

Working Paper Nr. 1

Microbial Forensics for the BWC

Anna Krin

Verifikation im Bereich des Biowaffenübereinkommens neu denken (VERIBIO)
Hamburg, Germany



Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

Abbreviations

- BW: Biological Weapons
- BWC: Biological Weapons Convention
- bp: base pairs
- CCS: Circular Consensus Sequencing
- CLR: Continuous Long Read
- CRISPR: Clustered Regularly Interspaced Short Palindromic Repeats
- CWC: Chemical Weapons Convention
- DNA: Deoxyribonucleic acid
- DLT: Distributed Ledger Technology
- ELISA: Enzyme-Linked Immunosorbent Assay
- ESI: Electrospray Ionization
- FBI: Federal Bureau of Investigation
- GC: Gas Chromatography
- HPLC: High-Performance Liquid Chromatography
- IP RP HPLC: Ion-Pair Reversed-Phase High-Performance Liquid Chromatography
- ISU: Implementation Support Unit
- LC: Liquid Chromatography
- LCN DNA: Low-Copy Number DNA
- MALDI: Matrix-Assisted Laser Desorption-Ionization
- MLST: Multilocus Sequence Typing
- MLVA: Multiple Locus Variable Number Tandem Repeat Analysis
- NGS: Next-Generation Sequencing
- NPT: Treaty on the Non-Proliferation of Nuclear Weapons
- ONT: Oxford Nanopore Technology
- OPCW: Organisation for the Prohibition of Chemical Weapons

- PCR: Polymerase Chain Reaction
- RNA: Ribonucleic acid
- RC-PCR: Reverse Complement Polymerase Chain Reaction
- SAB: Scientific Advisory Board
- SMRT: Single-Molecule Real-Time Sequencing
- SNP: Single-Nucleotide Polymorphisms
- S&T: Science and Technology
- TEM: Transmission Electron Microscopy
- TGS: Third-Generation Sequencing
- TOF: Time of Flight
- UKMFC: United Kingdom Microbial Forensics Consortium
- UNMOVIC: United Nations Monitoring, Verification, and Inspection Commission
- UNSCOM: United Nations Special Commission
- UNSGM: UN mechanism established by the Secretary General
- VNTR: Variable Numbers of Tandem Repeats
- VR: Virtual Reality
- WHO: World Health Organization
- WMD: Weapons of Mass Destruction
- WMS: Whole Metagenome Sequencing

Glossary

Antibodies: Proteins in the body that recognize structures (antigens) that trigger an immune response (e.g. pathogens). Can also be synthetically produced for biochemical purposes.

Artificial intelligence, AI: An umbrella term for various computational approaches to perform tasks that are generally associated with the capabilities of human cognitive processes.

Bacterial cultivation: Methods of growing bacteria in laboratory set-up under controlled conditions.

Bacterial staining: Methods of staining bacterial for microscopic observations.

Bacterial strains: Bacterial subtypes of the same species characterized based on their specific genetic traits.

Deoxyribonucleic acid, DNA: Genetic code, blue print of an organisms.

Distributed Ledger Technology, DLT : A decentralized cryptographically secured digital system for recording transactions across multiple computers (nodes), enabling immutable and transparent data sharing and storage. Blockchain is the most well-known type of DLT.

DNA denaturation: Process in which double-stranded DNA is unwound and separated into two single strands. Can be triggered, for example, by high temperature.

DNA replication: Process of DNA duplication in the cells, resulting in two identical DNA copies, e.g., during cellular division.

Enzymes: Biological molecules (proteins) involved in various biochemical processes, such as, e.g., DNA synthesis, etc.

Fluorescent labeling: Labeling of biochemical structures with fluorescent molecules resulting in a detectable fluorescent signal for characterization, identification or quantification purposes.

Mass spectrometry: Analysis technique for the unambiguous identification and quantification of sample components based on the measured mass-to-charge ratios. Includes many different set-up configurations to achieve high resolution.

Multilocus Sequence Typing, MLST: A technique to identify bacterial strains based on the variations in their essential genes (housekeeping genes).

Multiple Locus Variable Number Tandem Repeat Analysis, MLVA: Molecular typing method based on the characterization of unique repeats in genetic sequences (VNTRs).

Near-site inspections:

Nucleic acids: Collective term for DNA and RNA.

Nucleotides: Building blocks of DNA and RNA.

On-site inspections:

Off-site inspections:

“Omics”: Comprehensive study of large sets of biological molecules on the level of proteins (proteomics), genomes (genomics), metabolism (metabolomics), etc. in terms of their functionality instead of analyzing individual entities.

Polymerase Chain Reaction, PCR: Technique used in the laboratory to amplify DNA.

PCR primers: Short single stranded DNA fragments that anneal to complementary DNA sequences for amplification.

Sequencing methods: Methods for determining the exact order of DNA (or RNA) building blocks in a

genetic sequence and for obtaining genetic information. Various high-throughput sequencing methods are listed in the manuscript.

Ribosomes: Proteins that are involved in the process of protein biosynthesis, translating the genetic information into the sequence of interlinked building blocks of proteins (aminoacids). Consists of a small and a large subunits. The 16S ribosomal RNA is part of the small subunit of bacterial ribosom; the 18S ribosomal RNA is the part of the small ribosomal subunit of higher organisms (e.g. fungi); the 23S rRNA is part of large ribosomal subunit of bacteria.

Single-Nucleotide Polymorphisms, SNP: Genomic variations in which a DNA building block at a specific position in a gene sequence is replaced by another.

Variable Numbers of Tandem Repeats, VNTR: Segments in the genome sequence in which the same short sequences are repeated consecutively.

Summary

The Biological Weapons Convention (BWC) is one of the multilateral agreements that prohibits an entire category of weapons of mass destruction. However, it currently lacks a comprehensive and transparent verification mechanism. The Working Group established at the Ninth Review Conference (2022) is discussing, among other issues, the verification and compliance mechanism for the BWC. Regardless of the outcome of these debates, many questions remain concerning the technical and political foundations necessary for such a mechanism, should it be implemented in the future.

This manuscript is the first in a series of working papers analyzing available technical capabilities that could support future verification and compliance measures in the BWC. Specifically, it focuses on microbial forensics, a broad field that integrates areas like epidemiology, immunology, microbiology, and more, to identify and characterize pathogens. Unlike epidemiology, which has similar goals, microbial forensics mainly aims at attribution — the identification of responsible actor(s) in cases of biological weapon use. To accomplish this, it leverages advances in science and technology, including significant progress in genomic sequencing, computational tools, miniaturization, and automation of lab work-flows. These methods must be specific, sensitive, and reliable to meet the needs of microbial forensics.

Various options for a verification mechanism within the framework of the BWC are conceivable. Microbial forensics can be applied both for an ex-ante (early detection of the BWC provisions violation for preventing biological weapons development) and ex-post (identifying perpetrators in post-incident scenarios) verification concepts in the BWC. There are several documented examples of utilizing microbial forensics to investigate the cases biological weapons use, both by state and non-state actors. One prominent example is the investigation in the case of anthrax laced letters in the USA (Amerithrax, 2001) that to significant extend shaped and propelled the development of microbial forensics. Since then, the field has changed substantively, judging by the increase in international publications and some national initiatives (e.g. in UK Microbial Forensics Consortium) to harness expertise in this discipline.

Expertise and continuous training of specialists are essential to effectively apply microbial forensics and ensure that high standards and requirements are met throughout the entire chain from sample planning and collection through transfer, storage, analysis, data interpretation and reporting. Here, future guidelines and procedures that could be applied within the framework of the BWC can build on the experience gained from the United Nations Secretary-General's Mechanism (UNSGM) for investigating alleged use of biological and chemical weapons, which draws on a roster of experts, consultants, and laboratories, and includes regular training for professionals to maintain expertise and best practices.

Microbial forensics is one of the many disciplines that can be used in any future BWC verification and compliance frameworks. It should be viewed as support, not a replacement for other procedures such as classical forensics and information gathering to establish the chain of custody. Furthermore, the rapid evolution of technology introduces new applications that can benefit the Convention and could be incorporated into its future toolkit of methods. However, ongoing monitoring of these developments is essential. Therefore, establishing a mandated S&T advisory board alongside the BWC verification mechanism would be a prudent step.

Contents

1	Introduction	8
1.1	Verification in the BWC and objectives of the junior research group VERIBIO	8
1.2	Microbial forensics: Definition and promises for the BWC	10
2	Application of microbial forensics: historical cases	12
2.1	The Amerithrax case	12
2.2	Anthrax outbreak in Sverdlovsk	14
2.3	Confirmation of the presence of a bioweapon program in Iraq	16
3	UN Secretary General's Mechanism	18
4	Approaches and techniques used in microbial forensics	21
4.1	Next-generation sequencing	21
4.2	Further emerging technologies used in microbial forensics	28
4.3	Genetic engineering attribution using computational tools	29
4.4	Ribosomal ribonucleic acid (rRNA) detection	32
4.5	Mass spectrometry to analyze biomolecules	34
5	Conclusions	38
	References	40

List of Figures

2.1	Schematic representation of the DNA double helix	13
2.2	Single-nucleotide polymorphism (SNP)	14
2.3	Variable number tandem repeats	15
4.1	Scheme of the Polymerase Chain Reaction	22
4.2	Real-time PCR	23
4.3	Scheme of the amplicon-based sequencing.	24
4.4	Difference between the targeted amplicon-based sequencing and whole metagenome sequencing (WMS)	25
4.5	Nanopore sequencing	27
4.6	Principle behind sandwich ELISA	33
4.7	Matrix-assisted laser desorption-ionization (MALDI)- time of flight (TOF) spectrometry	34

Section 1

Introduction

1.1 Verification in the BWC and objectives of the junior research group VERIBIO

The Biological Weapons Convention (BWC) is one of the multilateral treaties that bans an entire category of weapons. Along with the Chemical Weapons Convention (CWC) and the Treaty on the Non-Proliferation of Nuclear Weapons (NPT), it forms a crucial part of arms control and disarmament efforts. However, unlike the CWC and NPT, it lacks a verification system to ensure compliance with its provisions. Currently, investigations into alleged biological weapon use can only be conducted under the UN mechanism established by the Secretary General (UNSGM). This mechanism, however, does not replace a transparent and comprehensive verification system. The need for more transparency in the current tense geopolitical realm has been illustrated by the recent speech (October 2025) of the NATO Secretary General at the NATO Biotechnology Community Conference where it was pointed out that “China has made rapid advancements in biotechnologies, including gene editing and synthetic biology for military and dual-use applications,” and that Russia continues to “leverage its expertise to conduct covert biological research.”¹

The discussion of a verification mechanism in the BWC is not a new phenomenon. One of the milestones on this obstacle-filled path was the decision during the Third Review Conference of the BWC (1991) to establish the Ad Hoc Group of Governmental Experts to “identify and examine potential verification measures from a scientific and technical standpoint”². This Group, further known as VEREX, has identified 22 possible measures in total that have been divided into two broad categories: the on-site and off-site ones.³ The off-site measures included, i.e., information monitoring, data exchange, remote sensing, and inspections. The on-site measures focused, e.g., on exchange visits, visual inspections, interviewing, and continuous monitoring by instruments and personnel. Furthermore, some measures have been evaluated in combination to determine whether they offer “enhanced capabilities and limitations that differ from a simple accumulation of the capabilities and limitations of the single measures involved (synergy)”³. However, despite the promising outcome from the published reports of VEREX, the negotiations on verification failed in 2001. It is assumed that one of the possible reasons for this failure was the opposition of the US biopharmaceutical industry. The industry’s reluctance was presumably based on concerns that regular inspections could lead to the disclosure and exploitation of proprietary data.

In 2022, over 20 years after the failure of the verification protocol in the BWC, during the Ninth Review Conference, it was decided to, i.e., establish a Working Group to address the measures on compliance and verification.⁴ However, many questions around any hypothetical future verification mechanism in the BWC remain unresolved, e.g.:

- What are the general framework conditions for a future BWC verification concept, considering current technical and geopolitical circumstances?

- How should BWC verification be designed in the industrial sector, considering the highly diversified biotechnological landscape where many compliance activities are carried out by academic or private stakeholders with a global presence?
- Which scientific and technical achievements can be effectively applied to the verification mechanism of the BWC and complement the methods already in place within the framework of the UNSGM?

This is just a brief list of potential questions that should be addressed in the preparation of any future verification concept for the BWC. Given the limited time available during the Working Group meetings, it seems unlikely that there will be in-depth discussions on these and other issues. Policymakers need an efficient exchange of information with industry and academia regarding scientific and technological advances, which have significantly changed since VEREX. The newly established junior research group VERIBIO (German: „Verifikation im Bereich des Biowaffenübereinkommens neu denken“), funded by the German Federal Ministry of Research, Technology and Space, will contribute to this informed discussion among various stakeholders and address some of the questions above. The main goal of VERIBIO is to develop a proposal for a potential ex-ante verification mechanism within the BWC integrating natural science and political science perspectives to create a comprehensive and effective framework. The core idea behind the ex-ante verification is to deter any actors from developing, producing, stockpiling, acquiring, or maintaining biological weapons in accordance with Article I of the Convention. It aims to prevent such activities rather than investigate them post-factum. This and subsequent manuscripts will focus on a detailed analysis of some developments in the science and technology field that should be considered when discussing appropriate technical measures for a potential BWC verification, inspired by VEREX achievements but guided by current state of the art.

It is undeniable that current advances in science and technology require careful assessment for their potential benefits in future BWC verification mechanisms, regardless of the political framework. The recent (December 2025) draft final report of the Working Group on the Strengthening of the Convention included following aspects regarding verification: consider “[s]cientific and technological developments relevant to the Convention that could have an impact on the effectiveness of proposed compliance and verification measures while also taking into account the relevant provisions of the Convention, as well as the unique nature of biological weapons.”⁵ It should also be noted that S&T requirements for an ex-ante verification process differ somewhat from those for an ex-post procedure. Early detection of violations in particular for preventing biological weapons development is crucial. The identification of perpetrators and attribution are priorities in post-incident scenarios.

For an ex-ante verification procedure, both on-site and off-site approaches and monitoring are possible. This is also reflected in the late draft final report of the Working Group, addressing “[p]otential on-site and off-site approaches to verification that are practicable, feasible, effective, cost-effective and complementary.”⁵ Here, a distinction should be met between the detection (necessary for the response to an outbreak but no in-depth analysis), identification (at the species level and beyond), characterization (includes information such as specific genetic and other particularities to distinguish natural from intentionally caused outbreaks), and attribution (involves search for specific signatures to link identified bioagent to its source) of bioagents. Depending on the questions addressed and the stage of the process, different methods are used for each of these tasks. Thus, on-site procedures would require expertise and equipment to reliably **detect** biothreat agents in small amounts, quickly, and without complex experimental procedures. Recently, significant advances have been made in nanotechnology, analytical methods, and microelectronics, resulting in portable devices meeting these criteria. Once the biothreat agent is detected, its unambiguous **identification** and **characterization** is possible using traditional techniques, such as mass spectrometry, chromatography, and whole genome sequencing. In particular, advances in the field of whole genome sequencing have a significant impact on resolving whether a biological threat agent has a natural origin or has been intentionally developed. These technologies are an essential part of microbial forensics. However, the most important aspect of microbial forensics is **attribution**. In fact, attribution is the element that distinguishes microbial forensics from epidemiological investigation, which shares many steps with microbial forensics when it comes to identifying and characterizing the infectious agent, but does not establish a chain of custody and differs in the rigor of the standards applied.⁶ Although attribution primarily reflects the objectives of the ex-post verification concept, this manuscript discusses a

number of methodological developments that can be used for attribution for the sake of completeness.

Microbial forensics has been crucial in investigations of biological weapons involving state and non-state actors. It can be used not only to investigate specific biothreat incidents but also as a “deterrent and support tool”,^{7,8} a property that aligns well with objectives of the ex-ante verification. It is therefore a promising tool that should be considered for the technical implementation of BWC verification from both an ex-ante and ex-post conceptual perspective. It has been highlighted as one of the two key biotechnology areas to attribute bioweapons in the case of the alleged BW use (the second being biosurveillance).⁹ As such, a more detailed analysis of its advancements will be provided in this manuscript. However, no single discipline suffices for the complex verification process, which demands a combination of approaches, robust technology, multiple information sources, and a solid political framework.

1.2 Microbial forensics: Definition and promises for the BWC

There is currently no canonical definition of microbial forensics. With some accuracy, it can be characterized as a “scientific discipline dedicated to analyzing evidence from a bioterrorism act, biocrime, or inadvertent microorganism/toxin release for attribution purposes”^{10,11}. This definition primarily focuses on bioterrorism or biocrime as the act of deliberately releasing bioagents to cause harm. However, state-sponsored bioweapon programs can be investigated or uncovered using microbial forensics (see also Section 2). It is therefore only reasonable that the capabilities and promises of microbial forensics have gained considerable attention in the BWC. The importance of determining “whether a disease outbreak is naturally occurring or deliberately caused, and cover diseases and toxins that could harm humans, animals, plants or the environment” was, e.g., emphasized at the meeting of the States Parties in 2010.¹² In this occasion, the importance of “forensic capability to provide evidence in the event of a prosecution involving the alleged use of a biological or toxin weapon” has also been recognized.¹²

In the document prepared by the Implementation Support Unit (ISU, the subsidiary body of the BWC) prior to the Seventh Review Conference a more specific definition of attribution has been provided. The core objective of microbial forensics is described as an investigative process that identifies and links a perpetrator to a biological weapons event.¹³ It should be noted that attribution in general does not rely only on microbial forensics, but also on other types of forensics and investigative methods that all together contribute to the holistic and context-based assessment of evidence leading to attribution. In the scope of this manuscript, however, we will focus solely on microbial forensics.

The highly interdisciplinary nature of microbial forensics has been emphasized, since a sophisticated characterization of genetic, physical, and chemical properties of the sample is required in the scope of an investigation.¹³ Microbial forensics indeed encompasses a wide range of disciplines such as epidemiology, microbiology, genetics, biochemistry, analytical chemistry, virology, statistics, plant pathology, veterinary medicine, physics, and bioinformatics on the one side,¹⁴ and the law enforcement, public health, policy, and intelligence community, on the other.¹¹ This also involves the engagement of various experts with deep knowledge in one or more of these fields. Considering the rapid developments across all these disciplines, it becomes clear that continuous and thorough assessment of scientific and technological progress is essential for the effective application of microbial forensics in future BWC verification of compliance. This fact highlights the need for a dedicated body within the BWC responsible for monitoring scientific and technological advancements related to the Convention’s objectives, similar to the Scientific Advisory Board (SAB) established under the CWC. Such a body should complement discussions on a feasible verification approach. The establishment of an S&T review mechanism under the BWC is currently on the Working Group’s agenda. However, the details of ongoing debates are beyond the scope of this manuscript.

In the compilation of measures identified by VEREX, microbial forensics was not explicitly mentioned. This is not surprising, since the area experienced its main surge in development primarily after 2001, following the investigation of the bioterrorism act in the USA involving letters laced with anthrax spores. This case will

be discussed in more detail below (see Section 2.1). Since then, microbial forensics has gained considerable attention, reflected by the increasing number of international publications from 1984 to 2022.¹⁵ The measures proposed by VEREX included “most sensitive techniques using PCR amplification, [...] and/or sequencing”³. Techniques based on polymerase chain reaction (PCR, see Box 4) and genome sequencing are still some of the core tools in modern microbial forensics. Since the introduction of VEREX, significant progress has been made in effectiveness, cost, and time reduction, which further emphasizes the importance of reviewing these methods from today’s perspective.

This manuscript is structured as follows: Section 2 reviews key historical cases where microbial forensics aided attribution in bioterrorism and state bioweapons investigations; Section 3 explains the role of the UNSGM in the investigation of the alleged use of biological weapons and the application of microbial forensics in this context; Section 4 examines modern methods and techniques used in microbial forensics; and finally, Section 5 summarizes the main points.

Section 2

Application of microbial forensics: historical cases

2.1 The Amerithrax case

The text of the BWC is primarily addressed to the States Parties and not to non-state actors, although Article IV obliges the States Parties to implement the provisions of the Convention at the national level in order to prevent non-state actors from engaging in activities prohibited under Article I of the Convention (another instrument designed to prevent the development, acquisition, production, possession, transportation, transfer or use of weapons of mass destruction by non-state actors is UN Security Council Resolution 1540[?] [?]). Any future verification mechanism under the BWC will therefore focus on the possible offensive bioweapon activities of the states, not on bioterrorism. However, the bioterror attack in the USA involving anthrax-causing biological agent *B. anthracis*, which became known under its codename as the Amerithrax case, has not only had a huge resonance in the community, but has also substantially contributed to the progress in the field of microbial forensics. Many lessons could be drawn during and after the Amerithrax investigation, and many techniques that at the time of the investigation were costly and available only to a handful of institutions became faster, more affordable, and widely spread. For this reason, a short overview of the microbial forensics procedure used in the course of the investigation will be given here. Although the Amerithrax case was a pivotal moment shaping microbial forensics in some ways, the discipline is deeply integrated into established epidemiological practices.¹¹ Therefore, the two fields of research are intertwined, even though they pursue different goals, as the attribution of causes is the prerogative of microbial forensics.

The Amerithrax case occurred in the USA in fall 2001, when letters laced with spores of *B. anthracis* were sent to several politicians and media representatives. Overall, 22 cases of anthrax were identified, with 5 fatalities, between October 4 and November 20.¹⁶ Additionally, there was widespread environmental contamination in the U.S. Postal Service building and other facilities involved in receiving or processing the letters.¹⁴ The investigation spanned nine years and involved analyzing a vast volume of collected data.¹⁷ As early as 1998, the FBI incorporated microbial forensics into its investigative toolkit.¹⁸ This step proved beneficial in the course of the Amerithrax investigation, since, in retrospect, traditional forensics alone could not provide the desired investigative leads.¹⁸ The origin of the disseminated spores and, above all, the person(s) responsible should be identified and prosecuted. This challenging procedure required substantial technical, financial, and human resources which were already in place at the commencement of the Amerithrex case.¹⁷

At the initial stage of investigation, clinical diagnostics involving some novel methods developed before Amerithrax, such as sensitive and quantitative immunoassay, identified *B. anthracis* as the sickness-causing agent in the first victim.¹⁸ Over the years, classical characterization of pathogens in microbiology included microscopy, staining methods, culture, etc. Some of these approaches have also been used in the Amerithrax case.¹⁴ However,

methods for better differentiation of bacterial strains were required. This can be achieved through the analysis of genomic sequences of the respective strains. In 1996, the so-called **variable number tandem repeats** (VNTR, Box 3) in the genome of *B. anthracis* were identified.^{19,20} VNTRs refer to segments in the genome, where short sequences of nucleotides (DNA building blocks) are repeated adjacent to each other.²¹ The exact number of repeats varies between strains, making these regions useful as genetic fingerprints to distinguish closely related bacteria. In 2001, Keim et al. developed a method known as **multiple locus variable number tandem repeat analysis** (MLVA, Box 3)²² that could subsequently be applied in the Amerithrax case to identify the strain of *B. anthracis* involved as the Ames strain, a laboratory strain used in vaccine research. This indicated an intentional spreading of spores,¹⁸ since the Ames strain is uncommon in nature.

Box 1. Basic structure of DNA

Deoxyribonucleic acid (DNA) contains the genetic information of an organism, its blueprint. The three-dimensional structure of DNA was deciphered in 1953.²³ The double-stranded DNA has a double helix structure resembling a twisted ladder and consists of building blocks called nucleotides. Each nucleotide consists of a sugar-phosphate backbone and one of the four nitrogen-containing bases adenine (A), guanine (G), cytosine (C), and thymine (T). The two DNA strands are held together through the base pairing between the bases. More precisely, adenine forms a base pair with thymine, and guanine with cytosine via hydrogen bonds, known as complementary base pairing. (see Figure 2.1).

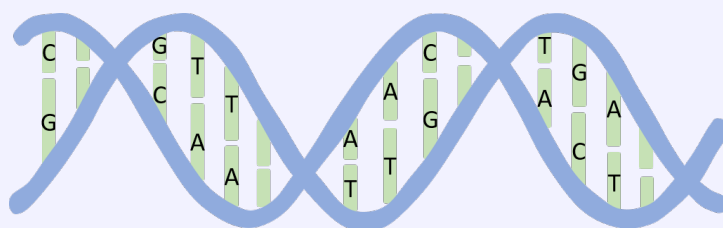


Figure 2.1: Schematic representation of the DNA double helix with base pairing between adenine (A) and thymine (T), and guanine (G) with cytosine (C).

Nevertheless, MLVA alone is insufficient for attribution. Additional information to this end can be achieved by genome sequencing, although costs for such a procedure were enormous at the time of the Amerithrax investigation, and the procedure itself was very time-consuming. A breakthrough was achieved, once morphological variants in the spores involved in the Amerithrax case were noticed. Laborious whole-genome sequencing of the archived culture of the Ames strain produced in 1981 (ancestor strain), together with the isolates featuring morphological variants were carried out. Based on the collected results, variation in the spores' morphology could be linked to genetic mutations and structural variations like insertions, deletions, or duplications, as well as **single-nucleotide polymorphisms** (SNPs²⁴, a genomic variant at a single base position in the genome see Box 2).^{14,20} Using this information, a detailed comparison of genotypic profiling of the isolates from Amerithrax case to the samples collected and characterized during the investigation from geographically spread regions could be performed.¹⁸ For this purpose, a **TaqMan® real-time polymerase chain reaction assay** (type of real-time PCR, see Box 5) was applied¹⁸. This procedure culminated in a clear identification of the source of the Amerithrax spores.^{14,18,20}

Although microbial forensics was a key part of the Amerithrax case, the investigation also used various other methods for gathering intelligence, including traditional forensics and witness interviews. Spores were examined using established chemical and physical techniques for detailed profiling, including C-14 dating.¹⁸ One of the questions in the scope of the case was whether the spores were coated with silicon for greater durability and resistance ("weaponized"), a hypothesis that could not be confirmed based on the **transmission electron microscopy** (TEM) results.²⁵ Additionally, the powder of dried spores retrieved from the letters has been

subject to minute analysis regarding its purity and quality.¹⁸

The Amerithrax investigation contributed to the development of many now-available methods.²⁰ Additionally, many techniques used during the investigation have been greatly improved. A key example is whole-genome sequencing, which has become less expensive and much faster due to the high level of parallelization of the process. This exemplifies the democratization of science. Whole genome sequencing would undoubtedly have been a game changer in the Amerithrax investigation today, although some known limitations — such as the need for an accurate, curated database and more user-friendly bioinformatics tools — still exist (for more details see Section 4). Sequencing helps determine whether the pathogenic strain used in the attack was genetically modified, eliminating natural causes for the incident. Moreover, several DNA microarrays are now available on the market that provide reliable genetic profiling as “a rapid and cost-effective alternative or supplement to sequencing”¹⁴ (although further testing of their robustness and effectiveness may be required in order to use them in court). All these methods must be kept up-to-date because genetic modifications — such as antibiotic resistance, increased virulence, or enhanced environmental stability — can now be more easily introduced through advances in biotechnology.

Box 2. Single-nucleotide polymorphisms (SNPs)

Single-nucleotide polymorphisms (SNPs) are genomic variations in which a DNA building block (nucleotide) at a specific position in the genome is substituted by another (see Figure 2.2).²⁶

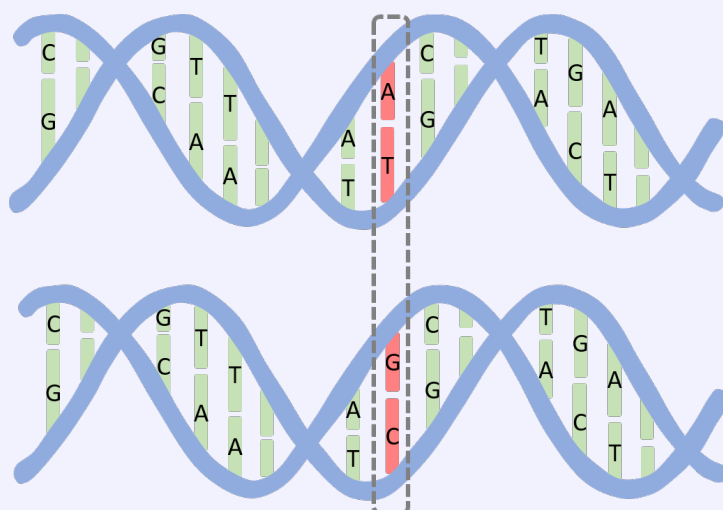


Figure 2.2: Single-nucleotide polymorphism (SNP) in a particular genomic region.

Alongside advanced and reliable methods, scientific expertise in microbiology, immunology, and related fields remains crucial, as demonstrated by the Amerithrax case. It was only through substantial knowledge of the bioagent *B. anthracis* that the morphological differences in the spores could be identified, which later enabled further characterization. This emphasizes that the success of any method relies on the skill and competence of the scientists using it.

2.2 Anthrax outbreak in Sverdlovsk

Another case study reexamines the analysis of *B. anthracis* using microbial forensics, this time within the scope of a large-scale state bioweapons program. The Soviet Union operated the largest secret bioweapons

program known in history, which apparently existed from the 1920s until at least 1992, involving thousands of people and dozens of institutions. During this program, various bacterial and viral agents were characterized and weaponized, including *B. anthracis* (anthrax), *Y. pestis* (plague), *C. botulinum* (botulinum toxin), and others.²⁷ In 1979, an unusual anthrax outbreak occurred in the city of Sverdlovsk (now Ekaterinburg), resulting in at least 68 human victims, although the exact number remains unknown.²⁸ Although Soviet authorities attributed the outbreak to the consumption of contaminated meat, a retrospective scientific analysis published in 1994 demonstrated that a plume of anthrax spores from the military facility was actually responsible for the outbreak.²⁸

Since the initial investigation of the Sverdlovsk incident, analyses of the victims have been conducted to identify the strain of anthrax involved. In 1998, DNA was extracted from tissue samples of 11 victims and examined using PCR. This analysis confirmed the presence of the two plasmids (circular extrachromosomal DNA) that encode the two anthrax toxins and the protective capsule essential for the pathogenicity of *B. anthracis*.²⁹ A characterization of a specific genomic region revealed the presence of multiple bacterial strains. As stated, an explanation for this phenomenon might be that the “victims were exposed to a mixture of different *B. anthracis* strains”²⁹.

In the mid-1990s, several studies investigated whether *B. anthracis*, which was identified in the Sverdlovsk case, was genetically modified. However, as mentioned above, biotechnology has made significant advances since the 90s. Therefore, in a 2016 study, researchers performed deep sequencing — a method of sequencing the same genome region multiple times — on two isolates from the Sverdlovsk outbreak to analyze their genomes.³⁰ No genetic engineering or drug resistance manipulation has been found, indicating that the anthrax strain in Sverdlovsk was the wild-type strain called strain 836 isolated from nature³¹ that was used for industrial spore production.³⁰ The Soviet 836 strain “was extremely virulent, stable in aerosol form, and persistent in the environment”³², in other words, an ultimate biological weapon. Thanks to microbial forensics, some light has finally been shed on the decades-long mystery surrounding the nature of this strain that was part of the Soviet Union’s biological weapons arsenal.

Box 3. Variable number tandem repeats (VNTRs)

Variable number tandem repeats (VNTRs) refer to segments in DNA in which short genome sequences are repeated multiple times (in tandem, see Figure 2.3).³³ The number of such repeats varies between different organisms. Several clusters of VNTRs are called multiple locus variable number tandem repeats (MLVAs).

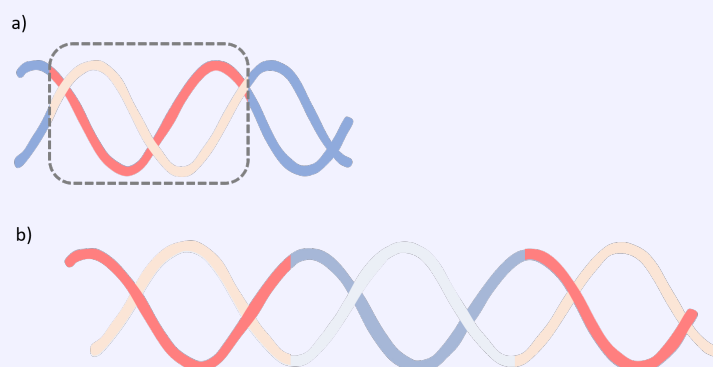


Figure 2.3: a) A cluster of variable number tandem repeats (VNTRs); b) multiple locus variable number tandem repeats (MLVA).

2.3 Confirmation of the presence of a bioweapon program in Iraq

The last case study presented here as an example of microbial forensics capabilities, concerns Iraq's biological weapons (BW) program. Iraq signed the BWC in 1972, once it had been opened for signatures, and ratified the treaty in 1991.³⁴ Despite this, Iraq possessed an extensive BW program since (at least) 1985 that was terminated in 1991.³⁵ Iraq's biological warfare included, e.g., anthrax, botulinum toxin, and aflatoxin. Details about this program (as well as those of the chemical weapons program in Iraq) have been uncovered by the United Nations Special Commission on Iraq (UNSCOM, 1991-1999), established by the UN Security Council. The first attempt to obtain reliable evidence of biological weapons was unsuccessful. However, as technology has improved, biotechnology methods have greatly advanced the investigation. For instance, **polymerase-chain reaction (PCR)** (see Box 4) analysis and new sampling techniques have enabled the detection of *B. anthracis* DNA traces in areas that initially tested negative.³⁶

The unique role of UNSCOM holds broader significance regarding the possibility of verifying the BWC in general. As noted in the publication dedicated to a retrospective analysis of UNSCOM's work, it demonstrated that "a multifaceted verification strategy" can successfully uncover clandestine BW programs.³⁷ Lessons learned from UNSCOM regarding the challenge of detecting activities prohibited by the BWC (and other treaties) due to their dual-use nature remain relevant today. However, due to disagreements among Security Council members about UNSCOM's activities, its mandate ended in 1999 and was replaced by the United Nations Monitoring, Verification, and Inspection Commission (UNMOVIC, November 2002-March 2003), which had similar responsibilities.³⁷

One of UNMOVIC's goals was to gain a more thorough understanding of the extent of Iraq's BW program. Several questions arose during the investigation, such as those posed by Dr. Casagrande, head of the Biological Analysis Lab for UNMOVIC: "Where are several tons of missing growth media? Why doesn't the declaration of biological agents produced match Iraq's production capacity?"³⁸ To achieve this, careful sample collection and analysis were essential. Samples were taken from multiple pieces of equipment believed to belong to the facility involved in bioweapons production. DNA was extracted manually and tested using **real-time PCR** (a PCR-based technique used to quantify DNA or RNA, see Box 5), which was the most common method due to its ease and sensitivity.¹⁴ The dual-use issue encountered during verification procedures related to compliance with international agreements has been vividly illustrated during the work of UNMOVIC. Investigations led to the conclusion that even brewery equipment could, in principle, be repurposed to covertly produce anthrax.³⁸

The search for possible biological agents present was guided by Iraq's submitted declaration and the lists of particular biological agents.¹⁴ In the course of the microbial forensics investigation, several species have been determined using this approach, such as *B. anthracis* and orthopoxvirus.¹⁴ On one of the occasions, testing for biological weapons was highly challenging due to the previous extensive decontamination of the analyzed equipment by Iraqis.³⁶ The samples have been sent to an international reference laboratory, where, despite the undertaken decontamination measures and the burial in the ground for 12 years, biological materials could be obtained and processed, revealing the presence of the DNA of the virulent *B. anthracis*.³⁶ It should be noted that this microorganism features a high environmental stability. Also, DNA is more stable than RNA, which can be subject to rapid degradation due to the omnipresence of RNA-degrading enzymes (RNases). Additionally, a broad-spectrum pathogen identification technology would have been highly beneficial for the process. One such technology currently available is, e.g., **whole metagenome sequencing** (Section 4) that can be applied to detect a wide range of microorganisms without prior knowledge of their identity, including rare, novel, or unculturable pathogens.³⁹ However, it cannot be performed during on-site inspections and requires a properly equipped laboratory and expertise.

Another challenge during inspections in Iraq was posed by protein toxins such as ricin, which is derived from the beans of castor oil plants.¹⁴ Using PCR, it was possible to identify genetic sequences from the castor beans.¹⁴ One goal that cannot be achieved with PCR is determining whether efforts have been made to extract the toxin and whether it was still active or made inactive by heat treatment.¹⁴

Inspections in Iraq faced enormous challenges due to technical limitations in terms of equipment, energy supply, and personnel, as well as reluctance and a lack of cooperation from Iraqi representatives. Nevertheless, they provided valuable insights into Iraq's past biological weapons program despite the difficult conditions and challenges related to collecting samples from partially decontaminated equipment and sources. This example highlights the strength of microbial forensic techniques combined with other investigative methods and demonstrates their usefulness in on-site inspection scenarios. Careful selection of the most user-friendly and sensitive methods, expertise in handling dual-use equipment, and thorough, informed sample collection are crucial in this context. As noted elsewhere³⁶, although science and technology continue to improve rapidly in sensitivity and specificity, efforts to decontaminate and erase traces of previous activities may also become more effective. After this brief historical overview of successful microbial forensic applications associated with biological weapons, the upcoming section will detail established and emerging techniques used in this field.

Section 3

UN Secretary General's Mechanism

BWC is an important cornerstone in the international disarmament regimes. However, despite its pivotal role, the BWC lacks a verification mechanism (see Section 1). Investigations into the use of biological weapons are only possible through the UN Secretary General's Mechanism (UNSGM), established by UN General Assembly Resolution 42/37C in 1987.⁴⁰ In case a comprehensive verification mechanism would become an integral part of the BWC, its technical components can draw on the expertise and knowledge acquired for the investigative and analytical procedures under the UNSGM.

The UNSGM mechanism is not formally linked to the BWC. Instead, it operates under the umbrella of the United Nations Office for Disarmament Affairs (UNODA).⁴¹ A request to invoke the mechanism can be submitted by any UN Member State in case of alleged biological weapons use (or chemical weapons for the states that are not members of the CWC), provided there is credible and convincing evidence supporting the claims.⁴⁰ In such case, the UN Secretary General is authorized to launch an investigation.⁴² The guidelines and procedures including technical aspects for the on-site or near-site investigation and reporting are detailed in UNGA Resolution A/44/561 (1989)⁴³ that were updated in 2007.⁹ Notably, the guidelines with respect to the final report of the UNSGM investigation do not specifically address attribution and do not require the assessment of the responsible actors.⁹ So far, the UNSGM has been used in the context of chemical weapons investigation, but has not been applied for bioagent incidents.⁴⁴

The UNSGM utilizes a roster of experts, expert consultants, and laboratories nominated by the UN Member States.^{40,42} Cooperations with other organisations such as the Organization for the Prohibition of Chemical Weapons (OPCW), the World Health Organization (WHO), and the World Organization for Animal Health are possible. As of now (2025), 637 qualified experts, 129 expert consultants, and 92 analytical laboratories representing different geographic regions are listed in the UNSGM roster.⁴² Analytical laboratories are expected to have high-security facilities, in-depth expertise in handling and analyzing potential biological hazards, and a robust quality management system and security procedures to protect data and materials.⁴¹

To date, several Member States have conducted a series of workshops and practical quality assurance exercises in order to build and improve the expertise required to investigate the alleged use of biological weapons. These initiatives are coordinated by the United Nations Office of Disarmament Affairs (UNODA) and through a Group of Friends of the UNSGM.⁴⁵ A series of such dedicated workshops to strengthen the UNSGM is organized by the Swiss SPIEZ laboratory, where results from a broad range of activities, including “laboratory exercise formats, national approaches to microbial forensics, the analysis of toxins, guidance and documentation”⁴⁶, etc. are discussed and exchanged. During the 2024 SPIEZ workshop, results of the German external quality assurance exercise for laboratories RefBio (including UNSGM roster laboratories) have been presented. RefBio is a part of the UNSGM project funded by the German Federal Foreign Office in order to support “the network and exchange on the UNSGM as well as the readiness of UNSGM-nominated experts”⁴⁷. The German UNSGM project provides, a.o., security and basic training for experts nominated to the UNSGM by Member States,

training on sampling and transport of infectious and toxic substances, a variety of simulation exercises for the UNSGM (e.g., the so-called Capstone exercise 2020-2022 on cooperation with the UNSGM mission team and relevant stakeholders).⁴⁷ The overarching goal of the RefBio project is to establish “a trusted international laboratory network with a broad geographical coverage”⁴¹. It is firmly anchored in the multilateral efforts to strengthen the UNSGM, which are supported by various UN Member States including Australia, Canada, China, Denmark, France, Germany, Sweden, Switzerland, the UK, and the USA.⁴¹

Since its launch in 2017, participation and geographical representation in RefBio has increased significantly.⁴⁶ Training scenarios focus on different biothreat agents (e.g., *B. anthracis*, *Y. pestis*, orthopoxviruses, ricin, etc.).⁴¹ In the recent exercise conducted in 2024, the scenario for the bacterial identification involved uncovering of a clandestine laboratory preparing for a deliberate release of *B. anthracis*. The tasks in the exercise included sample analysis comprehensive characterization of the respective agent (type of bacterial strain involved, its molecular profiling, presence of antibiotic resistance and of virulence genes, etc.), and the analysis of plasmid (small circular DNA) content.⁴⁶ Of the 29 laboratories (21 of them UNSGM roster laboratories) participating in the exercise, only one false-positive and one false-negative result were reported.⁴⁶ Also, the plasmid content was correctly identified by the majority of participants (86-100 %).⁴⁶ In bacterial characterization, there were larger deviations in the quality of the results obtained. The great challenge appeared to be the correct identification of the bacterial strain with the overall success rate of about 20-31 %.⁴⁶

The exercise also provided insights into techniques and methods recommendable to address such specific questions as formulated in the exercise: according to the reported results, Multi Locus Sequence Typing (MLST, a technique for an unambiguous characterization of bacterial isolates, Box 6)⁴⁸ and Single Nucleotide Polymorphism (SNP, see box 2) analysis were more successful than Multi Locus Variable Number of Tandem Repeats (MLVA, box 3) typing for molecular profiling.⁴⁶ With regard to the methods used by the participating laboratories, next-generation sequencing (NGS, Section 4.1) and amplification-based methods were predominant. Also several challenges and recommendations have been identified and formulated⁴⁶:

- improve the interpretation of the NGS results;
- address the current challenges with microbial profiling, strain identification, in-depth characterization;
- address the limited availability of sequencing data in public databases;
- conduct frequent exercises to maintain and steadily increase proficiency of laboratories.

Apart from bacterial agents, exercises to identify and characterize a variety of viral pathogens belonging to orthopoxviruses, *Coronaviridae*, haemorrhagic fever viruses and encephalitis viruses were conducted within the RefBio project between 2018 and 2022. In the last reported exercise (2023-2024), Mpox virus was selected as the target agent. It is noteworthy that the main challenge reported in the context of this task was the identification and characterization of genome modifications (both natural and artificial).⁴⁶ Overall, a high level of competence has been noted in the identification of species, while at the same time deficits have been observed in the identification of strains and variants as well as in the identification of mutations.⁴⁶ The importance to improve genome characterization proficiency “and the ability to detect and interpret signs of genetic engineering, which is key for forensic investigations and attributions”⁴⁶ has been emphasized. Genetic engineering characterization remains an important part of other external quality assurance exercises. For instance, in dry-lab exercises funded by the US Department of State since 2018 one of the important tasks is the identification and characterization of genetic modifications (including detection of artificial sequences, indicators of a strain produced in laboratory, etc.) in the target agents.⁴⁶ The wide range of participants in these dry-lab exercises is noteworthy: a total of 90 laboratories from 40 countries representing all geographical regions took part in the exercise organized in 2024.⁴⁶ The difference between the so-called dry and the wet labs is that in the dry lab environments the focus is on computational approaches and simulations/modeling for research purposes, while wet labs are oriented at procedures involving handling the liquids, chemicals, biological material, etc.

An accurate toxin identification is an integral part of the external quality assurance exercises in the RefBio project. Identifying toxins can pose significant challenges, especially when they are present in low concentrations and/or in a mixture of very similar toxins such as, e.g., ricin isoforms, or structurally similar protein toxins abrin and agglutinins that figured as the biothreat agents in the RefBio exercise conducted in 2019.⁴⁶ A wide range of methods has been utilized by the laboratories participating in the task, including immunological methods (Box 7), mass spectrometry (Section 4.4), and functional assays.⁴⁶ Notably, correct identification of ricin isoforms was only possible with mass spectrometry. The exercise demonstrate the need for further establishment and improvement of expertise and a better sharing of best practices.⁴⁶ Also, a stronger cooperation with the OPCW is required, since biological toxins are covered both by the BWC and the CWC.⁴¹

Any investigation into an incident involving biological warfare agents would face a multitude of challenges, ranging from political aspects to technical capabilities and communication strategies. A recent comprehensive study⁴⁴ provides insights into some of these issues and a roadmap for addressing them. In addition, it lists a range of technologies, including gene sequencing, environmental monitoring, omics-based technologies (e.g., data from transcriptomics and proteomics), and computer-assisted tools that can be used in the investigation process. Some of the methods included in this technology portfolio are discussed in more detail in Section 4. However, as the practice of various approaches to strengthening and supporting the UNSGM shows, the most important and limiting element is not technology, but primarily human expertise and experience.

Section 4

Approaches and techniques used in microbial forensics

The range of methods used in microbial forensics is highly versatile. Many of these techniques have been developed for epidemiology, infection biology, and biotechnology, and can assist in identifying and characterizing the biological threat agents. However, as already mentioned, microbial forensics differs from, e.g., epidemiological analysis due to its operation in the law enforcement and legal context and the resulting requirements and standards. This includes a multifaceted characterization of microorganisms at the focus of the investigative process, such as their specific strain, subtype, abundance, presence of any indications of genetic engineering, etc.,¹⁴ using genomic sequencing, DNA fingerprinting and other procedures. The field has also benefited from the progress in automation of laboratory workflows, bioinformatics tools, etc.⁴⁵ Increasing amount of collected biological data that encompass genomics, proteomics, transcriptomics, metabolomics (the so-called “omics”) transforms the possibilities in microbial data analysis. Due to microbial diversity and the complexity of biological systems, a combination of different approaches is often necessary for reliable data analysis and interpretation. A careful interplay of different methods and techniques, which should be sensitive, specific, and reproducible⁴⁹ is therefore more important for microbial forensics than any individual strategy. Findings from other sources of information should be incorporated into the analysis process. Ultimately, despite the availability of efficient and promising technological tools, human expertise remains the decisive factor that should be continuously evaluated and trained in order to strengthen the field of microbial forensics for all possible future scenarios.

It is not feasible to describe in detail all the methods used or potentially used in microbial forensics including the traditional approaches such as culture and staining techniques.⁴⁹ Instead, we provide an overview of some techniques that have already proven suitable or may be beneficial in this area in the near future. Some methods for nucleic acid analysis, such as different types of PCR, and protein characterization using antibody-based systems, are briefly outlined in information boxes. The reader should also refer to other publications for more details.^{14,49,50}

4.1 Next-generation sequencing

There is no doubt that sequencing techniques are essential in modern microbial forensics because of their high sensitivity and the ability to obtain high-throughput data for detailed microorganism characterization. Whole-genome sequencing enables microbial identification down to the substrain level⁶, a characteristics that became invaluable during the recent global Covid-19 pandemics. However, some other methods like, e.g., microarrays may offer a quick and cost-effective alternative to whole-genome sequencing in certain scenarios. Antibody-based systems (see Box 7) are the preferred choice when genetic information about the biological agent is unavailable. Although sequencing methods are often considered the “holy grail” of microbial forensics, they also have their limitations. These include data interpretation and expertise, validation and standardization,⁵⁰ necessity of a

curated database etc. Some suggestions for validation criteria for sequencing procedures are listed elsewhere.⁵¹

Box 4. Polymerase Chain Reaction (PCR) techniques Basic principle.

Polymerase Chain Reaction (PCR) is a frequently used technique in medical diagnostics and life sciences to amplify target DNA sequences in the sample. It was invented in 1983 by American biochemist K. Mullis⁵² and was awarded the Nobel Prize in Chemistry ten years later.⁵³

It involves a thermal cycling process that includes repeated heating and cooling steps. The first heating stage denatures the double-stranded DNA into two single strands. When cooled, short DNA sequences called primers bind to the complementary DNA sequence of each single strand. These primers serve as templates for the DNA-synthesizing enzyme, polymerase, which is added along with the primers, nucleotides (the building blocks of DNA), and other reagents like buffers to the reaction mixture. During the extension phase, the polymerase synthesizes complementary strands, resulting in two copies of the target DNA sequence. These steps are repeated multiple times to generate a large number of DNA copies for analysis (see Figure 4.1). Today, PCR can be performed automatically using commercially available instruments. Furthermore, advances in molecular biology have made it possible to amplify DNA under isothermal conditions (at a constant temperature of 60°C to 65°C), eliminating the need for expensive thermocycling equipment.⁵⁴

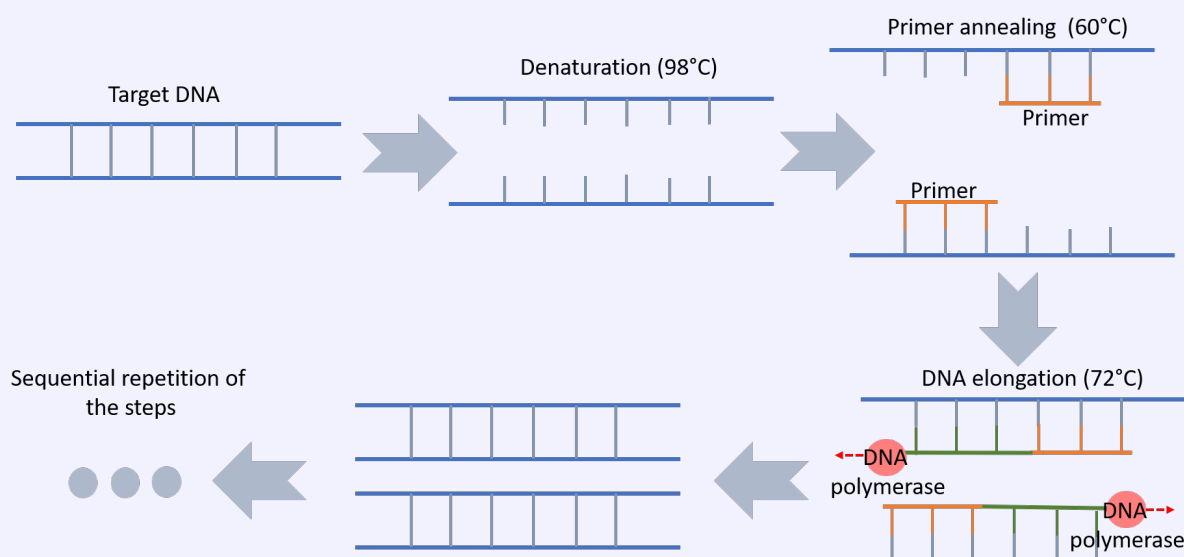


Figure 4.1: Scheme of the PCR procedure required for the amplification of the target DNA (see also⁵⁵). Please note that the temperature values given may vary depending on the primers used, etc., and are provided here to highlight the difference between the steps of a PCR process. For example, the annealing temperature can range from 50 to 65 °C.

The variation of this conventional procedure is multiplex PCR, which allows amplification of multiple DNA regions in a single process using multiple primers.⁵⁶ This advancement is significant for forensic applications (and specifically for microbial forensics), increasing sample throughput.⁵⁶ Furthermore, commercial kits to identify biological threat agents are available that remove the necessity of designing specific primers.⁵⁷

Box 5. TaqMan® real-time PCR

Real-time PCR is based on the conventional PCR technique and allows monitoring of the DNA amplification in the sample during the PCR process in real time due to the detection of the fluorescent signal produced during amplification.⁵⁸ The quantification occurs either using a fluorescent dye that binds nonspecifically to double-stranded DNA (e.g., SYBR® green) or a fluorescently labeled probe (TaqMan®).⁵⁸

In TaqMan® real-time PCR, a dual-labeled fluorogenic probe complementary to the target DNA segment is used in addition to the primers (see Figure 4.2). The two fluorescent dyes are the reporter dye and the quencher dye that suppresses the fluorescence of the reporter dye. The probe binds to the DNA strands close to the primers. DNA polymerase cleaves the probe during the elongation step. This results in the emission of a fluorescent signal. Fluorescent measurement during the PCR process allows quantification of the DNA in the sample.⁵⁹ The quantification occurs using either a standard curve or a comparative quantification.⁶⁰

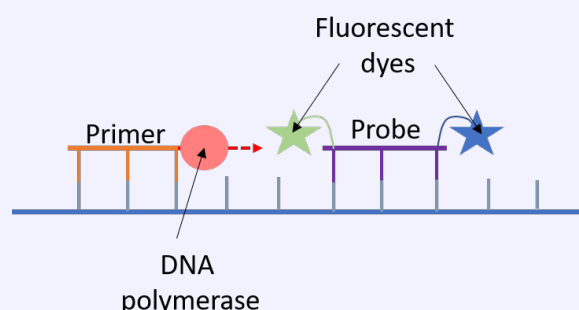


Figure 4.2: Construct for TaqMan® real-time PCR (see also⁶¹).

TaqMan® real-time PCR is a multiplex PCR technique that was used during the Amerithrax investigation (see Section 2.1). Development of real-time PCR assays to detect different *B. anthracis* genetic targets⁶² or biothreat pathogen *F. tularensis*⁶³ has been documented. Additionally, many commercial kits are available for performing real-time PCR in the lab.

One challenge for PCR is the presence of so-called PCR inhibitors — compounds that can be found in the environment (e.g., soil) or blood samples used to detect pathogens, which interfere with PCR and adversely affect its results⁶⁴. Commercial solutions have been created to address this issue. In one study, inhibitor-resistant PCR reagents for direct testing of samples containing biological warfare agent DNA were evaluated.⁶⁴ The evaluation included real-time PCR assays targeting *B. anthracis*, *B. melitensis*, Venezuelan equine virus, and orthopox virus. A similar study focused on the multiplex real-time PCR assays for the detection of *B. anthracis*, *F. tularensis*, and *Y. pestis* in blood samples.⁶⁵

Sequencing techniques have undergone a truly revolutionary process since the method was developed by the British biochemist F. Sanger and his colleagues in 1977.⁶⁶ These advancements led to several milestones in the field, highlighting what are known as the first, second, and third generations of sequencing technologies.⁶⁷ The second generation, also called next-generation sequencing (NGS), has numerous applications in microbial forensics due to its ability to decode genetic features of microorganisms up to the level of metagenomics.⁶⁷ **Targeted amplicon-based sequencing** and **whole metagenome sequencing** (WMS) are among the most promising strategies in NGS for microbial forensics.⁶⁷ The rapidly evolving approaches of third-generation sequencing (TGS), including **single-molecule real-time sequencing** and **nanopore sequencing**⁶⁸, are also

increasingly shaping this field. These techniques are described in more detail below.

- Targeted amplicon-based sequencing

The targeted amplicon-based sequencing aims to facilitate the analysis of a specific genomic region of interest without the necessity of whole-genome sequencing.⁶⁹ Thus, variations in a targeted region can be identified. The conventional workflow in NGS involves: 1.) extraction of DNA/RNA from the sample, 2.) construction of the library for sequencing, 3.) the actual sequencing, and 4.) data analysis.¹⁴ In case of the RNA as the source of genetic material, reverse transcription should be carried out to obtain cDNA. The construction of a library implies fragmenting the genetic material and attaching the adapter molecule to the ends of the fragments, which are then amplified before sequencing.¹⁴ The sequencing usually involves a process analogous to the cellular DNA replication (creating a copy of cellular DNA). Here, synthetic DNA building blocks (nucleotides) are sequentially attached to the target DNA, and each of the nucleotide bases (see Box 1) successfully incorporated in the growing copy of the single-stranded DNA, is detected. In the targeted amplicon sequencing, an additional step of amplifying the genetic region of interest using PCR is added before creating the library (see Figure 4.3).

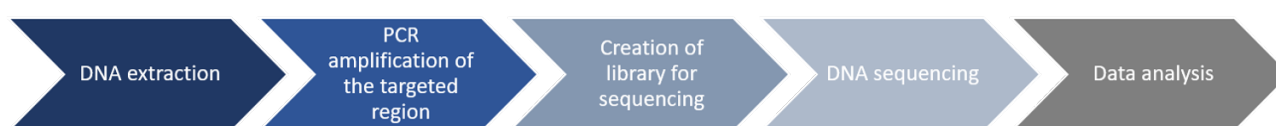


Figure 4.3: A schematic representation of sequential steps in the targeted amplicon-based sequencing.

In the context of microbial forensics, regions of interest that are often targeted include the 16S ribosomal RNA (a component of the small-subunit of a bacterial ribosome, part of the cellular machinery involved in protein biosynthesis), the 18S rRNA of fungi⁷⁰, and the internal transcribed spacer region (a non-coding spacer DNA situated between the small-subunit and the large-subunit of the rRNA⁷¹ genes). The 16S/18S rRNA contains both highly conserved regions and variable regions.^{69,72} The variable regions enable an accurate identification of the microorganism down to the genus level⁷³ due to their unique sequences that differ between species.⁶⁹ The variable sequences of internal transcribed spacer regions, on the other hand, can be used to distinguish closely related fungal species.^{67,69}

The technique of targeted amplicon-based sequencing offers many advantages compared to other sequencing methods. Since only a specific DNA region is sequenced, the method is faster and more cost-effective than whole-genome sequencing. Additionally, it allows reliable and efficient identification and screening of genetic variants while requiring only a small amount of DNA sample⁶⁹ due to the increased detection sensitivity⁷⁴. However, limitations of this approach include possible off-target amplification⁶⁹, the restriction of information to the specific region being examined, and bias towards targeted microorganisms⁷⁴. Other sequencing strategies, which are described below, help overcome these limitations but often come with higher costs and longer turnaround times.

- Whole metagenome sequencing (WMS)

The technical procedure in WMS corresponds to the four-step workflow described above, without the amplification step present in the targeted amplicon-based sequencing. The WMS method differs from the targeted amplicon-based sequencing approach in that it does not focus on a specific region, but provides information about genetic variations across the metagenome (the entire genetic content of any group of microorganisms⁷⁵) of the microorganism in question (see Figure 4.4). Therefore, a much more detailed genetic profile for the identification can be obtained down to the level of species or even strains.^{73,76} It is culture-independent and can be applied for the direct characterization of microorganisms in the samples (food, water, soil, etc.) and provides insights into the virulence potential of the microbial communities, information that cannot be retrieved with the targeted amplicon-based sequencing.⁷³ Furthermore, metagenomics allows the analysis of the genomes of non-cultivable organisms⁷⁷, the simultaneous analysis of various pathogens in a single assay⁷⁴, and the identification

of previously unknown or emergent microbes^{73,78}. All these characteristics of WMS are undoubtedly highly beneficial for microbial forensics.

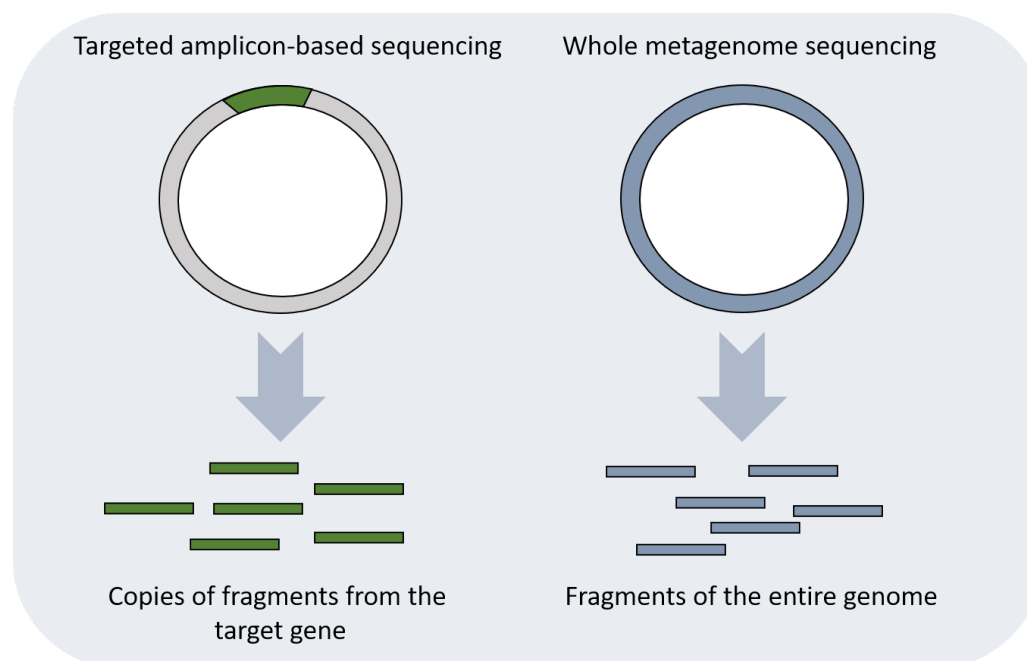


Figure 4.4: Difference between the targeted amplicon-based sequencing and whole metagenome sequencing (WMS)(adapted based on the information⁷⁹).

Besides the benefits listed above, the limitations of WMS should also be examined. For example, one limitation stems from its advantageous feature: the large amount of data generated. However, analyzing vast data sets puts heavy demands on the bioinformatics tools currently used in WMS. One way to address this challenge is to utilize machine learning techniques, such as deep learning, which can greatly simplify the analysis of the collected data.⁷⁴ Deep learning is a subset of machine learning with a more advanced neural network architecture. Still, there is a demand for a curated whole-genome-sequence database that can be used to interpret results.^{73,80} Additionally, the collected data needs adequate and expandable storage space that is readily available. Moreover, WMS remains quite expensive, despite the overall decrease in sequencing costs, and data collection can take days or weeks.⁷³

Other limitations reported in the context of WMS correspond to the information that can be retrieved with this method. Thus, it was noticed that WMS cannot differentiate between viable and non-viable microorganisms.⁷³ This information is important not only in a clinical context but also for microbiological forensic investigations to obtain clear evidence and establish a proper chain of custody. Additionally, the presence of interferences due to genetic material of non-microbial origin in samples has been reported.⁷³ Difficulties have also been faced in attributing genes related to virulence or antibiotic resistance to a specific bacterial strain in the sample⁷³ (differentiation between the pathogen analyzed and the background microbiome), which is crucial for microbial forensics to ensure a complete and accurate characterization of the pathogen.

Despite these known drawbacks of WMS, this technique remains one of the most promising for microbial forensics because of the unique and detailed information it provides about the biological material. In the Amerithrax investigation (see Section 2.1), whole genome sequencing was used as part of the microbial forensics process. Whole-genome sequence analysis can inform all three stages in microbial forensics (identification, characterization, attribution)⁸⁰, but contrary to metagenomics, analysis of genetic variations (such as MLVAs

and SNPs, etc., mentioned above) is limited to a single organism. WMS offers an additional level of information by extending analysis to multiple organisms present in the sample. It is expected that ongoing advances in this field will help develop strategies to address some of the challenges mentioned above.

- Single-molecule real-time sequencing (SMRT) and long-read sequencing

The trend in sequencing continues to move toward a highly parallelized process and the ability to generate long-read sequences. This trend enables the ongoing cost and time reductions in the process. Novel sequencing techniques (third-generation) have entered the market. One of them is the single-molecule real-time sequencing (SMRT) that was introduced in 2015.⁸¹ Currently, there is only one commercial platform (PacBio⁸²) available, utilizing this technique. The uniqueness of this technology lies in the fact that it enables both genetic and epigenetic characterization of the sample. Epigenetics refers to the way the genetic activity is regulated without changes in DNA sequencing.⁸³ The regulatory elements can be chemical groups (e.g., methyl groups $-\text{CH}_3$) attached to the DNA building blocks.

The procedure behind SMRT can be described as follows. DNA is fragmented to create the library, and adapters are added to each end of the double strand to create a circular template.⁸² No amplification step is required. The sequencing occurs on the special commercial chip (SMRT® cell) containing millions of nanostructure wells (called zero-mode waveguides⁸²) having very small detection volumes to detect single molecules with high signal-to-noise ratio.⁸⁴ Each single circularized fragment is immobilized at the bottom of a zero-mode waveguide together with the enzymes for the DNA synthesis.⁸² Due to the circular structure of the DNA template, the sequencing by synthesis occurs continuously, improving the accuracy of the results.⁸⁴ The SMRT utilizes fluorescently labeled DNA building blocks (nucleotides) that emit light, once the nucleotide is incorporated into the newly synthesized strand in the process of sequencing by synthesis.⁸¹ The signal is detected in real time. The emitted signal differs depending on whether the nucleotide base of the target DNA is modified (e.g., contains methylation), or not. It is possible to perform sequencing in two modes: the Circular Consensus Sequencing (CCS)⁸⁵ for highly accurate long reads, or Continuous Long Read (CLR) sequencing for the longest possible reads.⁸²

As mentioned earlier, fragmentation of the DNA is one of the steps in the sequencing process (creation of libraries). The size of the fragments created depends on the specifications of the sequencing platform used. A distinction is made here between so-called short reads and long reads. Generally, long reads can often be obtained directly from the sample containing biological material.⁸² However, long-read sequencing may have lower accuracy compared to short-read sequencing, which can be mitigated to some extent using consensus sequence generation,⁸⁶ and requires more computing resources. The SMRT technique belongs to the category of high-fidelity sequencing methods, delivering very accurate results with an accuracy of 99.9%⁸². Unlike short-read sequencing, where 50-300 base pairs (bp, unit for measuring genome length) DNA fragments had to be copied and assembled before they could be analyzed⁸¹, long reads in SMRT can be of 15,000 to 20,000 bases or more in length⁸², and can be used to reconstruct and sequence the genome in significantly less time. The importance of the long-read sequencing was recognized by awarding the technique the title of 2022 Method of the Year by the journal Nature Methods.⁸⁷

One advantage of this technical improvement is the possibility to identify large, complex variants more easily.^{81,82,85} Short-read sequencing is less applicable for tasks such as the detection of structural variants.⁸⁵ Additionally, as mentioned above, epigenetic profiling is possible with SMRT technology. As stated elsewhere⁸¹, the development of epigenetic profiling technology has far surpassed its use in microbial forensics. Nevertheless, epigenetics can provide more detailed information to determine a subpopulation of a microbial sample⁸¹ and also offer some other investigative approaches (e.g., identifying the age of the person(s) who left the genetic sample)⁸¹. In the future, further applications of epigenetics for microbial forensics might be identified.

- Nanopore sequencing

Another third-generation sequencing technique that utilizes long genomic reads is nanopore sequencing to characterize single long DNA and RNA molecules.⁸⁸ Similar to the SMRT method, this method does not

require amplification of genetic material. Here, we will focus on the Oxford Nanopore Technology (ONT).⁸⁹ The platform features a cell containing an array of protein nanopores embedded in an electrically resistant membrane⁸⁹. Sample preparation involves denaturing double-stranded DNA or RNA to produce a single strand. For RNA samples, reverse transcription is not required (in contrast to the SMRT).⁸⁶ A measurable disruption in the ionic current of the membrane occurs when the single-stranded DNA or RNA (with specific adaptors) passes through the nanopore in a controlled manner.^{86,89} This signal is then decoded using specialized algorithms and machine learning models⁸⁶ to determine the genomic sequence in real-time (see Figure 4.5). One of these recently released algorithms allows, e.g., accurate epigenetic insights into the analyzed biological sample.⁹⁰

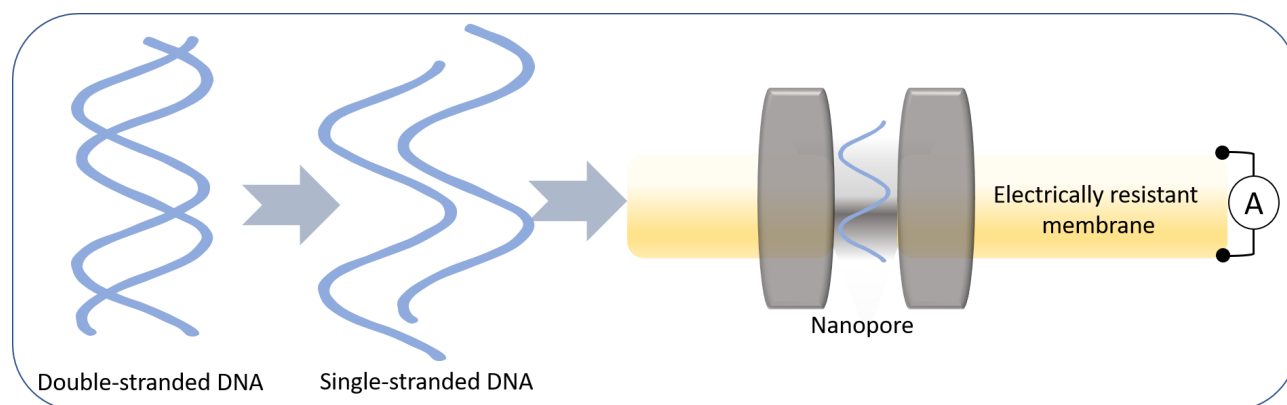


Figure 4.5: During nanopore sequencing, double-stranded DNA is converted to a single strand. Once the single-stranded DNA/RNA passes through a nanopore in the electrically resistant membrane, a change in the electrical current is measured (adapted based on the information⁸⁶).

Genome assembly⁹¹, structural variant detections⁹² and even complex metagenomic analysis⁹³ are possible with the nanopore sequencing. This technique has been reported to be applied for the whole genome sequencing of *B. anthracis*. Long-read sequencing enabled the pathogen-specific DNA sequence from the culture isolate to be detected after just 3.5 hours, although the genome assembly was error-prone compared to the results from short-read sequencing⁹⁴. Nevertheless, the method would deliver valuable information on the antibiotic resistance genes and indications of genetic engineering in a rather short period of time (about 8.5 hours).⁹⁴

In general, ONT nanopore sequencing is mentioned to have a higher error rate than other sequencing methods,⁹⁵ such as next-generation sequencing and SMRT sequencing.⁸⁶ However, various computational approaches exist to improve its accuracy.⁹⁵ Additionally, portable devices (MinION®) for in-field work are available, which promise to provide high-quality, low-cost sequencing.⁹⁶ Such devices have already been successfully applied for disease outbreak surveillance, such as the real-time genomic surveillance of the Ebola outbreak in Guinea in 2015⁸⁸, or the genomic surveillance of Zika virus⁹⁷, yellow fever virus⁹⁸, and Dengue virus⁹⁹ in Brazil.⁸⁸ Furthermore, such portable devices could be applied for metagenomic on-site characterization of pathogens, the presence of virulence genes, and antimicrobial resistance markers in the environmental setting.^{88,100} Recently, however, concerns have been raised about the cybersecurity of such portable sequencers, as they rely on external systems, creating a vulnerability that can be exploited by adversaries.¹⁰¹ A recent study analyzed vulnerabilities in Oxford Nanopore's portable sequencing devices, such as the circumvention of security measures and the installation of malware, which can potentially lead to unauthorized access to data and control over the sequencing device.¹⁰¹ The study was published after the issues had been addressed and mitigated by the manufacturer to avoid information risks. In general, cybersecurity is becoming increasingly important in the field of biotechnology, particularly due to the growing intersection between biological and digital domains.¹⁰²

Box 6. Multilocus sequence typing (MLST)

Multilocus sequence typing (MLST), was first described in 1998.⁴⁸ and throughout the years became one of the most commonly used bacterial typing methods¹¹. It is a multistep procedure that contains isolation of bacterial DNA; PCR amplification of housekeeping genes (genes that are essential for cellular processes and maintenance); sequencing; creation of the allelic profile (alleles are versions of genetic sequences at a given genomic location¹⁰³) of the sequenced data; comparisons to the data collected in dedicated open-access, curated database¹⁰⁴ for bacterial identification.¹⁰⁵ While in general seven housekeeping genes are used in the MLST analysis, optimization such as an augmented MLST method has been developed using 20 genes (or 21 genes¹⁰⁶) referred to as MLST-seq¹⁰⁷.¹¹ Additionally, automated workflow for data analysis (on the scale of minutes instead of hours) has been proposed.¹⁰⁵

4.2 Further emerging technologies used in microbial forensics

The new generations of sequencing methods have undoubtedly caused a fundamental shift in forensic applications. However, many other technologies are now emerging that enable rapid and efficient collection of biological samples and facilitate data analysis. These advances aim to reduce human involvement to minimize errors and to increase the reproducibility and reliability of results through streamlined workflows and automation. These include methods for automated DNA extraction, Rapid DNA analysis, Mobile DNA analysis, and 3D spatial DNA profiling. Many of these technologies are used in traditional DNA forensics during criminal investigations, but they can also be valuable tools for microbial forensics, especially in cases involving alleged biological weapons. Despite their great potential, challenges remain, such as the need for validation and standardization of procedures, issues in interpreting complex biological data, and others. Some of these new technologies are discussed in detail elsewhere.⁵⁰ Here, we briefly highlight their key applications in microbial forensics to provide insight into the rapid technological progress in the field of DNA forensics, which has transformed significantly since Amerithrax (Section 2.1).

Microbial forensics requires meticulous and high-standard sample collection to ensure chain of custody and result accuracy. This involves avoiding contamination and degradation of genetic material during crucial steps of DNA extraction and processing. Several commercially available kits for manual DNA extraction and purification have been developed for this purpose. Additionally, advancements in laboratory-based **automated extraction systems** are increasingly adopted⁵⁰ in various laboratory set-ups, suitability for microbial forensic tasks. These systems improve reproducibility and consistency, reduce cross-contamination risks, and greatly cut down the time needed for DNA extraction compared to manual methods.⁵⁰ Examples include magnetic bead-based kits for automated sample preparation, compatible with various biological targets such as DNA, RNA, proteins, and complete cells.¹⁰⁸ Another multiplex cartridge-based assay capable of sample processing and amplification is the Cepheid GeneXpert® System¹⁰⁹ that has been demonstrated to rapidly detect three biothreat agents: *F. tularensis*, *B. anthracis*, and *Y. pestis*.¹¹⁰ Commercial kits to extract DNA from a variety of sample matrices (blood, tissue, water, soil, etc.) have been developed.⁴⁹ Besides laboratory-based solutions, miniaturized, field-deployable automated DNA extraction systems have been introduced, often utilizing microfluidic technology to handle small amounts of samples.⁵⁰ Advances in DNA preservation technologies help to maintain sample integrity, including improved storage containers with stabilizers that absorb moisture and inhibit microbial growth.⁵⁰ Longer-term DNA archiving methods have also advanced, with innovations like DNA dry-state storage, which stores air-dried DNA samples at room temperature in a stable form.^{50,111}

Since the genetic material needed for comprehensive analysis in investigations is often in very low quantities, techniques like more sophisticated PCR protocols¹¹² and kits (such as GlobalFiler™ PCR amplification kit (Thermo Fisher™)) can facilitate **Low-Copy Number (LCN) DNA analysis**. However, this approach can produce artifacts that complicate interpretation, and impurities may also affect results.⁵⁰ To address this, post-PCR clean methods have been suggested such as the amplicon RX post-PCR clean-up kit, which has been evaluated for improving trace DNA profile recovery from samples amplified with the GlobalFiler™

PCR amplification kit.^{50,113} Quantification of LCN DNA can be performed with methods like **digital PCR** (dPCR).⁵⁰ In the process of dPCR the sample is separated into multiple independent partitions that each act as an independent PCR microreactor. The concentration of the target in the sample is determined using fluorescence detection. Unlike real-time PCR (Box 5) that is also used for quantification, dPCR allows for single molecule resolution and absolute quantification.^{114,115} Additional approaches such as reverse complement PCR (RC-PCR) can be used to amplify degraded DNA.¹¹⁶

Rapid on-site analysis of samples can be decisive for forensic investigations. For this purpose, **Rapid DNA technology systems**⁵⁰ have been developed offering a complete on-site workflow, including data analysis, in a portable “lab-in-a-box.”⁵⁰ These platforms often combine microfluidic technology, PCR, and capillary electrophoresis.⁵⁰ While primarily designed for forensic DNA profiling, they hold potential for future microbial forensic applications. However, these platforms are optimized for speed and may struggle with complex samples¹¹⁷, a limitation that can be addressed by **Mobile DNA analysis systems**.⁵⁰ These compact devices offer also a possibility of performing the entire workflow including sample extraction, amplifications and analysis in field conditions, but at the same time are designed to process a broader range of sample types compared to Rapid DNA analysis systems.¹¹⁷ Some of these systems utilize Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) and the associated protein Cas, CRISPR/Cas system,¹¹⁷ a groundbreaking technology for genome editing in molecular biology, often referred to as “genetic scissors”. CRISPR/Cas-based techniques have shown to be effective in pathogen detection,¹¹⁸ making these systems of interest for microbial forensics applications.

Advancements in **electrochemical sensing technology** to detect biothreat agents combine several advantages such as sensitivity, portability, cost-effectiveness, and ease of use.¹¹⁹ The basic principle behind these technologies is measuring change in physical parameters such as current, potential, conductance, resistance, or impedance due to a biochemical reaction and converting these measurements into the detection signal.¹¹⁹. One of the examples for such a platform is electrochemical sensor for the detection of *B. anthracis* incorporating graphene oxide-modified gold nanoparticle electrode.¹¹⁹ Furthermore, tested platforms for electrochemical sensing of *Y. pestis* revealed a significant advantage over conventional immunoassay technology (s. Box 7), due to its rapidity and comparable or better sensitivity.¹¹⁹ Another example is the portable electrochemical detection platform for the rapid and sensitive on-site identification of six biothreat agents (*B. anthracis*, *Y. pestis*, *F. tularensis*, *Brucella spp.*, *B. mallei*, orthopox viruses) in the automated manner.¹²⁰ Other platforms featuring electrochemical sensing technology to detect further biothreat agents including *Botulinum* neurotoxin, Ebolavirus, Marburg virus, yellow fever virus, etc., have been presented elsewhere.¹¹⁹

Another cutting-edge methods that finds application in forensics is **3D spatial DNA analysis** which characterizes a three-dimensional DNA structure in its biological context (e.g., in cellular environment).⁵⁰ This approach is especially useful for analyzing complex mixtures, as it provides information about the interactions between genetic material and its environment, helping to reconstruct the sample’s history — a valuable asset in investigative process.⁵⁰ One of the tools for this functional genomics is the so-called Hi-C. This method allows to map the three-dimensional genome architecture by combining techniques to detect interactions between molecules and deep sequencing.¹²¹ However, applying such methods requires careful sampling, the availability of specific equipment and expertise.⁵⁰ Additionally, the interpretation of data might be challenging due to the complexity of biological systems and their interactions. Computational approaches based on machine learning have been developed to facilitate such data analysis.^{50,122} Monitoring and evaluating these innovations are essential to determine how best to incorporate them into microbial forensic workflows. Advanced software can significantly assist in managing the complex data analysis needed for reliable results in investigations. Some promising computational tools for microbial forensics are briefly discussed in the next Section.

4.3 Genetic engineering attribution using computational tools

Efficient computational methods are crucial in many areas of daily life. Microbial forensics also requires robust and resilient software. Sequencing techniques discussed earlier depend on advanced bioinformatics tools

and algorithms. They enable complex analysis and interpretation of sequencing data, such as metagenomic profiling¹⁰⁰. Recently, several machine learning-based tools have been introduced into the areas of next-generation sequencing⁷⁴ or in nanopore sequencing¹²³. Also, other technologies such as 3D genomics, analysis of LCN DNA, and genome profiling in general benefit from AI-driven bioinformatic tools. Recently, specific microbiomes could be identified using computer image analysis technology combined with machine learning and deep learning approaches.^{124,125} However, when selecting suitable software, the limitations of AI-based tools should be considered. These include the need for curated, high-quality, labeled data required for training and testing the models, the risk that AI algorithms can deliver “hallucinated” results (factually fabricated content), and the lack of interpretability of these algorithms (“black box” characteristics of AI). AI algorithms need large amounts of curated data to produce meaningful and reliable results. In connection with supporting attribution, larger amounts of publicly available data are also required to enable a broader geographical and historical scope for data analysis. Also, a curated forensic microbial database and uniform standardization of sampling and sequencing procedures¹²⁵ are necessary to enable better data analysis in microbial forensics. Some of these limitations might be overcome in the future. In addition, further computational opportunities are emerging on the horizon. Looking ahead, quantum computing — still in its early development stages — offers the potential to greatly enhance sequencing data collection and analysis. Recent advances include using the so-called Quantum Images framework to accurately detect subtle genetic variations.¹²⁶

In following, we will focus on computational tools that can be used to facilitate the main task of microbial forensics, the attribution, more specifically, genetic engineering attribution. According to Lewis et al., sources informing attribution can be divided in three categories: non-technical indicators that might provide insights regarding the intent; intelligence; technical forensics.¹²⁷ Genetic engineering attribution falls into this last category. It can be understood as the ability to identify the “designer” of the modified biological sequence.¹²⁸ As mentioned, an indication of genetic engineering is an important clue for ruling out the natural occurrence of the biothreat incident. A detailed analysis of the “design” behind the genetic modifications can help link the sequenced sample to its country of origin and/or the original laboratory.¹²⁹ In the Amerithrax case (Section 2.1), identifying specific variations in certain regions of the *B. anthracis* genome ultimately allowed for pinpointing the biological source of the spores used in the attack. Although no genetic engineering was involved, this example illustrates how genome characterization can narrow down investigations to a particular lab.

Some ideas on identifying signatures of potential bacterial genetic engineering are discussed elsewhere.¹³⁰ An approach to detect signatures of genetic engineering even at a single-cell level has been reported.¹³¹ It was applied in the wake of the COVID-19 pandemic to address the question whether the virus was lab-engineered.¹³² This tool has been reported as applicable more broadly to identify biological weapons.¹³² However, despite their promise, such tools cannot identify the source of modified sequences without additional methods.

In principle, someone involved in the genetic engineering of a pathogen is seemingly presented with a myriad of options to achieve the desired outcome, although the constraints posed by the biological system itself have to be considered. However, the choice of a particular option is influenced by very specific factors, including the expertise level of the actor(s) and resources available, resulting in an individual “methodological signature”.¹²⁷ These fingerprint-like traits support the complex process of attribution. The selection of certain biochemical components (e.g., promoters, see Box 4), applied methods, choice of targeted genomic regions, etc., all leave a very unique pattern in the design of genetically modified constructs, similar to “clues in [computer] code”.¹²⁹ Computational tools, such as e.g., machine learning, in particular deep learning, can facilitate screening of the large amounts of highly complex sequencing data to assign these identified signatures to the lab of origin. As pointed out elsewhere¹³³, microorganisms grown under controlled lab conditions would have less genomic diversity compared to the species found in nature. Recording indicators such as the origin of the sample (laboratory vs. nature), genetic characterisation (stable genetic elements in terms of mutation rates), storage conditions, specific treatments (e.g., to induce mutagenesis), etc. contributes to a more accurate data interpretation.¹³³

The accuracy of developed computational attribution tools has improved substantially from about 50 %¹³⁴ to

over 70 %¹²⁹ attribution accuracy. Biological data used to train and evaluate the model included frequently recurring DNA motifs and a set of observable traits of an organism (phenotype¹³⁵).¹²⁹ Large non-profit plasmid repository¹³⁶ served as a source for some of these data and as a platform to test the deployment scenario for genetic engineering attribution.¹²⁹ The nation-of-origin prediction relevant in the context of the BWC can be achieved based on the metadata and on manual annotations provided within such a repository.¹²⁹ However, as pointed out elsewhere¹³⁷, machine-learning approaches have some limitations, including, e.g., a lack of explainability, high computational costs, and the reliance on the training data set to produce meaningful results. Therefore, alternative approaches should also be explored. For instance, an open-access algorithm based on the pan-genome (collection of various genomic regions that are either common or unique to research labs)¹³⁷ comparative process (alignment), combined with a lab score correction technique, was reported to have a higher accuracy (76 %)¹³⁷ than the machine-learning tool¹²⁹ mentioned above. This algorithm could be successfully applied for “predicting the lab-of-origin of previously unseen, newly engineered sequences”¹³⁷. This model can be easily extended to add new annotated sequences, is fully explicable, and requires only a few pre-processing steps.¹³⁷

The field of computational tools for genetic attribution is quite dynamic, as demonstrated by the recent competition for the best genetic attribution algorithms¹²⁸. Several approaches presented there have largely outperformed the previously reported models.¹²⁸ The prize-winning solutions included deep learning algorithms and an approach that was not based on artificial intelligence.¹²⁸ These developments in the area of genetic attribution are undoubtedly very promising. Further models and approaches can be envisaged to emerge in the near future. Also, combinations of various models might further help to push the boundaries of the individual approaches and validate findings.

Still, computational tools do not replace other forensic methods and knowledge collection for an unambiguous identification of the actor(s) responsible for (or planning) a biothreat incident. Other techniques are also required to avoid, at any cost, false-positive results. Additionally, genetic engineering attribution might face challenges with advances in synthetic biology and biotechnology, as the signatures required for attribution might become more obscured. Also, scenarios of intentionally creating false red flags to frame a specific institution and cover own malicious activities are feasible.¹³⁸ Nevertheless, these tools can provide valuable information that may support further investigation processes, though the models need to be improved to surpass their current limits: e.g., increasing their accuracy and expanding their scope to entire viral and bacterial genomes¹²⁹ (although it should be noted that microorganisms vary strongly in terms of their genome sizes). For this purpose, larger and more balanced datasets are required.¹²⁹ Some suggestions on how to improve genetic attribution have been published recently.¹³⁹ One of the recommendations is, e.g., to include engineered sequences from high-containment laboratories (Biosafety Level 3 and higher) in the data set in order to train computational systems.¹³⁹ However, databases or repositories containing sensitive data on pathogenic organisms should be equipped with sophisticated security measures including cryptographic methods to prevent potential data misuse or theft by malicious actors. Systems such as distributed ledger technology (DLT) can provide the necessary level of digital security through regulated data access and tamper-proof records. This approach has already been applied in various industrial branches where data integrity, controlled information exchange and information storage without the possibility of manipulation by third parties are desired. Use cases for the application of DLT in high-containment laboratories have recently been published.¹⁴⁰ Some experts therefore point out that computational attribution today can in principle work as an investigative lead, or as a tool to follow-up on another lead to confirm that a source identified is likely to be the real one. “But unless we know we have covered a fair portion of all possible sources worldwide, we have no idea whether we had a false-positive.”[‡]

Another aspect that might gain importance in the context of genetic engineering attribution is the usage of a so-called cryptographic approach to authenticate synthetic DNA sequences.¹⁴¹ The basic idea is to convert text information containing the specification of ownership into short genomic sequences by substituting letters of the alphabet with nucleotide sequences.¹⁴² This encrypted information is subsequently embedded into the target DNA, creating a specific watermark.¹⁴¹ The approach can be further combined with developed encryption

[‡]From conversation with Dr. R. Trapp, an independent consultant on chemical and biological weapons arms control and security.

protocols¹⁴¹ to gain an additional level of security and to camouflage the embedded information in the genome. Additionally, several digital signature techniques (analogous to ones used, e.g., in banking systems) for synthetic DNA have been introduced to further improve security and authenticate the origin of synthetic DNA. Thus, it is possible to verify that the sequence has not been altered and to confirm its original attribution for microbial forensics purposes.¹⁴¹ This level of authenticity is lacking in the genetic engineering attribution approaches presented above. Encrypted DNA signatures can help solve cases of possible theft and misuse of pathogen- or toxin-encoding genetic material from facilities that work with these biomaterials.

In this context, it is recommended to make the introduction of the DNA signature a widespread procedure¹⁴¹ in both industry and academia. Therefore, the absence of such authenticating marks would raise suspicion and justify further investigation. However, it should be noted that for all these tasks, there is a wide range of approaches with varying levels of security and reliability. Careful consideration is required when selecting a specific approach for verification purposes. Additionally, designing watermarks for complex biological systems without compromising their vitality can be challenging. There is also a risk of mutagenesis in the watermark region, which could affect the integrity of the original sequence.¹³⁸ More information regarding DNA barcoding will be provided in the upcoming manuscript. Further techniques and methods are available for the detection of pathogenic genetic material or proteins. Some of them are described in this section. However, the list is not exhaustive, merely offering a brief overview of available methods in microbial forensics.

4.4 Ribosomal ribonucleic acid (rRNA) detection

The ribosomal RNA (rRNA) is a useful target in pathogen detection assays due to the high number of copies of rRNA present in the cells, potentially reducing the time required for the amplification step.⁵⁴ Furthermore, rRNA genes are universal among bacteria and exhibit both conserved and variable regions.¹⁴³ In particular, the 16S rRNA mentioned above presents a popular target in the bacterial identification process.¹⁴³ High-throughput genotyping methods to analyse SNPs in 16S rRNA genes are available.⁶ According to Schmedes et al., the 16S rRNA gene is the most commonly used bacterial genetic marker for the general identification of bacteria.¹¹ However, a high degree of conservation can also be a drawback in bacterial identification, making it challenging to design species-specific probes.¹⁴³ In such cases, careful considerations regarding the choice of target (e.g. housekeeping genes as genetic markers) are necessary. One recently presented multiplex approach that does not solely rely on 16S rRNA is the so-called Germlex panel that targets species-specific and virulence-related genes in various bacteria, including bioterrorism agents utilizing PCR technology combined with massive parallel sequencing (or next-generation sequencing, NGS).¹⁴⁴

In addition to sequencing methods, several platforms have been developed to detect rRNA.¹⁴⁵ One involves a multi-step process consisting of three phases: target capture, amplification, and hybridization protection.¹⁴⁵ During target capture, short genomic sequences (oligonucleotides) attached to magnetic beads bind to the target RNA. These constructs of the capturing oligonucleotides and of the target RNA are entrapped in the tube using a magnet, while all interfering substances are washed out.^{54,145} Next, isothermal amplification occurs (see Box 4). This process uses both RNA polymerase (which synthesizes RNA) and reverse transcriptase (which synthesizes cDNA from RNA), increasing the assay's versatility. Finally, a specific probe labeled with a detection molecule binds either to the non-amplified rRNA or the amplification product, resulting in a detectable chemiluminescent signal.¹⁴⁵ Incidentally, there is a similar approach that focuses on antigens instead of nucleic acids, known as immunomagnetic separation. There, magnetic beads with immobilized antibodies can bind to a variety of targets to purify and concentrate biothreat agents in the sample.⁴⁹

This technology can detect messenger RNA and DNA and can be performed fully automatically. The assay has been primarily developed to detect food pathogens, including *Listeria*, *Salmonella*, and *Escherichia coli*¹⁴⁵, however, it can be applied for microbial forensics. Historical cases include the deliberate use of food pathogens for bioterrorism or biocrime. In one well-documented case, representatives of the religious Rajneeshee cult deliberately contaminated salad bars in Dallas, USA, with *Salmonella enterica* that causes inflammation of the

gastrointestinal tract, to incapacitate the local population and influence the outcome of the local elections.¹⁴ On a side note, this case also illustrates challenges that face microbial forensics in determining whether the pathogenic spread was natural or deliberately caused.¹⁴

Box 7. Immunoassay technology on the example of enzyme-linked immunosorbent assay (ELISA)

The principle behind antibody-based pathogen detection relies on the interaction between antibodies and antigens. The immune system typically produces antibodies in response to specific pathogen molecules. Immunological methods like enzyme-linked immunosorbent assay (ELISA) use synthetic antibodies to identify and measure pathogen-specific antigens. Various types of ELISA exist: direct, indirect, sandwich, and competitive, depending on the target. The sandwich ELISA, often used for antigen detection, first involves immobilizing an antibody against the specific antigen on a test plate. The sample containing the antigen is added; if present, the antigen binds to the immobilized antibody. A second enzyme-linked antibody then binds to the captured antigen, creating a “sandwich” structure. A substrate is added, which the enzyme converts into a measurable signal to identify and quantify the antigen (see Figure 4.6).¹⁴⁶

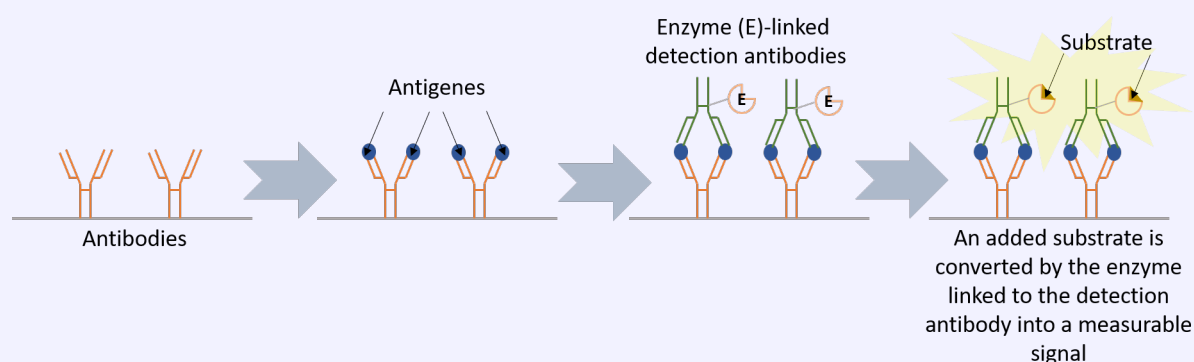


Figure 4.6: Principle behind sandwich ELISA (see also¹⁴⁷).

Immunoassay-based microarray platforms can be used for high-throughput testing. Portable devices employing immunoassay technology for rapid, automated on-site detection of biothreat agents have also been developed. Some of the available platforms are reviewed elsewhere¹⁴⁸. Also, protein toxins such as ricin or botulinum toxin can be detected using such platforms.¹⁴ The ability to produce rapid results is one of the main advantages of antibody-based detection techniques. It should be noticed, however, that for microbial forensics, they might lack the required level of specificity in order to discriminate beyond the level of species.¹⁴ Still, they present a useful alternative to genome-based methods for the first and rapid characterization of the biothreat agent(s).

Nowadays, microarray chips can be used to detect rRNA in a precise and sensitive manner.¹⁴³ For instance, 16S-rRNA-based microarrays that allow multiplex analysis of various pathogens have been developed.¹⁴⁹ Microarray chips are widely used in research and diagnostics. In general, they contain probes (short genomic sequences, proteins, sugars, antibodies, etc.) immobilized in the wells of a microarray chip.¹⁴³ Therefore, microarrays can detect both nucleic acids and protein signatures, broadening their applications in microbial forensics. Microarrays to detect biothreat agents have been described in literature.¹⁵⁰ As pointed out elsewhere¹⁴, this technology does not achieve the level of information provided by sequencing techniques, but offers a rapid and cost-effective alternative. One of the previously developed microarray platforms targeting the 16S rRNA was stated to “detect and identify thousands of biological species in a matter of hours”¹⁵¹, including biothreat agents, illustrating the applicability of such technology for microbial forensics.

Besides rRNA, microarray chips can be used, e.g., to characterize genomic variations such as single-nucleotide polymorphisms (SNPs, Box 3), variable numbers of tandem repeats (VNTRs, Box 3), or multilocus sequence typing (MLST, Box 6), or identify antibiotic resistance genes¹⁴³, among other applications. For example, a DNA microarray has been developed for detection of up to 90 antibiotic resistance genes in gram-positive bacteria such as *Staphylococcus* and *B. anthracis*.¹⁵² Furthermore, technology also allows the identification of artificially constructed pathogens.⁶ A detailed review of microarray chips and their forensic applications is beyond this paper's scope; the reader is advised to consult additional sources.¹⁵³

4.5 Mass spectrometry to analyze biomolecules

In principle, a wide range of biomolecules can be analyzed as part of a microbiological forensic process. Mass spectrometry, a highly sensitive analytical method (Box 8), plays a crucial role in this context.

Box 8 Mass spectrometry: basic principles and MALDI-TOF.

Mass spectrometry is a technique used to detect and analyze ionized molecules both qualitatively and quantitatively in the gas phase. Molecules are ionized, accelerated by electric and/or magnetic fields, and separated based on their mass-to-charge ratios.¹⁵⁴ Ionized molecules reach the detector at different speeds depending on their mass, with lighter ions of the same charge arriving faster than heavier ones. The resulting spectrum, which plots the mass-to-charge ratios, can be compared to a large database of molecular compounds, allowing for the unambiguous identification of molecules in the original sample in a fingerprint-like manner. Several search engines are currently used for computer-assisted data analysis in mass spectrometry. In proteomics (protein analysis), e.g., the bioinformatic tool Mascot¹⁵⁵ is widely recognized as the reference standard for protein identification.

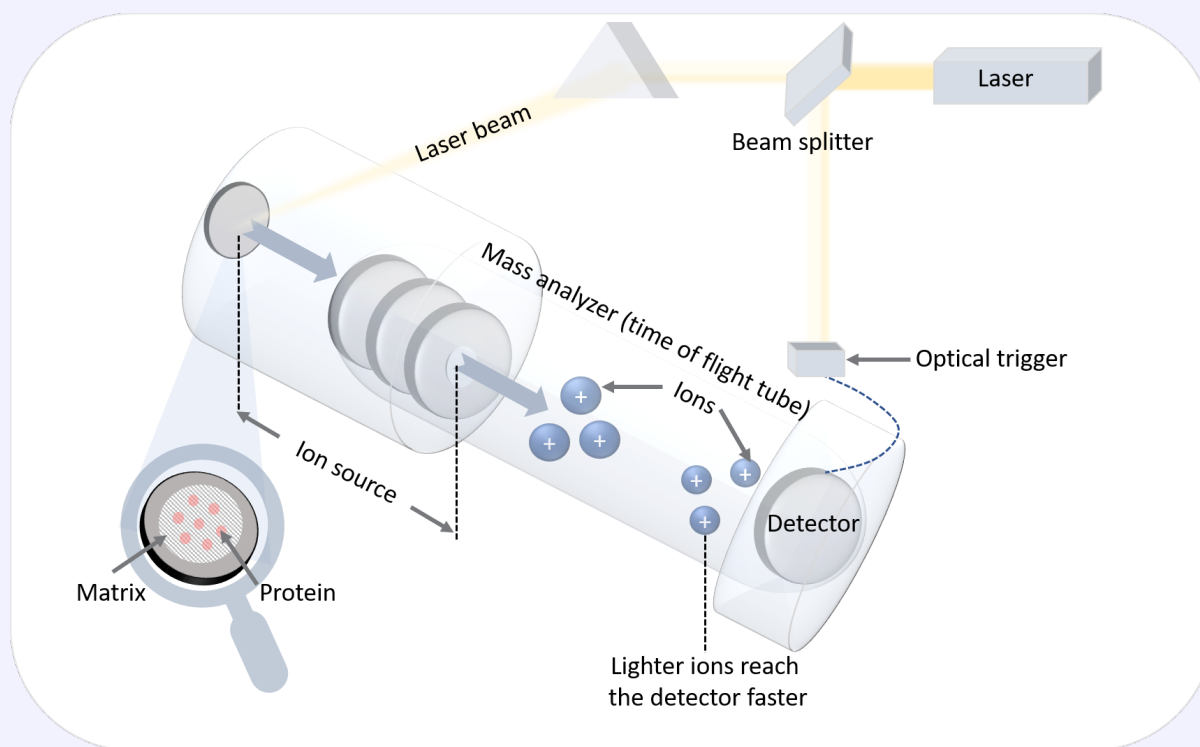


Figure 4.7: Matrix-assisted laser desorption-ionization (MALDI)- time of flight (TOF) spectrometry (see also ¹⁴⁷).

A variety of different instruments and techniques¹⁵⁶ are available in mass spectrometry, providing comprehensive information about the analyzed sample. However, the basic architecture of mass spectrometry devices consists of an ion source (for ionizing the molecules), a mass analyzer (where the molecules are separated according to their mass-to-charge ratios), and a detector.^{156,157} Due to its high sensitivity and specificity, mass spectrometry has a wide range of applications in medical diagnostics, basic research, and forensics. It is commonly employed to analyze proteins in samples and is therefore a key tool in microbial forensics for analyzing pathogenic proteins. However, proteins are non-volatile compounds, making vaporization into the gas phase difficult. To address this, methods such as matrix-assisted laser desorption-ionization (MALDI), or electrospray ionization (ESI), have been developed, each with its specific advantages and limitations.^{156,158} In a MALDI-TOF setup, proteins are embedded into a specific matrix and ionized with a laser beam, causing the matrix to evaporate and releasing the sample into the gas phase (soft ionization).¹⁵⁷ The resulting gas-phase ions enter the mass analyzer. One common type of mass analyzer is based on the time of flight (TOF) measurement, where the time it takes for the ion to reach the detector is measured. This timing starts when the laser beam ionizes the sample (Figure 4.7). The arrival time depends on the velocity of the ion, and therefore on its mass-to-charge ratio.¹⁵⁹

Mass spectrometry can be combined with other analytical methods, such as gas chromatography (GC) or liquid chromatography (LC), in particular high-performance liquid chromatography (HPLC), to achieve higher resolutions by separating the components of a sample prior to MS analysis. Additionally, tandem arrangements of multiple mass analyzers (MS/MS) are possible to better separate structurally similar molecules and enhance sensitivity. In proteomics, LC-MS/MS instruments have found wide application, since they can deliver information “that cannot be acquired from DNA sequencing alone”¹⁶⁰, efficiently complementing whole-genome sequencing. Additionally, devices combining the MALDI approach and TOF mass analyzers are common. For example, studies have shown MALDI-TOF’s effectiveness in analyzing entire bacterial cells.^{156,161}

Also, combination of mass spectrometry with PCR is possible. For instance the multiplex PCR, known as MassTag PCR¹⁶² leverages the high resolution of mass spectrometry and the sensitivity and specificity of PCR.¹⁴ MassTag PCR utilizes specifically labelled or tagged primers serving as barcodes.¹⁶² After the amplification, the labels are released from the PCR products using ultraviolet (UV) light, making them accessible for subsequent mass spectrometry. This process generates quantifiable spectrometric data, with each peak corresponding to the specific label or barcode.¹⁶² In microbial forensics, tags can be used to detect the presence and determine the abundance of pathogens in the sample. For example, MassTag PCR has been used to identify viral hemorrhagic fever viruses, including Ebola, Marburg, and Lassa viruses.¹⁶³

It can be used to analyze various biomolecules, encompassing nucleic acids, proteins, peptides, sugars, lipids, and metabolites. The advantages of mass spectrometry include the low amounts of sample required, speed, and accuracy.¹⁵⁶ These characteristics make this technology valuable for demanding microbiological forensic work alongside advanced sequencing techniques. Mass spectrometry and whole-genome sequencing can complement each other effectively. Note that microarray immunoassay technology (see Box 7) can also analyze peptides and other biomolecules. However, because of the superior sensitivity and specificity of mass spectrometry, we will briefly discuss its potential in microbial forensics. In the following, we will focus on several applications of mass spectrometry for the rapid detection, identification, and characterization of microorganisms.

As stated above, pathogenic **proteins** represent an important target in the context of microbial forensic analysis. Using high-resolution mass spectrometry instruments, such as a tandem arrangement of TOF analyzers coupled with MALDI ionization (MALDI-TOF/TOF), it is possible to characterize protein biomarkers in intact

microorganisms for quick identification of *Bacillus* spores, whether pure or mixed, without additional sample preparation.^{156,164} As pointed out by the authors of the study, such an approach, combined with bioinformatics algorithms, can potentially be used for the detection of genetically engineered organisms.¹⁶⁴ In another research study, a protein antigen for *B. anthracis* has been identified that is highly promising for the specific identification of this microorganism in unknown samples of various origins.^{164,165} This approach is a multistep procedure that involves purification of bacterial proteins using monoclonal antibodies (antibodies having defined specificity) for the selected bacterial antigens and the subsequent identification of bacterial antigens applying multi-dimensional protein identification technology¹⁶⁵ that involves mass spectrometry. As stated elsewhere¹⁵⁶, this methodology can be utilized to simultaneously identify other pathogens. Overall, there have been a number of studies based on the mass spectrometry analysis of proteins extracted from the spores of *B. anthracis* and related species that have been reviewed elsewhere.¹⁵⁶

One of the challenges in this area is that some proteins are common in closely related organisms, making unambiguous identification and determination of their origin quite difficult.¹⁵⁶ This phenomenon is observed in organisms belonging to the groups of *Yersinia* (*Y. pestis*, the causative agent of plague, and *Y. pseudotuberculosis*), *Bacillus* (such as *B. anthracis*, *B. cereus*, and *B. thuringiensis*), and *Brucella*.¹⁶⁰ According to the literature, in one of the analyses using MALDI-TOF, *Y. pseudotuberculosis* has been misidentified as *Y. pestis*, an error that could be corrected using genotyping methods.^{160,166} Also, cases when *B. cereus* or *B. thuringiensis* have been falsely identified as *B. anthracis* are known.¹⁶⁷ Such misidentification can hinder the investigation process, cause panic, and waste valuable resources. This demonstrates that “MALDI-TOF-MS identification of high-risk pathogens or potential bioterror agents could be a useful forensic and biosecurity tool, but that the results need to be interpreted”¹⁶⁰ using additional information. An intuitive approach is to use protein biomarkers that are unique to specific organisms. Interestingly, with the increase in available sequencing data constantly being added to databases, fewer such species-unique peptides remain.¹⁶⁰ For this reason, statistical methods and parameters, such as the so-called peptide strength,¹⁶⁸ a measure of the statistical frequency of a peptide sequence in the database, can be used to evaluate the data.¹⁶⁰ A database frequently used for the rapid and reliable identification of highly pathogenic bacteria is the openly available MALDI-TOF database of the Robert Koch Institute (RKI). The last version (4.2) of the database contains a total of 11,055 spectra from 1,601 microbial strains of bacteria belonging to the groups of *Bacillus*, *Brucella*, *Burkholderia*, *Francisella*, and *Yersinia*.¹⁶⁹ However, as stated elsewhere¹⁶⁷, it is strongly recommended to verify the obtained results by alternative methods, due to reported misidentifications of various bacterial species mentioned above.

Mass spectrometry methods can also help answer another key question in microbial forensics: whether a biological threat agent was intentionally released or resulted from a natural outbreak. Indications of intentional release include the lab origin of a pathogen (although biosafety breaches can also cause accidental outbreaks). Several studies demonstrate that protein expression (the process involving protein synthesis, modification, and regulation) differs between organisms grown in laboratories and wild-type isolates.¹⁶⁰ One study compared protein expression of two wild-type *Y. pestis* isolates to lab-grown samples using whole-genome sequencing, liquid chromatography tandem mass spectrometry (LC-MS/MS), and gas chromatography mass spectrometry (GC/MS).¹⁷⁰ Notably, no universal genomic changes were identified (excluding mutational hotspots related to loss of virulence); however, differences in protein expression patterns could be observed.¹⁶⁰ These differences may result from microorganisms adapting to laboratory conditions.¹⁶⁰ The study also aimed to identify signatures that could distinguish wild-type from laboratory-adapted strains. For this purpose, an extensive collection of published proteomic data on *Y. pestis* was analyzed. The evaluation was facilitated through machine learning. Overall, combining genomics and proteomics holds promise for accurately classifying wild-type and laboratory-adapted species, thus expanding the tools available for microbial forensics.^{160,170} This idea is reinforced by studies showing that mass spectrometry can be successfully applied to gather information about the composition of specific culture media for bacterial growth^{160,171}, or about virus preparation,^{160,172} providing detailed insights into pathogen production for thorough forensic analysis.

Besides proteins, mass spectrometry can analyze **nucleic acids**. It is a highly sensitive technique that can capture genetic variations in microbial sequences, as these variations lead to measurable differences in the

molecular mass of PCR products, which can be analyzed using ESI-MS (see Box 8) or MALDI-MS.¹⁵⁶ Mass spectrometry is capable of discriminating SNPs (see Box 2), and can be used to monitor post-transcriptional modifications (modifications that occur after the transcribing DNA segment into a complementary messenger mRNA) of bacterial RNA¹⁷³, or to analyze uncultured bacteria.¹⁵⁶ For instance, several VNTRs (see Box 3) and SNPs (see Box 2) of *B. anthracis* have been characterized using mass spectrometry¹⁷⁴ thus providing an alternative to the sequencing methods. In one reported study, high-throughput analysis of PCR products from the 16S rRNA region of 10 different bacterial species was conducted using ion-pair reversed-phase high-performance liquid chromatography (IP RP HPLC) coupled with ESI-MS, demonstrating high mass accuracy and reproducibility.¹⁷⁵ Additionally, mass spectrometry analysis of PCR products from the 16S-23S rRNA and interspace (ISR) regions can potentially be used to distinguish closely related species.¹⁵⁶ However, it was pointed out that the applicability of 16S rRNA beyond the bacterial genus level (e.g., *Bacillus*) is highly limited.¹⁷⁶

Lipids, lipopolysaccharides, and metabolites can also be analyzed using mass spectrometry. For instance, the structure of lipid A (a constituent of the outer membrane of gram-negative bacteria¹⁷⁷) from the biothreat agent *F. tularensis* has been characterized through various chemical and mass spectrometric techniques.^{156,178} Metabolomic profiling is highly challenging due to the large chemical diversity of the related compounds.¹⁵⁶ However, it can offer valuable information for microbial forensics. One study reports the detection and classification of deliberately released bioaerosols containing gram-positive (including *B. anthracis*) and gram-negative bacteria in outdoor field scenarios based on the identifying specific metabolites.^{156,179} Furthermore, mass spectrometry can analyze biological toxins such as ricin and botulinum toxin¹⁴ and is useful in particular when no genetic information is present. All these examples illustrate that mass spectrometry is indispensable in microbial forensics and, alongside with genomic techniques, can greatly aid in the accurate identification and detailed characterization of biomaterials. Further advances in this field are expected to significantly overcome current limitations of mass spectrometry.

Section 5

Conclusions

Because the need for microbial forensics is international, there should be internationally accepted practices.

Science Needs for Microbial Forensics:
Developing Initial International Research
Priorities. p.99, 2014.

Microbial forensics has already demonstrated its ability to provide valuable information in past incidents involving biological weapons. Therefore, this area is very promising for the technical implementation of any future verification mechanism and compliance under the BWC. It is relevant for both pre- and post-verification frameworks. It can a) serve as one of the key tools in the investigative process, b) help gather and evaluate information prior to any potential incident, and c) have deterrent effect on malicious actors. Nevertheless, as highlighted throughout the manuscript, microbial forensics alone is not sufficient for all these purposes. Instead, it must be consolidated with other methods and approaches in order to achieve the best possible results.

The methodological toolkit of microbial forensics is as versatile as the various scenarios related to biological weapons development, utilizing established and emerging technologies. It can be applied to analyze a broad spectrum of biomolecules: from nucleic acids to proteins and metabolites, among others. Progress in whole-genome sequencing methods has truly revolutionized this field. However, established immunological methods, advances in PCR, mass spectrometry, and other analytical techniques also form the pillars of this discipline.

Advances in computational approaches, including rapid developments in artificial intelligence, play an important role in analyzing complex biological data sets. Additionally, they receive significant attention for their potential application in attributing biological material to its source and/or method of production. Algorithmic improvements, the establishment of curated, harmonized databases for biological hazards, reviews of model evaluation metrics, reductions in computational costs, and other factors will further enhance the performance of bioinformatics tools for microbial forensics.

Technological and scientific advancements do not operate in a vacuum; they require sophisticated know-how and expertise. National microbial forensics capabilities should be developed similarly to the international practices established, e.g., in Switzerland (SPIEZ laboratory¹⁸⁰) and United Kingdom (United Kingdom Microbial Forensics Consortium (UKMFC)¹⁸¹). In the case of the UKMFC in particular, several working groups were set up, including a bioinformatics working group, a forensics and sampling working group, and a quality working group.¹⁸² The UKMFC has also an advisory board “to ensure that appropriate standards of work can be met in the event of an investigation of a biological attack”¹⁸². Expert education in microbial forensics should be made a priority for the State Parties. This should include in-depth training in both state-of-the-art approaches of microbial forensics and standard procedures of classical forensics to ensure correct and careful sample collection

and data acquisition. A new approach includes using virtual reality (VR) equipment for forensic training purposes in different simulated environments.¹⁸³

The toolkit of microbial forensics will continue to expand alongside advancements in related disciplines forming this field. Therefore, it is advisable to have a dedicated group of experts regularly reviewing these developments. The main goal is to guarantee that analytical methods keep pace with upcoming challenges and reflect the latest scientific standards. As a result, we believe that, in addition to the verification mechanism, a scientific advisory board should be established as a subsidiary body under the BWC. Further discussions among political leaders will help shape the roadmap for strengthening the Convention. It remains to be seen which mechanisms will be incorporated into the BWC in the future to ensure compliance with its provisions. However, scientific and technological advances relevant to the Convention should always be closely monitored, as political decisions lag significantly behind scientific progress that makes its way into everyday life.

Bibliography

- [1] NATO warns against losing biotech edge to China and Russia. <https://ukdefencejournal.org.uk/nato-warns-against-losing-biotech-edge-to-china-and-russia/> (accessed 2026).
- [2] The Third Review Conference of the Parties to the Convention on the Prohibition of Development Production, and Stockpiling of Bacteriological (Biological) and Toxin Weapons and on their Destruction. Final Document BWC/Conf.III/23. Geneva, 9-27 September 1991.
- [3] Ad Hoc Group of Governmental Experts to identify and examine potential Verification Measures from a scientific and technical standpoint. Report. BWC/CONF.III/VEREX/9. Geneva, 24 September 1993.
- [4] Ninth Review Conference of the States Parties to the Convention on the Prohibition of the Development, Production and Stockpiling of Bacteriological (Biological) and Toxin Weapons and on Their Destruction. Final Document of the Ninth Review Conference. BWC/CONF.IX/9. Geneva, 22 December 2022.
- [5] Working Group on the Strengthening of the Convention on the Prohibition of the Development, Production and Stockpiling of Bacteriological (Biological) and Toxin Weapons and on Their Destruction. Draft final report of the Working Group on the Strengthening of the Convention. BWC/WG/7/CRP.1. Geneva, 1 December 2025.
- [6] A. Castro and M. C. A. De Ungria. Methods used in Microbial Forensics and Epidemiological Investigations for Stronger Health Systems. *Forensic Sci Res*, 7(4):650–661, 2022.
- [7] A. Thomas and J. Archer. NATOs Weapons of Mass Destruction Forensics Conference. In *Workshop Report*, pages 1–8. NATO. https://www.nato.int/nato_static/assets/pdf/pdf_2011_02/20110615_WMD-Forensics-Conference-Report.pdf.
- [8] D. Perkins. Using Microbial Forensics to Strengthen Biosecurity and the Implementation of UN Security Council Resolution 1540. In *Science Needs for Microbial Forensics: Developing an International Science Roadmap*. Croatian Academy of Sciences and Arts, 3-16 October 2013. <https://www.un.org/ar/sc/1540/documents/expert-presentation-2013-63-Zagreb.pdf>.
- [9] T. Rissman and A. Prieto. *Attributing Biological Weapons Use: Strengthening Department of Defense Capabilities to Investigate Deliberate Biological Incidents*. RAND Corporation, Santa Monica, CA, 2024.
- [10] B. Budowle, S. E. Schutzer, A. Einseln, et al. Building Microbial Forensics as a Response to Bioterrorism. *Science*, 301(5641):1852–1853, 2003.
- [11] S. Schmedes, A. Sajantila, and B. Budowle. Expansion of Microbial Forensics. *J Clin Microbiol*, 54:1964–1974, 2016.
- [12] Meeting of the States Parties to the Convention on the Prohibition of the Development, Production and Stockpiling of Bacteriological (Biological) and Toxin Weapons and on Their Destruction. Report of the Meeting of States Parties. BWC/MSP/2010/6. page 4, Geneva, 17 December 2010.

- [13] Seventh Review Conference of the States Parties to the Convention on the Prohibition of the Development, Production and Stockpiling of Bacteriological (Biological) and Toxin Weapons and on Their Destruction. New scientific and technological developments relevant to the Convention. Background information document submitted by the Implementation Support Unit. Addendum. BWC/CONF.VII/INF.3/Add.1. page 19, Geneva, 23 November 2010.
- [14] 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. Science Needs for Microbial Forensics: Developing Initial International Research Priorities. National Academies Press (US), 25 July 2014.
- [15] X. Guo, L. Gu, Y. Luo, et al. A bibliometric analysis of microbial forensics from 1984 to 2022: progress and research trends. *Front Microbiol*, 14:1–13, 2023.
- [16] D. B. Jernigan, P. L. Raghunathan, B. P. Bell, et al. Investigation of bioterrorism-related anthrax, United States, 2001: epidemiologic findings. *Emerg Infect Dis*, 8(10):1019–1028, 2002.
- [17] Committee on Review of the Scientific Approaches Used During the FBI’s Investigation of the 2001 Bacillus Anthracis Mailings; National Research Council. Review of the Scientific Approaches Used during the FBI’s Investigation of the 2001 Anthrax Letters. Washington (DC): National Academies Press, 2011.
- [18] R. S. Decker and T. L. Kerns. Chapter 2 - The FBI’s Amerithrax Task Force and the advent of microbial forensics. In B. Budowle, S. Schutzer, and S. Morse, editors, *Microbial Forensics (Third Edition)*, pages 11–23. Academic Press, third edition edition, 2020.
- [19] G. L. Andersen, J. M. Simchock, and K. H. Wilson. Identification of a region of genetic variability among Bacillus anthracis strains and related species. *J Bacteriol*, 178(2):377–384, 1996.
- [20] P. J. Jackson. Chapter 3 - Microbial forensic investigation of the anthrax letter attacks: how the investigation would differ using today’s technologies. In B. Budowle, S. Schutzer, and S. Morse, editors, *Microbial Forensics (Third Edition)*, pages 25–32. Academic Press, third edition edition, 2020.
- [21] A. G. Fischer and G. Jocham. Chapter 33 - The role of serotonin in performance monitoring and cognitive control. In C. P. Müller and K. A. Cunningham, editors, *Handbook of the Behavioral Neurobiology of Serotonin*, volume 31 of *Handbook of Behavioral Neuroscience*, pages 571–588. Elsevier, 2020.
- [22] P. Keim, L. B. Price, A. M. Klevytska, et al. Multiple-Locus Variable-Number Tandem Repeat Analysis Reveals Genetic Relationships within *Bacillus anthracis*. *J Bacteriol*, 182(10):2928–2936, 2000.
- [23] J. Watson and F. Crick. Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid. *Nature*, 171:737–738, 1953.
- [24] Single Nucleotide Polymorphisms (SNPs). National Human Genome Research Institute. <https://www.genome.gov/genetics-glossary/Single-Nucleotide-Polymorphisms-SNPs> (accessed 2026).
- [25] Amerithrax investigative summary. Released Pursuant to the Freedom of Information Act. The United States Department of Justice, 19 February 2010.
- [26] What are single nucleotide polymorphisms (SNPs)? <https://medlineplus.gov/genetics/understanding/genomicresearch/> (accessed 2026).
- [27] M. Leitenberg, R. Zilinskas, and J. Kuhn. *The Soviet Biological Weapons Program: A History*. Harvard University Press, 2012.
- [28] M. Meselson, J. Guillemin, M. Hugh-Jones, et al. The Sverdlovsk Anthrax Outbreak of 1979. *Science (New York, N.Y.)*, 266(5188):1202–1208, 1994.

- [29] P. J. Jackson, M. E. Hugh-Jones, D. M. Adair, et al. PCR analysis of tissue samples from the 1979 Sverdlovsk anthrax victims: The presence of multiple *Bacillus anthracis* strains in different victims. *Proc Natl Acad Sci USA*, 95(3):1224–1229, 1998.
- [30] J. W. Sahl, T. Pearson, R. Okinaka, et al. A *Bacillus anthracis* Genome Sequence from the Sverdlovsk 1979 Autopsy Specimens. *mBio*, 7(5):1–8, 2016.
- [31] K. Alibek and S. Handelman. *Biohazard: the chilling true story of the largest covert biological weapons in the world—told from the inside by the man who ran it*. Random House Publishing Group, 1999.
- [32] J. B. Tucker. Biological weapons in the former Soviet Union: An interview with Dr. Kenneth Alibek. *The Nonproliferation Review*, 6(3):1–10, 1999.
- [33] Variable number tandem repeats. <https://www.nature.com/wls/definition/variable-number-tandem-repeats-vntrs-vntr-238/> (accessed 2026).
- [34] List of parties to the Biological Weapons Convention. https://en.wikipedia.org/wiki/List_of_parties_to_the_Biological_Weapons_Convention (accessed 2026).
- [35] R. A. Zilinskas. Iraq’s Biological Weapons: The Past as Future? *JAMA*, 278(5):418–424, 08 1997.
- [36] J. Tucker and G. Koblenz. The Four Faces of Microbial Forensics. *Biosecur Bioterror*, 7:389–397, 2009.
- [37] S. Black, H. Wilson, and F. Lentzos. UNSCOM’s work to uncover Iraq’s illicit biological weapons program: A primer. *Bull At Sci*, 77(4):199–202, 2021.
- [38] S. Wright. Proof of Iraqi weapons still elusive, U.N. inspector says. *MIT News*, 19 March 2003.
- [39] Y. Zhao, W. Zhang, and X. Zhang. Application of metagenomic next-generation sequencing in the diagnosis of infectious diseases. *Front Cell Infect Microbiol*, 14:1–15, 2024.
- [40] U. Jakob, S. Kloth, and I. Mergler. Investigation of Alleged Use of Biological Weapons – Strengthening the UN Secretary-General’s Mechanism. *PRIF Report*, 7:1–42, 2024.
- [41] S. Appelt, A.-M. Rohleder, C. Invernizzi, et al. Strengthening the United Nations Secretary-General’s Mechanism to an alleged use of bioweapons through a quality-assured laboratory response. *Nat Commun.*, 12(3078), 2021.
- [42] Secretary-General’s Mechanism for Investigation of Alleged Use of Chemical and Biological Weapons (UNSGM). <https://disarmament.unoda.org/en/our-work/weapons-mass-destruction/secretary-generals-mechanism-investigation-alleged-use-chemical> (accessed 2026).
- [43] Chemical and bacteriological (biological) weapons. Report of the Secretary-General, A/44/561. 4 October 1989 1991.
- [44] K. Warmbrod and G. Gronvall. A playbook on investigations into biological events of ambiguous origin. *The Nonproliferation Review*, 30:1–21, 10 2024.
- [45] R. Trapp. Chemical and microbiological forensics in investigations of alleged uses of chemical and biological weapons – a preliminary analysis. *CBWNet Working Paper*, 4:1–20, 02 2023.
- [46] UNSGM Designated Laboratories Workshop Report. Spiez, Switzerland, 11-13 September 2024.
- [47] UNSGM project. Contribution to Strengthening the United Nations Secretary-General’s Mechanism to Investigate Suspected Use of Biological Weapons. <https://www.rki.de/EN/Topics/Infectious-diseases/Biological-threats/Projects-networks/UNSGM/UNSGM.html>(accessed 2026).
- [48] M. C. J. Maiden, J. A. Bygraves, E. Feil, et al. Multilocus sequence typing: A portable approach to the identification of clones within populations of pathogenic microorganisms. *Proc Natl Acad Sci USA*, 95(6):3140–3145, 1998.

- [49] D. V. Lim, J. M. Simpson, E. A. Kearns, et al. Current and Developing Technologies for Monitoring Agents of Bioterrorism and Biowarfare. *Clin Microbiol Rev*, 18:583–607, 2005.
- [50] S. K. Alketbi. Emerging Technologies in Forensic DNA Analysis. *Perspect Legal Forensic Sc*, 1:1–24, 2024.
- [51] B. Budowle, N. Connell, A. Bielecka-Oder, et al. Validation of high throughput sequencing and microbial forensics applications. *Investig Genet*, 5:9, 2014.
- [52] K. B. Mullis, F. Faloona, S. J. Scharf, et al. Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. *Cold Spring Harb Symp Quant Biol*, 51 Pt 1:263–273, 1986.
- [53] The Nobel Prize in Chemistry 1993. <https://www.nobelprize.org/prizes/chemistry/1993/press-release/> (accessed 2026).
- [54] A. Van Stelten and K. Nightingale. Microbiological Analysis. DNA methods. In M. Dikeman and C. Devine, editors, *Encyclopedia of Meat Sciences (Second Edition)*, pages 294–300. Elsevier, 2014, 2014.
- [55] M. Gavrilov, J. Yang, R. Zou, et al. Engineered helicase replaces thermocycler in DNA amplification while retaining desired PCR characteristics. *Nat Commun*, 13(6312), 2022.
- [56] C. McDonald, D. Taylor, and A. Linacre. PCR in Forensic Science: A Critical Review. *Genes*, 15(4):1–22, 2024.
- [57] T. U. G. Mir, A. Wani, N. Akhtar, S. Sena, et al. Microbial forensics: A potential tool for investigation and response to bioterrorism. *J Health Sci Rev*, 5:100068, 2022.
- [58] Comparing SYBR Green I dye and TaqMan Assays for real-time PCR analysis. <https://documents.thermofisher.com/TFS-Assets/GSD/Flyers/sybr-green-vs-taqman-qpcr-infographic.pdf> (accessed 2026).
- [59] R. Maes, I. Langohr, A. Wise, R. Smedley, T. Thaiwong, and M. Kiupel. Beyond H&E: Integration of Nucleic Acid-Based Analyses Into Diagnostic Pathology. *Vet. Pathol.*, 51, 10 2013.
- [60] PCR/qPCR Data Analysis. A Technical Guide to PCR Technologies. <https://www.sigmaaldrich.com/DE/en/technical-documents/technical-article/genomics/qpcr/data-analysis?srsId=AfmBOoo1PUZjgkGHDY8mlBvs7edFhmG1QZG1NN41Hr02Q89o1kLSGEEQ> (accessed 2026).
- [61] What is PCR and qPCR? <https://www.youtube.com/watch?v=rpLSvEbOmqc> (accessed 2026).
- [62] T. Parsons, V. Cox, A. Essex-Lopresti, et al. Development of Three Real-Time PCR assays to Detect *Bacillus anthracis* and Assessment of Diagnostic Utility. *J Bioterr Biodef*, S3:1–6, 2013.
- [63] J. L. Mitchell, N. Chatwell, D. Christensen, et al. Development of real-time PCR assays for the specific detection of *Francisella tularensis ssp. tularensis*, *holarctica* and *mediaasiatica*. *Mol Cell Probes*, 24(2):72–76, 2010.
- [64] T. Minogue, P. Rachwal, A. Hall, et al. Cross-Institute Evaluations of Inhibitor-Resistant PCR Reagents for Direct Testing of Aerosol and Blood Samples Containing Biological Warfare Agent DNA. *Appl Environ Microbiol*, 80(4):1322–1329, 2014.
- [65] S. Weller, V. Cox, A. Essex-Lopresti, et al. Evaluation of two multiplex real-time PCR screening capabilities for the detection of *Bacillus anthracis*, *Francisella tularensis* and *Yersinia pestis* in blood samples generated from murine infection models. *J Med Microbiol*, 61:1546–1555, 2012.
- [66] F. Sanger, S. Nicklen, and A. R. Coulson. DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci USA*, 74(12):5463–5467, 1977.

- [67] M. Oliveira, K. Marszałek, M. Kowalski, et al. Sequencing Technologies in Forensic Microbiology: Current Trends and Advancements. *Forensic Sci*, 4:523–545, 2024.
- [68] X. Hu, J. Liu, T. Xu, et al. Research progress and application of the third-generation sequencing technologies in forensic medicine. *J Leg Med*, 71:102532, 2024.
- [69] Amplicon Sequencing: Principle, Steps, Types, Uses, Diagram. <https://microbenotes.com/amplicon-sequencing/> (accessed 2026).
- [70] C. Y. Chiu and S. A. Miller. Clinical metagenomics. *Nat Rev Genet*, 20:341–355, 2019.
- [71] M. Akhoundi, T. Downing, J. Votýpka, et al. Leishmania infections: Molecular targets and diagnosis. *Mol Asp Med*, 57:1–29, 2017.
- [72] J. Johnson, D. Spakowicz, B.-Y. Hong, et al. Evaluation of 16S rRNA gene sequencing for species and strain-level microbiome analysis. *Nat Commun*, 10, 2019.
- [73] E. Oniciuc, E. Likotrafti, A. Alvarez-Molina, et al. The Present and Future of Whole Genome Sequencing (WGS) and Whole Metagenome Sequencing (WMS) for Surveillance of Antimicrobial Resistant Microorganisms and Antimicrobial Resistance Genes across the Food Chain. *Genes*, 9, 05 2018.
- [74] P. Tripathi, R. Render, S. Nidhi, and V. Tripathi. Microbial genomics: a potential toolkit for forensic investigations. *Forensic Sci Med Pathol*, 21:417–429, 2025.
- [75] C. Méndez-García, R. Bargiela, M. Martínez-Martínez, et al. Chapter 2 - Metagenomic Protocols and Strategies. In M. Nagarajan, editor, *Metagenomics*, pages 15–54. Academic Press, 2018.
- [76] D. T. Truong, A. Tett, E. Pasolli, et al. Microbial Strain-Level Population Structure and Genetic Diversity from Metagenomes. *Genome Res*, 27:gr.216242.116, 2017.
- [77] A. Raza and M. S. Shahid. Chapter 10 - Next-generation sequencing technologies and plant molecular virology: a practical perspective. In L. Awasthi, editor, *Applied Plant Virology*, pages 131–140. Academic Press, 2020.
- [78] H. Browne, S. Forster, B. Anonye, et al. Culturing of ‘unculturable’ human microbiota reveals novel taxa and extensive sporulation. *Nature*, 533, 2016.
- [79] Amplicon and metagenomics overview. https://astrobiomike.github.io/misc/ampliconD_and_metagen (accessed 2026).
- [80] A. Sjödin, T. Broman, Ö. Melefors, et al. The Need for High-Quality Whole-Genome Sequence Databases in Microbial Forensics. *Biosecur Bioterror*, 11:S78–S86, 2013.
- [81] L. Gautier. Microbial Forensics: What We’ve Learned from Amerithrax and Beyond. *Biotechniques*, 75, 2023.
- [82] PacBio. Sequencing 101: long-read sequencing. <https://www.pacb.com/blog/long-read-sequencing/> (accessed 2026).
- [83] MedlinePlus. What is epigenetics? <https://medlineplus.gov/genetics/understanding/howgeneswork/epigenome/> (accessed 2026).
- [84] Single Molecule Real-Time sequencing Third-generation sequencing. <https://www.youtube.com/watch?v=yZjZPnbu4wo> (accessed 2026).
- [85] A. Wenger, P. Peluso, W. Rowell, et al. Accurate circular consensus long-read sequencing improves variant detection and assembly of a human genome. *Nat Biotechnol*, 37:1155–1162, 2019.
- [86] Nanopore Sequencing. <https://www.youtube.com/watch?v=FYEWrUVJ2as> (accessed 2026).

- [87] V. Marx. Method of the year: long-read sequencing. *Nat Methods*, 20:6–11, 2023.
- [88] Y. Wang, Y. Zhao, A. Bollas, et al. Nanopore sequencing technology, bioinformatics and applications. *Nat Biotechnol*, 39:1348–1365, 11 2021.
- [89] Oxford Nanopore Technologies. How nanopore sequencing works. <https://nanoporetech.com/platform/technology> (accessed 2026).
- [90] Oxford Nanopore Technologies. Oxford Nanopore releases new machine learning model to enable real-time, high-accuracy epigenetic insights in biological analyses. <https://nanoporetech.com/news/news-oxford-nanopore-releases-new-machine-learning-model-enable-real-time-high-accuracy> (accessed 2026).
- [91] H. Lu, F. Giordano, and Z. Ning. Oxford Nanopore MinION Sequencing and Genome Assembly. *Genom. Proteom. Bioinform*, 14, 2016.
- [92] W. De Coster, P. De Rijk, A. De Roeck, et al. Structural variants identified by Oxford Nanopore PromethION sequencing of the human genome. *Genome Res*, 29:issue = 4,, 2019.
- [93] L. Liu, Y. Yang, Y. Deng, et al. Nanopore long-read-only metagenomics enables complete and high-quality genome reconstruction from mock and complex metagenomes. *Microbiome*, 10, 2022.
- [94] H. McLaughlin, J. Bugrysheva, A. Conley, et al. Rapid Nanopore Whole-Genome Sequencing for Anthrax Emergency Preparedness. *Emerg Infect Dis*, 26:358–361, 2020.
- [95] F. Rang, W. Kloosterman, and J. De Ridder. From squiggle to basepair: Computational approaches for improving nanopore sequencing read accuracy. *Genome Biol*, 19, 2018.
- [96] Oxford Nanopore Technologies. MinION. <https://nanoporetech.com/products/sequence/minion> (accessed 2026).
- [97] N. Faria, J. Quick, I. Claro, et al. Establishment and cryptic transmission of Zika virus in Brazil and the Americas. *Nature*, 546, 2017.
- [98] N. R. Faria, M. U. G. Kraemer, S. C. Hill, et al. Genomic and epidemiological monitoring of yellow fever virus transmission potential. *Science*, 361(6405):894–899, 2018.
- [99] J. G. de Jesus, K. R. Dutra, F. C. D. S. Sales, et al. Genomic detection of a virus lineage replacement event of dengue virus serotype 2 in Brazil, 2019. *Mem Inst Oswaldo Cruz*, 115, 2020.
- [100] S. Hamner, B. L. Brown, N. A. Hasan, et al. Metagenomic Profiling of Microbial Pathogens in the Little Bighorn River, Montana. *Int J Environ Res Public Health*, 16, 2019.
- [101] C. Stillman, J. Bravo, C. Boucher, et al. Toward security-aware portable sequencing. *Nat Commun.*, 16, 2025.
- [102] L. Mazeaud and A. Kastelic. Cyberbiosecurity: A Matter of International Peace and Security? pages 1–20. Geneva: UNIDIR, 2025. <https://doi.org/10.37559/WMD/25/CBW/02>.
- [103] National Human Genome Research Institute (NIH). Allele. <https://www.genome.gov/genetics-glossary/Allele>(accessed 2026).
- [104] Public databases for molecular typing and microbial genome diversity. <https://pubmlst.org/>(accessed 2026).
- [105] Application Note MLST. A Fast, Accurate, and Automated Workflow for Multi Locus Sequence Typing of Bacterial Isolates. https://documents.thermofisher.com/TFS-Assets/LSG/Application-Notes/cms_042189.pdf(accessed 2026).

- [106] P. Singh, S. L. Foley, R. Nayak, et al. Massively parallel sequencing of enriched target amplicons for high-resolution genotyping of *Salmonella* serovars. *Mol Cell Probes*, 27(2):80–85, 2013.
- [107] P. Singh, S. L. Foley, R. Nayak, et al. Multilocus sequence typing of *Salmonella* strains by high-throughput sequencing of selectively amplified target genes. *J Microbiol Methods*, 88(1):127–133, 2012.
- [108] Advantages of Automated Sample Preparation. <https://www.thermofisher.com/de/en/home/life-science/dna-rna-purification-analysis/automated-purification-extraction/advantages-automated-extraction.html>(accessed 2026).
- [109] Cepheid GeneXpert Systems. <https://www.cepheid.com/en-GB/systems/genexpert-family-of-systems.html>(accessed 2026).
- [110] P. P. Banada, S. Deshpande, S. Banik, et al. Multiplex detection of three select agents directly from blood using the GeneXpert system. *J Clin Microbiol*, 57, 2019.
- [111] A. Kansagara, H. McMahon, and M. Hogan. Dry-state, room-temperature storage of DNA and RNA. *Nat Methods*, 5:iv–v, 2008.
- [112] P. Marshall, J. King, and B. Budowle. Utility of amplification enhancers in low copy number DNA analysis. *Int J Legal Med*, 129:43–52, 05 2014.
- [113] N. Aljanahi, S. K. Alketbi, M. Almheiri, et al. Enhancing trace DNA profile recovery in forensic casework using the amplicon RX post-PCR clean-up kit. *Sci Rep*, 15:3324, 2025.
- [114] Comparing real-time and digital PCR technologies. <https://documents.thermofisher.com/TFS-Assets/GSD/posters/comparing-realtime-digital-pcr-quantitation-poster.pdf>(accessed 2026).
- [115] P. Quan, M. Sauzade, and E. Brouzes. dPCR: A Technology Review. *Sensors (Basel)*, 18, 2018.
- [116] M. Bus, E. Jong, J. King, et al. Reverse Complement-PCR, an Innovative and Effective method for Multiplexing Forensically Relevant Single Nucleotide Polymorphism Marker Systems. *BioTechniques*, 71, 2021.
- [117] M. I. Hajam and M. M. Khan. Microfluidics: a concise review of the history, principles, design, applications, and future outlook. *Biomater Sci*, 12:218–251, 2024.
- [118] T. Huang, R. Zhang, and J. Li. CRISPR-Cas-based techniques for pathogen detection: Retrospect, recent advances, and future perspectives. *J Adv Res*, 50:69–82, 2023.
- [119] C. O’Brien, K. Varty, and A. Ignaszak. The electrochemical detection of bioterrorism agents: a review of the detection, diagnostics, and implementation of sensors in biosafety programs for Class A bioweapons. *Microsyst Nanoeng*, 7, 2021.
- [120] C. Pöhlmann, L. Bellanger, M. Drevinek, et al. Multiplex Detection of Biothreat Agents Using an Automated Electrochemical ELISA Platform. *Proc. Technol*, 27:104–105, 2017.
- [121] E. Lieberman-Aiden, N. L. van Berkum, L. Williams, et al. Comprehensive Mapping of Long-Range Interactions Reveals Folding Principles of the Human Genome. *Science*, 326(5950):289–293, 2009.
- [122] D. Plummer, X. Lang, S. Zhang, et al. A comprehensive benchmark of single-cell Hi-C embedding tools. *Nat Commun*, 16, 2025.
- [123] S. Dutt, B. Karawadeniya, Y. Bandara, et al. Nanopore sensing and machine learning: future of biomarker analysis and disease detection. *Future Sci OA*, 10, 2024.
- [124] P. Ma, C. Li, M. Rahaman, et al. A state-of-the-art survey of object detection techniques in microorganism image analysis: from classical methods to deep learning approaches., 2022.

- [125] H. Yuan, Z. Wang, Z. Wang, et al. Trends in forensic microbiology: From classical methods to deep learning. *Front Microbiol*, 14, 2023.
- [126] B. Kösoglu-Kind, R. Loredó, M. Grossi, et al. A biological sequence comparison algorithm using quantum computers. *Sci Rep*, 13, 2023.
- [127] G. Lewis, J. L. Jordan, D. A. Relman, et al. The biosecurity benefits of genetic engineering attribution. *Nat Commun*, 11, 2020.
- [128] O. M. Crook, K. L. Warmbrod, G. Lipstein, et al. Analysis of the first genetic engineering attribution challenge. *Nat Commun*, 13, 2022.
- [129] E. C. Alley, M. Turpin, A. B. Liu, et al. A machine learning toolkit for genetic engineering attribution to facilitate biosecurity. *Nat Commun*, 11, 2020.
- [130] T. Slezak, B. Hart, and C. Jaing. Chapter 20 - Design of genomic signatures for pathogen identification and characterization. In B. Budowle, S. Schutzer, and S. Morse, editors, *Microbial Forensics (Third Edition)*, pages 299–312. Academic Press, third edition edition, 2020.
- [131] Office of the Director of National Intelligence Intelligence Advanced Research Projects Activity (IARPA). FELIX. Finding Engineering-Linked Indicators. <https://www.iarpa.gov/index.php/research-programs/felix> (accessed 2026).
- [132] R. Griffin. Tool Used to Rule Out Foul Play With Covid-19 May Spot Bioweapons. *Bloomberg Newsletter*, October 20 2022.
- [133] B. Budowle. Genetics and attribution issues that confront the microbial forensics field. *Forensic Sci Int*, 146 Suppl:S185–S188, 2005.
- [134] A. A. K. Nielsen and C. A. Voigt. Deep learning to predict the lab-of-origin of engineered DNA. *Nat Commun*, 9, 2018.
- [135] National Human Genome Research Institute (NIH). Phenotype. <https://www.genome.gov/genetics-glossary/Phenotype> (accessed 2026).
- [136] Addgene. A Better Way to Share Science. <https://www.addgene.org/> (accessed 2026).
- [137] Q. Wang, B. Kille, T. R. Liu, et al. PlasmidHawk improves lab of origin prediction of engineered plasmids using sequence alignment. *Nat Commun*, 12, 2021.
- [138] W. Mo, C. A. Vaiana, and C. J. Myers. The need for adaptability in detection, characterization, and attribution of biosecurity threats. *Nat Commun*, 15, 2024.
- [139] O. M. Crook, A. Franz, and A. Maiwald. Tracing engineered biothreats with AI forensics: Five steps to improve attribution. *Bull At Sci*, 2025.
- [140] M. Smith. Use Cases for Emerging Technologies to Strengthen High-containment Laboratory Governance. *SIPRI Research Policy Paper*, 2025.
- [141] C.-T. Berezin, S. Peccoud, D. M. Kar, and ot. Cryptographic approaches to authenticating synthetic DNA sequences. *Trends Biotechnol*, 42(8):1002–1016, 2024.
- [142] C. Clelland, V. Risca, and C. Bancroft. Hiding messages in DNA microdots. *Nature*, 399:533–534, 1999.
- [143] A. Huyghe, P. Francois, and J. Schrenzel. Characterization of microbial pathogens by DNA microarrays. *Infect Genet Evol*, 9(5):987–995, 2009.
- [144] V. Dias, P. Gomes, C. Damaso, et al. Germplex PCR and Massively Parallel Sequencing; A Dynamic Duo for Detecting Bacteria of Biodefense and Public Health Importance. *Forensic Sci Int Genet*, 81:103372, 2025.

- [145] Roka Bioscience. Overview. The Power of Three Core Technologies. <https://rokabio.com/technology/overview/> (accessed 2026).
- [146] Nordic Biosite. ELISA Principles 101. <https://nordicbiosite.com/blog/elisa-principles-101> (accessed 2026).
- [147] J. Berg, J. L. Tymoczko, and S. L. *Stryer Biochemie*. Springer Berlin Heidelberg, 6 edition, 2010.
- [148] Board on Health Sciences Policy; Board on Life Sciences; Institute of Medicine; National Research Council; Hook-Barnard I, M. Posey Norris S, Alper J, editors. Technologies to Enable Autonomous Detection for BioWatch: Ensuring Timely and Accurate Information for Public Health Officials. Workshop Summary. Washington (DC): National Academies Press (US); 2013 Dec 31., State of the Art for Autonomous Detection Systems Using Immunoassays and Protein Signatures. <https://www.ncbi.nlm.nih.gov/books/NBK201364/> (accessed 2026).
- [149] B. H. Hwang, H. H. Shin, J. Seo, and Cothers. Specific Multiplex Analysis of Pathogens Using a Direct 16S rRNA Hybridization in Microarray System. *Anal Chem*, 84:4873–4879, 2012.
- [150] M. Uttamchandani, J. L. Neo, B. N. Z. Ong, et al. Applications of microarrays in pathogen detection and biodefence. *Trends Biotechnol.*, 27(1):53–61, 2009.
- [151] M. Theodore, J. Jackman, and W. Bethea. Counterproliferation with Advanced Microarray Technology. *Johns Hopkins APL Technical Digest*, 25:38–43, 2004.
- [152] V. Perreten, L. Vorlet-Fawer, P. Slickers, et al. Microarray-Based Detection of 90 Antibiotic Resistance Genes of Gram-Positive Bacteria. *J Clin Microbiol*, 43:2291–2302, 2005.
- [153] T. J. Gentry and J. Zhou. Microarray-Based Microbial Identification and Characterization. In *Advanced Techniques in Diagnostic Microbiology*, pages 276–290. Springer US, Boston, MA, 2006.
- [154] S. S. Rubakhin and J. V. Sweedler. A Mass Spectrometry Primer for Mass Spectrometry Imaging. In S. S. Rubakhin and J. V. Sweedler, editors, *Mass Spectrometry Imaging: Principles and Protocols*, pages 21–49. Humana Press, Totowa, NJ, 2010.
- [155] D. N. Perkins, D. J. C. Pappin, D. M. Creasy, et al. Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis*, 20(18):3551–3567, 1999.
- [156] Y.-P. Ho, P. M. Reddy, C. Chen, et al. Mass Spectrometry in Microbial Forensics. *Forensic Sci Rev*, 21:25–50, 2009.
- [157] MALDI TOF Mass Spectrometry. <https://www.youtube.com/watch?v=nccPXw2Avu8> (accessed 2026).
- [158] Creative Proteomics. Molecular Weight Characterization - Comparison of MALDI and ESI. <https://www.creative-proteomics.com/pronalyse/resource-molecular-weight-characterization-comparison-of-maldi-and-esi.html> (accessed 2026).
- [159] Masonaco. Time-of-Flight (TOF) mass analyzers. <https://www.masonaco.org/theoretical-background/mass-analyzers/time-of-flight> (accessed 2026).
- [160] E. D. Merkley. Chapter 9 - Proteomics for Microbial Forensics. In E. D. Merkley, editor, *Applications in Forensic Proteomics: Protein Identification and Profiling*, pages 143–160. ACS Symposium Series, 2019.
- [161] M. Strejček, T. Šmrhová, P. Junková, et al. Whole-Cell MALDI-TOF MS Versus 16S rRNA Gene Analysis for Identification and Dereplication of Recurrent Bacterial Isolates. *Front Microbiol*, 9, 06 2018.
- [162] An introduction to MassTag PCR. <https://bitesizebio.com/33408/introduction-masstag-pcr/> (accessed 2026).
- [163] G. Palacios, T. Briese, V. Kapoor, et al. MassTag Polymerase Chain Reaction for Differential Diagnosis of Viral Hemorrhagic Fevers. *Emerg Infect Dis*, 12(4):692–695, 2006.

- [164] P. A. Demirev, A. B. Feldman, P. Kowalski, et al. op-down proteomics for rapid identification of intact microorganisms. *Anal Chem*, 77:7455–7461, 2005.
- [165] T. Krishnamurthy, J. Hewel, N. J. Bonzagni, et al. Simultaneous identification and verification of *Bacillus anthracis*. *Rapid Commun Mass Spectrom*, 20(13):2053–2056, 2006.
- [166] P. Gérôme, P. Le Flèche, Y. Blouin, et al. *Yersinia pseudotuberculosis* ST42 (O:1) Strain Misidentified as *Yersinia pestis* by Mass Spectrometry Analysis. *Genome Announc*, 2, 2014.
- [167] P. Lasch, W. Beyer, A. Bosch, et al. A MALDI-ToF mass spectrometry database for identification and classification of highly pathogenic bacteria. *Sci Data*, 12.
- [168] K. Jarman, N. C. Heller, S. C. Jenson, et al. Proteomics Goes to Court: A Statistical Foundation for Forensic Toxin/Organism Identification Using Bottom-Up Proteomics. *J Proteome Res*, 17:3075–3085, 2018.
- [169] Lasch P., Stämmle M., Schneider, A. Version 4.2 (20230306) of the MALDI-ToF Mass Spectrometry Database for Identification and Classification of Highly Pathogenic Microorganisms from the Robert Koch-Institute (RKI). Zenodo. <https://zenodo.org/records/14562231>(accessed 2026).
- [170] O. P. Leiser, E. D. Merkley, B. H. Clowers, et al. Investigation of *Yersinia pestis* Laboratory Adaptation through a Combined Genomics and Proteomics Approach. *PLoS One*, 10:e0142997, 2015.
- [171] B. H. Clowers, D. S. Wunschel, H. W. Kreuzer, et al. Characterization of Residual Medium Peptides from *Yersinia pestis* Cultures. *Anal Chem*, 85:3933–3939, 2013.
- [172] D. Wunschel, E. Tulman, H. Engelmann, et al. Forensic proteomics of poxvirus production. *Analyst*, 138:6385–6397, 2013.
- [173] F. Kirpekar, S. Douthwaite, and P. Roepstorff. Mapping posttranscriptional modifications in 5S ribosomal RNA by MALDI mass spectrometry. *RNA*, 6:296–306, 2000.
- [174] M. N. Van Ert, S. A. Hofstadler, J. Yun, et al. Mass Spectrometry Provides Accurate Characterization of two Genetic Marker Types in *Bacillus Anthracis*. *Biotechniques*, 37(4):642–651, 2004.
- [175] B. M. Mayr, U. Kobold, M. Moczko, et al. Identification of Bacteria by Polymerase Chain Reaction Followed by Liquid Chromatography–Mass Spectrometry. *Anal Chem*, 77:4563–4570, 2005.
- [176] O. Bartoš, M. Chmel, and I. Swierczková. The overlooked evolutionary dynamics of 16S rRNA revises its role as the “gold standard” for bacterial species identification. *Sci Rep*, 14, 2024.
- [177] J. A. Sutcliffe and N. H. Georgopapadakou. In *Emerging Targets in Antibacterial and Antifungal Chemotherapy*. Springer New York, NY, 1992.
- [178] N. J. Phillips, B. Schilling, M. K. McLendon, et al. Novel Modification of Lipid A of *Francisella tularensis*. *Infect Immun*, 72:5340–5348, 2004.
- [179] P. A. Snyder, J. P. Dworzanski, A. Tripathi, et al. Correlation of Mass Spectrometry Identified Bacterial Biomarkers from a Fielded Pyrolysis-Gas Chromatography-Ion Mobility Spectrometry Biotector with the Microbiological Gram Stain Classification Scheme. *Anal Chem*, 76:6492–6499, 2004.
- [180] S. Reinert. B and C Incidents: The Growing Importance of Forensics. *CSS Analyses in Security Policy*, 358, 2025.
- [181] UK Microbial Forensics Consortium. <https://www.gov.uk/government/groups/uk-microbial-forensics-consortium> (accessed 2026).

- [182] UK Capabilities Relevant to Implementation and Operationalisation of Article VII of the Biological and Toxin Weapons Convention. BWC/WG/4/WP.5. In *Working Group on the Strengthening of the Convention on the Prohibition of the Development, Production and Stockpiling of Bacteriological (Biological) and Toxin Weapons and on Their Destruction*, Geneva, 19. August 2024.
- [183] B. Budowle, M. Bus, L. Dungan, et al. The Forensic Genomics Toolbox is Expanding. *BioTechniques*, 72:5–7, 2021.

Gefördert durch:



Bundesministerium
für Forschung, Technologie
und Raumfahrt