Establishing Novel RNA SNP Biomarkers via RNA Sequencing for Identifying Biological Agents for Combatting Warfare and Disease Control

Aleksandra Casazza-Ernst, DFS., Oklahoma State University
aleks.casazza@outlook.com

Abstract

Biological agents, often used as a source for biological warfare, are a threat to global health and public safety. Biological agents are among the most lethal weapons due to their versatility and ability to infect individuals. One way to combat such attacks is to have methodologies in place to identify biological agents efficiently, so that efforts to stop transmission and initiate treatment may begin as soon as possible. This practical application study examines journals and publications regarding genetics, forensic microbiology, and RNA single nucleotide polymorphism (SNP) assessments to determine if there is a method suitable for biological agent identification. It was found that RNA SNP analysis, which has been successful for forensic human identification testing and body fluid identification, is a useful tool for identification of biological agents.

Keywords

Biological agents, biological terrorism, forensic microbiology, Singular nucleotide polymorphism, database, RNA sequencing

1. Introduction

The use of biological weapons for bioterrorism or biowarfare events has been a global concern for centuries. Derived from plants or animals, these weapons are among the most lethal, due to their transmittable properties. Biological agents can spread and infect populations rapidly, so it is crucial that these agents are identified quickly [1]. While there has been great success in forensic genetics and the use of DNA technologies, the identification of biological agents requires a niche set of standards[2] . Methodologies commonly used for forensic microbiology can be used to combat this issue. Forensic microbiology is a discipline of forensic science that uses molecular techniques, microscopy, and genetics to identify microorganisms [3]. Studies have been conducted using forensic microbiology, specifically incorporating RNA single nucleotide polymorphism (SNP) analysis for identification purposes.[12] These methods can be applied to the identification of biological agents, creating a standard procedure that is needed. Having a regulated standard in which to identify agents is crucial in crisis situations. Not only would this system help to deescalate a catastrophic event, but it would also aid public health issues.

2. Materials and methods

As this was a practical application study, no materials were utilized. Contents of the manuscript are described by reading and evaluating previous publications. According to Oklahoma State University's submission policy to the Journal Oklahoma State Medical Proceedings, a practical application study is designed to explore and review relevant issues regarding health care, as it relates to a students' particular scope of expertise. This is accomplished via extensive research and review of published material and exploration of topics that have yet to addressed in these documents.

3. Analysis and explanations of work reviewed

a. Biological Warfare verse Bioterrorism

Biological warfare can be described as the use of pathogens, microorganisms, or toxins that are used to cause intentional harm, illness, or death in war. Biologically manufactured weapons are diverse, ranging from viruses, fungi, proteins, or bacteria. Each type will affect the body differently, impacting the spread of the pathogen (Kolblentz 2004). Compared to other weapons of mass destruction, biological agents are deemed most dangerous because they have the capability to harm the greatest number of people in a single attack. This is because bioweapons are synthesized from plants or animals, making them infectious and therefore more lethal for living organisms [1]. Biological agents can also be transmitted to others that were not involved in the original attack, extending the overall lethality. Like biological warfare, bioterrorism can be described as the deliberate use of pathogenic organisms to cause harm. However, bioterrorism attacks target a vulnerable group of people and are typically backed by political, religious, or criminal ideologies [4]. Due to the potential severity and deadliness of these attacks, it is imperative that forensic biologists dedicate time and research to identify biological threats.

b. Public Concerns Regarding Bioterrorism

Due to their transmissible properties, one of the most immediate concerns of the use of biological agents is their ability to infect individuals long after an attack. Although not used as a biological weapon, the Polio virus can be used as an example to study the spread of disease. Polio virus is an infectious disease that is transmitted from person to person via airborne and/or physical contact. In the early 2000s there was a re-emergence of Polio virus. The virus was confined to four countries at first but spread rapidly. Between the years 1998 and 2003, researchers were seeing more than 1000 new cases of Polio worldwide, eventually leading to the re-emergence of Polio in 14 more countries. The virus spread from the Middle East, into Africa, and then made its way across the Red Sea. Countries that had never been exposed to or had eradicated Polio virus, began experiencing cases at an alarming rate. This medical event showed that pathogens can spread quickly and through different countries. Furthermore, it is revealed that attacks involving biological agents will have a stronger effect on Third World countries. [5]. Alarming transmission rates, such as these, lead to the need for vaccines. However, vaccines cannot be created if the source of the virus is not known. Therefore, it is crucial to identify the biological agent as quickly as possible. Once identified, transmission between individuals can be evaluated and vaccines, if applicable, administrated. One area that can help combat the infection and transmission of disease via biological agents is forensic microbiology.

c. Forensic Microbiology and Forensic Genetics

Forensic microbiology covers a wide range of matters including microbiology, genetics, and forensic science and is used to investigate, analyze, and identify microbial agents. Microbiology can aid criminal investigations, public health crises, and bioterrorism threats. One way to identify biological agents is by their genetic composition. With the advancement of next generation sequencing (NGS), forensic biologists can sequence whole genomes and fragmented DNA sequences. Thus, NGS has become widely accepted in forensic microbiology [3]. There are currently two chief methods used for microbial identification: amplicon-based sequencing and whole metagenome sequencing (WMS). Despite the success with these methods, there are severe disadvantages that could impact public health in a terrorist event. These include bias for primer selection or amplified regions that can contain sequences from more than one organism. WMS is a rather lengthy process and poses time and cost issues during data analysis. Compared to metagenomic methods, such as next generation sequencing, WGS targets larger stretches of genetic materials, such as insertion and deletion sequences. Contrarily, metagenomic techniques uses random primers and target smaller sequences that can be pieced together, which yield a high throughput value. [6]. Despite the success these technologies have had in forensic microbiology, a simpler, cost-effective, and non-bias approach needs to be established. If researchers use forensic human identification (HID) techniques and apply them to forensic microbiology, biological threats can be more easily identified.

Historically, there have been a variety of tools used to identify microorganisms. These include molecular sequencing, biochemical analysis, microbial culture, microscopy, and mass spectroscopy. Most techniques currently involve the use of DNA technologies. [4] Since the 1980's, DNA has been used to revolutionize forensic case work and continues to be the for

human identification. The advancement of DNA technology has improved, increasing the specificity and reliability of DNA sources collected at crimes scenes. Current methods for human identification testing involve short tandem repeats (STRs) that are amplified using polymerase chain reaction (PCR) methods that define a particular region of DNA, known as a locus. STRs have been the driving source for forensic DNA analysis, which primarily include HID and family relationship testing. STR's are known as micro-satellites and are made of non-coding regions of DNA. Typically, they range from two to six base pairs in length. STRs can be used as biological markers for trace evidence, such as body fluids, hair, or skin. [13]. PCR methodologies have been generally consistent since their adoption into forensic casework and are generally inexpensive, which makes them a reliable resource to use for downstream DNA testing. Moreover, STR markers are multiallelic and thus increase the discriminatory power when performing statistics. With the rapid increase of sequencing technologies, researchers were able to characterize microbial agents, advancing forensic molecular biology [7]. A study titled, *A review on the field of forensic science and different techniques that can be applied to identify the pathogen and perpetrator of a bio crime* published in 2002, applied DNA testing to forensic microbiology. Researchers based their methods off the identification that was completed for the anthrax attack, in 2001. They found that older methods were taxing and costly and therefore turned to repetitive PCR, (rep-PCR). This variation of PCR targets repetitive DNA sequences traditionally in microbial species and works to identify different strains between organisms.

In addition to rep-PCR researchers also turned to whole genome sequencing. These methods were successful, but the authors noted ways to better their study. One of the topics mentioned was the need to identify the bio agent urgently and that there is no set protocol for the identification of biological agents. Authors suggested that this might be due to the versatility of bio agents[8] . Another study turned to replicating the identification methods being done in the medical field, using PCR techniques for the identifications of biological agents. While this approach was successful, the authors thought that a more developed and more efficient approach was needed [9]. Rather than looking at traditional sequencing techniques, Escheonwu B. and colleagues turned their focus to the identification of biomarkers in reference to forensic microbiology identification techniques.

Biomarkers include SNPs, insertions or deletion sequences (INDELs), mobile events, specific genes, and gene transfer events. found that it was important to focus on specific biomarkers because of evolutionary divergence in bacterial species, leading to the emergence of polymorphisms. Additionally, a lot of bacteria come from the same species, but differ genetically. Authors from the study noted that it was crucial to recognize the variation in genetic composition between bacterial species and strains. Specific typing methods can help to identify individuals that belong to the same species. For example, Y. pestis to Y. pseudotuberculosis are both genetically similar but can be differentiated via SNP analysis [4]. In summary, the sudues discussed above recommend that there needs to be a more specific approach using known biomarkers, instead of traditional DNA testing. Despite the success with DNA testing results, there have been issues with DNA profiling. It is imperative that there is as little room for error as possible in a biological event.

d. Issues with DNA testing

DNA fragmentation occurs due to environmental, chemical, and enzymatic mechanisms which ultimately lead to hydrolytic depurination, followed by beta elimination. Chemical alterations such as the deamination of cytosines to uracil can also take place[3]. Degraded biological samples prove challenging for forensic biologists because they are not suitable for traditional STR testing. Low quality DNA samples are detected by capillary electrophoresis at 85-400 base pairs. However, degraded DNA samples may not reach this amount, increasing the chance for allelic dropout. [10] This issue extends to the ability to perform PCR, which will fail if the DNA is too fragmented [11]. Recent studies have incorporated RNA sequencing, rather than DNA sequencing due to RNA's abundancy and specificity in the genome. Unlike nuclear DNA, single-stranded RNA is abundant in the cell and makes up part of the transcriptome, which is representative of genes that are expressed. RNA is also tissue-specific and is therefore commonly used for fluid identification. These characteristics of RNA may make this molecule useful for RNA-based HID on low template DNA.[12] Due to RNA's versatility and advantageous properties, its incorporation into future HID studies should be considered.

Discussion

a. How RNA has been used in the past

The incorporation of RNA for forensic testing became popular after a study conducted in Germany in the 1980s. Since the publication of the study, the use of mRNA for forensic applications has been an area of interest for many forensic investigators. It has been found that RNA is more stable than DNA, and not as susceptible to degradation and postmortem changes. Because RNA is transcribed solely for a specific tissue, RNA is used for body fluid identification [13]. Transcriptome investigations are further being investigated for biogeographical ancestry, age of the donor, and information regarding postmortem changes. The most recent studies are looking to investigate coding region SNPs (cSNPs) in tissue specific transcripts because these lead to the identification of body fluids. One method of sequencing that has become favorable for using RNA SNPs is referred to as whole transcriptome shotgun sequencing. Even though the transcriptome only makes up 2% of genes, referred to as, protein- coding genes, this method of sequencing still provides sufficient SNP coverage [14]. The study below outlines the use and convenience of RNA SNP analysis.

Alberte Jepsen and colleagues investigated low template DNA with shotgun sequencing techniques. However, this study focuses on the use of RNA, rather than DNA. The mentioned researchers discuss advancements in RNA sequencing technologies and the capabilities of the transcriptome in their paper called, *Identification of individuals from low template blood samples using transcriptome shotgun sequencing* in 2024. Despite not having a SNP RNA database, this study stresses that genotyping such SNPs may revolutionize HID for forensic applications. Given that RNA SNPs have a lower mutation rate in comparison to STRs, have potential for automated genotyping, and can determine phenotype information, RNA sequencing is an ideal candidate for further evaluation for HID. There is no published RNA-based SNP database for genotype comparison, making RNA shotgun sequencing a novel technology for HID. Thus, the purpose of this study is to shed light on this technology and explore how transcriptome shotgun sequencing can be used to construct an RNA-based SNP HID analysis from low template sources.

Once the DNA was run and the data collected, investigators determined parameters for what they believed to be probative conditions for SNP genotyping. It was determined that only bi-allelic SNPs with a minor allele frequency greater than 0.1 in various populations should be preserved. Additionally, that there had to be a concordance between RNA and DNA genotypes assessed at each SNP. If the concordance rate between the two was above 95% then the SNP was retained. In degraded samples, if the RNA SNP did not meet the 95% concordance rate, the SNP was valued read depth and genotype quality. Lastly, the SNPs were screened and removed if their source stemmed from being pathogenic.

The study concluded that RNA SNP analysis offers promising opportunities for forensic science. The success for HID and body fluid identification leads researchers to believe that these techniques can be extended to other areas in forensic biology. Researchers that took part in this study were able to develop a prototype RNA-based SNP selection that consisted of 24 RNA based HID SNPs. These SNPs were selected because they were highly expressed and were validated in low copy number samples. Due to the success and foundational knowledge of this study, techniques used can be applied to implementation of a RNA SNP database for biological agent identification.

b. Future Studies and conclusions

From what has been described above, it is imperative that forensic biologists and microbiologists determine a method for the identification of biological agents quickly and efficiently. In doing so, public trust and health would be restored between the government and the general people.

Trust
between these two populations is crucial, especially in times of war or mass disaster [15]. While these topics serve as the foundation of safer communities, it is imperative to acknowledge additional advancements that can be made.

Forensic research indicates that STR DNA testing is efficient for HID methods and kinship analysis, however, may lack qualifications for forensic microbiology purposes. To establish a method for the identification of biological threats, it is crucial that the method is fast, inexpensive, and can detect changes in bacterial strains. Reviewing the studies described, RNA SNP analysis can be the solution. RNA SNP analysis allows for the identification of polymorphisms of rapidly evolving species as well as different strains of the same pathogen. RNA analysis can lead to more probative results because mRNA is more abundant in the cell than most nuclear DNA and is less susceptible to degradation. It has also been proven to be an effective technique for body fluid identification [13]. Looking at the study by Jepsen and colleagues, it was proved that RNA SNP analysis was a successful method for HID and body fluid identification (Jepsen et al., 2024). Thus these methods shoul d be extended to identify biological agents. to serve as quick and reliable research for forensic and law enforcement investigators. During this investigation, Jepsen and colleagues worked back from commonly expressed genes, using the transcriptome for their RNA SNP selection[12] This is a valuable and reliable process in biomarker selection because choosing SNPs that directly relate to gene expression increase the specificity of the biological marker.

Moreover, these methods need to be extended to establish an RNA SNP database. As mentioned in the paper, there has yet to be an SNP RNA database. Lacking a database or suitable set of biological markers makes the identification of biological agents expensive and time consuming. In a time of distress and uncertainty, the last thing needed is a laborious process identifying the issue. It is crucial to have a reliable standard operating procedure to follow in times of public health crisis. Not only can will the construction of a database be useful for forensic biology identification and aid the field of forensic microbiology, it will serve as the first line of defense against biological attacks.

References:

1. Croddy Eric, Wirtz JJ. *Weapons of Mass Destruction : An Encyclopedia of Worldwide Policy, Technology, and History*. ABC-CLIO; 2005.
2. Courts C, Gosch A, Rothschild M. RNA Analysis in Forensic Molecular Biology. *Dtsch Arztebl Int.Deutscher Arzte-Verlag GmbH*. 2024;121(11):363-369. doi:10.3238/arztebl.m2024.0051
3. Kapp JD, Green RE, Shapiro B. A fast and efficient single-stranded genomic library preparation method optimized for ancient DNA. *Journal of Heredity*. 2021;112(3):241-249. doi:10.1093/jhered/esab012
4. Echeonwu B, Nwankiti O, Chollom S, Olawuyi K. Bioterrorism threat: A review of microbial forensics source-tracing of some bioterrorism agents. *Journal of Forensic Science and Medicine.Wolters Kluwer Medknow Publications*. 2018;4(3):161-166. doi:10.4103/jfsm.jfsm_5_18
5. Kosal ME. A new role for public health in bioterrorism deterrence. *Front Public Health.Frontiers Media S. A*. 2014;2(DEC). doi:10.3389/fpubh.2014.00278
6. Spatz S, Afonso CL. Non-Targeted RNA Sequencing: Towards the Development of Universal Clinical Diagnosis Methods for Human and Veterinary Infectious Diseases. *Vet Sci.Multidisciplinary Digital Publishing Institute (MDPI)*. 2024;11(6). doi:10.3390/vetsci11060239
7. Oliveira M, Marszałek K, Kowalski M, et al. Sequencing Technologies in Forensic Microbiology: Current Trends and Advancements. *Forensic Sciences.Multidisciplinary Digital Publishing Institute (MDPI)*. 2024;4(4):523-545. doi:10.3390/forensicsci4040035
8. Ponce M, Pylant M, Becket E. *A Review on the Field of Forensic Science and Different Tech-Niques That Can Be Applied to Identify the Pathogen and Perpe-Trator of a Bio Crime*.; 2022. https://journals.calstate.edu/cjugr
9. Mir T ul G, Wani AK, Akhtar N, Sena S, Singh J. Microbial forensics: A potential tool for investigation and response to bioterrorism. *Health Sciences Review*. 2022;5:100068. doi:10.1016/j.hsr.2022.100068
10. Andersen MM, Kampmann ML, Jepsen AH, et al. Shotgun DNA sequencing for human identification: Dynamic SNP selection and likelihood ratio calculations accounting for errors. *Forensic Sci Int Genet*. 2025;74. doi:10.1016/j.fsigen.2024.103146
11. Nguyen R, Kapp JD, Sacco S, Myers SP, Green RE. A computational approach for positive genetic identification and relatedness detection from low-coverage shotgun sequencing data. *Journal of Heredity*. 2023;114(5):504-512. doi:10.1093/jhered/esad041
12. Jepsen AH, Kampmann ML, Jacobsen SB, Børsting C, Andersen JD. Identification of individuals from low template blood samples using whole transcriptome shotgun sequencing. *Forensic Sci Int Genet*. 2024;72. doi:10.1016/j.fsigen.2024.103089
13. Courts C, Gosch A, Rothschild M. RNA Analysis in Forensic Molecular Biology. *Dtsch Arztebl Int.Deutscher Arzte-Verlag GmbH*. 2024;121(11):363-369. doi:10.3238/arztebl.m2024.0051

14. Kling D, Jepsen AH, Kampmann ML, et al. Forensic investigative genetic genealogy using genotypes generated or imputed from transcriptomes. *Forensic Sci Int Genet*. 2025;78. doi:10.1016/j.fsigen.2025.103277
15. Richards EP, State L, O'brien T, Rathbun KC. *The Uses of Fear in Preventive Medicine*. Vol 10.; 1910. http://biotech.law.lsu.edu/blaw/Bioterror.htm.http://biotech.law.lsu.edu/cphl/history/articles/rosenau_fear.htm