

Rapid and In-situ Detection of Biological Agents via Near-Infrared Spectroscopy and Cloud-Powered Neural Networks

Chloé Leseigneur,^a Isabelle Radgen-Morvant,^b César Metzger,^c Pierre Esseiva,^b Daniel Croll,^a Pascal Miéville,^{d,*}

^a *Laboratory of evolutionary genetics, Institute of Biology, University of Neuchâtel, CH-2000 Neuchâtel, Switzerland*

^b *Ecole des sciences criminelles, Faculté de droit, des sciences criminelles et d'administration publique, Université de Lausanne, CH-1015 Lausanne, Switzerland*

^c *Spiez Laboratory, Federal Office for Civil Protection, CH-3700 Spiez, Switzerland*

^d *Swiss Cat+ West Hub, Ecole Polytechnique Fédérale de Lausanne EPFL, 1015 Lausanne, Switzerland*

* *Communication address: pascal.mieville@epfl.ch*

ABSTRACT

The contagious nature of certain biological agents and the difficulty in treating infections render them a significant threat to public health and safety. In situations involving chemical, biological, radiological, and nuclear (CBRN) agents, effective detection is of paramount importance to prevent the undetected spread of these agents and enable swift and targeted responses tailored to the specific threat. Although portable detection tools are effective for chemical and nuclear detection, current biological detection methods face several challenges, including limited mobility, extended processing times, and varying accuracy. In the context of biological threats, where agents such as anthrax can be rapidly dispersed due to environmental factors and human activities, rapid detection is of paramount importance. It is imperative to develop field-applicable detection devices that are highly selective, capable of differentiating between biological and non-biological agents, as well as benign and harmful microorganisms. This study examines the potential of near-infrared (NIR) spectroscopy in conjunction with machine learning as a rapid in-situ biological detection method. The objective is to distinguish between biological agents and common white powders that are used as confounding agents in suspect letters. The non-pathogenic surrogates employed are safe and representative of typical biological warfare agents. The near-infrared (NIR) spectra of lyophilized bacterial and fungal surrogates, along with common white powders, were subjected to analysis and processing through the application of principal component analysis (PCA) and hierarchical clustering analysis (HCA). This resulted in the successful classification of the samples into distinct groups. The classification model demonstrated high accuracy in its prediction, thereby emphasizing the potential of the method for field detection of solid biological agents. These promising results suggest that NIR spectroscopy combined with machine learning could be further investigated as a rapid in-situ tool for biological detection in CBRN contexts.

INTRODUCTION

CBRN (Chemical, Biological, Radiological and Nuclear) agents can be created intentionally or by accident, causing risks to human health, environment and animals.¹ CBRN threats extends to accidents and attacks of varying scales and impacts, including those perpetrated by terrorists, pandemics and epizootics. Emergency responders, such as fire brigades, other specialized units (such as the EEVBS in Switzerland or the ATF in Germany) – and specialised military and civil protection/civil defence units, are responsible for CBRN protection and aim to prevent and reduce the impact of CBRN agents.² CBRN detection technologies need to be practical, rugged and applicable in the field to identify threats and exclude as much as possible false positives. While all CBRN agents pose considerable risks, biological threats are associated with particular challenges, mostly because they can affect large groups and instil an atmosphere of fear and panic.³ Existing methods of detection are slow (>1h) and cannot be performed onsite as they require culture, and polymerase chain reaction (PCR). Biological threats require both emergency preparedness and technology development.

In the context of CBRN scenarios, *Bacillus anthracis*, the etiological agent of the anthrax disease, is recognized as a significant biological warfare agent in the 20th century.⁴ Following the events of 9/11 in 2001, letters containing *Bacillus anthracis* spores were sent to various medias and Senate members, resulting in 11 cases of inhalation and five deaths.⁵ The anthrax powder was processed in order to to increase its propensity for aerosolization by adding electrostatic charges, reducing clumping, and uniformizing particle size to facilitate inhalation in a process named militarization.⁶ Since 2001,

numerous anthrax fallacious threats involving harmless common white powder resembling flour, demonstrated the ease of fabricating such threats.⁷ Detecting anthrax spores in-situ remains challenging due to the need for rapid, accurate and reliable identification under high-stress conditions. The main difficulties in the detection of *Bacillus anthracis*, include its appearance as a harmless-looking powder, its potential for rapid dispersion and the necessity for prompt action to prevent widespread harm.⁸ Given its high relevance as a common biological threat to which CBRN units are potentially exposed,⁹ *Bacillus anthracis* was chosen as the target biological agent via non-pathogenic surrogates for this study on biological detection.

The most prevalent biological detection techniques encompass cell culture, polymerase chain reaction (PCR), enzyme-linked immunosorbent assay (ELISA), mass spectrometry, and biosensors. While cell culturing and plating are highly reliable, the necessity for several days to obtain results increases laboratory workloads and reliance on laboratory facilities.¹⁰ PCR technology, particularly real-time PCR, offers the potential for faster results within hours; however, it still requires laboratory settings and special equipment, which limits its field use.^{11,12} Similarly, ELISA also demands specific reagents and technical expertise.¹³ Mass spectrometry methods (e.g., MALDI-TOF) can accurately identify pathogen but are constrained by the need for complex equipment.¹⁴ Biosensors, on the other hand, offer a more rapid detection process; however, these technologies still require further development to achieve the necessary sensitivity and selectivity for reliable field use.^{15,16} Therefore, the principal limitation for the use of these techniques for CBRN response work is (a) their speed of detection/analysis (between 1 hour and many days depending on the method), (b) the lack of ruggedness of the equipment so that these cannot necessarily be transported and used in any field situation, and (c) their complexity of handling that may not be adapted for use by non-specialist first responder, military and civil defence/protection personnel.

Infrared spectroscopy is a method sensitive to chemical composition and architecture of molecules,¹⁷ applicable to biological systems, for example in bacterial strains identification.¹⁸ It is effectively used in the study of various biological materials, including lipids, proteins, peptides, bio-membranes, animal tissues, plants, clinical sample and microbial cells.¹⁹ The near-infrared (NIR) wavelength region of the infrared spectrum, extends from 13'000 to 4000 cm^{-1} , covering a wave-length range from 769.23 nm¹ to 2500 nm.¹⁹ Discovered in 1800,²⁰ NIR spectroscopy has been employed analytically since the 1950s.²¹ The method's key advantages are a non-destructive and non-invasive technology.²² NIR technology has also been used for qualitative and quantitative control purposes in the pharmaceutical context²³ as well as for food sample analysis.²⁴ Once calibrations is correctly performed, NIR becomes a relatively straightforward technique to use and can be made portable for terrain use.²⁵ The device used in this study is the microNIR. A portable drug detection tool using NIR technology from VIAVI. The cloud-based software environment used to acquire the data is NIRLAB that promote the decentralization of forensic capabilities.²⁶ NIRLAB identification is based on NIR spectra post-processing and neural network training. Similar methods have been used in previous studies using NIR spectroscopy for bacteria detection. For example, Kammies et al. report the analysis and differentiation of strains *B. cereus* and *E.coli*, *S.epidermidis* by hyperspectral NIR and PCR post treatment.²⁷ Another study by Jiang et al. also demonstrated that PCA and HCA can effectively classify NIR spectra of different samples.²⁸ Multiple studies have emphasized the effectiveness of combining NIR spectroscopy with machine learning algorithms, highlighting the utility of supervised methods such as neural networks and kNN in spectra analysis.^{25,29,30}

The objective of this study is to assess the potential of portable near-infrared spectroscopy (NIR), as described above, in conjunction with cloud-based machine learning algorithms for the rapid and precise in-situ detection of biological agents of the bacillus type and discrimination with respect to commonly found white powders that may be used as fallacious threats and other biological agents such as fungi. In accordance with the preceding text, the *B. anthracis* has been selected as the target biological agent first due to its high importance in the CBRN context and second due to its compatibility with NIRLAB equipment, particularly its powdered weaponized form. In order to train the neural network and to ensure the safety of the experiment, non-pathogenic surrogates were selected based on their suitability for experimentation and their diversity. This selection allows for a comprehensive evaluation while avoiding the inherent dangers associated with highly pathogenic strains. The study encompasses an examination of common white powders, including baking soda, sugar, and talcum, which are pertinent in the context

of anthrax fallacious threats.³¹ This approach aims to enhance the application of NIR spectroscopy in public safety and CBRN scenarios, ensuring rapid and reliable detection of potential biological threats in situ.

METHODS

Biological non-pathogenic surrogates

Twenty-three biological non-pathogenic surrogates were selected, comprising 11 bacteria (see Table 1) and 12 fungi (see Table 2). The selection of each surrogate was made with due consideration to their safety, while maintaining proximity to biological agents. In this context, proximity is understood in terms of biochemical structure for the bacteria and physical appearance for fungal powders. The objective was to obtain a diverse range of samples and to prepare them in the closest possible powdered form.

Bacteria	Comments	References
<i>Bacillus cereus</i>	common soil and plant dwelling bacterium, spore-forming, can be aerosolized, opportunistic human pathogen (inter alia food safety threat), closely related to <i>Bacillus anthracis</i> * and <i>B. thuringiensis</i>	32
<i>Bacillus thuringiensis</i>	soil dwelling bacterium, common biocontrol agent, suspected opportunistic pathogen (rare occurrences), closely related to <i>Bacillus anthracis</i> * and <i>B. cereus</i>	33,34
<i>Bacillus amyloliquefaciens</i>	used as biocontrol agent in agriculture, synthesizes a narrow-spectrum natural antibiotic against <i>B. anthracis</i>	35,36
<i>Bacillus subtilis</i>	soil and gastrointestinal (e.g., humans, ruminants) bacterium, model species for the <i>Bacillus</i> genus, important industrial organism	37
<i>Escherichia coli</i>	normal part of the gastrointestinal microflora of humans, model bacterium, many pathotypes can cause disease in humans (food safety threat)**	38
<i>Pseudomonas putida</i>	soil dwelling bacterium, used as biocontrol agent in agriculture, used in bioremediation, opportunistic human pathogen	39
<i>Enterococcus faecalis</i>	normal part of the gastrointestinal microflora of humans, opportunistic human pathogen	40
<i>Staphylococcus epidermidis</i>	normal part of the skin microbiota of humans, opportunistic pathogen, closely related to human pathogen <i>Staphylococcus aureus</i> (of note: the CDC classifies staphylococcal enterotoxin B as a category B bioterrorism agent, other staphylococcal toxins are common causative agents of food poisoning)	41,42
<i>Paracoccus spp.</i>	i.e. <i>Paracoccus yeei</i> , opportunistic human pathogen (rare accounts)	43
<i>Burkholderia terrae</i>	soil dwelling bacterium, non-pathogenic, sister species to pathogenic <i>Burkholderia mallei</i> ** and <i>Burkholderia pseudomallei</i> **	
<i>Lactobacillus rhamnosus</i>	the <i>Lactobacillus</i> genus has been studied with NIR as food contaminant	44

Table 1. List of bacteria used in the present study as *b. anthracis* non-pathogenic surrogates with comments on selection criteria. * classified by the U.S. CDC as category A bioterrorism agent, ** classified by the U.S. CDC as category B bioterrorism agent⁴²

Fungi	Comments	References
<i>Saccharomyces cerevisiae</i>	model yeast species, opportunistic human pathogen	45
<i>Gliocladium catenulatum</i>	used as biocontrol agent in agriculture	46

<i>Ophiocordyceps heteropoda</i>	known instances of human poisoning cases	47
<i>Elaphocordyceps subsessilis</i>	soil fungus known to naturally produce ciclosporin A	48,49
<i>Candida albicans</i>	prevalent opportunistic human pathogen	50
<i>Trichoderma spp.</i>	soil dwelling fungi, several strains are used as biocontrol agents in agriculture, <i>Trichoderma longibrachiatum</i> (common house mold) can be toxic to humans, produces mycotoxins	51
<i>Clonostachys rosea</i>	used as biocontrol agent in agriculture	52
<i>Fusarium graminearum</i>	soil dwelling fungi, produces mycotoxins	51
<i>Aspergillus niger</i>	opportunistic human pathogen, related to the human pathogen <i>Aspergillus fumigatus</i>	53
<i>Exophiala xenobiotica</i>	an opportunistic black yeast, <i>Exophiala</i> is one of the rare fungi causing infection in immunocompromised people	54
<i>Penicillium commune</i>	common fungi contaminating cheese, it also produces mycotoxins	55,56
<i>Alternaria spp.</i>	can be plant pathogens, producing mycotoxins implicated in mammals' cancer, it has potential of emergence as a human pathogen	57

Table 2. List of fungi used in the current study as non-bacterial biological *b. anthracis* surrogates with comments on selection criteria.

All strains mentioned in Tables 1 and 2 were obtained from the Microbiology laboratory collection at the University of Neuchâtel. All bacteria were cultivated on agar plates, from which colonies were extracted and inoculated into a nutrient broth solution of 200ml. The same process was followed for fungi except they were inoculated into malt broth. The culturing process required different times for each biological species. The suspensions were distributed into four vials of 50ml, which were then centrifuged at 900 rpm for three minutes at 8600 g. Subsequently, the samples were lyophilized at -50°C and 0.05 mbar for 24 hours. Following lyophilisation, all specimens were crushed with a mortar prior to analysis. To compare to the biological spectra, 21 common white powders were chosen. The common white powders represent typical anthrax fallacious threats, mostly easy to access. The common white powders can be classified into four categories (Table 3.).

Categories	Powders	Formula	Comments
Cooking	Crystalized sugar	$C_{12}H_{22}O_{11}$	
Cooking	Cornstarch	$(C_6H_{10}O_5)_n$	Sugar polymer
Cooking	Kitchen salt	NaCl	
Salts	Sodium carbonate	Na_2CO_3	
Salts	Sodium bicarbonate	NaHCO ₃	
Salts	Potassium carbonate	K_2CO_3	
Salts	Dipotassium phosphate	K_2HPO_4	
Salts	Disodium hydrogen phosphate	Na_2HPO_4	
Salts	Potassium dihydrogen phosphate anhydrous	KH_2PO_4	Monopotassium phosphate
Salts	Magnesium chloride anhydrous	$MgCl_2$	
Salts	Ammonium chloride	NH_4Cl	
Household and industrial products	Talcum	$Mg_3Si_4O_{10}(OH)_2$	
Household and industrial products	Calcium carbonate	$CaCO_3$	Sampled from chalk
Household and industrial products	Anhydric citric acid	$C_6H_8O_7$	

Household and industrial products	Celite		Diatomaceous earth
Medicines and supplements	Aspirin		Acetylsalicylic acid
Medicines and supplements	Neocitran		Medicine for cold and flu symptoms
Medicines and supplements	L-isoleucine		Amino acid
Medicines and supplements	L-menthol		
Medicines and supplements	L-proline		Amino acid
Medicines and supplements	Propylparaben		Preservative
Medicines and supplements	Ethylparaben		Preservative

Table 3. Samples defined in the text as common white powder, mostly easily accessible, corresponding in the current study as non-biological fallacious threats encountered by CBRN responders.

NIR Analysis.

The MicroNIR Onsite W 1700 from VIAVI was employed as a portable near-infrared (NIR) tool, connected to the NIRLAB Mobile application via Bluetooth.²⁶ The system operates within the spectral region spanning 950 to 1650 nm. The biological samples (Tables 1 and 2) were lyophilised using the Lyoquest Telstar laboratory freeze dryer, while non-biological common white powders (Table 3) were crushed with a mortar in the case of grains (e.g., chalk) and directly sampled. The resulting lyophilised powders were weighed and transferred to aluminum plates (Figure 1) with a spatula. Aluminum plates are optimal to minimize reaction with the sample, interference, or beam absorption during the analysis with the microNIR device. The samples were arranged in a manner that ensured coverage of the device's beam, thereby allowing for an effective analysis.



Figure 1. The microNIR device used to analyse all samples on a typical aluminium plate. In this specific example, the sample is crystalized sugar.

The microNIR device was positioned in direct contact with the samples. Upon activation of the microNIR device, the resulting spectra were promptly recorded in the NIRLAB application. Each substance was subjected to three separate samplings and each sample was subjected to three replicate analyses for signal averaging, resulting in a total of nine spectra per powder. On some samples, depending on the acquired spectra quality, a fourth replicate analysis was performed.

Data processing and neural-network training

The combination of the sampling and measurement replication method described above with the list of samples (Tables 1, 2 and 3) yielded a total of 471 spectra, which were exported in Excel format (.csv) to Orange Data Mining software for subsequent data analysis. In order to identify a fingerprint of the substance with NIR, it is necessary to reduce the noise and increase the signal from the chemical information. This can be achieved through the application of data pre-processing. The pre-processing methods employed were the standard normal variate (SNV) to normalise the spectra and the Savitzky-Golay filter with and without a second derivative and a second-order polynomial. The selected spectral range was trimmed and retained between 950 and 1641.8 nm. The initial analysis was conducted using principal component analysis (PCA), with the objective of identifying any inconsistencies or outliers present in the data set.²⁵ PCA method enables the reduction of the dimensionality of multivariate data,⁵² while ensuring the maintenance of its intrinsic information content. This method accounts for the majority of the data variance. A preliminary PCA was conducted on all spectra. In order to account for the intra-variability within a given group (species or substance), an average spectrum is calculated for each fungal and bacterial species, as well as for each non-biological common white powder. Hierarchical clustering analysis (HCA) is a valuable method for organizing samples based on their pairwise similarity.⁵⁸ The

similarity between the samples was assessed using the HCA method, with the cosine distance measure employed as the metric. The pre-processed samples of all spectra were subjected to HCA. A distance map (heat map), based on the averaged spectra, was constructed to facilitate the visualization of clusters within the data set. The samples were randomized and divided into two groups: a calibration group comprising two-thirds of the samples and a prediction group comprising the remaining third. To assess the efficacy of the classification models, a series of machine learning algorithms, including neural networks, support vector machines (SVM), and k-nearest neighbors (kNN), were employed. The performance of these models was evaluated using a range of performance metrics, including the area under the receiver operating characteristic curve (AUC), classification accuracy (CA), F1 score, precision, recall, and the Matthews correlation coefficient (MCC). This comprehensive approach ensured a thorough evaluation of the models' performance.⁵⁹ A confusion matrix was constructed for the purpose of identifying specific instances that had been misclassified and of ascertaining the proportion of instances that had been predicted. The confusion matrix elucidates the patterns of errors inherent to the models.

RESULTS

As previously stated, a total of 471 spectra were acquired. Figure 2 depicts the NIR spectra after standard normal variate (SNV) and second derivative pre-processing, which demonstrate considerable chemical variability, particularly evident in the talcum samples, which exhibit elevated peaks at approximately 1370 and 1390 nm. The spectra of bacteria and fungi display greater clustering, with analogous peaks observed within each group.

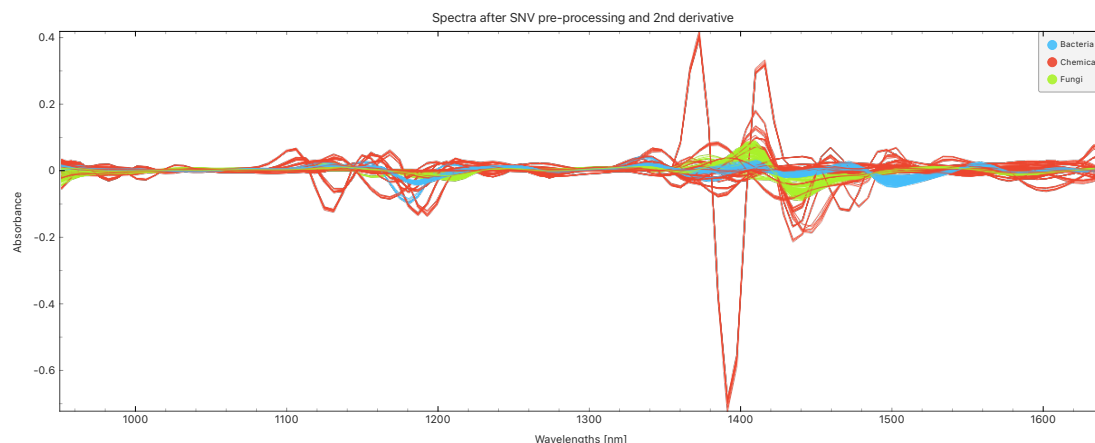


Figure 2. NIR spectra of bacteria (blue), fungi (green) and common white powders (Chemical, red) after SNV and second derivative pre-processing.

A PCA was conducted on all aforementioned spectra following pre-processing to reduce the dimensionality of the dataset while retaining the most significant variance. This enabled the discovery of patterns and relationships that might otherwise have remained obscured in the high-dimensional spectral data. The resulting data are presented in Figure 3 for the reader's convenience. A more segregative representation is provided by the PCA of the averaged spectra per sample. The averaging reduces the variability between individual samples and demonstrates a more pronounced separation between the groups (Figure 4). A clear segregation is evident in the PCA plot of mean spectra, with bacteria and fungi forming discrete clusters. However, the segregation is less pronounced regarding the common white powder substances, as several can be found within the fungal group, including cornstarch, sodium carbonate, sodium bicarbonate, and dipotassium phosphate. Furthermore, propylparaben was the only chemical substance found in the bacterial group.

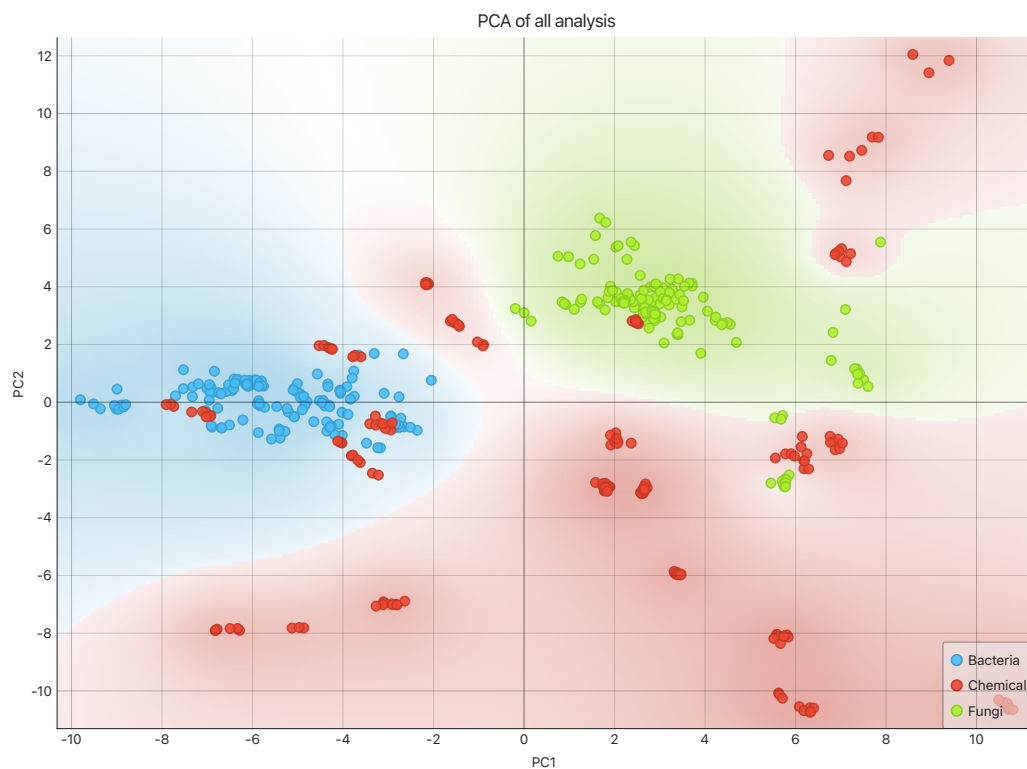


Figure 3. PCA of the 421 spectra collected on the samples listed in Tables 1 to 3, following the SNV and second derivative pre-processing as shown in Figure 2. The red zone represents the group of non-biological common white powders, the blue zone corresponds to the bacterial group of spectra, and the green zone corresponds to the fungal group.

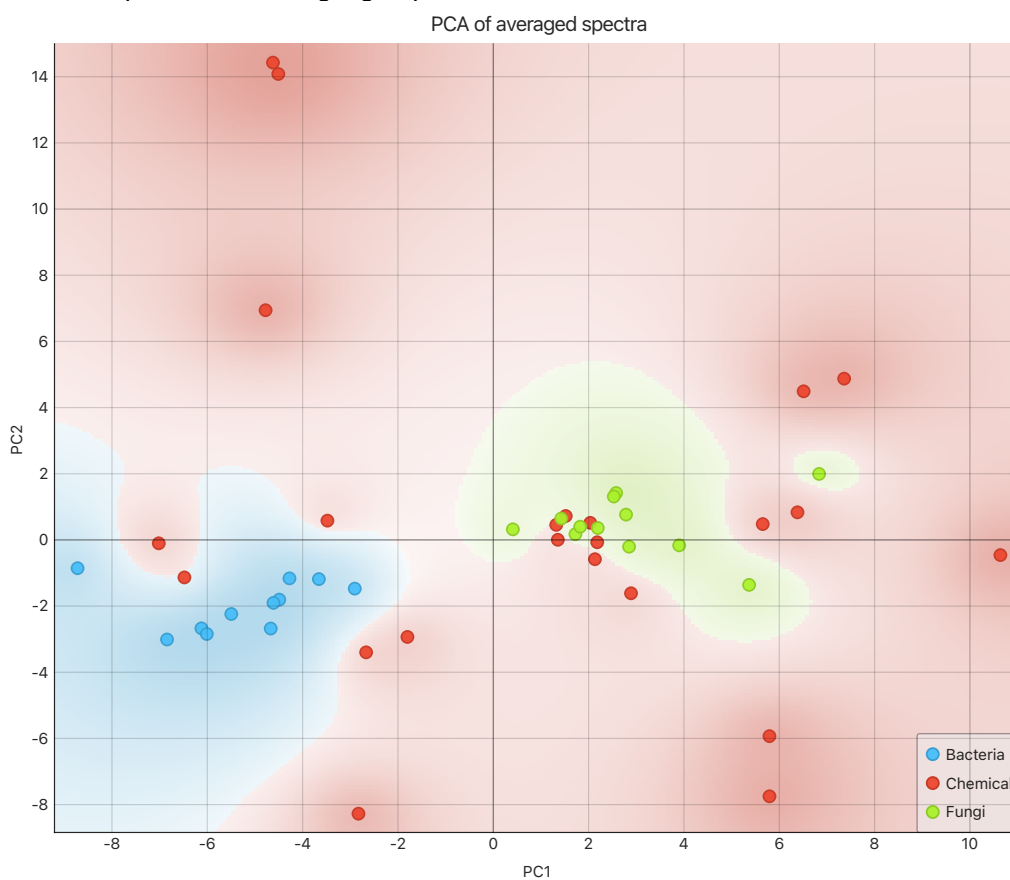


Figure 4. PCA of mean spectra of bacteria (blue), chemical (red) and fungi (green).

In order to evaluate the degree of spectral discrimination achieved through PCA, a hierarchical cluster analysis (HCA) was performed on the same set of 421 spectra. The aim is to conduct a comprehensive evaluation to ascertain whether the replicates within the dataset demonstrate discernible clustering patterns or whether there is a notable degree of overlap between the groups. The results of the HCA, presented as a distance map in Figure 5, indicate that bacterial species exhibit greater clustering tightness (see the main centered diagonal blue square), highlighting stronger within-group spectral similarity compared to other groups. It is notable that non-biological common white powder samples form discrete, smaller clusters than biological ones and are predominantly located in the lower half of the diagonal. Moreover, some common white powders, including magnesium chloride, ammonium chloride, and Neocitran, diverge from the primary common white powder clusters, forming smaller clusters at the top of the diagonal. This suggests that these compounds possess spectral characteristics analogous to those observed in other sample types. These observations are consistent with the PCA results, which identified two major bacterial and fungal clusters surrounded by greater variation in the common white powder spectra. Therefore, the combination of PCA and HCA provides a robust framework for evaluating the degree of spectral discrimination and clustering tendencies across the dataset.

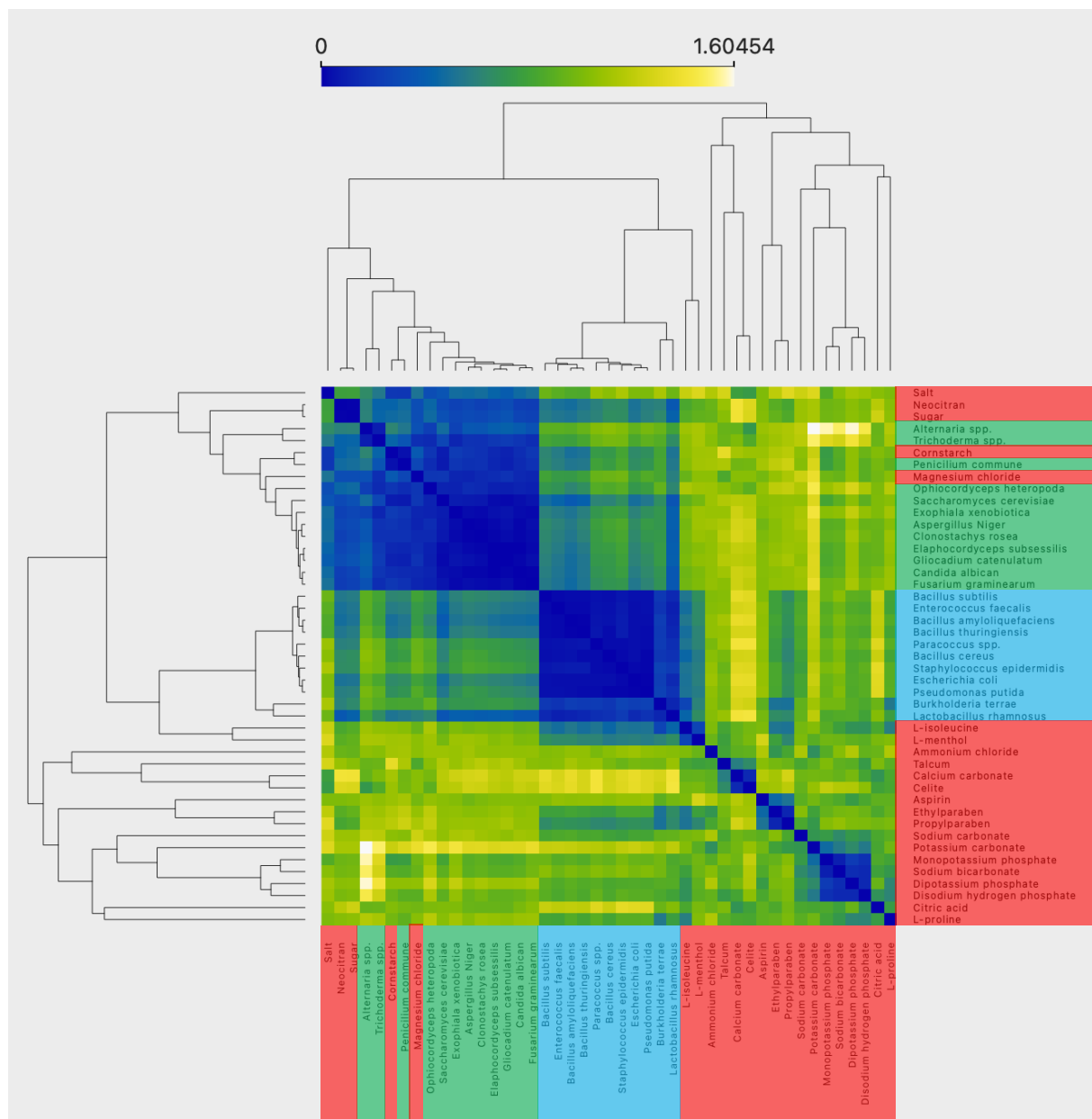


Figure 5. HCA with a distance map visualization. The mean spectra of the samples are used with cosine distance values, with dark blue representing a high degree of similarity and red, blue, and green representing progressively lower degrees of similarity. The legends have the same colors as the PCA, with red indicating chemicals, blue indicating bacteria, and green indicating fungi. The left top square is predominantly composed of fungal samples, while the middle square is composed of bacterial samples.

A model performance evaluation was conducted to test the robustness of the method. The performance of the classification of the dataset was evaluated using a neural network (NN), a support vector machine (SVM) and a k-nearest neighbour (kNN) model. The results demonstrated that all models exhibited strong performance. While SVM and kNN also performed well, the NN model achieved the best classification score across all metrics (Figure 6.a). The confusion matrix provided the proportion of instances between the predicted group and the calibration group for the NN model (Figure 6.b), confirming the accuracy of the chosen model.

Model	AUC	CA	F1	Prec	Recall	MCC
Neural Network	1.000	1.000	1.000	1.000	1.000	1.000
SVM	1.000	0.983	0.983	0.984	0.983	0.974
kNN	0.999	0.994	0.994	0.994	0.994	0.991

		Predicted			Σ
		Bacteria	Chemical	Fungi	
Actual	Bacteria	100.0 %	0.0 %	0.0 %	482
	Chemical	0.0 %	100.0 %	0.0 %	678
	Fungi	0.0 %	0.0 %	100.0 %	450
Σ		482	678	450	1610

Figure 6.a This table presents the performance scores of each model, NN, SVM and kNN. The performance metrics used are area under ROC curve (AUC), classification accuracy (CA), F1 score, precision, recall and Matthews correlation coefficient (MCC). **6.b** The second table shows the confusion matrix of the NN model.

DISCUSSION

The objective was to assess the potential of the portable near-infrared (NIR) spectrometer MicroNIR Onsite W 1700 from VIAVI, in conjunction with NIRLAB cloud-powered NIR spectroscopy (NN), as an in-situ rapid method for the detection of chemical, biological, radiological, and nuclear (CBRN) biological agents. This was achieved through the analysis of over 11 different lyophilised bacteria, 12 different lyophilised fungi, and 21 common, easily accessible white powders. In this context, lyophilised bacteria and fungi serve as non-pathogenic surrogates for typical militarised *B. anthracis*, allowing for the demonstration of the feasibility of rapid and significant discrimination between biological and non-biological common white powder groups of samples. The trained NN also demonstrates the possibility of automatically classifying a potential new powdery sample measured with the MicroNIR spectrometer with a high degree of confidence, confirming the possibility of using this apparatus in combination with the NIRLAB software to perform CBRN in-situ biological agent identification.

It is necessary to provide a commentary on the 100% discrimination achieved on the NN (Figure 6.b). The quality of the spectra acquisition and subsequent classification is undoubtedly influenced by the specific characteristics of the samples, including their granularity and homogeneity. These attributes can be subject to variability due to the inherent unpredictability of the real-world conditions employed in the preparation of microorganisms or common white powders.⁶⁰ For example, the lyophilisation, which usually preserves viability but can affect cellular structures and impact detection of surface protein,⁶¹ could impact the spectrum. It would also be relevant to explore the influence of the device temperature on the bacterial spectrum, as well as the operational temperature of the NIR devices in field conditions.⁶² Furthermore, historical instances of biological attacks utilising *B. anthracis* have involved the processing and treatment of the bacteria in a manner that was specifically designed to suit its intended use as a weapon.⁶ Such treatments could potentially impact the NIR analysis of the biological agent. Further experimentation with authentic *B. anthracis* and potentially other biological warfare agents is necessary to corroborate the findings presented in this study and to fully validate its use in operation.

The PCA facilitated the visualization of the differentiation between the biological groups, including bacteria and fungi. In addition, the analysis revealed a notable diversity of the common white powders. The results of the PCA visualization are promising and are corroborated by the HCA, which also reveals the existence of the three groups of common white powders. A number of the common white powders

exhibit a closer proximity to the fungi than to the other common white powders. Among these substances, cornstarch, sugar, salt, and magnesium chloride are included in the fungi cluster. The fact that cornstarch is a polysaccharide, which is quite similar in structure to chitin, itself a polysaccharide and a primary component of the cell wall of fungi, may explain why this powder is related to fungi. The same explanation is likely also valid for the proximity of other sugars. As previously noted, the neural network model yielded optimal predictions for the three groups, substantiating its resilience and underscoring prospective avenues for advancement. Nevertheless, it is important to exercise caution, as a 100% classification results may indicate overfitting, especially considering the small size of the dataset. A larger dataset would confirm the high specificity and sensitivity of the method by reducing the risk of overfitting.⁶³ These results nevertheless underscore the interest of advanced machine learning algorithms combined with portable NIR tools for rapid and accurate analysis of biological samples.

A further potential limitation of a spectral approach is that it does not separate substances and provides an average spectrum of the measured mixture. Consequently, the effect of an eventual matrix used for the culture or the dispersion of the biological agent can be significant.²⁶ Further research could investigate the discrimination of a biological agent within a non-biological dispersion matrix and the combination of different biological strains. The developed method is dependent on the ability to recognize substances within an existing database. It is therefore unable to identify unknown types of molecules or substances for which the NN has not been trained. This highlights the necessity of creating a larger database of biological agents and potential fallacious threats. Furthermore, the focus of this study is on bacterial detection, which overlooks significant threats like viruses.

As previously stated, the objective of this study is to illustrate the potential of rapid and in situ differentiation between biological and non-biological samples to alert CBRN responders. The initial objective was not to identify a specific biological agent. Nevertheless, the demonstrated differentiation between fungi and bacteria and the ability to distinguish each bacterial strain on the PCA provide a promising foundation for exploring the potential of selective identification. This could prove a valuable addition to CBRN operations, and it may also have applications in the detection of pathogenic strains in other fields, including medical microbiology, food science, and forensic sciences. To extend the potential of the method, it would be advantageous to create a database of NIR spectra that involves the various stakeholders addressing all aforementioned domains. In this context, the cloud-based neural network would facilitate the rapid updating of information on pathogens, ensuring that the database remains current and accessible to a large group of actors and that the NN is continuously updated.

CONCLUSION

This study demonstrates the potential of near-infrared portable spectroscopy in conjunction with cloud-based neural networks for expeditious, in-situ biological threat detection in contexts pertaining to chemical, biological, radiological, and nuclear (CBRN) incidents. The application of a cloud-based machine learning algorithm enables the effective classification of the studied groups and the discrimination of samples between biological and non-biological. It is proposed that NIR could serve as a rapid and simple-to-use biological detection device for biological CBRN emergencies. Moreover, the application of this technology could be extended to other fields, including microbiology, food analysis, and forensic science. In this context, the creation of a NIR spectra database of biological agents and non-biological potential fallacious threats involving the various stakeholders that address microbiological challenges would be beneficial.

Acknowledgments

The authors acknowledge the ETH Domain for its support through the Forschungsinfrastrukturen Program. They would also like to acknowledge the complementary support offered by Cap. Bernard Tschopp and Col. Nicolas Schumacher of the "Service Incendie et Secours" of Geneva. The authors would like to thank Florentin Coppey of the University of Lausanne and NIRLAB founder for providing the microNIR device and support. We thank Prof. Pilar Junier and lab assistant Ilona Palmieri of the microbiology lab of the University of Neuchâtel for their assistance.

Authors contributions

Conceptualization : P.M, C.L, C.M, P.E, I.R-M

Funding : P.M, D.C

Methodology : P.M, D.C, C.M, I.R-M

Resources : P.E, D.C

Software : P.E, I.R-M, C.L

Supervision: P.M, D.C
Writing: C.L, P.M, D.C
Revision: P.M, D.C

Bibliography

- (1) Gawlik-Kobylińska, M. Current Issues in Combating Chemical, Biological, Radiological, and Nuclear Threats to Empower Sustainability: A Systematic Review. *Appl. Sci.* **2022**, *12* (16), 8315. <https://doi.org/10.3390/app12168315>.
- (2) Alim, F.; Barrett, J.; Arcilesi, R.; Ewing, K.; Lattin, C. *An Introduction to Biological Agent Detection Equipment for Emergency First Responders*, NIJ GUIDE 101-00.; ISSUES IN INTERNATIONAL CRIME, 17 june 20215.
- (3) Ozaslan, D. M. Current Studies in Health and Life Sciences 2023. **2023**.
- (4) Spencer, R. C. Bacillus Anthracis. *J. Clin. Pathol.* **2003**, *56* (3), 182–187. <https://doi.org/10.1136/jcp.56.3.182>.
- (5) Jernigan, J. A.; Stephens, D. S.; Ashford, D. A.; Omenaca, C.; Topiel, M. S.; Galbraith, M.; Tapper, M.; Fisk, T. L.; Zaki, S.; Popovic, T.; Meyer, R. F.; Quinn, C. P.; Harper, S. A.; Fridkin, S. K.; Sejvar, J. J.; Shepard, C. W.; McConnell, M.; Guarner, J.; Shieh, W.-J.; Malecki, J. M.; Gerberding, J. L.; Hughes, J. M.; Perkins, B. A. Bioterrorism-Related Inhalational Anthrax: The First 10 Cases Reported in the United States - Volume 7, Number 6—December 2001 - Emerging Infectious Diseases Journal - CDC. <https://doi.org/10.3201/eid0706.010604>.
- (6) Matsumoto, G. Anthrax Powder: State of the Art? *Science* **2003**, *302* (5650), 1492–1497. <https://doi.org/10.1126/science.302.5650.1492>.
- (7) Leask, A.; Delpech, V.; McAnulty, J. Anthrax and Other Suspect Powders: Initial Responses to an Outbreak of Hoaxes and Scares. *New South Wales Public Health Bull.* **2003**, *14* (12), 218. <https://doi.org/10.1071/NB03059>.
- (8) Pugliese, G.; Favero, M. S. Anthrax as a Biological Weapon: Updated Recommendations for Management. *Infect. Control Hosp. Epidemiol.* **2002**, *23* (7), 411–411. <https://doi.org/10.1017/S0195941700083089>.
- (9) Kaszeta, D. The CBRN and Hazardous Materials Threat. In *CBRN and HAZMAT Incidents at Major Public Events*; John Wiley & Sons, Ltd, 2022; pp 1–18. <https://doi.org/10.1002/9781119743088.ch1>.
- (10) Lazcka, O.; Campo, F. J. D.; Muñoz, F. X. Pathogen Detection: A Perspective of Traditional Methods and Biosensors. *Biosens. Bioelectron.* **2007**, *22* (7), 1205–1217. <https://doi.org/10.1016/j.bios.2006.06.036>.
- (11) Chen, M.; Lan, X.; Zhu, L.; Ru, P.; Xu, W.; Liu, H. PCR Mediated Nucleic Acid Molecular Recognition Technology for Detection of Viable and Dead Foodborne Pathogens. *Foods* **2022**, *11* (17), 2675. <https://doi.org/10.3390/foods11172675>.
- (12) Levi, K.; Smedley and, J.; Towner, K. J. Evaluation of a Real-Time PCR Hybridization Assay for Rapid Detection of *Legionella Pneumophila* in Hospital and Environmental Water Samples. *Clin. Microbiol. Infect.* **2003**, *9* (7), 754–758. <https://doi.org/10.1046/j.1469-0691.2003.00666.x>.
- (13) Chen, C.; Ji, Y.; Liu, J.; Shi, H.; Ding, W. he. ELISA Immunological Reagents. In *In Vitro Diagnostic Industry in China*; Song, H., Dai, L., Eds.; Springer Nature Singapore: Singapore, 2023; pp 151–157. https://doi.org/10.1007/978-981-99-3110-1_7.
- (14) Clark, A. E.; Kaleta, E. J.; Arora, A.; Wolk, D. M. Matrix-Assisted Laser Desorption Ionization–Time of Flight Mass Spectrometry: A Fundamental Shift in the Routine

- Practice of Clinical Microbiology. *Clin. Microbiol. Rev.* **2013**, *26* (3), 547–603.
<https://doi.org/10.1128/cmr.00072-12>.
- (15) Bollella, P.; Katz, E. Biosensors—Recent Advances and Future Challenges. *Sensors* **2020**, *20* (22), 6645. <https://doi.org/10.3390/s20226645>.
- (16) Chen, C.; Wang, J. Optical Biosensors: An Exhaustive and Comprehensive Review. *Analyst* **2020**, *145* (5), 1605–1628. <https://doi.org/10.1039/C9AN01998G>.
- (17) Barth, A. Infrared Spectroscopy of Proteins. *Biochim. Biophys. Acta BBA - Bioenerg.* **2007**, *1767* (9), 1073–1101. <https://doi.org/10.1016/j.bbabi.2007.06.004>.
- (18) Naumann, D. Ft-Infrared and Ft-Raman Spectroscopy in Biomedical Research. *Appl. Spectrosc. Rev.* **2001**, *36* (2–3), 239–298. <https://doi.org/10.1081/ASR-100106157>.
- (19) Stuart, B. Infrared Spectroscopy. In *Kirk-Othmer Encyclopedia of Chemical Technology*; John Wiley & Sons, Ltd, 2015; pp 1–18.
<https://doi.org/10.1002/0471238961.0914061810151405.a01.pub3>.
- (20) Herschel, W. XIII. Investigation of the Powers of the Prismatic Colours to Heat and Illuminate Objects; with Remarks, That Prove the Different Refrangibility of Radiant Heat. To Which Is Added, an Inquiry into the Method of Viewing the Sun Advantageously, with Telescopes of Large Apertures and High Magnifying Powers. *Philos. Trans. R. Soc. Lond.* **1997**, *90*, 255–283. <https://doi.org/10.1098/rstl.1800.0014>.
- (21) Porep, J. U.; Kammerer, D. R.; Carle, R. On-Line Application of near Infrared (NIR) Spectroscopy in Food Production. *Trends Food Sci. Technol.* **2015**, *46* (2, Part A), 211–230. <https://doi.org/10.1016/j.tifs.2015.10.002>.
- (22) Blanco, M.; Villarroya, I. NIR Spectroscopy: A Rapid-Response Analytical Tool. *TrAC Trends Anal. Chem.* **2002**, *21* (4), 240–250. [https://doi.org/10.1016/S0165-9936\(02\)00404-1](https://doi.org/10.1016/S0165-9936(02)00404-1).
- (23) Ciurczak, E. W.; Igne, B. *Pharmaceutical and Medical Applications of Near-Infrared Spectroscopy*, 2nd ed.; CRC Press: Boca Raton, 2014.
<https://doi.org/10.1201/b17136>.
- (24) Li-Chan, E.; Chalmers, J.; Griffiths, P. *Applications of Vibrational Spectroscopy in Food Science*; Wiley, 2010.
- (25) Manley, M. Near-Infrared Spectroscopy and Hyperspectral Imaging: Non-Destructive Analysis of Biological Materials. *Chem. Soc. Rev.* **2014**, *43* (24), 8200–8214. <https://doi.org/10.1039/C4CS00062E>.
- (26) Coppey, F.; Bécue, A.; Sacré, P.-Y.; Ziemons, E. M.; Hubert, P.; Esseiva, P. Providing Illicit Drugs Results in Five Seconds Using Ultra-Portable NIR Technology: An Opportunity for Forensic Laboratories to Cope with the Trend toward the Decentralization of Forensic Capabilities. *Forensic Sci. Int.* **2020**, *317*, 110498. <https://doi.org/10.1016/j.forsciint.2020.110498>.
- (27) Kammies, T.-L.; Manley, M.; Gouws, P. A.; Williams, P. J. Differentiation of Foodborne Bacteria Using NIR Hyperspectral Imaging and Multivariate Data Analysis. *Appl. Microbiol. Biotechnol.* **2016**, *100* (21), 9305–9320. <https://doi.org/10.1007/s00253-016-7801-4>.
- (28) Jiang, Q.; Zhang, M.; Mujumdar, A. S.; Wang, D. Non-Destructive Quality Determination of Frozen Food Using NIR Spectroscopy-Based Machine Learning and Predictive Modelling. *J. Food Eng.* **2023**, *343*, 111374. <https://doi.org/10.1016/j.jfoodeng.2022.111374>.
- (29) Lorente, D.; Aleixos, N.; Gómez-Sanchis, J.; Cubero, S.; Blasco, J. Selection of Optimal Wavelength Features for Decay Detection in Citrus Fruit Using the ROC Curve

- and Neural Networks. *Food Bioprocess Technol.* **2013**, 6 (2), 530–541.
<https://doi.org/10.1007/s11947-011-0737-x>.
- (30) Zhang, W.; Kasun, L. C.; Wang, Q. J.; Zheng, Y.; Lin, Z. A Review of Machine Learning for Near-Infrared Spectroscopy. *Sensors* **2022**, 22 (24), 9764.
<https://doi.org/10.3390/s22249764>.
- (31) Tims, T. B.; Lim, D. V. Rapid Detection of *Bacillus Anthracis* Spores Directly from Powders with an Evanescent Wave Fiber-Optic Biosensor. *J. Microbiol. Methods* **2004**, 59 (1), 127–130. <https://doi.org/10.1016/j.mimet.2004.02.016>.
- (32) Bottone, E. J. *Bacillus Cereus*, a Volatile Human Pathogen. *Clin. Microbiol. Rev.* **2010**, 23 (2), 382–398. <https://doi.org/10.1128/cmr.00073-09>.
- (33) Tufts, J. A. M.; Calfee, M. W.; Lee, S. D.; Ryan, S. P. *Bacillus Thuringiensis* as a Surrogate for *Bacillus Anthracis* in Aerosol Research. *World J. Microbiol. Biotechnol.* **2014**, 30 (5), 1453–1461. <https://doi.org/10.1007/s11274-013-1576-x>.
- (34) Kolstø, A.-B.; Tourasse, N. J.; Økstad, O. A. What Sets *Bacillus Anthracis* Apart from Other *Bacillus* Species? *Annu. Rev. Microbiol.* **2009**, 63 (Volume 63, 2009), 451–476. <https://doi.org/10.1146/annurev.micro.091208.073255>.
- (35) Luo, L.; Zhao, C.; Wang, E.; Raza, A.; Yin, C. *Bacillus Amyloliquefaciens* as an Excellent Agent for Biofertilizer and Biocontrol in Agriculture: An Overview for Its Mechanisms. *Microbiol. Res.* **2022**, 259, 127016.
<https://doi.org/10.1016/j.micres.2022.127016>.
- (36) Deane, C. D.; Burkhart, B. J.; Saint-Vincent, P. M. B.; Tietz, J. I.; Lin, A.; Mitchell, D. A. *In Vitro* Biosynthesis and Substrate Tolerance of the Plantazolicin Family of Natural Products. *ACS Chem. Biol.* **2016**, 11 (8), 2232–2243.
<https://doi.org/10.1021/acscchembio.6b00369>.
- (37) Kovács, Á. T. *Bacillus Subtilis*. *Trends Microbiol.* **2019**, 27 (8), 724–725.
<https://doi.org/10.1016/j.tim.2019.03.008>.
- (38) Mido, T.; Schaffer, E. M.; Dorsey, R. W.; Sozhamannan, S.; Hofmann, E. R. Sensitive Detection of Live *Escherichia Coli* by Bacteriophage Amplification-Coupled Immunoassay on the Luminex® MAGPIX Instrument. *J. Microbiol. Methods* **2018**, 152, 143–147. <https://doi.org/10.1016/j.mimet.2018.07.022>.
- (39) Volke, D. C.; Calero, P.; Nickel, P. I. *Pseudomonas Putida*. *Trends Microbiol.* **2020**, 28 (6), 512–513. <https://doi.org/10.1016/j.tim.2020.02.015>.
- (40) Sava, I. G.; Heikens, E.; Huebner, J. Pathogenesis and Immunity in Enterococcal Infections. *Clin. Microbiol. Infect.* **2010**, 16 (6), 533–540. <https://doi.org/10.1111/j.1469-0691.2010.03213.x>.
- (41) Lin, S.; Sun, B.; Shi, X.; Xu, Y.; Gu, Y.; Gu, X.; Ma, X.; Wan, T.; Xu, J.; Su, J.; Lou, Y.; Zheng, M. Comparative Genomic and Pan-Genomic Characterization of *Staphylococcus Epidermidis* From Different Sources Unveils the Molecular Basis and Potential Biomarkers of Pathogenic Strains. *Front. Microbiol.* **2021**, 12.
<https://doi.org/10.3389/fmicb.2021.770191>.
- (42) CDC | *Bioterrorism Agents/Diseases (by category) | Emergency Preparedness & Response*. <https://emergency.cdc.gov/agent/agentlist-category.asp> (accessed 2024-11-15).
- (43) Aliste-Fernández, M.; Sanfeliu-Sala, I.; Sánchez-Delgado, J. Bacteremia caused by *Paracoccus yeei* in patient with compensated cirrhosis of the liver. *Enfermedades Infecc. Microbiol. Clínica* **2020**, 38 (9), 451–452.
<https://doi.org/10.1016/j.eimc.2020.01.013>.

- (44) Foca, G.; Ferrari, C.; Ulrici, A.; Sciutto, G.; Prati, S.; Morandi, S.; Brasca, M.; Lavermicocca, P.; Lanteri, S.; Oliveri, P. The Potential of Spectral and Hyperspectral-Imaging Techniques for Bacterial Detection in Food: A Case Study on Lactic Acid Bacteria. *Talanta* **2016**, *153*, 111–119. <https://doi.org/10.1016/j.talanta.2016.02.059>.
- (45) Morard, M.; Pérez-Través, L.; Perpiñá, C.; Lairón-Peris, M.; Collado, M. C.; Pérez-Torrado, R.; Querol, A. Comparative Genomics of Infective *Saccharomyces Cerevisiae* Strains Reveals Their Food Origin. *Sci. Rep.* **2023**, *13* (1), 10435. <https://doi.org/10.1038/s41598-023-36857-z>.
- (46) Mcquilken; Gemmell; Lahdenperä. *Gliocladium Catenulatum* as a Potential Biological Control Agent of Damping-off in Bedding Plants. *J. Phytopathol.* **2001**, *149* (3–4), 171–178. <https://doi.org/10.1046/j.1439-0434.2001.00602.x>.
- (47) Doan, U. V.; Mendez Rojas, B.; Kirby, R. Unintentional Ingestion of *Cordyceps* Fungus-Infected Cicada Nymphs Causing Ibotenic Acid Poisoning in Southern Vietnam. *Clin. Toxicol.* **2017**, *55* (8), 893–896. <https://doi.org/10.1080/15563650.2017.1319066>.
- (48) PubChem. *Elaphocordyceps subsessilis*. <https://pubchem.ncbi.nlm.nih.gov/compound/Elaphocordyceps-subsessilis> (accessed 2024-11-15).
- (49) Bushley, K. E.; Raja, R.; Jaiswal, P.; Cumbie, J. S.; Nonogaki, M.; Boyd, A. E.; Owensby, C. A.; Knaus, B. J.; Elser, J.; Miller, D.; Di, Y.; McPhail, K. L.; Spatafora, J. W. The Genome of *Tolypocladium Inflatum*: Evolution, Organization, and Expression of the Cyclosporin Biosynthetic Gene Cluster. *PLoS Genet.* **2013**, *9* (6), e1003496. <https://doi.org/10.1371/journal.pgen.1003496>.
- (50) Berman, J. *Candida Albicans*. *Curr. Biol.* **2012**, *22* (16), R620–R622. <https://doi.org/10.1016/j.cub.2012.05.043>.
- (51) Ueno, Y. Trichothecene Mycotoxins Mycology, Chemistry, and Toxicology. In *Advances in Nutritional Research*; Draper, H. H., Ed.; Springer US: Boston, MA, 1980; pp 301–353. https://doi.org/10.1007/978-1-4757-4448-4_10.
- (52) Sun, Z.-B.; Li, S.-D.; Ren, Q.; Xu, J.-L.; Lu, X.; Sun, M.-H. Biology and Applications of *Clonostachys Rosea*. *J. Appl. Microbiol.* **2020**, *129* (3), 486–495. <https://doi.org/10.1111/jam.14625>.
- (53) Sugui, J. A.; Kwon-Chung, K. J.; Juvvadi, P. R.; Latgé, J.-P.; Steinbach, W. J. *Aspergillus Fumigatus* and Related Species. *Cold Spring Harb. Perspect. Med.* **2015**, *5* (2), a019786. <https://doi.org/10.1101/cshperspect.a019786>.
- (54) Usuda, D.; Higashikawa, T.; Hotchi, Y.; Usami, K.; Shimosawa, S.; Tokunaga, S.; Osugi, I.; Katou, R.; Ito, S.; Yoshizawa, T.; Asako, S.; Mishima, K.; Kondo, A.; Mizuno, K.; Takami, H.; Komatsu, T.; Oba, J.; Nomura, T.; Sugita, M. *Exophiala Dermatitidis*. *World J. Clin. Cases* **2021**, *9* (27), 7963–7972. <https://doi.org/10.12998/wjcc.v9.i27.7963>.
- (55) Lund, F.; Nielsen, A. B.; Skouboe, P. Distribution of *Penicillium Commune* Isolates in Cheese Dairies Mapped Using Secondary Metabolite Profiles, Morphotypes, RAPD and AFLP Fingerprinting. *Food Microbiol.* **2003**, *20* (6), 725–734. [https://doi.org/10.1016/S0740-0020\(02\)00160-0](https://doi.org/10.1016/S0740-0020(02)00160-0).
- (56) Rundberget, T.; Skaar, I.; Flåøyen, A. The Presence of *Penicillium* and *Penicillium* Mycotoxins in Food Wastes. *Int. J. Food Microbiol.* **2004**, *90* (2), 181–188. [https://doi.org/10.1016/S0168-1605\(03\)00291-5](https://doi.org/10.1016/S0168-1605(03)00291-5).
- (57) Thomma, B. P. H. J. *Alternaria* Spp.: From General Saprophyte to Specific Parasite. *Mol. Plant Pathol.* **2003**, *4* (4), 225–236. <https://doi.org/10.1046/j.1364-3703.2003.00173.x>.

- (58) Kimes, P. K.; Liu, Y.; Hayes, D. N.; Marron, J. S. Statistical Significance for Hierarchical Clustering. *Biometrics* **2017**, *73* (3), 811–821. <https://doi.org/10.1111/biom.12647>.
- (59) Hastie, T.; Tibshirani, R.; Friedman, J. Overview of Supervised Learning. In *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*; Hastie, T., Tibshirani, R., Friedman, J., Eds.; Springer: New York, NY, 2009; pp 9–41. https://doi.org/10.1007/978-0-387-84858-7_2.
- (60) Schloss, P. D. Identifying and Overcoming Threats to Reproducibility, Replicability, Robustness, and Generalizability in Microbiome Research. *mBio* **2018**, *9* (3), 10.1128/mbio.00525-18. <https://doi.org/10.1128/mbio.00525-18>.
- (61) Cui, S.; Pan, Z.; Wu, S.; Mao, B.; Tang, X.; Zhang, Q.; Zhang, H.; Zhao, J. Improvement of the Lyophilization Survival Rate of *Lactobacillus Casei* via Regulation of Its Surface Substances. *Foods* **2022**, *11* (21), 3468. <https://doi.org/10.3390/foods11213468>.
- (62) Sørensen, K. M.; van den Berg, F.; Engelsen, S. B. NIR Data Exploration and Regression by Chemometrics—A Primer. In *Near-Infrared Spectroscopy: Theory, Spectral Analysis, Instrumentation, and Applications*; Ozaki, Y., Huck, C., Tsuchikawa, S., Engelsen, S. B., Eds.; Springer: Singapore, 2021; pp 127–189. https://doi.org/10.1007/978-981-15-8648-4_7.
- (63) Ying, X. An Overview of Overfitting and Its Solutions. *J. Phys. Conf. Ser.* **2019**, *1168* (2), 022022. <https://doi.org/10.1088/1742-6596/1168/2/022022>.