We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900
Open access books available

185,000
International authors and editors

200M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



Chapter 3

Modern Taxonomy for Microbial Diversity

Indira P. Sarethy, Sharadwata Pan and
Michael K. Danquah

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/57407

1. Introduction

Microorganisms are actually composed of very different and taxonomically diverse groups of communities: archaea, bacteria, fungi and viruses. The members of these groups or taxa are distinct in terms of their morphology, physiology and phylogeny and fall into both prokaryotic and eukaryotic domains. They constitute a broad group of life system inhabiting the known ecosystems on earth: terrestrial and marine; including geographical locations considered to be extreme or inimical to life. The latter comprise of such areas as habitats with high salinity, alkalinity, acidity, high and low temperatures, high pressure, and high radiation. Considering the adaptability of microorganisms to grow and survive under varied physico-chemical conditions and their contribution in maintaining the balance in ecosystems, it is pertinent to catalogue their diversity as it exists. The inability to visualize them with the naked eye precludes effective classification. As such, using the available tools, microorganisms are broadly classified into prokaryotes and eukaryotes and subsequently into various taxonomical units depending on the resources available and required.

The sustenance of life on earth depends on maintaining the diversity of microorganisms. Human intervention is resulting in depletion of biodiversity and many hotspots are also fast losing their endemic biodiversity. While specific data is hard to come by, it is likely that loss of macro life forms also results in loss of the associated microbial species: symbionts as well as the rhizosphere-colonizing microbes. The significant contribution made by microorganisms in ecosystem sustainability as well as the industrially important biomolecules obtained from them: antibiotics, anti-cancer drugs, enzymes, biofuel and various other compounds, implies that cataloguing them is imperative. However, a simple and effective microbial identification system is still far off. The available tools for classification and identification of microorganisms rely on a number of different technologies. This chapter provides an overview of taxonomy

INTECH
open science | open minds

© 2014 The Author(s). Licensee InTech. This chapter is distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/3.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

tools for understanding prokaryotic and eukaryotic microbial diversity. Taxonomy (or biosystematics) consists of three main parts: *classification* (arrangement of organisms based on similarity), *nomenclature* (naming of the organisms) and *identification* (determining whether an organism belongs to the group under which it is classified and named). Modern biosystematics also includes *phylogeny* as an integral part of the classification process [1].

2. Biogeography of microbial diversity

The diversity of microbial communities varies within habitats as much as between habitats [2]. This variation can even occur within a few millimetres, suggesting that microbial diversity encompasses more than the documented evidence available. Hence, biogeography is gaining importance as a field of study from microbial diversity point of interest. Many reasons have been postulated to explain this phenomenon. Due to the innately small size of the microorganisms, environmental complexity plays a major role in determining diversity. Spatial heterogeneity is likely to lead to the formation of many niches within a habitat [3]. Recent tools like metagenomics aid in biogeography studies by providing information on nucleic acid sequence data, thereby directly identifying microorganisms (see Section 9). Therefore the phylogenetic information can be used to compare microbial diversity profile across habitats [2].

Generally, diversity within a particular location and in a community is called alpha diversity. Beta diversity measures the community composition between two or more locations while gamma diversity applies to a region, across continents and biomes and is larger in size than that used for measuring alpha diversity [4].

3. Microbial evolution

The evolutionary relationship of microorganisms is called phylogeny. Understanding phylogenetic profiles of microbes becomes a daunting task because of their small size and the lack of particular indicators that could serve as markers. Some proteins and genes are considered as evolutionary chronometers which measure the evolutionary change [5]. Currently, the 16S rDNA sequence is considered to be most reliable for measuring evolutionary relationships in bacteria and archaea (detailed in Section 7.2.1) and the 18S sequence for fungi (see Section 8). However, it is necessary to choose the correct protein or gene for such studies. Such a gene or protein should have certain features which make it most appropriate for deriving evolutionary relationship. The most important criterion is that it should be present in all members of the target group and be functionally homologous in the organisms. The molecule must contain regions of conserved sequences for comparison purposes. The changes in sequence data must be at a slow enough rate to permit measurement so that it may also reflect evolutionary change for the entire group [5].

In the current system of classification, based on the 16S rDNA sequence, evolutionary relationships form the basis for division and three major domains have been recognized, out

of which two comprise of bacteria and archaea (prokaryotes) and the third domain is of eukaryotes [6]. It is important to understand evolution in the context of biodiversity. Evolution leading to new ecotypes/species is achieved in many ways. Some species with quick generation times also undergo mutation frequently leading to novel species or strains [3]. Horizontal gene transfer (HGT), via transformation, transduction or conjugation, also accounts for introduction of genes into distantly related organisms, thereby introducing new traits and also impacting on interaction between species and thereby ecosystem processes [7]. It has also been hypothesized that large population sizes of microbes and their low extinction rates may also play a role in maintaining biodiversity, though measurement of such extinction rates is difficult ([8] and references therein).

4. Microbial phylogeny

The phylogenetic tree representing all living organisms shows that, evolution of current forms of life occurred from a common ancestor (the universal ancestor), depicted by the root (see Figure 1). Two domains are of prokaryotic systems of life: the archaea and Eubacteria; in contrast to previous systems of classification, wherein, the prokaryotes were confined to a single kingdom. However, it is intriguing to note that, genomic studies have shown the archaea to contain unique gene sequences which are not present in bacteria or eukaryotes. Certain

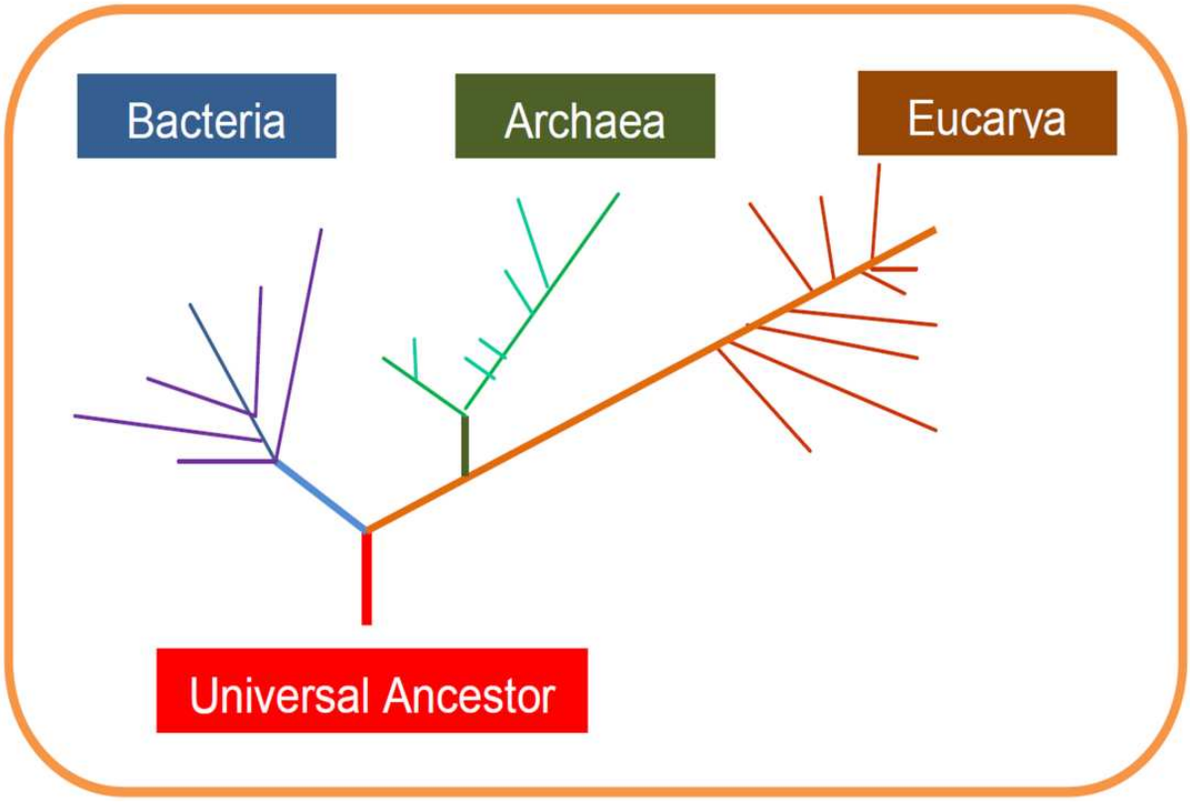


Figure 1. The universal phylogenetic tree based on comparative ribosomal RNA sequences (adapted from [6]).

genes are also shared between all the three domains. The genes required for core cellular functions are the ones which are necessary for survival of a cell and could have arisen from the common ancestor.

The divergence of the organisms represents the differences in genetic sequences which could have become fixed in each group as they evolved. It is also postulated that earlier, HGT played a key role in transfer of genes between organisms early in the evolutionary history [6]. It occurs as a response to any change in the environment and provides for better adaptation ([9] and references therein). Subsequently, reproductive isolation could have prevented extensive exchange of genes, though it continues to occur amongst prokaryotes.

5. The prokaryotic microbes — Bacteria and archaea

The bacteria and archaea have evolved along different lines though both are essentially prokaryotic. The archaea are considered to be the most primitive and are common inhabitants of the so-called extreme habitats (hot springs, deep sea hydrothermal vents, alkaline and acidic habitats). Though the bacteria and archaea share certain common features, the archaea also share similarities with eukaryotes which are further exemplified by the 16S rDNA-based phylogenetic analyses.

6. The species concept

An array of diverse definitions have been proposed to describe microbial species. Currently, a polyphasic approach is used to define a microbial species using phenotypic and genotypic properties [1, 10]. Whenever a new taxon is proposed, it is essential that the organism be isolated in pure culture and its characteristic features be tested under standard conditions [11]. Whether an organism constitutes a member of a common species is primarily based on whether its DNA-DNA re-association values are more than 70% and melting temperature (ΔT_m) is less than 5ºC, the experiments being performed under standard conditions [12]. All the strains within a species must show similar phenotypes. A designated type strain of a species constitutes the reference specimen for that species [13]. A species description must preferably be based on the characteristics of more than one type strain. To be assigned a different species name, members must show at least one and is governed by the [12]. If the 16S rDNA sequences of organisms are ≤ 98.7% or ≤ 97% identical, they are members of different species. This is considered even in the absence of DNA-DNA hybridization experiments since this level of divergence in 16S rDNA sequences constitutes less than 70% DNA-DNA similarity [14]. Uncultured microbes cannot be assigned to a definite species since their phenotype is not known; however, they can be assigned a '*Candidatus*' designation provided their 16S rRNA sequence subscribes to the principles of identity with known species [15]. A concept applying to a taxon lower than that of the strain is the ecotype – those microorganisms that occupy an ecological niche and are adapted to the conditions of that niche [16]. It is important to remember

here that the nomenclature of a taxon is very important as it serves to maintain effective communication across microbiological disciplines and it is governed by the Bacteriological Code [17, 18].

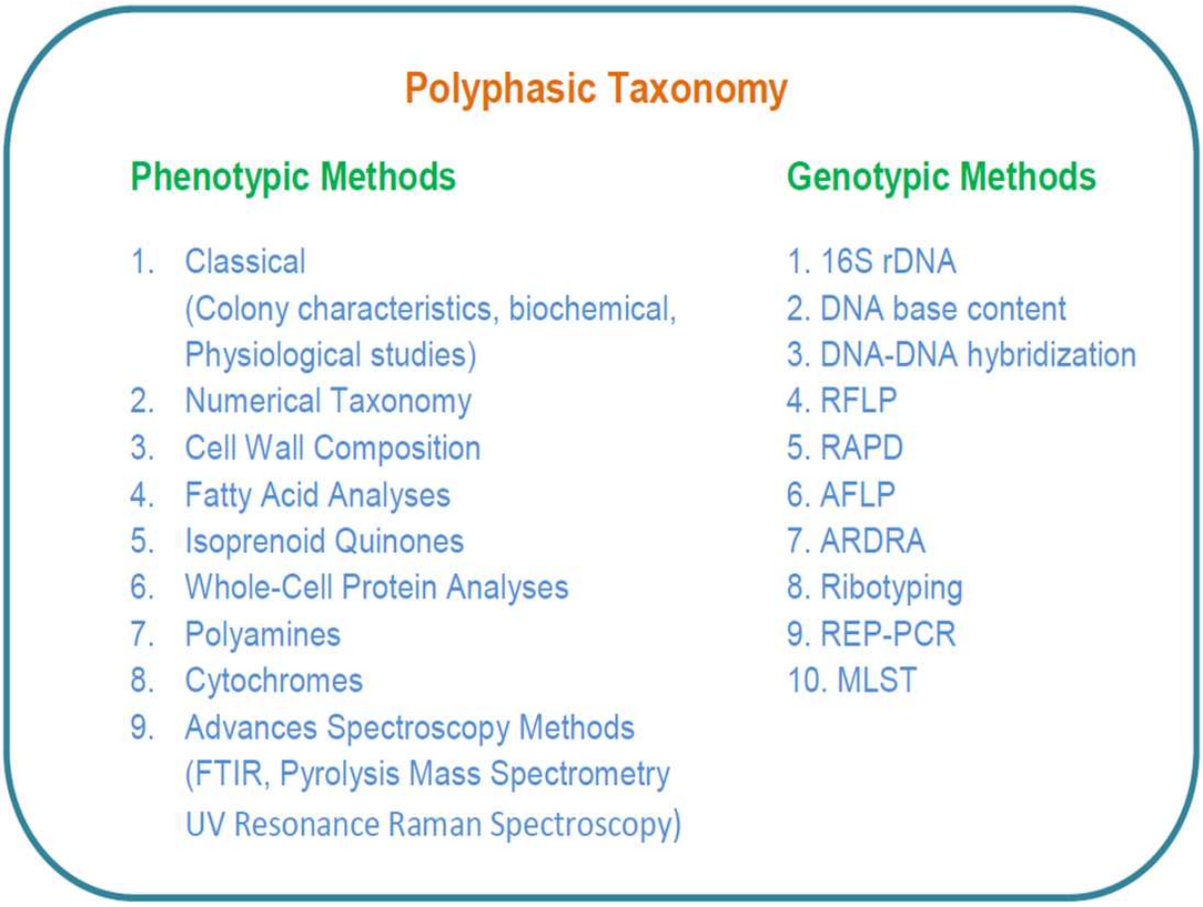


Figure 2. Various techniques used in polyphasic taxonomy for characterization of prokaryotes.

7. Polyphasic taxonomy

While the species is accepted as the basic unit of taxonomy (see Section 6), sub-species, strains and ecotypes occupy lower distinctive taxonomic levels for certain groups of organisms and are not mandatory for all. When classifying a new taxon, it is essential to describe phenotypic, genotypic and phylogenetic information as accurately as possible. This constitutes the polyphasic approach of taxonomy [1] and is shown in Figure 2. The phenotypic information comes from the colony characteristics, cell type, cell wall-type, pigmentation patterns, proteins and other chemotaxonomic markers while genotypic features are derived from the nucleic acids (DNA / RNA). Phylogenetic information is obtained from studying sequence similarities of the 16S rRNA or 23 S rRNA genes in case of bacteria and 18S rRNA in case of fungi. Many types of molecules are used for delineating and describing a taxon; some are mandatory (16S

rRNA genes, phenotypes, chemotaxonomy) while others are optional (amino-acid sequencing of certain protein products, DNA-DNA hybridization), unless required for appropriate description.

7.1. Phenotypic techniques

The phenotypic methods are all those that do not include the DNA/RNA sequencing or their typing methods. Study of morphological characteristics and chemotaxonomic profiles is broadly associated with phenotypic characterization.

7.1.1. Classical: Colony characteristics, biochemical and physiological analyses

The phenotypic features are the foundation for description of taxa. The morphological, biochemical and physiological characteristics provide in-depth information on a taxon. The morphology can include the colony characteristics (colour, shape, pigmentation, production of slime etc.). Further, the features of the cell are described as to shape, size, Gram reaction, extracellular material like capsule, presence of endospores, flagella presence and location, motility and inclusion bodies. Light microscopy is generally used to describe the broad cell features; however electron microscopy is recommended for high resolution images [18]. The biochemical and physiological features describe growth of the organism at different ranges of temperature, pH, salinity and atmospheric conditions, growth in presence of anti-microbial agents, production of various enzymes and growth in presence of different sole carbon and nitrogen sources [1]. These tests have to be carried out using standardized procedures to obtain results that are reproducible within and between laboratories [19].

7.1.2. Numerical taxonomy

Analyses of huge volumes of phenotypic data to derive meaningful relationships amongst a large number of microorganisms can be carried out using computer programs [20]. This system of analyses is called numerical taxonomy. Giving numerical weightage to each trait is followed by analyses of the data by the computer programs generating data matrices between each pair of isolates according to the degree of similarity. Based on the similarity data, cluster analyses are carried out (based on different algorithms) and dendrograms ('trees') are generated showing the overall pattern of similarity/dissimilarity amongst the various organisms being studied. While 16S rDNA sequences have garnered attention in recent times as sole means of bringing out the uniqueness of a species; numerical taxonomy (based on phenotypic traits of a large number of species) compares favourably with that of genotypic data and, indeed, is in alignment with the latter [1].

7.1.3. Cell wall composition

The peptidoglycan component of cell walls of bacteria does not provide much information except for classifying into Gram-positive, Gram-negative and acid-fast bacterial types. However, those in Gram-positive cells contain different types of peptidoglycan depending on the genus or species [21]. The peptidoglycan structure can be analysed by determining its type

(A or B), mode of cross-linking (whether it is directly linked or via interpeptide bridge and with amino acids in the bridge), and the composition of amino acids (especially the diaminoacid) of the side chain [18]. While the mode of cross-linkage can vary within a species and also between strains, the amino acid composition is common to all species within a genus. In archaea, pseudomurein is present where N-acetyl muramic acid is replaced by N-acetyl talosuronic acid [22].

7.1.4. Fatty acid analyses

Different types of lipids are present in bacterial cells. Polar lipids are present in the lipid bilayer of the cytoplasmic membrane. The diversity of polar lipids is known to be large and many are yet to be structurally elucidated. While in archaea, polar lipids are of types phospholipids, aminophospholipids, glycolipids and phosphoglycolipids, in bacteria, apart from the ones seen in archaea, there are also lipids derived from amino acids, capnines, sphingolipids (glycol or phosphosphingolipids) and hopanoids [18]. In Gram-negative bacteria, lipopolysaccharides are present in the outer membranes. The type of sugar present and the fatty acid type, the linkage of the fatty acid to the sugar (amide or ester linkage) provide information on characteristic of the cell [23–25]. However, determination of lipopolysaccharides is not routinely used in recent times. The total cellular fatty acids are extracted, esterified and the methyl ester content is analyzed by gas chromatography. Under standard conditions, this provides a reliable estimate of taxonomy up to genus and sometimes species level. The technique has been automated and the Sherlock MIS system (MIDI Inc.) has developed an extensive database. Though incomplete, it still is the most widely used system in recent times [18].

7.1.5. Isoprenoid quinones

The respiratory isoprenoid quinones are components that occur in cytoplasmic membranes of prokaryotes (archaea as well as bacteria). The naphthoquinones (with sub-types phylloquinone and menaquinone) and benzoquinones (ubiquinones, rhodoquinones and plastoquinones) are the major types of quinones. The variability they depict in their side chains in terms of length (5-15 isoprenoid units known till date), degree and position of saturation are of taxonomic significance and help in characterization to various levels of genus and species [26]. These features generally also mirror the 16S rDNA groupings. Isoprenoid ether-linked side chains are present in members of the archaea. They are of different types such as diethers, hydroxylated diethers, macrocyclic diethers, tetraethers, and polyol derivatives of the tetraether [18]. Non-isoprenoid based-ether-linked lipids are present in bacteria and can be straight chain or with simple branched side chains or with mono-unsaturated derivatives [18].

7.1.6. Other diagnostic methods

Other than the principal diagnostic methods described above, other techniques used to lesser levels or for comparison between species or strains comprise of the following:

a. Whole cell protein analyses, wherein protein is extracted from the cells and analysed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE). This can help in

comparison between related strains ([1] and references therein) and also shows congruency with DNA-DNA hybridization results [27]. However, since the identity of the protein bands is not known, this technique suffers from a drawback that is not associated with fatty acid analysis.

b. Polyamines are a group of compounds in the cytoplasm that provide stability to the DNA and maintain osmolarity in the cell. They are useful to distinguish above genus level and at species level too [28].

c. Cytochromes are associated with the cytoplasmic membrane and are involved in respiratory and photosynthetic electron transfer. They are 'heme' proteins with a 'heme' prosthetic group attached to a protein. These are not used alone in identification since there are limited types of distinct cytochromes [29].

7.1.7. Advanced spectroscopy and spectrometric methods

Many advanced analytical techniques such as Pyrolysis Mass Spectrometry, Fourier Transform Infrared spectroscopy and UV Resonance Raman Spectroscopy have been employed to determine chemical composition of bacterial cells, primarily the bioactive metabolites from drug discovery point of view and relate it to characteristics of microbes from which the metabolites are obtained [30, 31]. Pyrolysis Mass Spectrometry is a high-resolution technique, wherein microbial colonies are carefully picked and placed onto a iron-nickel foil, vacuum desiccated, heated rapidly and the pyrolysate bombarded with low-energy electrons. The ionized fragments are separated on the basis of their mass to charge ratio (*m*/*z*), detected and amplified by an electron multiplier [30]. Metabolites with *m*/*z* ratio of 51-200 constitute degradation products useful for taxonomical discrimination, since those with *m*/*z* less than 50 are produced by most biological materials and those with *m*/*z* ratio greater than 200 are not useful in discriminating taxons. The multivariate data is further analysed by Principal Components Analysis (PCA) to understand the variance while reducing the dimensionality of the data.

Fourier Transform Infrared spectroscopy (FTIR) is a simple and cost-efficient method and has been applied to identify discriminative features of strains. Most cellular components (fatty acids, proteins, carbohydrates and nucleic acids) can be analysed by this method to reveal strain-specific features [32]. Five IR spectral regions or 'windows': W1 (3000–2800 cm^{-1}) for fatty acids, W2 (1700–1500 cm^{-1}) for amide I and II bands of proteins and peptides, W3 (1500–1200 cm^{-1}) for a mixed region of fatty acid bending vibrations, proteins, and phosphate-carrying compounds, W4 (1200–900 cm^{-1}) for carbohydrates of cell walls and W5 (900–700 cm^{-1}) which is the 'fingerprint region' with unique absorbances specific for different taxa [33]. The differences in spectra are resolved using multivariate tools such as cluster analysis, discriminant analysis etc. [34].

UV Resonance Raman Spectroscopy (UVRR) uses a frequency of Raman spectra when biological materials are not subjected to fluorescent background while using IR or visible excitation [35]. Where an IR spectroscopy needs hundreds of cells, Raman spectroscopy does not require this [33, 36]. It can also be used with single cells, as demonstrated by Rösch et al.

[37], in combination with a data classification approach [38] and can also provide information about the Gram-type of a bacterium [39] as well as relate to moles G + C content [40].

7.2. Genotypic techniques

Modern taxonomy has been influenced by genetic methods and indeed, much of the classification and identification is predicated on specific gene sequences. All the techniques involving DNA or RNA fall under genotypic methods.

7.2.1. 16S rDNA-based analyses

The technique, which is very nearly a gold standard for taxonomic purposes today, is sequencing of the 16S rRNA gene of bacteria. The 23S rRNA gene sequence is also considered in many studies but lack of comprehensive databases for comparison is a drawback. Since the 16S rRNA is present in all bacteria, is functionally constant and is composed of conserved and variable regions, it has consistently served as a good taxonomic marker for deriving taxonomic relationships [1]. While it has proved to be the foundation for modern taxonomy, there are certain caveats and it has to be considered along with other techniques for formal identification purposes, especially at the species level. Generally universal sets of primers are used to amplify the 16S rRNA gene product gene (~1.5 kB). Subsequently, the amplified product is sequenced and quality of the sequence checked. The sequence is then aligned against high quality sequence data from curated databases [ARB (www.arb-home.de), RDP (http://rdp.cme.msu.edu/), SILVA (www.arb-silva.de) and LTP (www.arb-silva.de/projects/living-tree/)]. Multiple alignment programs such as CLUSTAL_X, CLUSTAL W, CLUSTAL X2, CLUSTAL W2, MEGA, T-COFFEE, MUSCLE) can also be used with manual editing [18].

As mentioned earlier, it has been shown that < 97% similarity of two 16S rDNA sequences implies a different species ([18] and references therein). This cut-off value is generally considered for ecological studies [41]. For actual taxonomic studies, a 98.5% cut-off value is considered [42]. However, the values should be based on high-quality and almost full-length sequences. When the similarity values are ~95%, other methods must also be included to justify the creation of a new genus. The descriptions must also provide information on the differences between the potentially new and existing genera. Subsequent to alignment, phylogenetic trees or dendrograms have to be constructed to reveal the taxonomic position of an organism. Different treeing algorithms such as maximum-parsimony, maximum-likelihood methods are preferred for evaluation of taxonomic position. Inclusion of type strain or type species is essential.

7.2.2. DNA base content

Determination of moles percent guanosine and cytosine constitutes a classical method of establishment of genomic content. This is now being used along with other genotyping methods to establish taxonomic position of an organism [1]. Within species, the G+C content ranges within 3% and within genera 10% [42]. Overall, the G+C content ranges from 24-76% in bacteria.

7.2.3. DNA-DNA hybridization

This method is an indirect measurement of sequence similarity between genomes. A cut-off value of 70% similarity is considered for establishment of species [1]. However, the method has to be reproducible between laboratories and performed under standardized conditions, which is often a drawback. Hence it is applied only where 16S rRNA gene sequences show similarity values above 98%. There have been reports where 16S rRNA gene sequence has shown 99% similarity and yet DNA-DNA hybridization values have been 60% or less. Hence, this method has to be used with caution and performed under highly standardized conditions.

7.2.4. Other genotyping methods

Earlier, sub-typing was done on the basis of biochemical profile (biotyping), serological profile (serotyping), phase susceptibility (phage typing) or antibiotic susceptibility. But currently DNA-typing methods are preferred due to their reproducibility, ease of performance and high level of discrimination between strains [1]. Genotyping methods such as Restriction Fragment Length Polymorphism (RFLP), Randomly Amplified Polymorphic DNA (RAPD), Amplified Fragment Length Polymorphism (AFLP), Amplified Ribosomal DNA Restriction Analysis (ARDRA), Repetitive Element-Polymerase Chain Reaction (REP-PCR), Ribotyping and Multi Locus Sequence Analyses (MLSA) are some of the newer methods to characterize a taxon.

RFLP was one of the earliest methods to be used and consisted of extraction of whole-genome DNA, restriction digestion using specific restriction enzymes and visualization of the DNA bands using gel electrophoresis. However, complex patterns can be generated, making comparison difficult. AFLP makes use of restriction analyses followed by PCR amplification to select for particular DNA fragments. Two restriction enzymes are used and fragments are amplified which can be genus and species-specific. Thus the method,though tedious, can be used for identification and typing purposes [1]. A PCR-based methodology makes use of random oligonucleotide primers 10 bases in length (RAPD), followed by amplification under specified conditions and analyses of the bands for similarity in size after gel electrophoresis. Ribotyping is another technique in which total genomic DNA is extracted, followed by restriction digestion and separation by electrophoresis [43]. Subsequently, the fragements are hybridized with a radiolabeled ribosomal operon probe (genes for 16S, 23S, tRNA and 5S rRNA) and analysed for presence of bands by autoradiography [44]. A simpler technique is amplifying the 16S rDNA (with or without spacer regions) using universal primers and restriction analysis (ARDRA). It can be used to discriminate at species level ([45]; [1] and references therein). Ribotyping and ARDRA have been shown to produce reproducible and congruent results in *Lactobacillus sp.* [46]. Consensus sequences complementary to repetitive sequences in the genomes of bacteria are used as primers, PCR-amplified and visualized as distinctive bands (REP-PCR). This method is discriminatory and rapid up to species or intraspecific level but is based on a library-based approach [47]. MLSA focuses on sequencing of housekeeping (6-8 protein-coding) genes and phylogenetic analyses of the same. It is a new method for characterizing a species [48]; however, databases are limited for realising the full-extent of utility of this method. Average Nucleotide Identity (ANI) of all orthologous genes in complete genome sequences has been proposed as a method to define species. Limited studies

show that 95% ANI corresponds roughly to 70% DNA-DNA similarity values [49]. However, more datasets are required to implement this method.

8. The eukaryotic microbes — Fungi

Fungi are important from industrial point of view as well as the increasing numbers of pathogens that are arising. Primarily, fungi are classified on the basis of appearance, the structure appearing above ground. The below-ground vegetative structures are difficult to identify [50]. The focus is on the asexual stage. The concept of species is somewhat difficult to interpret in the case of fungi since sexual mating does not occur in all fungi; meiosis occurs only in sexual fungi and even where mating occurs, the product of fusion could either be sterile or fertile. The biological concept of species can therefore be applied to sexual fungi. In the case of asexual fungi, similarities in characteristics provide a system for classification. Modern developments such as analyses of DNA by sequencing have brought about a paradigm shift in fungal taxonomy. The phylogenetic species concept is being favoured over the other definitions, especially where asexual fungi are considered. The evolutionary relationships amongst fungi have not been well-delineated. Traditionally, fungal phylogeny has been classified using morphology (involving primarily the fruiting body), cell wall composition [51], cell ultra-structure [52], cytology [53] and metabolism [54] and even based on the study of fossils [55]. Nomenclature also creates confusion since fungi can exist in sexual and asexual stages and can develop at different times in different substrates and relationship between the two states has to be established first.

Currently, modern developments underscoring the importance of phylogeny in bacterial classification has also been used for fungal taxonomy. Molecular genotyping methods such as restriction analysis of internal transcribed spacer (ITS) region, 18S rDNA and RFLP are being used to classify fungi [50]. However, the datasets are not extensive enough to permit effective identification [56]. Hibbett et al. [57] proposed a phylogenetic-based classification for fungi with 195 taxa. The 5S, 5.8S, 18S and 25S rDNAs are considered for phylogenetic studies. Of these, the 18S rDNA has been more extensively used for filamentous fungi [58]. The D1 and D2 variable regions of 25S rDNA are used for yeasts. Limited datasets are available where both 18S and 25 S rDNAs have been sequenced leading to difficulties in comparison [59]. The 25S rDNA allows for comparison to species level [60] while 18S rDNA has to be nearly completely sequenced to obtain pertinent information. The ITS region is suitable to reveal close relationships [61]. The 5S sequence provides information suitable for order level [62].

9. Uncultured microorganisms

Metagenomics has emerged as a promising field of interest, where identification of uncultured microorganisms is attempted. Since 99% of the microbial population is considered to be uncultivable, metagenomics assumes importance [63]. Next generation Sequencing (NGS) has

fuelled interest in this field. Classical metagenomics analyses samples by extracting environmental DNA followed by *de novo* sequencing, or amplification of 16S/18S rDNA using specific primers while functional metagenomics focuses on amplification of genes of interest (generally antibiotics, enzymes etc.) and their cloning into select target microorganisms to produce the metabolite of interest. Roche 454 pyrosequencing and Illumina are the most widely used NGS technologies [4]. DNA bar-coding approach is gaining popularity for assessing microbial diversity [64]. Though only limited datasets (especially for eukaryotic microbes) are currently available, the scenario is improving due to faster and cheaper sequencing methods. Inherent differences in microbial evolution rates, chimeric DNA sequences, artefacts generated during PCR or sequencing and non-universality of primers preclude derivation of a common threshold for taxonomic units [65]. However, bioinformatics handles some of these issues. Sequence quality is checked for series of Ns (nucleotides that are unresolved), errors in primer sequences are checked and verified, and sequences where length varies from the expected length [66] are assessed. Programs have been developed to remove pyrosequencing as well as PCR errors ([4] and references therein). After error-checking and trimming, the sequences are aligned, distance matrices calculated and used for clustering the Operational Taxonomic Units (OTUs) using programs such as MOTHUR [67]. OTUs represented by single sequences (singletons) are also documented and can overestimate diversity. Removal of singletons has not been shown to affect alpha diversity much [4] though more studies are required in this regard. Beta diversity remains conserved without singletons but diversity patterns may change in their presence [68].

10. Taxonomy of viruses

The definition of a virus 'species' is: "A virus species is a polythetic class of viruses that constitutes a replicating lineage and occupies a particular ecological niche" [69]. A virus isolate can refer to any virus as long as the virus has existed for some time. Viruses are not considered to be either prokaryotes or eukaryotes but have implication from health point of view; hence characterization of viruses has increased considerably. Where earlier, only electron microscopy was used, today sequencing of viral genomes constitutes advancements and the database is increasing. According to International Committee on Taxonomy of Viruses (ICTV), proposals are afoot to accept online descriptions of viral taxa based on taxonomical details such as : dsDNA, ssDNA, rtDNA, rtRNA, dsRNA, ssNRNA, ssPRNA, SAT (Satellites), VIR (Viroids), UN (unassigned).

11. Conclusions & perspectives

Zinger et al. state (see Pg. 2 of Ref. [4]): "In its broadest meaning, measuring biodiversity consists of characterizing the number, composition and variation in taxonomic or functional units over a wide range of biological organizations (from genes to communities)". The taxonomical classification of microorganisms has been difficult due to their small size, short

generation times and confounded by genetic exchange between unrelated organisms. These limitations have been largely overcome by modern developments of sequencing technologies and the recognition of rDNA sequences as a cornerstone for identification purposes. Overall, it is important to recognize that microbial diversity is intricately linked to its environment and this correlation has to be established by description of environmental parameters whenever sampling is carried out. It is also important to study the phenotypic characteristics and link them to the observations obtained from genotyping techniques. The link between habitat and diversity then becomes easier to understand for future studies.

Author details

Indira P. Sarethy[1], Sharadwata Pan[2] and Michael K. Danquah[3*]

*Address all correspondence to: mkdanquah@curtin.edu.my

1 Department of Biotechnology, Jaypee Institute of Information Technology, Noida (U.P.), India

2 Department of Chemical Engineering, Indian Institute of Technology Bombay, Mumbai, India

3 Department of Chemical and Petroleum Engineering, Curtin University of Technology, Sarawak, Malaysia

References

[1] Vandamme P, Pot B, Gillis M, De Vos P, Kersters K, Swings J. Polyphasic taxonomy, a consensus approach to bacterial systematics, Microbiological Reviews 1996; 60 (2): 407-438.

[2] Fierer N. Microbial biogeography: patterns in microbial diversity across space and time. In: Zengler K. (ed.), Accessing Uncultivated Microorganisms: from the Environment to Organisms and Genomes and Back. Washington DC: ASM Press; 2008. p95-115.

[3] Kassen R, Rainey P. The ecology and genetics of microbial diversity, Annual Review of Microbiology 2004; 58: 207 – 231.

[4] Zinger L, Gobet A, Pommier T. Two decades of describing the unseen majority of aquatic microbial diversity, Molecular Ecology 2012; 21(8): 1878–1896.

[5] Madigan MT, Martinko JM, Stahl DA, Clark DP. Brock biology of microorganisms 13th Edition; Ch. 16, pp. 446-474; Benjamin Cummings; 2010.

[6] Woese CR. (2000). Interpreting the universal phylogenetic tree, Proceedings of National Academy of Sciences U.S.A. 2000; 97(15): 8392-8396.

[7] Boucher Y, Douady CJ, Papke RT, Walsh DA, Boudreau MER, Nesbo CL, Case RJ, Doolittle WF. Lateral gene transfer and the origins of prokaryotic groups, Annual Review of Genetics 2003; 37: 283 – 328.

[8] Fierer N, Lennon JT. The Generation and Maintenance of Diversity in Microbial Communities, American Journal of Botany 2011; 98(3): 439–448.

[9] Jain R, Rivera MC, Moore JE, Lake JA. Horizontal Gene Transfer accelerates genome innovation and evolution, Molecular Biology and Evolution 2003; 20(10): 1598–1602.

[10] Achtman M, Wagner M. Microbial diversity and the genetic nature of microbial species, Nature Reviews 2008; 6: 431-440.

[11] Stackebrandt E, Frederiksen W, Garrity GM, Grimont PA, Kämpfer P, Maiden MC et al. Report of the ad hoc committee for the re-evaluation of the species definition in bacteriology, International Journal of Systematic and Evolutionary Microbiology 2002; 52(Pt 3): 1043–1047.

[12] Wayne LG, Brenner DJ, Colwell RR, Grimont PAD, Kandler O, Krichevsky MI et al. Report of the ad hoc committee on reconciliation of approaches to bacterial systematic, International Journal of Systematic Bacteriology 1987; 37(4): 463–464.

[13] Staley JT, Krieg NJ. Classification of prokaryotic organisms: an overview. In: Krieg NR, Holt JG. (eds.), Bergey's Manual of Systematic Bacteriology Vol. 1; Baltimore: The Williams & Wilkins Co. 1984. p1–3.

[14] Stackebrandt E, Ebers J. Taxonomic parameters revisited: tarnished gold standards, Microbiology Today 2006; 33: 152–155.

[15] Murray RG, Stackebrandt E. Taxonomic note: implementation of the provisional status Candidatus for incompletely described prokaryotes, International Journal of Systematic Bacteriology 1995; 45(1): 186–187.

[16] Koeppel A, Perry EB, Sikorski J, Krizanc D, Warner A, Ward DM. et al. Identifying the fundamental units of bacterial diversity: a paradigm shift to incorporate ecology into bacterial systematics, Proceedings of National Academy of Sciences USA 2008; 105(7): 2504–2509.

[17] Lapage SP, Sneath PHA, Lessel EF, Skerman VBD, Seeliger HPR, Clark WA. (eds.) International Code of Nomenclature of Bacteria (1990 Revision). Bacteriological Code. Washington, DC: American Society for Microbiology, 1992.

[18] Tindall BJ, Rosselló-Móra R, Busse HJ, Ludwig W, Kämpfer P. Notes on the characterization of prokaryote strains for taxonomic purposes, International Journal of Systematic and Evolutionary Microbiology 2010; 60(Pt 1): 249–266.

[19] On SL, Holmes B. Reproducibility of tolerance tests that are useful in the identification of campylobacteria, Journal of Clinical Microbiology 1991; 29(9): 1785–1788.

[20] Sneath P. Numerical taxonomy. In: Krieg NR, Holt JG. (eds.), Bergey's Manual of Systematic Bacteriology Vol. 1.; Baltimore: The Williams & Wilkins Co. 1984. p111-118.

[21] Schleifer KH, Ludwig W. Phylogenetic relationships of bacteria. In: Fernholm B, Bremer K, Jornvall H. (eds.), The hierarchy of life. Amsterdam: Elsevier Science Publishers B. V. 1989. p103-117.

[22] König H, Kralik R, Kandler O. Structure and modifications of the pseudomurein in Methanobacteriales, Zentralblatt fur Bakteriologie.Allgemeine Angewandte und Okologische Microbiologie Abt.1 Orig.C Hyg. 1982; 3(2): 179–191.

[23] Hase S, Rietschel ET. Isolation and analysis of the lipid A backbone. Lipid A structure of lipopolysaccharides from various bacterial groups. European Journal of Biochemistry 1976; 63(1): 101–107.

[24] Weckesser J, Mayer H. Different lipid A types in lipopolysaccharides of phototrophic and related non-phototrophic bacteria, FEMS Microbiology Letters 1988; 54(2): 143–153.

[25] Mayer H, Masoud H, Urbanik-Sypniewska T, Weckesser J. Lipid A composition and phylogeny of Gram-negative bacteria, Bull Japan Federation Culture Collections 1989; 5: 19–25.

[26] Collins MD, Jones D. Distribution of isoprenoid quinine structural types in bacteria and their taxonomic implications, Microbiological Reviews 1981; 45(2): 316–354.

[27] Costas M. (1992). Classification, identification, and typing of bacteria by the analysis of their one-dimensional polyacrylamide gel electrophoretic protein patterns. In: Charmbach A, Dunn MJ, Radola BJ. (eds.), VCH, Weinheim, Germany: Advances in Electrophoresis Vol. 5. 1992. p351–408.

[28] Busse J, Auling G. Polyamine pattern as a chemotaxonomic marker within the Proteobacteria, Systematic and Applied Microbiology 1988; 11(1): 1–8.

[29] Jones D, Krieg NR. In: Krieg NR. (ed.), Bergey's Manual of Systematic Bacteriology Vol. 1, Baltimore: Williams and Wilkins 1984. p15-18.

[30] Goodacre R, Kell DB. Pyrolysis mass spectrometry and its applications in biotechnology, Current Opinion in Biotechnology 1996; 7(1): 20-28.

[31] Magee J. Whole-organism fingerprinting. In: Goodfellow M, O'Donnell AG. (eds.), Handbook of new bacterial systematic. London: Academic Press Ltd. 1993. p.383-427.

[32] Dziuba B, Nalepa B. Propionic and LAB Identification by FTIR and ANN, Food Technology and Biotechnology 2012; 50(4): 399–405.

[33] Naumann D, Helm D, Labischinski H, Giesbrecht P. The characterization of microorganisms by Fourier-transform infrared spectroscopy (FT-IR). In: Nelson WH. (ed.), Modern techniques for rapid microbiological analysis. New York: VCH Publishers 1991. p.43-96.

[34] Samelis J, Bleicher A, Delbès-Paus C, Kakouri A, Neuhaus K, Montel MC. FTIR-based polyphasic identification of lactic acid bacteria isolated from traditional Greek Graviera cheese, Food Microbiology 2011; 28(1): 76–83.

[35] Spiro TG. Resonance Raman spectroscopy: a new structural probe for biological chromophores. Accounts of Chemical Research 1974; 7(10): 339-344.

[36] Naumann D, Keller S, Helm D, Schultz C, Schrader B. FT-IR spectroscopy and FT-Raman spectroscopy are powerful analytical tools for the noninvasive characterization of intact microbial cells, Journal of Molecular Structure 1995; 347: 399–405.

[37] Rösch P, Harz M, Schmitt M, Peschke KD, Ronneberger O, Burkhardt H. et al. Chemotaxonomic identification of single bacteria by micro-Raman spectroscopy: application to clean-room-relevant biological contaminations, Applied Environmental Microbiology 2005; 71(3): 1626-1637.

[38] López-Díez EC, Goodacre R. Characterization of microorganisms using UV resonance Raman spectroscopy and chemometrics, Analytical Chemistry 2004; 76(3): 585-591.

[39] Manoharan R, Ghiamati E, Dalterio RA, Britton KA, Nelson W H, Sperry JF. UV resonance Raman spectra of bacteria, bacterial spores, protoplasts, and calcium dipicolinate, Journal of Microbiological Methods 1990; 11(1): 1-15.

[40] Nelson WH, Manoharan R, Sperry JF. UV Resonance Raman Studies of Bacteria, Applied Spectroscopic Reviews 1992; 27(1): 67-124.

[41] Pedrós-Alió C. Marine microbial diversity: can it be determined? Trends in Microbiology 2006; 14(6): 257–263.

[42] Stackebrandt E, Liesack W. Nucleic acids and classification. In: Goodfellow M, O'Donnell AG. (eds.), Handbook of new bacterial systematics. London: Academic Press Ltd. 1993. p.151-194.

[43] Grimont F, Grimont PAD. Ribosomal ribonucleic acid gene restriction as potential taxonomic tools, Annales de l'Institut Pasteur/Microbiologie 1986; 137B(2): 165-175.

[44] Bouchet V, Huot H, Goldstein R. Molecular genetic basis of ribotyping, Clinical Microbiology Reviews 2008; 21 (2): 262-273.

[45] Kostman JR, Edlind TD, LiPuma JJ, Stull TL. Molecular epidemiology of Pseudomonas cepacia determined by polymerase chain reaction ribotyping, Journal of Clinical Microbiology 1992; 30(8): 2084-2087.

[46] Miteva V, Boudakov I, Ivanova-Stoyancheva G, Marinova B, Mitev V, Mengaud J. Differentiation of Lactobacillus delbrueckii subspecies by ribotyping and amplified

ribosomal DNA restriction analysis (ARDRA), Journal of Applied Microbiology 2001; 90(6): 909-918.

[47] Kon T, Weir SC, Howell ET., Lee H, Trevors JT. Repetitive element (REP)-polymerase chain reaction (PCR) analysis of Escherichia coli isolates from recreational waters of southeastern Lake Huron, Canadian Journal of Microbiology 2009; 55(3): 269-276.

[48] Hanage WP, Fraser C, Spratt BG. Sequences, sequence clusters and bacterial species, Philosophical Transactions of the Royal Society of London B 2006; 361P: 1917–1927.

[49] Goris J, Konstantinidis KT, Klappenbach JA, Coenye T, Vandamme P, Tiedje JM. DNA–DNA hybridization values and their relationship to whole-genome sequence similarities, International Journal of Systematic and Evolutionary Microbiology 2007; 57(Pt 1): 81–91.

[50] Horton TR. Molecular approaches to ectomycorrhizal diversity studies: variation in ITS at a local scale. Plant and Soil 2002; 244(1-2), 29– 39.

[51] Bartnicki-Garcia S. Cell wall composition and other biochemical markers in fungal phylogeny. In: Harbone JB. (ed.), Phytochemical phylogeny. London: Academic Press, Ltd. 1970. p.81-103.

[52] Heath IB. Nuclear division: a marker for protist phylogeny. Progress in Protistology 1986; 1: 115–162.

[53] Taylor FJR. Problems in the development of an explicit hypothetical phylogeny of the lower eukaryotes, BioSystems 1978; 10(1-2): 67–89.

[54] Vogel HJ. Distribution of lysine pathways among fungi: evolutionary implications, The American Naturalist 1964; 98(903): 435–446.

[55] Hawksworth DL, Kirk PM, Sutton BC, Pegler DN. Ainsworth and Bisby's dictionary of the fungi, 8th ed. International Mycological Institute, Egham, U.K. 1995.

[56] Kirk JL, Beaudette LA, Hart M, Moutoglis P, Klironomos JN, Lee H. et al. Methods of studying soil microbial diversity, Journal of Microbiological Methods 2004; 58 (2): 169-188.

[57] Hibbett DS, Binder M, Bischoff JF, Blackwell M, Cannon PF, Eriksson OE. et al. A higher-level phylogenetic classification of the Fungi, Mycological Research 2007; 111 (5): 509-547.

[58] Woese CR, Kandler O, Wheelis ML. Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya, Proceedings of National Academy of Sciences USA 1990; 87(12): 4576–4579.

[59] Guarro J, Gené J, Stchigel AM. Developments in fungal taxonomy, Clinical Microbiology Reviews 1999; 12 (3): 454-500.

[60] Guého E, Improvisi L, Christen R, de Hoog GS. Phylogenetic relationships of Cryptococcus neoformans and some related basidiomycetous yeasts determined from parti-

al large subunit rRNA sequences, Antonie van Leeuwenhoek, International Journal of General and Molecular Microbiology 1993; 63(2):175–189.

[61] de Hoog GS, Gerrits van den Ende AH. Molecular diagnostics of clinical strains of filamentous Basidiomycetes, Mycoses 1998; 41(5-6): 183–189.

[62] Walker WF, Doolittle WF. Redividing the basidiomycetes on the basis of 5S rRNA sequences, Nature 1982; 299: 723–724.

[63] Brady SF. Construction of soil environmental DNA cosmid libraries and screening for clones that produce biologically active small molecules. Nature Protocols 2007; 2: 1297-1305.

[64] Valentini A, Pompanon F, Taberlet P. DNA barcoding for ecologists, Trends in Ecology and Evolution 2009; 24(2): 110–117.

[65] Thornhill DJ, Lajeunesse TC, Santos SR. Measuring rDNA diversity in eukaryotic microbial systems: how intragenomic variation, pseudogenes, and PCR artifacts confound biodiversity estimates, Molecular Ecology 227; 16(24): 5326–5340.

[66] Huse SM, Huber JA, Morrison HG, Sogin ML, Mark Welch D. Accuracy and quality of massively parallel DNA pyrosequencing, Genome Biology 2007; 8: R143.

[67] Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB. et al. Introducing mothur: open-source, platform-independent, community supported software for describing and comparing microbial communities, Applied and Environmental Microbiology 2009; 75(23): 7537–7541.

[68] Agogué H, Lamy D, Neal PR, Sogin ML, Herndl GJ. Water mass-specificity of bacterial communities in the North Atlantic revealed by massively parallel sequencing, Molecular Ecology 2011; 20(2): 258–274.

[69] Fauquet CM, Fargette D. International Committee on Taxonomy of Viruses and the 3,142 unassigned species, Virology Journal 2005; 2: 64.

IntechOpen