lu, E. Lyon, K. V. Voelkerding, B. A. Zehnbauser, R. Agarwala, S. F. Bennett, B. Chen, E. L. H. Chin, J. G. Compton, S. Das, D. H. Farkas, M. J. Ferber, B. H. Funke, M. R. Furtado, L. M. Ganova-Raeva, U. Geigenmüller, S. J. Gunselman, M. R. Hegde, P. L. F. Johnson, A. Kasarskis, S. Kulkarni, T. Lenk, C. S. Liu, M. Manion, T. A. Manolio, E. R. Mardis, J. D. Merker, M. S. Rajeevan, M. G. Reese, H. L. Rehm, B. B. Simen, J. M. Yeakley, J. M. Zook, and I. M. Lubin. 2012. Assuring the quality of next-generation sequencing in clinical laboratory practice. *Nature Biotechnology* 30:1033-1036.
- González-Candelas, F., M. A. Bracho, B. Wróbel, and A. Moya. 2013. Molecular evolution in court: Analysis of a large hepatitis C virus outbreak from an evolving source. *BioMed Central* 11:76.
- Goris, J., K. T. Konstantinidis, J. A. Klappenbach, T. Coenye, P. Vandamme, and J. M. Tiedje. 2007. DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. *International Journal of Systematic and Evolutionary Microbiology* 57(Pt 1):81-91.
- Grunow, R., and E. J. Finke. 2002. A procedure for differentiating between the intentional release of biological warfare agents and natural outbreaks of disease: Its use in analyzing the tularemia outbreak in Kosovo in 1999 and 2000. *Clinical Microbiology and Infection* 8:510-521. Available at <http://onlinelibrary.wiley.com/doi/10.1046/j.1469-0691.2002.00524.x/pdf>. Accessed April 10, 2014.
- Grunow, R., S. R. Klee, W. Beyer, M. George, D. Grunow, A. Barduhn, S. Klar, D. Jacob, M. Elschner, P. Sandven, A. Kjerulf, J. S. Jensen, W. Cai, R. Zimmermann, and L. Schaade. 2013. Anthrax among heroin users in Europe possibly caused by same *Bacillus anthracis* strain since 2000. *Eurosurveillance* 18(13):pii=20437. Available at <http://www.eurosurveillance.org/ViewArticle.aspx?ArticleId=20437>. Accessed April 10, 2014.
- Guidi, R., P. Osimani, C. Azzari, M. Resti, and F. M. De Benedictis. 2011. Severe necrotizing pneumonia complicating influenza A (H1N1): The role of immunologic interaction. *International Journal of Immunopathological Pharmacology* 24(4):1093-1097.
- Gupta, M., N. Gupta, S. Trivedi, P. Patil, G. Gupta, V. K. Krishna, H. Gaudani, and V. S. Gomase. 2009. Immunogenomics: Recent discoveries. *International Journal of Genetics* 1(2):1-5.
- Hartung, T. 2011. Mapping the Human Toxome by Systems Toxicology. Description of ongoing research project funded by the National Institute of Environmental Health Sciences. Available at [http://tools.niehs.nih.gov/portfolio/index.cfm/portfolio/grantDetail/grant\\_number/R01ES020750](http://tools.niehs.nih.gov/portfolio/index.cfm/portfolio/grantDetail/grant_number/R01ES020750). Accessed April 10, 2014.
- Herndon, A. 2012. A curmudgeon's view of molecular phylogenetics (Part 1). *Bromeli Advisory*, pp. 3-4. Available at <http://www.bssf-miami.org/newsbulletins/July%202012.pdf>. Accessed April 10, 2014.
- Hugh-Jones, M. E., B. H. Rosenberg, and S. Jacobsen. 2011. The 2001 attack anthrax: Key observations. *Bioterrorism & Biodefense* S3:1-10.
- Human Microbiome Jumpstart Reference Strains Consortium. 2010. A catalog of reference genomes from the human microbiome. *Science* 328:994-999.
- Hunt, S. Y., N. G. Barnaby, and B. Budowle. 2009. Forensic microbiology. Pp. 22-34 in *Encyclopedia of Microbiology*, 3rd Ed. Burlington, MA: Academic Press.



- IOM/NRC (Institute of Medicine and National Research Council). 2014. *Technologies to Enable Autonomous Detection for BioWatch: Ensuring Timely and Accurate Information for Public Health Officials: Workshop Summary*. Washington, DC: The National Academies Press.
- JASON Advisory Panel. 2008. *Microbial Forensics*. JASON Report No. JSR-08-512. Report prepared for the National Counterproliferation Center. Available at <http://www.fas.org/irp/agency/dod/jason/forensics.pdf>. Accessed November 19, 2013.
- Jernigan, D. B., P. L. Raghunathan, B. P. Bell, R. Brechner, E. A. Bresnitz, J. C. Butler, M. Cetron, M. Cohen, T. Doyle, M. Fischer, C. Greene, K. S. Griffith, J. Guarner, J. L. Hadler, J. A. Hayslett, R. Meyer, L. R. Petersen, M. Phillips, R. Pinner, T. Popovic, C. P. Quinn, J. Reefhuis, D. Reissman, N. Rosenstein, A. Schuchat, W. J. Shieh, L. Siegal, D. L. Swerdlow, F. C. Tenover, M. Traeger, J. W. Ward, I. Weisfuse, S. Wiersma, K. Yeskey, S. Zaki, D. A. Ashford, B. A. Perkins, S. Ostroff, J. Hughes, D. Fleming, J. P. Koplan, and J. L. Gerberding. 2002. Investigation of bioterrorism-related anthrax, United States, 2001: Epidemiologic findings. *Emerging Infectious Diseases* 8(10):1019-1028.
- Johnson, R. C., S. R. Kalb, and J. R. Barr. 2011. Toxin analysis using mass spectrometry. Chapter 24 in *Microbial Forensics*, 2nd Ed., B. Budowle, S. E. Schutzer, R. G. Breeze, P. S. Keim, and S. A. Morse, eds. Burlington, MA: Academic Press.
- Jolley, K. A., C. M. Bliss, J. S. Bennett, H. B. Bratcher, C. Brehony, F. M. Colles, H. Wimalaratna, O. B. Harrison, S. K. Sheppard, A. J. Cody, and M. C. J. Maiden. 2012. Ribosomal multi-locus sequence typing: Universal characterisation of bacteria from domain to strain. *Microbiology* 158(4):1005-1015.
- Jünemann, S., F. J. Sedlazeck, K. Prior, A. Albersmeier, U. John, J. Kalinowski, A. G. Mellman, A. von Haeseler, J. Stoye, and D. Harmsen. 2013. Updating benchtop sequencing performance comparison. *Nature Biotechnology* 31:294-296.
- Kalb, S. R., W. I. Santana, I. N. Geren, C. Garcia-Rodriguez, J. Lou, T. J. Smith, J. D. Marks, L. A. Smith, J. L. Pirkle, and J. R. Barr. 2011. Extraction and inhibition of enzymatic activity of botulinum neurotoxins /B1, /B2, /B3, /B4, and /B5 by a panel of monoclonal anti-BoNT/B antibodies. *BMC Biochemistry* 12:58.
- Keim, P., L. B. Price, A. M. Klevytska, K. L. Smith, J. M. Schupp, R. Okinaka, P. J. Jackson, and M. E. Hugh-Jones. 2000. Multiple-locus variable-number tandem repeat analysis reveals genetic relationships within *Bacillus anthracis*. *Journal of Bacteriology* 182(10):2928-2936.
- Keim, P., K. L. Smith, C. Keys, H. Takahashi, T. Kurata, and A. Kaufmann. 2001. Molecular investigation of the Aum Shinrikyo anthrax release in Kameido, Japan. *Journal of Clinical Microbiology* 39(12):4566-4567.
- Keim, P., M. Van Ert, T. Pearson, A. Vogler, L. Hyunh, and D. M. Wagner. 2004. Anthrax molecular epidemiology and forensics: Using different markers for the appropriate evolutionary scales. *Infection, Genetics and Evolution* 4:205-213.
- Khan, A. S., and N. Pesik. 2011. Forensic public health: Epidemiologic and microbiologic investigations for biosecurity. Chapter 15 in *Microbial Forensics*, 2nd Ed., B. Budowle, S. E. Schutzer, R. G. Breeze, P. S. Keim, and S. A. Morse, eds. Burlington, MA: Academic Press.
- Kodama, Y., M. Shumway, and R. Leinonen. 2012. The Sequence Read Archive: Explosive growth of sequencing data. *Nucleic Acids Research* 40 (Database issue):D54-56.
- Koepfel, A., E. B. Perry, J. Sikorski, D. Krizanc, A. Warner, D. M. Ward, A. P. Rooney, E. Brambila, N. Connor, R. M. Ratcliff, E. Nevo, and F. M. Cohan. 2008. Identifying the fundamental units of bacterial diversity: A paradigm shift to incorporate ecology into bacterial systematics. *Proceedings of the National Academy of Sciences of the United States of America* 105(7):2504-2509.



- Köser, C. U., M. T. Holden, M. J. Ellington, E. J. Cartwright, N. M. Brown, A. L. Ogilvy-Stuart, L. Y. Hsu, C. Chewapreecha, N. J. Croucher, S. R. Harris, M. Sanders, M. C. Enright, G. Dougan, S. D. Bentley, J. Parkhill, L. J. Fraser, J. R. Betley, O. B. Schulz-Trieglaff, G. P. Smith, and S. J. Peacock. 2012. Rapid whole-genome sequencing for investigation of a neonatal MRSA outbreak. *New England Journal of Medicine* 366(24):2267-2275.
- Köser, C. U., J. M. Bryant, J. Becq, M. E. Török, M. J. Ellington, M. A. Marti-Renom, A. J. Carmichael, J. Parkhill, G. P. Smith, and S. J. Peacock. 2013. Whole-genome sequencing for rapid susceptibility testing of *M. tuberculosis*. *New England Journal of Medicine* 369:290-292.
- Larkin, M. 2001. What will the Human Genome Project mean for neurology? *Neurology Today* 1(2):22-24.
- Lipkin, W. I. 2013. The changing face of pathogen discovery and surveillance. *Nature Reviews Microbiology* 2:133-141.
- Loman, N. J., C. Constantinidou, J. Z. M. Chan, M. Halachev, M. Sergeant, C. W. Penn, E. R. Robinson, and M. J. Pallen. 2012. High-throughput bacterial genome sequencing: An embarrassment of choice, a world of opportunity. *Nature Reviews Microbiology* 10:599-606.
- Maiden, M. C., M. J. van Rensburg, J. E. Bray, S. G. Earle, S. A. Ford, K. A. Jolley, and N. D. McCarthy. 2013. MLST revisited: The gene-by-gene approach to bacterial genomics. *Nature Reviews Microbiology* 10:728-736.
- Markotić, A., I. Kuzman, K. Babić, A. Gagro, S. Nichol, T. G. Ksiazek, S. Rabatić, and D. Dekaris. 2002a. Double trouble: Hemorrhagic fever with renal syndrome and leptospirosis. *Scandinavian Journal of Infectious Diseases* 34(3):221-224.
- Markotić, A., S. T. Nichol, I. Kuzman, A. J. Sanchez, T. G. Ksiazek, A. Gagro, S. Rabatić, R. Zgorelec, T. Avsic-Zupanc, I. Beus, and D. Dekaris. 2002b. Characteristics of Puumala and Dobrava infections in Croatia. *Journal of Medical Virology* 66:542-551.
- Markotić, A., L. J. Žmak, and N. Turk. 2011. Detection of T-lymphocyte activation markers in differential diagnosis of infections caused by hantaviruses or leptospira. *Croatian Journal of Infection* 31(4):179-183.
- Mattocks, C. J., M. A. Morris, G. Matthijs, E. Swinnen, A. Corveleyn, E. Dequeker, C. R. Müller, V. Pratt, and A. Wallace. 2010. A standardized framework for the validation and verification of clinical molecular genetic tests. *European Journal of Human Genetics* 18:1276-1288.
- Mayr, E. 1942. *Systematics and the Origin of Species*. New York: Columbia University Press.
- Medved, M. M., A. Markotić L. Cebalo, B. Turković, and T. A. Zupanc. 2002. Haemorrhagic fever with renal syndrome in Croatia. *Lancet* 360(9330):415-416.
- Mellmann, A., M. Bielaszewska, R. Köck, A. W. Friedrich, A. Fruth, B. Middendorf, D. Harmsen, M. A. Schmidt, and H. Karch. 2008. Analysis of collection of hemolytic uremic syndrome-associated enterohemorrhagic *Escherichia coli*. *Emerging Infectious Disease* 14(8):1287-1290.
- Mellmann, A., D. Harmsen, C. A. Cummings, E. B. Zentz, S. R. Leopold, A. Rico, K. Prior, R. Szczepanowski, Y. Ji, W. Zhang, S. F. McLaughlin, J. K. Henkhaus, B. Leopold, M. Bielaszewska, R. Prager, P. M. Brzoska, R. L. Moore, S. Guenther, J. M. Rothberg, and H. Karch. 2011. Prospective genomic characterization of the German enterohemorrhagic *Escherichia coli* O104:H4 outbreak by rapid next generation sequencing technology. *PLOS One* 6:e22751.
- Meselson, M., J. Guillemin, M. Hugh-Jones, A. Langmuir, I. Popova, A. Shelokov, and O. Yampolskava. 1994. The Sverdlovsk anthrax outbreak of 1979. *Science* 266(5188):1202-1208.
- Michael, J. R., and P. G. Kotula. 2009. Elemental Microanalysis of *Bacillus Anthracis* Spores from the Amerithrax Case. Presentation to the National Academies of Science. Materials Characterization Department 1822, Sandia National Laboratory, Albuquerque, NM.

- Michael, J. R., L. N. Brewer, and P. G. Kotula. 2011. Electron beam-based methods for bio-forensic investigations. Chapter 25 in *Microbial Forensics*, 2nd Ed., B. Budowle, S. E. Schutzer, R. G. Breeze, P. S. Keim, and S. A. Morse, eds. Burlington, MA: Academic Press.
- Morelli, G., Y. Song, C. J. Mazzoni, M. Eppinger, P. Roumagnac, D. M. Wagner, M. Feldkamp, B. Kusecek, A. J. Vogler, Y. Li, Y. Cui, N. R. Thomson, T. Jombart, R. Leblois, P. Lichtner, L. Rahalison, J. M. Petersen, F. Balloux, P. Keim, T. Wirth, J. Ravel, R. Yang, E. Carniel, and M. Achtman. 2010. *Yersinia pestis* genome sequencing identifies patterns of global phylogenetic diversity. *Nature Genetics* 42(10):1140-1143.
- Murch, R. 2008. Bioterrorism: Investigation & Prosecution—Anthrax 2001 and Beyond. Presentation at Counterproliferation of Biological Threat Agents, 928th Wilton Park Conference, September 27, Wilton Park, United Kingdom. Available at <http://www.ncr.vt.edu/Highlights/Murch.html>. Accessed April 26, 2014.
- Murch, R. 2010. Exploring an International Microbial Forensics Capability to Support Attribution and Advance Global Biosecurity. Presentation at Trends in Science and Technology Relevant to the BWC, November 2, 2010, Beijing, China. Available at <http://dels.nas.edu/resources/static-assets/bls/miscellaneous/Randall%20Murch.pdf>. Accessed March 19, 2014.
- Murch, R. S., and E. L. Bahr. 2011. Validation of microbial forensics in scientific, legal, and policy contexts. Chapter 38 in *Microbial Forensics*, 2nd Ed., B. Budowle, S. E. Schutzer, R. G. Breeze, P. S. Keim, and S. A. Morse, eds. Burlington, MA: Academic Press.
- National Biosurveillance Advisory Subcommittee. 2009. *Improving the Nation's Ability to Detect and Respond to 21st Century Urgent Health Threats: First Report of the National Biosurveillance Advisory Subcommittee*. April. Available at <http://stacks.cdc.gov/view/cdc/12000>. Accessed April 4, 2014.
- National Biosurveillance Advisory Subcommittee. 2011. *Improving the Nation's Ability to Detect and Respond to 21st Century Urgent Health Threats: Second Report of the National Biosurveillance Advisory Subcommittee*. April. Available at [http://www.cdc.gov/about/pdf/advisory/nbasfinalreport\\_april2011.pdf](http://www.cdc.gov/about/pdf/advisory/nbasfinalreport_april2011.pdf). Accessed April 4, 2014.
- National Institute of Allergy and Infectious Diseases. 2009. Glossary. Available at <http://www.niaid.nih.gov/topics/pathogenGenomics/pages/definitions.aspx>. Accessed April 10, 2014.
- Netherlands National Institute for Public Health and the Environment. 2013. Combating the Superbug *Klebsiella Oxa-48* Outbreak in a Dutch Hospital. Available at [http://www.rivm.nl/en/Documents\\_and\\_publications/Common\\_and\\_Present/Newsmessages/2011/Combating\\_the\\_Superbug\\_Klebsiella\\_Oxa\\_48\\_Outbreak\\_in\\_a\\_Dutch\\_Hospital](http://www.rivm.nl/en/Documents_and_publications/Common_and_Present/Newsmessages/2011/Combating_the_Superbug_Klebsiella_Oxa_48_Outbreak_in_a_Dutch_Hospital). Accessed on April 10, 2014.
- NRC (National Research Council). 2003. *Countering Agricultural Bioterrorism*. Washington, DC: The National Academies Press.
- NRC. 2007. *The New Science of Metagenomics: Revealing the Secrets of Our Microbial Planet*. Washington, DC: The National Academies Press.
- NRC. 2011. *Review of the Scientific Approaches Used During the FBI's Investigation of the 2001 Anthrax Letters*. Washington, DC: The National Academies Press.
- NRC. 2012. *Biosecurity Challenges of the Global Expansion of High-Containment Biological Laboratories*. Washington, DC: The National Academies Press.
- NSTC (National Science and Technology Council). 2009. *National Research and Development Strategy for Microbial Forensics*. Washington, DC: Executive Office of the President. Available at <http://www.whitehouse.gov/files/documents/ostp/NSTC%20Reports/National%20MicroForensics%20R&DStrategy%202009%20UNLIMITED%20DISTRIBUTION.pdf>. Accessed on April 10, 2014.

- NSTC. 2013. *Biological Response and Recovery Science and Technology Roadmap*. Washington, DC: Executive Office of the President. Available at [http://www.whitehouse.gov/sites/default/files/microsites/ostp/NSTC/brrst\\_roadmap\\_2013.pdf](http://www.whitehouse.gov/sites/default/files/microsites/ostp/NSTC/brrst_roadmap_2013.pdf). Accessed on April 10, 2014.
- Orho-Melander, M. 2006. The metabolic syndrome: Genetics, lifestyle and ethnicity. *Diabetes Voice* 5:21-24. Available at [http://www.idf.org/sites/default/files/attachments/article\\_412\\_en.pdf](http://www.idf.org/sites/default/files/attachments/article_412_en.pdf). Accessed on April 10, 2014.
- Parla, J., M. Kramer, and W. R. McCombie. 2011. High-throughput sequencing. Chapter 27 in *Microbial Forensics*, 2nd Ed., B. Budowle, S. E. Schutzer, R. G. Breeze, P. S. Keim, and S. A. Morse, eds. Burlington, MA: Academic Press.
- Penin, F., J. Dubuisson, F. A. Rey, D. Moradpour, and J. M. Pawlotsky. 2004. Structural biology of hepatitis C virus. *Hepatology* 39:5-19.
- Perna, N. T., G. Plunkett, V. Burland, B. Mau, J. D. Glasner, D. J. Rose, G. F. Mayhew, P. S. Evans, J. Gregor, H. A. Kirkpatrick, G. Pósfai, J. Hackett, S. Klink, A. Boutin, Y. Shao, L. Miller, E. J. Grotbeck, N. W. Davis, A. Lim, E. T. Dimalanta, K. D. Potamousis, J. Apodaca, T. S. Anantharaman, J. Lin, G. Yen, D. C. Schwartz, R. A. Welch, and F. R. Blattner. 2001. Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7. *Nature* 409(6819):529-533.
- Pilo, P., and J. Frey. 2011. *Bacillus anthracis*: Molecular taxonomy, population genetics, phylogeny and patho-evolution. *Infection, Genetics, and Evolution* 11(6):1218-1224.
- Pont-Kingdon, G., F. Gedge, W. Wooderchak-Donahue, I. Schrijver, K. E. Weck, J. A. Kant, D. Oglesbee, P. Bayrak-Toydemir, and E. Lyon. 2012. Design and analytical validation of clinical DNA sequencing assays. *Archives of Pathology & Laboratory Medicine* 136:41-46.
- Price, E. P., M. L. Seymour, D. S. Sarovich, J. Latham, S. R. Wolken, J. Mason, G. Vincent, K. P. Drees, S. M. Beckstrom-Sternberg, A. M. Phillippy, S. Koren, R. T. Okinaka, W. K. Chung, J. M. Schupp, D. M. Wagner, R. Vipond, J. T. Foster, N. H. Bergman, J. Burans, T. Pearson, T. Brooks, and P. Keim. 2012. Molecular epidemiologic investigation of an anthrax outbreak among heroin users, Europe. *Emerging Infectious Diseases* 18:1307-1313.
- Primorac, D., and M. Schanfield, eds. 2014. *Forensic DNA Analysis: An Interdisciplinary Perspective*. Boca Raton, FL: Taylor and Francis.
- Quail, M. A., M. Smith, P. Coupland, T. D. Otto, S. R. Harris, T. R. Connor, A. Bertoni, H. P. Swerdlow, and Y. Gu. 2012. A tale of three next generation sequencing platforms: Comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers. *BMC Genomics* 13:341.
- Rasko, D. A., P. L. Worsham, T. G. Abshire, S. T. Stanley, J. D. Bannan, M. R. Wilson, R. J. Langham, R. S. Decker, L. Jiang, T. D. Read, A. M. Phillippy, S. L. Salzberg, M. Pop, M. N. Van Ert, L. J. Kenefic, P. S. Keim, C. M. Fraser-Liggett, and J. Ravel. 2011. *Bacillus anthracis* comparative genome analysis in support of the Amerithrax investigation. *Proceedings of the National Academy of Sciences of the United States of America* 108(12):5027-5032.
- Read, T. D., S. L. Salzberg, M. Pop, M. Shumway, L. Umayam, L. Jiang, E. Holtzapple, J. D. Busch, K. L. Smith, J. M. Schupp, D. Solomon, P. Keim, and C. M. Fraser. 2002. Comparative genome sequencing for discovery of novel polymorphisms in *Bacillus anthracis*. *Science* 296(5575):2028-2033.
- Rehm, H. L., S. J. Bale, P. Bayrak-Toydemir, J. S. Berg, K. K. Brown, J. L. Deignan, M. J. Friez, B. H. Funke, M. R. Hegde, and E. Lyon, for the Working Group of the American College of Medical Genetics and Genomics Laboratory Quality Assurance Committee. 2013. ACMG clinical laboratory standards for next-generation sequencing. *Genetics in Medicine* 15:733-747.

- Rinke, C., P. Schwientek, A. Sczyrba, N. N. Ivanova, I. J. Anderson, J. Cheng, A. Darling, S. Malfatti, B. K. Swan, E. A. Gies, J. A. Dodsworth, B. P. Hedlund, G. Tsiamis, S. M. Sievert, W. Liu, J. A. Eisen, S. J. Hallam, N. C. Kyrpides, R. Stepanauskas, E. M. Rubin, P. Hugenholtz, and T. Woyke. 2013. Insights into the phylogeny and coding potential of microbial dark matter. *Nature* 499:431-437.
- Rushing, J. W. 1994. Food Safety in the United States and Abroad: An Agriculturalist's Perspective. Transcribed and edited from an oral presentation at the fall meeting of the Southern Agromedicine Consortium, October 6-7, 1994, Charleston, SC. Available at [http://books.google.com/books?id=StyK\\_YDFe0C&pg=PA112&lpg=PA112&dq=cyanide+contaminated+grapes+from+Chile&source=bl&ots=q5GYjmicdv&sig=p7gauXurZVAikmai-3oJ-peGe4&hl=en&sa=X&ei=yMI6U--dB8PgsAtN2ID4Dw&ved=0CEMQ6AEwBA%23v=onepage&q=cyanide%20contaminated%20grapes%20from%20Chile&f=false#v=onepage&q&f=false](http://books.google.com/books?id=StyK_YDFe0C&pg=PA112&lpg=PA112&dq=cyanide+contaminated+grapes+from+Chile&source=bl&ots=q5GYjmicdv&sig=p7gauXurZVAikmai-3oJ-peGe4&hl=en&sa=X&ei=yMI6U--dB8PgsAtN2ID4Dw&ved=0CEMQ6AEwBA%23v=onepage&q=cyanide%20contaminated%20grapes%20from%20Chile&f=false#v=onepage&q&f=false). Accessed April 10, 2014.
- Saksida, A., D. Duh, B. Wraber, I. Dedushaj, S. Ahmeti, and T. Avsic-Zupanc. 2010. Interacting roles of immune mechanisms and viral load in the pathogenesis of Crimean-Congo hemorrhagic fever. *Clinical Vaccine Immunology* 17:1086-1093.
- Sboner, A., J. Mu, D. Greenbaum, R. K. Auerbach, and M. B. Gerstein. 2011. The real cost of sequencing: Higher than you think! *Genome Biology* 12:125. Available at <http://genomebiology.com/2011/12/8/125>. Accessed April 10, 2014.
- Scaduto, D. I., J. M. Brown, W. C. Haaland, D. J. Zwickl, D. M. Hillis, and M. L. Metzker. 2010. Source identification in two criminal cases using phylogenetic analysis of HIV-1 DNA sequences. *Proceedings of the National Academy of Sciences of the United States of America* 107(50):21242-21247.
- Schaudies, R. P. 2014. State of the art for autonomous detection systems using immunoassays and protein signatures. Pp. 173-196 in Institute of Medicine and National Research Council, *Technologies to Enable Autonomous Detection for BioWatch: Ensuring Timely and Accurate Information for Public Health Officials: Workshop Summary*. Washington, DC: The National Academies Press.
- Skowronski, E., and W. I. Lipkin. 2011. Molecular microbial surveillance and discovery in bioforensics. Chapter 11 in *Microbial Forensics*, 2nd Ed., B. Budowle, S. E. Schutzer, R. G. Breeze, P. S. Keim, and S. A. Morse, eds. Burlington, MA: Academic Press.
- Smith, J. 2011. Collection and preservation of microbial forensic samples. Chapter 22 in *Microbial Forensics*, 2nd Ed., B. Budowle, S. E. Schutzer, R. G. Breeze, P. S. Keim, and S. A. Morse, eds. Burlington, MA: Academic Press.
- Snitkin, E. S., A. M. Zelazny, P. J. Thomas, F. Stock, D. K. Henderson, T. N. Palmore, and J. A. Segre. 2012. Tracking a hospital outbreak of carbapenem-resistant *Klebsiella pneumoniae* with whole-genome sequencing. *Science Translational Medicine* 4(148):148ra116.
- Snyder, P., and R. E. Jabbour. 2014. State of the art for autonomous detection systems using mass spectrometry. Pp. 215-244 in Institute of Medicine and National Research Council, *Technologies to Enable Autonomous Detection for BioWatch: Ensuring Timely and Accurate Information for Public Health Officials: Workshop Summary*. Washington, DC: The National Academies Press.
- Tadin, A., N. Turk, M. Korva, J. Margaletić, R. Beck, M. Vucelja, J. Habuš, P. Svoboda, T. Avšič Županc, H. Henttonen, and A. Markotić. 2012. Multiple co-infections of rodents with hantaviruses, *Leptospira*, and *Babesia* in Croatia. *Vector-Borne and Zoonotic Diseases* 12(5):388-392.
- Tadin, A., L. Bjedov, J. Margaletić, B. Žibrat, L. Cvetko Krajcinović, P. Svoboda, I. C. Kurolt, Z. Štritof, N. Turk, O. Đaković Rode, R. Čivljak, I. Kuzman, and A. Markotić. 2014. High infection rate of bank voles (*Myodes glareolus*) with Puumala virus is associated with a winter outbreak of haemorrhagic fever with renal syndrome in Croatia. *Epidemiology and Infection*; doi: 10.1017/S095026881300321X.

- Tucker, J. B., and G. D. Koblentz. 2009. The four faces of microbial forensics. *Biosecurity and Bioterrorism* 7(4):389-397.
- U.K. Foreign and Commonwealth Office. 2013. Global Partnership Against the Spread of Weapons and Materials of Mass Destruction: President's Report for 2013. London.
- United Nations. 2011. Disarmament Yearbook, Vol. 36 (Part II). Available at <http://www.un.org/disarmament/HomePage/ODAPublications/Yearbook/2011/DYB2011-Part-II-web.pdf>. Accessed April 10, 2014.
- United Nations Security Council. 2004. 1540 Committee Resolution adopted under Chapter VII of the UN Charter. Available at [http://www.un.org/en/ga/search/view\\_doc.asp?symbol=S/RES/1540%20%282004%29](http://www.un.org/en/ga/search/view_doc.asp?symbol=S/RES/1540%20%282004%29). Accessed October 24, 2013.
- U.S. Department of Energy. 2012. Genome Glossary. Available at <http://genomicscience.energy.gov/glossary/index.shtml>. Accessed April 10, 2014.
- U.S. Department of Justice. 2010. *Amerithrax Investigative Summary*. Available at <http://www.justice.gov/amerithrax/docs/amx-investigative-summary.pdf>. Accessed April 10, 2014.
- U.S. Environmental Protection Agency. 2004. Data evaluation. Chapter 5 in *Risk Assessment Guidance for Superfund, Volume I: Human Health Evaluation Manual*. EPA/540/R/99/005. Available at <http://www.epa.gov/oswer/riskassessment/ragsa/pdf/ch5.pdf>.
- Van Ert, M. N., W. R. Easterday, L. Y. Huynh, R. T. Okinaka, M. E. Hugh-Jones, J. Ravel, S. R. Zanecki, T. Pearson, T. S. Simonson, J. M. U'Ren, S. M. Kachur, R. R. Leadem-Dougherty, S. D. Rhoton, G. Zinser, J. Farlow, P. R. Coker, K. L. Smith, B. Wang, L. J. Kenefic, C. M. Fraser-Liggett, D. M. Wagner, and P. Keim. 2007. Global genetic population structure of *Bacillus anthracis*. *PLOS One* 2(5):e461.
- Vandamme, A. M., and O. G. Pybus. 2013. Viral phylogeny in court: The unusual case of the Valencian anesthetist. *BMC Biology* 11:83-85.
- Velsko, S. P. 2011a. Nonbiological measurements on biological agents. Chapter 30 in *Microbial Forensics*, 2nd Ed., B. Budowle, S. E. Schutzer, R. G. Breeze, P. S. Keim, and S. A. Morse, eds. Burlington, MA: Academic Press.
- Velsko, S. P. 2011b. Inferential validation and evidence interpretation. Chapter 33 in *Microbial Forensics*, 2nd Ed., B. Budowle, S. E. Schutzer, R. G. Breeze, P. S. Keim, and S. A. Morse, eds. Burlington, MA: Academic Press.
- Wahl, K. L., D. S. Wunschel, and B. H. Clowers. 2011. Proteomics development and application for bioforensics. Chapter 26 in *Microbial Forensics*, 2nd Ed., B. Budowle, S. E. Schutzer, R. G. Breeze, P. S. Keim, and S. A. Morse, eds. Burlington, MA: Academic Press.
- Wang, X., P. E. Kilgore, K. A. Lim, S. Wang, J. Lee, W. Deng, M. Mo, B. Nyambat, J. Ma, M. O. Favorov, and J. D. Clemens. 2011. Influenza and bacterial pathogen coinfections in the 20th century. *Interdisciplinary Perspectives on Infectious Diseases*. Available at <http://www.hindawi.com/journals/ipid/2011/146376/>. Accessed April 10, 2014.
- Welch, R. A., V. Burland, G. Plunkett, P. Redford, P. Roesch, D. Rasko, E. L. Buckles, S. R. Liou, A. Boutin, J. Hackett, D. Stroud, G. F. Mayhew, D. J. Rose, S. Zhou, D. C. Schwartz, N. T. Perna, H. L. Mobley, M. S. Donnenberg, and F. R. Blattner. 2002. Extensive mosaic structure revealed by the complete genome sequence of uropathogenic *Escherichia coli*. *Proceedings of the National Academy of Sciences of the United States of America*. Online. Available at: <http://www.ncbi.nlm.nih.gov/ubmed/12471157>. Accessed July 2, 2014.
- White House. 2012. National Strategy for Biosurveillance. Available at [http://www.whitehouse.gov/sites/default/files/National\\_Strategy\\_for\\_Biosurveillance\\_July\\_2012.pdf](http://www.whitehouse.gov/sites/default/files/National_Strategy_for_Biosurveillance_July_2012.pdf).
- White House. 2014. Global Health Security Agenda: Toward a World Safe and Secure from Infectious Disease Threats. Available at <http://www.globalhealth.gov/global-health-topics/global-health-security/GHS%20Agenda.pdf>.
- Wolfe, N. D., C. P. Dunavan, and J. Diamond. 2007. Origins of major human infectious diseases. *Nature* 447(7142):279-283.

- World Health Organization. 2014. Zoonoses. Available at <http://www.who.int/topics/zoonoses/en/>. Accessed April 10, 2014.
- Wright, G. D. 2007. The antibiotic resistome: The nexus of chemical and genetic diversity. *Nature Reviews Microbiology* 5(3):175-186.
- Wu, D., P. Hugenholtz, K. Mavromatis, R. Pukall, E. Dalin, N. N. Ivanova, V. Kunin, L. Goodwin, M. Wu, B. J. Tindall, S. D. Hooper, A. Pati, A. Lykidis, S. Spring, I. J. Anderson, P. D'haeseleer, A. Zemla, M. Singer, A. Lapidus, M. Nolan, A. Copeland, C. Han, F. Chen, J. F. Cheng, S. Lucas, C. Kerfeld, E. Lang, S. Gronow, P. Chain, D. Bruce, E. M. Rubin, N. C. Kyrpides, H. P. Klenk, and J. A. Eisen. 2009. A phylogeny-driven genomic encyclopaedia of Bacteria and Archaea. *Nature* 462(7276):1056-1060.



# Appendix A

## Committee Biographies

**John D. Clements** (*Chair*) is professor and chair of microbiology and immunology, Tulane University School of Medicine. After receiving his doctorate in 1979 from the University of Texas Health Science Center at Dallas, Dr. Clements completed a National Research Council Associateship at Walter Reed Army Institute of Research in Washington, DC. In 1980, Dr. Clements was appointed as assistant professor in the Departments of Microbiology and Medicine at the University of Rochester School of Medicine in Rochester, New York. In 1982, Dr. Clements joined the faculty at Tulane University, where he has served as professor and chair of the Department of Microbiology and Immunology since 1999. Dr. Clements served as Vice Dean for Research from 2006 to 2009, and in 2009 was appointed as Director of the Tulane Center for Infectious Diseases. Dr. Clements is currently Director of the Tulane/Xavier Vaccine Development/Engineering Project supported by the Department of Defense, co-director of the South Louisiana Institute for Infectious Disease Research, and co-director of the Louisiana Vaccine Center, both collaborative projects between Tulane University and Louisiana State University Health Sciences Center in New Orleans.

Research in Dr. Clements's laboratory has resulted in more than 100 peer-reviewed publications and book chapters and 13 issued patents. Dr. Clements currently serves on the Defense Health Board and the Scientific Advisory Boards of the Western Regional Center for Excellence in Bio-defense Research and the PATH Enteric Vaccine initiative. In 2002, Dr. Clements chaired the committee to review all military infectious disease



research programs for the Department of Defense. In 2003, and again in 2004, Dr. Clements served as a member of the Iraq Survey Group in Baghdad as a subject-matter expert in weapons of mass destruction and dual-use equipment and programs. Dr. Clements is a veteran of the U.S. Marine Corps and served on active duty from 1966 to 1972. He was honorably discharged at the rank of Lieutenant Colonel from the U.S. Marine Corps Reserves in 1991.

**Munirul Alam** is senior scientist, International Center for Diarrhoeal Diseases Research, Bangladesh (ICDDR,B). Dr. Alam joined the ICDDR,B in late 2003 as a consultant scientist and was appointed to be a staff scientist in 2007 and senior scientist in 2010. Dr. Alam is the principal investigator (PI) of the National Institutes of Health–funded multiyear cholera study, Epidemiology and Ecology of *V. cholerae* in Bangladesh, which is running on its third consecutive term at ICDDR,B under collaborative agreements with the Johns Hopkins Bloomberg School of Public Health. Dr. Alam has been the coordinator of the Diarrheal Disease Epidemiology and Ecology research group and actively involved in molecular epidemiological and ecological surveillance of cholera in Bangladesh, including monitoring of the plankton community dynamics and climate variables in relation to the seasonal epidemics with the aim of developing a predictive model for cholera and other infectious diseases. Dr. Alam also conducts a hospital-based cholera intervention study, and his research interests include safe drinking water, bacterial community dynamics in human gut and environment, microbial forensics, drug resistance of enteric pathogens, and metagenomics.

Dr. Alam holds bachelor of science and master of science degrees from the University of Dhaka. He started his career as a pharmaceutical microbiologist in 1989 and was appointed lecturer of the Department of Microbiology at the University of Dhaka in 1991. Dr. Alam won the Monbukagusho scholarship and obtained a Ph.D. in microbiology from Okayama University, Japan, in 1996. He was an assistant professor and later an associate professor of Dhaka University. He was on postdoctoral fellowships for different terms in institutions in the United States, Japan, and Germany. He has actively attended more than 50 national and international scientific conferences and served as the oral, invited, and keynote speaker. Dr. Alam also served as chair and cochair of scientific sessions. Dr. Alam has supervised more than 80 research students including M.S., M.Phil., and Ph.D. candidates, and has authored 88 research articles published in peer-reviewed journals including *Proceedings of the National Academy of Sciences* and *The Lancet*. He has been a member of many national and international professional societies including the American Society for Microbiology. For his outstanding contributions in the field of micro-

biology, Dr. Alam has been elected a Fellow of the American Academy of Microbiology (2013). He has been a faculty member of the Johns Hopkins Bloomberg School of Public Health since 2011, and a visiting fellow of the University of Technology Sydney, Australia (2011-2014). He has been a member of international scientific committees and on the editorial boards of many national and international peer-reviewed journals. Dr. Alam is the former President of the Graduate Microbiologists Association of Bangladesh.

**Bruce Budowle** is director, Institute of Applied Genetics, and professor, Department of Molecular and Medical Genetics, University of North Texas Health Science Center. Dr. Budowle received a Ph.D. in genetics in 1979 from Virginia Polytechnic Institute and State University. From 1979 to 1982, Dr. Budowle was a postdoctoral fellow at the University of Alabama at Birmingham. Working under a National Cancer Institute fellowship, he carried out research predominately on genetic risk factors for such diseases as insulin-dependent diabetes mellitus, melanoma, and acute lymphocytic leukemia.

In 1983, Dr. Budowle joined the research unit at the Federal Bureau of Investigation (FBI) to carry out research, development, and validation of methods for forensic biological analyses. The positions he has held at the FBI include research chemist, program manager for DNA research, chief of the Forensic Science Research Unit, and the senior scientist for the Laboratory Division of the FBI. Dr. Budowle has contributed to the fundamental sciences as they apply to forensics in analytical development, population genetics, statistical interpretation of evidence, and quality assurance. Some of the methods he developed are (1) analytical assays for typing a myriad of protein genetic marker systems; (2) designing electrophoretic instrumentation; (3) developing molecular biology analytical systems to include restriction fragment length polymorphism typing of variable number tandem repeat (VNTR) loci and polymerase chain reaction-based single-nucleotide polymorphism assays, VNTR and short tandem repeat assays, and direct sequencing methods for mitochondrial DNA; (4) new technologies; and (5) designing image analysis systems. Dr. Budowle has worked on laying some of the foundations for the current statistical analyses in forensic biology and defining the parameters of relevant population groups. Dr. Budowle has been directly involved in developing quality assurance (QA) standards for the forensic DNA field. He has been a member of the Scientific Working Group on DNA Methods, chair of the DNA Commission of the International Society of Forensic Genetics, and a member of the DNA Advisory Board. He was one of the architects of the CODIS National DNA database, which maintains DNA

profiles from convicted felons, from evidence in unsolved cases, and from missing persons.

Some of Dr. Budowle's efforts over the last decade are in counterterrorism, primarily in efforts involving microbial forensics and bioterrorism. Dr. Budowle is heavily involved in the forensic applications on bioterrorism and has been involved in developing the field known as microbial forensics. In the area of microbial forensics, Dr. Budowle has been the chair of the FBI's Scientific Working Group on Microbial Genetics and Forensics, whose mission was to set QA guidelines, develop criteria for biologic and user databases, set criteria for a national repository, and develop forensic genomic applications. He also has served on the Steering Committee for the Colloquium on Microbial Forensics sponsored by the American Society of Microbiology and was the organizer of three Microbial Forensics Meetings held at The Banbury Center in the Cold Spring Harbor Laboratory.

In 2009, Dr. Budowle became executive director of the Institute of Applied Genetics and professor in the Department of Forensic and Investigative Genetics at the University of North Texas Health Science Center at Fort Worth. His current efforts focus on the areas of human forensic identification, microbial forensics, and emerging infectious diseases. He is working on microbial forensics topics such as attribution, QA, population genetics, next-generation sequencing technology, and sample collection. Dr. Budowle's commitment to helping families resolve missing persons cases led him to Fort Worth after a lifetime in the Virginia/Washington, DC area in order to collaborate with Health Science Center researchers and advance the knowledge and use of forensics and DNA to improve health and safety of the world's population. Dr. Budowle has also been instrumental in establishing the DNA-ProKids initiative to identify missing children on an international scale.

**Jongsik Chun** is associate professor of biology, Seoul National University. Dr. Chun serves on Seoul National University's Interdisciplinary Program in Bioinformatics Steering Committee. Dr. Chun is associated with a number of institutes at Seoul National University, including the Institute of Microbiology Research, the Genetic Engineering Combustion Institute, and the International Vaccine Institute. He also assists the Rural Development Administration and the Institute of Agriculture and Life Sciences. He has previously served as a research associate at the Center of Marine Biotechnology at the University of Maryland and as senior researcher at the Korea Research Institute of Bioscience and Biotechnology.

Dr. Chun received his B.Sc. from Seoul National University and completed his Ph.D. at Newcastle University School of Medicine. Dr. Chun completed his postdoctoral work at the Research Center for Molecular

Microbiology at Seoul National University. Dr. Chun is the Associate Editor of the *International Journal of Systematic and Evolutionary Microbiology*. He serves as an editorial board member for *Antonie van Leeuwenhoek* (Dutch Kluwer four issues) and *Microbes and Environments*.

**Rita R. Colwell** is distinguished university professor, University of Maryland at College Park, and distinguished university professor, Johns Hopkins University Bloomberg School of Public Health. Dr. Colwell's interests are focused on global infectious diseases, water, and health, and she is currently developing an international network to address emerging infectious diseases and water issues, including safe drinking water for both the developed and developing world. Dr. Colwell has shown how changes in climate, adverse weather events, shifts in ocean circulation, and other ecological processes can create conditions that allow infectious diseases to spread. In addition to her academic roles, Dr. Colwell is senior adviser and chairperson of Canon U.S. Life Sciences, and chairman and president of CosmosID, which is exploring the potential applications of molecular diagnostic technologies to the field of life sciences.

Dr. Colwell served as the 11th director of the National Science Foundation from 1998 to 2004. She has previously served as chairman of the Board of Governors of the American Academy of Microbiology and also as president of the American Association for the Advancement of Science, the Washington Academy of Sciences, the American Society for Microbiology, the Sigma Xi National Science Honorary Society, and the International Union of Microbiological Societies. Dr. Colwell has also been awarded 54 honorary degrees from institutions of higher education, including her alma mater, Purdue University. Dr. Colwell holds a B.S. in bacteriology and an M.S. in genetics from Purdue University, and a Ph.D. in oceanography from the University of Washington. She is a member of the Royal Swedish Academy of Sciences, the American Academy of Arts and Sciences, and the American Philosophical Society. She is the recipient of the Order of the Rising Sun bestowed by the emperor of Japan and the National Medal of Science bestowed by the president of the United States. She is a U.S. science envoy and a member of the National Academy of Sciences.

**Nancy D. Connell** is professor of medicine, Rutgers New Jersey Medical School, Rutgers University, and director, Rutgers New Jersey Medical School Center for Biodefense. Dr. Connell is professor in the Division of Infectious Disease in the Department of Medicine at Rutgers New Jersey Medical School (RNJMS) and the Rutgers Biomedical Health Sciences. A Harvard University Ph.D. in microbiology, Dr. Connell's major research focus is antibacterial drug discovery in respiratory pathogens such as

*Mycobacterium tuberculosis* and *Bacillus anthracis*. She is director of the Biosafety Level 3 (BSL-3) Facility of RNJMS's Center for the Study of Emerging and Re-emerging Pathogens and chairs the university's Institutional Biosafety Committee. Dr. Connell has been continuously funded by the National Institutes of Health (NIH) and other agencies since 1993 and serves on numerous NIH study sections and review panels. She has served on a number of committees of the National Academy of Sciences, for example, the Committee on Advances in Technology and the Prevention of Their Application to Next Generation Biowarfare Agents (2004), Trends in Science and Technology Relevant to the Biological Weapons Convention; an International Workshop (2010); the Committee to Review the Scientific Approaches Used in the FBI's Investigation of the 2001 *Bacillus anthracis* Mailings (2011) and the Educational Institute for Responsible Science (Malaysia) (2013).

**Paul Keim** is Cowden endowed chair in microbiology and Arizona regents professor, Northern Arizona University. Dr. Keim is the director of Northern Arizona University's Microbial Genetics and Genomics Center. He also directs the Pathogen Genomics Division at the Translational Genomics Research Institute (TGen), a nonprofit research institute. He maintains his Laboratory Affiliate at Los Alamos National Laboratory in the Division of Biosciences. Dr. Keim's current research interests include genomic analysis of bacterial pathogens and the application of genomic technology to clinical diagnostic problems.

Dr. Keim's laboratory has developed high-resolution strain typing analysis methods for the forensic analysis of *Bacillus anthracis*, *Yersinia pestis*, and *Francisella tularensis*. He has participated in collaborative projects with scientists from the former Soviet Union to understand the ecology and epidemiology of these pathogens.

Dr. Keim has served on grant review panels for the U.S. Department of Agriculture and the National Institutes of Health; on advisory groups for the Federal Bureau of Investigation (FBI), the U.S. Government Accountability Office, and the U.S. Department of Health and Human Services; and on three previous National Research Council committees. He received a B.S. in biology and chemistry from Northern Arizona University and a Ph.D. in botany from the University of Kansas. He has done postdoctoral work in genetics, genomics, and biotechnology. He is currently a member of the FBI's Scientific Working Group on Forensic Analysis of Chemical, Biological, Radiological and Nuclear Terrorism, the National Science Advisory Board for Biodefense, and the Executive Advisory Committee for the Pacific Southwest Regional Center for Biodefense. He is a fellow of the American Academy of Microbiology.

**Juncai Ma** is assistant director of the Institute of Microbiology at the Chinese Academy of Sciences (CAS), and deputy chairman of the Expert Committee on CAS Databases. Dr. Ma is the director of the Committee on Type Culture Collection CAS and commissioner of the CODATA Chinese National Committee, and executive of the World Federation for Culture Collections. Currently he is mainly engaged in the research work on bio-grid, parallel indexing, super-large-scaled full-text retrieval technology, a search engine of remote heterogeneous databases, Linux Cluster System, and comprehensive utilization of information technology in the field of biology. Meanwhile, he is in charge of the implementation of a variety of projects, including the China Microbial Resource Database, the Information Network System of CAS Biology Specimen Museum, the Microbial Information Gateway of National Scientific Digital Library, the E-Science Bio-Grid, the National Scientific Data Sharing Platform, as well as the Information Network of Chinese Biotechnology and Industry. In 2006, he received a Ph.D. from the Biological Resource Department of Mie University, Japan.

**Alemka Markotić** is professor at the Medical School of the University of Rijeka, associate member of the Croatian Academy of Sciences and Arts, and lecturer on bioterrorism and biodefense at the Forensic Sciences Study at the University of Split, Croatia. Dr. Markotić is head of the Research Department and head of the Department for Clinical Immunology, University Hospital for Infectious Diseases (UHID) in Zagreb, Croatia. She received her M.D. at the University of Sarajevo, Bosnia and Herzegovina (1989), an M.S. in medical microbiology and parasitology (1996), and a Ph.D. in infectious diseases (1999) from the University of Zagreb Medical School.

Dr. Markotić began her biomedical research career at the University of Sarajevo Medical School in Bosnia and Herzegovina studying ribavirin treatment of hantaviruses during which time she collaborated with the U.S. Army Medical Research Institute for Medical Diseases (USAMRIID) in Frederick, Maryland. She later received a National Academy of Sciences, National Research Council Postdoctoral Fellowship at USAMRIID to conduct research on the immunopathogenesis of hantaviruses. On the basis of this work, she received the Joel Dalymple Memorial Award (American Society of Virology) and the USAMRIID Coin.

Dr. Markotić's research on hantaviruses has earned her seven national and nine international awards. She has published more than 80 peer-reviewed papers, 14 book chapters and delivered over 90 presentations at national and international conferences. She has been the principal investigator on research projects studying immune responses to intracellular pathogens, zoonoses, and apoptosis in hantavirus-infected 293HEK



cells. At the UHID, Dr. Markotić established the Center for Emerging and Re-emerging Diseases with nine international and six national partners, and she is responsible for managing the first Croatian BSL-3 laboratory at the UHID. At the request of the European Union (EU) Commission, Dr. Markotić designed, organized, and presented a biosafety/biosecurity training workshop in Beijing, China, in May 2009. She worked for several years at the Institute of Immunology in Zagreb as a head of the Viral Vaccines and Interferon Quality Control Unit.

She is a member of the Council of the International Society for Hantaviruses and the Board for Allergy and Clinical Immunology at the Croatian Academy of Sciences and Arts, was a member of the Committee of the Croatian Science Foundation and the National Council for Science, and was vice president of the Scientific Council in the Scope of Biomedicine and Health. In 2004, 2005, and 2009, she was an expert evaluator for FP6 projects (EU Commission, Brussels) in immunology and emerging infectious diseases. In the 1990s, during the war in Bosnia and Herzegovina, she helped organize the Caritas Pharmacy and Health Care Unit that addressed the health needs of those affected by the conflict.

**Geoffrey Smith** is professor of pathology and Wellcome Trust Principal Research Fellow, Emmanuel College, Cambridge University. Dr. Smith is a British virologist and medical research authority in the area of Vaccinia virus and the family of Poxviruses. Dr. Smith completed his bachelor's degree at the University of Leeds in 1977, and in 1981 gained a Ph.D. in virology while in London. Between 1981 and 1984, while he was working in the United States at the National Institutes of Health, Dr. Smith developed and pioneered the use of genetically engineered live vaccines. Between 1985 and 1989, he lectured at the University of Cambridge.

Prior to 2002, he was based at the Sir William Dunn School of Pathology at the University of Oxford. Between 1988 and 1992, his work was funded by the Jenner Fellowship from the Lister Institute; he became a governor of the institute in 2003. In 1992 the Society for General Microbiology awarded Dr. Smith its Fleming Award for outstanding work by a young microbiologist. In 2002, he was elected as a fellow of the Academy of Medical Sciences. In 2003, he was invited to become a fellow of the Royal Society and in 2005 was awarded the Feldburg Foundation Prize for his work on poxviruses. Smith was editor-in-chief of the *Journal of General Virology* until 2008 and chairs the World Health Organization's Advisory Committee on Variola Virus Research. As of 2009, he remains head of the Department of Virology at Imperial College London and is president of the International Union of Microbiological Societies.

## Appendix B

### Convening Organizations

#### THE CROATIAN ACADEMY OF SCIENCES AND ARTS

Founded in 1861, the Croatian Academy of Sciences and Arts is the highest scientific and artistic institution in the Republic of Croatia.

The founder of the Croatian Academy of Sciences and Arts was Josip Juraj Strossmayer, bishop of Đakovo and Srijem, whose proposal to found the Academy of Sciences was unanimously approved by the Croatian Parliament in April 1861. Bishop Strossmayer was elected the patron of the Academy and the historian Canon Franjo Rački the first president. This national academy of sciences and arts took the name *Accademia Slavorum Meridionalium* (of the South Slavs).

The Academy promotes and organizes scientific research and encourages the application of the findings of this research, develops artistic and cultural activities, and is concerned with Croatian cultural heritage and its affirmation throughout the world. It also publishes the results of scientific research and artistic creation and makes proposals and gives its opinion on the promotion of sciences and arts in the fields that are of special importance to the Republic of Croatia.

The Croatian Academy's scientific and artistic activities are carried out through its nine departments (I—Department of Social Sciences, II—Department of Mathematical, Physical and Chemical Sciences, III—Department of Natural Sciences, IV—Department of Medical Sciences, V—Department of Philological Sciences, VI—Department of Literature, VII—Department of Fine Arts, VIII—Department of Music and Musicology, and IX—Department of Technical Sciences), as well as through its



scientific councils and committees. The research is performed through the scientific and research units (institutes) of the Croatian Academy in Zagreb and other Croatian towns. The Croatian Academy of Sciences and Arts collaborates with other academies of sciences and arts, international scientific organizations, universities, scientific institutions, state bodies, cultural and other institutions, as well as with individual scholars and artists from Croatia and abroad.

*The main bodies of the Academy are the Assembly, which includes all full members of the Academy, the Presidency, the executive body of the Assembly, which consists of the Management Board members, secretaries of the departments, and five full members of the Academy.*

The Academy consists of honorary, full, corresponding, and associate members. Full members reserve the right to bear the title of Fellow of the Croatian Academy (F.C.A.), and they are part of the permanent working structure of the Academy.

The Croatian Academy may elect as honorary members, persons who are exceptionally meritorious for the development and progress of sciences and arts. Among its deceased members, we should mention the Nobel Laureates Lavoslav Ružička, Vladimir Prelog, Ivo Andrić, and Linus Pauling; Nikola Tesla, world famous scientist and inventor; Vlaho Bukovac, famous Croatian painter; Ivan Meštrović, world famous sculptor; and Andrija Štampar, founder of the World Health Organization; Dmitrij Ivanovič Mendeljejev; and numerous other world-renowned scientists and artists.

The management board of the Academy for the period from 2011 to 2014 consists of the President, Professor Zvonko Kusić; two vice-presidents, Professor Velimir Neidhardt and Professor Jakša Barbić; Secretary-General, Professor Pavao Rudan; and Secretary, Marina Štancl.

More information is available at <http://info.hazu.hr/>.

### THE INTERNATIONAL UNION OF MICROBIOLOGICAL SOCIETIES (IUMS)

The Union is one of the 31 Scientific Unions of the International Council of Science (ICSU). It was founded in 1927 as the International Society of Microbiology, and became the International Association of Microbiological Societies affiliated with the International Union of Biological Sciences (IUBS) as a Division in 1967. It acquired independence in 1980 and became a Union Member of ICSU in 1982.

The objectives of the Union are to promote the study of microbiological sciences internationally: initiate, facilitate, and coordinate research and other scientific activities that involve international cooperation; ensure the discussion and dissemination of the results of international conferences,

symposia, and meetings and assist in the publication of their reports; represent microbiological sciences in ICSU; and maintain contact with other international organizations. The major goal of IUMS is to promote research and the open exchange of scientific information for advancement of the health and welfare of humankind and the environment and it strongly discourages any uses of knowledge and resources to the contrary. In particular, the IUMS strives to promote ethical conduct of research and training in the areas of biosecurity and biosafety so as to prevent use of microorganisms as biological weapons and therefore to protect the public's health and to promote world peace. IUMS seeks that all its member societies adopt or develop a Code of Ethics to prevent misuse of scientific knowledge and resources. The scientific activities of the Union are conducted by the three Divisions of Bacteriology & Applied Microbiology (BAM), Mycology, and Virology and by six specialist international committees, eight international commissions, and two international federations. Their major activities include the classification and nomenclature of bacteria, fungi, and viruses; food microbiology; medical microbiology and diagnostics; culture collections; education; and biological standardization. The Divisions are responsible for the organization of their International Congresses (International Congress of Bacteriology and Applied Microbiology, International Congress of Mycology, and International Congress of Virology), and the committees, commissions, and federations organize their own meetings.

More information is available at <http://www.iums.org/>.

### THE ROYAL SOCIETY

The Royal Society is a self-governing fellowship of many of the world's most distinguished scientists drawn from all areas of science, engineering, and medicine.

The Society's fundamental purpose, reflected in its founding charters of the 1660s, is to recognize, promote, and support excellence in science and to encourage the development and use of science for the benefit of humanity.

The Society has played a part in some of the most fundamental, significant, and life-changing discoveries in scientific history, and Royal Society scientists continue to make outstanding contributions to science in many research areas.

The Royal Society is the national academy of science in the United Kingdom, and its core is its Fellowship and Foreign Membership, supported by a dedicated staff in London and elsewhere. The Fellowship comprises the most eminent scientists of the United Kingdom, Ireland,

and the Commonwealth. Its current president is Sir Paul Nurse, a geneticist who was awarded the Nobel Prize for Physiology or Medicine in 2001.

A major activity of the Society is identifying and supporting the work of outstanding scientists. The Society supports researchers through its early and senior career schemes, innovation and industry schemes, and other schemes.

The Society facilitates interaction and communication among scientists via its discussion meetings, and disseminates scientific advances through its journals. The Society also engages beyond the research community, through independent policy work, the promotion of high-quality science education and communication with the public.

The Royal Society's Science Policy Centre provides independent, timely, and authoritative scientific advice to U.K., European, and international decision makers. It champions the contribution that science and innovation can make to economic prosperity, quality of life, and environmental sustainability and is a hub for debate about science, society, and public policy.

More information is available at [www.royalsociety.org](http://www.royalsociety.org).

### THE U.S. NATIONAL ACADEMIES

The National Academies of the United States comprise four organizations: the National Academy of Sciences, the National Academy of Engineering, the Institute of Medicine, and the National Research Council.

The National Academy of Sciences is a private, nonprofit, self-perpetuating society of distinguished scholars engaged in scientific and engineering research, dedicated to the furtherance of science and technology and to their use for the general welfare. Upon the authority of the charter granted to it by the Congress in 1863, the Academy has a mandate that requires it to advise the federal government on scientific and technical matters. Dr. Ralph J. Cicerone is president of the National Academy of Sciences.

The National Academy of Engineering was established in 1964, under the charter of the National Academy of Sciences, as a parallel organization of outstanding engineers. It is autonomous in its administration and in the selection of its members, sharing with the National Academy of Sciences the responsibility for advising the federal government. The National Academy of Engineering also sponsors engineering programs aimed at meeting national needs, encourages education and research, and recognizes the superior achievements of engineers. Dr. C. D. Mote, Jr. is president of the National Academy of Engineering.

The Institute of Medicine was established in 1970 by the National Academy of Sciences to secure the services of eminent members of appro-

priate professions in the examination of policy matters pertaining to the health of the public. The Institute acts under the responsibility given to the National Academy of Sciences by its congressional charter to be an adviser to the federal government and, upon its own initiative, to identify issues of medical care, research, and education. Dr. Victor J. Dzau is president of the Institute of Medicine.

The National Research Council was organized by the National Academy of Sciences in 1916 to associate the broad community of science and technology with the Academy's purposes of furthering knowledge and advising the federal government. Functioning in accordance with general policies determined by the Academy, the Council has become the principal operating agency of both the National Academy of Sciences and the National Academy of Engineering in providing services to the government, the public, and the scientific and engineering communities. The Council is administered jointly by both Academies and the Institute of Medicine. Dr. Cicerone and Dr. Mote are chair and vice chair, respectively, of the National Research Council.

More information is available at <http://www.national-academies.org>.



# Appendix C

## Agenda

13-16 October 2013  
Croatian Academy of Sciences and Arts

### SUNDAY, 13 OCTOBER

17:30      **Registration**

18:00–19:30   **Welcome Reception** (Palace Hotel)

Welcome from sponsoring organizations

- Croatian Academy of Sciences and Arts—
  - Zvonko Kusić, FCA, President
  - Dragan Dekaris, FCA, member, Department of Medical Sciences
- U. K. Royal Society and International Union of Microbiological Societies—Elinor Buxton for Geoffrey Smith
- U.S. National Academy of Sciences—John Clements

### MONDAY, 14 OCTOBER

8:15      **Registration**, Croatian Academy of Sciences and Arts  
\*\*All plenary sessions at Croatian Academy of Sciences and Arts

9:00

**Opening Plenary**

Chair: John Clements

Welcome/Meeting Goals: John Clements

Introduction: “The Emerging Field of Microbial Forensics”:  
Randall Murch

- Why is there a need for microbial forensics and why is it important? What is the current state of the art and today’s cutting-edge techniques and infrastructure? How do the forensics used for criminal investigations differ from epidemiological investigations for public health?
- What are the major research challenges for the field? How can basic science be used to solve the current challenges for microbial forensics and how might this help in other areas, such as public health?

Comments:

- Piers Millet
- Mats Forsman

Q&amp;A

10:30

**Break**

11:00

**Plenary: Microbial Ecology and Diversity—Microbial Forensics in the Context of Population Genetics**

Chair: Paul Keim

- What is known, in general, about the ecology of pathogens globally—Ruifu Yang
- Microbial ecology and diversity in the context of microbial forensics—Aaron Darling

Q&amp;A

12:00

**Lunch (Palace Hotel)****Lunch Speaker**

Chair: Geoffrey Smith

Talk: “Croatian Accomplishments in Forensics Genetics”—  
Dragan Primorac

**13:30 Plenary: Clinical and Forensic Approaches to Microbial Identification**

Chair: Dragan Primorac

- Technologies and techniques for forensics—Dana Kadavy
- Clinical diagnostic practices—Alemka Markotić

Q&A

**14:45 Plenary: Technologies and Approaches for Identifying Microbes for Law Enforcement—The 2001 Anthrax Letters**

After a talk about the anthrax letters case, a panel of experts will make short comments to amplify specific aspects or perspectives, offer lessons from other cases.

Chair: Gilles Vergnaud

Speaker: Paul Keim

Comments:

- Anthrax-contaminated heroin in Scotland and Germany—Richard Vipond
- Hepatitis C case in Spain—Fernando González-Candelas
- An international perspective: Inspections for biological weapons capabilities—Rocco Casagrande
- An international perspective: UN Security Council Resolution 1540—Dana Perkins

Q&A

**16:15 Poster Session and Reception, Croatian Academy of Sciences and Arts**

This session is intended to allow participants to discuss their research in an informal setting.

- Welcome from Pavao Rudan, FCA, Secretary General of the Croatian Academy of Sciences and Arts



**TUESDAY, 15 OCTOBER**

- 9:00**            **Plenary: Technologies and Approaches for Identifying Microbes in Public Health—The *E. coli* O104 Case**  
After a talk about the *E. coli* O104 case, a panel of experts will make short comments to offer lessons from other cases and amplify specific aspects or other perspectives.  
Chair: Munirul Alam  
Speaker: Dag Harmsen  
Comments:  
• Haruo Watanabe  
• Stephen Morse  
• Raymond Lin
- Q&A
- 10:30**            **Break**
- 11:00**            **Plenary: Sampling and Preservation Methods**  
A talk will address key issues, such as public health versus criminal investigation, transportation and storage, accepted protocols and practices, and the more general question of whether there is a need for standardized methods that are shared internationally. A panel of experts will make comments at points during the talk.  
Chair: Bruce Budowle  
Speaker: Adam Hamilton  
Comments:  
• Stephen Morse  
• Cerys Rees
- Q&A
- 12:15**            **Lunch (Palace Hotel)**
- Lunch Speaker**  
Chair: Herawati Sudoyo  
Talk: “The Importance of Reference Collections and the Role of the World Data Center for Microorganisms”—Juncai Ma

- 13:45**      **Plenary: Validation and Reference Materials for Microbial Forensics**  
A talk will address key issues, such as guidelines and components for validation, transportation and storage, test materials, and whether there is a need for internationally accepted standards for validation.  
Chair: Cindi Corbett  
Speaker: Bruce Budowle
- Q&A
- 14:45**      **Plenary: Bioinformatics and Data**  
Chair: Habib Bukhari
- Role and importance of bioinformatics and computational genomics in microbial forensics—Jongsik Chun
  - Managing large datasets—Aaron Darling
- Q&A
- 16:00**      **Break**
- 16:30**      **Breakout sessions**  
These sessions will address questions intended to set the stage for the concluding sessions on the final day, such as:
- If resources were not an issue, in what kinds of science should we invest?
  - Given that there are resource constraints, what critical investments should receive immediate priority?
- 17:30**      **Adjourn**
- 17:35**      **Visit to the Strossmayer Gallery of Old Masters of the Croatian Academy of Sciences and Arts**
- 19:30**      **Conference Dinner**

**WEDNESDAY, 16 OCTOBER**

- 9:00**            **Plenary: Reporting Back from the Breakout Sessions**  
Brief summaries of key points from the previous day  
Chair: John Clements
- Q&A
- 9:45**            **Breakout Sessions**  
These sessions will refine and develop further the ideas that emerged from the first breakout sessions and initial plenary discussion.
- 10:45**          **Break**
- 11:15**          **Concluding Plenary**  
Chair: John Clements  
Brief summaries of key points from the morning sessions  
General discussion  
Consideration of other issues, such as
- Starting the dialogue toward development of international standards
  - Creating opportunities for education and training
  - Assessing viability of an international body (e.g., UN Secretary General’s Investigation Mechanism) to provide first response and investigative assistance
- 12:30**          **Adjourn**

## Appendix D

### List of Participants

Munirul Alam  
International Center for Diarrheal  
Diseases Research  
Bangladesh

Anna Bielecka  
The General K. Kaczkowski  
Military Institute of Hygiene  
& Epidemiology  
Poland

Bruce Budowle  
University of North Texas Health  
Science Center  
United States

Habib Bukari  
COMSATS Institute of Information  
Technology  
Pakistan

Suzana Bukovski  
University Hospital for Infectious  
Diseases “Fran Mihaljevic,”  
Zagreb, and Medical School  
University of Osijek  
Croatia

Elinor Buxton  
The Royal Society  
United Kingdom

Rocco Casagrande  
Gryphon Scientific  
United States

Jongsik Chun  
Seoul National University  
Republic of Korea

John D. Clements  
Tulane University  
United States

Nancy Connell  
Rutgers University  
United States

Adam Hamilton  
Signature Science, LLC  
United States

Cindi Corbett  
Public Health Agency of Canada  
National Microbiology  
Laboratory  
Canada

Dag Harmsen  
University Hospital Münster  
Germany

Aaron Darling  
University of Technology Sydney  
Australia

Jo Husbands  
The National Academies  
United States

Peter Dees  
U.S. Department of State  
United States

Ninja Ivanus  
Croatian Academy of Sciences and  
Arts  
Croatia

Dragan Dekaris  
Croatian Academy of Sciences and  
Arts  
Croatia

Franca Jones  
United States

Jelena Dukic  
Croatian Academy of Sciences and  
Arts  
Croatia

Loic Josseran  
Université de Versailles Saint  
Quentin  
France

Meg Flanagan  
U.S. Department of State  
United States

Dana Kadavy  
Signature Science, LLC  
United States

Mats Forsman  
FOI, Swedish Defence Research  
Agency  
Sweden

Indrani Karunasagar  
Karnataka Veterinary, Animal and  
Fisheries Sciences University  
India

Fernando González Candelas  
Universidad de Valencia  
Spain

Paul Keim  
Northern Arizona University  
United States

Oleg I. Kiselev  
Research Institute of Experimental  
Medicine  
Russia

Jens H. Kuhn  
Integrated Research Facility at  
Fort Detrick  
NIH/NIAID/DCR  
United States

Zvonko Kusić  
Croatian Academy of Sciences and  
Arts  
Croatia

Raymond Lin Tzer Pin  
National Public Health Laboratory  
Singapore

Juncai Ma  
Institute of Microbiology  
Chinese Academy of Sciences  
China

Alemka Markotić  
University Hospital for Infectious  
Diseases “Fran Mihaljevic,”  
Zagreb and Medical School  
University of Rijeka  
Croatia

Carl N. Mayers  
Defence Science and Technology  
Laboratory  
Porton Down  
United Kingdom

Lorna Miller  
Defence Science and Technology  
Laboratory  
Porton Down  
United Kingdom

Piers Millet  
BWC Implementation Support  
Unit  
Geneva, Switzerland

Stephen Morse  
United States

Randall Murch  
VA Polytechnic Institute and State  
University  
United States

Zarko Nozica  
Polytechnics Zagreb  
Croatia

Marko Pećina  
Croatian Academy of Sciences and  
Arts  
Croatia

Dana Perkins  
Committee Established Pursuant  
to Security Council Resolution  
1540  
United Nations  
United States

Christine Pourcel  
Institut de Génétique et  
Microbiologie  
UFR des Sciences—Université  
Paris-Sud 11  
France

Dragan Primorac  
The Pennsylvania State University  
and University of New Haven,  
United States  
University of Split and University  
of Osijek, Croatia

Cerys Rees  
Defence Science and Technology  
Laboratory  
Porton Down  
United Kingdom

Gi-eun Rhie  
Korea Centers for Disease Control  
and Prevention (KCDC)  
Republic of Korea

Pavao Rudan  
Croatian Academy of Sciences and  
Arts  
Croatia

Ben Rusek  
The National Academies  
United States

Marina Santic  
Medical School University of  
Rijeka  
Croatia

Frances Sharples  
The National Academies  
United States

Herawati Sudoyo  
Eijkman Institute for Molecular  
Biology and Indonesian  
Academy of Sciences  
Indonesia

Ante Tadin  
University Hospital for Infectious  
Diseases, Zagreb  
Croatia

Gabriel Trueba  
Universidad San Francisco de  
Quito  
Ecuador

Gilles Vergnaud  
Institut de Génétique et  
Microbiologie  
UFR des Sciences—Université  
Paris-Sud 11  
France

Adriana Vince  
University Hospital for Infectious  
Diseases, Zagreb, and Medical  
School University of Zagreb  
Croatia

Richard Vipond  
Public Health England  
United Kingdom

Natalia Vynograd  
National Medical University  
Ukraine

Haruo Watanabe  
National Institute of Infectious  
Diseases  
Japan

Kristin White  
The National Academies  
United States

Ruifu Yang  
Beijing Institute of Microbiology  
and Epidemiology  
China

# Appendix E

## List of Presentations

**Bruce Budowle**, Director, Institute of Applied Genetics and Professor, Department of Molecular and Medical Genetics, University of North Texas Health Science Center

- Sampling and Preservation: Commentary Using Food and Agriculture as Examples of Targets for an Attack with a Bioweapon
- Validation and Reference Materials for Microbial Forensics

**Rocco Casagrande**, Founder and Managing Director, Gryphon Scientific

- Technologies and Approaches for Identifying Microbes for Law Enforcement: Tools Needed to Support Biological Disarmament

**Jongsik Chun**, Associate Professor of Biology, Seoul National University

- Bioinformatics Challenges for Microbial Forensics

**Aaron Darling**, Associate Professor in Computational Genomics and Bioinformatics, University of Technology Sydney Faculty of Science's ithree institute

- Big Data and Computing Challenges in Microbial Forensics
- Microbial Ecology and Diversity in the Context of Forensics

**Mats Forsman**, Research Director for Biological Analyses, Swedish Defense Research Agency (FOI)

- Commentary: Microbial Forensics—A Swedish Perspective



**Fernando González-Candelas**, Full Professor of Genetics, Department of Genetics, University of Valencia

- Molecular Evolution in Court: Analysis of a Large Hepatitis C Virus Outbreak from an Evolving Source

**Adam Hamilton**, President and CEO, Signature Science

- Sampling and Handling for Microbial Forensics Applications

**Dag Harmsen**, Head of Research, Periodontology Department, University Hospital Münster

- The *E. coli* O104 Case

**Dana R. Kadavy**, Senior Microbiologist, Signature Science

- Forensic Approaches to Microbial Identification

**Paul Keim**, Cowden Endowed Chair in Microbiology and Arizona Regents Professor, Northern Arizona University

- The FBI Amerithrax Investigation

**Raymond Lin Tzer Pin**, Head and Senior Consultant, Division of Microbiology; Clinical Director, Molecular Diagnostic Centre

- Commentary: Technologies and Approaches for Identifying Microbes in Public Health

**Juncai Ma**, Assistant Director of Institute of Microbiology at the Chinese Academy of Sciences (CAS); Deputy Chairman of the Expert Committee on CAS Databases

- The Importance of Reference Collections and the Role of the World Data Center for Microorganisms

**Alemka Markotić**, Head, Department for Research, University Hospital for Infectious Diseases, Zagreb, Croatia; Professor at the Medical School of the University of Rijeka and Associate Member of the Croatian Academy of Sciences and Arts

- Clinical and Forensic Approaches to Microbial Identification: Clinical Diagnostic Practices

**Piers Millet**, Deputy Head of the Implementation Support Unit, Biological Weapons Convention, United Nations Office for Disarmament Affairs

- Commentary: The Emerging Field of Microbial Forensics

**Stephen A. Morse**, Associate Director for Environmental Microbiology, National Center for Emerging and Zoonotic Infectious Diseases, Centers for Disease Control and Prevention, Atlanta, Georgia

- Sampling and Preservation Methods: Public Health Aspects
- Technologies and Approaches for Identifying Microbes in Public Health

**Randall Murch**, Professor in Practice, Virginia Polytechnic Institute and State University, Visiting Professor, Department of War Studies, King's College London, UK

- The Trajectory of Microbial Forensics: From Origins to “Grand Challenges”

**Dana Perkins**, Expert, Group of Experts, Committee established pursuant to United Nations Security Council Resolution 1540

- Using Microbial Forensics to Strengthen Biosecurity and the Implementation of UN Security Council Resolution 1540

**Dragan Primorac**, Adjunct Professor, Eberly College of Science, The Pennsylvania State University, and Henry C. Lee College of Criminal Justice and Forensic Sciences, University of New Haven; Professor, Medical Schools, Split and Osijek, Croatia

- Croatian Accomplishments in Forensics Genetics

**Cerys Rees**, Capability Lead, CB Analysis & Attribution, Detection Department, Dstl Porton Down

- Commentary: Sampling and Preservation Methods

**Richard Vipond**, Operations Manager for the Rare and Imported Pathogens Laboratory (RIPL), Public Health England, Porton Down

- Heroin-Associated Anthrax Cases 2009/10, 2012 and . . .

**Haruo Watanabe**, Director General, National Institute of Infectious Diseases, Tokyo, Japan

- Commentary: Technologies and Approaches for Identifying Microbes in Public Health

**Ruifu Yang**, Beijing Institute of Microbiology and Epidemiology

- What Is Known, in General, About the Ecology of Pathogens Globally



## Appendix F

### Speaker Biographies

**Bruce Budowle** (see Appendix A)

**Rocco Casagrande** is founder and managing director, Gryphon Scientific. Dr. Casagrande holds a B.A. in chemistry and biology from Cornell University and a Ph.D. in experimental biology from the Massachusetts Institute of Technology. From December 2002 to March 2003, Dr. Casagrande worked as a United Nations biological weapons inspector in Iraq, where he participated in more than 50 inspections and acted as the chief of the U.N. biological analysis laboratory. In a prior post, Dr. Casagrande led a team of biologists and engineers at Surface Logix, a Boston-based biotechnology firm, to develop and test real-time detectors for biological agents. In his most recent position, he led the homeland security practice at Abt Associates, one of the nation's largest public policy research and consulting firms. He has published numerous articles on biological defense and has consulted on chemical and biological warfare and defense for several congressional offices and government agencies.

**Jongsik Chun** (see Appendix A)

**Aaron Darling** is associate professor in computational genomics and bioinformatics, University of Technology Sydney Faculty of Science's ithree institute. Dr. Darling has over a decade of experience developing computational methods for comparative genomics and evolutionary modeling, and in 2013, moved from the University of California, Davis

(UC Davis) to start a computational genomics group at the University of Technology Sydney.

Dr. Darling embarked on his research career at the University of Wisconsin–Madison (UW-Madison). Following a bachelor's degree in computer science, he worked with members of the UW-Madison Genome Center to sequence and analyze the first genomes of pathogenic *Escherichia coli*. During this time, Dr. Darling led the development of some widely used computational methods for analyzing genomic data, including the mpiBLAST open-source parallel BLAST software and the Mauve software for comparing multiple genome sequences. Following the award of a Ph.D. at UW-Madison, Dr. Darling received a fellowship from the U.S. National Science Foundation to pursue postdoctoral studies at the University of Queensland (UQ). After 2 years at UQ, he returned to UC Davis to develop a research program in computational metagenomics—the study of uncultivated microorganisms from the environment using computational methods. Dr. Darling now brings his experience to understand the relationship between humans and microorganisms in collaboration with microbiologists at the ithree institute.

**Mats Forsman** is research director for biological analyses, Swedish Defense Research Agency (FOI). After finishing his Ph.D. in microbiology, he joined the Swedish Defence Research Agency (FOI), Division for CBRN Defence and Security. In 2002, he was appointed associate professor, Department of Cell and Molecular Biology, Umeå University. Dr. Forsman has represented Sweden for more than 10 years in various Economic Development Administration and NATO research projects and working groups concerning various biodefense issues. He is currently research director and group leader at FOI for biological analyses, a group involved in both research and applied activities in the field of molecular diagnostics, bioinformatics, ecology, and epidemiology of pathogenic bacteria.

**Fernando González-Candelas** is full professor of genetics, Department of Genetics, University of Valencia. Dr. González-Candelas' main research interests are in population and evolutionary genetics, molecular and evolutionary epidemiology, molecular systematics and genomics, bioinformatics, and conservation biology. Dr. González-Candelas is currently working on the molecular evolutionary epidemiology of different pathogens, mainly RNA viruses, such as hepatitis C virus (HCV) and human immunodeficiency virus (HIV), and bacteria, such as *Legionella pneumophila*. The basic approach is the analysis of nucleotide sequence variability at different levels, from inpatient to worldwide samples, depending on the specific goals of the different projects.

He has previously studied the population and evolutionary biol-

ogy of Mediterranean endemic *Limonium* (Plumbaginaceae) species, by using an array of genetic markers (random amplified polymorphic DNAs, amplified fragment length polymorphisms, microsatellites, isozymes) and including the analysis of quantitative traits. In addition, he has been very involved in the establishment of the Bioinformatics Service at the University of Valencia, of which he served as head from 1994 to 2000. He has also helped in the development of the Automated Sequencing Service. Dr. González-Candelas also served as scientific expert on the Spanish National Biosafety Council from 1998 until 2010.

**Adam Hamilton** is president and CEO, Signature Science. Mr. Hamilton's background includes experience in engineering (B.S., M.S.), science (chemistry and microbiology), training, statistics, and quality assurance. He is a registered professional engineer (1993-present) and has been heavily involved in applications of science and engineering for public safety and homeland security for more than 25 years. Mr. Hamilton has attained the highest level of certification in homeland security from the American College of Forensic Examiners and is Federal Emergency Management Agency-certified in the National Response Plan and the National Incident Management System. Mr. Hamilton is on the Training Advisory Board for the Texas Attorney General and is a board member of the Greater Austin Crime Commission. He is also a graduate of the FBI Citizen's Academy and has previously served as the HAZMAT Committee Chairman of the Texas Commission on Law Enforcement Officer Standards and Education. He also serves as the Future Plans Officer (N5) for the Texas State Guard Maritime Regiment and is a certified concealed handgun instructor, wild-land firefighter, diver, and armorer.

Mr. Hamilton graduated with honors from the University of Texas at Austin with a bachelor of science in engineering and received his master of science in engineering from the University of Texas.

**Dag Harmsen** is head of research, Periodontology Department, University Hospital Münster. He is an M.D. specializing in microbiology and the epidemiology of infectious diseases. He commenced his training in 1991 at the Institute of Hygiene and Microbiology, University of Würzburg (Germany). From September 2000 to January 2002 he headed the Research and Development Diagnostics division of CREATOGEN AG (Augsburg, Germany). From February 2002 to September 2004 was senior research scientist at the Institute of Hygiene, University of Münster (Germany). From June 2005 until July 2008, he was temporary head of the Department of Periodontology, University of Münster. He is now head of research of this department. His scientific interests focus on molecular diagnostics,

epidemiology, and the phylogeny of microorganisms. Furthermore, he is specialized in applied bioinformatics in microbiology.

**Dana R. Kadavy** is senior microbiologist, Signature Science. Dr. Kadavy joined Signature Science in 2003, serving as a microbiologist and principal investigator to multiple microbiological and molecular research projects. She has developed unique field deployment and sampling, development, and validation assays for molecular detection of microorganisms and alternative collection TTP to support national security and law enforcement scientific applications. Leveraging her academic teaching experience, Dr. Kadavy has developed very popular and effective training programs in Hazardous Materials Response and Applied Foundational Biology (for response units). Prior to joining Signature Science, Dr. Kadavy served as Chief of Bacteriology with the Armed Forces Institute of Technology.

Dr. Kadavy earned a doctorate in microbiology from the University of Nebraska–Lincoln in 2001. She has published multiple professional papers and presents regularly at professional conferences.

**Paul Keim** (see Appendix A)

**Raymond Lin Tzer Pin** is head and senior consultant, Division of Microbiology; Clinical Director, Molecular Diagnostic Centre. Dr. Lin is a senior consultant at the Department of Laboratory Medicine, head of the Microbiology Division, and clinical director of the Molecular Diagnosis Centre. He is adjunct associate professor with the Department of Microbiology, Yong Loo Lin School of Medicine, National University of Singapore. He studied medicine at the National University of Singapore and obtained his M.Sc. with distinction in medical microbiology at the London School of Hygiene and Tropical Medicine (University of London). He is a Fellow of the Royal College of Pathologists of Australia and of the Academy of Medicine, Singapore. Professor Lin has wide-ranging interests related to medical microbiology, including antibiotic resistance, strain typing, molecular diagnostics, and emerging infections.

Dr. Lin is actively engaged in infection control issues and has had experience in several hospitals in this respect. He also believes strongly in promoting quality systems in the laboratory and the use of informatics in microbiology practice. He has published in many refereed journals and has coauthored a monograph on enteroviruses. He sits on committees on the hospital and national level, and is involved in two committees of the Royal College of Pathologists of Australasia.

**Juncai Ma** (see Appendix A)

**Alemka Markotić** (see Appendix A)

**Piers Millet** is deputy head of the Implementation Support Unit, Biological Weapons Convention (CBWC), United Nations Office for Disarmament Affairs, in Geneva, Switzerland. His duties there include acting as deputy secretary to meetings of the BWC, liaising with international, regional, and expert bodies as well as developments in science and technology relevant to the treaty regime. Dr. Millett has served as a member of the Secretariat for all meetings of the BWC since 2001.

He trained originally as a microbiologist and is a chartered biologist in the United Kingdom. He has a doctorate from the University of Bradford on the past, present, and future of anti-animal biological warfare which focused heavily on the impact of developments in the life sciences on biological weapons.

Dr. Millet also holds postgraduate degrees in international politics and security studies as well as research methodology. He is widely published on issues related to preventing the acquisition and use of biological weapons and is a regular speaker at conferences around the world. His efforts have seen him collaborate with a range of other intergovernmental organizations, including the World Health Organization, the World Organisation for Animal Health, the Food and Agriculture Organization, the International Committee for the Red Cross, INTERPOL, the Organization for the Prohibition of Chemical Weapons, the U.N. Institute for Disarmament Research and the U.N. Interregional Crime and Justice Research Institute. He is also a member of the Advisory Board of the Virtual Biosecurity Center and a founding member of the Safety Committee of the International Genetically Engineered Machines Competition.

**Stephen A. Morse** is associate director for environmental microbiology, National Center for Emerging and Zoonotic Infectious Diseases, Centers for Disease Control and Prevention, Atlanta, Georgia. Dr. Morse's previous post at the CDC was associate director for science. He is the coauthor of the *Atlas of Sexually Transmitted Diseases and AIDS, 4th Edition*. Before he joined the CDC, Dr. Morse was a professor of microbiology and infectious diseases at the Harvard School of Public Health and at Oregon Health Sciences University. In 1984, the CDC recruited him to direct their sexually transmitted disease research program. He was subsequently named deputy director of the bioterrorism program. In 2008, he became associate director for environmental microbiology at the National Center for Preparedness, Detection and Control of Infectious Diseases at the CDC, a post that also encompassed bioterrorism issues.

Dr. Morse earned his M.S.P.H. and Ph.D. degrees in microbiology and public health at the University of North Carolina at Chapel Hill. He is cur-



rently a member of the Public Health Foundation Board at the University of North Carolina's Gillings School of Global Public Health. In 2000, he received the school's Harriet Hylton Barr Distinguished Alumni Award, which recognizes the achievements of alumni and their outstanding contributions to public health.

**Randall Murch** is associate director, Research Program Development, National Capital Region, Virginia Tech, adjunct professor, School of Public and International Affairs, and adjunct professor, Department of Plant Pathology, Virginia Tech, and visiting professor, Department of War Studies, King's College London, United Kingdom. Following graduate school and brief service in the U.S. Army Reserve, Dr. Murch's first career was with the Federal Bureau of Investigation (FBI), where he was a special agent. In his early years with the FBI, he was assigned to the Indianapolis and Los Angeles Field Offices where he performed counterterrorism, counterintelligence, and other investigations. During his career, he was assigned to the FBI Laboratory as a forensic biologist, research scientist, department head, and deputy director at various times. Interspersed with his laboratory assignments were four assignments in the bureau's technical investigative program: as a program manager for complex operations planning, Intelligence Division; unit chief for a technology development and deployment group, Technical Services Division; squad supervisor, New York Field Office; and deputy director, Investigative Technology Division. Between his last two FBI assignments, he was detailed to the Defense Threat Reduction Agency (DTRA), Department of Defense, where he was the director of the Advanced Systems and Concepts Office, where he led advanced studies on complex current and future challenges dealing with weapons of mass destruction (WMDs). While in the FBI, he created the FBI's WMD forensic investigative program, served as the FBI's science advisor to the 1996 Olympic Games, led forensic investigative aspects of a number of major terrorism cases, and initiated a number of new programs for both the FBI Laboratory and technical investigative program. In 1996, Dr. Murch created the FBI's Hazardous Materials Response Unit, the nation's focal point for the forensic investigation of WMD threats, events, and hoaxes; this laid the foundation for the creation of new fields in nuclear, chemical, and biological weapons forensics. Throughout his FBI career, he also was involved with extensive liaison at the national and international levels in furthering science and technology for law enforcement, counterterrorism, and national security purposes. He retired from the FBI in November 2002 after nearly 23 years of service, and as a member of the Senior Executive Service for the last 7 of those years.

From December 2002 to December 2004, Dr. Murch was employed as a research staff member, Institute for Defense Analyses (IDA), a lead-

ing federally funded research and development center, where he led and participated in studies for the defense, intelligence, and homeland security communities. He is still an adjunct staff member at IDA. He joined Virginia Tech in December 2004, where he now works in the areas of life science research program development, microbial systems biology, microbial forensics, biosecurity, and university strategic planning. He has served or still serves on the Board of Life Sciences, National Research Council; DTRA's Threat Reduction Advisory Committee; the Defense Intelligence Agency's BioChem 2020, the FBI's Scientific Working Group on Microbial Genomics and Forensics, and a new standing committee of the National Academy of Sciences for the Department of Homeland Security's National Biodefense Analysis and Countermeasures Center. He has also been or is a member of, or has advised study committees of, the National Research Council, National Academy of Sciences, Institute of Medicine, Defense Science Board, and Threat Reduction Advisory Committee.

Dr. Murch received his bachelor of science in biology from the University of Puget Sound, Tacoma, Washington, in 1974; his master of science degree in botanical sciences from the University of Hawaii in 1976, and his Ph.D. in plant pathology from the University of Illinois, Urbana-Champaign in 1979.

**Dana Perkins** is a member of the Group of Experts established pursuant to United Nations Security Council Resolution 1540. Dr. Perkins earned a master's degree in biochemistry from the University of Bucharest, Romania. She also earned a Ph.D. in pharmacology and experimental therapeutics in 2002 from the University of Maryland, Baltimore, where she specialized in microbiology/neurovirology. In her prior position, she led the Biological Weapons Nonproliferation and Counterterrorism Branch in the Office of Policy and Planning, Office of the Assistant Secretary for Preparedness and Response (ASPR), U.S. Department of Health and Human Services (DHHS). At DHHS/ASPR, some of her responsibilities and duties included providing subject-matter expertise, interagency coordination, and senior-level policy advice on the scientific (biodefense and biosecurity) and public health aspects of national and international emergency preparedness and response; directing and coordinating national and international progress on issues related to biodefense and biosecurity; developing and reviewing policies on biosecurity, biological weapons nonproliferation, and health security; and performing expert analysis and preparing implementation plans to support the U.S. government biodefense and biosecurity policy.

Currently, Dr. Perkins serves in a U.S. government-seconded position as a member of the Group of Experts supporting a subsidiary body

of the U.N. Security Council, the 1540 Committee. The 1540 Committee was established pursuant to Resolution 1540 (2004) to monitor the implementation of this resolution worldwide. U.N. Security Council Resolution 1540 imposes binding obligations on all states to prevent the proliferation of nuclear, chemical, and biological weapons, related materials, and their means of delivery to terrorists and other non-State actors. It also encourages enhanced international cooperation on such efforts.

**Dragan Primorac** is a pediatrician, forensic expert, and geneticist, adjunct professor, Eberly College of Science, The Pennsylvania State University, and Henry C. Lee College of Criminal Justice and Forensic Sciences, University of New Haven, Professor at Medical Schools in Split and Osijek, Croatia; founder of the Special Hospital for Orthopaedics, Neurology and Physical Medicine and Rehabilitation "St. Catherine," Croatia; and cofounder of Department of Forensic Sciences, University of Split, Croatia.

Professor Primorac is a pioneer in the application of DNA analysis for identification of bodies in mass graves and one of the founders of forensic DNA analysis in the region. He authored more than 100 scientific papers and abstracts in clinical medicine, molecular genetics, forensic science, population genetics, genetic legacy of *Homo sapiens sapiens*, and education, science, and technology policy. His papers have been cited more than 1,500 times and currently he is the most cited clinician in Croatia in his age group in the field of biomedicine. Currently, he is involved in two FP-7 projects (diagnostic and prognostic biomarkers for inflammatory bowel disease IBD-BIOM and multidimensional omics approach to stratification of patients with low back pain), altogether worth over 12 million euros.

Professor Primorac has been an invited speaker at 60 conferences around the world and he is the cofounder of International Society of Applied Biological Sciences (ISABS) in which the Scientific Committee for Nobel Prize Laureates is involved. Several renowned media outlets, both electronic and print, have reported on the results of his work, such as the *New York Times*, *USA Today*, *Chicago Tribune*, *Hartford Courant*, *JAMA*, *The Lancet*, *Science*, NBC, and Channel 8 (Connecticut TV station).

Professor Primorac has received 21 domestic and international awards including the Young Investigator Award of the American Society for Bone and Mineral Research in 1992, the Michael Geisman Fellowship Award of the Osteogenesis Imperfecta Foundation in 1993, the Life Time Achievement Award from the Henry C. Lee's Institute of Forensic Science in 2002, the Award of the Italian Region Veneto for Special Achievements in Promoting Science in the EU in 2007, the University of New Haven's International Award for Excellence in 2010, and the Presidential Award by the president of the International Association of Forensic Sciences for

his contribution to forensic sciences in 2011. He is an honorary citizen of five cities in and outside of Croatia.

From 2003 to 2009 he served as minister of science, education and sports, Republic of Croatia. According to the International Republican Institute survey of October 1, 2007, he was rated as the most successful minister in the Croatian government with 31 percent approval rate. As the minister of science, education and sports, Dr. Primorac launched a series of successful reforms in primary, secondary, and tertiary education as well as in science, technology, and sports that significantly improved the system. The award for numerous efforts made in the Croatian educational system is the survey of *Newsweek* (2010) which rated Croatia 22nd in education, ahead of 12 countries from the G20 group.

**Cerys Rees** is capability lead, CB Analysis and Attribution, Detection Department, Dstl Porton Down, and is responsible for both the operational and research aspects in this area. She is a microbiologist with 15 years of experience handling dangerous pathogens, and for the last nine years has been the technical leader for the development of the CB analysis capability for defense and security (law enforcement) purposes in the United Kingdom.

**Richard Vipond** is operations manager, Rare and Imported Pathogens Laboratory (RIPL), Public Health England, Porton Down, and project manager for groups providing diagnostic support for RIPL and research into related pathogens. He is responsible for assessment and application of new diagnostics, genotyping, and genomic methods to study high-containment pathogens. His interests include bacterial pathogenicity, genomics, proteomics, diagnostics, epidemiology, and high-containment bacterial pathogens. Dr. Vipond has published on bacterial pathogenicity, vaccine evaluation, vaccine characterization, genome analysis, and epidemiology.

**Haruo Watanabe** is director general, National Institute of Infectious Diseases, Tokyo, Japan, and professor, Faculty of Medicine, Tokyo University, Japan. Dr. Watanabe's research interest areas are in understanding the molecular pathogenesis of enteric bacteria including enterohemorrhagic *Escherichia coli*, *Salmonella*, and *Shigella*, and development and application of molecular epidemiological methods for outbreak investigations. Dr. Watanabe holds M.D. and Ph.D. degrees from Gunma University in microbiology.

**Ruifu Yang** is a professor at the Beijing Institute of Microbiology and Epidemiology, China. His research is focused on bacterial genomics, evo-

lution, and pathogenesis. He has published more than 150 papers in peer-reviewed journals, including the *New England Journal of Medicine*, *Nature*, *Nature Genetics*, *PNAS*, *PLOS Genetics*, and *Clinical Infectious Diseases*. He is vice director of the Chinese Society of Microbiology and a member of the National Evaluation Committee for Food Safety Risk Assessment. He is also associate editor for *PLOS NTD*, *Microbiologica Sinica*, *Chinese Journal of Preventive Medicine*, *Journal of Zoonosis*, and editor for *Pathogens and Diseases*, *New Microbes and New Infections*, and *Chinese Journal of Microbiology and Immunology*.