Overview of Microbial Forensics and the Concepts of Validation

Bruce Budowle
Executive Director of the Institute of Investigative Genetics
Professor in Department of Forensic and Investigative Genetics
University of North Texas Health Science Center
Fort Worth, Texas USA
Task:

- Overview of Microbial Forensics
- Validation
- If I were a committee member and what I want to consider
Anthrax Attacks

Florida
New York
Washington, DC

Demonstrated the need for a coordinated national response for threat assessments, countermeasures, and forensics for biological agents
Four Mission Areas

Bioterrorism or other crime

- PREVENT
  Terrorism by tackling underlying causes

- PURSUE
  Terrorists and those that sponsor them

- PROTECT
  The public and Nation interests

- PREPARE
  For the consequences

To reduce the

THREAT

To reduce the

RISK

To reduce the

VULNERABILITY
Forensic Science

• Application of science in the investigation of legal matters

• Scientific knowledge and technology are used to serve as witnesses in both criminal and civil matters

• Science may not offer definitive solutions for all scenarios; it does provide a special investigative role

• Goal is “attribution” – i.e., who committed the crime
Microbial Forensics

- Analysis of evidence from a bioterrorism act, biocrime, hoax, or inadvertent microorganism/toxin release for attribution purposes
- Essentially the same as any other forensic discipline
- Multi-disciplinary
Current Situation

- The FBI is the lead investigative agency in response to acts of terrorism

- The FBI Laboratory currently cannot conduct forensic examinations of hazardous CBRN agents at Quantico

- The task is monumental and must rely on partnerships

- Robust capabilities for the forensic exploitation of hazardous CBRN evidence are still developing

- The knowledge base required to interpret and assign weight to some CBRN analyses is lacking
A Complex Problem

- Potential terrorist weapons include a wide range of chemical, biological, and radiological / nuclear agents

  - **Biological:**
    - Human, plant and animal Pathogens

  - **Chemical:**
    - Chemical warfare agents, toxic industrial chemicals

  - **Radiological / Nuclear:**
    - Actual material, dispersive device (RDD), low – high yield device

- Must Consider ALL Related Evidence and Matrices
Magnitude of Problem

• Any infectious agent can be used as a biological weapon
• Emerging pathogens

• Over 1000 agents known to infect humans*
  – 217 virus species
  – 538 bacterial species
  – 307 fungi
  – 66 parasitic protozoa

• Additional plant and animal pathogens not counted
• Numerous strain variations
• Potential bio-engineered organisms

This does not address forensic signatures
Biosynthetic Technology

• New risks are accruing today due to advances in DNA synthesis

• These risks are vastly outweighed by the benefits of synthesis technology

• Biology will become, relative to current capabilities, extraordinarily easy to engineer

• There may be no limit to the number of harmful biological agents that humans could produce

• Rapid, accurate and thorough diagnostics are needed
Being entirely prepared is not possible

We must acknowledge this limitation

So what should be done?
Proactive

Build relationship / infrastructure
Stress ---

Epidemiology and Forensics

Build relationships with public health, agriculture and laboratory assets (government, academia, and private)
Overt Attack

Courthouse receives letter labeled “Anthrax”

Covert Attack

Unusual, Disease Clusters, Signs & Symptoms

State & Local Public Health

Law Enforcement ↔ CDC/FDA
Epidemiologic considerations that may signal a bioterrorist attack

- Disease caused by an uncommon agent (such as smallpox)
- Unusual, atypical, genetically engineered or antiquated strain of agent
- High morbidity or mortality associated with a common disease or syndrome
- Failure of patients to respond to usual therapy
- Disease with an unusual seasonal or geographic distribution
- Increase in normal incidence
- Atypical disease transmission (such as shigella in muffins)
- Illness in people who are exposed to same ventilation system
Epidemiologic considerations that may signal a bioterrorist attack

• More than one unusual or unexplained disease existing in a person

• Illness that affects a large disparate population

• Illness that is unusual for a population or age group

• Unusual death or pattern of illness in animals preceding or accompanies death or illness in humans and vice versa

• A number of ill persons seeking treatment or medicine at the same time

• Same strain or genetic type from spatially or temporally disparate sources

• Simultaneous cluster of disease in noncontiguous areas

• Large number of unexplained diseases or deaths
Post-Harvest background: Human Food-borne Illness

- ~ 75 Million Illnesses
- More than 300 Thousand Hospitalizations
- ~ 5 Thousand Deaths
- A Cost of $5-15 Billion per Year
Surveillance

• The ongoing, systematic collection, analysis, and interpretation of health data essential to the planning, implementation, and evaluation of public health practice, closely integrated with the timely feedback of these data to those who need to know
Success Through Partnerships

- National BioForensics Analysis Center at Fort Detrick, Maryland is the FBI Partner Lab Hub for Biological Agents
Crime Scene

- Control crime scene
- Document – Administrative Log
- Coordinate personnel and make assignments
- Ensure properly equipped
- Survey scene
- Photograph and sketch
- Plan and search
- Collect evidence
- Collect samples
- Package
- Transport
- Etc.
Evidence Collection

Sometimes Incompatible with Other Forensic Techniques
Sample Analyses Flow Chart

Issues:
- Control material
- Control environment
- Maximize recovery

Bacteria ID:
- Classic bacteriology

Strain ID:
- Genetic Engineering

Spore density:
- Cfu/ml

Percentage viability:
- Dead vs live spores

Particle sizing

Analytical chemistry
- Inorganic
  - Silica
  - Silicates
  - Cations and anions
  - Heavy metals
- Organic
  - Carbohydrates
  - Agar (Agarobiose)
  - Peptones
  - Headspace

Electron Microscopy (EM)
- Scanning EM
  - EDX Analysis
- Transmission EM

Recover Evidence

Evidence

Transmission EM
Technologies

• Many

• Different levels of sensitivity, specificity, .....
Quality Management Guidelines for Laboratories Performing Microbial Forensic Work

- Goal is to promote development of a microbial forensics program that is scientifically valid and rigorous

- Define criteria for development and validation of microbial forensics methods that will support attribution for criminal investigations

- Establish national working guidelines for quality assurance and quality control as applied to microbial forensics
Validation

- Necessary

- Varies

- Ill defined

- Challenge to define better...

- Documents – QA & Validation
Validation

- Collection
- Shipping and storage
- Extraction
- Analysis
- Interpretation
Example Validation Criteria List

- Sensitivity
- Specificity
- Reproducibility
- Precision
- Accuracy
- Resolution
- Reliability
- Robustness
- Specified samples
- Purity

- Input values
- Quantitation
- Dynamic range
- Limit of detection
- Controls
- Window of performance for operational steps of assay
- Critical equipment calibration
- Critical reagents
- Databases

Note: Not all these need apply and others may be necessary
Validation of Testing Procedures

- Developmental Validation
- Internal Validation
- Preliminary Validation***
Validation

• Developmental validation is the acquisition of test data and determination of conditions and limitations. Developmental validation should be appropriately documented and should address specificity, sensitivity, reproducibility, bias, precision, false-positives, false-negatives, and determine appropriate controls. Any reference database used should be documented.

• Internal validation is an accumulation of test data within the laboratory to demonstrate that established methods perform as expected.
Validation

- Preliminary validation is the acquisition of limited test data to enable an evaluation of a method used to provide investigative support to investigate a biocrime or bioterrorism event. If the results are to be used for other than investigative support, then a panel of peer experts, external to the laboratory, should be convened to assess the utility of the method and to define the limits of interpretation and conclusions drawn.

- SOPs are for routine work
- But so locked in – restricts analytical thinking and possibly ignores both inculpatory and exculpatory evidence

- Stress Panel – met substantial resistance
- Government players have yet to relinquish “control”
- Not about control – but about access to best and brightest
Validation

• “Validation Plan” should be prepared

• To facilitate, guide, and educate

• Establish the range of conditions for which the interpretation of the analytical results is valid

• Equally as important, the conditions where results or the standard interpretation are not valid
Validation

A validation protocol should at a minimum include:

• Description of goal or purpose of the assay
• Critical steps
• Critical reagents
• Critical equipment
• Parameters and conditions to be evaluated
• Reference and test materials needed
• Sufficient number of replicate analyses to demonstrate reproducibility and reliability
• Aspects unique to the system that require specified validity testing
Minimum Validation Criteria

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

• Recovery
• Stability
• Integrity
• Target – either organism or analyte (e.g., toxin)
• Influence of sample matrix
Collection Validation

- Sampling strategy – hypothesis or circumstance driven, targeted or randomized/statistical collection
- Describe controls such as field blanks
- Consider resource limitations
- Type of collection material or tool
- Efficiency of collection
- Efficiency of recovery
- Substrate - inert or interacts with the target
- Stability and preservation - inhibitors, viability…
- Recovery from matrix
- Packaging and storage strategies
Sampling Objectives

- Real-time monitoring
- Screening
- Bulk material
- Questionable article
- Extent of contamination
- Effectiveness of decontamination
- Clearance for re-occupancy
- Transitional
- Crime scene / forensic
Sampling Approach

- Logical and systematic
- Scheduled
- Risk-based
- Targeted
- Statistical/Random
Sample Collection and Preservation for Plant Pathogens

- Develop plan with consultation
- Use experience of current collectors for how to collect (pattern and cutting and bagging)
- Chain of custody
- Targeted and statistical sampling plans – symptomatic/asymptomatic
- Collect vectors
- Collection from multiple plant parts, multiple plants
- Collection from other nearby plants
- Stability over time
- Preservation and transport of sample material (“breathing” of bags)
- On ice - 4C? Glycerol? Dessicant?
Extraction of Analyte

- Specific – virus, bacteria, fungal, toxin
- Spore v vegetative cell
- Active v 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
Extraction of Analyte

- Maintenance of original state – activity (toxin)
- Minimum input
- Yield, recovery
- Purity
- Critical reagents
- Controls
Analysis Phase Validation

- SOP
- controls – positive, negative and inhibition
- Specificity
- Dynamic range
- Reproducibility
- Reliability
- Precision
- Accuracy
- Predictive value
Analysis Phase Validation

- Input 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
• Sensitivity claims of 1 copy
• Is this reliable?
• What does it mean?
• Should there be an “inconclusive” category?
WHAT LEMMINGS BELIEVE
Interpretation of Results

• What did it mean when Biowatch had some positive hits for *F. tularensis*?

• Can Intelligence Analysts assess the data effectively?
Missing Data

• Second World War

• Study of planes returning from bombing Germany

• Created a rough diagram where bullet holes were and recommended those areas be reinforced

• Abraham Wald (1980) pointed out that essential data were missing from the sample studied

• What about the planes that did not return?

• Wings and tail – not the body
Missing data

• Biology is mutable

• Many things can change the organisms composition

• It is vast and unknown

• The full nature of the biological world is unknown

• We can speculate

• But never conclude

• Yet, we must conclude
Randomness

- Randomness does not create regularity
- Randomness creates the opposite
- Be cautious of the use of the word “random”
Interpretation of Results

• Statements – qualitative, quantitative, semi-quantitative
• Database – type, relevance, representative, quality of data
• Background data – normal values, reference range, endemicity
• Does a result require follow up or further analysis – temporal/spatial analysis, effect of passage
Interpretation of Results

• Alternate (reasonable) explanations
• Limits of interpretation
• Statistical approach – match, similarity, most recent common ancestor, identical
• Thresholds
• Software
Questions

• Four genetic markers (A1, A3, D, and E) for identifying the evidentiary materials
• Metadata – strength and weaknesses
• A large number of samples from known sources in repository
• Significance of the presence of these markers in the evidentiary samples???
• Statistical strength (cluster analysis) of finding the same signatures in specific repository samples
• Analytical weaknesses – e.g., false negatives

• Purpose - possible source (or lineage) attribution of the pathogen(s) detected in the evidentiary samples
Standard Operating Protocols

• Contain sufficient detail about the procedure so one can carry out the assay and include, if appropriate:

1) all steps in the procedure
2) proper controls
3) all reagents and preparations
4) calibration
5) criteria for analysis of results
6) interpretation of results
Miscellaneous Needs

- Need a panel of isolates to test (validate) the assays (a “type” set)
- Type of panel materials? – appropriate to assay
- Validity of panel
- Literature compendium - especially of old literature
Forensic Questions

• What is the agent?
  – Species, strain, or more

• Was the event intentional?
  – Obvious in the Anthrax case

• How was it made?

• Where did it come from?

• Who did it?
Non DNA-based tools for the microbial forensics toolbox

- Characterization of physical attributes acquired during preparation
- Isotope analyses to approximate the age and source
- Physiologic methods (e.g., fatty acid composition, phage typing, serotyping)
- Analysis of growth media and media components adhering to the microorganisms
- Analysis of stabilizers and additives used in the preparation of a sample
Non DNA-based tools for the microbial forensics toolbox

- Identification of incidental biocontaminants, such as environmental pollen and fungi, for location and time of year of preparation
- Better understanding of bacterial endemism for identifying unique strains that may exist in only one location or few locations
- Monitoring changes in the immunological response of a host to a pathogen or toxin, such as temporal IgG and IgM responses
- Improvements in immunoassays (and antibodies) for more effective rapid detection and field deployable assays.
Forensic Genetic Questions

- 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 the 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?
Forensic Genetic Questions

- Are the genetic differences to 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 recent sources of the isolate?
- Can contaminants be informative?
- Are there alternative explanations for the results that were obtained?
Comparison of Human DNA Forensics and Microbial Forensics

<table>
<thead>
<tr>
<th>Similarities</th>
<th>Differences</th>
</tr>
</thead>
<tbody>
<tr>
<td>QA/QC</td>
<td>Database size and composition</td>
</tr>
<tr>
<td>Databases</td>
<td>Statistical Interpretation methods</td>
</tr>
<tr>
<td>Qualitative conclusions of test results</td>
<td>Confidence/uncertainty of outcome of interpretation</td>
</tr>
<tr>
<td>Quantitative conclusions</td>
<td></td>
</tr>
<tr>
<td>General acceptance, admissibility,…</td>
<td></td>
</tr>
</tbody>
</table>

Main limitation --- Scant Data
Components of a bioinformatic genetic toolbox for microbial forensics

• Algorithm(s) for DNA marker alignment encompassing pattern heterogeneity of the various types of genetic markers and for detecting genes, such as for pathogenicity and antibiotic resistance

• 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
Components of a bioinformatic genetic toolbox for microbial forensics (continued)

• Capability to relate diversity to function

• Capability for comparative and functional genomics

• Contain or access curated (genetic marker) databases on pathogens and near neighbors and their background occurrence with epidemiological history, when available

• 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)
Conclusion
Capabilities?

- Identify experts – build relationships
- Scientific working groups and establish guidelines
- Standards -- Standardization
- Define better and encourage validation and peer review of the science
- Address and formalize preliminary validation
- Share information and capabilities within the law enforcement community
- Foster partnerships
- Test/challenge system for reliability