Microbial Identification Framework for Risk Assessment

May 2017

Canada

Cat. No.: En14-317/2018E-PDF

ISBN 978-0-660-24940-7

Aussi disponible en français

Summary

The *New Substances Notification Regulations (Organisms)* (the regulations) of the *Canadian Environmental Protection Act*, 1999 (CEPA) are organized according to organism type (micro-organisms and organisms other than micro-organisms) and by activity.

The Microbial Identification Framework for Risk Assessment (MIFRA) provides guidance on the required information for identifying micro-organisms. This document is intended for those who deal with the technical aspects of information elements or information requirements of the regulations that pertain to identification of a notified micro-organism. It is intended to help notifiers with the choice of methodology and the analysis of scientific data required for adequate microbial identification. The document is technical in nature and relies heavily on concepts of microbiology, microbial taxonomy and methods of microbial identification.

Table of contents

Introduction

The *New Substances Notification Regulations (Organisms)* (the regulations) of the *Canadian Environmental Protection Act*, 1999 (CEPA) are organized according to organism type (micro-organisms and organisms other than micro-organisms) and activity. A common information element of the micro-organism schedules is information substantiating the identification of the notified micro-organism[1]. The Microbial Identification Framework for Risk Assessment (MIFRA) provides guidance on this information element. The MIFRA aims to set a common frame of reference to be used in microbial identification for the purposes of risk assessment related to a New Substance Notification (NSN). This frame of reference should help in the choice of methodology and the analysis of data required for adequate microbial identification. It describes specific methodology used in microbial identification and incorporates 13 case studies demonstrating the application of this framework, as well as a suggested template for their presentation. Methods for identifying and characterizing micro-organisms are constantly evolving and new methods with improved specificity and reliability are continuously being developed; therefore, the New Substances Program (NSP) will update this document periodically and as required.

Genetic modifications made to a notified micro-organism are not within the scope of this document as these are addressed under specific information elements in the regulations. The MIFRA should be used in conjunction with the *Guidance for the Notification and Testing of New Substances: Organisms* which contains guidance on completing all other information elements of the regulations. In addition, the scope of this document does not extend to synthetic micro-organisms that have been constructed with significant genomic content from potentially unrelated species (for example, the experimental micro-organism *Mycoplasma laboratorium*). Such artificial micro-organisms may not have relevant taxonomic relationships with natural species, and are to be considered a special case in the context of identification for a use that falls under the scope of CEPA. A pre-notification consultation (PNC) should be requested with the NSP to discuss the notification of a synthetic micro-organism.

[1] Information in respect of the micro-organism: its identification and the information substantiating its identification (paragraph 1(a) of Schedules 1, 2, 3, 4 of the **regulations (**http://laws-lois.justice.gc.ca/eng/regulations/SOR-2005-248/index.html)

Pre-Notification Consultation (PNC)

A PNC is an option for notifiers who wish to consult with the NSP during the planning or preparation of their NSN package to discuss any questions or concerns they have about the required prescribed information.

To request a PNC, please contact the NSP:

Substances Management Information Line
Telephone: 1-800-567-1999 (Toll Free in Canada) or 1-819-938-3232 (Outside of Canada)
Facsimile: 1-819-938-5212
E-mail: eccc.substances.eccc@canada.ca

This document has been prepared to assist notifiers responsible for complying with the *New Substances Notification Regulations (Organisms)* of CEPA. Any obligations under these regulations are independent of any obligations that notifiers may have under other Regulations or Acts.

Definitions of technical terms used in this document are provided in Appendix A.

1. Who should use the MIFRA?

Anyone who deals with the technical aspects of information elements or information requirements of the regulations that pertain to identification of a notified micro-organism: notifiers (sometimes also referred to as 'proponents' or 'applicants'), including their business partners, Canadian agents, foreign suppliers, technical contacts, third-party commercial laboratories, etc.

This document is technical in nature as it provides detailed guidance on a technical information element; it relies heavily on concepts of microbiology, microbial taxonomy and methods of microbial identification.

2. The importance of microbial identification for the purpose of risk assessment

A valid and well-supported microbial identification is the cornerstone of the risk assessment of micro-organisms notified under the regulations.

The accurate identification of a micro-organism allows known characteristics of a taxonomic group to be used in the risk assessment. Information on this taxonomic group can also help to identify "closely-related" organisms (at species or strain levels), that can be used in providing complementary information to determine the hazard of the notified micro-organism and discriminate it from potential pathogens of clinical or environmental significance. Also, information on closely-related organisms could be used to provide surrogate information for other information elements of the regulations.

Inaccurate identification can lead to an inaccurate determination of the micro-organism's hazard level, potentially leading to a risk assessment conclusion that will not be applicable to the actual micro-organism being notified. This may result either in potential negative impacts to human health and the environment, or to unnecessary risk management action for low hazard micro-organisms.

3. Acceptable taxonomic level of identification

The objective of microbial identification is to use phenotypic and genotypic characteristics of the notified micro-organism to allow its placement within a recognized taxonomic group. The notifier must establish the identity of the notified micro-organism.

The level of taxonomic designation will vary depending on the micro-organism, but in general, a designation to the species level is expected. In the context of a NSN, the acceptable level of identification is not required to be the lowest taxonomic level known for the micro-organism (for example, subspecies, strain or serovar). However, designation to the subspecies, strain or serovar may be appropriate where the notified organism is closely-related to clinical or environmentally pathogenic micro-organisms. For example, *Escherichia coli* as a species includes a number of pathogenic and non-pathogenic serotypes. In this case, if an *E. coli* species is notified and is claimed to be a non-pathogenic strain, then the level of identification required to differentiate the notified *E. coli* will be to a strain or serotype level. The data provided to substantiate the identification should allow the differentiation of the notified micro-organism from pathogenic *E. coli* serotypes. If this is not done, there will be uncertainties in the risk assessment because of the possibility that the notified *E. coli* may be pathogenic.

If a micro-organism can only be identified to a genus or sub-generic level (for example, a clade with multiple species) due to the lack of available information or suitable methods to identify it to species level, the members of that taxonomic group will be considered as potential candidates for the identity of the notified micro-organism. This will result in uncertainty related to the identity of the micro-organism, which the notifier would have to address through literature searches and appropriate choice of surrogate micro-organisms.

The NSP acknowledges that microbial taxonomy is in constant flux. Taxonomic designation of a notified micro-organism should follow the most current international codes of nomenclature and standard taxonomic sources. International codes of nomenclature are those that are officially recognized and accepted by international committees.

International committees

- For bacteria: International Committee on Systematics of Prokaryotes (ICSP) http://www.the-icsp.org/)
- For algae and fungi: International Association for Plant Taxonomy (IAPT) http://www.iapt-taxon.org/nomen/main.php?page=pf
- For viruses: International Committee on Taxonomy of Viruses (ICTV) http://www.ictvonline.org/codeOfVirusClassification.asp

Acceptable standard resources used for microbial identification and taxonomic classification are those that are deemed acceptable by recognized authorities or by the scientific community. These include:

(i) reference texts such as Bergey's Manual of Systematics of Archaea and Bacteria (2015) or The Yeasts, A Taxonomic Study (Kurtzman *et al,* 2011)
(ii) peer-reviewed articles in scientific journals; and
(iii) online resources.

In general, several sources of corroborating information are preferred, when possible, to a single source.

Examples of online resources

- NCBI (National Center for Biotechnology Information) Taxonomy (http://www.ncbi.nlm.nih.gov/taxonomy)
- *List of Prokaryotic names with Standing in Nomenclature* (LPSN) (http://www.bacterio.net/)
- CBS-KNAW Fungal Biodiversity Centre (http://www.cbs.knaw.nl/)
- Mycobank (http://www.mycobank.org/)

4. General approach to microbial identification

Microbial species are part of a continuum of diversity and species delineation could be ambiguous (Konstantinidis *et al.,* 2006; Liti *et al.,* 2006). Despite this ambiguity, micro-organisms have been classified for practical purposes into different species based on the coherence of their genetic and biological characteristics.

"Despite the lack of information on speciation mechanisms and genomic sequence similarities, members of most prokaryotic species do form a genomically coherent entity." *Stackebrandt* (*2011*).

While the issues with the concept of species are acknowledged by the NSP, it is nevertheless important to consider the species as a defined taxonomic entity and a practical component of a risk assessment. Some of the factors that influence microbial identification and taxonomic classification include the concept of speciation, the pluralistic nature of microbial taxa, topological incongruence, horizontal gene transfers, and presence of multiple copies of conserved gene regions. As such, a polyphasic approach for accurate identification and taxonomic designation of micro-organisms is recommended by experts (Stackebrandt, 2011, 2014; CBS-KNAW Fungal Biodiversity Centre, Utrecht, Netherlands).

Various methods and tools are available to identify a micro-organism to the species level and to resolve its phylogenetic and taxonomic relationships within genera/clade/species, usually to a taxonomic level adequate to conduct a robust risk assessment under the regulations.

Polyphasic tiered approach for identification of a micro-organism

A polyphasic tiered approach should be applied for accurate identification of micro-organisms. The **polyphasic** principle uses multiple lines of evidence (data) from a combination of different methodologies, while the **tiered** principle allows for sequential selection of these methodologies, in order to identify a micro-organism efficiently.

A polyphasic approach enables the substantiation of the identification and reduces the uncertainty related to the results generated by individual methods. Certain characteristics or traits of the micro-organism are useful in its identification. Depending on the micro-organism, the combination of complementary methods used to examine these characteristics will vary. An effective polyphasic approach will combine phenotypic and genotypic characteristics such as those outlined in Tables 1 and 2.

Table 1: list of phenotypic characteristics of a micro-organism that can be used in a polyphasic approach to identification

Characteristics	Examples (not exhaustive list)	Methods (not exhaustive list)
a) Morphological	➢ Colony morphology (colour, shape, presence of halo, etc.), fruiting bodies, mycelia and hyphal structures ➢ Cell and spore morphology (shape, cluster type, type of cell wall, etc.), cell staining, motility (pili, fimbria, flagella and their quantity), presence or absence of envelope structures such as a capsule or slime layer, etc.	Growing the micro-organism on specific media with or without supplements, plating, staining, and microscopic observation.
b) Physiological and Biochemical	➢ Growth temperature (optimum, maximum, minimum and range) ➢ pH optimum and range ➢ Requirement for nutrients and growth supplements ➢ Enzymatic activities ➢ Carbohydrate utilization ➢ Acid production from carbohydrates ➢ Utilization of sources of carbon, nitrogen, etc.	Growing the micro-organism, under different conditions and in specific media.

Characteristics	Examples (not exhaustive list)	Methods (not exhaustive list)
	➢ Oxygen requirement ➢ Salt tolerance ➢ Growth on selective, differential or enriched media ➢ Susceptibility/resistance to antibiotics, antifungal or antiviral agents ➢ Susceptibility/resistance to heavy metals or other substances. ➢ Pigment production	
c) Serological	➢ Agglutination ➢ Immunodiffusion ➢ ELISA ➢ Detection of specific proteins (western blotting)	Testing the microbial antigens against specific antibodies.
d) Toxin/metabolite production	➢ Endotoxins, exotoxins, etc. ➢ Mycotoxins ➢ Other metabolites	Molecular methods, HPLC, ELISA.
e) Chemotaxonomic	➢ Fatty acid methyl ester (FAME) ➢ Lipopolysaccharides, peptidoglycan type, whole cell sugars, cell wall sugars, mycolic acids, diaminoacids, quinone system, polyamine content, cell wall amino acids, etc. ➢ Cellular proteome, metabolome, etc.	Analysis of the extract from the micro-organism (Fatty acid, protein, etc.)

Table 2: list of genotypic characteristics of a micro-organism that can be used in a polyphasic approach to identification

Characteristics	Examples (not exhaustive list)	Methods (not exhaustive list)
a) Conserved genes or hypervariable regions in the conserved gene sequence	➢ 16S rRNA, chaperonin-60 (cpn60) for bacteria; 16S rRNA, type II chaperonin for archaea ➢ 18S, 5.8S, 28S rRNA operon (along with ITS1, ITS2, D1/D2 regions) for eukaryotes ➢ Multi-Locus Sequence Alignment (MLSA) or Typing (MLST) using house-keeping genes (for example, gyrase A, gyrase B, translation initiation factor 1, translation initiation factor 2, transcription elongation factor 1, recombinase A, recombinase B, cytochrome C oxidase, β-subunit of ATP-synthase, etc.)	DNA sequencing and DNA alignment
b) Whole Genome Sequence	➢ Full genome analysis	DNA sequencing, DNA alignment and genome annotation (for example, useful in the detection of specific genes that may contribute to the identification, such as virulence factors)
c) DNA polymorphism	➢ DNA base ratio (G+C content) ➢ Random Amplification of Polymorphic DNA (RAPD) ➢ Restriction Fragment Length Polymorphism (RFLP) ➢ Pulsed-field gel electrophoresis (PFGE) ➢ Southern and northern blotting ➢ Cellular transcriptome	PCR and DNA-based Typing and Hybridization
d) DNA Hybridization	➢ DNA:DNA hybridization or DNA:RNA hybridization	Hybridization

A tiered approach will be helpful in the selection of appropriate methods listed in Table 1 and Table 2. Not all methods described above will be needed for the identification of a micro-organism. Ideally, the selection of methodologies for the substantiation of the identity of the

micro-organism for the purposes of the NSN should start with the ones that are most discriminatory. The tiered approach should help to organize the multiple lines of evidence (data) collected during the identification and characterization of the micro-organism.

A polyphasic tiered approach should start with methods that allow a putative placement of the notified micro-organism in a taxonomic group. These methods frequently involve sequence analysis of a conserved gene region as the first step (other methods specific for a group of micro-organisms can also be used).

This placement within a taxonomic group, based on results from the first tier, will inform the choice of methods to be used in the subsequent steps of the tiered approach, until an acceptable taxonomic level of identification, for the purpose of the risk assessment, is reached. Based on the type of the micro-organism (bacteria, archaea, fungi, yeast, virus), the tiered approach could vary. For certain organisms, the components of the polyphasic tiered approach and the methods to be used are well-established in the literature (for examples see Section 5 - case studies).

A schematic for a polyphasic tiered approach for microbial identification of a notified micro-organism is provided below (Figure 1). Decisions on the number of tiers, amount of details in each tier and type of additional information and/or tests required, will depend on:

- the source of the micro-organism (for example, strain from a recognized culture collection, clinical or environmental isolate, etc.);
- the relatedness or similarity of the notified micro-organism to other micro-organisms suspected or demonstrated to pose hazards to humans, plants, or animals; and
- how well the characteristics used for identification within the taxonomic group are described in the literature.

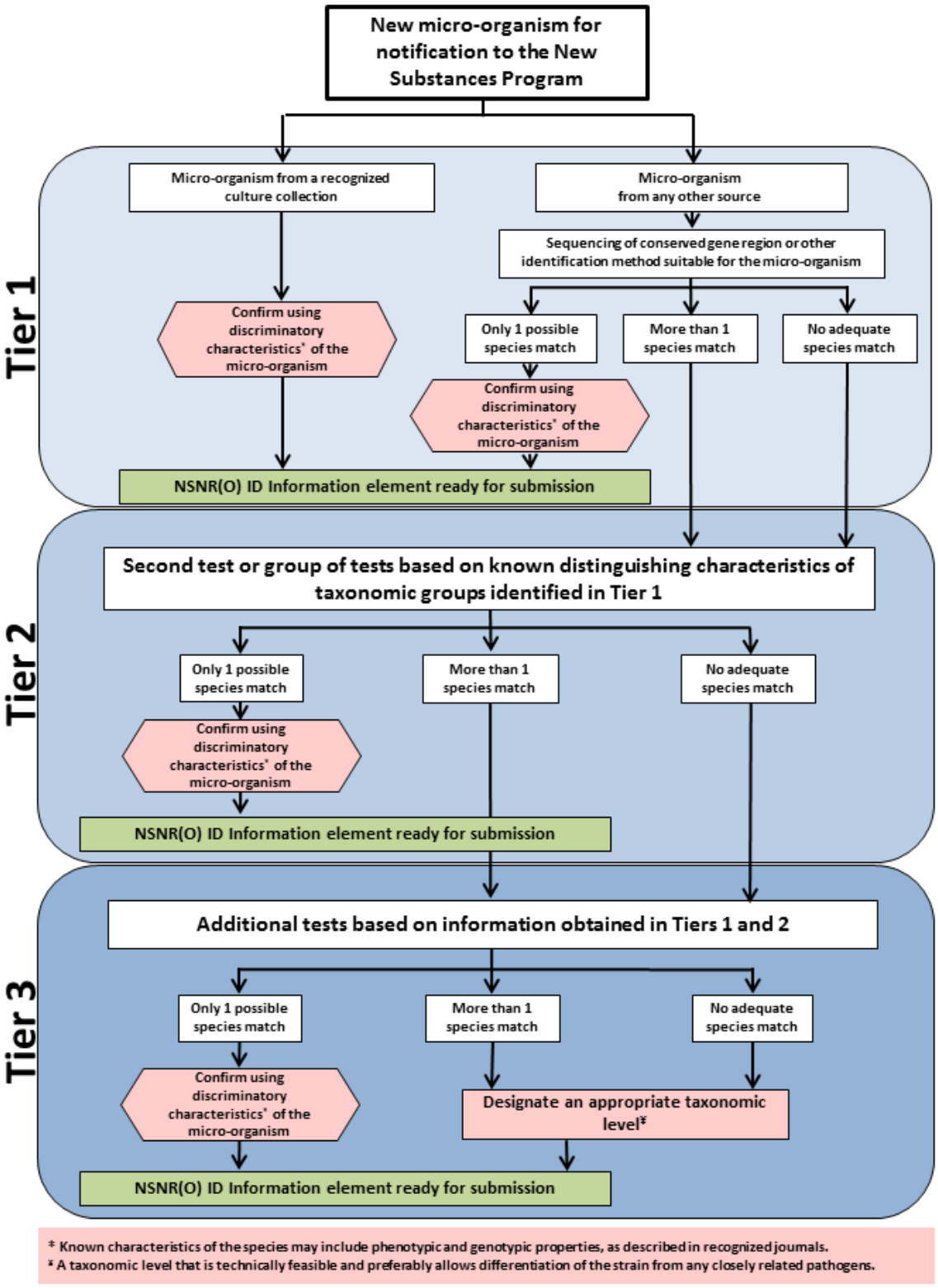

Figure 1. A polyphasic tiered approach for identification of the notified micro-organism and substantiation of its identity.

As depicted in the flowchart (Figure 1), if a strain was obtained from a recognized culture collection, a certificate of analysis from the culture collection, along with data confirming the characteristics of the strain, will be required to adequately substantiate its identity.

For micro-organisms isolated from the environment, the identification should ideally start with sequencing of conserved gene regions, like ribosomal RNA genes, or any other established identification methods specific for a group of micro-organisms. Analysis of these results in Tier 1 could lead to three different possible outcomes:

(1) Notified micro-organism matching to a single species: In this case, confirmation of the identity with other known characteristics of the species will be required.

(2) Notified micro-organism matching to more than one possible species: In this case, additional tests (Tier 2) to refine the taxonomic placement of the micro-organism should be performed based on the data collected and on information from the literature on taxonomic group identified in Tier 1.

(3) No adequate match between the notified micro-organism and any known species: Where the data collected is not conclusive to any species or genus, additional tests (Tier 2) to designate the taxonomic placement of the notified micro-organism should be performed based on the data collected and on information from the literature on taxonomic groups identified in Tier 1.

Analysis of the results in Tier 2 could lead to similar outcomes as from Tier 1, which can either designate the notified micro-organism conclusively to a species or identify additional tests required for Tier 3. Analysis of results in Tier 3, would mostly lead to species-level designation. If species ambiguity still exists at this stage, the notifier can designate the notified micro-organism to a higher taxonomic level (for example, clade level) resulting from Tier 3 analysis.

Note: In case of uncertainties related to the choice of methods, the number of tiers and the conclusion on the identification of the notified micro-organism, notifiers can request a PNC.

Recognized methods for identification of a micro-organism

Many methods are commonly used to identify and characterize micro-organisms, each with their own strengths and limitations. A list of methods, along with suggested best practices, is provided in Appendix B.

Whether the identification of the micro-organism is performed by the notifier or a third-party laboratory, the principles of a polyphasic tiered approach and best practices outlined in Appendix B must be followed.

In the NSN package, a list of tests used to arrive at the taxonomic designation of the notified micro-organism must be provided, along with the results and any other information used to reach the taxonomic designation. A template to organize and analyze the data in accordance with the polyphasic tiered approach is provided in Appendix C, for suggested use by notifiers.

5. Case studies for identification of different micro-organisms

Several case studies, highlighting the polyphasic tiered approach used for the identification of different micro-organisms among those most often notified under the regulations (bacteria, viruses, fungi and algae), are provided in Appendix D. The methods and analyses used in the case studies are based on scientific literature supporting the identification of those micro-organisms.

- The case studies provided are to illustrate the appropriate taxonomic levels sufficient for the purpose of risk assessment for different micro-organisms.
- The cases studies are intended to help notifiers in choosing and structuring the data generated to substantiate the identification of the notified micro-organism.

Table 3: list of case studies highlighting the type of micro-organism and methodologies used for identification

Case Study	Highlights	Page
Arcobacter sp.	✓ Identification of a **bacterium** that could not be assigned to a specific species. ✓ Use of 16S rRNA and gyrB gene analysis, as well as phenotypic characteristics.	39
Aspergillus niger	✓ Identification of a **fungus** using morphological characteristics, extrolite production and genetic analysis (D2 region, ITS region and the calmodulin gene).	44
Aurantiochytrium limacinum	✓ Identification of a **microalga** using genetic analysis of the 18S rRNA gene, morphological characteristics, FAME analysis and pigment profiling.	47
Bacillus amyloliquefaciens	✓ Identification of a **bacterium** using phenotypic properties and polymorphism in a variable region of the 16 rRNA gene.	50
Candida tropicalis	✓ Identification of a **yeast** obtained from a culture collection. ✓ Use of genetic analysis of the D1D2 region of the 28S rRNA gene and morphological characteristics.	55
Deinococcus proteolyticus	✓ Identification of a **bacterium** using morphological characteristics and genetic analysis of 16S rRNA and cpn60.	57
Influenza virus	✓ Identification of a **virus** using whole genome sequencing and immunological data.	60
Komagataella phaffii	✓ Identification of a **yeast** using morphological characteristics and genetic analysis of the D1D2 LSU rRNA gene region.	62
Listeria monocytogenes	✓ Identification of a **bacterium** using genome sequencing of the parental strain, morphological characteristics, genetic analysis and protein expression.	64
Rhodococcus aetherivorans	✓ Identification a **bacterium** using genetic analysis of 16S rRNA, FAME analysis, enzyme activity and morphological characteristics.	69

Case Study	Highlights	Page
Saccharomyces cerevisiae	✓ Identification of a **yeast** using genetic analysis of the D2 LSU region, as well as morphological and phenotypic characteristics.	76
Shewanella indica	✓ Identification of a **bacterium** that can be assigned to a particular species but is closely-related to a strain from another species. ✓ Use of genetic analysis of 16S rRNA and gyrB, as well as phenotypic properties.	80
Trichoderma reesei	✓ Identification of a **fungus** using morphological and physiological characteristics, as well as genetic analysis of ITS1 and ITS2 regions of the rRNA operon and of tef1.	84

6. Updates to the MIFRA

Methods for identifying and characterizing micro-organisms are constantly evolving and new methods with improved specificity and reliability are continuously being developed. Although many protocols currently exist for identifying microbial strains of medical, agricultural and environmental significance, it is also recognized that no internationally accepted standards currently exist for determining what constitutes an accurate identification of a micro-organism. Therefore, the NSP intends to update this document periodically and as required. The NSP will continue to recognize that notifiers may use identification methods and tools that may not be mentioned in this document; however, it remains the responsibility of the notifier to demonstrate the validity of these methods, consistent with the polyphasic tiered approach described in this document.

Appendices

Appendix A – Glossary

Terms

Clade – A group of organisms that cluster together in a phylogenetic analysis which have a common ancestor. For the purpose of the MIFRA, clade designation will be accepted at the genus or species level.

Genotypic methods – Techniques that analyze the genetic make-up of an organism and determine its relationship with others.

Morphology – The study of the form or shape of an organism or part thereof, either directly observed by the naked eye and/or using specific procedures or instruments (e.g., microscope, staining procedure, growth on selective media, etc.) (OECD, 2003).

Phenotypic methods – Techniques that directly or indirectly detect, measure or characterize, features of an organism resulting from the observable expression of its genetic constitution. Phenotypic characteristics include morphological, physiological and biochemical features (OECD, 2003).

Phylogenetic analysis – The study of evolutionary relationships among the organisms within a taxonomic group and with the members of other taxa, normally performed using sequence data.

Recognized culture collection – A recognized culture collection is one that is publicly accessible and adheres to the World Federation for Culture Collections' "Guidelines for Establishment and Operation of Cultures of Microorganisms" (3rd Edition, February 2010). Examples of recognized culture collections include, American Type culture Collections (ATCC), Leibniz-Institut DSZM-German culture Collection of Microorganisms and Cell Cultures, and the CBS Fungal Biodiversity Centre (CBS-KNAW), etc.

Speciation – The formation of new and distinct species in the course of evolution.

Serotype – A distinct group of strains within a species (or a sub-species) of bacteria or virus, classified based on the cell surface antigens (for example, O, H or K antigens) of the micro-organism.

Taxonomic group – Taxonomic group refers to a rank (such as species, genus, family, order, etc.) in the taxonomic hierarchy, where organisms within a rank are grouped based on phenotypic or genotypic similarities.

Taxonomy – Taxonomy is the science of identifying, classifying and assigning a name to an organism, in order to determine the relationship of the organism with others, as well as the genetic variations within and among the different taxonomic groups and the evolutionary aspects of a taxonomic group.

Topological incongruence – This may be observed when a species contains genes, from a horizontal gene transfer event from an unrelated species, that may have evolved differently than other genes present in its genome.

Appendix B – Recognized methods for identification of micro-organisms

Phenotypic methods

Table B-1.1: morphological properties

Methods	Colony morphology includes: - Shape: circular, filamentous (fibrous), rhizoid (thick fibers), irregular. - Elevation: raised, flat, convex (rounded), umbonate (peaked), crateriform (indented). - Colony margin (edge): entire (smooth), undulating (wavy), lobular (finger-like projections), filiform (fibrous projections), curled (swirled) - Surface refraction: smooth, dull, glistening (mucoid), rough (ground glass), rugose (wrinkled) - Opacity and color: transparent, translucent, opaque, fluorescent, iridescent, pigmentation. Cell morphology includes: Bacteria - Size, shape (bacilli, cocci, spiral, etc.), Gram staining (Gram negative or positive), and sporulation including endospores. Surface features: Staining of flagella (monotrichous, amphitrichous, lophotrichous or peritrichous - the presence of flagella also indicates motility), and capsule or slime layer. Fungi and microalgae - shape and size of vegetative cells (for example, mycelia, hyphae) and fruiting bodies (conidiospores, teliospores, sporangia, etc.). Viruses - envelope (i.e. enveloped or naked) and capsid shape (i.e. polyhedral, spherical, filamentous, etc.).
Level of identification	Useful in the preliminary identification of a micro-organism, or as confirmatory data. For certain organisms like fungi, colony and cell morphology could be an important discriminating factor.
Strengths	Bacteria and fungi are directly observable by the naked eye or under a light microscope. For viruses or for higher magnification and resolution of bacterial or fungal structures, electron microscope (EM)/Scanning electron microscope (SEM) is required. Simple, quick and inexpensive.

Table B-1.1: morphological properties	
Limitations	Suitable only for cultivable organisms. Highly dependent on culture conditions. Low value for identification at species-level when variations among strains are observed. Similar morphological properties can be observed between unrelated species. May require expensive equipment like EM/SEM and technical knowledge. May require special staining reactions. Several bacterial species stain poorly with the Gram stain (for example, Gram staining does not work for *Mycoplasma* spp., *Mycobacteria* spp., *Chlamydia* spp., *Rickettsia* spp., *Campylobacter* spp., *Fusobacterium* spp., and *Brucella* spp., etc.). Alternate staining or counterstaining may be required.
Best practices	Ideally only the documented culture conditions specific for the test and species-of-interest should be used. The colony morphology of the notified micro-organism should be described and compared with the type strain of the species.
Tips for notification	All relevant information (protocol, results, analysis and interpretation) of the tests must be submitted, along with data and culture conditions for the notified micro-organism and all comparators used.
References	Busse et al., 1996 ; Jackman, 2012; Janda and Abbott, 2002; OECD, 2003; UK Standards for Microbiology Investigation 2016; Tshikhudo et al., 2013; Gram stain, 2010; and McClelland, 2001

Table B-1.2: physiological and biochemical properties	
Methods	These include: - Growth temperature (optimum, maximum, minimum, range) - pH optimum and range of growth - Oxygen requirement (aerobic, anaerobic, facultative) - Water activity requirement - Salt tolerance - Growth on selective, differential or enriched media - Susceptibility/resistance to antibiotics, antifungal or antiviral agents - Susceptibility/resistance to heavy metals or other substances (for example, bile, chlorine, etc.). - Nutritional requirements and growth factors - Carbon and nitrogen source(s) utilization - Enzymatic activities - Carbohydrate utilization - Acid production from carbohydrates - Pigment production
Level of identification	Its value to microbial identification should be considered to be auxiliary and supplementary, unless justified with detailed data with sufficient discriminatory power. Often useful as the confirmatory data for certain micro-organisms. Allows species level identification for certain micro-organisms, when complemented with other identification methods.
Strengths	Cost-effective, easy to perform the tests and interpret the results. Certain physiological and biochemical properties can discriminate different species within a genus, for certain micro-organisms. Standardized commercial kits are available for many of these properties. Automated commercial systems for specific microbial groups are available to ensure a more reliable identification.
Limitations	Highly dependent on culture conditions. Suitable only for cultivable micro-organisms. Low value for identification at species-level when variations among strains are observed. A mutation will generally impact a physiological or biochemical property. Databases often biased towards species of clinical importance. The reliability

Table B-1.2: physiological and biochemical properties	
	depends on the richness and quality of the database used. Species or taxonomic groups not included in the database will not be identified or could be misidentified.
Best practices	Physiological and biochemical properties of the notified micro-organism should be described and compared with the type strain of the species. When commercial test kits are used, selection of the kit should be relevant to the group of micro-organisms that the notified micro-organism is suspected to belong to. The notified strain and the reference strains must be tested at the same time.
Tips for notification	All relevant information (protocol, results, analysis and interpretation) of the tests must be submitted, along with data and culture conditions for the notified micro-organism and all comparators used. Information about the reference database and software, used to support the identification of the micro-organism and similarity indices, along with alternate species possibilities must be provided. The specificity and richness of the database should be clearly conveyed, as well as its limitations. Provides important information specific to the notified micro-organism that may be used in other information elements in the notification.
References	Anders et al., 2007; Busse et al., 1996; Gillis et al., 2005; Colwell and Grigorova, 1987; Janda and Abbott, 2002; and UK Standards for Microbiology Investigation 2016

Table B-1.3: serological properties (immunological-based assays)	
Methods	Allow detection of phenotypic markers (surface antigens) with antibodies or antisera that are specific for cellular components such as cell walls, capsules, flagella, receptors, etc. Allow determination of antigenic homologies among related micro-organisms.
Level of identification	Their value to microbial identification should be considered to be auxiliary or supplementary, unless justified with detailed data with sufficient discriminatory power. Identification of serotypes or serovars is possible for some genera and species. Detection of specific proteins and immunological activity of homologous proteins can also contribute to strain-level identification.
Strengths	Reliable if the antigens selected for serotyping are suitable for the group of micro-organisms. Antibody-based assays are in general sensitive and highly specific. Commercial kits and standardized procedures are available.
Limitations	Requires the serotypes within the taxonomic group to be already established. Serological techniques require expression of an antigenic molecule. Availability and cost of the commercial antibodies. The production process of the required antibodies can be difficult when they are not commercially available. Methods may require optimization. Sensitivity and specificity of immunoassays can be affected by cross-reactive antibodies. Technical expertise is required. For species that cannot be cultured, additional processes such as the production of semi-synthetic antigens may be required.
Best practices	Serological profile of the notified strain should be described and compared with known serotypes described for the taxonomic group. When commercial test kits are used, the notified strain and the reference strains must be tested at the same time as positive and negative controls.
Tips for notification	All relevant information (protocol, results, analysis and interpretation) of the tests must be submitted, along with data and culture conditions for the notified micro-organism and all comparators used.
References	Andreotti et al., 2003; Gasanov et al., 2005; Jones and Krieg, 1984; Oskam et al., 2000; and Stanier et al., 1970

Table B-1.4: toxin/ metabolite production	
Methods	Absence or presence of toxins or other metabolites can be used in identification of certain micro-organisms. Some methods can target the toxins/metabolites (analysis by chromatography or ELISA), or the genetic determinants of those products.
Level of identification	Can identify to genus, species or strain level depending on the micro-organisms when complemented with other methods of identification. For example, the detection of specific toxins contributes to the assignment of an *Escherichia coli* strain to categories of clinical concern.
Strengths	Some toxins are specific to certain genera and species (for example, aflatoxin production by *Aspergillus* spp. may be useful in their identification). Standardized methods and/or commercial kits are available for certain toxins or secondary metabolites.
Limitations	Highly dependent on culture conditions. Low value for identification at species-level when variations among strains are observed. Growth media, conditions and analytical procedures must be standardized for testing. Some commercial kits yield only preliminary results; additional confirmatory testing is therefore required to confirm the production of toxins.
Best practices	Selection of toxins or secondary metabolites to be tested should be based on the profile of toxins or secondary metabolites documented for the relevant taxonomic group of the notified micro-organism.
Tips for notification	All relevant information (protocol, results, analysis and interpretation) of the tests must be submitted, along with data and culture conditions for the notified micro-organism and all comparators/reference standards used. Those methods provide important information specific to the notified micro-organism that may be used in other information elements in the notification.
References	Arenas-Hernandez et al., 2012; Cecchini et al., 2016; Nielson et al., 2011; Samson et al., 2014; and Touzet et al., 2007

Table B-1.5: chemotaxonomic properties - analysis of fatty acids (FAs)	
Method	The analysis of fatty acids (FA) allows the identification and the classification of micro-organisms based on the composition of FA of the phospholipid bilayer of cell membranes. Depending on the type of micro-organism, different types of FA profiling are possible, including FAME analysis, phospholipids fatty acids (PLFA) and polyunsaturated fatty acids (PUFA), PLEL (phospholipid ether lipids) and LPS-HYFA (hydroxy fatty acids of lipopolysaccharides), etc.
Level of identification	FA profiles are generally useful to identify bacterial or fungal genera; in certain cases, identification can be done at species or sub-species levels. FA profiles are frequently used for microalgae species identification. Its value to microbial identification should be considered to be auxiliary or supplementary, unless justified with detailed data with sufficient discriminatory power.
Strengths	The FA composition appears to be highly conserved, and is therefore reliable. A commercial identification system (MIDI Sherlock® system) is available for FA analysis. It provides access to an established database which could be customized for species of interest.
Limitations	Highly dependent on culture conditions. Suitable only for cultivable organisms. Requires gas chromatography and technical expertise. FA misidentification is common, and new unknown FAs may need to be identified by mass spectrometry. The reliability depends on the richness and quality of the database used. Identification is limited only to the micro-organisms included in the database. Species or taxonomic groups not included in the library will not be identified or could be misidentified.
Best practices	The reference database should be up-to-date.
Tips for notification	All relevant information (protocol, results, analysis and interpretation) of the tests must be submitted, along with data and culture conditions for the notified micro-organism and all comparators/reference standards used. Information about the reference database and software used to support the identification of the micro-organism and similarity indices, along with alternate species possibilities must be provided. The specificity, richness and limitations of the database should be clearly conveyed.
References	Busse et al., 1996; da Costa et al., 2011a; Kunitsky et al., 2006; Purcaro et al., 2010; Spiegelman et al., 2005; and Tshikhudo et al., 2013

Genotypic methods

Table B-2.1: sequencing of conserved gene regions	
Method	In most cases, sequence analysis of conserved gene regions, sometimes in combination with highly variable regions in the same gene, often allows for putative placement of a micro-organism in a taxonomic group in relation to other members of the group. 16S rRNA and chaperonin-60 (*cpn*60) genes are commonly used for Bacteria; 16S rRNA and type II chaperonin genes are commonly used for Archaea; ribosomal RNA operons are commonly used for eukaryotes (usually portions of these genes, such as 18S/ITS1, 5.8S, ITS2, partial 28S rRNA).
Level of identification	Mostly to genus- and species-level for certain micro-organisms (depending on the variable regions of the gene sequenced). Its value to microbial identification should be considered to be significant and primary.
Strengths	Because of the highly conserved regions, these sequences are effective universal phylogenetic markers. This method is the most commonly used for identification and phylogenetic classification of micro-organisms. Depending on the conserved gene or regions used and group of micro-organisms, well-supported online tools are available and databases and interpretation of analyses is straightforward. Sequencing of conserved gene regions is simple and cost-effective.
Limitations	Low discriminatory power of 16S rRNA gene at species level for several taxonomic groups because of the highly conserved function (for example, some species of *Bacillus* and *Pseudomonas*). Results may vary depending on: - quality of sequence generated; - clustering methods and/or sequence alignment scores used; - multiple copy numbers, inter-gene variations and the need to analyze a consensus sequence; - the richness of the databases (sequence coverage will influence the results); - the level of curation of the database.
Best practices	The choice of the conserved gene(s) for identification, other than rRNA or *cpn*60, must be explained. Sequence conservation of the chosen gene/region should be well-documented for the species.

Table B-2.1: sequencing of conserved gene regions	
	Must choose primers with specificity for archaea, bacteria or fungi. Must choose suitable reference database (for example, cpnDB, Ribosomal Database Project (RDP) database, CBS-KNAW, SGD, etc.) depending on the notified micro-organism. For 16S rRNA gene analysis, full length sequence analysis is required, unless the use of a shorter sequence has been shown to be discriminatory in the literature. The cut-off for species identification is generally accepted at 99% for full length ribosomal gene sequences (and ideally 99.5% sequence similarity should be used). However, lower cut-off values may be used if references are provided to substantiate the cut-off value for the species in question.
Tips for notification	Sequence in FASTA format must be provided to the NSP, and/or deposited in a database accessible by the NSP. Information about the reference databases and the software used to support the identification of the micro-organism must be provided. Sequence alignment results must include percent identities and alternate species possibilities should be provided. The method, software or algorithm used in the alignment must be provided as well as the length of the sequence used in the alignment.
References	Anderson and Cairney, 2004; Azevedo et al., 2014; Debourgogne et al., 2012; Hanson et al., 2015; Hill et al., 2004; Hill et al., 2006; Hirsch et al., 2010; Hirsch et al., 2013; Janda and Abbott, 2007; Janke et al., 2013; Kurosawa and Itoh, 1993; Links et al., 2012; Parlapani and Boziaris, 2016; Schoch et al., 2012; Srinivasan et al., 2015; Tindall et al., 2010; and Zeaiter et al., 2002

Table B-2.2: Multi-Locus Sequence Alignment (MLSA) or Typing (MLST)	
Method	Phylogenetic analysis of multiple internal fragments of housekeeping genes (typically up to 8), that are ubiquitous to a taxonomic group, present as a single copy within the genome and not subject to selective pressures. Examples include: gyrase A, gyrase B, translation initiation factor 1, translation initiation factor 2, transcription elongation factor 1, recombinase A, recombinase B, cytochrome C oxidase, β-subunit of ATP-synthase, etc.
Level of identification	Analysis of the linked sequences may lead to strain level identification; however, depending on the gene combinations used, discrimination may be possible only at sub-species, species, clade, or genus levels. Its value to microbial identification should be considered to be significant and primary.
Strengths	Housekeeping genes involve cellular metabolism so they can be very discriminatory and reliable if they are appropriately selected. Avoids effects of recombination and horizontal transfer occurring in a single gene. Distinguishes between highly related species and strains, where the analysis of universal target genes shows low resolution. MLST and MLSA are well-supported by online tools and databases.
Limitations	Discriminatory power depends on the housekeeping genes used. For MLSA, linked sequence alignments may not be universally compatible with all software. Limited number of software is available to handle multilocus data sets. Analysis and interpretation of data can be complex when literature is not available to support the use of the chosen MLST scheme in the taxonomic group being studied. Analysis may be complicated for diploid and polyploid organisms.
Best practices	Multiple markers improve the resolution among closely-related species and the success rate of identification. Genes used for MLSA should have: - single copy number in genome - ideally protein-encoding functions - minimal distribution distance of 100 kb between genes - sufficient discriminatory power - nucleotide length to allow convenient sequencing (900 to 2 250 pb)

Table B-2.2: Multi-Locus Sequence Alignment (MLSA) or Typing (MLST)	
	- must predict whole-genome relationships with acceptable precision and accuracy
Tips for notification	Sequences in FASTA format must be provided to NSP, and/or deposited in a database accessible by NSP. Phylogenetic grouping along with percent identities and alternate species possibilities must be provided. Information about the reference databases (for example, PubMLST and MLST.net) and the software used to support the identification of the micro-organism must be provided.
References	Azevedo et al., 2014; Das et al., 2014; Fan et al., 2014; Kurosawa and Itoh, 1993; Macheras et al., 2011; Maiden et al., 1998; Pascual et al., 2010; Thompson et al., 2005; Tindall et al., 2010; and Zeigler, 2003

Table B-2.3: Whole Genome Sequencing (WGS)	
Method	Suitable for identification of species and strains of bacteria, archaea, fungi, virus, microalgae, etc. Comparative genome sequence analysis allows identification of genes or gene complexes that encode putative virulence factors, enzymes, toxins, metabolites, antibiotic or antifungal resistance, abiotic and biotic stress resistance, survival and persistence, etc.
Level of identification	Species- and strain-level identification possible. Its value to microbial identification should be considered to be significant and primary.
Strengths	Highly suitable for micro-organisms that cannot be identified using traditional culture based or single gene sequencing methods. Lower sequencing error due to long sequence reads and overlapping scaffolds. Provides important information specific to the notified micro-organism. Allows rapid MLST typing of various genomic regions of interest (for example, MLST of ribosome protein subunits (*rps* genes) in the genome). WGS analysis and combined bioinformatics platforms and databases are becoming publicly available for clinical diagnostics, epidemiology and surveillance purposes.

Table B-2.3: Whole Genome Sequencing (WGS)	
Limitations	Heterozygous positions in the genome and polyploidy can be challenging. Tedious process involved in collection of large amount of high quality DNA, library preparation and sequencing, etc. Large amount of sequence data generated. Computationally demanding to assemble and annotate the genome, identify the multilocus sequence types, align sequences and use WGS data for species identification. Technically advanced network systems, data handling pipelines and bioinformatics expertise are needed to annotate the results into useful information for microbial identification, as well as, for biological and ecological relevance. Results depend on the richness and quality of reference genome databases used. Currently, no standards available for sequencing results, coverage depth and assembly quality.
Best practices	Best practices depend on platform and methods used for sequencing, including adequate genome coverage.
Tips for notification	Annotated sequences must be provided to the NSP in a machine-readable format and/or deposited in an accessible database. Raw sequence data should be available upon request. If only key genetic sequences are used for identification purposes, then only those annotated sequences need to be provided. Phylogenetic grouping along with percent identities and alternate species possibilities should be provided. Information about the reference databases (for example, NCBI Genome, RDP database, CBS-KNAW, SGD, etc.) and the software used to support the identification of the micro-organism must be provided. Provides important information specific to the notified micro-organism that may be used in other information elements in the notification.
References	Larsen et al., 2012; Thomsen et al., 2016; Ronholm et al., 2016; Salvetti et al., 2016; Saputra et al., 2015; Schatz et al., 2012; and Zhang et al., 2015

Table B-2.4: PCR- and DNA-based typing and hybridization (RAPD, RFLP, Ribotyping, PFGE)	
Methods	Analysis of polymorphisms, including variable regions, non-coding and repetitive sequences of the genome, copy numbers of various genes of interest (for example, virulence factors, antimicrobial resistance, etc.). Methods include: - DNA base ratio (G+C %) - DNA polymorphism studies such as Random Amplification of Polymorphic DNA (RAPD) - Restriction Fragment Length Polymorphism (RFLP) - Pulsed-field gel electrophoresis (PFGE) - Denaturing gradient gel electrophoresis (DGGE) - Variable number tandem repeat (VNTR) - Multiple-Locus Variable number tandem repeat analysis (MLVA) - Amplified restriction fragment length polymorphism (AFLP), etc. - Southern and northern blotting
Level of identification	Species- and strain-level identification may be possible in taxonomic groups in which a specific method has been established as discriminatory. Often used to analyze interspecies variations and establish phylogeny among related species. Their value to microbial identification may be considered to be significant and primary in taxonomic groups in which a specific method has been established for identification purposes.
Strengths	Moderate to high resolution and specificity. Suitable for cultivable micro-organisms. Most methods have standardized protocols and are commercially available.
Limitations	Most of these methods require technical expertise, specialized equipment (PFGE or DGGE) to perform and analyze the results. Analysis and interpretation of data can be complex when literature is not available to support the use of chosen method for identification and taxonomic grouping of similar micro-organisms. For certain techniques like PCR-RFLP, knowledge of genome sequences pertaining to specific endonucleases is required. For methods such as RFLP, standardization of process is important and results vary depending on PCR conditions. Labour- and time-intensive procedure (i.e. as much as 4-7 days for PFGE).

Table B-2.4: PCR- and DNA-based typing and hybridization (RAPD, RFLP, Ribotyping, PFGE)	
Best practices	Suitable reference strains must be included while conducting any of these methods. In the case of PCR-based methods, the choice of the primers used should be justified (i.e., rationale for basing the identification of the particular DNA region or gene) and their nucleotide sequences must be provided.
Tips for notification	All relevant information (protocol, results, analysis and interpretation) of the tests must be submitted, along with data for the notified micro-organism and all comparators/reference standards used.
References	Olive and Bean, 1999; and Raengpradub, 2009

Table B-2.5: DNA-DNA hybridization	
Method	DNA-DNA hybridization is a measure of similarity between genomes. Used to confirm the assignment of a strain to a specific species.
Level of identification	Species level. Applied for taxonomic classification and to identify a new species. Its value to microbial identification should be considered to be significant and primary.
Strengths	High level of confidence in the species designation. Laboratories providing the service are available.
Limitations	Technique is considered difficult due to variability of conditions and results. Suitable for cultivable micro-organisms. Must be optimized at each laboratory. Results are influenced by genome size, presence of large plasmids and purity of the DNA.
Best practices	The most closely-related taxa should be examined at the same time. A rationale for the choice of the most closely-related taxa must be provided.
Tips for notification	All relevant information (protocol, results, analysis and interpretation) of the tests must be submitted, along with data for the notified micro-organism and all comparators/reference standards used. The temperature (indicative of stringency) at which the hybridization is conducted must be indicated along with a rationale for that choice.
References	Stackebrandt *et al.,* 2011; and Wayne *et al.,* 1987

Other methods useful for microbial identification

Table B-3.1: methods that depend on mass spectrometry

Method	Powerful analytical technique. MALDI-TOF (Matrix-Assisted Laser Desorption Ionization-Time of Flight) MS has been reported for a number of purposes including microbial identification and strain typing. Both MALDI-TOF MS and Electrospray ionization-MS (ESI-MS) have been used to analyze bacterial proteins or toxins for rapid species identification. Gel-free protein analysis using iTRAQ, ICAT can be used for cellular proteome and metabolome analysis that can be compared to a library of reference spectra.
Level of identification	Both species identification and strain typing are possible, depending on the species and the methodology. Its value to microbial identification could be considered to be significant and primary, if supported by literature for the taxonomic group.
Strengths	Fast, accurate, sensitive and less expensive than molecular and immunological-based detection methods, as long as the reference library is rich and comprehensive to analyze the species and strain of interest. The MALDI-TOF MS spectrum of an individual microbe is the taxon-specific property of that organism, which is independent of its geographical location, culture conditions (which should not be drastically different) or sample preparation methodology.
Limitations	Data acquisition and analysis is time consuming and requires technical expertise, high cost of the equipment. Identification of new isolates is possible only if the spectral database contains peptide mass fingerprints of the type strains of specific genus/ species/ subspecies/ strains. Lack of appropriate references. Low coverage of existing databases.
Best practices	Mass spectrogram of the notified strain must be compared to that of a representative species/strain. At least 5 to 10 biomarkers are required for species identification; for strain or sub-species identification, either a much higher number of reproducible peaks is required or the presence of peaks specific to a sub-species or serotype.

Table B-3.1: methods that depend on mass spectrometry	
Tips for notification	All relevant information (protocol, results, analysis and interpretation) of the tests must be submitted, along with data for the notified micro-organism and all comparators/reference standards used. The database used for the identification of the notified strain must be clearly described.
References	Cheng *et al.,* 2016; Karger, 2016; Singhal *et al.,* 2015; Suarez *et al.,* 2015; and Wunschel *et al.,* 2012

Table B-3.2: analysis of respiratory lipoquinones	
Method	The characterization of different types of quinones (ubiquinone, menaquinone, dihydro-menaquinone, demethylmenaquinone, rhodoquinone), the length of the isoprenoid side chain and the number of saturated isoprenoid units is used in the identification of Bacteria and Archaea, and in determination of phylogenetic relationships.
Level of identification	Identification of bacterial genus and in certain cases, to species level. Its value to microbial identification should be considered to be auxiliary or supplementary, unless justified with detailed data with sufficient discriminatory power.
Strengths	Can be used to identify different microbial populations in environmental samples like activated sludge. Reliable method if the type of quinones and isoprenoid chain length are already established for the taxonomic group of the notified micro-organism.
Limitations	High level of technical expertise and knowledge of respiratory quinones is required to do the analysis and the interpretation. Requires gas chromatography or HPLC or UPLC to analyze the samples.
Best practices	Chromatography studies may yield only preliminary results and additional confirmatory testing (nuclear magnetic resonance spectra or mass spectrograms) are required to confirm the identification of the quinones.
Tips for notifications	All relevant information (protocol, results, culture conditions, analysis and interpretation) of the tests must be submitted, along with data for the notified micro-organism and all comparators/reference standards used. Chromatographs must be submitted.
References	Busse *et al.,* 1996; Collins and Jones, 1981; Hiraishi, 1997; Spiegelman *et al.,* 2005; and Tindall, 2010

Table B-3.3: analysis of peptidoglycans	
Method	Peptidoglycans are divided into two main types (A and B) based on their cross-linkages
Level of identification	Peptidoglycan structure is uniform in Gram-negative bacteria but varies greatly in some groups of Gram-positive bacteria. Its value to microbial identification should be considered to be auxiliary or supplementary, unless justified with detailed data with sufficient discriminatory power.
Strengths	Peptidoglycan structure is an important taxonomic criterion for characterization of Gram-positive bacteria, in particular for members of the suborder *Micrococcineae*, staphylococci and aerobic endospore-forming bacteria. Simple, fast and requires only a small amount of cells, and inexpensive equipment. A number of methods exist for extraction of peptidoglycans from the cell wall or using whole bacterial cells.
Limitations	The discriminatory power of peptidoglycan structure is restricted to Gram-positive bacteria. No variations among the *Proteobacteria* and *Bacteroidetes* phyla. Not known to be of taxonomic interest in Archaea, fungi, algae, or other groups of micro-organisms.
Best practices	The most closely-related taxa must be examined at the same time. Analysis of 2,6-diaminopimelic acid (Dpm) and OH-Dpm in whole-cell hydrolysates is more reliable than analysis of whole-cell sugars (their composition may vary depending on culture conditions whereas Dpm and OH-Dpm do not change as they originate from the peptidoglycan).
Tips for notifications	All relevant information (protocol, results, analysis and interpretation) of the tests must be submitted, along with data for the notified micro-organism and all comparators/reference standards used.
References	Busse *et al.,* 1996; Schumann, 2011; and Tindall, 2010

Appendix C – Template

Table C-1: template to organize the data generated and the analysis conducted for polyphasic identification of a notified micro-organism

Notified micro-organism designation	Specify the unique strain designation for the notified micro-organism.
Source	Specify source of the strain/isolate.
Test methods Tier 1 Tier 2 Tier 3	Describe the test methods used.
Data Tier 1 Tier 2 Tier 3	Provide clearly documented results of tests conducted on the notified strain and any comparator strains (controls) used. Add any attachment as required.
Analysis	Describe how the data presented above has been analyzed to arrive to a conclusion on microbial identification and substantiation of the notified micro-organism. Relate the characteristics observed in the notified micro-organism to those of the genus, species and strain as described in the reference material. Explain, if any, the discrepancies between the characteristics of the notified strain and comparator strains used.
Conclusion	Report the identity to species-level, or other appropriate taxonomic level.

Appendix D – Microbial identification: Case studies

Case study *Arcobacter* sp.

Table D-1.1: strain information and identification methodology for *Arcobacter* sp. strain W34m

Notified micro-organism designation	*Arcobacter* sp. strain W34m
Source	Enrichment for bacterial growth from a hydrocarbon-contaminated aquifer
Test methods	
Tier 1	1. Genotypic methods: sequence analysis of the 16S rRNA gene (~1400 bp fragment).
Tier 2	2. Genotypic methods: sequence analysis of the gyrase B (*gyrB*) gene (2100 bp fragment).
Tier 3	3. a) Morphological properties (Gram staining, colony colour and size, cell size and shape). b) Physiological and biochemical properties (growth requirements and characteristics, motility).
Data	
Tier 1	1. Primary analysis showed high sequence identity to several species in the *Arcobacter* genus in the NCBI. Refined secondary analysis: comparison of test sequence with complete 16S rRNA gene sequences of the type strain of each of the 16 recognized species in the *Arcobacter* genus from NCBI (Table D-1.2). Using information from the literature (Collado *et al.,* 2011; Levican *et al.,* 2012), it was possible to assign this strain to one of three clades in this genus, clade 1 which contains five species: *Arcobacter marinus*, *Arcobacter halophilus*, *Arcobacter mytili*, *Arcobacter molluscorum* and *Arcobacter bivalviorum*.
Tier 2	2. *gyrB* gene sequence analysis showed similar clade structure as 16S rRNA gene sequence analysis(see Table D-1.3 for identity between strain W34m and other *Arcobacter* species).
Tier 3	3. a) Gram negative, beige colonies ~ 1 mm in diameter, arc-shaped cells between 0.5 and 2 microns in size, observed after 96 hours growth on marine broth agar. b) Optimal growth in 3% NaCl, growth in 2% to 10% NaCl, aerobic and anaerobic growth, moderate motility, catalase positive, oxidase negative, negative for indoxyl acetate hydrolysis.

Analysis	Although phylogenetic analysis of two genes was used to unambiguously assign this strain to a clade containing five species, no definitive species determination was possible. *A. marinus* was initially considered as a possible identity because of 99.6% identity for 16S rRNA gene sequence. However, the *gyrB* gene sequence showed only 94.1% identity with *A. marinus*. A comparison of phenotypic properties of the notified strain with reference strains from the *Arcobacter* genus was carried out and could not help to assign strain W34m to any of these species (Table D-1.4).
Conclusion	Strain W34m could not be assigned to a specific species in the *Arcobacter* genus; however, based on phylogenetic analysis it can be assigned to *Arcobacter* clade 1, a higher taxonomic designation containing five species. All species in clade 1 are considered adequate surrogates for the risk assessment of this strain when strain-specific information is not available. In this case, the assignment to clade 1 helped to reduce the uncertainty regarding the hazard profile generally associated with *Arcobacter* species, as no pathogenic species belongs to this clade (in comparison, clade 3 contains foodborne pathogens). Furthermore, the risk assessment of this strain was robust as the strain itself was well-characterized for a number of important properties.

Table D-1.2: comparison of *Arcobacter* sp. strain W34m 16S rRNA gene sequence with type strains of *Arcobacter* species

		Sequence identity to *Arcobacter* sp. strain W34m
Clade 3	*Arcobacter thereius* CCUG 56902	92.7%
	Arcobacter trophiarum CCUG 59229	92.3%
	Arcobacter skirrowii ATCC 51132	92.4%
	Arcobacter cryaerophilus ATCC 43158	92.1%
	Arcobacter cibarius CCUG 48482	92.6%
	Arcobacter butzleri ATCC 49616	92.7%
Clade 2	*Arcobacter defluvii* CECT 7697	94.6%
	Arcobacter ellisii CECT 7837	94.3%
	Arcobacter venerupis CECT 7836T	94.8%
	Arcobacter nitrofigilis ATCC 33309	94.4%
Clade 1	*Arcobacter marinus* CECT 7277	99.6%
	Arcobacter molluscorum CECT 7696	97.5%
	Arcobacter halophilus ATCC BAA 1022	96.4%
	Arcobacter mytili CECT 7386	94.3%
	Arcobacter bivalviorum CECT 7835T	92.8%
Unassigned to a clade	*Arcobacter sulfidicus* (unspecified strain name)	92.4%

Table D-1.3: comparison of *Arcobacter* sp. strain W34m gyrB gene sequence with type strains of *Arcobacter* species

		Sequence identity to *Arcobacter* sp. strain W34m
Clade 3	*Arcobacter thereius* CCUG 56902	81.0%
	Arcobacter trophiarum CCUG 59229	79.5%
	Arcobacter skirrowii ATCC 51132	81.9%
	Arcobacter cryaerophilus ATCC 43158	80.2%
	Arcobacter cibarius CCUG 48482	83.0%
	Arcobacter butzleri ATCC 49616	81.9%
Clade 2	*Arcobacter defluvii* CECT 7697	84.0%
	Arcobacter ellisii CECT 7837	83.5%
	Arcobacter venerupis CECT 7836T	83.0%
	Arcobacter nitrofigilis ATCC 33309	82.9%
Clade 1	*Arcobacter marinus* CECT 7277	94.1%
	Arcobacter molluscorum CECT 7696	88.8%
	Arcobacter halophilus ATCC BAA 1022	88.0%
	Arcobacter mytili CECT 7386	87.2%
	Arcobacter bivalviorum CECT 7835T	84.0%

Table D-1.4: comparison of *Arcobacter* sp. strain W34m properties to those of other *Arcobacter* species

Characteristics	*Arcobacter sp. strain W34m*	*Arcobacter marinus* CECT 7277 (clade 1)	*Arcobacter halophilus* ATCC BAA 1022 (clade 1)	*Arcobacter nitrofigilis* ATCC 33309 (clade 2)	*Arcobacter butzleri* ATCC 49616 (clade 3)	*Arcobacter cryaerophilus* ATCC 43158 (clade 3)
Cell morphology	arc	arc	arc	arc	helical	helical
Motility	+	+	+	++	+++	++
Growth at 1% salinity	-	+	-	+	+	+
Growth at up to 10% salinity	+	-	+	-	-	-
Growth at 37°C on CCDA	-	-	-	-	+	+
Catalase	+	-	-	+	+	+
Oxidase	-	+	+	+	+	+
Indoxyl acetate hydrolysis	-	+	+	+	+	+
Anaerobic growth	+	-	+	variable	variable	variable

CCDA: *Campylobacter* charcoal deoxycholate agar

Case study *Aspergillus niger*

Table D-2.1: strain information and identification methodology for *Aspergillus niger* VMZ

Notified micro-organism designation	*Aspergillus niger* VMZ
Source	This strain was obtained from a collaborator's collection. It was claimed to be from a culture collection and was originally isolated from bran.
Test methods	
Tier 1	1. Morphological properties: colour, shape and size of colony, conidia and conidiophore
Tier 2	2. Toxin/metabolite production: extrolite profile
Tier 3	3. Genotypic methods: sequence analysis of the large subunit of the ribosomal RNA gene (D2 region and ITS region) and of the calmodulin gene
Data	
Tier 1	1. - Black colonies with a white leading edge on Casitone Yeast Agar (CYA) for 7 days at 25°C. - Globose conidial head with a size of 50.9 ±17.2 µm. - Smooth and colorless conidiophore/stipe. - Conidia that forms chains, smooth, globose and indented center with a size of 4.2 ± 0.5 µm.
Tier 2	2. Extrolites produced: ochratoxin A, fumonisin B, unalenone (kotanins), naphtho-γ-pyrones, pyranonigrin A, pyrophen, tensidol A and B.
Tier 3	3. Comparison of the D2 region using a proprietary database showed that the sequence of strain VMZ matched to a number of sequences of *Aspergillus* species with the same percent identity, including *A. awamori, A. niger* and *A. foetidus*. - Comparison of the ITS region using the Nite Biological Resource Centre (NBRC) ribosomal database showed that the consensus sequence of strain VMZ was identical to a number of database entries for *A. niger* and *A. phoenicis*. - Phylogenetic analysis (Figure D-2.1) showed that the calmodulin gene of strain VMZ is rooted near other calmodulin genes from *A. niger* strains.
Analysis	Morphological properties of *strain* VMZ are consistent with *A. awamori* and *A. niger* and are similar to other species belonging to *Aspergillus* section *Nigri* such as *A. brasiliensis*, *A. tubingensis*

	and *A. acidus,* as reported in the scientific literature (Varga *et al.,* 2011). Morphological properties by themselves are not sufficient for species-level identification. A comparison with publicly available scientific literature shows a similar extrolite profile between *A. niger* and *A. awamori* (Frisvad *et al.,* 2011; Perrone *et al.,* 2011). Therefore, the two species, *A. niger* and *A. awamori*, cannot reliably be distinguished by morphological characteristics or extrolite profiles (Perrone *et al.,* 2011). The D2 region and the ITS region do not contain enough variation to discriminate species in the section *Nigri*. Therefore, secondary markers are required such as calmodulin, β-tubulin or the RNA polymerase II second largest subunit (Samson *et al.,* 2014). The scientific literature reports that the species in section *Nigri* can be distinguished using calmodulin gene sequence data (Samson *et al.,* 2007a; Samson *et al.,* 2007b; Varga *et al.,* 2011). A comparison of a partial calmodulin gene sequence of strain VMZ with calmodulin genes selected from GenBank showed that strain VMZ sequence matched more closely with *A. niger* sequences than with *A. awamori* sequences.
Conclusion	Properties of strain VMZ, data analysis and comparison to descriptions in recent publications support the conclusion that the appropriate identity of strain VMZ is *A. niger*.

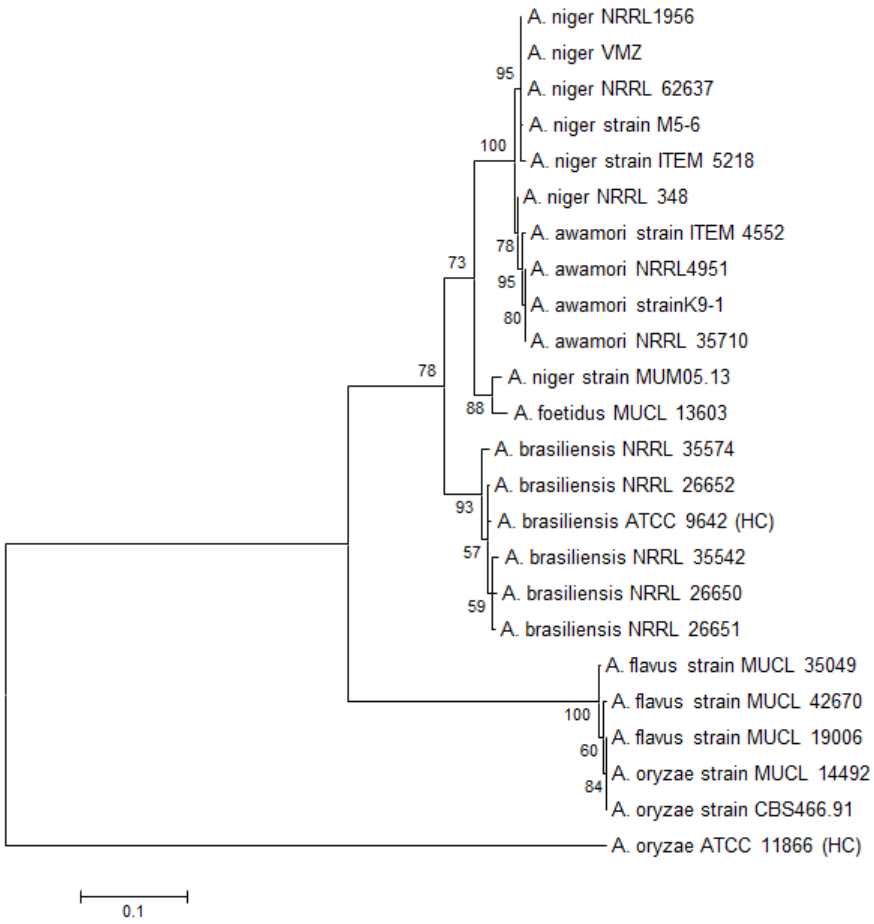

Figure D-2.1: phylogenetic tree generated using partial calmodulin gene sequences of strain VMZ alongside *Aspergillus* sp. calmodulin genes selected from GenBank. The alignment was generated by Muscle and analyzed using the Kimura 2-parameter distance model, which was then used to construct a phylogenetic tree using MEGA version 5.2 (Tamura *et al.,* 2011).

Case study *Aurantiochytrium limacinum*

Table D-3.1: strain information and identification methodology for *Aurantiochytrium limacinum* EnX1

Notified micro-organism designation	*Aurantiochytrium limacinum* EnX1
Source	Environmental isolate from coastal waters, selected for extracellular secretion of Enzyme X1
Test methods	
Tier 1	1. Genotypic methods: sequence analysis of the 18S rRNA gene (~1700bp) using publicly available sequences in NCBI.
Tier 2	2. Morphological properties: cell morphological characteristics based on light and electron microscopy (as described in Yokohoma and Honda, 2007 and Manikan *et al.,* 2015)
Tier 3	3. Chemotaxonomic properties: FAME analysis for polyunsaturated fatty acid profile (as described in Manikan *et al.,* 2015)
Tier 4	4. Physiological and biochemical properties: pigment profile (as described in Yokohoma and Honda, 2007)
Data	
Tier 1	1. Based on 18S rRNA gene sequence analysis (Table D-3.2), strain EnX1 clearly positioned as a member of *Thraustochytriaceae*, with 99.6% similarity to the type strain *Aurantiochytrium limacinum* ATCC MYA-1381 (SR21) and 99.4% *Aurantiochytrium sp.* strain SW1 (deposited name: *Schizochytrium limacinum* ATCC MYA-1381). It also showed > 99% similarity with different strains belonging to *Aurantiochytrium* species and a closely-related genus, *Schizochytrium*.
Tier 2	2. Strain EnX1 formed small colonies on Sea Water Nutrient agar medium, measuring up to 3-4 mm in diameter after 1 week of growth. Its thallus was thin-walled, globose and orange. Vegetative cells (sporangia) were spherical, dispersed as single cells of 8-15 µm diameter; and showed continuous or successive binary divisions to form diads, tetrads and clusters. Mature sporangia appeared to form cell clusters developing into zoosporangia. The motile amoeboid protoplast stage of zoosporangia was observed before the production of zoospores. Zoospores were ovoid in shape and of the biflagellate heterokont morphological type.

Tier 3	3. FAME analysis showed that strain EnX1 cells produced high levels of docosahexaenoic acid (DHA, C22:6) (48%) and hexadecanoic acid (DHA, C16:0) (28%) along with detectable levels of docosahexaenoic acid (DHA, C22:6) (8%), octadecanoic acid (C18:0) (4%) and trace amounts of 6 docosapentaenoic acid (EPA, C20:5) (0.6%) and arachidonic acid. AA, C20:4*n*6 (0.5%).
Tier 4	4. The pigment profile of strain EnX1 showed that its cells possess astaxanthin, phoenicoxanthin, canthaxanthin, echnenone and beta carotene (Table D-3.3).
Analysis	Genotypic analysis of the 18S rRNA gene showed that strain EnX1 belongs either to an *Aurantiochytrium* species of the thraustochytrid family or to the closely-related genus *Schizochytrium.* The morphology, fatty acid and pigment profiles of strain EnX1 were found to be consistent with what has been published in the scientific literature and provided sufficient evidence to support the identification of strain EnX1 as *A. limacinum*: ➢ Strain EnX1 formed small colonies and did not develop ectoplasmic net elements, two critical distinguishing features of the genus *Aurantiochytrium*. In comparison, *Schizochytrium sensu stricto* does not tend to form small colonies regardless of media and develops ectoplasmic net elements (Yokohoma and Honda, 2007). ➢ Strain EnX1 released amoeboid cells in nutrient media, a characteristic feature of the type species *A. limacinum* SR21 (Yokohoma and Honda, 2007) and another *Aurantiochytrium sp.* strain SW1 (Manikan *et al.,* 2015). ➢ The fatty acid profile of strain EnX1 was found to be similar to those of the type species *A. limacinum* SR21 (Yokohoma and Honda, 2007) and *Aurantiochytrium sp.* SW1 (Manikan *et al.,* 2015). ➢ The pigment profile of strain EnX1 was also similar to that of *A. limacinum* type species and distinct from that reported for *Schizochytrium sensu stricto* (Yokohoma and Honda, 2007). Taken together, strain EnX1 can be confirmed as a species of *A. limacinum.*
Conclusion	Morphological, phenotypic and genotypic properties were used to reliably identify the new strain as a strain of *A. limacinum*. The strain was designated *A. limacinum* EnX1.

Table D-3.2: comparison of strain EnX1 18S rRNA gene sequence using NCBI (Nucleotide database - nt)

Descriptions	Query coverage %	% identity
Aurantiochytrium limacinum ATCC MYA-1381 (SR21)	100	99.6
Aurantiochytrium species Strain SW1	99	99.4
Aurantiochytrium species Strains BL11, KRS101, TF23, YLH70	>98	>99.0
Schizochytrium species isolate OUC174, OUC166, LY02012	>98	>99.0

Table D-3.3: carotenoid pigment profile of strain EnX1 compared to a strain of *A. limacinum* and of *Schizochytrium aggregatum*

Micro-organisms	Astaxanthin	Phoenicoxanthin	Canthaxanthin	Echinenone	Carotene
strain EnX1[a]	+	+	+	+	+
Aurantiochytrium limacinum ATCC MYA-1381 (SR21)[b]	+	+	+	+	+
Schizochytrium aggregatum ATCC 28209[b]	-	-	-	-	+

[a] Pigment tests conducted on strain EnX1

[b] Data from Yokohoma and Honda, 2007

Case study *Bacillus amyloliquefaciens*

Table D-4.1: strain information and identification methodology for *Bacillus amyloliquefaciens* strain 74-57

Notified micro-organism designation	*Bacillus amyloliquefaciens* strain 74-57
Source	Lake sediment located in the Lanaudière region in the province of Québec
Test methods	
Tier 1	1. Genotypic methods: sequence analysis of the 16S rRNA gene (~1400 bp fragment).
Tier 2	2. Physiological and biochemical properties: hydrolysis and degradation of carbohydrates, nitrate reduction, Voges-Proskauer reaction, growth on sole carbon source, acid/gas production from carbohydrates, growth temperature and NaCl tolerance.
Tier 3	3. Genotypic methods: V3 region polymorphism of the 16S rRNA gene, based on Jeyaram *et al.,* (2011).
Data	
Tier 1	1. The full length 16S rRNA gene sequence showed 99% identity to sequences from several *B. subtilis* and *B. amyloliquefaciens* strains in NCBI (Table D-4.2).
Tier 2	2. Strain 74-57 is a Gram-positive, rod shaped bacterium positive for esculin, gelatin, casein and Tween 20 degradation. It is positive for the Voges-Proskauer reaction. It reduces nitrate to nitrite, can use citrate as sole carbon source and produces acid without gas from glucose, fructose, maltose, ribose, sucrose and trehalose. Growth optimum at 30°C. NaCl tolerance is up to 8% (Table D-4.3).
Tier 3	3. Sequence alignment of the V3 region of the 16S rRNA gene of strain 74-57 with that from other strains of *B. subtilis* and *B. amyloliquefaciens* showed that strain 74-57 has the same polymorphism as *B. amyloliquefaciens* (Table D-4.4).
Analysis	The 16S rRNA gene sequence showed more than 98% identity to 16S rRNA gene sequences of several strains of *B. subtilis* and *B. amyloliquefaciens* and other *Bacillus* species, including *B. mojavensis*, *B. atropheus*, *B. vallismortis*, etc. (Table D-4.2). The phenotypic properties were able to refine the identity of strain 74-57 to either *B. subtilis* or *B. amyloliquefaciens*, ruling out other *Bacillus* species.

	B. subtilis and *B. amyloliquefaciens* share many characteristics making them notoriously difficult to differentiate based on phenotypic properties. Differentiation between *B. subtilis* and *B. amyloliquefaciens* can be resolved at the species-level by examining polymorphisms within the V3 region of 16S rRNA gene sequence (Jeyaram *et al.,* 2011). The alignment of the V3 region of the 16S rRNA gene sequences of strain 74-57 with those from the *B. subtilis* and *B. amyloliquefaciens* strains confirmed the identity of strain 74-57 as *B. amyloliquefaciens.*
Conclusion	These genotypic, biochemical and physiological properties support the identification of strain 74-57 as *B. amyloliquefaciens*.

Table D-4.2: samples of sequences producing significant alignments

Description	Query coverage	E value	% identity
B. amyloliquefaciens strain ATCC 23350 16S ribosomal RNA gene, partial sequence	100%	0.0	100%
Bacillus sp. SDLI1, complete genome	100%	0.0	99%
B. subtilis strain ATCC 13952, complete genome	100%	0.0	99%
B. amyloliquefaciens XH7, complete genome	100%	0.0	99%
B. amyloliquefaciens strain RD7-7, complete genome	100%	0.0	99%
Bacillus sp. BH072, complete genome	100%	0.0	99%
B. amyloliquefaciens subsp. plantarum NAU-B3, complete genome	100%	0.0	99%
Uncultured *Bacillus sp*. clone Filt.87 16S ribosomal RNA gene, partial sequence	99%	0.0	99%
B. subtilis strain B10 16S ribosomal RNA gene, partial sequence	99%	0.0	99%
B. amyloliquefaciens strain BS5582 16S ribosomal RNA gene, partial sequence	100%	0.0	99%
B. amyloliquefaciens strain Ab-525 16S ribosomal RNA gene, partial sequence	99%	0.0	99%
B. subtilis strain IHB B 1516 16S ribosomal RNA gene, partial sequence	99%	0.0	99%
B. subtilis gene for 16S rRNA, partial sequence, strain: M14K	99%	0.0	99%
B. subtilis strain ET 16S ribosomal RNA gene, partial sequence	99%	0.0	99%
B. amyloliquefaciens strain GXBA-4 16S ribosomal RNA gene, partial sequence	99%	0.0	99%
B. amyloliquefaciens strain LCEP-1 16S ribosomal RNA gene, partial sequence	99%	0.0	99%
Bacillus sp. BIHB 335 16S ribosomal RNA gene, partial sequence	99%	0.0	99%
B. subtilis strain ZJ06 16S ribosomal RNA gene, partial sequence	99%	0.0	99%
B. subtilis isolate G8 16S ribosomal RNA gene, partial sequence	99%	0.0	99%
Etc.			

Table D-4.3: morphological and biochemical properties of strain 74-57 compared to possible species matches in the *B. subtilis* group

Characteristics	*Bacillus subtilis* ATCC 6051	*Bacillus amyloliquefaciens* ATCC 23350	Notified strain 74-57
Gram staining	Gram positive	Gram positive	Gram positive
Cell shape and size	Rod 0.7-0.8×2.0-3.0 µm	Rod 0.7-0.9×1.8-3.0 µm	Rod
Spore	Ellipsoidal to cylindrical spores	Ellipsoidal, central, paracentral or terminal spores	Yes, shape not provided
Casein	+	+	+
Esculin	+	+	+
Gelatin	+	+	+
Starch	+	+	+
Nitrate reduction	+	+	+
Tween 20	+	+	+
Urea	-	-	-
Voges-Proskauer	+	+	+
Citrate as a sole carbon source	+	+	+
Acid without gas is produced from			
glucose	+	+	+
fructose	+	+	+
maltose	+	+	+
ribose	+	+	+
sucrose	+	+	+
trehalose	+	v	+
Catalase	+	Not available	-
Growth temperature	Optimum 28-30°C	Optimum 30-40°C	Positive at 25°C
NaCl 7%	+	+	+

Table D-4.4: alignment of V3 region of 16S rRNA gene sequence of strain 74-57 compared to strains of *B. amyloliquefaciens* and *B. subtilis*. Yellow and blue highlights show the polymorphisms within the V3 region of the 16S rRNA

Designation	Strain	Sequence	GenBank#
B. amyloliquefaciens	ATCC 23842	TTGTTAGGGAAGAACAAGTGCCGTTCAAATAGGGCGGCACCTTG	JF749277
	At4	TTGTTAGGGAAGAACAAGTGCCGTTCAAATAGGGCGGCACCTTG	AY211486
	At1	TTGTTAGGGAAGAACAAGTGCCGTTCAAATAGGGCGGCACCTTG	AY211483
	ATCC 23350	TTGTTAGGGAAGAACAAGTGCCGTTCAAATAGGGCGGCACCTTG	EF433406
Notified strain	74-57	TTGTTAGGGAAGAACAAGTGCCGTTCAAATAGGGCGGCACCTTG	Not applicable
B. subtilis	ATCC 6633	TTGTTAGGGAAGAACAAGTACCGTTCGAATAGGGCGGTACCTTG	EF433403
	DSM 10	TTGTTAGGGAAGAACAAGTACCGTTCGAATAGGGCGGTACCTTG	AJ276351
	JN-1	TTGTTAGGGAAGAACAAGTACCGTTCGAATAGGGCGGTACCTTG	AB07253.1
	ATCC 6051	TTGTTAGGGAAGAACAAGTACCGTTCGAATAGGGCGGTACCTTG	EF423592

Case study *Candida tropicalis*

Table D-5.1: strain information and identification methodology for *Candida tropicalis* ATCC 13803

Notified micro-organism designation	*Candida tropicalis* ATCC 13803 (original designation FDA PCI M-59)
Source	Purchased from ATCC in October 2016. Strain was deposited to American Type Culture Collection (ATCC) by the United States Food and Drug Agency, from an unspecified origin. The datasheet and certificate of analysis are attached.
Test methods	Since the identity of the strain was substantiated by ATCC, the recipient needs only to confirm the identity for quality assurance purposes:
Tier 1	1. Morphological properties: check for purity by plating and confirm colony colour and aspect, and cellular aspect. 2. Genotypic methods: sequence analysis of the 28S rRNA gene (592 bp fragment of the D1D2 region) and comparison to the 521 bp GenBank nucleotide sequence KU729171 which is the reference sequence used by ATCC.
Data	
Tier 1	1. Colonies on Yeast Peptone Galactose agar (YPGA) are cream-coloured, soft and wrinkled near the margin. Budding cells are ellipsoidal. Pseudomycelium is abundant, consisting of long, poorly branched elements. Conidia are arranged in small groups around the middle of each cellular element. 2. The sequence of the D1D2 region was found to be identical (i.e., 100% similarity) to that of the GenBank nucleotide sequence KU729171 (Figure D-5.1).
Analysis	Both morphological and genetic properties conform to the expected characteristics for this strain as described in the datasheet and the certificate of analysis.
Conclusion	The strain received from ATCC is confirmed as *Candida tropicalis* ATCC 13803.

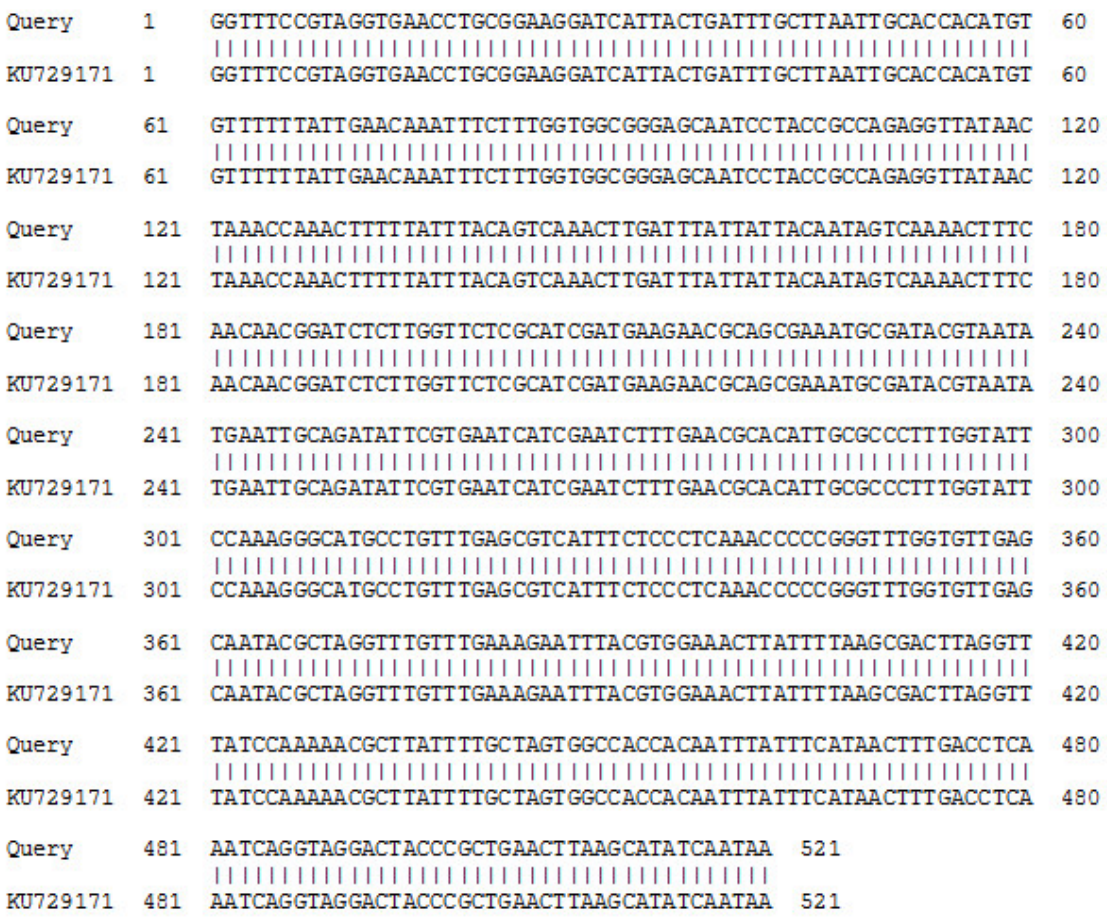

Figure D-5.1: comparison of D1D2 region of 28S rRNA gene sequence of the notified strain with GenBank nucleotide sequence KU729171

Case study *Deinococcus proteolyticus*

Table D-6.1: strain information and identification methodology for *Deinococcus proteolyticus* strain Alpha1

Notified micro-organism designation	*Deinococcus proteolyticus* strain Alpha1
Source	Soil sample near a nuclear generating station in Ontario (Canada).
Test methods	
Tier 1	1. Morphological, physiological and biochemical properties: cell aspect, colony colour, Gram stain, catalase and resistance to irradiation (this particular test was done because the micro-organism was obtained near a nuclear power plant).
Tier 2	2. Genotypic methods: sequence analysis of the 16S rRNA gene (~1100 bp)
Tier 3	3. Genotypic methods: sequence analysis of the cpn60 gene (554 bp)
Data	
Tier 1	1. Strain Alpha1 produces red/pink colonies when grown on solid culture medium; is Gram positive, catalase positive, resistant to 10 kGy gamma irradiation. Analysis of the organism by electron microscopy reveals spherical cells of size between 1.0 and 3.0 µm.
Tier 2	2. A 1.1 Kb amplicon from the 16S rRNA gene of Strain Alpha1 was analyzed using NCBI BLASTN (Table D-6.2). The results suggest that strain Alpha 1 is a member of the genus *Deinococcus*. The highest identity scores reported (95%) are towards two sequences from *Deinococcus proteolyticus* strain MRP. Similarity scores of 93 and 94% are reported with the 16S rRNA sequences from other *Deinococcus* species. In addition, there was a 2% gap in the alignment with 16S rRNA from *Deinococcus proteolyticus* strain MRP.
Tier 3	3. A region of the cpn60 gene from strain Alpha1 (corresponding to positions 274-828 of the cpn60 gene from *Escherichia coli*) was amplified following published protocols (Hill et al, 2006) and its nucleotide sequence was determined. A sequence similarity search in the cpnDB database (http://cpndb.ca/seqComp.php - Hill et al, 2004) showed that it shares 99.3% sequence similarity to the cpn60 gene from the type strain of *D. proteolyticus.* Similarity to the cpn60 gene from other *Deinococcus* species is much lower, at 88.2% to 84.2% (Table D-6.3).

Analysis	The resistance to gamma irradiation strongly suggests that strain Alpha 1 belongs to the *Deinococcus* genus. The data from the morphological, physiological and biochemical tests do not allow the assignment of strain Alpha1 to a particular *Deinococcus* species since those characteristics are shared by several species of that genus. Sequence analysis indicates that the 16S rRNA gene of Strain Alpha1 shares similarity with several species of the *Deinococcus* genus. These similarity scores cannot be relied upon to assign strain Alpha 1 to a given *Deinococcus* species because: ➢ more than 99% sequence similarity is recommended for the assignment to a species (Janda and Abbott, 2002; Tindall *et al*., 2010), and ➢ only 1097 bp of the gene has been sequenced and the percentage of ambiguities (gaps) is above 1%. The length of sequencing is below the ideal threshold of 1300 bp and the percentage of sequence ambiguities (gaps) is above the recommended threshold of 1% (Janda and Abbott, 2002). The sequence analysis of the 554 bp fragment of the cpn60 gene of strain Alpha1 showed it is highly similar to the cpn60 gene from *D. proteolyticus*. While the cpn60 gene of strain Alpha1 also shares sequence similarity with the same gene from other species of *Deinococcus*, the percentage is significantly lower.
Conclusion	These genotypic, biochemical and physiological properties support the identification of strain Alpha1 as *Deinococcus proteolyticus*.

Table D-6.2: comparison of strain Alpha1 16S rRNA gene sequence using NCBI BLAST

Descriptions	Query coverage %	% identity	Gaps %
Deinococcus proteolyticus strain MRP 16S ribosomal RNA gene, complete sequence	98	95	2
Deinococcus proteolyticus MRP, complete genome	98	95	2
Deinococcus proteolyticus MRP 16S ribosomal RNA gene, partial sequence	98	94	2
Deinococcus sp. 14 pro 16S ribosomal RNA gene, partial sequence	96	94	2
Deinococcus sp. Grk2 16S ribosomal RNA gene, partial sequence	97	93	
Deinococcus piscis strain 3ax 16S ribosomal RNA gene, partial sequence	98	93	

Table D-6.3: comparison of a 552 bp fragment from strain Alpha1 cpn60 gene against the cpnDB database

Descriptions	cpnID	Query coverage %	E value	% identity
NC_015161 *Deinococcus proteolyticus* MRP (type strain)	b18672	552	38	99.3
NZ_KB899708 *Deinococcus aquatilis* DSM 23025	b27375	552	1.7e+02	88.2
AY453859 *Deinococcus grandis* DSMZ 3693	b9497	552	2e+02	87.1
CP002191 *Deinococcus gobiensis* I-0	b20546	552	2e+02	86.8
CP001114 *Deinococcus deserti* VCD115	b13538	552	2.1e+02	86.4
NZ_ATTJ01000001 *Deinococcus* sp. 2009	b28170	552	2.3e+02	85.8
NC_001263 *Deinococcus radiodurans* R1	b1273	552	2.4e+02	85.7
CP000359 *Deinococcus geothermalis* DSM 11300	b7527	552	2.8e+02	84.4
NZ_APCS01000080 *Deinococcus wulumuqiens* R12	b26963	552	2.9e+02	84.2

Case study *Influenza* virus

Table D-7.1: strain information and identification methodology for *Influenza virus* A/Vancouver/35/2016 vaccine strain (ca A/Vancouver)

Notified micro-organism designation	*Influenza virus* A/Vancouver/35/2016 vaccine strain (ca A/Vancouver)
Source	Live attenuated cold adapted reassortant influenza virus derived from the cold adapted virus A/Ann Arbor/6/60 and the wild-type influenza virus A/Vancouver/35/2016 (H3N2)
Test methods	
Tier 1	1. Genotypic methods: whole genome sequence
Tier 2	2. Serological properties: hemagglutination-inhibition (HAI) of influenza H3 viruses
Data	
Tier 1	1. Genomic sequencing of the notified strain, which is a 6:2 reassortant, showed that its genome segments are nearly identical to those of its donor organisms: cold-adapted (ca) A/Ann Arbour/6/60 master donor virus, which contributed segments PB1, PB2, PA, NP, M and NS, and wild-type A/Vancouver/35/2016 (H3N2), which contributed segments HA and NA (see Table D-7.2 for the sequence comparison between donor viruses and notified virus). Three point mutations were observed where two translated into an amino acid change.
Tier 2	2. ca A/Vancouver reacts with ferret antisera raised to the reference strain giving an HAI titer equal to or within two-fold of the HAI titer of the wild type reference strain (Table D-7.3).
Analysis	The sequences of ca A/Vancouver genome segments PB2, PB1, PA, NP, M, and NS are identical with those of the corresponding segments of A/Ann Arbour/6/60 except for one mutation in the Matrix segment. The NA genome segment of the notified virus is identical to the corresponding DNA segment of the wild type A/Vancouver/35/2016 and two point mutations were observed in the HA segment. The 2 amino acid changes on the HA antigen did not reduce its antigenicity which was confirmed by serotyping by the HAI test and confirmed the H3 serotype of ca A/Vancouver.
Conclusion	Genomic segment analysis and HAI data unambiguously identify the notified virus strain as an influenza virus with antigenic properties identical to the wild-type A/Vancouver/356/2016, as well

	as cold adapted properties. Therefore, the designation of the notified virus is accepted as cold adapted A/Vancouver/35/2016.

Table D-7.2: sequence comparison of the notified cold adapted ca A/Vancouver vaccine strain and the ca A/Ann Arbor/6/60 and the wild-type influenza virus A/Vancouver/35/2016 (H3N2)

Influenza segment	ca A/Vancouver (modified nucleotide position in bracket)	Wild type A/Vancouver/35/2016 (H3N2)	ca A/Ann Arbor/6/60
PB1	Identical	No data	Identical
PB2	Identical	No data	Identical
Pa	Identical	No data	Identical
NP	Identical	No data	Identical
M	A (Thr) (at 256)	No data	T (Thr) (at 256)
NS	Identical	No data	Identical
HA	G (Gly) (at 637) G(Gly) (at 735)	A (Asp) (at 637) A (Gly) (at 735)	No data
NA	Identical	Identical	No data

Table D-7.3: hemagglutination reaction of Influenza H3 viruses (using reference ferret sera)

	A/HongKong/4801/2014	A/California/7/2004	A/New Caledonia/20/99	A/Vancouver/35/2016	ca A/Vancouver
Reference Antigen					
A/HongKong/4801/2014	320	160	80	160	160
A/California/7/2004	160	320	40	160	320
A/New Caledonia/20/99	40	40	1280	40	40
A/Vancouver/35/2016	160	160	80	320	160
Test Antigen					
ca A/Vancouver	160	160	80	640	320

Case study *Komagataella phaffii*

Table D-8.1: strain information and identification methodology for *Komagataella phaffii* strain C345

Notified micro-organism designation	*Komagataella phaffii* strain C345
Source	Environmental isolate from black oak trees in California, U.S.A
Test methods	
Tier 1	1. Morphological properties: cellular and colony morphology
Tier 2	2. Genotypic methods: sequence analysis of the LSU rRNA operon (D1/D2 region) using CBS-KNAW database and NCBI GenBank, as described in Kurtzman (2009)
Data	
Tier 1	1. Cells of strain C345 are spherical to oval. Growth on agar is tannish-white, has a dull surface and is butyrous in texture. Pseudohyphae and hyphae are absent. Colony margins are finely to moderately lobate. Ascospores are hat-shaped.
Tier 2	2. A pairwise sequence alignment of the D1/D2 region of the LSU rRNA operon of strain C345 in the CBS-KNAW database shows 99.8% similarity with *K. phaffii* strain CBS 2612^{T}, 98.6% similarity with *K. kurtzmanii* CBS 12817^{T}, 97.9% with *K. ulmi* CBS 12361^{T} and 97.6% similarity with *K. pastoris* CBS 704^{T}. The BLAST searches using NCBI Genbank also showed similar results (Table D-8.2).
Analysis	The morphological description is consistent with the reported morphology of *K. phaffii* (previously known as *Pichia pastoris*). However, *K. phaffii* cannot be separated from other closely-related *Komagataella* species by cell and colony morphology or their reactions on standard fermentation and assimilation tests (Kurtzman, 2005; Kurtzman, 2009). Pairwise sequence alignment of the D1/D2 region of LSU rRNA showed only one possibility at more than 99% identity to *K. phaffi* CBS 2612^{T}. This clearly identifies this strain as *K. phaffii.*
Conclusion	Based on the morphological properties and sequence analysis of D1/D2 LSU rRNA region, strain C345 was identified as *Komagataella phaffii* species. The notified strain has been designated as *Komagataella phaffii* strain C345. Given the taxonomic reclassification of *K. phaffii,* the name *P. pastoris* should also be used when performing literature searches.

Table D-8.2: comparison of strain C345 D1/D2 LSU rRNA gene sequence to other *Komagataella* species using CBS-KNAW database

Sequence identity of *Komagataella phaffii* C 345 to	% Identity
K. phaffii CBS 2612 = NRRL Y-7556	99.8
K. kurtzmanii CBS 12817 = NRRL Y-63667	98.6
K. ulmi CBS 12361 = NRRL YB-407	97.9
K. pastoris CBS 704 = NRRL Y-1603	97.6

Case study *Listeria monocytogenes*

Table D-9.1: strain information and identification methodology for *Listeria monocytogenes* strain WEX 321

Notified micro-organism designation	*Listeria monocytogenes* strain WEX 321
Source	Genetically modified organism derived from the wild type *L. monocytogenes* strain 10403S isolated from human skin lesions. Strain WEX 321 is highly attenuated as a result of deletions of sigma factor (*sigL*) gene and is capable of expressing human Interferon-γ (IFN-γ) as a result of insertion of an IFN-γ expression cassette into its genome.
Test methods	
Tier 1	1. a) Genotypic methods: genome sequence of the parental strain b) Morphological properties: (Gram staining, colony and cell properties)
Tier 2	2. a) Genotypic methods: PCR analysis of the integration locus of the expression cassette b) Genotypic methods: DNA sequencing of expression cassette c) Serological properties: western blot, Protein expression IFN-γ
Data	
Tier 1	1. a) Comparison of sequence of the parental strain (accession number NC_017544) to other *L. monocytogenes* genomes gave between 96-99% symmetrical identity, and comparison of the 16S rRNA gene sequence gave 99% identity to more than 50 strains of *L. monocytogenes* in GenBank using MegaBlast alignment tool and the nr/nu database and 100% identity to *L. monocytogenes* NCTC 10357, 99% identity to *Listeria innocua* NCTC 11288, 98% identity to *Listeria welshimeri* NCTC 11857, and 97% identity to *Listeria seeligeri* NCTC 11856 (Table D-9.2). b) Gram-positive and rod-shaped, growth of blue colonies without a yellow halo in the chromogenic Rapid'L.mono (RLM) agar plates. Identical data was obtained for the parental strain and strain WEX 321 (Figure D-9.1).
Tier 2	2. a) Amplification of *sigL* gene regions in strain WEX 321 (900 bp) showed a deletion compared to the same amplification in the parental strain (2545 bp). Specific amplification of the IFN-γ expression cassette in strain WEX 321 compared to the same amplification in the parental strain and showed a DNA bands of the appropriate size for the expression cassette in WEX 321 (1555 bp) (Figure D-9.2).

	b) The DNA sequence of the IFN-γ expression cassette in strain WEX 321 was done and showed 100% identity to the predicted *in silico* sequence. c) IFN-γ expression was shown by the presence of a distinct band in the western blot results. Appropriate controls and weight ladder were included showing the specificity to IFN-γ (Figure D-9.3).
Analysis	Based on the full genome and the 16S rRNA gene sequences of the parental strain of strain 10403S using a cut-off value of 98.9% (Stackebrandt 2011) and colony morphology on RLM plates (confirming phosphatidylinositol phospholipase C activity and its inability to metabolize xylose), the identity of WEX 321 is confirmed to be *L. monocytogenes.* Data provided allows distinction of strain WEX 321 from other *Listeria* pathogenic species: *Listeria ivanovii, Listeria innocua* and *Listeria welshimeri*. Data obtained on the amplification of the *sigL* region and the DNA sequencing of the expression cassette of the parental strain and comparison to the DNA of the strain WEX 321 as well as the western blot analysis, confirmed that the notified micro-organism is *L. monocytogenes* strain WEX 321.
Conclusion	Analysis of all data confirmed that the notified strain belongs to *L. monocytogenes* taxon and it can specifically be identified as strain WEX 321.

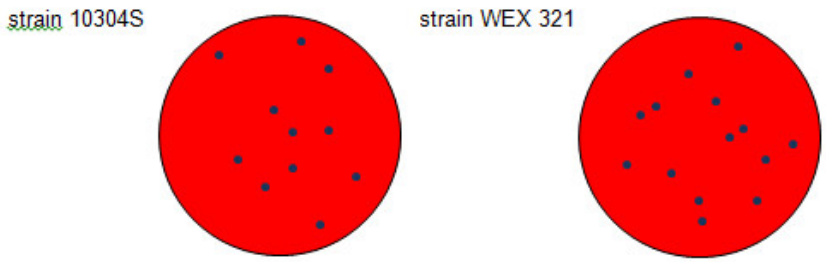

Figure D-9.1: growth of strain 10304S and strain WEX 321 on Rapid'L.mono agar plates for 24h at 30˚C.

Table D-9.2: comparison of strain 10403S 16S rRNA gene sequence in NCBI to rRNA type strain database

Genus species strain	Query coverage %	E value	% identity
Listeria monocytogenes NCTC 10357	100	0.0	100
Listera innocua NCTC 11288	100	0.0	99
Listeria welshimeri NCTC 11857	100	0.0	98
Listeria seeligeri NCTC 11856	98	0.0	97

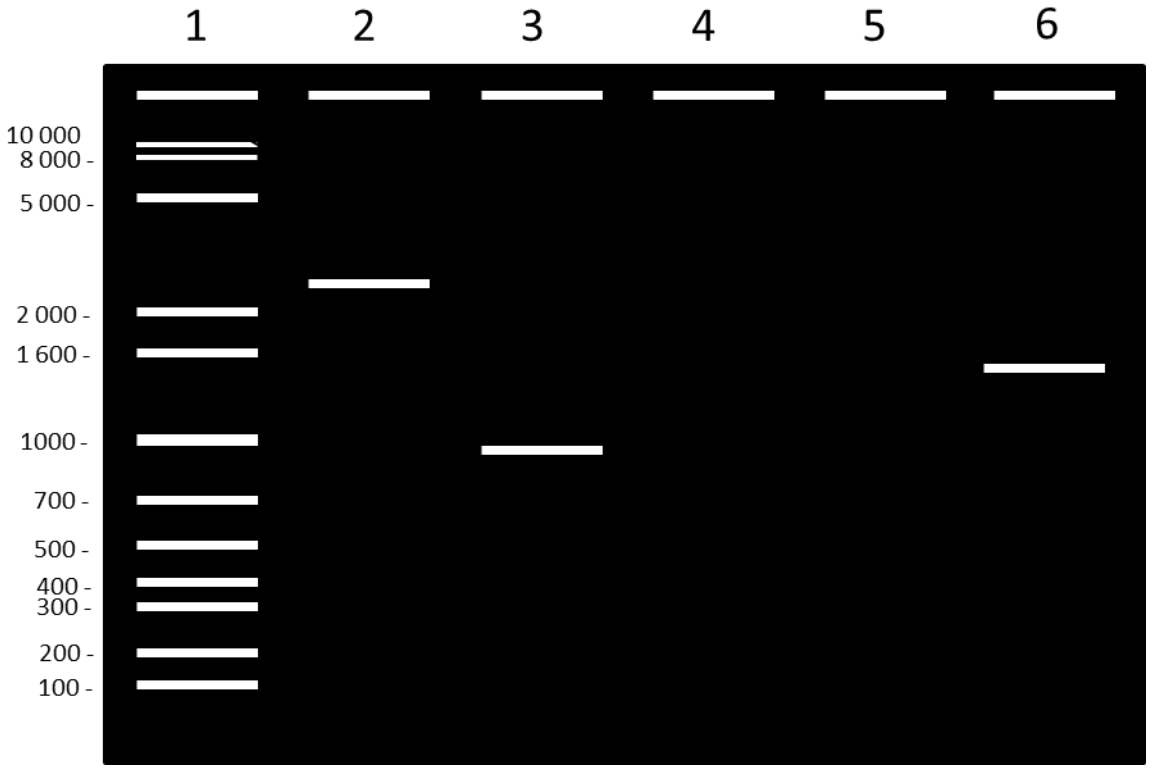

Figure D-9.2: agarose gel electrophoretic picture of PCR results for strain 10403S and strain WEX 321. Lane 1: Molecular weight ladder (kb); lane 2: PCR amplification of *sigL* locus in strain 10403S; lane 3: PCR amplification of *sigL* locus in strain WEX 321; lane 4: PCR amplification of *sigL* locus in *S. cerevisiae* (negative); lane 5: PCR amplification of hIFN-γ in strain 10403S (negative); lane 6: PCR amplification of hIFN-γ in strain WEX 321.

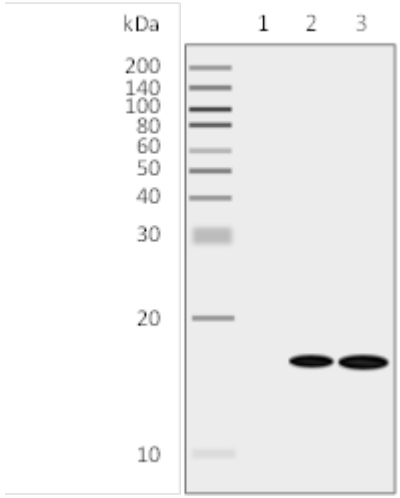

Figure D-9.3: western blot analysis of hIFN- γ using mouse anti-hIFN-γ. Lane 1: Protein extract from strain 10403S, lane 2: Protein extract from strain WEX 321; lane 3: 10 ng of hIFN- γ.

Case study *Rhodococcus aetherivorans*

Table D-10.1: strain information and identification methodology for *Rhodococcus aetherivorans* strain Rae1

Notified micro-organism designation	*Rhodococcus aetherivorans* strain Rae1
Source	Enrichment of a hydrocarbon activated sludge obtained from an industrial wastewater treatment site in Canada.
Test methods	
Tier 1	1. Genotypic methods: sequence analysis of the 16S rRNA gene (~1421 bp) using NCBI BLAST and RDP database.
Tier 2	2. Chemotaxonomic properties: FAME analysis by gas chromatography according to the Microbial Identification System (MIDI; v 6.2) with the RTSBA6 database using whole cells grown on tryptic soy broth agar (TSBA) for 5 days at 25°C.
Tier 3	3. a) Physiological and biochemical properties: Assay using API 20NE kit (BioMérieux). b) Morphological properties: Colony and cell morphology on Yeast Extract Agar.
Data	
Tier 1	1. NCBI BLAST analysis of the 16S rRNA gene sequence of strain Rae1 showed that this strain is affiliated with the genus *Rhodococcus* and falls within the *Rhodococcus rhodochrous* 16S rRNA clade (Table D-10.2). Strain Rae1 showed 99.7% similarity with *Rhodococcus aetherivorans* DSM 44752^{T}, 99.5% similarity with the type strain of *Rhodococcus ruber* KCCM 41053^{T}, 97.8% with *Rhodococcus zopfii* DSM 44108 ^{T} and 97.2% with *Rhodococcus phenolicus* DSM 44812^{T} (Table D-10.3). Also, sequence analysis using the RDP database for type strains revealed high similarity index of > 0.999 with *R. aetherivorans* and *R. ruber* (Tables D-10.4 and D-10.5).
Tier 2	2. The fatty acid profile of strain Rae1 was very similar to those of *R. aetherivorans* KCCM 41053^{T}, especially for some unsaturated fatty acids ($C_{18:1}$ω9*c cis*-9-Octadecenoic acid; 10-Methyl $C_{18:0}$ Tuberculostearic acid), which was discriminatory between *R. aetherivorans* and *R. ruber*. However, the composition of other fatty acids of strain Rae1 was similar to those of both (Table D-10.6).
Tier 3	3. API 20NE tests showed that strain Rae1 is capable of utilizing glucose, mannose, cellobiose and D-galactose; positive for nitrate reduction; and negative for gelatin and urea hydrolysis (Table D-

	10.7). Colony and cell morphology: Strain Rae1 appeared rough and pinkish on yeast extract agar. It is a Gram-positive rod (approximately 0.5–1.0 µm wide and 1.5–5.0 µm long) and a non-spore-forming bacterium.
Analysis	Based on 16S rRNA gene sequence analysis, strain Rae1 can be assigned to the *Rhodococcus rhodochrous* 16S rRNA clade. Strain Rae1 is closely related to two species in that clade, namely, *R. aetherivorans* and *R. ruber.* The high 16S rRNA gene sequence similarity of strain Rae1 with *R. aetherivorans* and *R. ruber*, suggested that both species are possible candidates. The FAME analysis also revealed that strain Rae1 had higher similarity with *R. aetherivorans* (for unsaturated fatty acids composition for $C_{18:1}\omega 9c$ *cis*-9-Octadecenoic acid; 10-Methyl $C_{18:0}$ Tuberculostearic acid) than with *R. ruber*; however, other fatty acids were not discriminatory. The ability of the strain Rae1 to utilize glucose, mannose, cellobiose, D-galactose and N-acetyl glucosamine, and its inability to hydrolyse gelatin and urea and ability to reduce nitrate, is similar to *R. aetherivorans*. The colony and cell morphology of strain Rae1 are consistent with the type strain of *R. aetherivorans* DSM 44752^T, although morphological properties do not differentiate *R. aetherivorans* and *R. ruber*.
Conclusion	Taken together, the data indicates that strain Rae1 belongs to the *R. aetherivorans* species.

Table D-10.2: comparison of strain Rae1 16S rRNA gene sequence using NCBI BLAST-nr/nt datadase

Descriptions	Query coverage %	E value	% identity
Rhodococcus sp. WB1 *; Rhodococcus* sp. USA-AN012 ; *Rhodococcus* sp. L3	100	0.0	100
R. aetherivorans strains icdP1, IAR1, 126189	100	0.0	100
R. aetherivorans strain DSM 44752 $^{T=}$10bc312 = JCM 14343 = NCIMB 13964.	98	0.0	99.7
R. ruber DSM 43338 ^{T}	100	0.0	99.5
R. ruber isolate OUCZ91B	99	0.0	100
R. ruber strain M2	100	0.0	99.6
Several uncultured *Rhodococcus* species	100	0.0	>99

Table D-10.3: comparison of strain Rae1 16S rRNA gene sequence using NCBI BLAST-ref-seq_rna database (limited to type strains)

Descriptions	Query coverage %	E value	% identity
R. aetherivorans DSM 44752 ^{T}	98	0.0	99.7
R. ruber DSM 43338 ^{T}	100	0.0	99.5
R. zopfii DSM 44108 ^{T}	100	0.0	97.8
R. phenolicus DSM 44812 ^{T}	98	0.0	97.2

Table D-10.4: comparison of strain Rae1 16S rRNA gene sequence using RDP database

Sequence ID	Similarity score	S_ab Score	Unique common oligomers	Sequences full name
S000015815	1.000	1.000	1267	*Rhodococcus ruber*; AS4.1187; AF350248
S000393887	1.000	1.000	1404	*Rhodococcus* sp. USA-AN012; AF420413
S000539645	0.998	0.998	1231	*Rhodococcus ruber*; IV11; AJ833916
S000893761	0.998	1.000	1299	*Rhodococcus aetherivorans*; AK44; EU004422
S000965726	1.000	1.000	1316	*Rhodococcus sp.* 9camb; EF151233
S001014613	0.999	0.995	1410	*Rhodococcus sp.* L3; EF426447
S001155726	0.999	0.995	1405	*Rhodococcus sp.* XQ-K; EU876664
S001187762	0.999	1.000	1307	*Rhodococcus sp.* NCIMB 9784; EU445342
S001572358	1.000	1.000	1328	*Rhodococcus aetherivorans*; IAR1; AB453385
S002165055	1.000	1.000	1397	*Rhodococcus aetherivorans*; IR34-DHCE-402; AB546298
S002907533	1.000	1.000	1221	*Rhodococcus aetherivorans*; BW38; HE578785
S002949467	1.000	1.000	1244	*Rhodococcus aetherivorans*; M8; AB610652
S002957974	1.000	1.000	1196	*Rhodococcus ruber*; W3; JN613346
S004052277	0.999	1.000	1306	*Rhodococcus aetherivorans* I24; KF410351
S004052290	0.999	1.000	1307	*Rhodococcus aetherivorans*; DSM 44752; KF410364
S004071425	1.000	1.000	1440	*Rhodococcus sp.* BCP1; CM002177
S004071426	1.000	0.996	1440	*Rhodococcus sp.* BCP1; CM002177
S004091284	1.000	0.995	1338	*Rhodococcus aetherivorans*; 8; KJ571061
S004232883	1.000	0.995	1399	*Rhodococcus sp.* ADA-2; KM210251
S004449071	1.000	0.995	1394	*Rhodococcus sp.* FCL1; KM461685

Table D-10.5: comparison of Strain Rae1 16S rRNA gene sequence using RDP database (limited to type strains)

Sequence ID	Similarity score	S_ab Score	Unique common oligomers	Sequences full name
S000010863	0.971	0.872	1384	*Rhodococcus rhodochrous* (T); X79288
S000126126	0.961	0.838	1396	*Rhodococcus* equi (T); DSM 20307T; AF490539
S000322868	0.999	0.957	1344	*Rhodococcus aetherivorans* (T); 10bc312; AF447391
S000359172	0.973	0.873	1295	*Rhodococcus phenolicus* (T); G2P; AY533293
S000364376	0.974	0.846	1386	*Rhodococcus rhodnii* (T); type strain: DSM43336; X80621
S000364380	1.000	0.985	1400	*Rhodococcus ruber* (T); type strain: DSM43338; X80625
S000388573	0.972	0.882	1390	*Rhodococcus pyridinivorans* (T); PDB9; AF173005
S000388867	0.976	0.904	1389	*Rhodococcus zopfii* (T); DSM 44108 (T); AF191343
S000394065	0.970	0.857	1400	*Rhodococcus corynebacterioides* (T); DSM 20151; AF430066
S000403334	0.966	0.848	1326	*Rhodococcus gordoniae* (T); W4937; AY233201
S000424749	0.962	0.838	1369	*Rhodococcus kroppenstedtii* (T); K07-23; AY726605
S000438867	0.965	0.837	1398	*Rhodococcus coprophilus* (T); JCM 3200; U93340
S000544287	0.971	0.852	1394	*Rhodococcus triatomae* (T); type strain: IMMIB RIV-085; AJ854055
S002918508	0.965	0.838	1312	*Rhodococcus nanhaiensis* (T); SCSIO 10187; JN582175

Table D-10.6: cellular fatty acid contents of strain Rae1 and closely-related *Rhodococcus* species

Fatty acid	strain Rae 1 [a]	*R. aetherivorans* DSM 44752[b]	*R. ruber* KCCM 41053[b]	*R. ruber* DSM 43338 [c]
Saturated				
$C_{14:0}$ tetradecanoic acid	2.0	1.8	2.4	1.5-2.0
$C_{15:0}$ pentadecanoic acid	2.5	2.6	1.6	2.9-4.0
$C_{16:0}$ hexadecanoic acid	22.1	23.5	28.5	25-27.4
$C_{17:0}$ septadecanoic acid	4.5	3.0	1.9	3.0-4.3
$C_{18:0}$ octadecanoic acid	3.0	2.2	3.4	1.0-2.4
Unsaturated $C_{18:1}\omega 9c$ *cis*-9-octadecanoic acid	16.8	15.6	22.7	16.0-20.7
Methyl				
10-Methyl $C_{17:0}$	2.6	1.7	trace	1.8-3.0
10-Methyl $C_{18:0}$ Tuberculostearic acid	26.6	27.6	17.7	15.6-18.0
Summed features‡				
3 ($C_{16:1}\omega 6c$ and/or $C_{16:1}\omega 7c$)	22.4	20.1	16.2	Not available

Values are percentages of total fatty acids.

[a] FAME analysis conducted on strain Rae1;

[b] Results published in Hwang *et al.,* 2015;

[c] Results published in Jones and Goodfellow, 2012;

‡Summed features represent groups of two or three fatty acids that could not be separated by gas chromatography with the MIDI system

Table D-10.7: differential biochemical characteristics of strain Rae1 and its closely-related species

Characteristics	strain Rae1[a]	*R. aetherivorans* DSM 44752[b c]	*R. ruber* KCCM 41053[b c]
Gelatin hydrolysis	-	-[b]	-[b]
Nitrate reduction	+	+[b]	-[b]
Urea hydrolysis	-	-[b]	+[b]
Utilization of :			
N-Acetyl-D-glucosamine	+	+[b]	-[b]
Cellobiose	+	+[b] and -[c]	-[b]
Galactose	+	+[b]	+[b] and -[c]

[a] API 20NE test conducted on strain Rae1;

[b] Results published in Hwang *et al.,* 2015;

[c] Results published in Jones and Goodfellow, 2012

Case study *Saccharomyces cerevisiae*

Table D-11.1: strain information and identification methodology for *Saccharomyces cerevisiae* strain BioEt

Notified micro-organism designation	*Saccharomyces cerevisiae* strain BioEt
Source	A wild wine strain isolated from a vineyard and maintained in a culture collection. The strain was adapted for high temperature tolerance.
Test methods	
Tier 1	1. Morphological properties: colony colour and shape, cell aspect.
Tier 2	2. Physiological and biochemical properties: temperature, pH, salinity, carbon utilization, growth without vitamins and antibiotic susceptibility.
Tier 3	3. Genotypic methods: sequence analysis of the rRNA operon (~2300 bp including ITS1, ITS2, D1/D2/D3 regions).
Data	
Tier 1	1. Colonies are butyrous, cream-coloured, opaque with a smooth surface occasionally raised or folded, when plated on yeast extract peptone dextrose (YPD) agar. Cells are ovoid with some budding observed.
Tier 2	2. Physiological and biochemical properties of strain BioEt: • capable of growth between 27 - 42°C. • growth observed in the pH range of 3.0 - 6.0. • capable of growth at NaCl concentrations of 0 - 3.0% • forms pseudohyphae on corn meal or low nitrogen SLAD medium. • positive for maltose fermentation and growth without vitamins; no growth on mannitol or glycerol as sole carbon sources (Table D-11.2). • susceptible to amphotericin B and 5-fluorocytosine; and resistant to griseofulvin, itraconazole and terbinafine.
Tier 3	3. Multiple sequence alignment using a proprietary Microseq® ID fungal D2 LSU database showed that strain BioEt shares 100% identity with S. cerevisiae ATCC 18824 (type strain) and 99.52% identity with *S. cerevisiae* ATCC 9763 (Table D-11.3). To confirm this with a higher coverage database, a multiple sequence alignment was conducted using the ITS1/5.8S/ITS2 regions of

	strain BioEt and publicly available sequences from NCBI. Strain BioEt showed 99% identity with other strains of *S. cerevisiae* (including clinical isolate *S. cerevisiae* YJM 451) and *S. cerevisiae* var. *boulardii* (Isolates Biocodex, Unique 28); and 98% identity with *S. paradoxus*, *S. pastorianus* and *S. bayanus* (Table D-11.4).
Analysis	The morphological properties of strain BioEt are consistent with *S. cerevisiae* strains as reported in the literature (Barnett *et al.,* 2000; Vaughan-Martini and Martini, 2011). Comparison of physiological and biochemical properties of strain BioEt and possible candidate species showed that strain BioEt has properties that are common to *S. cerevisiae, S. bayanus*, *S. paradoxus* and *S. pastorianus*. Based on rRNA operon sequence analysis, the identity of strain BioEt can be narrowed down to *S. cerevisiae*. The antifungal susceptibility profile of Strain BioEt is also similar to that of other *S. cerevisiae* strains as reported in the literature.
Conclusion	Based primarily on genotypic information, with confirmation from morphological, physiological and biochemical properties, strain BioEt can be identified with certainty as *S. cerevisiae.*

Table D-11.2: comparison of biochemical properties of strain BioEt with *Saccharomyces* species

Characteristics	Maltose fermentation	Melibiose fermentation	Inulin utilization	Growth without vitamins	Mannitol	Glycerol
strain BioEt	+	-	-	+	-	-
S. cerevisiae ATCC 18824 [a]	+	-	-	-	-	-
S. cerevisiae var. *boulardii* ATCC MYA 796 [a]	+	-	-	+	-	-
S. cerevisiae YJM 309 [a]	+	-	-	+	-	-
S. cerevisiae [b]	v	v	-	v	-	v
S. paradoxus [b]	v	-	-	v	+	v
S. bayanus [b]	v	v	-	v	v	v
S. pastorianus [b]	+	v	-	v	-	v

+ indicates positive; **-** indicates negative; v indicates variable;

[a] Data generated along with Strain BioEt;

[b] Data compiled from The Yeasts, a Taxonomic Study (Vaughan-Martini and Martini, 2011)

Table D-11.3: comparison of strain BioEt rRNA gene sequence for ITS1/5.8S/ITS2 regions using Microseq® ID fungal D2 LSU database

Sequence Entry	% identity
Saccharomyces cerevisiae (ATCC 18824)	100.00
Saccharomyces cerevisiae (ATCC 9763)	99.55
Saccharomyces bayanus (ATCC 76513)	98.12
Saccharomyces pastorianus (ATCC 12752)	97.96
Zygosaccharomyces microellipsoides (ATCC 10605)	95.85

Table D-11.4: Comparison of strain BioEt rRNA gene sequence for ITS1/5.8S/ITS2 regions using NCBI BLAST

Descriptions	Query coverage %	E value	% identity
S. cerevisiae ATCC 18824	99	0.0	99
S. cerevisiae strain CBS 1171	99	0.0	99
S. cerevisiae var. boulardii strain biocodex	99	0.0	99
S. cerevisiae var. boulardii strain Unique 28	99	0.0	99
S. cerevisiae YJM 451	99	0.0	99
S. paradoxus strain CBS 432	98	0.0	98
S. pastorianus strain NRRL-Y-17217	98	0.0	98
S. mikatae ATCC MYA-4448	98	0.0	98
S. bayanus, strain CBS 380	98	0.0	98

Case study *Shewanella indica*

Table D-12.1: strain information and identification methodology for *Shewanella indica* strain ECM3

Notified micro-organism designation	*Shewanella indica* strain ECM3.
Source	Enrichment for bacterial growth from a hydrocarbon-contaminated aquifer
Test methods	
Tier 1	1. Genotypic methods: sequence analysis of the 16S rRNA gene (~1400 bp, following procedures and primers suggested in Ruimy *et al.,* 1994).
Tier 2	2. Genotypic methods: sequence analysis of the gyrase B (*gyr*B) gene (~1200 bp, following procedures and primers suggested in Yamamoto and Harayama, 1995).
Tier 3	3. a) Morphological properties: Gram staining, colony and cell properties. b) Physiological and biochemical properties: growth requirements and characteristics.
Data	
Tier 1	1. A BLAST search sequence analysis compared the 16S rRNA gene sequence of strain ECM3 with publicly available sequences from NCBI GenBank. It showed 99.6% sequence identity with *Shewanella indica* type strain KJW27*,* 99.5% with *Shewanella algae* strain BrY (ATCC 51181)*,* 99.0% with *S. algae type* strain ATCC 51192, 98.7% sequence identity with *Shewanella upenei* strain 20-23*,* and 98.1% sequence identity with *Shewanella haliotis* type strain DW01 and *Shewanella chilikensis* type strain JC5 (Table D-12.2).
Tier 2	2. A BLAST search sequence analysis compared the gyrB gene sequence of strain ECM3 with publicly available sequences from NCBI GenBank. Strain ECM3 showed highest (99.9%) *gyr*B gene sequence identity with *S. indica* KJW27, 99.7% with *S. algae* BrY (ATCC 51181)*,* 95.2% with *S. haliotis* DW01, 94.7% with *S. chilikensis* JC5, and 93.1% with *S. algae* type strain ATCC 51192 (Table D-12.2).
Tier 3	3. a) Strain ECM3 colonies are smooth, opaque and pale tan in colour, 2.0 - 3.0 mm in diameter. Strain ECM3 cells are Gram-negative and rod-shaped with a single polar flagellum. b) Optimum growth was observed at 37°C, in 3% NaCl, and pH of 7.5. Growth was observed at a temperature range of 10 - 42°C and at NaCl concentrations of 2% and 5%. Aerobic and

	anaerobic growth was observed; strain ECM3 showed β-haemolysis on sheep blood agar. Further physiological tests were chosen such that the relatedness of strain ECM3 with *S. indica* KJW27, *S. algae* BrY (ATCC 51181) and *S. algae* ATCC 51192, the two closely-related species, can be determined (as suggested in Verma *et al.,* 2011 and Zhao *et al.,* 2007, Table D-12.3).
Analysis	Based on the sequence analyses of 16S rRNA and *gyr*B genes, strain ECM3 is closely related to two *Shewanella* strains, the type strain *S. indica* KJW27 and *S*. *algae* BrY (ATCC 51181). However, the *gyr*B gene of strain ECM3 had poor association with the type strain of *S. algae* ATCC 51192. Therefore, based on genetic data strain ECM3 was determined to be closely related to *S. indica* KJW27 and *S*. *algae* BrY (ATCC 51181), and less so to the *S. algae* type strain. The analysis of physiological and biochemical properties of these four strains likewise indicated that strain ECM3 is closely-related to the type strain of *S. indica*, as well as to strain BrY. It is possible that strain BrY is in fact a strain of *S. indica* and not of *S. algae*. Venkateswaran *et al.,*, (1999) had reported that although *S. algae* BrY falls within the *S. algae* clade based on 16S rRNA gene sequence analysis, it formed a separate branch for *gyr*B gene sequence.
Conclusion	Considering both gene sequence analysis and its phenotypic profile, strain ECM3 is accepted as a strain of the species *S. indica.* Nevertheless, when strain-specific information is not available, information from *S*. *algae* BrY (ATCC 51181) can also be considered as appropriate to fulfill specific information requirements.

Table D-12.2: comparison of strain ECM3 16S rRNA and *gyr*B gene sequences with *Shewanella* species

Sequence identity of *Shewanella Strain ECM3 to:*	For 16S rRNA gene	For gyrB gene
Shewanella indica KJW27 [T]	99.6%	99.9%
Shewanella algae BrY (ATCC 51181)	99.5%	99.7%
Shewanella algae ATCC 51192 [T]	99.0%	93.1%
Shewanella upenei 20-23 [T]	98.7%	94.2%
Shewanella chilikensis JC5 [T]	98.1%	94.7%
Shewanella haliotis DW01 [T]	98.1%	95.2%
Shewanella aquamarina SW-120 [T]	96.6%	91.6%
Shewanella loihica PV-4 [T]	96.3%	91.3%
Shewanella piezotolerans WP3 [T]	94.8%	No data
Shewanella literosedimensis SMK1-12 [T]	93.4%	No data

[T,] Type strains

Table D-12.3 : phenotypic properties of *Shewanella* strain ECM3 in comparison with *Shewanella indica* KJW27^{T}, *Shewanella algae* BrY (ATCC 51181) and *Shewanella algae* type strain ATCC 51192^{T}

Characteristics	strain ECM3[a]	*S. algae* BrY ATCC 51181[b]	*S. indica* KJW27 [c]	*S. algae* ATCC 5119 [d]
Growth at 4°C	-	-	-	-
Growth at 45°C	+	No data	+	-
Growth with 10 % NaCl (w/v)	+	+	+	-
Substrate Utilization:				
cis-Aconitic acid	+	No data	+	-
Citric acid	+	-	+	-
Formic acid	+	-	+	-
Pyruvic acid	+	+	No data	No data
Lactic acid	+	+	+	No data
Succinic acid	-	-	+	No data
Fumaric acid	-	-	No data	No data
a-Ketobutyric acid	-	-	-	+
a-Ketovaleric acid	-	No data	-	+
L-Asparagine	-	No data	-	+
L-Aspartic acid	-	No data	-	+
L-Histidine	-	No data	-	-
L-Serine	-	No data	-	+
L-Proline	+	No data	+	-

[a] Test results from studies conducted on strain ECM3;

[b] Data from Caccavo *et al.,* 1992, Venkateswaran *et al.,* 2011;

[c] Data from Verma *et al.,* 2011; [d] Data from Verma *et al.,* 2011;

Case study *Trichoderma reesei*

Table D-13.1: strain information and identification methodology for *Trichoderma reesei* strain Xyl123

Notified micro-organism designation	*Trichoderma reesei* strain Xyl123
Source	A mutant derivative of *T. reesei* RUT-C30 (parental strain) was obtained from a university scientist. The scientist had purchased RUT-C30 from ATCC (ATCC 56765). Further genetic modifications and selective breeding were undertaken to the parental strain to increase xylanase production and the strain was designated as *T. reesei* strain Xyl123.
Test methods	
Tier 1	1. Morphological, physiological and biochemical properties: colony and cell morphology and colour (Bissett, 1984, 1991 a,b,c, 1992; Samuels, 1996);
Tier 2	2. Genotypic methods: sequence analyses of the rRNA operon (ITS1 and ITS2 regions) and 420 bp of the translation elongation factor gene (*tef*1) were performed as suggested in the literature (Kuhls *et al.,* 1997; Druzhinina *et al.,* 2006; Hoyos-Carvajal *et al.,* 2009; Samuels and Ismaiel, 2009; Druzhinina *et al.,* 2012; Cummings *et al.,* 2016). Sequence analysis was performed using NCBI and TrichOKEY 2.0 databases.
Data	
Tier 1	1. Morphological characteristics of strain Xyl123 were verified on Difco's corn meal agar. Strain Xyl123 appeared smooth with white colonies, and possessed ellipsoidal to oblong conidia that are green in color.
Tier 2	2. Sequence analysis of the ITS1 and ITS2 regions of the rDNA (Kuhls, 1997; Kubecek *et al.*, 2003; Druzhinina *et al.*, 2006) using NCBI showed 99.9% similarity of strain Xyl123 with *T. reesei* QM 6a (100% coverage) and *Trichoderma parareesei* ATCC MYA-4777 (with 97% coverage) (Table D-13.2). Sequence analysis of the *tef*1 gene showed 100% identity of strain Xyl123 with *T. reesei* CBS 836.91, and less than 90% similarity to other species of *T. longibrachiatum* clade (Table D-13.3).

Analysis	The properties of strain Xyl123 are consistent with what has been described in the literature for the clade *Longibrachiatum.* This clade is characterized by sparsely branched conidiophores having a high proportion of solitary phialides. The absence of distinctive features unique to individual species and the presence of teleomorphs belonging to related genera (e.g., *Hypocrea*) complicate the taxonomy of *Trichoderma* leading to many misclassifications. For example, within the *Longibrachiatum* clade*, Trichoderma longibrachiatum* is morphologically synonymous with *T. reesei.* According to Druzhinina *et al.,* (2012), most species in the *Longibrachiatum* clade have smooth, ellipsoidal to oblong conidia except for *Trichoderma ghanense, Trichoderma saturnisporum* and *Trichoderma sp. TR 175* with tuberculate conidia. ITS sequencing could not differentiate strain Xyl123 from the members of clade *Longibrachiatum* of the *Trichoderma* genus, as the strain shares > 99% similarity with several species of this clade. However, sequencing of *tef*1 was able to clearly discriminate strain Xyl123 from closely-related species of the *Longibrachiatum* clade (Gazis *et al.,* 2011; Druzhnina *et al.,* 2012) and attribute strain Xyl123 to the species *T. reesei*. Sequence analyses of both ITS1/ITS2 and *tef*1 gene regions using TrichOKEY2.0 also showed a “high” identification reliability index for the species identification: *Hypocrea jecorina/Trichoderma reesei.*
Conclusion	Based on morphological data and sequence analysis (using ITS1/2 and *tef*1), strain Xyl123 was confirmed as *T. reesei.*

Table D-13.2: comparison of strain Xyl123 ITS1/2 regions with *Trichoderma* species using NCBI BLAST limited to type strain sequences

Descriptions	Query coverage %	E value	% identity
T. reesei QM 6a	100	0.0	99.9
T. parareesei ATCC MYA-4777	97	0.0	100
T. saturnisporum ATCC 18903	100	0.0	99.3
T. citrinoviridae strain CBS 258	100	0.0	99.3
T. longibrachiatum ATCC 18648	100	0.0	99.2
Hypocrea orientalis GJS 88-81	99	0.0	98.9
T. pseudokoningii DAOM 167678	99	0.0	98.8
T. ghanense ATCC 208858	100	0.0	98.2
T. novae-zelandiae CBS 639-92	94	0.0	98.0

Table D-13.3: comparison of strain Xyl123 *tef1* gene with *Trichoderma* species using NCBI BLAST limited to type strain sequences

Descriptions	Query coverage %	E value	% identity
T. reesei CBS 836.91	100	0.0	100
T. reesei DAOM 167654	100	0.0	99.6
T. parareesei TUB F-1066	99	0.0	87.2
T. saturnisporum ATCC 18903	99	0.0	84.5

Appendix E – References

Anders, H.-J., M. Keller, M. Berchtold, and W. Hecker. 2007. Polyphasic approach to microbial identification. Amer. Pharm. Rev. Published October 1, 2007.

Anderson, I.C. and J.W.G Cairney. 2004. Diversity and ecology of soil fungal communities: increased understanding through the application of molecular techniques. Environ. Microbiol. 6:769-779.

Andreotti, P. E., G. V. Ludwig, A. Harwood Peruski, J. J. Tuite, S. S. Morse, and L. F. Peruski. 2003. Immunoassay of infectious agents. Biotech. 35:859.

Arenas-Hernandez, M. M. P., Y. Martinez-Laguna, and A. G. Torres. 2012. Clinical implications of enteroadherent *Escherichia coli*. Curr. Gastroenterol. Rep. 14:386-394.

Azevedo, H., F.M. Lopes, P.R. Silla, and M. Hungria. 2015. A database for the taxonomic and phylogenetic identification of the genus Bradyrhizobium using multilocus sequence analysis. BMC Genomics 16;S10.

Barnett, J. A. 2000. Commentary on a new major reference work Yeasts: Characteristics and Identification, 3rd edn. By JA Barnett, RW Payne and D. Yarrow. Cambridge University Press: Cambridge.£ 200.00, ix+ 1139 pp., 1376 photomicrographs by Linda Barnett. ISBN 0 521 57396 3. Published 25 May 2000. Yeast, 16: 1445–1447.

Battista, J.R. and Rainey, F.A. (2012) Genus I. *Deinococcus*, pages 396-403, in Bergey's Manual of Systematic Bacteriology, 2nd ed., Vol. 1. The Archaea and the Deeply Branching and Phototrophic Bacteria. Boone, D.R. and Casternholz, R.W. (eds. For Vol. 1). Springer, New York. 721 p

Bergey's Manual of Systematics of Archaea and Bacteria. 2015. Whitman, W.B. (ed) http://onlinelibrary.wiley.com/book/10.1002/9781118960608

Bissett, J., Gams, W., Jaklitsch, W., and G.J. Samuels. 2015. Accepted *Trichoderma* names in the year 2015. IMA Fungus 6: 263.

Bissett, J. 1992. *Trichoderma atroviride.* Can. J. Botany 70: 639-641.

Bissett, J. 1991a. A revision of the genus *Trichoderma*. II. Infrageneric classification. Can. J. Bot. 69: 2357-2372.

Bissett, J. 1991b. A revision of the genus *Trichoderma*. III. Section *Pachybasium*. Can. J. Bot. 69: 2373-2417.

Bissett, J. 1991c. A revision of the genus *Trichoderma*. IV. Additional notes on section *Longibrachiatum*. Can. J. Bot. 69: 2418-2420.

Bissett, J. 1984. A revision of the genus *Trichoderma*. I. Section *Longibrachiatum* sect. nov. Can. J. Bot. 62: 924-931.

Busse, H. J., E. B. M. Denner, and W. Lubitz. 1996. Classification and identification of bacteria: Current approaches to an old problem. Overview of methods used in bacterial systematics. J. Biotechnol. 47:3-38.

Caccavo, F., R.P. Blakemore, and D.R. Lovley. 1992. A hydrogen-oxidizing, Fe (III)-reducing microorganism from the Great Bay Estuary, New Hampshire. Appl. Environ. Microbiol. 58: 3211-3216.

Cecchini, F., L. Fajs, S. Cosnier, and R.S. Marks. 2016. Vibrio cholerae detection: Traditional assays, novel diagnostic techniques and biosensors. TrAC Trends Anal. Chem. 79:199-209.

Cheng, K., H. Chui, L. Domish, D. Hernandez, and G. Wang . 2016. Recent development of mass spectrometry and proteomics applications in identification and typing of bacteria. Prot. Clin. Appl. 10:346-357.

Clarke, S. C. 2005. Pyrosequencing: Nucleotide sequencing technology with bacterial genotyping applications. Expert Rev. Mol. Diag. 5:947-953.

Collado, L., A. Levican, J. Perez, and M. J. Figueras. 2011. Arcobacter defluvii sp. nov., isolated from sewage samples. Int. J. Syst. Evol. Microbiol. 61: 2155-2161.

Collins, M. D. and D. Jones. 1981. Distribution of isoprenoid quinone structural types in bacteria and their taxonomic implications. Microbiol. Rev. 45:316-354.

Colwell, R. R. and R. Grigorova. 1987. Current methods for classification and identification of microorganisms. *In* Methods in microbiology, London : Academic Press.

Cummings, N. J., Ambrose, A., Braithwaite, M., Bissett, J., Roslan, H. A., Abdullah, J., Stewart, A., Agbayani, F.V., Steyaert, J. and R.A. Hill. 2016. Diversity of root-endophytic *Trichoderma* from Malaysian Borneo. Mycol. Prog. 15: 1-14.

Da Costa, M. S., L. Albuquerque, M. F. Nobre, and R. Wait. 2011a. The identification of fatty acids in bacteria. Methods Microbiol. 38:183-196.

Da Costa, M. S., L. Albuquerque, M. F. Nobre, and R. Wait. 2011b. The Extraction and Identification of Respiratory Lipoquinones of Prokaryotes and Their Use in Taxonomy. Methods Microbiol. 38:197-206.

Das, S., H.R. Dash, N. Mangwani, J. Chakraborty, and S. Kumari. 2014. Understanding molecular identification and polyphasic taxonomic approaches for genetic relatedness and phylogenetic relationships of microorganisms. J Microbiol Meth 103:80-100.

Debourgogne, A., C. Gueidan, S. de Hoog, A. Lozniewski, and M. Machouart. 2012. Comparison of two DNA sequence-based typing schemes for the Fusarium solani Species Complex and proposal of a new consensus method. J Microbiol Meth 91:65-72.

Diggle, M. A. and S. C. Clarke. 2004. Pyrosequencing™: Sequence typing at the speed of light. Appl. Biochem. Biotechnol. - Part B Mol. Biotech. 28:129-138.

Druzhinina, I. S., A. G. Kopchinskiy, and C.P. Kubicek. 2006. The first 100 *Trichoderma* species characterized by molecular data. Mycosci. 47: 55-64.

Druzhinina, I. S., M. Komon-Zelazowska, A. Ismaiel, W. Jaklitsch, T. Mullaw, G. J. Samuels, and C. P. Kubicek. 2012. Molecular phylogeny and species delimitation in the section Longibrachiatum of Trichoderma. Fungal Genetics and Biology 49:358-368.Fan, L., J.H.L. Hui, Z.G. Yu, and K.H. Chu. 2014. VIP Barcoding: composition vector-based software for rapid species identification based on DNA barcoding. Mol. Ecol. Resour. 14:871-881.

Frisvad, J.C., T.O. Larsen, U. Thrane, M. Meijer, J. Varga, R.A. Samson, and K.F. Nielsen. 2011. Fumonisin and ochratoxin production in industrial *Aspergillus niger* strains. PLoS ONE *6:* e23496.

Gasanov, U., Hughes,D. and P. M. Hansbro. 2005. Methods for the isolation and identification of Listeria spp. and Listeria monocytogenes: A review. FEMS Microbiol. Rev. 29:851-875.

Gazis, R., S. Rehner, and P. Chaverri. 2011. Species delimitation in fungal endophyte diversity studies and its implications in ecological and biogeographic inferences. Mol. Ecol. 20: 3001-3013.

Gillis, M., P. Vandamme, P. De Vos, J. Swings, and K. Kersters. 2005. Polyphasic Taxonomy, p. 43-48. *In* D. J. Brenner, N. R. Krieg, J. T. Staley, and G. M. Garrity (eds.), Bergey's Manual of Systematic Bacteriology: Volume Two: The Proteobacteria, Part A Introductory Essays. Springer US, Boston, MA.

Gram Strain. 2010. Staining Procedures, p. 91-123. *In* L. S. Garcia (ed.), Clinical Microbiology Procedures Handbook. American Society for Microbiology.

Hanson, B., Y. Zhou, E.J. Bautista, *et al.,* 2016. Characterization of the bacterial and fungal microbiome in indoor dust and outdoor air samples: a pilot study. Environ. Sci.: Proc. Impacts 18:713-724

Hill, J. E., S. L. Penny, K. G. Crowell, S. H. Goh, and S. M. Hemmingsen. 2004. cpnDB: A Chaperonin Sequence Database. Genome Res. 14:1669-1675.

Hill, J. E., A. Paccagnella, K. Law, P. L. Melito, D. L. Woodward, L. Price, A. H. Leung, S. M. Hemmingsen, and S. H. Goh. 2006. Identification of Campylobacter spp. and discrimination from Helicobacter and Arcobacter spp. by direct sequencing of PCR-amplified cpn60 sequences and comparison to cpnDB, a chaperonin reference sequence database. J. Med. Microbiol. 55:393-399.

Hiraishi, A., K. Inagaki, Y. Tanimoto, M. Iwasaki, N. Kishimoto, H. Tanaka. 1997. Pylogenetic characterization of a new thermoacidophilic bacterium isolated from hot springs in Japan. J. Gen. Appl. Microbiol. 43:295-304.

Hirsch, P.R., T.H. Mauchline, and I.M. Clark. 2010. Culture-independent molecular techniques for soil microbial ecology. Soil Biol. Biochem. 42:878-887.

Hirsch, P.R., T.H. Mauchline, and I.M. Clark. 2013. Culture-independent molecular approaches to microbial ecology in the soil and the rhizosphere. In FJ de Bruijn ed, Molecular Microbial Ecology of the Rhizosphere, Volume 1 pp. 45-55. Rothamsted Research, UK: John Wiley & Sons.

Hoyos-Carvajal, L., S. Orduz, and J. Bissett. 2009. Genetic and metabolic biodiversity of *Trichoderma* from Colombia and adjacent neotropic regions. Fungal Genet. Biol. 46: 615-631.

Hong, S., Lee, M., Kim, D., Varga, J., Frisvad, J.C., Perrone, G., Gomi, K., Yamada, O., Machida, M., Machida, M., and R.A. Samson. 2013. *Aspergillus luchuensis*, an Industrially Important Black *Aspergillus* in East Asia. PLoS ONE *8:* e63769.

Hwang, C. Y., I. Lee, Y. Cho, Y. M. Lee, K. Baek, Y.-J. Jung, Y. Y. Yang, T. Lee, T. S. Rhee, and H. K. Lee. 2015. Rhodococcus aerolatus sp. nov., isolated from subarctic rainwater. Int. J. Syst. Evol. Microbiol. 65:465-471.

Jackman, J. 2012. The Microbe: The Basics of Structure, Morphology, and Physiology as They Relate to Microbial Characterization and Attribution, p. 13-34. *In* J. B. Cliff, H. W. Kreuzer, C. J. Ehrhardt, and D. S. Wunschel (eds.), Chemical and Physical Signatures for Microbial Forensics. Springer New York, New York, NY.

Janda, J. M. and S. L. Abbott. 2002. Bacterial identification for publication: When is enough enough? J. Clin. Microbiol. 40:1887-1891.

Janke, T., K. Schwaiger, M. Ege, C. Fahn, E. von Mutius, J. Bauer, and M. Mayer. 2013. Analysis of the Fungal Flora in Environmental Dust Samples by PCR–SSCP Method. Curr. Microbiol. 67:156-169.

Jeyaram, K., Romi, W., Singh, T.A., Adewumi, G.A., Basanti, K., and F.A. Oguntoyinbo. 2011. Distinct differentiation of closely related species of *Bacillus subtilis* group with industrial importance. J. Microbiol. Methods 87: 161-164.

Jones, A.L. and M. Goodfellow. 2012. Genus IV. *Rhodococcus*, pages 437-464, *in* Bergey's Manual of Systematic Bacteriology Vol. 5. The Actinobacteria. Whitman, E. *et al.,* (eds.). Springer-Verlag, New York. 2083 p.

Jones, D. and N. R. Krieg. 1984. Bacterial Classification V: Serology and Chemotaxonomy *In* N. R. Krieg (ed.), Bergey's Manual of Systematic Bacteriology Vol. 1. Williams & Wilkins.

Karger, A. 2016. Current developments to use linear MALDI-TOF spectra for the identification and typing of bacteria and the characterization of other cells/organisms related to infectious diseases. Prot. Clin. Appl. 10:982-993.

Kim, M., Oh, H.-S., S.-C. Park, and J. Chun. 2014. Towards a taxonomic coherence between average nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of prokaryotes. Int. J. Syst. Evol. Microbiol. 64: 346–351.

Konstantinidis, K. T., A. Ramette, and J. M. Tiedje. 2006. The bacterial species definition in the genomic era. Philosophical Transactions of the Royal Society B: Biol. Sci. 361:1929-1940.

Kubicek, C. P., J. Bissett, I. Druzhinina, C. Kullnig-Gradinger, and G. Szakacs. 2003. Genetic and metabolic diversity of *Trichoderma*: a case study on South-East Asian isolates. Fungal Genet. Biol. 38: 310-319.

Kuhls, K., E. Lieckfeldt, G. J. Samuels, W. Meyer, C. P. Kubicek, and T. Börner. 1997. Revision of *Trichoderma* sect. *Longibrachiatum* including related teleomorphs based on analysis of ribosomal DNA internal transcribed spacer sequences. Mycol. 89: 442-460.

Kunitsky, C., G. Osterhout, and M. Sasser. 2006. Identification of microorganisms using fatty acid methyl ester (FAME) analysis and the MIDI Sherlock® Microbial Identification System *In* M. Miller (ed.), Encyclopedia of rapid microbiological methods, Volume 3.

Kurosawa, N. and Y.H. Itoh. 1993 Nucleotide sequence of the 16S rRNA gene from thermoacidophilic archea Sulfolobus acidocaldarius ATCC909. Nuc. Acids Res. 21:357.

Kurtzman, C. P., J.W. Fell, and T. Boekhout. 2011. The Yeasts, a Taxonomic Study, Fifth Edition.

Kurtzman, C.P. 2009. Biotechnological strains of *Komagataella* (*Pichia*) *pastoris* are *Komagataella phaffii* as determined from multigene sequence analysis. J. Ind. Microbiol. Biotechnol. 36: 1435-1438.

Kurtzman, C.P., 2005. Description of *Komagataella phaffii* sp. nov. and the transfer of *Pichia pseudopastoris* to the methylotrophic yeast genus *Komagataella*. Int. J. Syst. Evol. Microbiol. 55: 973-976.

Larsen, M.V., S. Cosentino, S. Rasmussen, C. Friis, H. Hasman, R.L. Marvig, L. Jelsbak, T. Sicheritz-Pontén, D.W. Ussery, F.M. Aarestrup, and O. Lunda. 2012. Multilocus Sequence Typing of Total-Genome-Sequenced Bacteria. J. Clin. Microbiol. 50:1355-1361.

Levican, A., L. Collado, C. Aguilar, C. Yustes, A. L. Dieguez, J. L. Romalde, and M. J. Figueras. 2012. Arcobacter bivalviorum sp. nov. and Arcobacter venerupis sp. nov., new species isolated from shellfish. Syst. Appl. Microbiol. 35: 133-138.

Links, M. G., T. J. Dumonceaux, S. M. Hemmingsen, and J. E. Hill. 2012. The Chaperonin-60 Universal Target Is a Barcode for Bacteria That Enables De Novo Assembly of Metagenomic Sequence Data. PLoS ONE 7:e49755.

Liti, G., D.B.H. Barton, and E. J. Louis. 2006. Sequence diversity, reproductive isolation and species concepts in saccharomyces. Genet. 174:839-850.

Macheras, E., A.L. Roux, S. Bastian, *et al.,* 2011. Multilocus sequence analysis and rpoB sequencing of Mycobacterium abscessus sensu lato strains. J. Clin. Microbiol. 49:491-499.

Maiden, M.C.J., J.A. Bygraves, E. Feill, G. Morelli, J.E. Russell, R. Urwin, Q. Zhang, K. Zhou, K. Zurth, D.A. Caugant, I.M. Feavers, M. Achtmans, and B.G. Spratt. 1998. P. Natl. A. Sci. USA 95:3140-3145.

Manikan, V., M.Y.M. Nazir, M.S. Kalil, M.H.M. Isa, A.J.A. Kader, W.M.W. Yusoff, and A.A. Hamid. 2015. A new strain of docosahexaenoic acid producing microalga from Malaysian coastal waters. Algal Res. 9: 40-47.

McClelland, R. 2001. Gram's stain: the key to microbiology. MLO: medical laboratory observer 33:20-28.

Nielsen, M. L., J. B. Nielsen, C. Rank, M. L. Klejnstrup, D. K. Holm, K. H. Brogaard, B. G. Hansen, J. C. Frisvad, T. O. Larsen, and U. H. Mortensen. 2011. A genome-wide polyketide synthase deletion library uncovers novel genetic links to polyketides and meroterpenoids in Aspergillus nidulans. FEMS Microbiol. Lett. 321:157-166.

OECD (Organisation for Economic Co-operation and Development). Guidance document on the use of taxonomy in risk assessment of micro-organisms: Bacteria. ENV/JM/MONO(2003)13. 19-8-2003.

Olive, D.M. and P. Bean. 1999. Principles and applications of methods for DNA-based typing of microbial organisms. J. Clin. Microbiol. 37:1661-1669.

Oskam, L., E. Slim, and S. Buhrer-Sekula. 2003. Serology: Recent developments, strengths, limitations and prospects: A state of the art overview. Lep. Rev. 74:196-205.

Parlapani, F.F. and I.S. Boziaris. 2016. Monitoring of spoilage and determination of microbial communities based on 16S rRNA gene sequence analysis of whole sea bream stored at various temperatures. LWT – Food. Sci. Technol. 66:553-559.

Pascual, J., M.C. Macián, D.R. Arahal, E. Garay, and M.J. Pujalte. 2010. Multilocus sequence analysis of the central clade of the genus Vibrio by using the 16S rRNA, recA, pyrH, rpoD, gyrB, rctB and toxR genes. Int. J. Syst. Evol. Microbiol. 60:154-165.

Perrone, G., G. Stea, F. Epifani, J. Varga, J.C. Frisvad, and R.A. Samson. 2011. *Aspergillus niger* contains the cryptic phylogenetic species *A. awamori*. Fungal Biol. *115:* 1138-1150.

Purcaro, G., P. Q. Tranchida, P. Dugo, E. La Camera, G. Bisignano, L. Conte, and L. Mondello. 2010. Characterization of bacterial lipid profiles by using rapid sample preparation and fast comprehensive two-dimensional gas chromatography in combination with mass spectrometry. J. Sep. Sci. 33:2334-2340.

Raengpradub, S. 2009 Microbial Identification: Tracking the Great Unknown with Innovative and Advanced Technologies. J. Lab. Automat. 14:232-234.

Rogers, G.N., T.J. Pritchett, J.L. Lane, and J.C. Paulson. 1983. Differential sensitivity of human, avian, and equine influenza a viruses to a glycoprotein inhibitor of infection: Selection of receptor specific variants. Virol. 131: 394-408.

Ronholm, J., N. Nasheri, N. Petronella, and F. Pagotto. 2016. Navigating Microbiological Food Safety in the Era of Whole-Genome Sequencing. Clin. Microbiol. Rev. 29:837-857.

Ruimy, R., V. Breittmayer, P. Elbaze, B. Lafay, O. Boussemart, M. Gauthier, and R. Christen. 1994. Phylogenetic analysis and assessment of the genera *Vibrio*, *Photobacterium*, *Aeromonas*, and *Plesiomonas* deduced from small-subunit rRNA sequences. Int. J. Syst. Evol. Microbiol. 44: 416-426.

Salvetti, E., L. Orru, V. Capozzi, A. Martina, A. Lamontanara, D. Keller, H. Cash, G. E. Felis, L. Cattivelli, S. Torriani, and G. Spano . 2016. Integrate genome-based assessment of safety for

probiotic strains: Bacillus coagulans GBI-30, 6086 as a case study. Appl. Microbiol. Biotechnol. 100:4595-4605.

Samson, R. A., C.M. Visagie, J. Houbraken, S.B. Hong, V. Hubka, C.H.W. Klaassen, G. Perrone, K. A. Seifert, A. Susca, J. B. Tanney, J. Varga, S. Kocsube, G. Szigeti, T. Yaguchi, and J. C. Frisvad. 2014. Phylogeny, identification and nomenclature of the genus Aspergillus. Stud. Mycol. 78:141-173.

Samson, R.A., P. Noonim, M. Meijer, J. Houbraken, J.C. Frisvad, and J. Varga. 2007a. Diagnostic tools to identify black aspergilli. Stud. Mycol. 59: 129-145.

Samson, R.A., J. Varga, S.M. Witiak, and D.M. Geiser. 2007b. The species concept in *Aspergillus*: Recommendations of an international panel. Stud. Mycol. 59: 71-73.

Samuels, G. J. 1996. *Trichoderma*: a review of biology and systematics of the genus. Mycol. Res. 100: 923-935.

Samuels, G. J., and A. Ismaiel. 2009. *Trichoderma evansii* and *T. lieckfeldtiae*: two new *T. hamatum*-like species. Mycol. 101: 142-156.

Saputra, D., S. Rasmussen, M. V. Larsen, N. Haddad, M. M. Sperotto, F. M. Aarestrup, O. Lund, and T. Sicheritz-Ponten. 2015. Reads2Type: a web application for rapid microbial taxonomy identification. BMC Bioinform. 16:1-10.

Schatz, M.C., J. Witkowski, and W.R. McCombie. 2012 Current challenges in de novo plant genome sequencing and assembly. Genome Biol. 13:243,

Schoch, C.L., K.A. Seifert, S. Huhndorf, V. Robert, J.L. Spouge, C.A. Levesque, W. Chen, and Fungal Barcoding Consortium. 2012. Nuclear ribosomal internal transcribed spacer ITS region as a universal DNA barcode marker for Fungi. PNAS 109:6241-6246.

Schumann, P. 2011. Peptidoglycan Structure. Methods Microbiol. 38 ;, 101-129.

Singhal, N., M. Kumar, P. K. Kanaujia, and J. S. Virdi. 2015. MALDI-TOF mass spectrometry: An emerging technology for microbial identification and diagnosis. Front. Microbiol. 6.

Siqueira, J., A. F. Fouad, and I. N. Rocas. 2012. Pyrosequencing as a tool for better understanding of human microbiomes. J. Oral Microbiol. 4.

Spiegelman, D., G. Whissell, and C. W. Greer. 2005. A survey of the methods for the characterization of microbial consortia and communities. Can. J. Microbiol. 51:355-386.

Srinivasan, R., U. Karaoz, M. Volegova, J. MacKichan, M. Kato-Maeda, S. Miller, R. Nadarajan, E.L. Brodie, and S.V. Lynch. 2015. Use of 16S rRNA Gene for Identification of a Broad Range of Clinically Relevant Bacterial Pathogens. PLoS ONE 10:e0117617.

Stackebrandt, E. 2011. Report of the *ad hoc* committee for the re-evaluation of the species definition in bacteriology, p. 99-104. *In* F. J. de Bruijn (ed.), Handbook of Molecular Microbial Ecology I: Metagenom. Compl. Approach.

Stackebrandt, E., D. Smith, S. Casaregola, G. C. Varese, G. Verkleij, N. Lima, and P. Bridge. 2014. Deposit of microbial strains in public service collections as part of the publication process to underpin good practice in science. SpringerPlus 3:1-4.

Stanier, R. Y., D. Wachter, C. Gasser, and A. C. Wilson. 1970. Comparative immunological studies of two Pseudomonas enzymes. J. Bacteriol. 102:351-362.

Suarez, S., X. Nassif, and A. Ferroni. 2015. Applications of MALDI-TOF technology in clinical microbiology. Pathol. Biol. 63:43-52.

Sutton, S. 2006. The Gram Stain. Pharm. Microbiol. Forum 12:4-8.

Tamura, K., D. Peterson, N. Peterson, G. Stecher, M. Nei, and S. Kumar. 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol. Biol. Evol. 28: 2731-2739.

Thompson FL, Gevers D, Thompson CC, Dawyndt P, Naser S, Hoste B, et al. 2005. Phylogeny and molecular identification of vibrios on the basis of multilocus sequence analysis. Appl Environ Microbiol. 71: 5107-5115.

Thomsen, M. C. F., J. Ahrenfeldt, J. L. B. Cisneros, V. Jurtz, M. V. Larsen, H. Hasman, F. M. Aarestrup, and O. Lund. 2016. A Bacterial Analysis Platform: An integrated system for analysing bacterial whole genome sequencing data for clinical diagnostics and surveillance. PLoS ONE 11:e0157718.

Tindall, B. J., R. Rossello-Mora, H. J. Busse, W. Ludwig, and P. Kampfer. 2010. Notes on the characterization of prokaryote strains for taxonomic purposes. Int. J. Syst. Evol. Microbiol. 60:249-266.

Touzet, N., J. M. Franco, and R. Raine. 2007. Characterization of Nontoxic and Toxin-Producing Strains of Alexandrium minutum (Dinophyceae) in Irish Coastal Waters. Appl. Env. Microbiol. 73:3333-3342.

Tshikhudo, P., R. Nnzeru, K. Ntushelo, and F. Mudau. 2013. Bacterial species identification getting easier. Afri. J. Biotechnol. 12:5975-5982.

UK Standards for Microbiology Investigations. 2016. Identification of aerobic actinomycetes. Public Health England. Issue 2.1. 28 p (PDF).

(https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/509336/ID_10i2.1.pdf)

Varga, J., J.C. Frisvad, S. Kocsubé, B. Brankovics, B. Tóth, G. Szigeti, and R.A. Samson. 2011. New and revisited species in *Aspergillus* section *Nigri*. Stud. Mycol. 69: 1-17.

Vaughan-Martini, A., and A. Martini. 2011. Chapter 61 - Saccharomyces Meyen ex Reess (1870), pp. 733-746. *In* The Yeasts Taxonomy (Fifth Edition), Kurtzman, C.P., Fell, J.W. and T. Boekhout (eds.). Elsevier, London. 2354 p.

Venkateswaran, K., D.P. Moser, M.E. Dollhopf, D.P. Lies, D.A. Saffarini, B.J. MacGregor, D.B. Ringelberg, D.C. White, M. Nishijima, H. Sano, and J. Burghardt. 1999. Polyphasic taxonomy of the genus *Shewanella* and description of *Shewanella oneidensis* sp. nov Int. J. Syst. Evol. Microbiol. 49: 705-724.

Verma, P., P.K. Pandey, A.K. Gupta, H.J. Kim, K.S. Baik, C.N. Seong, M.S. Patole, and Y.S. Shouche. 2011. *Shewanella indica* sp. nov., isolated from sediment of the Arabian Sea. Int. J. Syst. Evol. Microbiol. 61: 2058-2064.

Wayne, L. G., D. J. Brenner, R. R. Colwell, P. A. D. Grimont, O. Kandler, M. I. Krichevsky, L. H. Moore, W. E. C. Moore, R. G. E. Murray, E. Stackebrandt, M. P. Starr, and H. G. Truper. 1987. Report of the Ad Hoc Committee on Reconciliation of Approaches to Bacterial Systematics. Int. J. Syst. Evol. Microbiol. 37:463-464.

Wunschel, D. S., A. Fox, and K. L. Wahl. 2012. Glycoprotein and Protein Markers for Strain Differentiation and Growth Environment or Media Attribution, p. 61-70. *In* J. B. Cliff, H. W. Kreuzer, C. J. Ehrhardt, and D. S. Wunschel (eds.), Chemical and Physical Signatures for Microbial Forensics. Springer New York, New York, NY.

Yamamoto, S. and S. Harayama. 1995. PCR amplification and direct sequencing of *gyrB* genes with universal primers and their application to the detection and taxonomic analysis of *Pseudomonas putida* strains. Appl. Environ. Microbiol. 61: 1104-1109.

Yokoyama, R. and D. Honda. 2007. Taxonomic rearrangement of the genus *Schizochytrium* sensu lato based on morphology, chemotaxonomic characteristics, and 18S rRNA gene phylogeny (Thraustochytriaceae, Labyrinthulomycetes): emendation for *Schizochytrium* and erection of *Aurantiochytrium* and *Oblongichytrium* gen. nov. Mycosci. 48: 199-211.

Zeaiter, Z., P.-E. Fournier, H. Ogata, and D. Raoult. 2002. Phylogenetic classification of Bartonella species by comparing groEL sequences. Int. J. Syst. Evol. Microbiol. 52:165-171.

Zeigler, D. R. 2003. Gene sequences useful for predicting relatedness of whole genomes in bacteria. Int. J. Syst. Evol. Microbiol. 53:1893-1900.

Zhou, D. and R. H. White. 1989. Biosynthesis of caldariellaquinone in *Sulfolobus* spp. J. Bacteriol. 171:6610-6616.