Organisation de Coopération et de Développement Economiques
Organisation for Economic Co-operation and Development

19-Aug-2003

English - Or. English

ENVIRONMENT DIRECTORATE
JOINT MEETING OF THE CHEMICALS COMMITTEE AND THE WORKING PARTY ON CHEMICALS, PESTICIDES AND BIOTECHNOLOGY

ENV/JM/MONO(2003)13
Unclassified

Series on Harmonisation of Regulatory Oversight in Biotechnology, No. 29

GUIDANCE DOCUMENT ON THE USE OF TAXONOMY IN RISK ASSESSMENT OF MICRO-ORGANISMS: BACTERIA

JT00148130

English - Or. English

[

Also published in the Series on Harmonisation of Regulatory Oversight in Biotechnology:

No. 4, *Industrial Products of Modern Biotechnology Intended for Release to the Environment: The Proceedings of the Fribourg Workshop* (1996)

No. 5, *Consensus Document on General Information concerning the Biosafety of Crop Plants Made Virus Resistant through Coat Protein Gene-Mediated Protection* (1996)

No. 6, *Consensus Document on Information Used in the Assessment of Environmental Applications Involving* Pseudomonas (1997)

No. 7, *Consensus Document on the Biology of* Brassica napus *L.* (*Oilseed Rape)* (1997)

No. 8, *Consensus Document on the Biology of* Solanum tuberosum *subsp.* tuberosum *(Potato)* (1997)

No. 9, *Consensus Document on the Biology of* Triticum aestivum *(Bread Wheat)* (1999)

No. 10, *Consensus Document on General Information Concerning the Genes and Their Enzymes that Confer Tolerance to Glyphosate Herbicide* (1999)

No. 11, *Consensus Document on General Information Concerning the Genes and Their Enzymes that Confer Tolerance to Phosphinothricin Herbicide* (1999)

No. 12, *Consensus Document on the Biology of* Picea abies *(L.) Karst (Norway Spruce*) (1999)

No. 13, *Consensus Document on the Biology of* Picea glauca *(Moench) Voss (White Spruce)* (1999)

No. 14, *Consensus Document on the Biology of* Oryza sativa *(Rice)* (1999)

No. 15, *Consensus Document on the Biology of* Glycine max *(L.) Merr. (Soybean)* (2000)

No. 16, *Consensus Document on the Biology of* Populus *L. (Poplars)* (2000)

No. 17, *Report of the OECD Workshop on Unique Identification Systems for Transgenic Plants, Charmey, Switzerland, 2-4 October 2000* (2001)

No. 18, *Consensus Document on the Biology of* Beta vulgaris *L. (Sugar Beet)* (2001)

No. 19, *Report of the Workshop on the Environmental Considerations of Genetically Modified Trees, Norway, September 1999.* (2001)

No. 20, *Consensus Document on Information used in the Assessment of Environmnetal Applications Involving Baculovirus* (2002)

No. 21, *Consensus Document on the Biology of* Picea sitchensis *(Bong.) Carr. (Sitka Spruce)* (2002)

No. 22, *Consensus Document on the Biology of* Pinus strobus *L. (Eastern White Pine)* (2002)

No. 23, *OECD Guidance for the Designation of a Unique Identifier for Transgenic Plants* (2002)

No. 24, *Consensus Document on the Biology of* Prunus *sp. (Stone Fruits)* (2002)

No. 25, *Module II: Herbicide Biochemistry, Herbicide Metabolism and the Residues in Glufosinate-Ammonium (Phosphinothricin)-Tolerant Transgenic Plants* (2002)

No. 26, *Output on the Questionnaire on National Approaches to Monitoring / Detection / Identification of Transgenic Products* (2003)

No.27,*Consensus Document on the Biology of* Zea mays *subsp.*mays *(Maize)* (2003)

No. 28, *Consensus Document on the Biology of European White Birch (*Betula pendula *Roth)* (2003)

OECD Environment, Health and Safety Publications

Series on Harmonisation of Regulatory Oversight in Biotechnology

No.29

Guidance Document on the Use of Taxonomy in Risk Assessment of Micro-organisms: Bacteria

Environment Directorate

Organisation for Economic Co-operation and Development

Paris 2003

ABOUT THE OECD

The Organisation for Economic Co-operation and Development (OECD) is an intergovernmental organisation in which representatives of 30 industrialised countries in North America, Europe and the Pacific, as well as the European Commission, meet to co-ordinate and harmonise policies, discuss issues of mutual concern, and work together to respond to international problems. Most of the OECD's work is carried out by more than 200 specialised Committees and subsidiary groups composed of Member country delegates. Observers from several countries with special status at the OECD, and from interested international organisations, attend many of the OECD's Workshops and other meetings. Committees and subsidiary groups are served by the OECD Secretariat, located in Paris, France, which is organised into Directorates and Divisions.

The Environment, Health and Safety Division publishes free-of-charge documents in nine different series: Testing and Assessment; Good Laboratory Practice and Compliance Monitoring; Pesticides; Risk Management; Harmonisation of Regulatory Oversight in Biotechnology; Safety of Novel Foods and Feeds; Chemical Accidents; Pollutant Release and Transfer Registers; and Emission Scenario Documents. More information about the Environment, Health and Safety Programme and EHS publications is available on the OECD's World Wide Web site (http://www.oecd.org/ehs/).

This publication is available electronically, at no charge.

For the complete text of this and many other Biosafety publications, consult the OECD's World Wide Web site (http://www.oecd.org/biotrack/)

or contact:

OECD Environment Directorate,
Environment, Health and Safety Division

2 rue André-Pascal
75775 Paris Cedex 16
France

Fax: (33-1) 45 24 16 75

E-mail: ehscont@oecd.org

FOREWORD

The OECD's Working[1] Group on Harmonisation of Regulatory Oversight in Biotechnology decided at its first session, in June 1995, to focus its work on the development of ***consensus documents*** which are mutually acceptable among Member countries. These consensus documents contain information for use during the regulatory assessment of a particular product.

On reviewing a published consensus document and drafting other consensus documents on micro-organisms, the Working Group felt that these documents did not focus in a straightforward way on questions that are relevant to risk/safety assessment issues. Responding to the concern, the Working Group decided to take an alternative approach, namely the development of ***guidance documents***, in the micro-organisms area. Guidance documents are intended to provide guidance on specific topics and issues, such as taxonomy and detection techniques that are relevant to risk/safety assessment in biotechnology.

This guidance document addresses the use of microbial taxonomy in assigning or confirming the identity of a subject micro-organism. It is primarily intended for use by risk assessors, but it may also be useful for applicants and other stakeholders in the regulatory process.

Canada and the United States served as lead countries in the preparation of this document. It has been revised on a number of occasions based on the input from other member countries.

The Joint Meeting of the Chemicals Committee and the Working Party on Chemicals, Pesticides and Biotechnology subsequently recommended that this document be made available to the public.

[1] In August 1998, following a decision by OECD Council to rationalise the names of Committees and Working Groups across the OECD, the name of the "Expert Group on Harmonisation of Regulatory Oversight in Biotechnology" became the "Working Group on Harmonisation of Regulatory Oversight in Biotechnology".

TABLE OF CONTENTS

AN INTRODUCTION TO OECD'S WORKING GROUP

This document was prepared by the Working Group on Harmonisation of Regulatory Oversight in Biotechnology. The group is comprised of individuals from government ministries or agencies, which have responsibility for the environmental risk/safety assessment of products of modern biotechnology (including genetically modified organisms).

Regulatory harmonisation has been the primary goal of the Working Group since it was established in 1995. It is the attempt to ensure that the information used in risk/safety assessment, as well as the methods used, are as similar as possible. It could lead to countries recognising or even accepting information from one another's assessments. The benefits of harmonisation are clear. It increases mutual understanding among Member countries, which avoids duplication and increases efficiency, which in turn improves safety.

The Working Group is also focusing on outreach activities, particularly through its information exchange mechanism, BioTrack Online. This mechanism includes information on regulatory developments in OECD Member countries, including details of laws, regulations and the contact points of the responsible ministries and agencies. It also has a database of field trials in OECD Member countries, as well as a database of products that have been commercialised.

To ensure that scientific and technical developments are taken into account, OECD countries have agreed that guidance documents will be updated regularly. Additional areas relevant to the subject of each guidance document will be considered at the time of updating.

Users of this document are therefore invited to provide the OECD with relevant new scientific and technical information, and to make proposals concerning additional areas that might be considered in the future. ***A short, pre-addressed questionnaire is included at the end of this document. The information requested should be sent to the OECD at one of the addresses shown.***

PREFACE

For risk assessments of micro-organisms used in biotechnology there is, in a general sense, a significant amount of commonality in methods that are used. Regardless of the organism employed or the uses of the organism that are evaluated, certain basic issues always need to be addressed during the course of an assessment. This document addresses one of the basic elements: the use of microbial taxonomy in assigning or confirming the identity of a subject micro-organism. Since the methods of taxonomy and the rules for naming organisms are different for prokaryotes than for eukaryotes or viruses, this document will be limited in scope to the use of taxonomy in the assessment of Eubacteria and Archea (simplified as "bacteria").

In general, the document is primarily intended for risk assessors who must deal with technical information to substantiate the identification of a micro-organism, who must make decisions on the acceptability of such information, and who must ultimately relate the information to the risk assessment of the micro-organism. It is presumed that the reviewer has a limited, but not expert, understanding of taxonomy, or at least has access to resources that can provide a basic background in this subject. The document may also be useful for applicants and other stakeholders in the regulatory process.

More specifically, the document explains why the discipline of bacterial systematics is important to risk assessment. It reviews some of the basic principles of microbial taxonomy and how they might be employed to determine potential risks of a micro-organism used in biotechnology applications. Methods applying these principles are listed and described. However, no single method is best suited for all types of bacteria. Therefore, the document also addresses some of the complexities and limitations that need to be considered in employing these principles. Finally, it addresses the prospect of future developments expected to have a significant effect on how one classifies bacteria and how those classifications may affect risk assessments.

EXECUTIVE SUMMARY

The taxonomic identification of a subject micro-organism is a key element in any risk assessment for a biotechnology product. The uses of taxonomy in risk assessment may be seen as having two components, 1) providing a common frame of reference and 2) use in predictive analysis. In order that predictive analyses can take place, good identification of both the subject and a comparison micro-organism is needed. Inferences derived from a comparison bacterium's characteristics may be used to help formulate questions for risk assessment of the subject micro-organism. Data for subject or comparison bacteria may be acquired directly through testing, or indirectly via interpretation of published, or otherwise available, information relevant to the issues of the case at hand. Selection of a comparison bacterium may be complex but they can be used in risk assessment, given a good understanding of bacterial systematics and the relationships between the comparison and subject bacteria.

Identifying an unknown micro-organism is a two-step process requiring methods to characterize the traits of an organism, and approaches to interpret the characterization data. Phenotypic methods include techniques that directly or indirectly detect, measure or characterize features of an organism resulting from the observable expression of its (total) genetic constitution. Phenotypic characteristics of bacteria include morphological, physiological and biochemical features and require growth of the organism in pure culture under appropriate conditions. Chemotaxonomic methods examine phenotype by using quantitative analysis of the organism's chemical constituents. Genotypic methods directly compare sequences, rather than rely on gene expression.

Approaches to interpret data include determinative approaches, numeric taxonomic approaches and phylogenetic approaches. There are many specific items that risk assessors need to consider to determine the adequacy of data for bacterial identification. Sometimes neither genotypic nor phenotypic methods alone suffice for either classification or identification of some bacteria, but it is possible to combine these methods using polyphasic taxonomy. Experience with all of the above techniques reveals that no single method is perfect for all taxa and all levels of taxonomic hierarchy.

Using bacterial identification in risk assessment is an inexact science and requires significant interpretive work by an assessor. There are constraints on the use of taxonomic identification methods in support of risk due to limits specific to the methods chosen, horizontal gene transfer and its affect on evolution of bacteria, variation in species concepts for different kinds of bacteria, inexact comparisons resulting from use of comparison organisms, and the overall ability to relate specific risk issues to the identification of a bacterium. Thus, interpretation of taxonomic data for use in risk assessment is not trivial, but the complexities should not be construed as preventing its use.

In general one should use methods appropriate for the organism with the objective of ensuring that the subject micro-organism cannot be confused with a member of a different species. Limitations of the techniques make it very difficult to use the simplest of methods and still obtain a reliable identification. In cases where the desired species-level assignment may not be achievable, a designation to the lowest level permissible (usually genus or subgenus) is needed. Sometimes, the best uses of taxonomic methods can narrow down an isolate to a species complex level yet still fail to provide a definitive name assignment. However, providing a single species name for a subject micro-organism, while strongly preferred, is not absolutely essential for risk assessment, provided there exist close taxonomic relatives that have been well characterized.

Knowledge of bacterial taxonomy has improved dramatically, but there is much more to be done. There is an improved ease of classification and identification of bacteria, but new knowledge has highlighted inadequacies in older techniques that may have led to some taxa being misclassified and misidentified in the past. Methods that give precise and unequivocal identifications for some genera of bacteria exist, but for many genera of current interest in biotechnology, only approximations of species assignments can be made with any confidence. To make better use out of taxonomic information, however, several advances will be needed in our knowledge, such as the application of genomics information, a better understanding of the nature of speciation in bacteria and relating taxonomic standing with risk related features.

SECTION I: INTRODUCTION

What is Taxonomy?

Taxonomy is a means for organizing elements of groups of things in an orderly reproducible manner. Vandamme et al. (1996) illustrate that taxonomy for organisms has three components: (i) classification, (ii) nomenclature, and (iii) identification.

Classification involves the organization of relationships among related taxa (*i.e.*, the ordering of organisms in groups) and the creation of schemes for interpreting those relationships.

Nomenclature is a means by which standardized approaches to the formulation of names ensures that people use the same terms in referring to a single taxonomic entity (*i.e.*, the labelling of units in those groups).

Identification is the determination that an unknown isolate belongs to one of the labelled groups (taxa) and is included within the classification schemes.

While classification provides the foundation upon which to base conclusions of relatedness, proper identification is necessary to apply the knowledge of relatedness. Because in systematics (the science of taxonomy) there are many different ways to organize information, systematists can, and often do, create more than one taxonomy for groups of organisms. Thus, there is no "official" classification of prokaryotes (Staley and Krieg, 1984). Nonetheless, classifications of most taxa become accepted through usage by the microbiology community.

However, there is an official nomenclature. Since 1980 there has been an Approved List of Bacterial Names (Skerman et. al., 1980) that is maintained by the International Committee on Systematics of Prokaryotes and updated with each publication of the *International Journal of Systematic and Evolutionary Microbiology* (IJSEM). Names of bacteria not previously listed in the "Approved List" in 1980, have been added through publication, either published directly in the IJSEM (or its predecessor the *International Journal of Systematic Bacteriology*, IJSB), or through valid publication in other journals, with subsequent listing through validation lists in IJSB/IJSEM.

Great strides are being made in refining the classification process for bacteria. As a result of the application of new methods, especially those that probe the heart of the bacterial genome, significant revisions in traditional taxonomies have been made over the past two decades. These continue to be made, and the new understandings of relationships among bacterial genera and species have allowed a better appreciation for bacterial evolution and the roles bacteria have played in the evolution of life in general. In this light, the modern tools of taxonomy and newly emerging tools of genomics have revealed that many, if not most, prokaryotes exhibit different modes of inheritance as compared with eukaryotic organisms. Whereas the eukaryotic genetic apparatus favors lineal decent with variation provided by mutation and within-species recombination, prokaryotic genomics has shown considerable complex (even chimeric) composition of the bacterial and archeal genomes studied so far (Bult et al., 1996; Galibert et al., 2001; and Wood et al., 2001) that is strongly suggestive of a significant impact of horizontal gene transfer (i. e.,

between species movement of genetic material) on prokaryotic evolution. The extent of this phenomenon may not be consistent across all bacterial genera but it appears that lineal decent is only one of two major mechanisms of gene inheritance in bacteria. The phenomenon of horizontal gene transfer and its effects on bacterial evolution has an impact on bacterial identification as is discussed in Section III.

In spite of the refinements in methods of taxonomic study, it should be recognized that the use of taxonomic information in risk assessment focuses primarily on proper utilization of identification information, not the creation of new taxonomies through the classification process. Nonetheless, it is helpful to know how organisms are assigned to categories through the classification process in order to understand the basis for the identification methods used for specific taxa. Risk assessors need to know what name to use when referring to a subject organism and how an organism bearing that name relates to other similar bacteria having valid bacterial names.

Relevance of taxonomy to risk assessment from a regulatory perspective

The risk of any new biotechnology product, particularly those derived through relatively new technologies, must be characterized prior to use. The common components of any risk assessment include a consideration of a number of factors relating to the potential for an organism to cause adverse effects (i. e., hazard) and the persistence and fate (i. e., exposure) of the organism and/or any of its by-products (e. g., toxins). These components are considered in combination to characterize the overall risk of releasing the organism to the environment.

In this context, the relevance of taxonomy as a necessary component of risk assessment is well established. Determining the identification of a micro-organism is usually the first step of a regulatory risk assessment. The uses of taxonomy in risk assessment may be seen as having two components, 1) providing a common frame of reference and 2) use in predictive analysis.

1. Providing a common frame of reference

One function of proper identification of a subject micro-organism is to establish a common label for the micro-organism. This label, the name assigned to the micro-organism, is to the risk assessment of biological products what the chemical structure and identity are to reviews of chemical substances. Not only is it essential to the basic characterization of the organism during pre-release phases, it also forms the basis for subsequent hazard and exposure assessments, especially those employing information gathered from scientific literature rather than data provided by an applicant. When identifications are reliably done, all who need to refer to a subject micro-organism, whether immediately during pre-release assessment or later after commercialization and widespread use, can feel assured that they are referring to the same entity.

If an organism is improperly or inexpertly identified at the onset of development, misunderstandings of expected or predicted features will be carried throughout the review process and if undiscovered or discovered belatedly, may result in misinterpretations of information and inadequate risk analysis. Ultimately this may lead to undesirable consequences that the risk assessment is designed to protect against.

For example, one *Pseudomonas* having a particularly difficult identification when first reviewed for a release, became a research organism several years after evaluation (Gagliardi et al., 2001). In this case extra effort was applied to get as accurate a name as possible with the tools identification available at the time. However, if the results of an initial assessment are used as support for subsequent assessments of the latter variations, and if the initial identification is in error, all the subsequent analyses will carry forward the same error. It may be fortuitous that the error is inconsequential for the initial assessments, but this

good fortune may not carry through as more complex genetic modifications are made at latter stages of product development.

2. Use in predictive analysis

Proper identification can also enable the predictive analysis of a subject micro-organism in the absence of information or direct testing. Characterization information and data from biosafety tests, usually from laboratory studies, is expected upon receipt of an application for review. However, there may be times when information from micro-organisms similar to the subject micro-organism can be used to provide risk-related information when no direct data or information concerning the subject micro-organism exists. This information can be used to identify potential concerns and in this way be used a basis for further inquiries and/or control actions for the subject micro-organism (*e.g.*, making sure that monitoring parameters are set appropriately based on expected environmental behaviours).

By having a well supported name for a subject organism one may confidently select appropriate related micro-organisms for comparison and have confidence that the use of data from such related organisms will be meaningful in support of assessing the potential risk of the subject micro-organism. The relevance of comparison organisms in this context is further described in Section III.

SECTION II: METHODS AND APPROACHES FOR CLASSIFYING AND IDENTIFYING MICRO-ORGANISMS

Classifying micro-organisms is essentially a two-step process that requires methods to characterize the traits of an organism, and approaches to interpret the characterization data to group organisms with similar traits together. Identifying an unknown micro-organism follows essentially the same two-step process (methods to characterize the traits of an organism, and approaches to interpret the characterization data); the difference being that the data from an unknown organism is compared against an existing classification scheme to arrive at a taxonomic designation.

Methods to generate characterization data range from traditional culture-based phenotypic and biochemical tests to more elaborate molecular techniques. Approaches can reflect the evolutionary inheritance of traits (*e.g.*, lineal decent), the intrinsic properties of the organism regardless of how they were acquired, or a combination of both. The basic approaches to classification and identification have evolved as the science of bacteriology has become more sophisticated, and the methods used to identify bacteria have evolved with these approaches. The intent of this section is to provide a general overview of these approaches and methods, and not to extensively review all of them. For this type of review, the reader is directed to more comprehensive reviews such as those by Rosello-Mora and Amann (2001) and VanDamme et al. (1996).

Methods

Phenotypic Methods

Phenotypic methods include techniques that directly or indirectly detect, measure or characterize features of an organism resulting from the observable expression of its (total) genetic constitution. Hence, traits expressed from plasmid-borne genes can be used, along with traits expressed from the organisms's chromosome, to determine taxonomic designations using phenotypic methods. These methods have long been employed to identify unknown organisms, and despite the advent of newer molecular technologies, still have utility in determining taxonomic designations. In part, this is because some phenotypic features are basic and critical to grouping organisms into large classes of similar organisms.

Classical phenotypic characteristics of bacteria have been described as comprising morphological, physiological and biochemical features (Van Damme et al., 1996; Rosello-Mora and Amann, 2001), and require growth of the organism in pure culture on appropriate media. Because phenotypic characteristics are culture dependent, it is important that the observed features are attributed to expression of genetic differences and not the conditions in which the organism is cultured. The basis for this limitation is discussed further in Section IV under General Issues.

While not exhaustive, the following provides examples of classical phenotypic features and methods to measure or detect them.

Morphological

Morphological features of bacteria are directly observable by the naked eye or under a microscope. While most features do not require specific methodology to observe, some require specific instrumentation

(a light or electronic microscope) and/or specific procedures (*e.g.*, spore and flagellar staining). Table 1 outlines some morphological features and associated methods to detect or characterize them.

TABLE 1	
feature	method
colony shape, colour, surface texture, margin shape	-direct observation from culture plate
cell shape and size	-cell stain (*e.g.*, Gram stain, acid-fast stain) and microscopic observation - micrometer -accurate sizing requires observation under scanning electronic microscope
Inclusion bodies (*e.g.*, polyhydroxybutyrate)	-microscopic observation
spore production and morphology	-spore staining and microscopic observation
Flagella	-flagella staining and (electronic) microscopic observation

Physiological

Physiological features characterize how and under what conditions bacteria grow, survive and perpetuate. When a micro-organism is able to grow only under specific (sometimes extreme) conditions, these features are considered robust and can be very useful in arriving at taxonomic designations. At the very least, highly restrictive physiological features can be used to narrow the identity of micro-organisms to just a few options. In general, the number of tests that are applied should be limited to those characters known to have distinguishing value (Steffen, 1998). Table 2 outlines some physiological features and associated methods to measure them.

TABLE 2	
feature	method
growth temperature (minimum, maximum, optimum)	establish growth curve over various temperatures
pH range of growth	pH strips, pH meter, colorimetric analysis
oxygen tolerance (aerobic, anaerobic, facultative)	incubate at various oxygen tensions; commercial anaerobic jars are available
carbon dioxide tolerance	incubate at various CO_2 tensions
salt tolerance	grow in various salt concentrations
requirement for growth supplements	grow in presence of growth supplements (*e.g.*, sheep blood)
antimicrobial susceptibility or resistance	grow in a concentration gradient of antimicrobial agents
susceptibility or resistance to other substrates (*e.g.*, heavy metals, sulfur)	grow in presence of heavy metals or other substrates

Metabolic

Metabolic features, for the most part, are indirectly observed because they are based on the cellular metabolism of an organism (*e.g.*, biochemical reactions or metabolic activities). The methods used to characterize metabolic features usually involve growth on various substrates, assays for enzymes in biochemical pathways or assays for metabolic by-products resulting from enzyme activity. Direct

observation techniques can also be used (*e.g.*, immunological detection of an enzyme or molecular detection of genes encoding enzymes); however, indirect methods are often used because carbohydrate fermentation, metabolic reaction, enzymatic and substrate utilization assays are fairly common, cost-efficient and simple to use. Furthermore, many commercial companies have developed miniaturized identification systems comprised of multiple assays on a single plate which makes characterizing metabolic features much simpler (see Phenetic Approaches/Numeric Taxonomy). While there are numerous enzymes and substrates that can be assayed, Table 3 outlines some enzymes and substrates with discriminating properties, and associated methods to detect them.

TABLE 3

feature	method
catalase activity	catalase test (H_2O_2 $\vert H_2O + O_2$)
oxidase activity	oxidase test (*e.g.*, Swabzyme-oxidase)
Acid production from carbohydrates	various tests depending on carbohydrate source (*e.g.*, API strips, VITEK, BBL, enterotube, etc., or whole tube analysis)
pigment production	observation of propagation media, Pseudomonas F and P agar
carbohydrate utilization	Hugh and Leifson tests depending on various sugars source
carbon source utilization	ability to grow on a sole carbon source(s) (*e.g.*, BIOLOG, or whole tube analysis)
enzymatic activities	*e.g.*, hydrolase, lipase, proteinase, amylase

Chemotaxonomic Methods

Chemotaxonomic methods involve quantitative analysis of the organism's chemical constituents. Because the techniques used in chemotaxonomy are not directed toward DNA or RNA, they are considered phenotypic methods. Consequently, they have the same culture dependent limitations as do phenotypic methods. Growing cultures under carefully standardized conditions is therefore required before any sort of comparative analysis is done. Table 4 outlines some key chemotaxonomic features and methods to characterize them.

TABLE 4	
feature	method
Gram behaviour	Gram stain
teichoic acid	extraction and purification followed by gas-liquid chromatography
peptidoglycan type	acid hydrolysis, HPLC, TLC
Fatty acid	chromatography (*e.g.*, fatty acid methyl ester [FAME] analysis by the MIDI system)
polar lipids	chromatography
lipopolysaccharides	chromatography, gel electrophoresis
isoprenoid quinones (*e.g.*, ubiquinones)	chromatography
polyamines	gas chromatography, HPLC
whole cell proteins	gel electrophoresis, SDS-PAGE

Typing

Typing methods often rely on the analysis of chemotaxonomic features of bacteria, permit the identification of features that discriminate micro-organisms below species level (*e.g.*, strains) and can be useful to understand intraspecific variability. Table 5 outlines some of the features that can be used to delineate sub-species, and associated methods to detect or characterize them.

TABLE 5	
feature	method
Antigenic cellular components (*e.g.*, capsules, flagella, fimbria, etc.)	serotyping
total cellular protein profiles	extraction followed by PAGE, multi-locus enzyme electrophoresis
Toxins	ELISA, cell line testing, molecular probes
lipopolysaccharide profiles	extraction followed by PAGE
total chemical composition	pyrolysis mass spectrophotometry, Fourier transform infrared spectroscopy, UV resonance Raman spectroscopy

Genotypic Methods

Genotypic methods are based on modern technologies and provide alternative techniques that avoid some of the problems associated with methods that depend upon phenotypic expression. Since genotypic methods directly compare sequences, rather than rely on gene expression, they are thought to be more reliable than numeric taxonomic approaches that use phenotypic characters. An important scientific reason for using these methods is that for the first time they provide a potential for development of phylogenetic taxonomies (see Phylogenetic Approaches).

DNA Base Ratios and DNA Hybridization

The DNA base ratio is one of the first nucleic acid technology applied to bacterial systematics (Lee et al. , 1956), and is calculated by the relative abundance of guanine (G) and cytosine (C) compared to the total genomic content of the micro-organism, *i.e.*, [G+C]/[A+T+C+G]. This technique follows on the premise that the nucleotide ratios [G+C]/[A+T] differ from genome to genome and hence can be used to distinguish phenotypically similar but genomically different organisms at all taxonomic levels (*e.g.*, genus, species or sub-species). In general, it is recognized that the greater the differences between the G+C content of two organisms are, the less closely related they are. It has been experimentally shown that organisms differing by greater than 10 mol% do not belong to the same genus and that 5 mol% is the average range for a species (Rosselo-Mora and Amann, 2001).

Other initial efforts focused on DNA-DNA hybridization of whole genomes. For some time, the "standard" comparison for species has been whether genomes were shared at the 70% homology level with a maximum 5º C difference in melting temperature (T_m). However, the scientific basis for this number has never been firmly established and its value seems more indicative than absolute (Vandamme et al., 1996). The basis for this standard is presumably that genomes that hybridize with the greatest affinity are likely to share the greatest sequence similarity and thus be most similar.

DNA-based Typing Methods

Like phenotypic typing methods, DNA-based typing methods allow the delineation of intraspecific boundaries between closely related organisms, that is, below species level. The methods involved in molecular typing evolved as the technology evolved, and two basic techniques are recognized. Early methods concentrated on the digestion of whole genomes using restriction enzymes followed by analysis of restriction patterns in polyacrylamide gels (restriction fragment length polymorphisms, RFLP). Later, the polymerase chain reaction (PCR) was used to amplify specific genome fragments. In this case, primers of 10-20 oligonucleotides in length are used to generate PCR products that vary depending on the organism. When the primers are very short (*e.g.*, 10 base pairs), the amplification is random and the technique is called "randomly amplified polymorphic DNA" analysis (RAPD). When the primers are based on amplification of interspersed, highly repetitive, non-coding elements, and the patterns result from the separation of fragments generated from restriction digestion, the technique is called "repetitive PCR" (rep PCR).

RNA-based Typing Methods

The general conservation of ribosomal RNA (rRNA) genes amongst bacteria, and the presence of certain hypervariable sequences in the genes within different species, can permit discrimination of bacteria to the genera, species and sometimes sub-species level (Woese, 1987; Stackebrandt and Goebel, 1994; Ge and Taylor, 1998). Ribotyping methods take advantage of the variability in rRNA genes. By restriction digest of amplified genomic RNA that contain all or part of the 16S or 23S rRNA genes, a fingerprint of the genetic material unique to the tested organism can be obtained and compared to patterns from other organisms.

Sequencing of House Keeping Genes

The advent of direct sequence analysis has allowed other methods to be developed. Some of these molecular approaches provide a focused insight into those conserved regions of the genome of bacteria that change at a rate fast enough that some variation can be observed in order to discriminate between closely related taxa, yet not so frequently as to provide too wide a separation between otherwise similar organisms. Most seek to sequence all, or portions, of housekeeping genes that fit these criteria. The most commonly

used example of this is that part of the genome coding for 16S rRNA, for the reason listed previously. Usually the portion coding for the small subunit (16S) is utilized, but the intergenic spacer region (ITS) between the 16S and 23S rRNA genes have also been used (Fisher and Triplett, 1999). Sequencing of the ITS region can be used to discriminate bacteria to the sub-species level. Other methods utilize different housekeeping genes, such as DNA gyrase (Yamamoto and Harayama, 1995). Gupta (1998) has suggested using protein sequences rather than nucleic acid sequences for these same purposes.

Approaches

Determinative Approaches

Traditional methods for identifying both bacteria and fungi can be classified as determinative methods. Using these schemes, attempts were made to classify organisms based on a few key characteristics that could easily be observed. Taxonomic designation of an unknown organism is determined following sequential analysis of relatively invariant features characteristic of a taxon based on classification keys derived from known organisms. These approaches have always relied on phenotypic as opposed to genotypic methods.

Keys are often helpful in clinical situations where rapid identification of a pathogen is needed in order to expedite therapy, and where the subset of probable organisms is limited. However, under such circumstances it is possible that, since the paramount need has been to ensure that an identification leads to proper treatment, some bacteria that are otherwise closely related, but require different therapy, may appear separated taxonomically, while others that seem to provide similar host responses, but are distantly related, may be grouped.

As more micro-organisms from diverse environments have been observed, the determinative methods have proved to have limited utility. In part, this has been due to the inability to find enough reliably invariant characteristics that are appropriate for one and only one taxon. Without these, unambiguous identifications in other than clinical situations were found to be increasingly more difficult.

Phenetic Approaches/Numeric Taxonomy

Numeric taxonomic methods have been employed for bacterial classification and identification for many decades. These are essentially statistical methods that use groups of traits that, taken together, point to specific taxa. Application of statistical approaches provides a mechanism for using a wide range of metabolic, biochemical and structural features, each given equal weight (Sneath, 1984) rather than stressing single features over others that may be present and thus have been useful for bacteria where distinguishing morphological features are absent. This analysis is referred to the “unweighted-pair group method with arithmetic mean” (UWPGA) technique. From this perspective, it has been generally accepted that an isolate must have at least 80-85% similarity to belong to a given species based on unweighted-pair group method analysis (Janda and Abbott, 2002).

This approach became the dominant one for classification of bacteria in the latter third of the Twentieth Century. In this method, researchers chose characteristics that strongly differentiated among taxa when strains were directly compared. Thus, as with the determinative approach, phenotypic methods were primarily employed. However, numeric taxonomy is really the application of statistics to any combination of features, phenotypic and genotypic, as long as they are not weighted to favor any given set of features. The intent of avoiding weighting individual elements is to avoid having single features bias the analysis. In practice, it is often difficult to give equal weight to single phenotypic traits with equivalent single genotypic results.

Usually, a battery of many tests are applied to a large number of strains of bacteria known or suspected to be closely related. Most tests are based on binary (yes or no) responses. Features are considered clearly discriminatory and are retained in the test set when strains of the same taxonomic group respond the same way (either positive or negative) the preponderance of time (*e.g.*, >75%). Those features that are equivocal (*e.g.*, gave a response of one test state versus its opposite ~50% of the time) are found not useful for the taxa in question. The responses of strains believed to belong to the same taxon are then compared statistically and a matrix of similarities devised. Taxonomic distance using this approach is thus based on an objective numerical value of similarity. Because a large number of tests are used, the consequence of error in interpreting any one test is minimized, and the discriminatory power of each test is reinforced by that of the others.

The successful application of numeric approaches has permitted development of automated methods for identification that have proved helpful in some circumstances. These automated approaches utilize simplified biochemical and physiological tests adapted to machine reading. Standard cultures are extensively tested and the results placed in a reference database. Unknowns receive the same treatment as the reference strains and the results for the tests are compared to those found in the database. Statistical analyses provide a presumptive identification or set of identifications for the isolate in question. Often the identification is accompanied by an estimation of confidence that the isolate matches a species in the database used by the method.

These automated methods have proved useful for certain taxa where a wide range of cultures have been provided for testing and inclusion in the database. Most clinical identifications are done in this fashion. Attempts have been made to extend this to “environmental” isolates, with less success. It may be that the range of variation for features significant for taxonomic placement is greater for environmental isolates than is accommodated by the databases used for clinical ones. It may simply be that too few different isolates of the former type have been collected as reference strains for the comparisons. Or, unlike some clinical isolates which might have pathogenicity factors such as a toxin, it may be that environmental isolates do not have specific, easily measurable defining traits associated with them.

For whatever reason, use of automated methods may be less successful for identification of environmental isolates than clinical ones. This is often reflected in relatively low confidence values for identifications or in multiple “likely” species names, each with “high probability” rankings, for single isolates. Users sometimes fail to understand the implications of the statistical analyses forming the basis of the identifications and will choose the highest ranked name from among those listed, rather than acknowledge that the isolate could be any of the top ranked species coming from the analysis.

Phylogenetic Approaches

Unlike higher organisms, the lack of distinguishing morphological features and a sequence of fossils that illustrate genetic variation has frustrated the development of a phylogeny of bacteria. However, with the advent of gene sequencing and the recognition that certain conserved regions of the bacterial genome could serve as slowly varying reference points in lieu of morphological change, efforts to classify bacteria by phylogenetic approaches has move quickly forward and has revolutionized the way classification of all organisms, prokaryotes and higher forms alike, are now viewed.

The phylogenetic approach is based on genotypic methods in which certain genes are sequenced, compared with sequences of other micro-organisms, and then placed in relative standing to other organisms. Housekeeping genes, or predominantly conserved genes with minor areas of hyper-variability, are often used.

In the past it was noted that the structure of ribosomes and their RNA sequences was significantly different between two large groups of prokaryotes, now called Bacteria and Archaea (Woese, 1987). As the study of the differences among the members of these large groups developed, the use of ribosomal RNA/DNA was found to be able to differentiate among finer scale taxonomic units down to the genus and sometimes species level (Palleroni, 1993). Using a statistical approach called cladistics, "trees" have been constructed that show taxonomic relationships found using comparisons of ribosomal nucleic acid sequences. While there are no consensus guidelines to delineate species based on ribosomal RNA/DNA sequences, it is generally accepted that sequences >97% in similarity belong to the same taxon (Rosello-Mora and Amann, 2001; Janda and Abbott, 2002). However, this figure has been challenged by a number of studies which indicate that unidentified isolates defined a species match with as little as 0.2-1% divergence (Drancourt et al., 2000; Woo et al., 2001; Janda and Abbott, 2002).

While such molecular approaches can give an insight into the evolutionary relationships existing among bacteria, they really represent the evolution of specific stretches of the genome in the different taxa studied. As suggested by Doolittle in Huynen et al. (1999), because there may be different rates of evolution for different parts of a bacterial genome, these methods may fail to provide a complete overview of the genetic basis for differentiation among bacterial species. One reason this may be so is that genes chosen as targets for molecular approaches to bacterial taxonomy are often found on the "chromosome" of bacteria, and are thus stably associated with a particular genome, while many of the important features of bacteria are found on extrachromosomal elements that vary in their ability to transfer among genomes. Genes introduced recently into a species through plasmid exchange and recombination thus have "evolved" differently than genes that originate in a lineal ancestor to the current species and are subsequently modified through a succession of mutational events. Using the evolution of a conserved housekeeping gene as a surrogate for whole genome evolution can thus be misleading.

Polyphasic Approaches

Sometimes neither genotypic nor phenotypic methods alone suffice for either classification or identification of some bacteria. Current methods do not rely solely on single molecular features (Steffen, 1998). It is possible to combine these methods to create a more robust method, called polyphasic taxonomy. This approach permits the combination of the information obtained from phenotypic and genotypic methods useful for the taxonomic group of concern.

As Schloter et al. (2000) note, most bacteriologists now consider that bacterial species should be defined by a polyphasic taxonomic approach integrating results from numerical pheno- and genotyping and rRNA gene homology studies. Usually this means that a species is delineated from others and can be defined if pheno- and genotyping analyses differentiate the investigated group of strains from related species in consensus, if diagnostic phenotypic characters for the new species are found, and type strains are deposited in culture collections. DNA/DNA hybridization is still preferentially considered over 16S rDNA sequencing.

The polyphasic approach utilizes various methods such as 16S rDNA molecular sequencing methods or gas chromatography/fatty acid methyl esther (GC-FAME) as a form of "range-finding" to provide putative placements of related strains in a classification diagram, usually a dendrogram or "tree". Confirmation of the placements of the members of the grouping is accomplished by independent use of structural, physiological and biochemical phenotypic tests as well as other molecular approaches (such as hybridization). Some may use multilocus enzyme assays in this approach. Others may use more than one set of conserved gene sequences as components in this methodology. Whatever method is used, it is important to be cognizant of what each component is measuring to avoid using complementary methods, which may really be providing the same answers to the same taxonomic questions.

In general, methods useful for placing an unknown into an appropriate genus- or species-level taxon are not as useful for identifying a unique isolate at the strain-level. Often single markers, or groups of unique identifiers are needed for the latter. While this is important, for example, as a means of tracing a released organism in the environment, strain- or isolate-specific methods are not the subject of this document.

Summary

Identifying an unknown micro-organism is a two-step process requiring methods to characterize the traits of an organism, and approaches to interpret the characterization data. Phenotypic methods include techniques that directly or indirectly detect, measure or characterize features of an organism resulting from the observable expression of its (total) genetic constitution. Phenotypic characteristics of bacteria include morphological, physiological and biochemical features and require growth of the organism in pure culture under appropriate conditions. Chemotaxonomic methods examine phenotype by using quantitative analysis of the organism's chemical constituents. Genotypic methods directly compare sequences, rather than rely on gene expression.

Approaches to interpret data include determinative approaches, which involve sequential analysis of relatively invariant features characteristic of a taxon based on classification keys derived from known organisms; numeric taxonomic methods which are essentially statistical methods that use groups of traits that, taken together, point to specific taxa; and phylogenetic approaches, which are based on genotypic methods in which certain genes or proteins are sequenced, compared with sequences of other micro-organisms, and then placed in relative standing to other organisms. The successful application of numeric approaches has permitted development of automated methods for identification that have proved helpful in some circumstances. Sometimes neither genotypic nor phenotypic methods alone suffice for either classification or identification of some bacteria, but it is possible to combine these methods using polyphasic taxonomy. Bacteriologists now consider that bacterial species should be defined by a polyphasic taxonomic approach integrating results from numerical pheno- and genotyping and rRNA gene homology studies. However, experience with all of the above techniques reveals that no single method is perfect for all taxa and all levels of taxonomic hierarchy.

SECTION III: CONCEPTUAL ISSUES RELATED TO IDENTIFICATION METHODS

Section II described the various methods and approaches currently available for bacterial identification. While the methods described there are well advanced and have been shown to be useful in providing identifications for biotechnology risk assessment, there are some scientific issues that arise that can affect the way reviewers interpret the results of such analyses. Some of the more significant ones are described in this section.

Limits to the ability to assign names to some taxa

Sometimes, the best uses of taxonomic methods fail to provide a definitive name assignment. For example, the genus *Pseudomonas*, is a particularly complex genus. Embedded within this genus is a "supercluster" of species of *Pseudomonas* grouped, in part, based on the production of fluorescent pigments. This subgroup is so complex that one species, *Pseudomonas fluorescens* has been subdivided into at least five biovars, each of which may deserve species rank (Barrett et al., 1986). Yet even with this fine subdivision of one species, it is not uncommon that new isolates believed to be part of this grouping cannot be definitively assigned to just one member of this supercluster. This is well illustrated by a diagram showing the relationships of many isolates of “species” of fluorescent *Pseudomonas* (Figure 1).

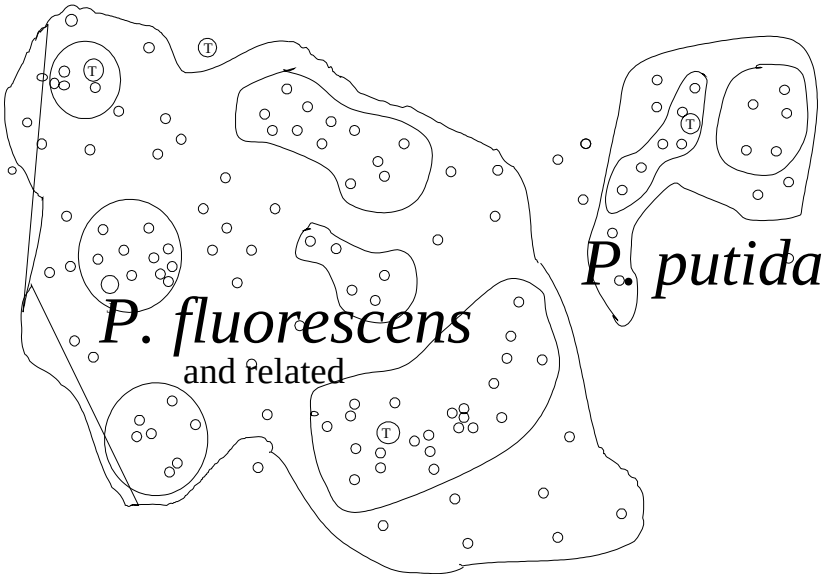

Figure 1: Modified from Smirnov and Kiprianova, 1990 (Figure 30). Scheme showing relationships among species of fluorescent *Pseudomonas.* The left hand grouping are species and biovars associated with *Pseudomonas fluorescens* and the right hand shows the two primary biovars of *Pseudomonas putida.* Note the large number of isolates that are intermediate between the "species" of the *fluorescens* cluster and the several that are located between the *fluorescens* and the *putida* groups. Individual isolates are shown by circles, type strains by "T".

Even though no existing name can be assigned, the ability to place an isolate close to characterized species is useful information for risk assessment. If there are no pathogens on either "side" of the taxonomic boundary (*e.g.*, at the genus, species or sub-species level), one would not expect the subject micro-organism to be pathogenic. If any of the micro-organisms on either "side" of the boundary is a pathogen, then one might use this as a basis for further inquiry regarding the potential pathogenicity of the subject micro-organism.

Conversely, if there are characteristics found in more distant species, but not in any of the nearby taxa, it implies that there is a far lower probability that those features from the distant species might be found in the isolate. Again to illustrate, lactose utilization is not associated with members of the fluorescent pseudomonads (Palleroni, 1984). One would not expect this feature in any member of that cluster of species. This observation has even been used as a basis for introducing a *lac zy* marker gene into one such micro-organism as a means of tracing movement in the environment (Gagliardi et al., 2001) since no *Pseudomonas* found naturally occuring at a release site for this GMM should have that feature and thus be confused with the GMM .

Effects of the dependence of some methods on stable gene expression

Phenotypic and chemotaxonomic methods measure and depend on the consistent expression of specific sets of genes. An important consideration is that gene regulation can be very complex and thus expression can be variable and environment-dependent. One possible source of apparent inconsistent expression is variability in the conditions under which expression is observed. The use of standard protocols is intended to minimize variation in expression, but this is not always accomplished. Thus, phenotypic methods can be dependent on specific laboratory associated variables. This may lead to some methods being unique to a particular analyzing laboratory and these may be difficult to replicate without

experience in that facility, and thus inter-laboratory comparisons can be difficult. While *in vitro* micro-environments can affect the operation of the genomic regulatory components, gene expression can also be impacted when multiple versions of a promoter, or even a whole operon, may be present in a genome. So even if there are scrupulous laboratory controls to limit method variability, some isolates may exhibit variable responses to some tests.

A second source of inconsistent expression is variability in the gene complement of the genome under examination. The phenotype of bacteria, being based on the observation of expressed genes, is dependent on the stability of the genes that produce the features that are measured. Stability, in this case, refers not to changes in degree of expression of the individual genes of interest, but to the presence or absence, and the frequency of occurrence in the genome of the bacterium in question, of the genes for the feature. Some genes may be mobile via the well known mechanisms of conjugation, transduction or transformation. If mobile, such genes may be lost from the original bacterium entirely, if not replicated before transfer, or may be transferred, so as to appear expressed in genomes of bacteria of other taxa. Loss can also occur if the genes are on extrachromosomal elements which themselves are lost (cured) from the genome in question. In either case, whether genes are lost or transferred, unless all features used in creating a phenotype for a taxon are stable, identification based solely on phenotype may be affected by gene mobility. While the mechanism is different than for the effects of variable gene regulation mentioned previously, the result is, similarly, reduced reliability of a measure of identification.

Horizontal gene transfer

Horizontal gene transfer and its effect on identification of bacteria

Horizontal gene transfer represents a special impediment to interpreting taxonomic data in the context of risk assessment. At the time of preparation of this document a separate Guidance Document addressing this topic in a comprehensive way was being prepared. The current document only considers the importance of horizontal gene transfer to the methods of taxonomy. For other considerations, the reader may want to consult the other document as it becomes available.

Horizontal gene transfer has an important effect on both phenotypic and molecular approaches to taxonomy. In the first case, it reduces the stability of expression of phenotypic features, because they may be gained or lost from the genome by this phenomenon. In conjugation, some plasmids are freely self-transmissible, making those genes located on such plasmids relatively unstable for that genome. Even when genes are located on the main replicon (*i.e.*, the "chromosome") of bacteria, they have the potential to be mobilized if associated with certain genomic features. Transposons and insertion sequences (IS elements) can cause genes to move within a genome, including to and from plasmids. Thus transposition has the potential to expedite transfer via a two step process; first within the bacterium, from a portion of the genome on the main replicon to a location on a smaller, often less stable, genetic element such as a plasmid, and next from one bacterium to another if the plasmid itself is transferred. Transduction may also mobilize traits. This is generally a single step process whereby a viral genetic element captures a gene or genes and allows those genes to move out of the bacterium via a subsequent "defective" viral infection of an alternate host bacterium. This may sometimes result in modification of an analogous part of the recipient chromosome, but it may be a random insertion in the recipient genome, depending on the type of tranducing phage involved in the transfer. Transduction's importance for taxonomy has not yet been fully evaluated.

To examine the potential for gene transfer within a species, one can consider similarities for specific features among closely related near neighbors of that species. One way to do this is to determine if there is a feature which, if it were to be found in the subject bacterium, would cause concern. If there is substantial experimental evidence to show that the feature of concern is never known to occur in any

member of the taxonomic grouping to which closely related species belong, such as the group of fluorescent *Pseudomonas* species in the example above, that would generally be an indication that there may be a barrier to transfer of that feature to, or maintenance in, those taxa. This would suggest that the subject bacterium may also have limited capability to acquire the feature in question via gene transfer by natural mechanisms. Conversely, if the feature is present in related taxa, then it follows that one would expect that such barriers are not so effective, and the feature could be mobilized to the subject bacterium in a natural setting.

In the second case, molecular methods as described in Section II often are based on the use of a small portion of the genome that includes relatively conserved, usually lineally inherited, housekeeping genes. Results from these methods may reflect the similarities of those components only, while not reflecting other relatedness among taxa due to a common source of horizontally, rather than lineally, distributed genes. This may not be important for certain taxa, especially those with single large replicons and small amounts of genomic material found in plasmids. However, some bacteria of importance to biotechnology, such as certain *Burkholderia* and some legume symbionts, have multiple large replicons with substantial genetic material apparently derived via horizontal transfer (Lessie et al., 1996; Galibert et al., 2001). Thus, the potential problem of seemingly different evolutionary pathways for different parts of the same genome makes drawing inferences about the total relationships between bacteria sharing those housekeeping genes used in molecular identification methods much less secure than once thought.

Horizontal gene transfer and its effect on evolution of bacteria

The complexities of classification of bacteria, and thus of their identification, are related to the processes of evolution that affect bacterial speciation. One such process is horizontal gene transfer. Horizontal gene transfer is now at the heart of a controversy over the reality of the species concept in Bacteria and Archea. Genomics research has shown that bacteria may be comprised of associations of functional subunits which have evolved from common ancestral metabolic pathways and gene sequences. Over geological time, natural multiple infusions of foreign DNA may have been responsible for the creation of new genera of bacteria, in addition to those that may have developed through mutational drift away from a common source genome. The evolution of bacteria may include acquisition or rearrangement of these components.

The presence of orthologous genes, those with structural similarity and functional relatedness even when found in different taxa, are now used as ways of illustrating evolutionary relationships among sequenced genomes (Eisen, 2000). Specialized characteristics such as those producing insect-toxic protein crystals or dinitrogen fixation give the appearance of having moved from genome to genome horizontally (Chien and Zinder, 1994; Galibert et al., 2001). Other, less obvious, features, no doubt have moved also. For details of this issue see the various reports of annotations of whole bacterial genome sequencing such as Bult et al. (1996), Galibert et al. (2001) and Wood et al. (2001).

Doolittle (1999) has argued that gene transfer precludes establishing a universal tree of life and has further indicated his belief (Doolittle, 1999; Huynen, et al. 1999) that speciation in Bacteria and Archea is not meaningful in a systematics sense due to the “chimeric” nature of prokaryote genomes. That is, even if one uses a variety of measures to evaluate the whole of a genome, the different components of the genome may have different “histories” and thus are not comparable in a phylogenetic sense. Others counter with their belief that bacterial genomes contain a core of genes not affected by horizontal transfer and thus subject to lineal descent. These genes would provide an anchor to the concept of species. It appears that this controversy will not be settled until many whole bacterial and archeal genomes are fully sequenced and the ancestry of core genes, if they exist, are worked out.

Most of these issues were addressed in a recent conference on the identification of members of the genus *Pseudomonas* (Workshop on Identification Methods for *Pseudomonas,* 1997, http://www.bif.atcc.org/epa_web/). This is a complex genus, which has undergone extensive reorganization in the past two decades. On the basis of molecular methods, it was split into multiple genera. However, even members of the remaining species assigned to this genus cannot all be readily distinguished using current technologies. What is apparent is that some members of this genus, such as the type species, *P. aeruginosa,* provide coherent species by almost any method, including traditional phenotypic ones. Conversely, as illustrated at the beginning of this Section, other species like *P. fluorescens, P. putida, P. tolassii, P. marginalis,* etc. are better viewed as a species complex, with subunits as broad as species level taxonomic units as described for other genera, yet so diverse that the boundaries among these subunits cannot be defined. The size, proportion included in extrachromosomal elements, and plasticity (including transposons and insertion sequences) of the genomes of these species makes for a nearly impossible task of clear species separation. A polyphasic approach was considered essential for the identification of members of this genus, but even after proper application of such a method, it was concluded that many of the isolates that fall within the species complexes, like *fluorescens* or *stutzeri,* will not be separable into single named species.

Regardless of how this debate over speciation in bacteria is resolved, gene transfer is seen as having the potential to have some influence over the evolution of bacterial taxa, and thus on bacterial classification and identification. Even if Doolittle's proposal is shown to be correct, there will remain pragmatic approaches, such as polyphasic taxonomy, to revealing relationships among groups of organisms we now consider unique taxa.

Relevance of comparison organisms

Unless a subject micro-organism has been previously released into the environment, pre-release data specifically on a subject micro-organism's behavior after release cannot be obtained directly. Yet such test data on very similar organisms may exist, and some extrapolations may be made from these similar organisms. Hence, in some cases, a comparison micro-organism can be used for assessing the potential risks of a subject micro-organism prior to deliberate release into the environment.

In order that predictive analyses can take place, good identification of both the subject and comparison bacterium is needed. In addition, for the comparison bacterium, there must be additional useful information beyond that otherwise available for review of the subject organism, *i.e.*, the information on the comparison bacterium must provide added value to the assessment of the subject micro-organism. This capability will always be limited by the degree of functional similarity between the two organisms, but it is often the case that close functional and taxonomic relatedness occur together. One should not expect that analyses of a comparison micro-organism will automatically provide direct answers to risk assessment questions about a subject organism. Nevertheless, inferences derived from a comparison's characteristics may then be used to help formulate questions for risk assessment of the subject micro-organism.

Data for comparison bacteria may be acquired directly through testing, or indirectly via interpretation of published, or otherwise available, information relevant to the issues of the case at hand. Such issues as toxin production or pathogenicity potential are obvious ones for evaluation. Other examples of data may include experience with field releases under controlled conditions or microcosm studies (*e.g.*, Gagliardi et al., 2001) for strains related to the subject organism.

Use of taxonomy as a basis for further inquiries

The concept of using taxonomic relatedness to choose a comparison bacterium as an aid in risk assessment is predicated on the assumption that most of the features of concern for the subject organism

are inherited in a common fashion in both the subject and comparison bacterium. If lineal descent controls the inheritance of the key characteristics of one organism, but not the other, extrapolations based on observations of the comparison bacterium will not be as meaningful as they would otherwise be. However, horizontal gene transfer may affect both organisms similarly, in which case assessments based in part on observations of a comparison bacterium would retain their validity.

No matter how the speciation problems for bacteria are resolved the risk assessor will still be faced with the reality that use of comparison organisms will provide only inexact comparisons in most cases. Unless the comparison bacterium chosen is a direct precursor of the subject micro-organism, there will usually be, at best, an approximate mapping of expressed features from the comparison bacterium to the subject.

One often uses comparisons among members of the same genus since these would be expected to share more functions than distantly related taxa. Consider the example of members of closely related species within the genus *Pseudomonas*. As mentioned earlier, these species are often difficult to separate taxonomically. For example a new isolate that falls somewhere within the "*fluorescens* supercluster" would be expected to share a limited set of features exhibited by all the members of that group, even though it may not be possible to tell within which species or biovar of that group the new isolate belongs. As an illustration from a risk assessment perspective, a new isolate that is identified as a biovar of *P. fluorescens* might merit further inquiry and/or testing, since the subgroup, *P. fluorescens* biovar II, has embedded within it some members previously known as *P. marginalis*, which is a known plant pathogen (Janse et al., 1992). Since not all *P. marginalis*-like isolates fit within the boundaries of *P. fluorescens* biovar II, one cannot automatically assume that only biovar II isolates are likely to be pathogenic, nor can one assume that biovar II isolates must be pathogenic, but the direction of inquiry can be focused by knowing that the isolate falls within the "supercluster" and can be enhanced if own can narrow down that information to one of the categorized biovars of *P. fluorescens*.

In a similar way, many members of the fluorescent *Pseudomonas* group share production of biologically active pigments as a feature of their metabolism (Palleroni, 1984). Therefore a new isolate found to closely resemble members of this cluster of bacteria, but not assignable to just one existing taxon, would nevertheless be expected to possess the capability of producing bioactive pigment molecules related to those made by the other similar bacteria in the supercluster.

Sharing of features is not limited to species in the same genus. As indicated above, one of the revelations of modern bacterial systematics is that some features of bacterial genomes may be derived from distantly related species. This may result in some functional characteristics of bacteria being shared among seemingly remotely related taxa. For example, one current method of classification places three important bacterial species previously seen as distantly related, *Sinorhizobium meliloti*, *Agrobacterium tumefaciens* and *Brucella abortus*, in the same *alpha*-2 subclass of the group of Gram-negative organisms included in the Proteobacteria (Stackebrandt et al., 1988). The three species represent a nodule symbiont, a plant pathogen and an animal pathogen respectively; three different ecological niches. Yet all three share a common type of feature, namely, host/bacterium interaction.

Furthermore, the plant pathogenic species *Agrobacterium tumefaciens* and the symbiotic species *Sinorhizobium meliloti* share much genomic sequence material located on the largest replicon, usually called the chromosome, of their respective genomes. Wood et al. (2001) found these two genomes so similar that they postulated that they were derived from a recent common ancestor. In addition research has revealed that in both species a gene was independently discovered, and uniquely named, that codes for a 1,2-glucan synthetase which was found in each to have some function in host interactions (Inon De Iannino et al., 1998). The respective genes from each species were found to complement defective versions in one of the other species. Apparently a related module of carbohydrate synthesis has been conserved in these

related species, even though they are found in distinct genera. Evolution has allowed each to vary the genes which specify the type of host with which they interact, but the production of certain carbohydrates was apparently retained by each as a component of the interactions.

Other examples of features that appear across a wide range of genera considered distantly related by some measures, but which might be closely related by other criteria, include complex metabolic functions important to biotechnology exemplified by the nitrogenase complex genes or aromatic biodegradation genes (Chien and Zinder 1994; Hirsch et al., 1995; Harwood and Parales, 1996). How this may happen may be explained by recent developments in genomics, which is the study of an organism's genetic material.

By careful consideration of the relevance of the relatedness between these species, one could conceivably use each as a comparison for the other for those functions known to be shared. To do so, however, requires a sophisticated understanding of the principles of bacterial taxonomy and the limitations of this discipline. Extrapolations from comparisons of distantly related species must only be undertaken when there is knowledge of specific common functionality that is understood to be derived from a common genetic source.

Differences in species concepts for different kinds of bacteria

The same criteria for determining what is a species are not applied to all bacteria. Schloter et al. (2000) addresses this briefly by illustrating that "inconsistencies in systematics of bacteria arise from the simultaneous application of different species concepts. A serious obstacle to a unified species concept is subjective consideration of practical usefulness for species definition. Particularly in the group of human and animal pathogens, many species are delineated primarily on phenotypic traits such as host range preference and pathogenicity. *Brucella* species, for example, show interspecific DNA relatedness above 98%." It was pointed out that a single species concept (i. e. *Brucella melitensis*) comprising six biovars and respective biotypes for *Brucella* strains, was proposed, but this concept is not accepted in the scientific community working on pathogenic micro-organisms. Several other genera of well known pathogens might also be seen as comprised of a single species, but cannot be reclassified because of the confusion it might cause among public health workers.

Conversely, several genera of well known bacteria have undergone, or are undergoing, a process of ever finer subdivision, rather than consolidation, that is resulting in frequent classification and nomenclature changes. *Burkholderia cepacia* was once considered a single species but recently (Vandamme, et al., 1997) underwent the first of many revisions that resulted in designation of many new species level taxa (Coenye, LiPuma et al., 2001; Coenye, Mahenthiralingam et al., 2001) called genomovars. The frequency of such changes in classification are very difficult to keep up with and could result in assignment of an outmoded species name, based on current nomenclature, to a subject organism during risk assessment. It could mean that different authorities assessing the same bacteria at different times might use different names for the same object, thus defeating the use of taxonomy in providing a common term of reference for subject organisms.

Summary

Using bacterial identification in risk assessment in an inexact science. It requires significant interpretive work by an assessor. There are constraints on the use of taxonomic identification methods in support of risk due to limits specific to the methods chosen, horizontal gene transfer and its affect on evolution of bacteria, variation in species concepts for different kinds of bacteria, inexact comparisons resulting from use of comparable organisms, and the overall ability to relate specific risk issues to identification of a bacterium.

SECTION IV: CONSIDERATIONS WHEN SUBSTANTIATING A TAXONOMIC DESIGNATION

General Considerations

Experienced risk assessors have successfully used bacterial taxonomy as an aid in assessment for many years. Nonetheless, the use of taxonomic identification methods in support of risk assessments may be constrained for many reasons. Some of those constraints are imposed by limits specific to the methods chosen. For each method, the strengths and weaknesses must be considered, as some will work well for certain isolates but not others. From this perspective, the purpose of identification must always be kept at the forefront, whether it is to provide a common focus of discourse and information exchange regarding a subject bacterium, or to help choose a comparison for a test micro-organism and, from observations for that comparison bacterium, devise an appropriate question set for developing a risk assessment for the subject organism. Therefore, identifications of micro-organisms are conducted on a case-by-case basis, applying the most appropriate methods for the subject organism.

The previous section provided some illustrations of scientific questions that indicate the need for caution in the use of taxonomic data for risk assessment. While the issues presented suggest that interpretation of taxonomic data for use in risk assessment is not trivial, they nevertheless should not be construed as preventing its use. On the contrary, such information can be very useful provided that care is taken in the generation and use of relevant data. This section is meant to provide guidance on how to identify an unknown micro-organism by indicating the type of information useful in determining an identification, and describing general and specific considerations in generating and applying the data.

Inherent methodological limitations

For many taxa, phenotypic numeric taxonomic (NT) methods have proved especially helpful. However, there are significant problems in its application to bacteria. Numeric taxonomic approaches that include only phenotypic methods sometimes do not provide adequate resolution for species level classification. Since numeric taxonomy often uses a large set of often independently expressed features for developing a taxonomy for bacterial groupings, it is generally not the size of the test set that is the problem with phenotypic NT. Rather, the problems often lie in the requirement of dependable expression of each test under comparable conditions.

There are several reasons why expression of such features may not be dependable. Laboratory variability can lead to unreliable expression. This is illustrated by one study in which split samples from the human periodontal pocket, identified by the same laboratory using immunologic probes and traditional pure culture techniques, yielded levels of agreement between 0% and 81% (M. I.. Krichevsky, personal communication). Hence, because taxonomies are based on test systems devised by individual researchers, other laboratories that choose to replicate the test systems must do so exactly or risk observing different responses than expected. Slight differences in variables such as media composition, temperature, inoculum size and incubation time may result in opposite test results from seemingly identical cultures. Similar, but non-identical test batteries may be used to do comparisons. All of these variables may lead to different results for the same set of strains done at different laboratories.

There may be other reasons that lead to variable test results. Some features may simply not be expressed all the time in certain strains, even when observational variables are kept constant. Thus, variable responses can occur even when the same tests are done by the same technician in the same laboratory. Regulation of the pathways in some strains of bacteria may be unpredictable, possibly because regulatory elements are defective or are controlled by intracellular variables not known to researchers. Usually reliable features may, therefore, vary even if the structural genes for the features are fully functional. It is commonly noted that long term maintenance of cultures can lead to loss of features that are fully functional in new isolates of the same strain. An understanding of why these test responses vary requires an understanding of the biochemical basis for gene function. Unfortunately, such specific genetic research has been performed for a limited number of key pathways in a small set of bacteria. For these reasons, better methods for classifying bacteria that do not rely on potentially unpredictable expression of genes have been sought.

Molecular methods may avoid the problems of phenotypic methods since they have the advantage of directly analyzing sequences, rather than relying on potentially variable gene expression. These methods, such as those that employ housekeeping gene sequencing, including 16S rDNA/DNA or *gyr*B provide a potential for development of phylogenetic taxonomies. However, the molecular methods, by being limited to comparison of a few conserved genes, generally do not provide as broad a comparison as rigorous phenotypic methods. While such molecular approaches can give an insight into the evolutionary relationships existing among micro-organisms, especially bacteria, they really represent the evolution of specific, relatively invariant, stretches of the genome in the different taxa studied.

Interpretation of molecular data

Molecular data, such as gene sequencing and Southern blot analysis using gene probes, can provide information regarding the presence or absence of a particular gene. However, such analysis does not necessarily provide the complete picture with respect to the expression of such genes and hence of the phenotypic characteristic of the organism, *i.e.*, the presence of a gene does not necessarily mean that it is expressed. In recent studies, a number of *B. cereus* hemolytic and non-hemolytic enterotoxin genes were found in non-cereus *Bacillus* species, but these were not expressed (Hansen and Hendriksen, 2000; Rivera et al., 2000; P. Gillevet, personal communication 2002). Concluding a taxonomic designation of these non-*cereus* isolates in this situation would be incorrect if based solely on molecular data. Therefore, data such as this should be interpreted with caution by risk assessors. This is particularly important where the expression of structural proteins are encoded by a series of genes (*e.g.*, operons, pathogenicity islands). In this instance, the utilization of a probe for one gene in, for example, the operon, may not necessarily mean that the other genes in the operon are present in the genome or that the gene is even transcribed.

The potential limitations of molecular analyses can be mitigated by using follow-up methods, such as reverse transcriptase polymerase chain reaction (RT-PCR) to determine if the gene is transcribed. If RNA transcripts can be detected, the potential for expression of functional formed proteins could be investigated by immunologic probing to determine if the protein is synthesized and cytotoxicity assays to determine if the protein is functional.

Pragmatic Considerations

Dealing with uncertainty

Reasonable efforts to obtain a taxonomic designation for a subject micro-organism should use methods appropriate for the organism. For risk assessment purposes, taxonomic designations should minimally be to the species level, and should follow international codes of nomenclature and standard taxonomic sources where they exist. The objective is to ensure that the subject micro-organism cannot be

confused with a member of a different species, especially with relevance to ones having undesirable attributes.

However, providing a species name for a subject micro-organism, while preferred, is not absolutely essential for risk assessment, provided there exist close taxonomic relatives that have been well characterized. For various reasons, even when very sophisticated methods are used, a species-level assignment may not always be achievable. This usually manifests itself when an isolate is incompletely classifiable because it is shown to be very close to the boundaries defining two named species, but not close enough to deserve either name. This knowledge generally suggests that there is a reasonable probability that some characteristics of the two nearby species may also be found in the isolate, but in a combination unique to the isolate.

In these cases, a designation at the lowest level permissible (usually genus or subgenus) is needed. Often, the isolate in such cases may belong to a "species complex", which concept, while generally reflecting the reality of systematic proximity, has no nomenclature standing in bacterial taxonomy. Nevertheless, identification to the level of such groupings, by showing relationships to closely related species, can still be useful in the risk assessment of the micro-organism. As previously discussed, information on functional properties can be implied from knowledge of approximate taxonomic placement, and be useful in predicting potential risks of an unknown micro-organism, if similar properties are found in comparison organisms.

Using appropriate rigor in performing identifications

While it may be economically tempting to use a simple, often automated, approach to identify an unknown micro-organism to the species level, the limitations of the methods (see above and Section III) make it very difficult to use the simplest of methods and still obtain a reliable identification. One must carefully consider and understand the basis for choosing one identification over another, including the limitations of an automated system, in order to ensure that appropriate designations are chosen.

Over the years, reviewer experience has shown that automated systems using phenotypic methods infrequently provide useful identifications for risk assessments of subject organisms derived from environmental isolates. Such methods are dependent upon the strength of the computer database on which the statistics of these methods are based. Even using the most rigorous phenotypic methods, many environmental isolates fall outside of well defined taxa. This problem seems to be amplified when the identification methods are of the automatic variety. These methods are useful, nevertheless, in "range funding", *i.e.*, providing an initial indication of the plausible taxonomic neighborhood to which an isolate may belong. Obtaining a single name for a subject bacterium using these methods is not readily accomplished, except for certain specialized sets of taxa (*e.g.*, clinically important bacteria) for which the databases behind the methods are most robust.

Simple, automated methods may work for some species, provided the species is one for which the method has an adequately robust database and the strains to be identified are typical of the species. The API 20E strip, for example, is still considered the "gold standard" commercial system for the identification of species in the family *Enterobacteriaceae* (O'Hara et al., 1992). Some, however, question even the values of these tests for certain micro-organisms. Given the substantial increase in new described taxa since 1975 (Euzeby, 1997), Janda and Abbott (2002) make the point that many new taxa added to existing commercial databases are based on the results of tests from pre-configured panels, even though the best tests available to identify these new isolates are not on the panels or are not amendable to automated commercial kits.

Frequently, however, organisms used in biotechnology are not in that category and the simple-to-use automated identification systems may not be adequate. Difficulties most often arise when the subject organism is an environmental isolate with no apparent connection to the more common clinically important species, which often dominate the population of the database used in automated systems. Many of the species used in bioremediation, for example, are members of complex genera (*e.g.* fluorescent pseudomonads) in which member species are notoriously difficult to separate and identify. As noted previously in Section IV, for these types of bacteria, some automated methods provide apparently positive species identifications when in fact equivocal results are more appropriate.

Even more commonly, these methods provide equivocal results, but the presentation of the results by commercial suppliers of this information may lead users to misinterpret them as giving unique positive identifications. Often the results of an identification will be a list of several taxa, not all of which need be closely related to each other. Many automated identification systems will often list the most probable identifications in rank order, with the most probable at the top. The difficulty arises when attempting to rely on the "preferred" name that is provided by such systems. Simply choosing the top name is inappropriate since many systems will list only those names that are contained in the system database. It is possible that in some cases, other more similar species have never been tested using the systems method, and thus have no relevant data contained in the database used.

For example, a method report might indicate that an unknown tested micro-organism has a 70% probability of being species "X", and a 50% probability of being species "Y", while the organism is in fact a species "Z". If it turns out that species "Z" is unknown to the database used for the methods of statistical analyses, choosing "X" would be wrong in this case. Simply because there is a strong separation between the most probable and next most probable does not require that one must choose from the list supplied. Rather, one should be sure that the identification was made in an absolute sense as well as a relative one. Users are generally better served by beginning with the list provided and obtain additional data that can be used to discriminate among the candidate bacterial names listed, rather than to choose the most probable name in the list and assume it is correct.

Therefore, such identifications should be examined carefully, particularly when more than one possible species name is designated for a single culture. In general, if methods used are unable to resolve the identification of an unknown, it is better to provide more than one possible identification than to arbitrarily choose one name from a list of options.

Unless rigorous analyses needed for publishing a new species has been done, however, one should be cautious before declaring an unknown isolate to be a new species. When those who perform identifications do not show strong confidence in a name resulting from an identification of a subject organism, even after extensive follow-up test as suggested by Steffen (1998), it is better to consider all the options revealed by the testing, even those which point to the lack of an existing taxon, and the probabilities associated with those options, than to choose just one.

Who does an identification?

From a practical standpoint, the identification of a genetically modified micro- organism may be affected by the manner in which precursor organisms are obtained and identified. There are three basic ways in which bacteria are obtained and characterized prior to genetic modification: 1) from a service culture collection, 2) from a research culture collection, or 3) from original isolation from a native source. There are considerations which vary among these modes that could affect the confidence in the name assigned to the precursor organism and thus the identification of the modified micro-organism.

Most service collections, whose primary function is generally to store and distribute cultures, usually for a fee, perform extensive characterizations of their cultures. Those who construct modified micro-organisms starting with cultures obtained from such collections usually select cultures with most of the features desired in the final micro-organism and proceed to make stepwise modifications of the precursor. This procedure most often guarantees a significant degree of characterization of the final micro-organism that includes information provided by the service collection and acquired by the developer. The identification of the final organism is usually dependent on the work of the culture collection respective of the supplied culture. Most such identifications are of high quality. Questions may arise for cultures that have been in a collection for many decades and were originally characterized by methods that currently might be construed as less reliable, but many collections are going back to their older materials and re-characterizing the cultures using modern techniques. Many "older" cultures are also characterized by modern techniques when new batches are made for "replenishing" isolates for which there are no more cultures left in stock.

In some ways, research collections can be even more reliable than service collections for some micro-organisms. That is because researchers working with a very few taxa may have done more extensive characterization for each culture than a service collection with thousands of cultures. Sometimes a research collection will use the most sophisticated methods for identifying cultures in which they have an interest, methods that may not be generally available to service collections. However, some of those who maintain a research collection of cultures do so for other than taxonomic purposes. Such research collections may not be as interested in the taxonomic status of their cultures as the performance of the cultures with respect to the investigative interests of the researcher. In these cases, original isolates may be identified by less sophisticated means, to obtain a convenient label, even though other components of characterization are highly sophisticated. Thus the focus of research, and its effects on the methods of taxonomic characterization of cultures in the collection, needs to be taken into account when obtaining micro-organisms to be modified for use in biotechnology.

When micro-organisms are obtained from an original habitat, identification can be obtained in two ways. Some will send a culture to an organization that performs identification of unknowns as a service. Such organizations include service culture collections, research collections or commercial companies using automated identification methods. Others will perform self-characterizations with facilities that they maintain themselves. Such facilities may range from sophisticated methods and equipment equivalent to any in a research or service collection to simple commercial kits for automated analyses.

For the person performing a risk assessment on a micro-organism there are some concerns that must be dealt with in order to evaluate the reliability of a name assigned to a micro-organism used in biotechnology. It is usually appropriate to accept identifications performed by service collections because of their experience and the need for them to provide accurate information to customers about cultures in their collections. However, as mentioned earlier, older holdings in the collection may not have had the benefit of characterization using sophisticated modern methods. This does not necessarily put into question an identification, since many species are as readily identified by older methods rigorously applied as by newer techniques, but some caution is in order for such cases. Good documentation by a service collection should help dispel any such concerns.

Similarly, research collections that specialize in classification of taxa of interest also should provide identifications that need not be questioned. Research collections that do not publish on the systematics of their cultures may use appropriate techniques nevertheless, but it might be helpful to inquire about the methods used to verify this.

Identifications done in-house by an applicant or under contract may need to be scrutinized closely before organism names are accepted. Some commercial services provide fine identifications for some taxa

and questionable ones for others. Those using automated systems are dependent on the quality of databases used during the identification procedures. Some of these databases are excellent for a few taxa, but have only a few examples of others. Years of experience by some authorities who have performed risk assessments on GMMs used in biotechnology has revealed that environmental isolates identified by these automated methods are often mis-named or left un-named due to limitations of these databases.

What is the "best" approach?

Experience with all of the above mentioned techniques reveals that no single method is perfect for all taxa and all levels of taxonomic hierarchy (Janda and Abbott, 2002). It is true, however, that great strides in applying molecular methods have been made in the past few decades. When regulatory evaluation of biotechnology products was relatively new in the 1980's, few laboratories were equipped to perform ribosomal nucleic acid analyses. Now rDNA assays are nearly routine. Chemotaxonomic methods such as FAME are also more commonplace. These advances permit the application of more sophisticated analytical methods to the problems of bacterial identification.

Nonetheless, the sophistication of these methods, alone, is insufficient to make bacterial identification trivial in most cases for biotechnology product micro-organisms. The scientific reasons for this have been cited in earlier sections. The best advice that one can use in choosing methods is to have the work performed by those familiar with the presumptive genus of the bacterium and who are prepared, if needed, to perform methods in combination, "polyphasic" methods, to resolve identification problems.

Some years ago Palleroni attempted to illustrate which methods were best for certain purposes (Palleroni, 1993), but an analysis of this work reveals that each method has its limitations and that it usually is best to combine methods in a polyphasic approach. However, there seems to be no single polyphasic taxonomy methodology that is best for all bacteria. Steffen (1998) reported on the approach used by the German Collection of Micro-organisms and Cell Cultures (DSMZ). An extensive set of procedures was outlined. It was reported that for a reliable identification result most bacteria require simultaneously performed identification methods combined with secondary, and in most cases some tertiary, biochemical tests. By experience of the DSMZ Identification Service, a combination of primary, secondary and tertiary biochemical tests, one or two partly automated commercially available test systems (API, BIOLOG), and the sequencing of the 16S rDNA bacterial gene usually leads to the affiliation of an isolate to a certain species. However, Steffen stated that due to the fact that different organisms or different taxonomic groups have been studied and classified by a wide variety of methods, standardization of identifications cannot be currently achieved. One may have to deviate from a rigid scheme intended for most bacteria. An attempt to devise an alternate scheme only for the genus *Pseudomonas* was described earlier (ATCC, 1997). It was suggested that initial efforts using 16S rDNA and GC-FAME methods often provided a reasonable range-finding for the initial part of an identification scheme. Participants at the workshop that developed this scheme acknowledged that for certain complex species in that genus, even this pragmatic approach might not suffice.

Apparently, as previously mentioned, gene transfer has a role to play in this uncertainty. The significance of the effects of gene transfer on bacterial evolution is that many taxonomies are based on the inheritance of a few stable, rarely mobilized characteristics. Many of these are essential functions, called housekeeping genes. While use of methods that employ these genes, such as 16s rDNA, certainly provide strong evidence for understanding how the main portion of a genome has evolved, this approach might have reduced utility for risk assessment if the features of concern for a subject micro-organism are not part of that main portion and are not primarily inherited in a lineal fashion. A reviewer, therefore, would want to know if, and to what degree, horizontal gene transfer is known to have affected members of the genus of a subject organism before selecting a method for identification of the bacterium to the species level. This is

true especially if the methodological options include some that depend on an element of part of the genome affected by this phenomenon.

In short, there is no "best" method. Those who need to identify their cultures must be aware of the advantages and limitations of each type of approach with respect to the presumptive classification of the strain in question. That usually means an initial round of basic "range-finding" tests that rapidly narrow the identification to a few genera, followed by selection of a method that is most likely to give a useful answer. Useful, in this context, may mean a single, species name, arrived at with great confidence, or it could mean an approximation, entailing a choice of several species known to be related in a "complex", but not always distinguishable from each other. Knowledge that the unknown belongs within the complex may impart enough information to complete a risk assessment, even if the exact species name cannot be determined. The choice of method may be dependent on how precise an identification needs to be for assessment purposes. For example, a method that allows an approximate placement within a complex, where all members of the complex are innocuous or beneficial bacteria, may be sufficient, depending on the intended use of the organism. On the other hand, even one that provides a unique species name may not be enough, if subspecies or individual strains of a named species differ significantly in potential for detrimental effects, such as pathogenicity.

Data needs for the reviewer

The following describes some information that risk assessors find helpful when substantiates the taxonomic designation of a bacterium:

1. Tests and Databases

    - list of tests used to arrive at the taxonomic designation, and a brief description of test conditions, when such conditions are not established as standard for the methods used.
    - data from the tests used to arrive at the taxonomic designation; and second choice
    - any database against which the test data was compared
    - other test data to differentiate the notified micro-organism from close relatives and/or pathogens

2. Molecular identification (modified from the Points to Consider for Identifying a Pseudomonad, ATCC 1997 - http://www.bif.atcc.org/epa_web/)

    a) 16S rRNA

    - the method used
    - the sequence used to determine the isolate was in the concluded genus and/or species and the basis for comparison
    - the measure of similarity and the value obtained

    b) DNA homologies

    - description of sequences used (*e.g.*, homologous sequences, coding sequences)

- the method used (*e.g.*, PCR, RFLP, fingerprinting, DNA/DNA hybridization)
- the homology results that led to the conclusion

3. Phenotypic tests

a) Morphological

- cellular (*e.g.*, shape, Gram stain, size of bacterium, spore production/morphology)
- colonial (*e.g.*, shape, colour, surface texture, margin)

b) Physiological

- growth conditions (*e.g.*, temperature, type of media, pH, oxygen requirement)
- metabolic products

c) Metabolic

- biochemical reactions (*e.g.*, catalase and oxidase activity)
- substrate utilization (*e.g.*, glucose, sucrose, formic acid, lactic acid)

4. Chemotaxonomic tests

- cellular components (*e.g.*, fatty acids, polyamines)
- cell surface components (*e.g.*, antigens, lipopolysaccharides, cell wall components)

Interpretation of "positive" identifications

Not only is risk assessment predictive in nature, it is also an art and often times involves the informed judgement of the risk assessor. The current state of knowledge for some micro-organisms, however, sometimes makes it difficult to make a determination of risk despite the positive identification of the micro-organism.

A case in point here is the *Burkholderia cepacia* complex (Bcc), which is comprised of more than seven distinct genomovars (phenotypically similar but genotypically distinct organisms) [Vandamme, et al., 1997; Coenye, et al., 2001 (a,b)]. Some members of this complex have beneficial biotechnological applications (environmental isolates) while others cause adverse human health effects, sometimes death, in patients with cystic fibrosis (clinical isolates). These predominant isolates are generally assigned to specific genomovars (*e.g.*, Genomovars I and IV are predominantly environmental strains whereas Genomovars II and III are predominantly clinical strains); however, all genomovars have been found in clinical settings. Hence, the positive identification of a particular genomovar does not necessarily preclude the micro-organism from being one of clinical importance. Furthermore, a number of virulence factors and markers have been identified with clinical strains but these factors and markers are not present in all

clinical strains. Hence, the absence of a virulence factor or marker also does not necessarily preclude the micro-organism from one of being clinical importance.

The risk assessor must therefore take particular caution in the interpretation of these results. Further information in this case, such as testing in appropriate and validated animal models for pathogenicity endpoints, is essential to aid the risk assessor in characterizing risk. It becomes evident, then, that the information requirements required under various legislation, regulations and/or guidelines are critical for the characterization of a micro-organism, and strengthens the fact that identification is a critical element in the risk assessment but not the only element.

Issues for selecting comparison micro-organisms

GMMS present a special case for bacterial identification during pre-release phases because data may not have been yet gathered on the subject micro-organism that best describes that organism, especially when introduced outside the laboratory. This necessitates use of a closely related comparison organism to acquire relevant data. One should choose a comparison bacterium that is expected to most closely mimic the behaviour or characteristics of the subject micro-organism. For GMMs, this comparison bacterium ideally and most likely will be a direct precursor of the subject micro-organism (*i.e.*, the naturally occurring parental organism of the genetically modified micro-organism). Provided proper use of systematics is employed, taxonomic relatedness may then be used as a selection criterion for obtaining a comparison bacterium other than a direct precursor.

It must always be understood that taxonomic similarity is not an exact equivalent to functional similarity (Achenbach and Coates, 2000). Useful information does not necessarily have to come from the closest relative of the subject organism. The farther one gets taxonomically from the subject organism, usually, the lower the confidence that features relevant to risk assessment of the subject organism will be present in the comparison bacterium.

Post-release issues

Use of taxonomic data and its application to selection of comparison bacteria may also be used to help with selection of monitoring and testing methods, should further work on the subject organism be deemed necessary. If the intended use or environmental testing of a micro-organism being readied for release is such that dispersal away from the site of application is a concern, use of test data previously obtained from a comparison bacterium can help in the selection of monitoring methods, test site design and mitigation strategies for dealing with undesirable outcomes. For example, if it is known that a closely related micro-organism has high mobility and potential for dispersal from an application site, then placement of sampling devices can be adjusted to ensure capture of mobilized subject bacteria, which would be expected to behave similarly to its close relative. Knowledge of such things as heat, drying or oxygen tolerances may help establish where, when and, how to sample for the released organism. The presence of resistant forms such as spores, or the observed viable-but-not-culturable (VBNC) state in a comparison bacterium would lead the investigator to prepare for longer and more specialized monitoring regimes than otherwise might be planned. That is, methods that require culturing of micro-organisms would fail to detect a released bacterium that had entered a VNBC state, but certain molecular methods, not needing cultivation of the bacteria, might be able to do so long after the other methods could not. Thus, an extended monitoring period and the use of methods not dependent on culturing, would be called for if it were known that such features were present in the gene pool of the subject organism. Many of these features might not be observed under laboratory testing schemes used to prepare a subject organism for risk review.

Summary

Interpretation of taxonomic data for use in risk assessment is not trivial, but the complexities should not be construed as preventing its use. This section provided both general and specific guidance on generating and applying relevant data. In general one should use methods appropriate for the organism with the objective of ensuring that the subject micro-organism cannot be confused with a member of a different taxon. Limitations of the techniques make it very difficult to use the simplest of methods and still obtain a reliable identification. In cases where the desired species-level assignment may not be achievable, a designation to the lowest level permissible (usually genus or subgenus) is needed. A list of specific considerations for risk assessors, in consideration of the adequacy of data for bacterial identification, was provided in this section. Finally, a practical concern is locating a person or organisation to perform identifications. Factors in making such a decision were described.

SECTION V: ISSUES FOR THE FUTURE

Much has changed since regulation of bacteria produced through biotechnology began in the 1980's. Techniques for modifying organisms have improved and the knowledge and experience needed to perform risk assessments of these new bacteria have improved significantly as well. Specifically for this document, knowledge of bacterial taxonomy has improved dramatically, especially within the last decade. However, this improved knowledge has not necessarily made classification and identification of bacteria easier. Some of what the new knowledge has revealed is that it has been easier, in the past, to misclassify and misidentify bacteria through use of techniques that were not as accurate and useful as presumed at the time of their application. The new knowledge has not always brought us closer to understanding speciation in bacteria.

This is not to say that the use of current concepts of bacterial taxonomy is not appropriate when applied to risk assessment. Rather, it means that risk assessors need to be very cautious in taking taxonomic information at face value and to extrapolate with great care. Methods that give precise and unequivocal identifications for some genera of bacteria exist, but for many genera of current interest in biotechnology, only approximations of species assignments can be made with any security.

Because most genetically modified micro-organisms cannot reveal their full potential until release to the environment, and because some testing to observe characteristics relevant to risk assessment also cannot be done until releases take place, any mechanism for anticipating what those characteristics might be and ways of planning to observe them after release requires predictive information that often can only be obtained by observing related bacteria under conditions of interest. Unless some way exists for the use of information on related organisms, this presents a difficult situation for the developer of a new biotechnology bacterial strain - one can’t test without release, but one can’t get permission to release without testing first.

Selection of an appropriate comparison organism for which data exists requires some way of predicting relatedness. Taxonomic relatedness, while acknowledged as an imperfect predictor, is still a useful indicator. To make better use out of taxonomic information, however, several advances will be needed in our knowledge.

Understanding the nature of speciation in bacteria

The current debate over what constitutes a bacteria species and even whether a species is definable for prokaryotes, needs to be advanced. Specifically, there needs to be a resolution of the issues regarding the existence of a “core” genome in bacteria. If such cores exist, as revealed by whole genome sequencing and comparative genomics, there is hope for developing new molecular techniques that allow for taxonomies to be based upon the core genomes. It also will help in understanding the role of horizontal gene transfer in bacterial speciation, possibly reducing the confusing effect of this phenomenon on bacteria systematics.

Relating taxonomic standing with risk related features

Although taxonomically closely related bacteria are presumed to be related in a more general sense, this is not necessarily the case. Some means needs to be devised that can illustrate when

taxonomically related organisms are also similar in their risk potential. When phenotypic information is used to develop taxonomies, it would be helpful to understand if these same features have any bearing on risk. Some such features, such as growth at mammalian body temperature or versatility in use of carbon substrates, may be shown to reflect survival potential in certain environments. There is a need to systematically review these identification characteristics, correlate them with environmental parameters and/or known effects of bacteria on hosts, and develop a scheme to better utilize this information in risk assessment.

Use of genomics

Some of the issues above may be resolvable through the use of genomics. Genomics is an emerging tool for use during risk assessments of micro-organisms. Whole genome sequencing and the use of sophisticated bioinformatics techniques may enable reviewers to begin to answer questions that have been problematic up until now. For taxonomy, comparisons of whole genomes for multiple examples of related bacteria may overcome the impediments to classification imposed by standard identification methods. Early use of genomics has, for example, shown how specific portions of some genomes must have been derived via horizontal, versus lineal, inheritance (Garcia-Vallve et al., 2000).

The number of microbial genomes is expanding rapidly. About sixty microbial genomes were published by early 2002 and about 175 genomes were undergoing sequencing at that time (TIGR Microbial Database; http://www.tigr.org/tdb/mdb/mdbinprogress.html). Several of these included species from the same genus and strains from the same species, allowing for comparative genomics to take place. From a microbial systematics perspective, once sufficient comparisons are completed it may be possible, for some taxa at least, to determine what may be the "core" genome that is shared by related taxa and thus stabilize some of the rapidly evolving classification schemes. Should this be successful, the next step of developing molecular probes for these "cores" is already within the capabilities of current technology. Thus, identification could, in future, be linked directly to the genome of a micro-organism via a specific molecular probe for a unique taxon, or group of taxa, rather than through the use of indirect measurements of indicator features, as is now the case.

There needs to be restraint to the enthusiasm for genomics, however. Given the way taxonomy may be used in risk assessment as a predictive tool to assist in the absence of direct measurement of microbial function during pre-release phases of review, it is important to understand that genomics only can describe the genetic potential of a micro-organism. It is evident that gene expression is highly complex in organisms and that the regulatory networks even of bacteria may be so large and redundant that describing the genetic potential of a bacterium through genomics may not be sufficient to predict post-release functioning of the organism. What genomics may be able to do directly is reveal the absence of a genetic potential for a particular function and provide a measure of assurance if that particular feature is of concern for risk.

This potential is emphasized by a recent report (Stahl and Teidje, 2002) issued by the American Academy of Microbiology, which is comprised of a prestigious set of microbiologists selected for their career accomplishments. A colloquium sponsored by the Academy focused on the effects of advances in genomics on the discipline of microbial ecology. A significant portion of the report dwelled on the importance of microbial systematics to microbial ecology and the influence of advances in genomics on systematics. As the report pointed out:

> "(t)oday, traditional taxonomic concepts *(i.e.*, species, genus, family) do not serve microbial systematics, in which problems of horizontal gene transfer and mechanisms of speciation and evolution are varied and complex. A new framework for taxonomy, one better adapted to genomic information and

> microbial taxa, needs to be derived. ... (but) taxonomy is far more than an outdated means of classification. It provides a common language for describing microbial forms in the context of a rich literature about their physiology, metabolism, and life history. Molecular phylogenetics has forced us to reevaluate how organisms are related without requiring us to discard traditional taxonomic views....An important feature of sequence-based classification schemes is that they provide a universally applicable and cost-effective method, eliminating much of the ambiguity arising from earlier systems. In addition to providing information about evolution and phylogenetic relationship, sequences will ultimately be mapped to specific phenotypic and ecological characteristics of an organism."

It is especially this last feature of the potential of genomics that provides a tie from taxonomy to risk assessment. The new tools that may elucidate systematic relations among bacteria may simultaneously provide insight into functional relationships. That capability is at the heart of the use of comparison organisms when direct measurement of features is not possible, so that genomics approaches promise to provide significant refinement to methods used to deal with this problem.

In any case the incorporation of genomics into microbial risk assessment is not that far off, with simplistic application of data from the few currently sequenced micro-organisms already being used by some reviewers, though not yet as a taxonomic tool. If the subject micro-organism, or its precursor, has been, fortuitously, the subject of a sequencing project, much speculation about its genetic potential can be eliminated by querying the information published or otherwise available from the genome sequence.

Summary

Knowledge of bacterial taxonomy has improved dramatically, but there is much more to be done. There is an improved ease of classification and identification of bacteria, but new knowledge has highlighted inadequacies in older techniques that may have led to some taxa being misclassified and misidentified in the past. Methods that give precise and unequivocal identifications for some genera of bacteria exist, but for many genera of current interest in biotechnology, only approximations of species assignments can be made with any security. To make better use out of taxonomic information, however, several advances will be needed in our knowledge, such as understanding the nature of speciation in bacteria and relating taxonomic standing with risk related features.

REFERENCES

Achenbach, L. A. and J. D. Coates. 2000. Disparity between bacterial phylogeny and physiology: Comparing 16S rRNA sequences to assess relationships can be a powerful tool, but its limitations need to be considered. *ASM News*. 66:714-715

American Type Culture Collection. 1997. Workshop on Identification Methods for Pseudomonas.

Barrett, E., R. E. Solanes, J. S. Tang and N. Palleroni. 1986. Pseudomonas fluorescens biovar V: its resolution into distinct component groups and the relationship of these groups to other P. fluorescens biovars, to P. putida, and to psychrotrophic pseudomonads associated with food spoilage. J. Gen. Microbiol. 132:2709-2721

Bult, C. J., O .White, G. J. Olsen, L. Zhou, R. D., Fleischmann G. G. Sutton, J. A. Blake, L. M FitzGerald, R. A. Clayton, J. D. Gocayne, A. R., Kerlavage B. A .Dougherty, J. F. Tomb, et al. 1996. Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*. *Science*. 273:1058-1073

Chien Y. T. and S. H. Zinder. 1994. Cloning, DNA sequencing, and characterization of a nifD-homologous gene from the archaeon *Methanosarcina barkeri* 227 which resembles nifD1 from the eubacterium *Clostridium pasteurianum*. *J. Bacteriol*. 76:6590-6598

Coenye, T., J. LiPuma, D. Henry, B. Hoste, K. Vandemeulebroecke, M. Gillis, D. Speert and P. Vandamme. 2001. *Burkholderia cepacia* genomovar VI, a new member of the *Burkholderia* cepacia complex isolated from cystic fibrosis patients. *Int. J. Syst. Evol. Microbiol.* 51: 271–279

Coenye,T., E. Mahenthiralingam, D. Henry, J. LiPuma, S. Laevens, M. Gillis, D. P. Speert and, P. Vandamme . 2001. *Burkholderia ambifaria* sp. nov., a novel member of the *Burkholderia cepacia* complex including biocontrol and cystic fibrosis-related isolates. *Int. J. Syst. Evol. Microbiol*. 51: 1481–1490

Doolittle, W. F. 1999. Phylogenetic Classification and the Universal Tree. *Science*. 284: 2124-2128.

Drancourt, M. , C. Bollet, A. Carlioz, R. Martelin, J.-P. Gayral and D. Raoult. 2000. 16S ribosomal DNA sequence analysis of a large collection of environmental and clinical unidentifiable bacterial isolates. *J. Clin. Microbiol.* 38: 3623-3630

Eisen, J. A.. 2000. Assessing evolutionary relationships among microbes from whole-genome analysis. *Curr. Opinion Microbiol*. 3:475–480

Euzeby, J. P. 1997. List of bacterial names with standing in nomenclature: a folder available on the internet. *Int. J. Syst. Bacteriol.* 47: 590-592

Fisher, M. M. and E.W. Triplett. 1999. Automated approach for ribosomal intergenic spacer analysis of microbial diversity and its application to freshwater bacterial communities. *Appl. Env. Microbiol.* 65(10): 4630-4636.

Gagliardi, J. V., J. S. Angle, J. J. Germida, R. C. Wyndham, C. P. Chanway, R. J. Watson, C. W. Greer, T. McIntyre, H. H. Yu, M. A. Levin, E. Russek-Cohen, S. Rosolen, J. Nairn, A. Seib, T. Martin-Heller and G. Wisse. 2001. Intact soil-core microcosms compared with multi-site field releases for pre-release testing of microbes in diverse soils and climates. *Can. J. Microbiol.* 47: 237-252.

Galibert F., T. Finan, S. Long, A. Puhler, P. Abola, F. Ampe, F. Barloy-Hubler, M. Barnett, A. Becker, P. Boistard, G. Bothe, M. Boutry, L. Bowser, J. Buhrmester, E. Cadieu, D. Capela, P. Chain , A. Cowie, R. Davis, S. Dreano, N.Federspiel, R. Fisher, S. Gloux, T. Godrie, A. Goffeau, B. Golding, J. Gouzy, M. Gurjal, I. Hernandez-Lucas, A. Hong, L. Huizar, R. Hyman, T. Jones, D. Kahn, M. Kahn, S. Kalman, D. Keating, E .Kiss, C .Komp, V. Lelaure, D. Masuy, C. Palm, M. Peck, T. Pohl, D. Portetelle, B. Purnelle, U. Ramsperger, R. Surzycki, P. Thebault, M. Vandenbol , F. Vorholter, S .Weidner, D. Wells, K. Wong , K.Yeh and J. Batut. 2001. The composite genome of the legume symbiont *Sinorhizobium meliloti. Science* 293:668-672

Garcia-Vallve, S., A. Romeu and J. Palau. 2000. Horizontal Gene Transfer in Bacterial and Archaeal Complete Genomes. *Genome Research* 10: 1719-1725

Gaviria Rivera, A. M., P. E. Granum and F. G. Priest. 2000. Common occurrences of enterotoxin genes and enterotoxicity in *Bacillus thuringiensis*. *FEMS Microbiol. Lett.* 190: 151-15

Ge, Z. and D. E. Taylor. 1998. *Helicobacter pylori*: Molecular genetics and diagnostic typing. *Br. Med. Bull*. 54(1): 31-38

Gillevet, P. 2002. Personal Communication.

Gupta, R. 1998. Protein phylogenies and signature sequences: a reappraisal of evolutionary relationships among archaebacteria, eubacteria and eukaryotes. *Microbiol. Molec. Biol. Rev.* 62:1435-1491

Hansen, B. M. and N. B. Hendriksen. 2000. Detection of Enterotoxic *Bacillus cereus* and *Bacillus thuringiensis* Strains by PCR Analysis. *Appl. Environ. Microbiol*. 67: 185-189.

Harwood, C. S. and R. E. Parales. 1996. The beta-ketoadipate pathway and the biology of self-identity. *Ann. Rev. Microbiol.* 50:553-590

Hirsch A. M., H. I. McKhann, A .Reddy, J. Liao, Y. Fang and C. R. Marshall. 1995. Assessing horizontal transfer of nifHDK genes in eubacteria: nucleotide sequence of nifK from Frankia strain HFPCcI3. *Mol. Biol. Evol*. 12:16-27

Huynen, M. B. Snel, P. Bork, J. W. Stiller, B. D. Hall and W. F. Doolittle. 1999. Technical Comments: Lateral Gene Transfer, Genome Surveys, and the Phylogeny of Prokaryotes. *Science* 286:1443a

Inon De Iannino, N., G. Briones, M. Tolmasky, and R. Ugalde. 1998. Molecular cloning and characterization of the *Brucella abortus* cyclic (1-2) glucan synthetase gene: genetic complementation of *Rhizobium meliloti* ndvB and *Agrobacterium tumefaciens* chvB mutants, *J. Bacteriol*. 180:4392-4400

Janda, J. M. and S. L. Abbott. 2002. Bacterial identification for publication: when is enough enough? *J. Clin. Microbiol.* 40(6): 1887-1891

Janse, J. D., J. H. J. Derks, B. E. Spit, and W. R. van der Tuin. 1992. Classification of fluorescent soft rot Pseudomonas bacteria, including P. marginalis strains, using whole cell fatty acid analysis. System. Appl. Microbiol. 15:538-555

Kiprianova, E. A., A. V. Panichev, O.I. Boyko and A.D. Garagulya. 1979. Numerical taxonomy of bacteria belonging to the genus *Pseudomonas. Mikrobiologia.* 48:1023-1032

Krichevsky, M. I. 2001. Personal Communication from the First Meeting of Participants in the Double Blind International Cooperative Study for the Identification of *Pseudomonas* Species. March 8-9, 2001.

Lee, K. Y., R. Wahl and E. Barbu. 1956. Contenu en bases purique et pyrimidiques des acides deoxyribonucleiques des bacteries. *Ann. Inst. Pasteur.* 91: 212-224

Lessie, T., W. Hendrickson, B. Manning and R. Devereux. 1996. Genome complexity and plasticity in *Burkholderia cepacia. FEMS Microbiol. Lett.* 144:117-128

O'Hara, C. M. D. L. Rhoden and J. M. Miller. 1992. Re-evaluation of the API 20E identification system versus conventional biochemicals for identification of members of the family *Enterobacteriaceae*: a new look at an old product. *J. Clin. Microbiol.* 90: 1S-15S

Palleroni, N. 1984. *Genus* 1. *Pseudomonas Migula 1894, 237*^{AL} *In:*Bergey's Manual of Systematic Bacteriology: Vol. 1.; N. R. Krieg and J.G. Holt, *eds.* Williams and Wilkins, Baltimore *pp. 141-199*

Palleroni. N. 1993. *Pseudomonas* classification: A new case history in the taxonomy of Gram-negative bacteria. *Antonie v. Leeuwenhoek.* 64:231-251

Rossello-Mora, R. and R. Amann. 2001. The species concept for prokaryotes. *FEMS Microbiol. Rev.* 25: 39-67

Schloter, M., M. Lebuhn, T. Heulin and A. Hartmann. 2000. Ecology and evolution of bacterial microdiversity. *FEMS Microbiol. Rev*. 24:647-660

Skerman , V., V. Mcgowan , and P. Sneath 1980, Approved Lists of Bacterial Names. *Int. J.Syst. Bacteriol.*, 30:225-420

Smirnov V. V. and E. A. Kiprianova. 1990. "Bacteria of the Genus Pseudomonas". Kiev Council of Sciences, Kiev, Ukraine. p. 158

Sneath, P. 1984. Numerical taxonomy. *In* N.R. Krieg and J. G. Holt (eds.) *Bergey's Manual of Systematic Bacteriology*. Vol. 1. The Williams and Wilkins Co., Baltimore. *p*. 111-118

Stackebrandt, E., R. G. E. Murray, and H. G. Trüper. 1988. Proteobacteria classis nov., a name for the phylogenetic taxon that includes the "purple bacteria and their relatives". *Int. J. Syst. Bacteriol.* 38:321-325

Stackebrandt, E. and B. M. Goebel. 1994. Taxonomic note: a place for DNA-DNA reassociation and 16S rRNA sequence analysis in the present species definition in bacteriology. *Int. J. Syst. Bacteriol*. 44:846-849

Stahl, D. and J. Tiedje. 2002. Microbial Ecology and Genomics: A Crossroads of Opportunity- A Report from the American Academy of Microbiology. Amer. Acad. Microbiol., Washington DC, 28 pp.

Staley, J. and N. Krieg. 1984. Classification of prokaryotic organisms: an overview. In: *Bergeys Manual of Systemtatic Bacteriolgy*. Williams and Wilkins, Baltimore. pp. 1-4.

Steffen, M. 1998 Characterisation and identification of bacteria. *In*: Proceedings: Microbiological Plant Protection Products - Workshop on the Scientific Basis for Risk Assessment National Chemicals Inspectorate, Sweden. pp 39-48.

Vandamme, P., B. Holmes, M. Vancanneyt, T. Coenye, B. Hoste, R. Coopman, H. Revets, S. Lauwers, M. Gillis, K. Kersters and J. Govan. 1997. Occurrence of Multiple Genomovars of *Burkholderia cepacia* in Cystic Fibrosis Patients and Proposal of *Burkholderia multivorans* sp. Nov. *Int. Jour. Systematic Bacteriol.* 47:1188–1200

Vandamme, P., B. Pot, M. Gillis, P. De Vos, K. Kersters and J. Swings. 1996. Polyphasic taxonomy, a consensus approach to bacterial systematics. *Microbiol. Rev*. 60:407-438

Woese, C. R. 1987. Bacterial evolution. *Microbiol. Rev.* 18:221-271

Woo, P. C. Y., E. Y. L. Cheung, K.-W.Leung and K.-Y. Yuen. 2001. Identification by 16S ribosomal RNA gene sequencing of an *Enterobacteriaceae* species with ambiguous biochemical profile from a renal transplant recipient. *Diagn. Microbiol. Infect. Dis.* 39: 85-93

Wood, D., J. Setubal, R. Kaul, D. Monks, J. Kitajima, V. Okura, Y. Zhou, L. Chen, G. Wood, N. Almeida L. Woo, Y. Chen, I. Paulsen, J. Eisen, P. Karp, D. Bovee, P. Chapman, J. Clendenning, G. Deatherage, W. Gillet, C. Grant, T. Kutyavin, R. Levy, M.-J.Li, E. McClelland, A. Palmieri, C. Raymond, G. Rouse, C. Saenphimmachak, Z. Wu, P. Romero, D. Gordon, S. Zhang, H. Yoo, Y. Tao, P. Biddle, M. Jung, W. Krespan, M. Perry, B. Gordon-Kamm, L. Liao, S. Kim, C. Hendrick, Z.-Y.Zhao, M. Dolan, F. Chumley, S. Tingey, J.-F.Tomb, M.Gordon, M.Olson and E.Nester. 2001. The genome of the natural genetic engineer *Agrobacterium tumefaciens* C58. *Science* 294:2317-2323

Yamamoto, S.and S. Harayama. 1995. PCR amplification and direct sequencing of *gyr*B genes with universal primers and their application to the detection and taxonomic analysis of *Pseudomonas putida* strains. *Appl. Environ. Microbiol.* 1:110 -1109

QUESTIONNAIRE TO RETURN TO THE OECD

This is one of a series of OECD Guidance Documents that provide information for use during regulatory assessment of particular micro-organisms, or plants, developed through modern biotechnology. The Consensus Documents have been produced with the intention that they will be updated regularly to reflect scientific and technical developments.

Users of Guidance Documents are invited to submit relevant new scientific and technical information, and to suggest additional related areas that might be considered in the future.

The questionnaire is already addressed (see reverse). **Please mail or fax this page (or a copy) to the OECD, or send the requested information by E-mail:**

OECD Environment Directorate
Environment, Health and Safety Division
2, rue André-Pascal
75775 Paris Cedex 16, France

Fax: (33-1) 45 24 16 75
E-mail: ehscont@oecd.org

For more information about the Environment, Health and Safety Division and its publications (most of which are available electronically at no charge), consult ***http://www.oecd.org/ehs/***

==========

1. Did you find the information in this document useful to your work?

   ☐ Yes ☐ No

2. What type of work do you do?

   ☐ Regulatory ☐ Academic ☐ Industry ☐ Other (please specify)

3. Should changes or additions be considered when this document is updated?

4. Should other areas related to this subject be considered when the document is updated?

Name: .....
Institution or company: .....
Address:.....
City:.....Postal code: ..... Country:.....
Telephone:.....Fax: ..... E-mail: .....
Which Document are you commenting on? .....

FOLD ALONG DOTTED LINES AND SEAL

.....

PLACE
STAMP
HERE

OECD Environment Directorate
Environment, Health and Safety Division
2, rue André Pascal
75775 Paris Cedex 16
France

.....