Journal of Microbiological Methods 103 (2014) 80–100

Contents lists available at ScienceDirect

Journal of Microbiological Methods

journal homepage: www.elsevier.com/locate/jmicmeth

Review

Understanding molecular identification and polyphasic taxonomic approaches for genetic relatedness and phylogenetic relationships of microorganisms

Surajit Das *, Hirak R. Dash, Neelam Mangwani, Jaya Chakraborty, Supriya Kumari

Laboratory of Environmental Microbiology and Ecology (LEnME), Department of Life Science, National Institute of Technology, Rourkela 769 008, Odisha, India

ARTICLE INFO

Article history:
Received 21 November 2013
Received in revised form 22 May 2014
Accepted 22 May 2014
Available online 2 June 2014

Keywords:
Bacteria
Fungi
Molecular identification
Gene sequencing
Polyphasic taxonomy

ABSTRACT

The major proportion of earth's biological diversity is inhabited by microorganisms and they play a useful role in diversified environments. However, taxonomy of microorganisms is progressing at a snail's pace, thus less than 1% of the microbial population has been identified so far. The major problem associated with this is due to a lack of uniform, reliable, advanced, and common to all practices for microbial identification and systematic studies. However, recent advances have developed many useful techniques taking into account the house-keeping genes as well as targeting other gene catalogues (*16S rRNA*, *rpoA*, *rpoB*, *gyrA*, *gyrB* etc. in case of bacteria and 26S, 28S, β-tubulin gene in case of fungi). Some uncultivable approaches using much advanced techniques like flow cytometry and gel based techniques have also been used to decipher microbial diversity. However, all these techniques have their corresponding pros and cons. In this regard, a polyphasic taxonomic approach is advantageous because it exploits simultaneously both conventional as well as molecular identification techniques. In this review, certain aspects of the merits and limitations of different methods for molecular identification and systematics of microorganisms have been discussed. The major advantages of the polyphasic approach have also been described taking into account certain groups of bacteria as case studies to arrive at a consensus approach to microbial identification.

Contents

Abbreviations: VBNC, viable but nut culturable; MLST, multi-locus sequence typing; MALDI-TOF, matrix assisted laser desorption ionisation-time of flight; MS, mass spectrometry; SMM, shotgun mass mapping; PCR, polymerase chain reaction; ERIC, enterobacterial repetitive intergenic consensus sequences; REP, repetitive element palindromic sequences; bp, base pair(s); HS, heat shock; ATP, adenosine triphosphate; ST, sequence type; SIP, stable isotope probing; IPP, intact protein profiling; LC, liquid chromatography; FAME, fatty acid methyl ester(s); nMDS, non-metric multidimensional scaling; ANOSIM, analysis of similarities; DFA, discriminant function analysis; FID, flame ionisation detection; MIS, microbial identification system; ATR, attenuated total reflection; FTIR, fourier transform infrared; DCIMS, direct chemical ionisation mass spectrometry; THM, thermal hydrolysis/methylation; TMAH, tetra methyl ammonium hydroxide; rDNA, ribosomal DNA; ITS, internal transcribed spacer; SSU, small subunit; LSU, large subunit; REVSEQ, reverse and complementary sequence; SPSS, statistical package for the social sciences; DGGE, denaturing gradient gel electrophoresis; TGGE, temperature gradient gel electrophoresis; CGGE, constant gradient gel electrophoresis; FC, flow cytometry; FACS, fluorescence activated cell sorting; TEFAP, Tag-encoded FLX-amplicon pyrosequencing; bTEFAP, bacterial tag encoded FLX titanium amplicon pyrosequencing; CBC, citrus bacterial canker; SDS, sodium dodecyl sulphate; PAGE, polyacrylamide gel electrophoresis; RPLP, restriction fragment length polymorphism; IVS, intervening sequences; TTC, triphenyl-tetrazolium chloride; ATCC, American type culture collection; RPHPLC, reverse phase high performance liquid chromatography; TOF-MS, time-of-flight mass spectra; IF, initiation factor; c553, cytochrome533; Trx, thioredoxin; VP, voges proskauer; ARDRA, amplified rDNA restriction analysis.

* Corresponding author. Tel.: +91 661 2462684; fax: +91 661 2462022.
E-mail addresses: surajit@nitrkl.ac.in, surajit@myself.com (S. Das).

http://dx.doi.org/10.1016/j.mimet.2014.05.013
0167-7012/

1. Introduction

Microorganisms are omnipresent, extraordinarily diverse and perform specialized roles in the environment. They impart many harmful effects such as spoilage and health damages besides their beneficial role in the production of antimicrobials and bioactive compounds, bioremediation of toxic chemicals and involvement in food, beverages and pharmaceutical industries. Microbes have been present for over 3.8 billion years, however, their existence became obvious 300 years ago and yet still only <1% of total microbes are known. Taxonomic information of an unknown microbe is highly essential to establish its biodiversity, relationship among other organisms in the ecosystem and its functional aspects (Gevers et al., 2005). Thus, proper isolation and identification is mandatory before deducing the novel characteristic features of any microbial isolate. With the advent of genomics, the complexity of microbial world is largely understandable. In this regard, recent advancements in microbial systematics have led to a 'polyphasic taxonomic approach' which aims to generate all genotypic, phenotypic and phylogenetic information of a microbial taxon (Vandamme et al., 1996). The resulting biological data of the organisms find relevance and implications in biotechnological research such as production of bioactive molecules, microbial transformations, ecology and bioremediation using microbial isolate or consortium.

The prevalent conventional techniques are not sufficient to provide a complete draft for microbial taxonomy as these conventional techniques describe only shape, colour, size, staining properties, motility, host-range, pathogenicity and assimilation of carbon sources (Prakash et al., 2007). However, a comprehensive approach is required to furnish

the information and subsequent derivation of a microbial lineage. Considerable advancement in the study of microbial taxonomy in the past few decades has taken place such as chemotaxonomy (Minnikin et al., 1975), numerical taxonomy and DNA–DNA hybridization (Johnson, 1991), DNA amplification and sequencing (Konstantinidis and Tiedje, 2007) and whole genome sequencing (Janssen et al., 2003). Despite of this tremendous progress and development of genome based techniques, any individual technique cannot be relied upon solely as a source of taxonomic information. Another problem associated with microbial taxonomic analysis is the existence of some common taxa that cannot be cultured probably due to their protection strategy — viable but not culturable (VBNC) state (Rompre et al., 2002). In contrast, culture independent techniques targeting a single gene (e.g. 16S rRNA gene) also cannot provide adequate resolution for proper microbial identification (Jany and Barbier, 2008; Koser et al., 2012). Hence, no single tool infers definitive assessment of the microbial community, however, it can be studied under two headings i.e. culture dependent and culture independent approaches.

Therefore, the use of a polyphasic approach involving a combination of molecular biology techniques and conventional microbiological methods is necessary to obtain a better understanding of microbial diversity. However, there are many techniques used to determine "genetic relatedness" of microorganisms but this review aims to discuss the techniques used for identification using the polyphasic taxonomic approach with the pros and cons of the currently available techniques used for molecular identification of isolates.

2. Why use the molecular methods for identification?

Microorganisms present in the environment can be enumerated, isolated and characterised by various culture dependent classical techniques such as pour plate and spread plate methods followed by Gram's staining and biochemical tests to decipher their physiological characteristics. Based upon colony morphology on growth medium, using microscopic observation and biochemical tests isolated bacteria can be assigned to specific genera. However, these techniques are time consuming as well as dependent upon many environmental factors (Rastogi and Sani, 2011). Therefore, the advanced techniques like sequence based, gel based and protein based systems have become advantageous due to their fast reactions, high specificity and less chance of error. The development of these advanced techniques has changed drastically the definition of microorganisms as in many cases sequence diversity within a species reveals multiple sequence clusters that are ecologically distinct (Cohan, 2002). By using the house-keeping gene sequences for phylogenetic studies, domains of Archaea, Bacteria and Eukarya have been developed in contrast to the traditional classification of organisms i.e. prokaryotes and eukaryotes. In addition, fungi are the most laborious to work with due to their slow growing nature and the conventional practices of identification procedures lead to many ambiguities. Thus, a high level of expertise is necessary for fungal identification (Siqueira and Rocas, 2005). Though the technologies have progressed far beyond the level needed for identification of microorganisms in most laboratories, each procedure has its own lacuna.

3. Catalogue of gene targets for microbial identification

To document different phylogenetic lineages for correlation within genomes of microbes, specific genotypic markers are essential. Varieties of house-keeping genes are targeted in different microbial entities to confer precise genetic fingerprints. There are several genetic markers for rapid cataloguing of gene families having conserved sequences falling within definite microbial clades (Fig. 1). Different genetic targets for microbial identification are *16S rDNA*, *gyrB*, *rpoA*, *rpoB*, *rpoC*, *rpoD* genes, etc. (Glazunova et al, 2009). 16S rRNA gene is the primary gene target for identification of bacteria as the gene sequences contain conserved, variable and hypervariable regions (Clarridge, 2004). Genetic markers such as DNA metabolic enzyme *gyrA* and *gyrB* are also important for identifying microbial diversity in the environment. *rpoA*, *rpoB*, *rpoC*, *rpoD* are also conserved genes in the ribosomal region effectively used as molecular markers. Genes for DNA repair mechanisms like *recA* and *recN* present in maximum microbial genera have high sequence similarity and are feasible markers in molecular identification (Karlin et al., 1995). Another gene *ppk1* encoding polyphosphate kinase can be an efficient marker for molecular identification, the functions of which includes inhibition of RNA degradation, activation of Lon protease, membrane channel formation, stress resistance, motility, virulence as well as quorum sensing (Shi et al., 2004). *dnaJ* gene is used as a molecular marker for alpha-proteobacteria for molecular diversity analysis (Alexandre et al., 2008). Many other genes which are used as molecular markers to study the diversity of different groups of microorganisms are: *gapA* for *Escherichia coli* and *Mycoplasma gallisepticum*, *pyrH* for classifying *Vibrio vulnificus* and *Vibrio* sp., *luxABE* gene for *Aliivibrio logei*, *Aliivibrio salmonicida* and *Photobacterium mandapamensis*, *ectBC* or classifying *Halomonas variabilis* and *Halomonadaceae*, *dnaA* for *Mycobacterium* sp. and *Rhizobium meliloti*, *pheS* for classifying *Enterococcus* sp. and *Lactobacilli* and *dnaK* gene for *Ensifer* sp., *Sinorhizobium* and *Nitrosomonas europaea.*

Multi-locus sequence typing (MLST) is considered to be the "gold standard" for typing of bacterial and fungal species. The MLST scheme identifies internal nucleotide sequences of approximately 400 to 500 bp in multiple house-keeping genes. Unique sequences (alleles) are assigned as a random integer number and a unique combination of alleles at each locus. This is designated as an "allelic profile" which specifies the sequence type (Larsen et al., 2012).

There are also certain protein and fatty acid based cataloguing techniques used for deducing better functional diversity among microbial flora. MALDI-TOF MS technique proves to be a dependable approach for molecular identification (Sauer et al., 2008). Shotgun mass mapping (SMM) and protein stable isotope probing are also primary techniques for creating diversity maps between microbial communities. Fatty acid composition in microbial species has also been identified as an important biomarker for taxonomic studies due to slight variations in bond formation, chain length and functional group binding (Dawyndt et al., 2006).

There are also various PCR based fingerprinting techniques like ERIC, REP, and BOX-PCR for creating genotypic imprints (Frye and Healy, 2006). 26S, 28S and β-tubulin genes are studied for representing molecular relationship within fungal species.

4. Culture dependent identification of bacteria and fungi

4.1. Sequence based techniques for bacteria

Variations in house-keeping gene sequences can be explored for the identification of bacterial species which may be attributed to the fact that, the sequence of these genes vary slowly during the process of evolution. To get around the pitfalls of the conventional techniques, amplification and sequencing of DNA offers a good alternative. In this context, the most common house-keeping genes used for bacterial identification purpose are *16S rRNA*, *gyrB*, *rpoA*, *rpoB*, *rpoC*, *rpoD* etc. (Glazunova et al., 2009). Intraspecies variations, sequevars and subspecies variant genotypes are generally used to express micro heterogeneity within a species which denotes the difference of only a few base pairs of the sequences or less than 0.5% (Clarridge, 2004). Thus, gene sequence analyses discriminate several strains of microorganisms in a more precise manner rather than the conventional practices. As in the case of 80% of isolates, a close match occurs within a described species and 10% represent a new species within a genus and the remaining 10% may represent novel taxa which can only be deduced through the sequence based identification procedures in contrast to the conventional practices (Drancourt et al., 2000). The steps used for the sequence based

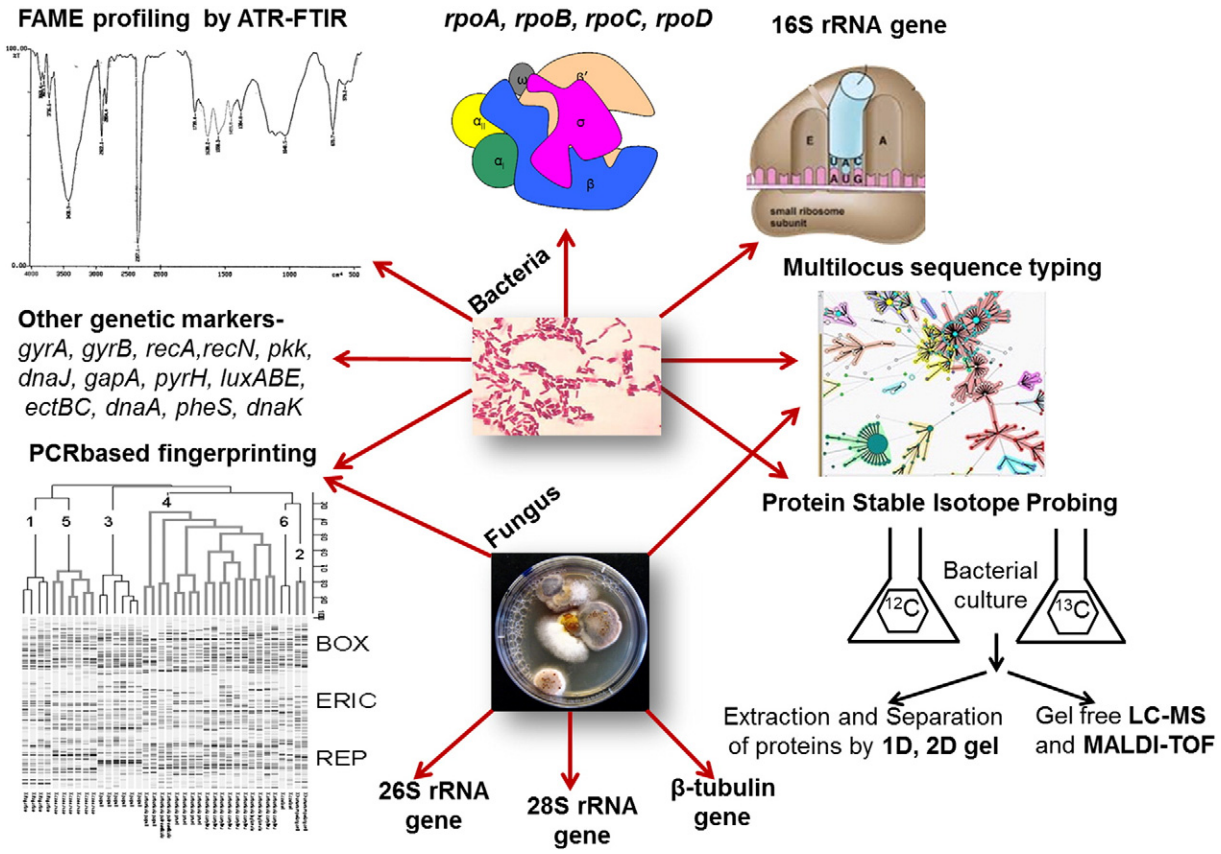

Fig. 1. Catalogue of gene targets for molecular identification of bacteria and fungi.

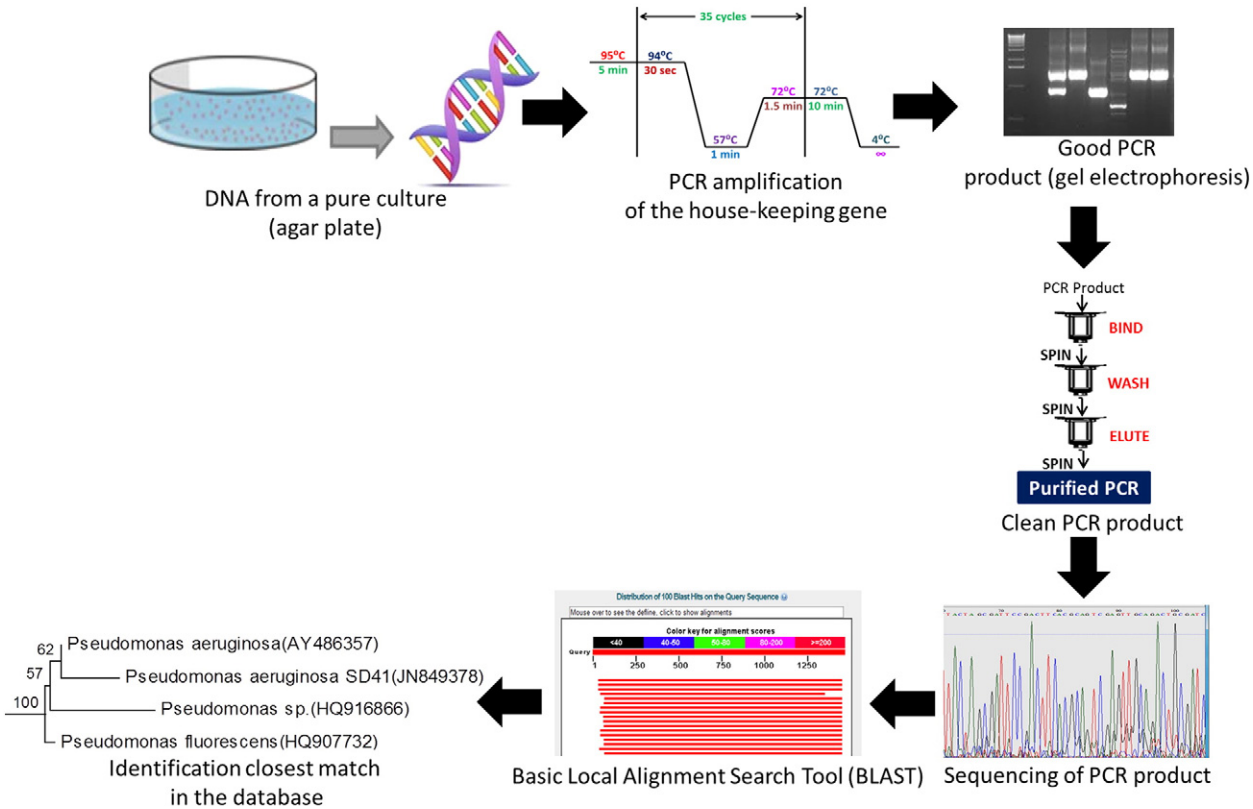

Fig. 2. Flow diagram showing the steps involved in bacterial identification using housekeeping sequencing approach.

identification are the same in spite of the variation of the gene sequence (Fig. 2).

4.1.1. 16S rRNA gene sequence based identification

As huge varieties of bacteria are present in the environment, the range and complexity of the techniques to be utilised for their identification purpose is highly bewildering (Spratt, 2004). However, the use of nucleotide sequence data from 16S rRNA gene has been regarded to be the most suitable practice not only to identify but also to draw the phylogenetic relationship for all microorganisms on earth. The reasons behind the use of 16S rRNA gene to be utilised for identification purpose include i) occurrence of the gene in all organisms performing the same function, ii) the gene sequence is conserved sufficiently containing conserved, variable and hyper-variable regions, and iii) around 1500 bp of sequence size which is relatively easy to sequence and large enough to contain sufficient information for identification and analysis of phylogeny (Clarridge, 2004). This technique became popular only after sufficient deposition of 16S rRNA gene sequences in the database as well as the availability of suitable primers for gene amplification (Table 1).

There are a number of advantages as well as disadvantages regarding the comparative sequencing and analysis of the 16S rRNA gene. The major advantage of this technique includes rapid and accurate identification. In conventional techniques certain groups of bacteria are difficult to identify and require special equipment and expertise for their identification such as gas chromatography and mass spectroscopy for anaerobes whereas sequence based identification does not rely on them (Clarridge, 2004). Certain slow growing bacteria are present which take considerable time for their growth and hence their identification by conventional techniques becomes tedious. In this regard, 16S rRNA gene sequencing represents a unique technology yielding the result in considerably less time. Similarly, some isolates with atypical morphological characteristics can also be detected using this technique (Gee et al., 2004). Novel bacterial genera as well as the discovery of new species can be accomplished by combining the phenotypic characteristics, 16S rDNA sequences, sequences of other genes as well as the results of DNA–DNA hybridization tests (Staley, 2006).

4.1.2. rpoA, rpoB, rpoC and rpoD gene sequencing

The most common entity among ribosomal RNA encoding genes is their universality and *rpoB* encodes the β-subunit of the bacterial RNA polymerase (Mollet et al., 1997). The comparison of *rpoB* sequences have been explored for the phylogenetic analysis of archea and bacteria as well (Rowland et al., 1993; Klenk and Zillig, 1994). However, the comparison of *rpoB* sequences for species based identification has yet not been explored completely. Though it can provide a useful tool for molecular identification of bacterial species, the incomplete database constituting lesser amount of sequences which does not cover as many species is of great concern. Some recent developments regarding the *rpoB* based identification system includes the identification of four hyper-variable zones along the gene sequences, i.e. two zones on the *rpoB* gene of 350 and 450 base pairs (bp) in size and two more flanking spacers at both the ends of the same gene, i.e. *rpoL–rpoB*, 301–310 bp, and *rpoB–rpoC*, 86–177 bp (Gundi et al., 2009). Hence, species-specific primers can be designed from the hyper-variable regions among the gene for rapid identification of any bacterial species as this gene has been shown to be more discriminative than the 16S rRNA gene. Similarly, *rpoB* gene is found to exist as a single copy in the bacterial genome which limits the nonspecific amplification and chances of experimental errors.

Similarly, *rpoA* gene sequences that code for the α-subunit of RNA polymerase have been shown to be useful for many bacterial species like *Enterococcus* sp. and lactic acid bacteria (Naser et al., 2005). However, *rpoA* gene sequences possess a lower resolution having a 5% interspecies gap and 2% level of intra-species variation (Naser et al., 2007). In some cases, *rpoC* gene is also taken into account for the identification of many species of bacteria. Just like *rpoB* gene, *rpoC* gene codes for beta subunits of DNA dependent RNA polymerase and is organised in an operon where *rpoB* always precedes *rpoC* (Severinov et al., 1997). There are many reports of using *rpoC* gene sequences along with other gene sequences for determining the relatedness among rifampin resistant *Mycobacterium tuberculosis* (de Vos et al., 2013). *rpoD* gene in bacteria codes for a 70-kDa protein and the conserved amino acid regions of σ^{70} is found in Pp σ^{70}. It is transcribed as a monocistronic mRNA of 2.1 kb which is under the control of HS response (Fujita et al., 1995). The conserved regions of *rpoD* sequences have been employed widely for the identification procedures of many Actinobacteria species like *Frankia* (Berneche-D'Amours et al., 2011) as well as *E. coli* (Burton et al., 1981). The primer sequences that have been used for the amplification and further analysis of *rpoD* are listed in Table 2.

4.1.3. gyrA and gyrB sequencing

Among the DNA metabolic enzymes altering its topology, type II DNA topoisomerases/DNA gyrase is essential and ubiquitous. DNA gyrase is encoded by both *gyrB* and *gyrA* which belongs to the single gene family (Onodera and Sato, 1999). The presence of highly conserved motifs in these gene sequences provides a useful tool for the designing of universal primers for the study of bacterial identification and diversity. As higher genetic variation is observed among the protein coding genes, they can be used for the identification and classification of closely related taxa. Variation of *gyrA* gene sequences has been found to discriminate among *Bacillus subtilis* group whereas *gyrB* is useful to discriminate among members of *Bacillus cereus* group. Two conserved amino acid sequences in *gyrB* of *E. coli*, *Pseudomonas putida* and *B. subtilis* were reverse transcribed and two universal primers have been designed for the purpose of identification. The PCR product with a predicted size of 1.2 kb is able to identify successfully the various taxonomic groups of α, β- and γ-proteobacteria, *gyrB* specific primers have also been designed which do not amplify topoisomerase IV and thus avoids the paralogous comparison (Wang et al., 2007). The greatest advantage of using *gyrB* sequences in identification practices is that, the average base substitution rate of 16S rRNA gene is 1% per 50 million years, whereas, the rate is estimated to be 0.7–0.8% per one million years in the case of *gyrB* (Chun and Bae, 2000). Hence, those species having completely identical 16S rDNA sequences

Table 1
Suitable primer sets for several conserved regions of 16S rRNA gene amplification in bacteria.

Primer name	Sequence (5′–3′)	Reference
8F	AGA GTT TGA TCC TGG CTC AG	Frank et al. (2008)
U1492R	GGT TAC CTT GTT ACG ACT T	
928F	TAA AAC TYA AAK GAA TTG ACG GG	Weidner et al. (1996)
336R	ACT GCT GCS YCC CGT AGG AGT CT	
1100F	YAA CGA GCG CAA CCC	Eden et al. (1991)
1100R	GGG TTG CGC TCG TTG	
337F	GAC TCC TAC GGG AGG CWG CAG	Weidner et al. (1996)
907R	CCG TCA ATT CCT TTR AGT TT	
785F	GGA TTA GAT ACC CTG GTA	Eden et al. (1991)
805R	GAC TAC CAG GGT ATC TAA TC	
533F	GTG CCA GCM GCC GCG GTA A	Frank et al. (2008)
518R	GTA TTA CCG CGG CTG CTG G	

Table 2
Primers reported to be used for the amplification of *rpo* genes in bacteria.

Primer name	5′ to 3′ sequences	Reference
rpoAF	GCAGTTAGCAGAGCGGACAG	Guo et al. (2012)
rpoAR	GGAAACCAACGGCACAAT	
rpoBF	GGTTCGCCGCGCTGGCGCGAAT	Rigouts et al. (2007)
rpoBR	GACCTCCTCGATGACGCCGCTTTCT	
rpoC-uni	GTGCACTCGGTCCACAG	Ventura et al. (2006)
rpoC-rev	CATGCTCAACAACGAGAAG	
rpoDF	ACGAGACTGCTGTTGCGCITICCGAACA	Fujita et al. (1995)
rpoDR	TGACGTGATTTCCTGAGCAG	

can be differentiated using *gyrB* gene sequences in addition to the 16S rDNA sequence data.

4.1.4. recA and recN gene sequencing

16S rRNA gene sequencing method has been undoubtedly regarded as an accurate molecular identification method. However, in some cases it is not able to distinguish preferably at interspecies level as found in *Mycobacterium* (Rogall et al., 1990). In this regard, *recA* gene sequence provides a useful tool for differentiation at interspecies level. *recA* gene apparently occurs in all bacteria and is responsible for DNA damage repair, induction of SOS repair and homologous DNA recombination. As a part of the SOS response, it coordinates more than 20 genes that are involved in DNA repair, DNA recombination, DNA synthesis and cell division (Miller and Kokjohn, 1990). Many Gram-negative and Gram-positive bacteria have been reported to be classified based on the result obtained from the *recA* gene sequences (Karlin et al., 1995). Comparative phylogenetic analysis of *recA* and 16S rRNA gene sequence demonstrates a highly similar branching pattern suggesting the advantage of using this gene sequence for molecular systematics as well as species level identification (Eisen, 1995). Species of *Mycobacterium* (Blackwood et al., 2000) and *Burkholderia* (Payne et al., 2005) have been reported for *recA* gene sequence based identification procedures. When the amplification of *recA* gene is combined with restriction fragment length polymorphism methods, it becomes more accessible for identification and typing of any bacterial entity. However, recently Glazunova et al (2010) developed a novel approach of identification of genus *Streptococcus* using partial *recN* gene sequencing. As per this report, the partial *recN* gene sequences provide 56.4–98.2% interspecies nucleotide similarity whereas, 89.8–98% similarity was obtained at inter subspecies level. However, 16S rRNA gene (94.1%), *rpoB* (84.6%), *sodA* (74.8%), *groEL* (78.1%) and *gyrB* (73.2%) showed higher mean DNA sequence similarity than the *recN* gene which was found to be 66.6% (Glazunova et al., 2010). Now-a-days *recN* gene sequence based identification has been found to be widely applicable in species level identification of many bacterial species like *Streptococcus*, *E. coli* (Finch et al., 1985), *Neisseria gonorrhoeae* (Skaar et al., 2002) and *Amycolatopsis* (Everest et al., 2011).

4.1.5. Analysis of the ppk gene and protein

When biochemical tests are considered for the identification of a bacterial strain they are highly influenced by genetic as well as environmental factors. Therefore, many genes are responsible for the diverse phenotype of the bacteria and *ppk* genes are the best model to study the phenotypic diversity due to their conserved sequence nature. One widely studied among these genes is the *ppk1* gene which codes for polyphosphate kinase gene 1 in many bacteria. The Ppk1 enzyme encoded by the *ppk1* gene is responsible for the synthesis of inorganic polyphosphate (poly P) which contains hundreds of orthophosphate residues linked by phosphoanhydride bonds like that of ATP (Brown and Kornberg, 2004). Though many bacterial species also contain another enzyme Ppk2 coded by the *ppk2* gene which is highly conserved among bacterial species, it is not ubiquitous in nature (Zhang et al., 2002). The *ppk2* gene, 1074 bp in length encoding for a protein with 357 amino acids and 40.8 kDa, is found to be homologous in *Pseudomonas aeruginosa*, archea, rhizobia, cyanobacteria, *Streptomyces* and several other pathogenic bacteria (Zhang et al., 2002). Ppk proteins may have many functions related to inhibition of RNA degradation, activation of Lon protease, membrane channel formation, stress resistance, motility, virulence as well as quorum sensing (Shi et al., 2004). Mutation studies on these genes revealed the role of these genes in genetic diversity, points to connectedness, redundancy as well as resilience in metabolic pathways (Tan et al., 2005), which can be used for the analysis of microbial diversity and identification practices. *ppk* gene sequences and their coded proteins have been used widely for the diversity study of many bacterial species including *Microbacterium*, *Aureobacterium* (Richert et al., 2007), *Azotobacter vinelandii* (Lloret et al., 2006) and *Helicobacter pylori* (Tan et al., 2005). In addition, the fraction of informative sites of total length was 32.9% for *ppk* in comparison to *rpoB* (23.8%) and *recA* (26.8%) sequences which confirms the advantage of using *rpoB* as a gene target for bacterial identification (Richert et al., 2007).

4.1.6. dnaJ gene sequencing for Mycobacterium identification

Among the multilocus sequence analyses now-a-days used, *dnaJ* gene sequence is the most recent addition to the list. *dnaJ* chaperone gene has been investigated widely to be used as a molecular marker for alpha-proteobacteria (Alexandre et al., 2008). As there is no evidence of horizontal gene transfer of this gene in proteobacteria, it can be used for detection in an efficient manner. There are several domains present in DnaJ protein containing highly conserved J domain, a conserved region of four repeats with consensus sequences (CxxCxGxG), G/F rich domain and a C terminal region of variable length which is not conserved (Siegenthaler et al., 2004). In some cases, *dnaJ* gene (1131 bp) is found to be more than one copy having high G + C content in Gram-positive *Actinobacteria* which are having different physiological functions and distinct evolutionary paths (Ventura et al., 2005). *dnaJ* gene sequencing has been widely used for identification of many species of *Mesorhizobium* (Alexandre et al., 2008), *E. coli* (Georgopoulos et al., 1980), *Staphylococcus* (Shah et al., 2007) and *Mycobacterium* (Takewaki et al., 1994).

4.1.7. Others: gapA, pyrH, luxABE, ectBC, dnaA, pheS and dnaK

Many other relatively new genes have been discovered for bacterial identification practices and are listed in Table 3. When an organism is classified into a new genus, it cannot be confirmed relying on the sequence result of any single gene. At this point in time, the basis of phylogenetic analysis should be based on the multiple independent loci along with the 16S rRNA gene (Ast et al., 2009). Similarly, some universal protein coding genes have also been applied for the identification and construction of phylogeny as well as deciphering the global phylogeny and consistency of the triplicate division of eubacteria, archaea, and eukarya. Due to continuous explosion of sequence information many numbers of individual genes and their amino acid sequences are analysed by various models and algorithms. The most emerging field of sequence based identification systems include the gene sequences of RNA polymerase core subunits, DNA polymerases, translation elongation factors, EF-Tu and EF-G, glutamine synthetase, ribosomal protein genes, RecA protein as well as heat shock proteins of the family Hsp60 and Hsp70. *gapA* is the gene present in all microbial species that encodes for glyceraldehyde-3-phosphate dehydrogenase, the enzyme that catalyses the oxidation and phosphorylation of glyceraldehyde-3-phosphate to 1,3-biphosphoglycerate and the gene sequences are highly conserved in both bacteria as well as eukaryotes (Nelson et al., 1991). Similarly, pyrH is mostly present in virulent isolates and it has a role in induction of antigenic property by encoding UMP kinase that catalyses UMP phosphorylation (Lee et al., 2007). *lux* genes are arranged in a single operon as *luxCDABE* and are essential for luminescence activity. A fatty acid reductase complex involved in the synthesis of fatty aldehyde substrate for the luminescence reaction catalysed by the LuxAB subunits of luciferase is encoded by *luxCDE* (Winson et al., 1998). A similar study showed that genes involved in the synthesis of the major compatible solute, ectoine: diaminobutyric acid aminotransferase gene (*ectB*) and ectoine synthase gene (*ectC*) can also be complemented to the 16S rDNA based phylogenetic relationship (Okamoto et al., 2004). Now-a-days sequencing of DNA and proteins have become very easy and fast by the use of automated tools thus increasing the efficiency and efficacy of the techniques, hence, only selection of the suitable model remains the prerequisite for the identification and phylogenetic analysis of any bacterium.

4.2. Multilocus sequence typing

The molecular biology technique dealing with typing of multiple loci is called MLST which deals with the sequencing of internal fragments of

Table 3
List of some uncommon genes used in bacterial identification practices.

Targeted gene sequence	Function of the gene/product of the gene	Reported to be used in identification for	Reference
gapA	Production of D-glyceraldehyde-3-phosphate dehydrogenase	*E. coli*, *Mycoplasma gallisepticum*	Charpentier and Branlant (1994), Ferguson et al. (2005)
pyrH	UMP kinase catalysing UMP phosphorylation	*Vibriovulnificus*, *Vibrio* sp.	Lee et al., 2007, Thompson et al. (2005)
luxABE	Production of luciferase enzyme that catalyses light emitting reaction	*Aliivibriologei*, *Aliivibrio salmonicida*, *Photobacterium mandapamensis*	Manukhov et al. (2011), Kaeding et al. (2007)
ectBC	ectoine: diaminobutyric acid aminotransferase	*Halomonasvariabilis*, *Halomonadaceae*	Arahal et al. (2007), Okamoto et al. (2004)
dnaA	Encodes for protein that activates initiation of DNA replication	*Mycobacterium* sp., *Rhizobium meliloti*	Margolin et al. (1995), Mukai et al. (2006)
pheS	phenylalanyl-tRNA synthase	*Enterococcus* sp., *Lactobacilli*	Naser et al. (2005), Naser et al. (2007)
dnaK	Initiation of DNA replication,	*Ensifer* sp., *Sinorhizobium*, *Nitrosomonas europaea*	Iizumi and Nakamura (1997), Martens et al. (2008)

multiple house-keeping genes and was first developed in *Neisseria meningitidis* (Maiden et al., 1998). The principle involves the measurement of DNA sequence variations using a set of house-keeping genes to characterise the strains from which these genes are derived from by their unique allelic profile. Depending upon the desired degree of discrimination, nucleotide differences between the strains can be checked in the number of genes. In this context, data analysis plays an important role in which allele numbers are assigned to all unique sequences and subsequently combined to assign an allelic profile and a sequence type (ST). The relatedness of the isolates is performed by comparing the allelic profile from the data base. In this case the major challenge involves the suitable selection of the house-keeping genes and maintaining the balance for an acceptable identification power. Hence, in case of *Staphylococcus aureus* seven house-keeping genes have been standardized i.e. carbamate kinase (*arcC*), shikimate dehydrogenase (*aroE*), glycerol kinase (*glpF*), guanylate kinase (*gmk*), phosphate acetyltransferase (*pta*), triosephosphate isomerase (*tpi*) and acetyl coenzyme A acetyltransferase (*yqiL*) for MLST (Saunders and Holmes, 2007). Similarly for *V. vulnificus*, the used house-keeping genes are glucose-6-phosphate isomerase (*glp*), DNA gyrase subunit B (*gyrB*), malate–lactate dehydrogenase (*mdh*), methionyl-tRNA synthetase (*metG*), phosphoribosyl aminoimidazole synthetase (*purM*), threonine dehyrogenase (*dtdS*), diaminopimelate decarboxylase (*lysA*), transhydrogenase alpha subunit (*pntA*), dihydroorotase (*pyrC*) and tryptophanase (*tnaA*) (Bisharat et al., 2007) which differ from species to species. Though the MLST scheme is widely used for pathogenic bacteria, it can also be applied in the case of many pathogenic fungi as well as non-pathogenic bacteria (Maiden, 2006).

As MLST analyses from six to eight genetic loci, it may not always provide sufficient resolution among closely related bacteria, hence, each MLST scheme has to be developed for a group of related bacteria (Achtman, 2008). Much advancement has taken place so far in this technique and ribosomal multilocus sequence typing (rMLST) provides a universal reference point to complement existing methods in the interpretation of bacterial diversity (Jolley et al., 2012). Whole genome sequencing techniques can also be used along with sequence typing for a better understanding of identification, phylogenetic analysis, antimicrobial resistance and virulence gene characterisation (Larsen et al., 2012). The MLST technique can be further improved by the addition of computer based techniques for the development of internet services as well as interlinking with software (Chan et al., 2001).

4.3. Protein profiling for the identification of bacteria

Apart from the genetic techniques, most advanced protein profiling techniques have also been proposed as alternatives for microbial identification. In the past, extraction of bacterial proteins and their sequencing using MS/MS have been followed by matching the results of MS spectrum to database for their identification. Most recently, the protein profiling of intact bacterial colonies used for the microbial identification using MALDI-TOFMS has proven to be a reliable approach (Sauer et al., 2008). By this method, bacteria of independently prepared groups can be most accurately classified and identified using unsupervised (hierarchical clustering analysis) and supervised (genetic algorithm) approaches which possess the capability to identify bacteria in the minimal number of cells and mixed flora (Fig. 3). Hence, applying bacterial colony directly to MALDI-TOF MS is a simple and reliable method of protein profiling and the selection of a panel of m/z values instead of the whole spectra reduces the dependence on spectrum consistency which ultimately achieves higher bacterial identification rates (Carbonnelle et al., 2011). Similarly, different bacterial species in a mixed flora can be identified with a detection limit lower than the environmentally significant value.

In addition to that, carbon and nitrogen fluxes in microbial communities can be analysed by protein based stable isotope probing (Protein-SIP). Intact protein profiling (IPP) can also be used for the identification of microbial species, where microbial species can be grown on C^{13} and N^{15} containing substrates for identification of IPP. In the case of SMM peptide fragmentation and their sequences can be used additionally for identification of species (Jehmlich et al., 2009). In an alternative manner, protein processing and isolation of peptide mixtures of bacterial samples can be analysed by LC–MS/MS with the use of in-house algorithm BACid by comparing with the proteome database for identification of bacteria at genus, species and strain levels (Jabbour et al., 2011).

4.4. FAME profiling for the identification of bacteria

Fatty acid composition in bacterial species shows a qualitative conserved pattern and significant change can only take place over a long period of time. Till date, over 300 different fatty acids have been detected in bacteria and their position of double bonds, binding of functional

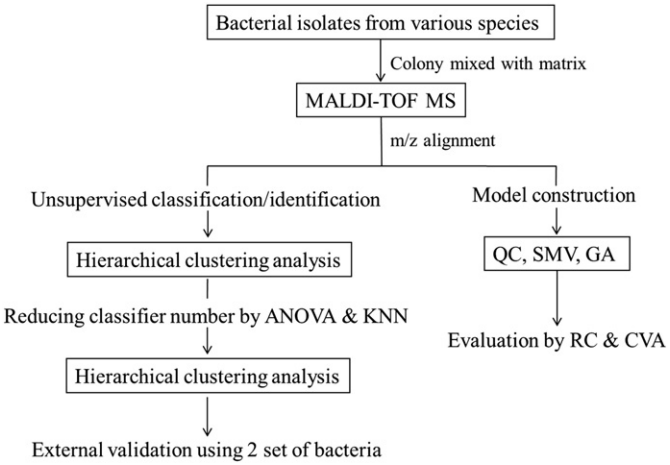

Fig. 3. Work flow of the study for bacterial identification using protein profiling.

groups, difference in chain length make them a considerable taxonomic marker (Dawyndt et al., 2006). Whole cell FAME analysis in association with machine learning techniques is in wide application for bacterial community analysis (Inglis et al., 2003; Van den Velde et al., 2006; Pineiro-Vidal et al., 2008; Slabbinck et al., 2009). A variety of statistical techniques are used additionally to analyse FAME biomarker profiling which includes nMDS, ANOSIM, and DFA (Ehrhardt et al., 2010). Besides identification practices, fatty acid analysis of bacterial species can provide information on the methods to grow microorganisms of forensic interest, differentiation between vegetative cells and free spores and many more (Song et al., 2000; Das et al., 2007). Much advancement has taken place for use of fatty acid profiling techniques in bacterial identification practices i.e. profiling by capillary gas chromatography coupled with FID (Song et al., 2000), the much more automated system MIDISherlock MIS that identifies based upon the FAME pattern of known organisms (Sasser, 1990) or analysis of mixture of cellular lipidomes by electrospray ionisation MS (Han and Gross, 2003). Many microorganisms including Gram-positive bacteria *S. aureus*, *Listeria monocytogenes*, *Bacillus anthracis* and *B. cereus* and Gram-negative bacteria *Yersinia enterocolitica*, *Salmonella typhimurium*, *Shigella sonnei* and *E. coli*, and non-Enterobacteriaceae *Vibrio cholerae*, *V. vulnificus* and *Vibrio parahemolyticus* were reported to be analysed for FAME by using ATRFTIR spectroscopic procedure as well as discrimination by multivariate analysis (Whittaker et al., 2003). DCIMS profiling of FAMEs from in situ THM of whole bacterial cells with TMAH has also been reported as a potential technology for real time detection and identification of microorganisms (Xu et al., 2003).

4.5. Sequence based techniques for fungi

Many techniques have been employed so far for the identification of fungi mainly by using morphological as well as biochemical characteristics such as utilisation of carbon and nitrogen compounds. However, these techniques have less significance due to the occurrence of different morpho/biotypes present within a single species, in addition to being time consuming (Sugita and Nishikawa, 2003). In order to overcome these problems, many advanced techniques have been developed and are described below.

4.5.1. 26S rRNA gene sequence

Genetic approaches of studying fungal phylogeny mostly include nuclear rDNA markers. In the case of fungi, rDNA consists of the SSU 18S, ITS, ITS1 + 5.8S + ITS2 and the LSU 25-28S regions (Hibbett et al., 2007). However, ITS has been considered to be the best fungal barcode for identification purposes (Porter and Golding, 2012). It has been reported that the sequence comparison of internal transcribed spacer and D1/D2 26S rDNA spacer sequences (Fig. 4) using any of the primer sets (Table 4) provide a useful tool for the identification of fungi. Analysis of D1/D2 26S rDNA sequences not only have the advantage of species identification but also permitphylogenetic analysis. However, the major disadvantage is the absence of registered fungal DNA sequences in the DNA data libraries. It is very difficult to phylogenetically analyse ITS sequences as they are more diverse in comparison to the sequences of the 26S rDNA region.

4.5.2. 28S rRNA gene sequence

Though many studies of sequence based identification of fungi deal either with the 18S rRNA gene, internal transcribed spacers or the 5′ end of the 28S rRNA gene, the highly conserved 2900 base pair region of the 28S rRNA gene is largely unexplored due to their undefined nature in the genomic database. However, the sequences of the 28S rRNA gene beyond the D1–D2 hypervariable region is a useful tool for the identification procedure of many pathogenic fungi like *Aspergillus candidus*, *Aspergillus flavus*, *Aspergillus fumigatus*, and *Candida albicans*. The broad range of fungal rRNA primers (Table 5) can be used in end-point PCR for amplification and identification of fungi. In this regard, the analytical sensitivity and cross-reactivity testing should be performed for proper validation of the result.

Table 4
Primer sets for the amplification of the 26S rRNA gene for identification of fungi.

Primer	5′ to 3′ sequence	Reference
N-nc26S1	CGACCCCAGGTCAGGCG	Kuzoff et al. (1998)
N-nc26S2	GAGTCGGGTTGTTTGGGA	
N-nc26S3	AGGGAAGCGGATGGGGGC	
N-nc26S4	AGGGAAGCGGATGGGGGC	
N-nc26S5	CGTGCAAATCGTTCGTCT	
268rev	GCATTCCCAAACAACCCGAC	
641rev	TTGGTCCGTGTTTCAAGACG	
950rev	GCTATCCTGAGGGAAACTTC	
1229rev	ACTTCCATGACCACCGTCCT	
1499rev	ACCCATGTGCAAGTGCCGTT	

Gene sequence based identification of fungi has many advantages over the conventional methods as it does not require viable organisms or sporulation i.e. moulds which are helpful for a rapid identification procedure. Thus the optimal targets for fungal identification are internal transcribed spacer regions i.e. ITS1 and ITS2 and they have been proven to be useful for the identification of yeasts such as *Candida*, *Cryptococcus*, *Trichosporon*, and mold *Aspergillus* (Petti, 2007). As the ITS regions have many limitations, the alternative gene targets have also been proposed to be useful for fungal identification such as the D1 and D2 domains of the 28S rRNA gene, and elongation factors α-tubulin and β-tubulin (Hall et al., 2003).

4.5.3. β-tubulin gene sequences

Biochemical analysis has suggested the occurrence of two α-tubulin and two β-tubulin genes in fungi. *tubA* gene is responsible for the production of two alpha tubulin polypeptides i.e. alpha 1 and alpha 2, where alpha 2 is produced by *tubB* gene. Similarly, beta-tubulin polypeptides beta 1 and beta 2 are produced by *benA* gene, and beta 3 is produced by *tubC* gene (Seip et al., 1990). It has been reported that 3.5 times higher phylogenetic relationship can be deduced by analysing β-tubulin genes than by analysing mitochondrial SSU rRNA genes (Craven et al., 2001). Recent analysis of fungal beta tubulin sequences shows that the intron regions of these gene sequences may vary between different lineages (Begerow et al., 2004). However, the gene sequence remains conserved among *Hymenomycetes* and *Ustilaginomycetes*. Even though the most conserved regions are highly susceptible to introns, as splicing and branching sites of introns are highly variable in the basidiomycetes group, this problem can be overcome by analysing only the amino acid sequences because introns are removed before translation during the process of splicing (Begerow et al., 2004).

4.6. PCR based finger printing techniques

There are various PCR fingerprinting techniques available for the amplification of polymorphic DNA through specific selection of primer

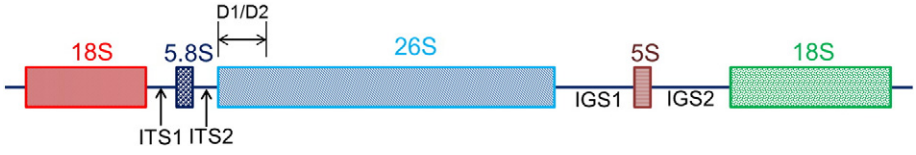

Fig. 4. Schematic representation of the fungal rRNA gene. ITS: internal transcribed spacer; IGS: intergenic spacer; D1/D2: Domain 1 and 2.

Table 5
Broad-range fungal rRNA gene primers for identification procedure.

Primer	5′ to 3′ sequence	Reference
End 18S	GTA AAA GTC GTA ACA AGG TTT C	Khot et al. (2009)
5.8S	GTG AAT CAT CGA RTC TTT GAA C	
28S-1	TAC CCG CTG AAC TTA AGC ATA	
28S-2	GCA TAT CAA TAA GCG GAG GAAA	
28S-3	AGT ARC GGC GAG TGA AGC GG	
28S-4	AGC TCA AAT TTG AAA SCT GG	
28S-5	CTT CCC TTT CAA CAA TTT CAC RT	
28S-6	GAG GTA AAG CGA ATG ATT AG	
28S-7	CTT GTT RCT TAR TTG AAC GTG	
28S-8	ACC ACA AAA GGT GTT AGT WCA TC	

annealing sites. These molecular typing methods of bacteria are extremely reliable, rapid and highly discriminative as well as reproducible. The fundamentals behind these techniques are the use of PCR with primer sets (Table 6) that are complimentary to the highly conserved, naturally occurring, repetitive DNA sequences present in multiple copies in distinct intergenic positions throughout the genome of most Gram-negative and Gram-positive bacteria. The REP-PCR based molecular typing method was first used by Versalovic et al. (1994) and has been advanced dramatically in the past couple of years to become completely automated (Healy et al., 2005). In comparison to REP-PCR, ERIC-PCR is highly sensitive and useful in detecting microorganisms from any environment. BOX-PCR is the superior to all the techniques creating distinct fingerprinting patterns, however, ERIC and REP-PCR are the methods which are primarily used for genotyping (Frye and Healy, 2006). REP-PCR has a consensus sequence of 38 bp in addition to the 5 bp in the stem loop of the palindrome structure. Similarly, ERIC-PCR has a consensus sequence of 126 bp and is also found in extragenic regions. However, BOX-PCR has three subunits, BOX-A, BOX-B and BOX-C having 59, 45 and 50 nucleotide lengths respectively. Most importantly, BOX-PCR does not share any sequence homology with either ERIC or REP-PCR (Olive and Bean, 1999). The arbitrarily primed polymerase chain reaction (AP-PCR) can amplify fragments of DNA from any genome varying the size distribution of amplified fragments among species. Thus, the closely related taxa possess similar fragment distributions, while that of the distantly related taxa are more divergent, hence, providing considerable phylogenetic information (Espinosa and Borowsky, 1998).

Classically, the comparison of DNA fingerprinting banding pattern distribution among taxa is based on the assumption that, bands of the same size are homologous. However, this assumption is correct for the closely related taxa and becomes less reliable for the increasingly separate ones. Thus the assumption of homology can be confirmed by DNA hybridisation (Pillay and Kenny, 1995). In order to enhance the efficiency of PCR fingerprinting techniques, these can be conjugated with hybridisation and southern blotting. The major advantage of this technique involves the considerable easier handling than probing with a single band; however, the disadvantage involves the sharing of sequence similarity with other amplified products of different size (Pillay and Kenny, 1995).

4.7. Phylogenetic information from whole genome sequences

Molecular approaches of microbial identification rely heavily on the analysis of a single gene of interest. However, there exists a huge debate as to whether a single gene represents the true evolutionary lineage of a microorganism? It has been reported that the conserved genes also can be transferred through the natural phenomenon of horizontal gene transfer (Ochman et al., 2000). Besides, the degradation of phylogenetic signal due to saturation for amino acid substitutions (Forterre and Philippe, 1999) has also led scientists to develop the 'super-tree' approach. This approach takes into account the whole genome sequences for microbial classification using alignments and analysis of a large number of conserved genes. In addition, whole genome sequences reveal much more information regarding the presence or absence of orthologous or families of genes, specific protein folds, conserved insertions or deletions, differences of overall gene content, conservation of gene order and biases of nucleotide compositions of genomes (Coenye and Vandamme, 2003).

In this approach, the whole genome of the culturable bacteria can be obtained by automated pyrosequencer followed by detection of putative orthologous genes and comparison of genes and amino-acid sequences of various house-keeping genes including 16S rRNA, *gyrB*, *rpoD*, *sodA*, *dnaK*, *recA*, *gki*, *ddl*, *alaS* and *ileS* (Coenye and Vandamme, 2003). The salient advantage of comparing the amino-acid sequences over gene sequences are the silent mutations which occur more frequently than replacement mutations to randomise the third codon position of protein coding genes. Thus, the third codon base composition varies systematically between species which indicates the lineage specific subjection to a selective force (Swofford and Olsen, 1990). The 'super-tree' can be constructed based on 16S rRNA gene sequences and other house-keeping amino-acids by combining the similarity matrices of the individual comparisons using various software. In a similar approach, dinucleotide relative abundance values and codon usage can be determined by using the inverted complementary sequences using software tools like REVSEQ, YANK, UNION and EMBOSS (Rutherford et al., 2000). This can be followed by the analysis of colinearity of genomes among the regions displaying substantial sequence similarity. Finally, the results obtained from the whole genome sequence data can be analysed using statistical tools such as Bionumerics and SPSS.

With the increased availability of completely sequenced genome data, it has become easy to predict the phylogenetic relationship among microorganisms. Though most of the phylogenetic approaches use maximum information available from the whole genome sequence data, there is still the need for selecting the phylogenetic marker genes for individual microorganism (Capella-Gutierrez et al., 2014). Whole genome sequencing approaches can also be applied for analysis of fungal phylogeny as there are now a growing number of complete fungal genomes available (Galagan et al., 2005).

Table 6
PCR techniques used for the fingerprinting analysis of bacterial isolates.

PCR	Primers to be used	Amplicon Size	Preferred organisms	References
ERIC	5′-ATGTAAGCTCCTGGG-3′ 5′-AAGTAAGTGACTGGGGG-3′	150–5500 bp	*E. coli*, *Salmonella typhimurium*	Wilson and Sharp (2006)
REP	5 ′-ACGTCTTATCAGGCCTAC-3′ 5′-IIIGCGCCGICATCAGGC-3′	150–10,000 bp	*Bacillus* sp., *E. coli*, *S. typhimurium*	Albufera et al. (2009)
BOX	5′-CTACGGCAAGGCGACGCTGACG-3′	140–1230 bp	*Salmonella enterica*	Mohapatra et al. (2007)
AP	5′-CCGGAAGAAGGGGCGCCAT-3′ 5′-CGATTTAGAAGGACTTGCACAC-3′	Any length	*Leptospira interrogans*	Ramadass et al. (2002)
$(GTG)_5$	5′-GTGGTGGTGGTGGTG-3′	Any length	*Enterococcus* sp.	Svec et al. (2005)

5. Culture independent identification procedures

The vast diversity of living microorganisms in nature are both beneficial and deleterious for human beings and a detailed compact understanding is required for their community structure to deal with their activities, response to environmental stress and climate change (Hirsch et al., 2010). However, only 1% of the total number of microorganisms can be cultured under laboratory conditions and in nutrient rich conditions it can be increased to a mere 10% (Sorensen, 1997). Apart from the environmental point of view, the clinical sectors also face the hurdle of less cultivable microflora. For example, gastrointestinal tract commensal microbes show analogous function as that of the indigenous microbiota of other sites of the body. Some fastidious and/or anaerobic microorganisms that grow easily in a community overshadow the culture of other microorganisms (Frank et al., 2003). Hence, a new field 'metagenomics' has emerged that deals with the investigation of collective microbial genomes obtained directly from the environmental sample without cultivation and without the prior knowledge of the constituent communities (Riesenfeld et al., 2004).

5.1. Gradient gel electrophoresis

In order to study the bacterial community structure in a particular environment, culture based approaches don't provide comprehensive information because they fail in cultivation of the predominant microorganisms. DGGE as the culture independent method for quick study of the bacterial community was first developed by Muyzer et al (1993), which is helpful in the direct measurement of a microbial community in a specific ecosystem. There are various other methods available for the culture independent approaches of microbial identification (Fig. 5).

In culture independent approaches, the total DNA can be isolated from the sample and the desired conserved gene can be amplified by PCR reaction which is of the same size but of different sequences with the aim of generating a distinguishing banding pattern. With this principle, DGGE was first developed by Muyzer et al. (1993) for the rapid culture independent study of bacterial communities in specific ecosystems. In DGGE, the organism specific fingerprint is obtained and based upon the banding pattern, the community structure in terms of their quality as well as quantity can be analysed (Fig. 6). DGGE is the most commonly used method among the culture independent fingerprinting techniques. Besides DGGE, there are various other methods available for the culture independent approaches of microbial identification. The application of PCR-DGGE is highly versatile and successfully used in many fields of microbial ecology to explore many uncultivable bacterial species. Many applications of PCR-DGGE have been described (Bano and Hollibaugh, 2002; Sekiguchi et al., 2002; Crump et al., 2003; Nicol et al., 2003) in the field of soil, sea, river, lake ecosystems as well as clinical samples. Identification of species is achieved by purifying and sequencing the bands in the DGGE profile. However, PCR-DGGE analysis is not suitable for the identification of all species, but can be used to screen and group the microbes thus reducing their numbers to be identified by other molecular methods. Along with the DGGE technique, certain other methods like TGGE and CGGE can also be used for the same purpose of identification of microorganisms and their community structure.

5.2. Flow cytometry for bacterial identification

Due to advances in technology, flow cytometry (FC) has been introduced in the field of microbiology to investigate microbial communities, identify and quantitate populations using DNA content and to identify populations of interest using molecular probes (Lange et al., 1996). The technology involves the mixture of fluorescent nucleic acid stains for bacterial identification as well as calibrated suspension of beads for accurate sample volume measurement. The most important advantage of using flow cytometry is the rapidity of the analysis that makes it

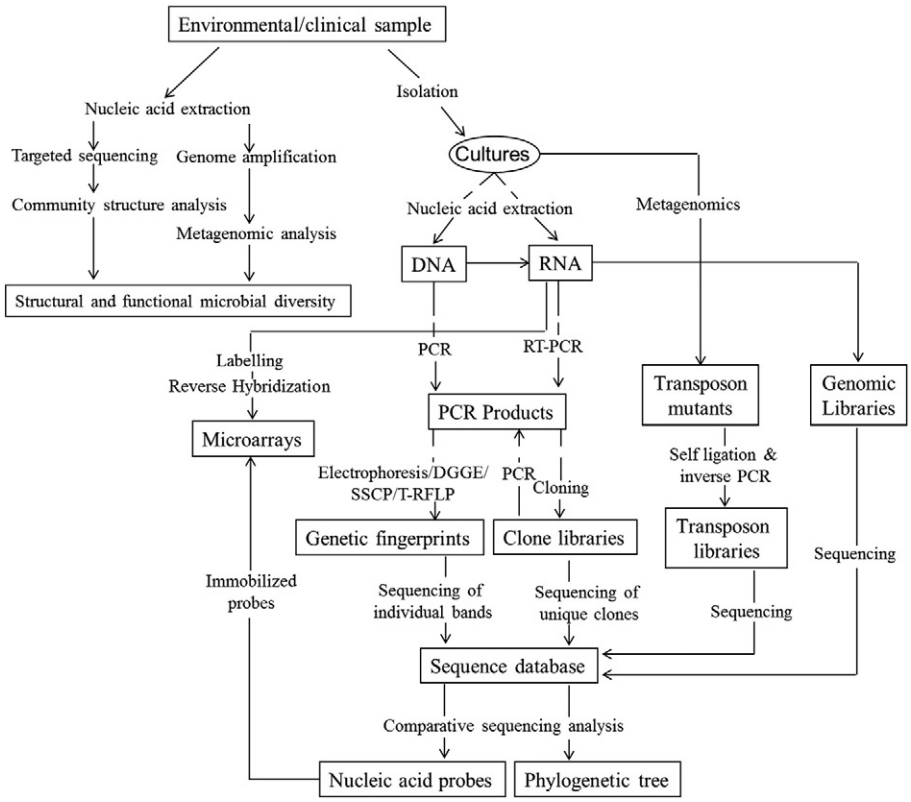

Fig. 5. Molecular approaches for detection and identification of bacteria from environmental samples.

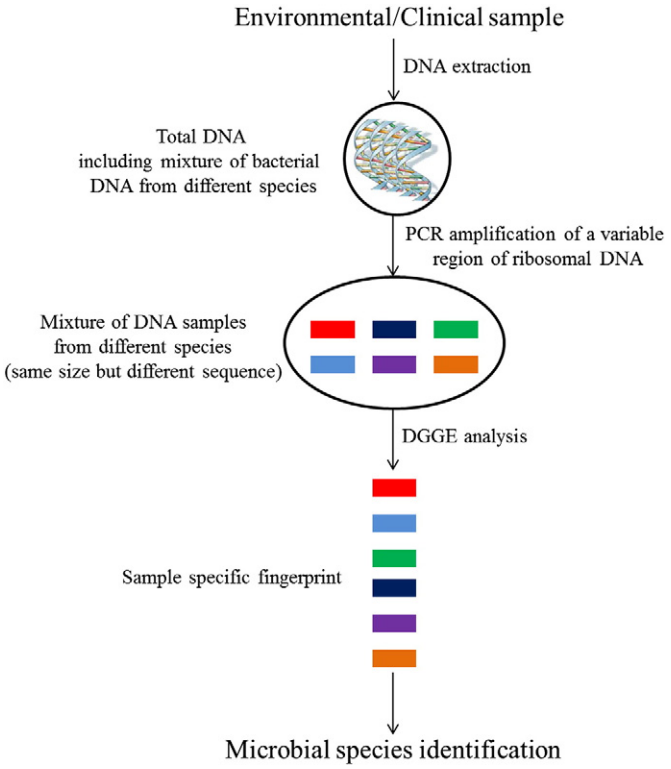

Fig. 6. Flow diagram of the application of DGGE to environmental samples.

possible for both routine and exploratory analysis of microbial populations (Duhamel and Jacquet, 2006). Though the application of FC and FACS are abundant for the application of sorting and identifying microbial populations (Fig. 7), the technology still has been under utilised for the investigation of microbial systems. The major drawbacks of this technique for investigating microbial systems are their small cell size, reduced per cell content of proteins, nucleic acids and other biological molecules, and lack of experience in bacteria in applying probes developed for higher eukaryotes. In addition to that, bacterial membranes are less permeable to fluorochromes than mammalian cells and most of the fluorochromes may be pumped out by many bacterial cells due to the presence of efficient efflux pumps (Vives-Rigo et al., 2000).

5.3. Bacterial tag encoded FLX titanium amplicon pyrosequencing (bTEFAP)

Microbial diversity is much more complex in environmental conditions than originally expected and to deal with such huge diversity more advanced and reliable techniques have to be used. In this regard, the recently discovered pyro-sequencing technique may be employed. With the availability of increased speed and efficiency in genomic data generation, next generation sequencing has become one of the highly efficient technologies most relied upon (Dini-Andreote et al., 2012). In this context, automated bioinformatics tools have also been discovered to achieve rapid, consistent, accurate taxonomic analysis by reading sequencing results (Liu et al., 2008). Next generation sequencing has made comprehensive evaluation of microbial diversity possible in any environment. TEFAP has been proved to be useful for evaluating archaeal, bacterial, fungal and algal diversity (using 16S and 18S conserved primer sets) as well as analysing functional genes (using specific primer sets) (Sun et al., 2011). The most recent advancement in this field is the use of bTEFAP which has been used successfully in environmental conditions (Sundarakrishnan et al., 2012) as well as clinical conditions (Dowd et al., 2008) for assessment of microbial richness.

6. Polyphasic taxonomic approach: Case studies

Though the scientific rationale of all the techniques available for bacterial identification cannot be questioned, the possible errors in each of the techniques cannot be underestimated. Hence, there is no phylogenetic standard for family, genus or species level demarcation. This is the major reason why the polyphasic approach of bacterial identification comes into play. Any polyphasic approach for classification and identification of bacterial species includes methods available to allocate bacterial species phylogenetically, to compare and group large number of strains into clusters of similar groups, DNA–DNA hybridisation to compare the relationship between representatives as well as the

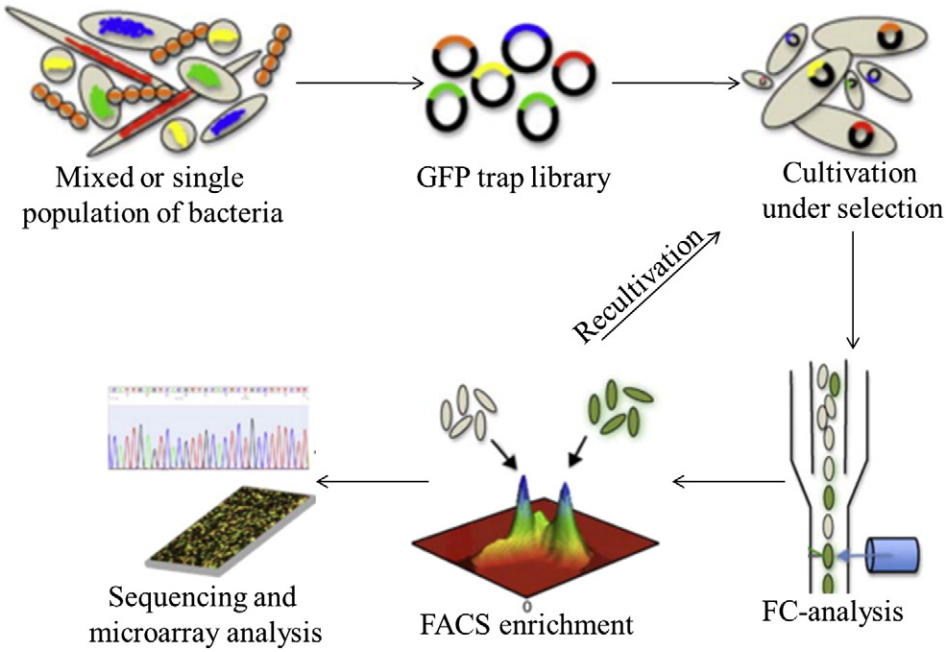

Fig. 7. Flow cytometry analysis of bacterial sorting and identification.

descriptive methods to provide further genotypic and phenotypic information. All possible methods which inform the biological nature of the strain should be given equal importance and attention (Vandamme et al., 1996). Polyphasic classification is purely empirical which does not follow any strict rules or guidelines and may integrate any significant information in the organism to provide a consensus type of classification. Though this procedure is not hampered by any conceptual prejudice, the outcome reflects its biological activity. When a new bacterium is isolated from a new source any single method is not appropriate to assign a particular taxonomy and in this regard a polyphasic approach should be considered (Prakash et al., 2007). For the identification of fungi, the species concept has evolved by the use of morphological data to polyphasic scrutiny thus incorporating a powerful molecular component. This results in traditional classification being less practical, as holotypes are not presently useful as DNA cannot be extracted. Thus, a polyphasic approach may be adopted for introduction of new fungal species or designing of epitypes which includes molecular, morphological, physiological as well as pathogenic data. In spite of the advancement of the technologies, automation, robotisation and the increasing strength of databases, there should be an international interconnection and accessibility of the technologies available for the development of a better, reliable and sustainable protocol for the identification of any unknown microbial (bacterial, fungal and algal) species (Kampfer and Glaeser, 2012).

6.1. Polyphasic taxonomy of the genus Xanthomonas

The members of genus *Xanthomonas* are aerobic, Gram-negative, non-spore forming rods (Leyns et al., 1984) and are the major known causative agents for bacterial spot on pepper and tomato as well as for at least 124 monocotyledons and 268 dicotyledons plant species. Due to their remarkable phytopathogenic diversity in contrast to the phenotypic uniformity, the taxonomy and classification of the genus *Xanthomonas* remained unclear until polyphasic taxonomy was applied (Vauterin et al., 2000). To sort out the problem of classification of the genus *Xanthomonas*, a series of studies has been undertaken in the last decade to address the species delineation within the genus which includes individual biochemical and physiological tests (Van den Mooter and Swings, 1990), electrophoresis of whole cell proteins (Vauterin et al., 1993a,b), and gas chromatography of cellular fatty acids (Yang et al., 1993). Thus, in order to derive at a strong conclusion, more than 1000 isolates from diverse origins were analysed using the fast and sensitive fingerprinting technique of DNA hybridisation (Vauterin et al., 2000). The most recent classification of *Xanthomonas* spp. based on the sequence similarity of 16S rRNA gene sequences deciphers the genus with three phylogenetic lineages i.e. first one including *X. arboricola*, *X. axonopodis*, *X. bromi*, *X. campestris*, *X. cassavae*, *X. cucurbitae*, *X. codiaei*, *X. fragariae*, *X. hortorum*, *X. melonis*, *X. oryzae*, *X. pisi*, *X. populi*, *X. vasicola*, and *X. vesicatoria*; the second including *X. albilineans*, *X. hyacinthi*, *X. theicola*, and *X. translucens*; and the third consisting of only *X. sacchari* (Hauben et al., 1997).

6.1.1. DNA–DNA hybridization

DNA–DNA hybridisation and the decrease in thermal stability of the hybrids have provided a useful tool for the demarcation of species within the genus *Xanthomonas*. There are many pitfalls regarding the use of this technique for molecular taxonomic studies for *Xanthomonas* spp. because of the limitations of converting DNA hybridisation value into percentage of whole genome sequence similarity. However, extensive DNA–DNA hybridisation reveals the natural relationships between *Xanthomonas* strains which demonstrate that a number of pathovars are composed of two or more unrelated genotypes and should not be regarded as biological entities (Vandamme et al., 1996). Though, first *Xanthomonas* was described as a monotypic genus, later it was divided into groups A and B (Stall et al., 1994) and subsequently 183 reported strains have been classified into 20 different species based on the DNA–DNA hybridisation technique (Rodriguez-R et al., 2012). Still, there exist many conflicts regarding the use of this technique for phylogeny purposes, for example, classification of *Xanthomonas fuscansaurantifolii* (Schaad et al., 2006) also known as *X. axonopodis* pv. "aurantifolii" was originally identified to be the DNA homology group of *X. axonopodis* (Vauterin et al., 1995); however, later it was designated to be *X. fuscans* based on other DNA sequence based molecular techniques and production of water soluble brown pigment and host range (Schaad et al., 2005).

6.1.2. 16S rRNA gene sequences

Analysis of 16S rRNA gene sequences in *Xanthomonas* species reveals very limited diversity because of the high level of sequence similarity in this region and thus the 16S rRNA gene is considered to be less suitable for distinguishing between closely related strains (Hauben et al., 1997). However, recent studies revealed the identification of *Xanthomonas* strains into two groups *X. axonopodis* pv. *citri* strains causing type A CBC were placed in one group, while *X. axonopodis* pv. *aurantifolii* strains causing type B and C CBC in citrus species were more closely related to a strain of *Xanthomonas*. The mean level of sequence similarity between the subgroups is 99.6% whereas, the maximum number of base difference was 12 (Cubero and Graham, 2002).

Furthermore, many other house-keeping genes have also been targeted for the identification of *Xanthomonas* species which include *gyrB* sequences. The advantage of using *gyrB* sequences is that the gene sequence can discriminate all the *Xanthomonas* species type strains except those of *X. euvesicatoria*, *X. perforans* and *X. alfalfae*, which form a species complex with identical *gyrB* sequences and are referred as the *X. euvesicatoria* species complex (Parkinson et al., 2009).

6.1.3. DNA base ratio

Even though the use of G + C ratio is a very old and simple analysis for bacterial classification, it is still useful for distinguishing between very closely related species. For two closely related genera *Xanthomonas* and *Stenotrophomonas* percentage of the G + C ratio varies between 63 and 70% and 65 and 68% respectively (Vandamme et al., 1996). However, recently, the G + C ratio of some individual genes have been taken into account for a clear understanding of the molecular identification. GC contents of the variable T3SP genes based on the genes from the three sequenced strains were significantly lower (56.3%) than in the rest of the genes (65%) or the core type III secretome (62%) which is useful for the phylogeny of *Xanthomonas* strains (Guy et al., 2013).

6.1.4. Numerical analysis of phenotypic features

The most extensive work regarding the numerical analysis of phenotypic features of *Xanthomonas* species have been carried out by Van den Mooter and Swings (1990). Based upon their analysis of 295 phenotypic features in 266 *Xanthomonas* or related strains, nine phena were proposed. However, a recent study has taken into account uncommon phenotypic features previously not taken into account, such as absorption spectra of pigments, biopolymer production and other phenotypic features to distinguish *X. campestris* from other mucoid yellow pigmented soil bacteria (Soudi et al., 2011).

6.1.5. Monoclonal antibodies

Monoclonal antibodies have been reported for many *Xanthomonas* species and their pathovars that allow the identification of a large number of isolates. This technique has proven to be a much simpler way to identify pathovars among the *Xanthomonas* species (Vauterin et al., 1993a,b). From the point of view of agricultural microbiology, this technique is important not only to identify the strains causing black rot in the field, but also to monitor the progress of black rot caused by serologically distinct strains originating from distinct inoculum sources. Many pathovars of *Xanthomonas* species can be detected using this technique like *X. c.* pv. *campestris*, *X. campestris* pv. *vesicatory*, *X. c.* pv. *amoraciae*, *X. c.* pv. *phaseoli var.*

fuscans (Yuen et al., 1986; Franken et al., 1992; Leite et al., 1994). Though many advanced techniques have been employed for identification of *Xanthomonas* sp., use of monoclonal antibodies still persists as one of the major tools for identification (Robene-Soustrade et al., 2006; Park et al., 2009).

6.1.6. Whole cell protein analysis

SDS-PAGE is one of the important molecular techniques for identification up to species level by analysis of whole cell proteins and the technique is relatively simple and rapid to perform. However, for proper identification, extensive data is required to cover all known target species. In SDS-PAGE, whole cell proteins are separated on the basis of their polypeptide length to estimate their molecular weight. Fatima et al (2012) distinguished seven isolates of *Xanthomonas* species into four different groups i.e. *X. campestris*, *X. maltophilia*, *X. axonopodis* and *X. nematophilus* on the basis of whole cell protein analysis. Similarly, *X. arboricola* pv. *corylina* was clustered into three groups by protein analysis studies (Scortichini et al., 2002). These protein electrophoretic clusters can be used to select strains for subsequent DNA–DNA hybridisation.

6.1.7. Cellular fatty acid analysis

Composition of cellular fatty acid of *Xanthomonas* strains is very complex and over 65 different fatty acids have been found that vary as per the occurrence of branched and hydroxyl branched fatty acids (Alexandrino et al., 2001). Many studies have been conducted by using whole-cell FAMEs to identify 975 strains that include seven representatives of *Xanthomonas* (*X. albilineans*, *X. axonopodis*, *X. campestris*, *X. fragariae*, *X. maltophilia*, *X. oryzae* and *X. populi*). In 99% of the *Xanthomonas* species nine sets of fatty acids have been reported as 11:0 iso, 11:0 iso 3OH, 12:0 3OH, 13:0 iso 3OH, 15:0 iso, 16:1 cis 9, 16:0, 17:1 iso F and 17:0 iso (Yang et al., 1993). However, one of the major disadvantages of this technique involves the discrepancy between RFLP and FAME analyses that was not apparent when the DNA re-association method was used (Roberts et al., 1998).

6.2. Polyphasic taxonomy of the genus Campylobacter

Members of the genus *Campylobacter* are Gram-negative, spiral, microaerophilic, motile either with unipolar or bipolar flagella, which include species that are among the major causes of food borne diseases in developing and developed countries. Though the genomic structure of this genus is very complex, significant taxonomic advances have been achieved for its correct and rapid identification. It may produce inflammatory responses, bloody diarrhoea, periodontitis or dysentery syndrome that includes cramps, fever and pain. Based upon phylogenetic studies, *Campylobacter* has been separated into three genera *Campylobacter*, *Arcobacter* and *Helicobacter* (Linton et al., 1997). Several techniques and parameters used in phylogenetic studies of *Campylobacter* and related genera have been discussed below.

6.2.1. rRNA gene homology

Identification of *Campylobacters* up to exact species level using phenotypic tests is difficult due to the insufficiency of these criteria to adequately identify all species. Though many molecular methods have been reported for rapid reliable identification of *Campylobacter* species, most of them rely on 16S rRNA gene sequencing. There are reports of 16 species and 6 sub-species of *Campylobacter* whereas the rRNA gene sequence reveals many other *Campylobacter* spp., especially non-jejuni species (Kawasaki et al., 2008). In many studies, certain *Campylobacter* spp. possess IVS and ITS in 16S rRNA gene mostly for *C. sputorum*, *C. curvus*, *C. rectus*, *C. helveticus* and *C. hyointestinalis*, the base sequences of which ranges between 227 and 229 bp (Etoh et al., 1998; Man et al., 2010) which may be explored for more appropriate rapid identification of the *Campylobacter* spp. The ITS region of 16S rRNA gene in *Campylobacter* spp. has been reported to produce the highest mean pairwise percentage difference (35.94%) compared to the other molecular markers i.e. 16S (5.34%) and 23S rRNA gene (7.29%) alone (Man et al., 2010). Though 16S rRNA gene sequence is of wide use for the identification of *Campylobacter* species, *rpoB* gene is reported to have higher resolution than 16S rRNA gene which is sufficient for the identification of many closely related species as well as most subspecies except *C. coli* and *C. jejuni* (Korczak et al., 2006; Hansson et al., 2008).

6.2.2. DNA–DNA hybridization

Significant DNA hybridisation has been reported between *C. fetus* and *C. hyointestinalis*; between *C. jejuni*, *C. coli*, *C. lari*, and *C. upsaliensis*; and between *C. rectus* and *C. showae*, however, most of the DNA–DNA hybridization studies have been performed before the phylogenetic relationships were established by rRNA sequence based studies (Vandamme et al., 1996). A recent study reveals different hybridisation patterns between members of two species and also between individual strains of the same species, which suggests that, DNA finger printing with the 16S rRNA gene probe might be a potentially useful technique for identification of *Campylobacter* epidemiological studies (Romaniuk and Trust, 2006). However, DNA–DNA hybridisation studies should be coupled with 16S rRNA gene sequencing and other biochemical differentiations for the description of a novel species of *Campylobacter* (Logan et al., 2000).

6.2.3. DNA base ratio

The (G + C)% values in *Campylobacter*, *Arcobacter* and *Helicobacter* species vary between 30 and 46%, 27 and 31% and 35 and 44% respectively (Vandamme et al., 1996).

6.2.4. Classical phenotypic characteristics

Campylobacter and its relatives exhibit a diverse cellular morphology with the corkscrew like morphology considered to be a typical form. Cells of species other than *C. jejuni*, *C. coli*, *C. lari*, and *C. upsaliensis* are predominantly S-shaped rods, bent rods or straight rods. Several other biochemical characteristics used for *Campylobacter* identification include hippurate hydrolysis and sensitivity to nalidixic acid, metronidazole, and 2,3,5-TTC (Walder et al., 1983). Similarly, flagellation types can also be used for this purpose as most of the *Campylobacter* species are characterised by a single polar or bipolar flagellum. Hence, number and flagellation type are not relevant taxonomic parameters whereas, structure of the flagellum are the important characteristics for the purpose of identification. This is evident from the fact that, all taxa of the *Campylobacter* and *Arcobacter* rRNA homology groups have unsheathed flagella whereas a flagellar sheath is found in all *Helicobacter* species (Vandamme et al., 1991).

6.2.5. Repository quinone components

Unlike most other Gram-negative bacteria, the repository chain of *Campylobacter* consists of only menaquinones. However, the other species consist of menaquinone-6 and thermoplasmaquinone as the major components. Hence, the analysis of repository quinone compounds can be used for fruitful characterisation of *Campylobacter* species up to strain level (Vandamme et al., 1996). RPHPLC chromatogram of menaquinones can be used for the identification and characterisation of *Campylobacter* isolates by comparison with the ATCC standard cultures. The functional groups targeted for this purpose may include individual fatty acids, aldehydes, dimethylacetals, and quinones (Moss et al., 1990). Though advanced techniques of identification and phylogenetic analysis of *Campylobacter* species have been developed, analysis of repository quinone components cannot be ruled out completely when the polyphasic taxonomic approach is taken into account.

6.2.6. Cellular fatty acid analysis

Many studies have been carried out so far for the use of fatty acid analysis for identification and differentiation of *Campylobacter* species.

Thus, most of the known *Campylobacter* taxa have been subdivided into many gas–liquid chromatography groups by the difference in their cellular fatty acid components. Advances in taxonomic studies of *Campylobacter* species reveals that this technique is not very useful as the difference in fatty acid composition is no longer considered to be genus specific. However, the fatty acid analysis should be coupled with other tests like DNA hybridisation, 16S rRNA gene sequencing and phenotypic characteristics for correct assessment of a novel genus or species of *Campylobacter* (Vandamme et al., 2010).

6.2.7. Protein analysis

Many advances have taken place for the use of proteomics to identify *Campylobacter* species. Several protein biomarkers have been identified by TOF MS of bacterial cell mass ionised by MALDI. These protein biomarkers include DNA-binding protein HU, translation IF-1, c553, a transthyretin-like periplasmic protein, chaperonin GroES, Trx, and ribosomal proteins (Fagerquist et al., 2006). The whole cell protein profiles of all the strains are almost similar besides the presence of a dense band in thermotolerant *Campylobacter* species whereas, the low molecular mass region of the protein profile has additional heterogeneity (Duim et al., 2004). Recently, Moolhuijzen et al (2009) targeted various toxin and membrane bound proteins for the identification of *Campylobacter* species up to subspecies and strain level by targeting proteins responsible for bacterial adherence, motility, two-component system, toxin and resistance and membrane proteins (Moolhuijzen et al., 2009).

6.3. Polyphasic taxonomy of the genus Bacillus

The genus *Bacillus* is a systematically diverse assemblage of aerobic, Gram-positive spore forming organisms which exhibit a wide range of nutritional requirements, growth conditions, DNA base compositions as well as major amino acid compositions of the cell wall (Claus and Berkeley, 1986). 16S rRNA gene comparative analysis (Priest, 1993), phenotypic features (Priest et al., 1988), multilocus sequence typing (Helgason et al., 2004) and amplified fragment length polymorphism (Mignot et al., 2004) revealed five to six phylogenetically distinct clusters of *B. cereus* group that includes *B. anthracis*, *B. mycoides*, *B. pseudomycoides*, *B. thuringiensis*, *B. weihenstephanensis* and *B. cereus* (Ash et al., 1991; Rasko et al., 2005). Many additional groups were also discovered subsequently belonging to alkaliphilic bacilli (Nielsen et al., 1994) and a new species named *B. cytotoxicus* (Lapidus et al., 2008). The phylum *Firmicutes*, class *Bacilli*, order *Bacillales*, includes seven families that lie within the taxonomy of the genus *Bacillus* and related organisms (Ludwig et al., 2007; Logan et al., 2009). However, the newly classified aerobic, endospore forming bacteria have been reclassified as i) *Bacillaceae*, ii) *Alicyclobacillaceae*, iii) *Paenibacillaceae*, iv) *Planococcaceae*, v) *Thermoactinomycetaceae*, vi) *Pasteuriaceae* and viii) *Sporolactobacillaceae* (Cihan et al., 2012).

6.3.1. DNA–DNA hybridization

Morphological and physiological characteristics may be used to classify microorganisms up to a certain degree of confidence, but genetic and genomic features enable a deeper resolution for the differentiation, thus DNA–DNA hybridisation has been placed in a key position in microbial species demarcation. The prerequisites for DNA–DNA reassociation requires (i) shearing the gDNA of the assayed organism and the gDNA of the reference organism(s) into small fragments of 600–800 bp (ii) heating the mixture of DNA fragments from both strains to dissociate the DNA double-strands and (iii) subsequently decreasing the temperature until the fragments reanneal (Peak et al., 2011). The DNA–DNA hybridisation test has clear advantages over other tests for the identification of *Bacillus* species, e.g. *B. subtilis* like organisms can never be distinguished from each other by conventional phenotypic tests and they possess almost identical 16S rRNA gene sequences (99.2–99.6%), but they can be differentiated by DNA–DNA hybridisation results (Chun and Bae, 2000).

6.3.2. 16S rRNA gene sequences

As *Bacillus* is a highly extensive heterogeneous group with more than 222 validly described species, it needs an efficient method for species differentiation. In this regard, sequencing of the 16S rRNA gene has proven to be a major breakthrough for the generation of large public sequence databases and the determination of the exact nature of the sequence data (Blackwood et al., 2004). The hypervariable region (nucleotides 70 to 344) is adequate to discriminate between all the *Bacillus* species except between *B. cereus* and *B. anthracis* and between *B. mojavensis* and *B. atrophaeus*. In order to overcome these problems, DNA sequencing of certain house-keeping genes provide more sensitive DNA sequencing subtyping than 16S rRNA sequencing for many *Bacillus* species that may include the protein coding genes like *rpoB*, *gyrB*, *nifD*, *recA*, *atpD* (Ki et al., 2009). Although 16S rRNA gene sequencing is extremely useful for classification of *Bacillus* species, it has low discriminatory power at species level and for some genera of *Bacillus*. The type strains of *B. globisporus* and *B. psychrophilus* share 99.5% sequence similarity with regard to their 16S rRNA genes, and yet at the DNA level exhibit only 23 to 50% relatedness in reciprocal hybridisation reactions (Janda and Abbott, 2007).

6.3.3. DNA base ratio

Gram-positive bacteria are known for their low G + C content, however, there are some exceptions in Firmicutes as some of them may have aG + C content as high as 55% as in the case of *Geobacillus thermocatenulatus*. The average G + C ratio in *Bacillus licheniformis* is 46.2% (Veith et al., 2004) whereas in case of *B. subtilis* it is reported to be 43.46% (Yeo et al., 2012).

6.3.4. Numerical analysis of phenotypic features

The signature phenotypic and physiological properties for proper identification of *Bacillus* species include VP test, lecithinase production, mannitol utilisation, starch hydrolysis, nitrate reduction, decomposition of L-tyrosine, β-haemolysis, length and width of spores and cells (Lopez and Alippi, 2007). The most distinguishable characteristic features of *Bacillus* species include the use of citrate as sole carbon source, production of hydrogen sulphide, utilisation of esculin and D-ribose and least utilisation of inulin and potassium gluconate (Freitas et al., 2008). Common *Bacillus* species infer a common line of biochemical test patterns that include catalase and oxidase positive, spore-forming rods. Most of the strains utilised dextrose, D-fructose, gelatine, casein, urea, mannose, trehalose, sucrose whereas, negative for indole production, L-rhamnose, D-melzitose phenylalanine deamination, hydrogen sulphide production (Mohanty et al., 2011).

6.3.5. Whole cell protein analysis

Use of SDS-PAGE profiles of whole cell proteins offers many advantages for identification of *Bacillus* species. It is a fairly fast and easy method of identifying a large number of strains and it has an adequate level of taxonomic resolution at both species and subspecies levels (Tae-Woon et al., 2010). Protein profiles produced by SDS-PAGE of whole cell extract correlates with the results of DNA–DNA hybridisation suggesting its use in rapid bacterial identification (Berber, 2004). However, for identification of *Bacillus* species this technique may be coupled with others like serotyping, bacteriophage typing, bacteriocin activities, antibiogram and biotyping, plasmid typing, analysis of cellular fatty acid content, native-PAGE, small-subunitribosomal RNA sequencing and genome analysis for a better resolution and proper identification (Berber and Cokmus, 2001).

6.4. Polyphasic taxonomy of the family Comamonadaceae

Most of the saprophytic, phytopathogenic bacteria with diverse biochemical properties have been placed under the family Comamonadaceae. There have been many reports of the existence of members of the Comamonadaceae in soil, water, industrial wastes,

clinical samples as well as infected plant materials. Due to their physiological flexibility, they play a major role in biodegradation of xenobiotics, recalcitrant, oil-derived wastes and many other pollutants. Consistent with recent classification, this family includes the following genera *Acidovorax*, *Alicycliphilus*, *Brachymonas*, *Caldimonas*, *Comamonas*, *Curvibacter*, *Delftia*, *Diaphorobacter*, *Doohwaniella*, *Giesbergeria*, *Hydrogenophaga*, *Hylemonella*, *Lampropedia*, *Macromonas*, *Malikia*, *Ottowia*, *Panaciterramonas*, *Pelomonas*, *Polaromonas*, *Ramlibacter*, *Rhodoferax*, *Schlegelella*, *Simplicispira*, *Variovorax*, *Verminephrobacter*, and *Xenophilus* (http://www.ncbi.nlm.nih.gov/nuccore/?term=Comamonadaceae). However, due to the expanding number of members under this family, the phylogenetic relationships within the family Comamonadaceae is yet to be established completely (Wen et al., 1999).

6.4.1. rRNA gene similarities

The major problem associated with the classification of Comamonadaceae is the inadequate number of 16S rRNA gene sequences in the database. Five genera have been proposed with nine rRNA subbranches. *Comamonas* encompasses four rRNA subbranches and *Acidovorax* encompasses two, while *Variovorax*, *Hydrogenophaga* and *Xylophilus* each represent a single subbranch. Each of the four remaining rRNA subbranches contains a misclassified *Aquaspirillum* species: *A. delicatum*, *A. gracile*, *A. metamorphum* and *A. anulus*. *A. sinuosum* and *A. giesbergeri* are closely related and belong to the last rRNA subbranch (Vandamme et al., 1996). Comparative analysis of the 16S rRNA gene has been reported in many bacterial species including *Giesbergeria voronezhensis*, *Giesbergeria kuznetsovii*, *Aquaspirillum annulus*, *Giesbergeria annulus*, and *Simplicispira metamorpha* (Grabovich et al., 2006). However, compared to other groups of bacteria, use of the 16S rRNA gene for identification of Comamonadaceae needs further research and development.

6.4.2. DNA–DNA hybridizations

For delineation of species placed within the family Comamonadaceae, DNA relatedness has proven to be a useful tool for speciation. In general, DNA–DNA cross-hybridisation of >70% indicates that the two bacteria belong to the same species (Hagstrom et al., 2002), however, in case of Comamonadaceae, 40% DNA binding can ascertain speciation (Willems et al., 1990). The lacuna behind this technique is the non-correlation of DNA–DNA hybridisation data with phenotypic differences, thus nomenclature changes may not be possible by using hybridisation results to describe and differentiate between taxa (Vandamme et al., 1996).

6.4.3. Amplified rDNA restriction analysis

Ribosomal rDNA fragments comprising 16S rDNA, the 16S–23S rDNA spacer region and a part of the 23S rDNA may be amplified using conserved primers followed by restriction digestion with a suitable restriction enzyme. This forms a rapid, reliable identification practice for members of the Comamonadaceae family (Vaneechoutte et al., 1992). This technique popularly called as ARDRA has been implemented for many members of this family including *Polaromonas*, *Hydrogenophaga*, *Acidovorax*, *Macromonas*, *Comamonas* (Mukherjee et al., 2013), *Paenibacillus* (Weid et al., 2002), *Proteobacterium* (Lagace et al., 2004) and novel species of *Comamonas nitrativorans* sp. nov. (Etchebehere et al., 2001).

6.4.4. DNA base ratio

The DNA base ratio varies widely among the members of the family Comamonadaceae which is found to be one of the unique characteristic features of this group of organisms. The average (G + C) % of all the members belonging to this group has been determined and found to vary between 57 and 70% (Hildebrand et al., 2010). Certain misclassified species have lower G + C values (57–58%) as observed in case of *Aquaspirillium* species (Mechichi et al., 2003).

6.4.5. Whole cell protein analysis

Whole cell protein analysis and numerical interpretation can group large numbers of genotypically related Comamonadaceae strains, which corroborates with DNA hybridization groups. In some exceptions, certain genera like *Xylophilus* and *Comamonas* show aberrant profiles due to the presence of one or more heavy protein bands that may affect the numerical interpretation (Willems et al., 1991). Thus, during the polyphasic taxonomic approach for the identification of Comamonadaceae, in addition to other genetic approaches, the whole cell protein approach is a useful technique as reported for the identification of environmentally important strains (Boon et al., 2001) viz. *Comamonas odontotermitis* sp. nov. (Chou et al., 2007) and *Acidovorax caeni* sp. nov. (Heylen et al., 2008).

6.4.6. Numerical analysis of phenotypic features

Classical physiological and morphological characteristics including carbon assimilation tests followed by computer assisted numerical interpretation distinguishes various genera of Comamonadaceae. These clusters are correlated with other techniques and the phenotypic results may be used to discriminate the taxonomic levels. The major advantage is that in current practice, when the phenotypic results do not match with the other genomic and genetic techniques, the phenotypic characteristics play a decisive role for the identification of members of the Comamonadaceae group (Khan et al., 2002).

6.4.7. Immunotyping

The immunotyping technique has been described for two early members of the Comamonadaceae group i.e. *Comamonas* and *Hydrogenophaga* to unravel their internal structure (Wauters et al., 2003). The usefulness of this technique is still being debated due to its failure to correlate with other molecular techniques like DNA–DNA hybridisation and rRNA gene sequence information.

6.4.8. Cellular fatty acid analysis

For identification of the members of Comamonadaceae family, cellular fatty acid analysis has been applied to a few genera and other methods have been used for more superior identification and taxonomic resolution. Nevertheless, a general conclusion may not be drawn based on the observation of this single technique and it needs to be coupled with other techniques. The major fatty acids which are detected during the fatty acid profiling of members of Comamonadaceae include palmitoleic acid, palmitic acid, cis-vaccenic acid and the absence of 3-hydroxypalmitic acid (Bruland et al., 2009; Gomila et al., 2010).

6.4.9. Polyamine pattern

Polyamines can be analysed using HPLC equipped with a fluorescent detector and a reverse phase column, however, its ability to discriminate among the members of Comamonadaceae group is very unclear and all the members of proteobacter β-subclass are characterised by the presence of 2-hydroxyputrescine.

7. Limitations of the advanced molecular techniques

Though various culture independent as well as culture dependent techniques are available now-a-days for molecular identification of bacteria from the environment (Fig. 8), each of them contains certain limitations and none of them is perfect regarding the identification of an unknown bacterium. The 16S rRNA gene sequencing based identification procedure is followed widely in both culturable as well as unculturable approaches but there also remain certain questions regarding its use from the identification point of view. The recent discovery of horizontal gene transfer, questions the usefulness of this procedure of identification (Harrington et al., 2012). If horizontal transfer of the 16S rRNA gene prevails in nature it will cause severe complications regarding the identification of the bacterial isolates. Similarly, though we consider the integrity of the 16S rRNA gene in the bacterial genome, its sequence also tends to vary by mutation. Detection of a

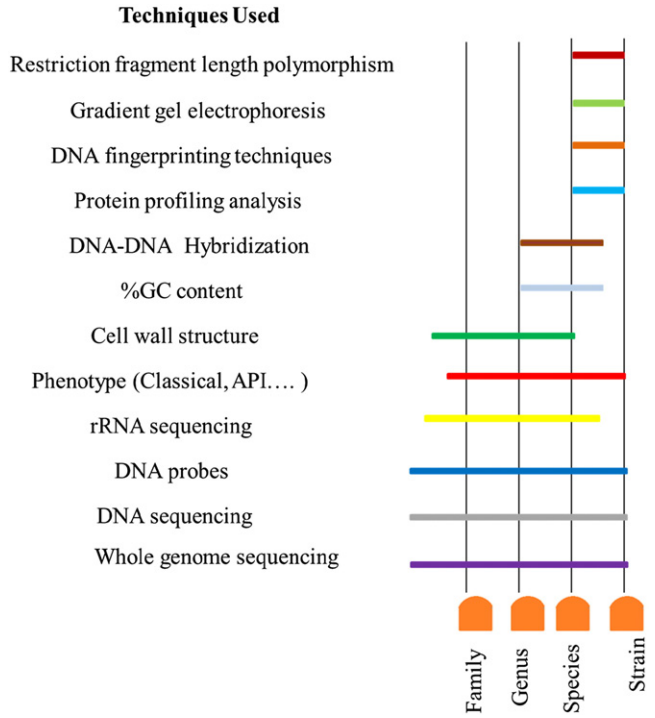

Fig. 8. Taxonomic resolutions of the currently used techniques.

point mutation at position 857 (G to A) at the 16S rRNA gene of *Neisseria gonorrhoeae* confirms the fact that this gene may also be affected by mutational events (Gurtler et al., 2012). If mutation occurs in the 16S rRNA gene of any bacterium, it no longer remains as the conserved region and the ambiguity regarding this gene sequence based identification is lost. In addition, after sequencing when it is matched with the database, many of them fall under the unculturable group and many species cannot be identified due to the discrepancy in the database as well as unavailability of the sequences in the database. Moreover, this technique alone becomes less reliable when species level identification is concerned. Earlier, >97% of similarity level has been proposed for bacterial speciation and >0.5% difference is the indication of a new species (Stackebrandt and Goebel, 1994; Palys et al., 1997). However, with the advancements of molecular tools different cutoffs have been set up for different studies, e.g. >99 and >97% of sequence similarity has been proposed for the identification of species and genus respectively (Janda and Abbott, 2007). Thus, 16S rRNA gene sequencing could not differentiate reliably between two closely related bacterial species. Similarly, the entire sequence of 16S rRNA gene (~1500 bp) having less than 1% undetermined positions are required for the identification of a new taxon, which is not possible practically. Though the sequencing based procedures have been advanced drastically, minimising the sequencing error at any level, the sequences submitted over a decade may not all be accurate (Clayton et al., 1995). It is possible that the sequences deposited recently are more accurate than those sequences submitted earlier and this should be noted when identifying any isolate.

Similar problems appear with the other house-keeping gene sequencing for identification practices. In addition to that, the relatively newer targeted gene sequences cannot be matched with the ample database containing the huge amount of sequences to be matched.

In contrary to sequence based identification, PCR based identification is extremely sensitive, can detect a single target molecule and can also be applied to culture independent samples. PCR fingerprinting techniques allow discrimination of isolates that would otherwise not be distinguishable and it introduces less bias in comparison to culture followed by PCR or other analysis. However, when these fingerprinting techniques are applied to environmental samples, they do not discern between live or dead cells and simply detect the nucleic acid, hence enumeration is not possible by these methods. The chances of contamination and nonspecific priming may lead to false positive results. During ERIC, REP or any PCR fingerprinting technique, conditions such as DNA quality and PCR temperature can result in non-reproducible fingerprints. Hence, care should be taken for either types of PCR for comparison.

Gel based detection techniques like gradient gel electrophoresis are extremely useful for the identification of any microbial community both in culturable as well as unculturable approaches because of their high detection rate and extreme sensitivity. As this method is very simple and nonradioactive, PCR fragments can be readily isolated from the gel and used for sequencing reactions. This method also has some lacuna. This procedure is completely dependent on PCR and the sensitivity and accuracy of PCR has a huge impact on the performance and results of this kind of electrophoresis. Similarly, those genes which are exceptionally rich in their GC content cannot be analysed easily by DGGE technique (Temmerman et al., 2003). In addition, the decrease in sensitivity of the results occurs when large DNA fragments are amplified and analysed as the technique has become less useful when the amplicon size is more than 400 bp. As the genes targeted for identification purpose like 16S rRNA gene is of larger size, the full sequence of these genes cannot be used for this purpose and incorrect or inaccurate identification occurs. These techniques are also not applicable for extremely complex communities Gafan and Spratt (2005).

Flow cytometry based detection of microorganisms is the most advanced method of detection now-a-days when the accuracy as well as detection time is concerned. Antibodies are currently changing the way in which we identify microbes, thus making it easier and faster. Their specificity and use of fluorochrome labelled antibodies to specific antigens renders them one of the most powerful tools for the identification of microbes. The microorganisms no longer need to be cultivated and can be directly identified both qualitatively and quantitatively from the samples studied. However, the major disadvantage of this technique is the limited availability of antibodies directed against any particular microorganism (Gunasekera et al., 2000). As a huge variety of both known and unknown bacteria are present in the environment, it is practically impossible to develop the antibodies against them which severely limit this technique. However, this technique is quite useful for the detection of the targeted, physiologically important microorganisms directly from environmental samples.

Regarding the protein profiling based identification system, it is much more useful for biological samples as well as for those groups which cannot be specified based on their sequence dissimilarities and by other available techniques. It is an exclusively culture dependent practice and the group of bacteria which cannot be cultured in the normal laboratory conditions cannot be identified following this practice. At the same time, it is a time consuming process as a relatively long labelling incubation time is required for this process.

Similarly, the use of MALDI-TOF MS for protein profiling is an expensive approach and during esterification it is not specific to the carboxy terminal of a peptide, hence may lead to ambiguity in the result. Isotope labelling of proteins is not always a practical possibility, labelling with stable isotopes is very expensive, sometimes the isotopic levels exhibit chromatography shifts which can make quantification of differentially labelled peptides difficult computationally and there may not be enough different isotopes for simultaneous quantification of proteins from different samples (De Bruyne et al., 2011). One more major disadvantage of peptide quantification by signal intensity includes obtaining experimental variation and signal noise that affects the quantitative value and accuracy.

8. Concluding remarks

Technological progress has led the advanced techniques to completely dominate the conventional techniques of microbial identification.

Despite this, many microorganisms from the environment are yet to be identified and assigned a perfect nomenclature. In the future, more data will be available and more new species of microorganisms will be discovered because of the automation, software development and combination of different databases. Now-a-days there is an increasing access to microbial genomes due to the accumulation of numerous DNA sequences which ultimately leads to a higher level of accuracy and reliability of results. Though polyphasic taxonomy is a useful technique to meet the challenge of identifying any unknown strain, new mathematical and informative strategies should be developed for the possible development of a synthetic taxonomy. Thus, the conventional biochemical basis of identification practices cannot be replaced completely and the vast majority of the data should lead to a perfectly reliable and stable identification and classification system. Though many advanced techniques are available now-a-days for the identification of any unknown bacterium, a far less number of the total number of microbial species have been discovered and identified till now; many of them are yet to be cultured under laboratory conditions and some of them may possess certain unique characteristic features. Microbial taxonomy and biosystematics is a major modern discipline which needs further financial and intellectual support to determine its role in biotechnology, biodiversity, agriculture, medical science and environmental science.

Acknowledgements

Authors would like to acknowledge the authorities of NIT, Rourkela for providing facilities and fellowships to H.R.D., N.M. and J.C. S.K. thanks Department of Biotechnology, Government of India for a research fellowship. S.D. acknowledges research grants from the Department of Biotechnology, Ministry of Science and Technology and National Fund for Basic, Strategic and Frontier Application Research in Agriculture (NFBSFARA), Government of India.

References

Achtman, M., 2008. Evolution, population structure, and phylogeography of genetically monomorphic bacterial pathogens. Annu. Rev. Microbiol. 62, 53–70.
Albufera, U., Bhugaloo-Vial, P., Issack, M.I., Jaufeerally-Fakim, Y., 2009. Molecular characterization of *Salmonella* isolates by REP-PCR and RAPD analysis. Infect. Genet. Evol. 3, 322–327.
Alexandre, A., Laranjo, M., Peter, J., Young, W., Oliveira, S., 2008. dnaJ is a useful phylogenetic marker for alphaproteobacteria. Int. J. Syst. Evol. Microbiol. 58, 2839–2849.
Alexandrino, M., Knief, C., Lipski, A., 2001. Stable-isotope-based labeling of styrene-degrading microorganisms in Biofilters. Appl. Environ. Microbiol. 67, 4796–4808.
Arahal, D.R., Vreeland, R.H., Litchfield, C.D., Mormile, M.R., Tindall, B.J., Oren, A., Bejar, V., Quesada, E., Ventosa, A., 2007. Recommended minimal standards for describing new taxa of the family Halomonadaceae. Int. J. Syst. Evol. Microbiol. 57, 2436–2446.
Ash, C., Farrow, J.A.E., Wallbanks, S., Collins, M.D., 1991. Phylogenetic heterogeneity of the genus *Bacillus* revealed by comparative analysis of small-subunit-ribosomal RNA sequences. Lett. Appl. Microbiol. 13, 202–206.
Ast, J.C., Urbanczyk, H., Dunlap, P.V., 2009. Multi-gene analysis reveals previously unrecognized phylogenetic diversity in *Aliivibrio*. Syst. Appl. Microbiol. 32, 379–386.
Bano, N., Hollibaugh, J.T., 2002. Phylogenetic composition of bacterioplankton assemblages from the Artic Ocean. Appl. Environ. Microbiol. 68, 505–518.
Begerow, D., John, B., Oberwinkler, F., 2004. Evolutionary relationships among beta-tubulin gene sequences of basidiomycetous fungi. Mycol. Res. 108, 1257–1263.
Berber, I., 2004. Characterization of *Bacillus* species by numerical analysis of their SDSPAGE protein profiles. J. Cell Mol. Biol. 3, 33–37.
Berber, I., Cokmus, C., 2001. Characterization of *Bacillus sphaericus* strains by Native-PAGE. Bull. Pure Appl. Sci. 20, 17–21.
Berneche-D'Amours, A., Ghinet, M.G., Beaudin, J., Brzezinski, R., Roy, S., 2011. Sequence analysis of *rpoB* and *rpoD* gene fragments reveals the phylogenetic diversity of actinobacteria of genus *Frankia*. Can. J. Microbiol. 57, 244–249.
Bisharat, N., Cohen, D.I., Maiden, M.C., Crook, D.W., Peto, T., Harding, R.M., 2007. The evolution of genetic structure in the marine pathogen, *Vibrio vulnificus*. Infect. Genet. Evol. 7, 685–693.
Blackwood, K.S., He, C., Gunton, J., Turenne, C.Y., Wolfe, J., Kabani, A.M., 2000. Evaluation of *recA* sequences for identification of *Mycobacterium* species. J. Clin. Microbiol. 38, 2846–2852.
Blackwood, K.S., Turenne, C.Y., Harmsen, D., Kabani, A.M., 2004. Reassessment of sequence-based targets for identification of Bacillus species. J. Clin. Microbiol. 42, 1626–1630.
Boon, N., Goris, J., Vos, P.D., Verstraete, W., Top, E.M., 2001. Genetic diversity among 3-chloroaniline- and aniline-degrading strains of the Comamonadaceae. Appl. Environ. Microbiol. 67, 1107–1115.
Brown, M.R.W., Kornberg, A., 2004. Inorganic polyphosphate in the origin and survival of species. Proc. Natl. Acad. Sci. U. S. A. 101, 16085–16087.
Bruland, N., Bathe, S., Willems, A., Steinbuchel, 2009. *Pseudorhodoferax soli* gen. nov., sp. nov. and *Pseudorhodoferax caeni* sp. nov., two members of the class Betaproteobacteria belonging to the family Comamonadaceae. Int. J. Syst. Evol. Microbiol. 59, 2702–2707.
Burton, Z., Burgess, R.R., Lin, J., Moore, D., Holder, S., Gross, C.A., 1981. The nudeotide sequence of the cloned *rpoD* gene for the RNA polymerase sigma subunit from *E. coil* K12. Nucl. Acids Res. 9, 2889–2903.
Capella-Gutierrez, S., Kauff, F., Gabaldon, T., 2014. A phylogenomics approach for selecting robust sets of phylogenetic markers. Nucl. Acids Res. http://dx.doi.org/10.1093/nar/gku071.
Carbonnelle, E., Mesquita, C., Bille, E., Day, N., Dauphin, B., Beretti, J.L., Ferroni, A., Gutmann, L., Nassif, X., 2011. MALDI-TOF mass spectrometry tools for bacterial identification in clinical microbiology laboratory. Clin. Biochem. 44, 104–109.
Chan, M.S., Maiden, M.C., Spratt, B.G., 2001. Database-driven Multi Locus Sequence Typing (MLST) of bacterial pathogens. Bioinformatics 17, 1077–1083.
Charpentier, B., Branlant, C., 1994. The *Escherichia coli gapA* gene is transcribed by the vegetative RNA polymerase holoenzyme E sigma 70 and by the heat shock RNA polymerase E sigma 32. J. Bacteriol. 176, 830–839.
Chou, J.H., Sheu, S.Y., Lin, K.Y., Chen, W.M., Arun, A.B., Young, C.C., 2007. *Comamonaso dontotermitis* sp. nov., isolated from the gut of the termite *Odontotermes formosanus*. Int. J. Syst. Evol. Microbiol. 57, 887–891.
Chun, J., Bae, K.S., 2000. Phylogenetic analysis of *Bacillus subtilis* and related taxa based on partial *gyrA* gene sequences. Antonie Van Leeuwenhoek 78, 123–127.
Cihan, A.C., Tekin, N., Ozcan, B., Cokmus, C., 2012. The genetic diversity of genus *Bacillus* and the related genera revealed by 16S rRNA gene sequences and ardra analyses isolated from geothermal regions of turkey. Braz. J. Microbiol. http://dx.doi.org/10.1590/S1517-83822012000100037.
Clarridge III, J.E., 2004. Impact of 16S rRNA gene sequence analysis for identification of bacteria on clinical microbiology and infectious diseases. Clin. Microbiol. Rev. 17, 840–862.
Claus, D., Berkeley, C.W., 1986. The genus *Bacillus*. Bergey's Manual of Systematic Bacteriology. Sneath pHA, vol. 2. Williams, Wilkins, Baltimore, pp. 1105–1139. ISBN: 0-683-07893-934.
Clayton, R.A., Sutton, G., Hinkle Jr., P.S., Bult, C., Fields, C., 1995. Intraspecific variation in small-subunit rRNA sequences in GenBank: why single sequences may not adequately represent prokaryotic taxa. Int. J. Syst. Bacteriol. 45, 595–599.
Coenye, T., Vandamme, P., 2003. Extracting phylogenetic information from whole-genome sequencing projects: the lactic acid bacteria as a test case. Microbiology 149, 3507–3517.
Cohan, F.M., 2002. What are bacterial species? Annu. Rev. Microbiol. 56, 457–487.
Craven, K.D., Hsiau, P.T.W., Leuchtmann, A., Hollin, W., Schardl, C.L., 2001. Multigene phylogeny of *Epichloe* species, fungal symbionts of grasses. Ann. Mo. Bot. Garden 88, 14–34.
Crump, B.C., Kling, G.W., Bahar, M., Hobbie, J.E., 2003. Bacterioplankton community shifts in an arctic lake correlate with seasonal changes in organic matter source. Appl. Environ. Microbiol. 69, 2253–2268.
Cubero, J., Graham, J.H., 2002. Genetic relationship among worldwide strains of Xanthomonas causing canker in citrus species and design of new primers for their identification by PCR. Appl. Environ. Microbiol. 68, 1257–1264.
Das, S., Lyla, P.S., Khan, S.A., 2007. Fatty acid profiles of marine benthic microorganisms isolated from the continental slope of Bay of Bengal: A possible implications in the benthic food web. Ocean Sci. J. 42, 247–254.
Dawyndt, P., Vancanneyt, M., Snauwaert, C., DeBaets, B., DeMeyer, H., Swings, J., 2006. Mining fatty acid databases for detection of novel compounds in aerobic bacteria. J. Microbiol. Methods 66, 410–433.
De Bruyne, K., Slabbinck, B., Waegeman, W., Vauterin, P., De Baets, B., Vandamme, P., 2011. Bacterial species identification from MALDI-TOF spectra through data analysis and machine learning. Syst. Appl. Microbiol. 34, 20–29.
de Vos, M., Muller, B., Borrell, S., Black, P.A., van Helden, P.D., Warren, R.M., Gagneux, S., Victor, T.C., 2013. Putative compensatory mutations in the *rpoC* gene of rifampin-resistant *Mycobacterium tuberculosis* are associated with ongoing transmission. Antimicrob. Agents Chemother. 57, 827–832.
Dini-Andreote, F., Andreote, F.D., Araujo, W.L., Trevors, J.T., Van-Elsas, J.D., 2012. Bacterial genomes: habitat specificity and uncharted organisms. Microb. Ecol. 64, 1–7.
Dowd, S.E., Sun, Y., Wolcott, R.D., Domingo, A., Carroll, J.A., 2008. Bacterial tag-encoded FLX amplicon pyrosequencing (bTEFAP) for microbiome studies: bacterial diversity in the ileum of newly weaned *Salmonella*-infected pigs. Foodborne Pathog. Dis. 5, 459–472.
Drancourt, M., Bollet, C., Carlioz, A., Martelin, R., Gayral, J.P., Raoult, D., 2000. 16S ribosomal DNA sequence analysis of a large collection of environmental and clinical unidentifiable bacterial isolates. J. Clin. Microbiol. 38, 3623–3630.
Duhamel, S., Jacquet, S., 2006. Flow cytometric analysis of bacteria- and virus-like particles in lake sediments. J. Microbiol. Methods 64, 316–322.
Duim, B., Wagenaar, J.A., Dijkstra, J.R., Goris, J., Endtz, H.P., Vandamme, P.A.R., 2004. Identification of distinct *Campylobacter lari* genogroups by amplified fragment length polymorphism and protein electrophoretic profiles. Appl. Environ. Microbiol. 70, 18–24.
Eden, P.A., Schmidt, T.M., Blakemore, R.P., Pace, N.R., 1991. Phylogenetic analysis of *Aquaspirillum magnetotacticum* using polymerase chain reaction-amplified 16S rRNA-specific DNA. Int. J. Syst. Evol. Bacteriol. 41, 324–325.
Ehrhardt, C.J., Chu, V., Brown, T.C., Simmons, T.L., Swan, B.K., Bannan, J., Robertson, J.M., 2010. Use of fatty acid methyl ester profiles for discrimination of *Bacillus cereus* T-strain spores grown on different media. Appl. Environ. Microbiol. 76, 1902–1912.
Eisen, J.A., 1995. The RecA protein as a model molecule for molecular systematic studies of bacteria: comparison of trees of RecAs and 16S rRNAs from the same species. J. Mol. Evol. 41, 1105–1123.

Espinosa, L., Borowsky, R., 1998. Evolutionary divergence of AP-PCR (RAPD) patterns. Mol. Biol. Evol. 15, 408–414.
Etchebehere, C., Errazquin, M.I., Dabert, P., Moletta, R., Muxi, L., 2001. Comamonas nitrativorans sp. nov., a novel denitrifier isolated from a denitrifying reactor treating landfill leachate. Int. J. Syst. Evol. Microbiol. 51, 977–983.
Etoh, Y., Yamamoto, A., Goto, N., 1998. Intervening sequences in 16S rRNA genes of *Campylobacter* sp.: diversity of nucleotide sequences and uniformity of location. Microbiol. Immunol. 42, 241–243.
Everest, G.J., Cook, A.E., Kirby, B.M., Meyers, P.R., 2011. Evaluation of the use of *recN* sequence analysis in the phylogeny of the genus *Amycolatopsis*. Antonie Van Leeuwenhoek 100, 483–496.
Fagerquist, C.K., Bates, A.H., Heath, S., King, B.C., Garbus, B.R., Harden, L.A., Miller, W.G., 2006. Sub-speciating *Campylobacter jejuni* by proteomic analysis of its protein biomarkers and their post-translational modifications. J. Proteome Res. 5, 2527–2538.
Fatima, S., Anjum, T., Bajwa, R., Shahid, A.A., 2012. Identification of indigenous *Xanthomonas* isolates through 16S rRNA gene amplification and SDS PAGE analyses. Afr. J. Microbiol. Res. 6, 3651–3655.
Ferguson, N.M., Hepp, D., Sun, S., Ikuta, N., Levisohn, S., Kleven, S.H., Garcia, M., 2005. Use of molecular diversity of *Mycoplasma gallisepticum* by gene-targeted sequencing (GTS) and random amplified polymorphic DNA (RAPD) analysis for epidemiological studies. Microbiology 151, 1883–1893.
Finch, P.W., Chambers, P., Emmerson, P.T., 1985. Identification of the *Escherichia coli recN* gene product as a major SOS protein. J. Bacteriol. 164, 653–658.
Forterre, P., Philippe, H., 1999. Where is the root of the universal tree of life? Bioessays 21, 871–879.
Frank, D.N., Spiegelman, G.B., Davis, W., Wagner, E., Lyons, E., Pace, N.R., 2003. Culture-independent molecular analysis of microbial constituents of the healthy human outer ear. J. Clin. Microbiol. 41, 295–303.
Frank, J.A., Reich, C.I., Sharma, S., Weisbaum, J.S., Wilson, B.A., Olsen, G.J., 2008. Critical evaluation of two primers commonly used for amplification of bacterial 16S rRNA genes. Appl. Environ. Microbiol. 74, 2461–2470.
Franken, A.A.J.M., Zilverentant, J.F., Boonekamp, P.M., Schots, A., 1992. Specificity of polycllonal and monoclonal antibodies for the identification of *Xanthomonas campestris* pv. *campestris*. Netherlands J. Neth. J. Plant Pathol. 98, 81–94.
Freitas, D.B., Reis, M.P., Lima-Bittencourt, C.I., Costa, P.S., Assis, P.S., Chartone-Souza, E., Nascimento, A.M.A., 2008. Genotypic and phenotypic diversity of *Bacillus* spp. isolated from steel plant waste. BMC Res. Notes. http://dx.doi.org/10.1186/1756-0500-1-92.
Frye, S.R., Healy, M., 2006. Molecular strain typing using repetitive sequence-based PCR, In: Tang, Y.-W., Stratton, C.W. (Eds.), Advanced Techniques in Diagnostic Microbiology, 1st ed. Springer, New York, NY, pp. 444–471.
Fujita, M., Hanaura, Y., Amemura, A., 1995. Analysis of the *rpoD* gene encoding the principal sigma factor of *Pseudomonas putida*. Gene 167, 93–98.
Gafan, G.D., Spratt, D.A., 2005. Denaturing gradient gel electrophoresis gel expansion (DGGEGE)- An attempt to resolve the limination of co-migration in the DGGE of complex polymicrobial communities. FEMS Microbiol. Lett. 253, 303–307.
Galagan, J.E., Henn, M.R., Ma, L.J., Cuomo, C.A., Birren, B., 2005. Genomics of the fungal kingdom: insights into eukaryotic biology. Gen. Res. 15, 1620–1631.
Gee, J.E., De, B.K., Levett, P.N., Whitney, A.M., Novak, R.T., Popovic, T., 2004. Use of 16S rRNA gene sequencing for rapid confirmatory identification of *Brucella* isolates. J. Clin. Microbiol. 42, 3649–3654.
Georgopoulos, C.P., Lundquist-Heil, A., Yochem, J., Feiss, M., 1980. Identification of the *E. coli* dnaJ gene product. Mol. Gen. Genet. 178, 583–588.
Gevers, D., Cohan, F.M., Lawrence, J.G., Spratt, B.G., Coenye, T., Feil, E.J., Stackebrandt, E., Van de Peer, Y., Vandamme, P., Thompson, F.L., Swings, J., 2005. Opinion: re-evaluating prokaryotic species. Nat. Rev. Microbiol. 3, 733–739.
Glazunova, O.O., Raoult, D., Roux, V., 2009. Partial sequence comparison of the *rpoB*, *sodA*, *groEL* and *gyrB* genes within the genus *Streptococcus*. Int. J. Syst. Evol. Microbiol. 59, 2317–2322.
Glazunova, O.O., Raoult, D., Roux, V., 2010. Partial *recN* gene sequencing: a new tool for identification and phylogeny within the genus *Streptococcus*. Int. J. Syst. Evol. Microbiol. 60, 2140–2148.
Gomila, M., Pinhassi, J., Falsen, E., Moore, E.R.B., Lalucat, J., 2010. *Kinneretia asaccharophila* gen. nov., sp. nov., isolated from a freshwater lake, a member of the Rubrivivax branch of the family Comamonadaceae. Int. J. Syst. Evol. Microbiol. 60, 809–814.
Grabovich, M., Gavrish, E., Kuever, J., Lysenko, A.M., Podkopaeva, D., Dubinina, G., 2006. Proposal of *Giesbergeria voronezhensis* gen. nov., sp. nov. and *G. kuznetsovii* sp. nov. and reclassification of [*Aquaspirillum*] *anulus*, [A.] *sinuosum* and [A.] *giesbergeri* as *Giesbergeria anulus* comb. nov., *G. sinuosa comb*.nov. and*G. giesbergeri comb*. nov., and [*Aquaspirillum*] *metamorphum* and [A.] *psychrophilum* as *Simplicispira metamorpha* gen. nov., comb. nov. and *S. psychrophila* comb. nov. Int. J. Syst. Evol. Microbiol. 56, 569–576.
Gunasekera, T.S., Attfield, P.V., Veal, D.A., 2000. A flow cytometry method for rapid detection and enumeration of total bacteria in milk. Appl. Environ. Microbiol. 66, 1228–1232.
Gundi, V.A.K.B., Dijkshoorn, L., Burignat, S., Raoult, D., Scola, B.L., 2009. Validation of partial *rpoB* gene sequence analysis for the identification of clinically important and emerging *Acinetobacter* species. Microbiology 155, 2333–2341.
Guo, W., Hao, H., Dai, M., Wang, Y., Huang, L., Peng, D., Wang, X., Wang, H., Yao, M., Sun, Y., Liu, Z., Yuan, Z., 2012. Development of quinoxaline 1,4-dioxides resistance in *Escherichia coli* and molecular change under resistance selection. PLoS One 7, e43322. http://dx.doi.org/10.1371/journal.pone.0043322.
Gurtler, V., Mayall, B.C., Wang, J., 2012. HRM confirmation of Neisseria gonorrhoeae in clinical specimens by G → A (position 857) mutation detection in the 16S rRNA gene before sequencing and after porA confirmation. J. Microbiol. Methods 89, 148–152.
Guy, E., Genissel, A., Hajri, A., Chabannes, M., David, P., Carrere, S., Lautier, M., Roux, B., Boureau, T., Arlat, M., Poussier, S., Noela, L.D., 2013. Natural genetic variation of *Xanthomonas campestris* pv. *campestris* pathogenicity on arabidopsis revealed by association and reverse genetics. MBio 4. http://dx.doi.org/10.1128/mBio.00538-12.
Hagstrom, A., Pommier, T., Rohwer, F., Simu, K., Stolte, W., Svensson, D., Zweifel, U.L., 2002. Use of 16S ribosomal DNA for delineation of marine bacterioplankton species. Appl. Environ. Microbiol. 68, 3628–3633.
Hall, L., Wohlfiel, S., Roberts, G.D., 2003. Experience with the MicroSEQ D2 large subunit ribosomal DNA sequencing kit for identification of filamentous fungi encountered in the clinical laboratory. J. Clin. Microbiol. 42, 622–626.
Han, X., Gross, R.W., 2003. Global analyses of cellular lipodomes directly from crude extracts of biological samples by electrospray ionization mass spectrometry: a bridge to lipidomics. J. Lipid Res. 44, 1071–1079.
Hansson, I., Persson, M., Svensson, L., Engvall, E.O., Johansson, K.E., 2008. Identification of nine sequence types of the 16S rRNA genes of *Campylobacte rjejuni* subsp. *jejuni* isolated from broilers. Acta Vet. Scand. 50, 10. http://dx.doi.org/10.1186/1751-0147-50-10.
Harrington, C., Del Casale, A., Kennedy, J., Neve, H., Picton, B.E., Mooij, M., O'Gara, F., Kulakov, L.A., Larkin, M.J., Dobson, A.D., 2012. Evidence of bacteriophage-mediated horizontal transfer of bacterial 16 s rRNA genes in the viral metagenome of the marine sponge *Hymeniacidon perlevis*. Microbiology 158, 2789–2795.
Hauben, L., Vauterin, L., Swings, J., Moore, E.R.B., 1997. Comparison of 16S ribosomal DNA sequences of all *Xanthomonas* species. Int. J. Syst. Bacteriol. 47, 328–335.
Healy, M., Reece, K., Walton, D., Huong, J., Frye, S., Raad, I.I., Kontoyiannis, D.P., 2005. Use of the diversilab system for species and strain differentiation of *Fusarium* species isolates. J. Clin. Microbiol. 43, 5278–5280.
Helgason, E., Tourasse, N.J., Meisal, R., Caugant, D.A., Kolstø, A.B., 2004. Multilocus sequence typing scheme for bacteria of the *Bacillus cereus* group. Appl. Environ. Microbiol. 70, 191–201.
Heylen, K., Lebbe, L., Vos, P.D., 2008. *Acidovoraxcaeni* sp. nov., a denitrifying species with genetically diverse isolates from activated sludge. Int. J. Syst. Evol. Microbiol. 58, 73–77.
Hibbett, D.S., Bindera, M., Bischoff, J.F., Blackwell, M., Cannon, P.F., Eriksson, O.E., Huhndorf, S., James, T., Kirk, P.M., Lücking, R., Lumbsch, H.T., Lutzoni, F., Matheny, P.B., McLaughlin, D.J., Powell, M.J., Redhead, S., Schoch, C.L., Spatafora, J.W., Stalpers, J.A., Vilgalys, R., Aime, M.C., Aptroot, A., Bauer, R., Begerow, D., Benny, G.L., Castlebury, L.A., Crous, P.W., Dai, Y.C., Gams, W., Geiser, D.M., Griffith, G.W., Gueidan, C., Hawksworth, D.L., Hestmark, G., Hosaka, K., Humber, R.A., Hyde, K.D., Ironside, J.E., Kõljalg, U., Kurtzmana, C.P., Larssona, K., Lichtwardta, R., Longcorea, J., Miądlikowska, J., Millera, A., Moncalvoa, J., Mozley-Standridgeag, S., Oberwinkler, F., Parmastoa, E., Reeb, V., 2007. A higher-level phylogenetic classificaiton of the Fungi. Mycol. Res. 111, 507–547.
Hildebrand, F., Meyer, A., Walker, A.E., 2010. Evidence of selection upon genomic GC-content in bacteria. PLOS Gen. http://dx.doi.org/10.1371/journal.pgen.1001107.
Hirsch, P.R., Mauchline, T.H., Clark, I.M., 2010. Culture-independent molecular techniques for soil microbial ecology. Soil Biol. Biochem. 42, 878–887.
Iizumi, T., Nakamura, K., 1997. Cloning, nucleotide sequence, and regulatory analysis of the *Nitrosomonas europaea dnaK* gene. Appl. Environ. Microbiol. 63, 1777–1784.
Inglis, T.J.J., Aravena-Roman, M., Ching, S., Croft, K., Wuthiekamm, V., Mee, B.J., 2003. Cellular fatty acid profile distinguishes *Burkholderia pseudomallei* from a virulent *Burkholderia thailandensis*. J. Clin. Microbiol. 41, 4812–4814.
Jabbour, R.E., Deshpande, S.V., Stanford, M.F., Wick, C.H., Zulich, A.W., Snyder, A.P., 2011. A protein processing filter method for bacterial identification by mass spectrometry-based proteomics. J. Proteome Res. 10, 907–912.
Janda, J.M., Abbott, S.L., 2007. 16S rRNA gene sequencing for bacterial identification in the diagnostic laboratory: pluses, perils, and pitfalls. J. Clin. Microbiol. 45, 2761–2764.
Janssen, P., Audit, B., Cases, I., Darzentas, N., Goldovsky, L., Kunin, V., Lopez-Bigas, N., Peregrin-Alvarez, J.M., Pereira-Leal, J.B., Tsoka, S., Ouzounis, C.A., 2003. Beyond 100 genomes. Genome Biol. 4, 402.
Jany, J.L., Barbier, G., 2008. Culture-independent methods for identifying microbial communities in cheese. Food Microbiol. 25, 839–848.
Jehmlich, N., Schmidt, F., Taubert, M., Seifert, J., von Bergen, M., Richnow, H.H., Vogt, C., 2009. Comparison of methods for simultaneous identification of bacterial species and determination of metabolic activity by protein-based stable isotope probing (Protein-SIP) experiments. Rapid Commun. Mass Spectrom. 23, 1871–1878.
Johnson, J.L., 1991. DNA reassociation experiments. In Nucleic Acid Techniques in Bacterial Systematics. In: Stackebrandt, E., Goodfellow, M. (Eds.), Wiley, Chichester, pp. 21–44.
Jolley, K.A., Bliss, C.M., Bennett, J.S., Bratcher, H.B., Brehony, C., Colles, F.M., Wimalarathna, H., Harrison, O.B., Sheppard, S.K., Cody, A.J., Maiden, M.C.J., 2012. Ribosomal multilocus sequence typing: universal characterization of bacteria from domain to strain. Microbiology 158, 1005–1015.
Kaeding, A.J., Ast, J.C., Pearce, M.M., Urbanczyk, H., Kimura, S., Endo, H., Nakamura, M., Dunlap, P.V., 2007. Phylogenetic diversity and cosymbiosis in the bioluminescent symbioses of *Photobacterium mandapamensis*. Appl. Environ. Microbiol. 73, 3173–3182.
Kampfer, P., Glaeser, S.P., 2012. Prokaryotic taxonomy in the sequencing era — the polyphasic approach revisited. Environ. Microbiol. 14, 291–317.
Karlin, S., Weinstock, G.M., Brendel, V., 1995. Bacterial classifications derived from *recA* protein sequence comparisons. J. Bacteriol. 177, 6881–6893.
Kawasaki, S., Fratamico, P.M., Wesley, I.V., Kawamoto, S., 2008. Species-specific identification of Campylobacters by PCR-restriction fragment length polymorphism and PCR targeting of the GyraseB gene. Appl. Environ. Microbiol. 74, 2529–2533.
Khan, S.T., Horiba, Y., Yamamoto, M., Hiraishi, A., 2002. Members of the family Comamonadaceae as primary poly(3-hydroxybutyrate-co-3-hydroxyvalerate)-degrading denitrifiers in activated sludge as revealed by a polyphasic approach. Appl. Environ. Microbiol. 68, 3206–3214.
Khot, P.D., Ko, D.L., Fredricks, D.N., 2009. Sequencing and analysis of fungal rRNA operons for development of broad-range fungal PCR assays. Appl. Environ. Microbiol. 75, 1559–1565.

Ki, J.S., Zhang, W., Qian, P.Y., 2009. Discovery of marine *Bacillus* species by 16S rRNA and rpoB comparisons and their usefulness for species identification. J. Microbiol. Methods 77, 48–57.
Klenk, H.P., Zillig, W., 1994. DNA dependent RNA polymerase subunit B as a tool for phylogenetic reconstructions: branching topology of the archaeal domain. J. Mol. Evol. 38, 420–432.
Konstantinidis, K.T., Tiedje, J.M., 2007. Prokaryotic taxonomy and phylogeny in the genomic era: advancements and challenges ahead. Curr. Opin. Microbiol. 10, 504–509.
Korczak, B.M., Stieber, R., Emler, S., Burnens, A.P., Frey, J., Kuhnert, P., 2006. Genetic relatedness within the genus *Campylobacter* inferred from *rpoB* sequences. Int. J. Syst. Evol. Microbiol. 56, 937–945.
Koser, C.U., Ellington, M.J., Cartwright, E.J.P., Gillespie, S.H., Brown, N.M., Farrington, M., Holden, M.T.G., Dougan, G., Bentley, S.D., Parkhill, J., Peacock, S.J., 2012. Routine use of microbial whole genome sequencing in diagnostic and public health microbiology. PLoS Pathog. 8.
Kuzoff, R.K., Sweere, J.A., Soltis, D.E., Soltis, P.S., Zimmer, E.A., 1998. The phylogenetic potential of entire 26S rDNA sequences in plants. Mol. Biol. Evol. 15, 251–263.
Lange, M., Guillou, L., Vaulot, D., Simon, N., Amann, R.I., Ludwig, W., Medlin, L.K., 1996. Identification of the class Prymnesiophyceae and the genus *Phaeocystis* with ribosomal RNA targeted nucleic acid probes detected by flow cytometry. J. Phycol. 32, 858–868.
Lapidus, A., Goltsman, E., Auger, S., Galleron, N., Segurens, B., Dossat, C., Land, M.L., Broussolle, V., Brillard, J., Guinebretiere, M.H., Sanchis, V., Nguen-The, C., Lereclus, D., Richardson, P., Wincker, P., Weissenbach, J., Ehrlich, S.D., Sorokin, A., 2008. Extending the *Bacillus cereus* group genomics to putative food-borne pathogens of different toxicity. Chem. Biol. Interact. 171, 236–249.
Lagace, L., Pitre, M., Jacques, M., Roy, D., 2004. Identification of the bacterial community of maple sap by using Amplified ribosomal DNA (rDNA) restriction analysis and rDNA sequencing. Appl. Environ. Microbiol. 70, 2052–2060.
Larsen, M.V., Cosentino, S., Rasmussen, S., Friis, C., Hasman, H., Marvig, R.L., Jelsbak, L., Sicheritz-Ponten, T., Ussery, D.W., Aarestrup, F.M., Lunda, O., 2012. Multilocus sequence typing of total-genome-sequenced bacteria. J. Clin. Microbiol. 50, 1355–1361.
Lee, S.E., Kim, S.Y., Kim, C.M., Kim, M.K., Kim, Y.R., Jeong, K., Ryu, H.J., Lee, Y.S., Chung, S.S., Choy, H.E., Rhee, J.H., 2007. The *pyrH* gene of *Vibrio vulnificus* is an essential in vivo survival factor. Infect. Immun. 75, 2795–2801.
Leite Jr., R.P., Minsavage, G.V., Bonas, U., Stall, R.E., 1994. Detection and identification of phytopathogenic *Xanthomonas* strains by amplification of DNA sequences related to the *hrp* genes of *Xanthomonas campestris* pv. *vesicatory*. Appl. Environ. Microbiol. 60, 1068–1077.
Leyns, F., De Cleene, M., Swings, J., De Ley, J., 1984. The host range of the genus *Xanthomonas*. Bot. Rev. 50, 308–356.
Linton, D., Lawson, A.J., Owen, R.J., Stanley, J., 1997. PCR detection, identification to species level, and fingerprinting of *Campylobacter jejuni* and *Campylobacter coli* direct from diarrheic samples. J. Clin. Microbiol. 35, 2568–2572.
Liu, Z., DeSantis, T.Z., Andersen, G.L., Knight, R., 2008. Accurate taxonomy assignments from 16S rRNA sequences produced by highly parallel pyrosequences. Nucleic Acids Res. 36, e120.
Lloret, L., Barreto, R., Leon, R., Moreno, S., MartInez-Salazar, J., Espin, G., Soberon-Chavez, G., 2006. Genetic analysis of the transcriptional arrangement of *Azotobactervinelandii* alginate biosynthetic genes: identification of two independent promoters. Mol. Microbiol. 21, 449–457.
Logan, J.M.J., Burnens, A., Linton, D., Lawson, A.J., Stanley, J., 2000. *Campylobacterlanienae* sp. nov., a new species isolated from workers in an abattoir. Int. J. Syst. Evol. Microbiol. 50, 865–872.
Logan, N.A., Berge, O., Bishop, A.H., Busse, H.J., de Vos, P., Fritze, D., Heyndrickx, M.R.A., Kampfer, P., Rabinovitch, L., Salkinoja-Salonen, M.S., Seldin, L., Ventosa, A., 2009. Proposed minimal standards for describing new taxa of aerobic, endospore-forming bacteria. Int. J. Syst. Evol. Microbiol. 59, 2114–2121.
Lopez, A.C., Alippi, A.M., 2007. Phenotypic and genotypic diversity of *Bacillus cereus* isolates recovered from honey. Int. J. Food Microbiol. 117, 175–184.
Ludwig, W., Schleifer, K.H., Whitman, W.B., 2007. Revised Road Map to the Phylum Firmicutes. Bergey's Manual Trust Website, (http://www.bergeys.org/outlines/Bergeys,Vol 3,Outline.pdf).
Maiden, M.C., 2006. Multilocus sequence typing of bacteria. Annu. Rev. Microbiol. 60, 561–588.
Maiden, M.C., Bygraves, J.A., Feil, E., Morelli, G., Russell, J.E., Urwin, R., Zhang, Q., Zhou, J., 1998. Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms. Proc. Natl. Acad. Sci. U. S. A. 95, 3140–3145.
Man, S.M., Kaakoush, N.O., Octavia, S., Mitchell, H., 2010. The internal transcribed spacer region, a new tool for use in species differentiation and delineation of systematic relationships within the *Campylobacter* genus. Appl. Environ. Microbiol. 76, 3071–3081.
Manukhov, I.V., Khrul'nova, S.A., Baranova, A., Zavilgelsky, G.B., 2011. Comparative analysis of the lux Operons in *Aliivibrio logei* KCh1 (a Kamchatka Isolate) and *Aliivibrio salmonicida*. J. Bacteriol. 193, 3998–4001.
Margolin, W., Bramhill, D., Long, S.R., 1995. The *dnaA* gene of *Rhizobium meliloti* lies within an unusual gene arrangement. J. Bacteriol. 177, 2892–2900.
Martens, M., Dawyndt, P., Coopman, R., Gillis, M., De Vos, P., Willems, A., 2008. Advantages of multilocus sequence analysis for taxonomic studies: a case study using 10 housekeeping genes in the genus *Ensifer* (including former Sinorhizobium). Int. J. Syst. Evol. Microbiol. 58, 200–214.
Mechichi, T., Stackebrandt, E., Fuchs, G., 2003. Alicycliphilus denitrificans gen. nov., sp. nov., a cyclohexanol-degrading, nitrate-reducing β-proteobacterium. Int. J. Syst. Evol. Microbiol. 53, 147–152.
Mignot, T., Couture-Tosi, E., Mesnage, S., Mock, M., Fouet, A., 2004. In vivo *Bacillus anthracis* gene expression requires PagR as an intermediate effector of the AtxA signalling cascade. Int. J. Med. Microbiol. 293, 619–624.
Miller, R.V., Kokjohn, T.A., 1990. General microbiology of *recA*: environmental and evolutionary significance. Annu. Rev. Microbiol. 44, 365–394.
Minnikin, D.E., Alshamaony, L., Goodfellow, M., 1975. Differentiation of *Mycobacterium*, *Nocardia* and related taxa by thin-layer chromatographic analysis of whole-organism methanolysates. J. Gen. Microbiol. 88, 200–204.
Mohanty, S., Choudhury, P.K., Dash, A., Samanta, M., Maiti, N.K., 2011. Genotypic and phenotypic diversity of *Bacillus* spp. isolated from freshwater ecosystems. J. Aquac. Res. Dev. 2. http://dx.doi.org/10.4172/2155-9546.1000112.
Mohapatra, B.R., Broersma, K., Nordin, R., Mazumder, A., 2007. Evaluation of repetitive extragenic palindromic-PCR for discrimination of fecal *Escherichia coli* from humans, and different domestic- and wild-animals. Microbiol. Immunol. 51, 733–740.
Mollet, C., Drancourt, M., Raoult, D., 1997. *rpoB* sequence analysis as a novel basis for bacterial identification. Mol. Microbiol. 26, 1005–1011.
Moolhuijzen, P.M., Lew-Tabor, A.E., Wlodek, B.M., Aguero, F.G., Comerci, D.J., Ugalde, R.A., Sanchez, D.O., Appels, R., Bellgard, M., 2009. Genomic analysis of *Campylobacter fetus* subspecies: identification of candidate virulence determinants and diagnostic assay targets. BMC Microbiol. 9, 86. http://dx.doi.org/10.1186/1471-2180-9-86.
Moss, C.W., Lambert-Fair, M.A., Nicholson, M.A., Guerrant, G.O., 1990. Isoprenoid quinones of *Campylobacter cryaerophila*, *C. cinaedi*, *C. fennelliae*, *C. hyointestinalis*, *C. pylori*, and *C. upsaliensis*. J. Clin. Microbiol. 28, 395–397.
Mukai, T., Miyamoto, Y., Yamazaki, T., Makino, M., 2006. Identification of *Mycobacterium* species by comparative analysis of the *dnaA* gene. FEMS Microbiol. Lett. 254, 232–239.
Mukherjee, S., Kumar, D., Nanda, A.K., Chakraborty, R., 2013. 16S rRNA gene sequence analyses of the metagenome derived from waters of river Mahananda at Siliguri: an approach to understand bacterial diversity. Indian J. Biotechnol. 12, 80–87.
Muyzer, G., De Waal, E.C., Uitterlinden, A.G., 1993. Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. Appl. Environ. Microbiol. 59, 695–700.
Naser, S.M., Thompson, F.L., Hoste, B., Gevers, D., Dawyndt, P., Vancanneyt, M., Swings, J., 2005. Application of multilocus sequence analysis (MLSA) for rapid identification of *Enterococcus* species based on *rpoA* and *pheS* genes. Microbiology 151, 2141–2150.
Naser, S.M., Dawyndt, P., Hoste, B., Gevers, D., Vandemeulebroecke, K., Cleenwerck, I., Vancanneyt, M., Swings, J., 2007. Identification of lactobacilli by *pheS* and *rpoA* gene sequence analyses. Int. J. Syst. Evol. Microbiol. 57, 2777–2789.
Nelson, K., Whittam, T.S., Selander, R.K., 1991. Nucleotide polymorphism and evolution in the glyceraldehyde-3-phosphate dehydrogenase gene (*gapA*) in natural populations of *Salmonella* and *Escherichia coli*. Proc. Natl. Acad. Sci. U. S. A. 88, 6667–6671.
Nicol, G.W., Glover, L.A., Prosser, J.I., 2003. The impact grassland management on archaeal community structure in upland pasture rhizosphere soil. Environ. Microbiol. 5, 152–162.
Nielsen, P., Rainey, F.A., Outtrup, H., Priest, F.G., Fritze, D., 1994. Comparative 16S rRNA sequence analysis of some alkaliphilic bacilli and the establishment of a sixth rDNA group within the genus *Bacillus*. FEMS Microbiol. Lett. 117, 61–66.
Ochman, H., Lawrence, J.G., Groisman, E.A., 2000. Lateral gene transfer and the nature of bacterial innovation. Nature 405, 299–304.
Okamoto, T., Maruyama, A., Imura, S., Takeyama, H., Naganuma, T., 2004. Comparative phylogenetic analyses of *Halomonas variabilis* and related organisms based on 16S rRNA, *gyrB* and *ectBC* gene sequences. Syst. Appl. Microbiol. 27, 323–333.
Olive, D.M., Bean, P., 1999. Principles and applications of methods for DNA-based typing of microbial organisms. J. Clin. Microbiol. 37, 1661–1669.
Onodera, Y., Sato, K., 1999. Molecular cloning of the *gyrA* and *gyrB* genes of *Bacteroides fragilis* encoding DNA gyrase. Antimicrob. Agents Chemother. 43, 2423–2429.
Palys, T., Nakamura, L.K., Cohan, F.M., 1997. Discovery and classification of ecological diversity in the bacterial world: the role of DNA sequence data. Int. J. Syst. Bacteriol. 47, 1145–1156.
Park, D.S., Shim, J.K., Kim, J.S., Lim, C.K., Shrestha, R., Hahn, J.H., Kim, H.G., 2009. Sensitive and specific detection of *Xanthomonas campestris* pv. vesicatoria by PCR using pathovar-specific primers based on rhs family gene sequences. Microbiol. Res. 164, 36–42.
Parkinson, N., Cowie, C., Heeney, J., Stead, D., 2009. Phylogenetic structure of *Xanthomonas* determined by comparison of *gyrB* sequences. Int. J. Syst. Evol. Microbiol. 59, 264–274.
Payne, G.W., Vandamme, P., Morgan, S.H., LiPuma, J.J., Coenye, T., Weightman, A.J., Jones, T.H., Mahenthiralingam, E., 2005. Development of a *recA* gene based identification approach for the entire *Burkholderia* genus. Appl. Environ. Microbiol. 71, 3917–3927.
Peak, K.K., Duncan, K.E., Luna, V.A., King, D.S., McCarthy, P.J., Cannons, A.C., 2011. *Bacillus* strains most closely related to *Bacillus nealsonii* are not effectively circumscribed within the taxonomic species definition. Int. J. Microbiol. http://dx.doi.org/10.1155/2011/673136.
Petti, C.A., 2007. Detection and identification of microorganisms by gene amplification and sequencing. Clin. Infect. Dis. 44, 1108–1114.
Pillay, M., Kenny, S.T., 1995. Anomalies in direct pairwise comparison of RAPD fragments for genetic analysis. Biotechniques 19, 694–698.
Pineiro-Vidal, M., Pazos, F., Santos, Y., 2008. Fatty acid analysis as a chemotaxonomic tool for taxonomic and epidemiological characterization of four fish pathogenic *Tenacibaculum* species. Lett. Appl. Microbiol. 46, 548–554.
Porter, T.M., Golding, G.B., 2012. Factors that affect large subunit ribosomal DNA amplicon sequencing studies of fungal communities: classification method, primer choice, and error. PLoS ONE 7. http://dx.doi.org/10.1371/journal.pone.0035749.
Prakash, O., Verma, M., Sharma, P., Kumar, M., Kumari, K., Singh, A., Kumari, H., Jit, S., Gupta, S.K., Khanna, M., Lal, R., 2007. Polyphasic approach of bacterial classification — an overview of recent advances. Indian J. Microbiol. 47, 98–108.
Priest, F.G., 1993. Systematics and ecology of *Bacillus*. In: Sonenshein, A.L., Hoch, J.A., Losick, R. (Eds.), *Bacillus subtilis* and other Gram-positive bacteria — biochemistry, physiology, and molecular genetics. ASM press, American Society for Microbiology, Washington, D.C. ISBN: 1-55581-053-5.

Priest, F.G., Goodfellow, M., Todd, C., 1988. A numerical classification of the genus *Bacillus*. J. Gen. Microbiol. 134, 1847–1882.

Ramadass, P., Latha, D., Senthilkumar, A., Srinivasan, P., Saranya, N., 2002. Arbitrarily primed PCR — a rapid and simple method for typing of leptospiral serovars. Indian J. Med. Microbiol. 20, 25–28.

Rasko, D.A., Altherr, M.R., Han, C.S., Ravel, J., 2005. Genomics of the *Bacillus cereus* group of organisms. FEMS Microbiol. Rev. 29, 303–329.

Rastogi, G., Sani, R.K., 2011. Microbes and microbial technology. Agric. Environ. Applal., 2007Richert, K., Brambilla, E., Stackebrandt, E., 2007. The phylogenetic significance of peptidoglycan types: molecular analysis of the genera Microbacterium and Aureobacterium based upon sequence comparison of *gyrB*, *rpoB*, *recA* and *ppk* and 16SrRNA genes. Syst. Appl. Microbiol. 30, 102–108.

Riesenfeld, C.S., Schloss, P.D., Handelsman, J., 2004. Metagenomics: genomic analysis of microbial communities. Ann. Rev. Gen. 38, 525–552.

Rigouts, L., Nolasco, O., de Rijk, P., Nduwamahoro, E., Van Deun, A., Ramsay, A., Arevalo, J., Portaels, F., 2007. J. Clin. Microbiol. 45, 252–254.

Robene-Soustrade, I., Laurent, P., Gagnevin, L., Jouen, E., Pruvost, O., 2006. Specific detection of *Xanthomonas axonopodis* pv. dieffenbachiae in *Anthurium* (*Anthurium andreanum*) tissues by nested PCR. Appl. Environ. Microbiol. 72, 1072–1078.

Roberts, P.D., Hodge, N.C., Bouzar, H., Jones, J.B., Stall, R.E., Berger, R.D., Chase, A.R., 1998. Relatedness of strains of *Xanthomonas fragariae* by restriction fragment length polymorphism, DNA–DNA reassociation, and fatty acid analyses. Appl. Environ. Microbiol. 64, 3961–3965.

Rodriguez-R, L.M., Grajales, A., Arrieta-Ortiz, M.L., Salazar, C., Restrepo, S., Bernal, A., 2012. Genomes-based phylogeny of the genus *Xanthomonas*. BMC Microbiol. 12, 43–57.

Rogall, T., Wolters, J., Flohr, T., Bottger, E.C., 1990. Towards a phylogeny and definition of species at the molecular level within the genus *Mycobacterium*. Int. J. Syst. Bacteriol. 40, 323–330.

Romaniuk, P.J., Trust, T.J., 2006. Identification of *Campylobacter* species by southern hybridization of genomic DNA using an oligonucleotide probe for 16S rRNA genes. FEMS Microbiol. Lett. 43, 331–335.

Rompre, A., Servais, P., Baudart, J., de-Roubin, M.R., Laurent, P., 2002. Detection and enumeration of coli forms in drinking water: current methods and emerging approaches. J. Microbiol. Methods 49, 31–54.

Rowland, G.C., Aboshkiwa, M., Coleman, G., 1993. Comparative sequence analysis and predicted phylogeny of the DNA-dependent RNA polymerase β subunits of *Staphylococcus aureus* and other eubacteria. Biochem. Soc. Trans. 21, 40S.

Rutherford, K., Parkhill, J., Crook, J., Horsnell, T., Rice, P., Rajandream, M.A., Barell, B.G., 2000. Artemis: sequence visualisation and annotation. Bioinformatics 16, 944–945.

Sasser, M., 1990. Identification of Bacteria by Gas Chromatography of Cellular Fatty Acids. MIDI Technical Note No. 101. MIDI, Newark, De, (Revised Feb 1997).

Sauer, S., Freiwald, A., Maier, T., Kube, M., Reinhardt, R., 2008. Classification and identification of bacteria by mass spectrometry and computational analysis. PLoS One 3, e2843. http://dx.doi.org/10.1371/journal.pone.0002843.

Saunders, N.A., Holmes, A., 2007. Multilocus sequence typing (MLST) of *Staphylococcus aureus*. Methods Mol. Biol. 391, 71–85.

Schaad, N.W., Postnikova, E., Lacy, G.H., Sechler, A., Agarkova, I., Stromberg, P.E., Stromberg, V.K., Vidaver, A.K., 2005. Reclassification of *Xanthomonas campestris* pv. citri (ex Hasse 1915) Dye 1978 forms A, B/C/D, and E as *X. smithii* subsp. citri (ex Hasse) sp. nov. nom. rev. comb. nov., *X. fuscans* subsp. *aurantifolii* (ex Gabriel 1989) sp. nov. nom. rev. comb. nov., and *X. alfalfae* subsp. *citrumelo* (ex Riker and Jones) Gabriel et al., 1989 sp. nov. nom. rev. comb. nov.; *X. campestris*pv*malvacearum* (ex smith 1901) Dye 1978 as *X. smithii* subsp. *smithii* nov. comb. nov. nom. nov.; *X. campestris* pv. *alfalfae* (ex Riker and Jones, 1935) dye 1978 as *X. alfalfae* subsp. *alfalfae* (ex Riker et al., 1935) sp. nov. nom. rev.; and "var. fuscans" of *X. campestris*pv. phaseoli (ex Smith, 1987) Dye 1978 as *X. fuscans* subsp. *fuscans* sp. nov. Syst. Appl. Microbiol. 28, 494–518.

Schaad, N.W., Postnikova, E., Lacy, G., Sechler, A., Agarkova, I., Stromberg, P.E., Stromberg, V.K., Vidaver, A.K., 2006. Emended classification of xanthomonad pathogens on citrus. Syst. Appl. Microbiol. 29, 690–695.

Scortichini, M., Rossi, M.P., Marchesi, U., 2002. Genetic, phenotypic and pathogenic diversity of *Xanthomonas arboricola* pv.*corylina* strains question the representative nature of the type strain. Plant Pathol. 51, 374–381.

Seip, E.R., Woloshuk, C.P., Payne, G.A., Curtis, S.E., 1990. Isolation and sequence analysis of a β-tubulin gene from *Aspergillus flavus* and its use as a selectable marker. Appl. Environ. Microbiol. 56, 3686–3692.

Sekiguchi, H., Watanabe, M., Nakahara, T., Xu, B., Uchiyama, H., 2002. Succession of bacterial community structure along the Changjiang River determined by denaturing gradient gel electrophoresis and clone library analysis. Appl. Environ. Microbiol. 68, 5142–5150.

Severinov, K., Mooney, R., Darst, S.A., Landick, R., 1997. Tethering of the large subunits of *Escherichia coli* RNA polymerase. J. Biol. Chem. 272, 24137–24140.

Shah, M.M., Iihara, H., Noda, M., Song, S.X., Nhung, P.H., Ohkusu, K., Kawamura, Y., Ezaki, T., 2007. *dnaJ* gene sequence-based assay for species identification and phylogenetic grouping in the genus *Staphylococcus*. Int. J. Syst. Evol. Microbiol. 57, 25–30.

Shi, X., Rao, N.N., Kornberg, A., 2004. Inorganic polyphosphate in *Bacillus cereus*: motility, biofilm formation, and sporulation. Proc. Natl. Acad. Sci. U. S. A. 101, 17061–17065.

Siegenthaler, R.K., Grimshaw, J.P.A., Christen, P., 2004. Immediate response of the DnaK molecular chaperone system to heat shock. FEBS Lett. 562, 105–110.

Siqueira, J.F., Rocas, I.N., 2005. Exploiting molecular methods to explore endodontic infections: part 1 — current molecular technologies for microbiological diagnosis. J. Endod. 31, 411–423.

Skaar, E.P., Lazio, M.P., Seifert, H.S., 2002. Roles of the *recJ* and *recN* genes in homologous recombination and DNA repair pathways of *Neisseria gonorrhoeae*. J. Bacteriol. 184, 919–927.

Slabbinck, B., Baets, B.D., Dawyndt, P., Vos, P.D., 2009. Towards large-scale FAME-based bacterial species identification using machine learning techniques. Syst. Appl. Microbiol. 32, 163–176.

Song, Y., Yang, R., Guo, Z., Zhang, M., Wang, X., Zhou, F., 2000. Distinctness of spore and vegetative cellular fatty acid profiles of some aerobic endospore-forming bacilli. J. Microbiol. Methods 39, 225–241.

Sorensen, J., 1997. The rhizosphere as a habitat for soil microorganisms. In: van Elsas, J.D., et al. (Eds.), Modern Soil Microbiology. Marcel Dekker, Inc., New York, pp. 21–45.

Soudi, M.R., Alimadadi, N., Ghadam, P., 2011. Minimal phenotypic test for simple differentiation of *Xanthomonas campestris* from other yellow-pigmented bacteria isolated from soil. Iran. J. Microbiol. 3, 84–91.

Spratt, D.A., 2004. Significance of bacterial identification by molecular biology methods. Endod. Top. 9, 5–14.

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

Staley, J.T., 2006. The bacterial species dilemma and the genomic-phylogenetic species concept. Philos. Trans. R. Soc. Lond. Biol. Sci. 361, 1899–1909.

Stall, R.E., Beaulieu, C., Egel, D., Hodge, N.C., Leite, R.P., Minsavage, G.V., Bouzar, H., Jones, J.B., Alvarez, A.M., Benedict, A.A., 1994. Two genetically diverse groups of strains are included in *Xanthomonas campestris* pv. Vesicatoria. Int. J. Syst. Evol. Microbiol. 44, 47–53.

Sugita, T., Nishikawa, A., 2003. Fungal identification method based on DNA sequence analysis: reassessment of the methods of the pharmaceutical society of Japan and the Japanese pharmacopoeia. J. Health Sci. 49, 531–533.

Sun, Y., Wolcott, R.D., Dowd, S.E., 2011. Tag-Encoded FLX amplicon pyrosequencing for the elucidation of microbial and functional gene diversity in any environment. Methods Mol. Biol. 733, 129–141.

Sundarakrishnan, B., Pushpanathan, M., Jayashree, S., Rajendran, J., Sakthivel, N., Jayachandran, S., Gunasekaran, P., 2012. Assessment of microbial richness in pelagic sediment of Andaman sea by bacterial tag encoded FLX titanium amplicon pyrosequencing. Indian J. Microbiol. 52, 544–550.

Svec, P., Vancanneyt, M., Seman, M., Snauwaert, C., Lefebvre, K., Sedlacek, I., Swings, J., 2005. Evaluation of (GTG)5-PCR for identification of *Enterococcus* spp. FEMS Microbiol. Lett. 247, 59–63.

Swofford, D.L., Olsen, G.J., 1990. Phylogeny reconstruction. In: Hillis, D.M., Moritz, C. (Eds.), Molecular Systematics. Sinauer Associates, Sunderland, MA.

Tae-Woon, K., Kim, Y.H., Kim, S.E., Lee, J.H., Park, C.S., Kim, H.Y., 2010. Identification and distribution of *Bacillus* species in Doenjang by whole-cell protein patterns and 16S rRNA gene sequence analysis. J. Microbiol. Biotechnol. 20, 1210–1214.

Takewaki, S., Okuzumi, K., Manabe, I., Tanimura, M., Miyamura, K., Nakahara, K., Yazaki, Y., Ohkubo, A., Nagai, R., 1994. Nucleotide sequence comparison of the mycobacterial dnaJ gene and PCR-restriction fragment length polymorphism analysis for identification of mycobacterial species. Int. J. Syst. Evol. Microbiol. 44, 159–166.

Tan, S., Fraley, C.D., Zhang, M., Dailidiene, D., Kornberg, A., Berg, D.E., 2005. Diverse phenotypes resulting from polyphosphate kinase gene (ppk1) inactivation in different strains of *Helicobacter pylori*. J. Bacteriol. 187, 7687–7695.

Temmerman, R., Scheirlinck, I., Huys, G., Swings, J., 2003. Culture-independent analysis of probiotic products by denaturing gradient gel electrophoresis. Appl. Environ. Microbiol. 69, 220–226.

Thompson, F.L., Gevers, D., Thompson, C.C., Dawyndt, P., Naser, S., Hoste, B., Munn, C.B., Swings, J., 2005. Phylogeny and molecular identification of vibrios on the basis of multilocus sequence analysis. Appl. Environ. Microbiol. 71, 5107–5115.

Van den Mooter, M., Swings, J., 1990. Numerical analysis of 295 phenotypic features of 266 *Xanthomonas* strains and related strains and an improved taxonomy of the genus. Int. J. Syst. Evol. Bacteriol. 40, 348–369.

Van den Velde, S., Lagrou, K., Desmet, K., Wauters, G., Verhaegen, J., 2006. Species identification of corynebacteria by cellular fatty acid analysis. Diagn. Microbiol. Infect. Dis. 54, 99–104.

Vandamme, P., Falsen, E., Rossau, R., Hoste, B., Segers, P., Tytgat, R., Ley, J.D., 1991. Revision of *Campylobacter*, *Helicobacter*, and *Wolinella* taxonomy: emendation of generic descriptions and proposal of *Arcobacter* gen. nov. Int. J. Syst. Bacteriol. 41, 88–103.

Vandamme, P., Pot, B., Gillis, M., de Vos, P., Kersters, K., Swings, J., 1996. Polyphasic taxonomy, a consensus approach to bacterial systematics. Microbiol. Rev. 60, 407–438.

Vandamme, P., Debruyne, L., De Brandt, E., Falsen, E., 2010. Reclassification of bacteroides ureolyticus as *Campylobacter ureolyticus* comb. nov. Int. J. Syst. Evol. Microbiol. 60, 2016–2022.

Vaneechoutte, M., Rossau, R., De Vos, P., Gillis, M., Janssens, D., Paepe, N., De Rouck, A., Fiers, T., Claeys, G., Kersters, K., 1992. Rapid identification of bacteria of the Comamonadaceae with amplified ribosomal DNA-restriction analysis (ARDRA). FEMS Microbiol. Lett. 72, 227–233.

Vauterin, L., Hoste, B., Yang, P., Alvarez, A., Kersters, K., Swings, J., 1993a. Taxonomy of the genus *Xanthomonas*. In: Swings, J., Civerolo, E.L. (Eds.), Xanthomonas. Chapman & Hall, Ltd., London, pp. 157–192.

Vauterin, L., Swings, J., Kersters, K., 1993b. Protein electrophoresis and classification. In: Goodfellow, M., O'Donnell, M.G. (Eds.), Handbook of New Bacterial Systematics. Academic Press, London, pp. 251–280.

Vauterin, L., Hoste, B., Kersters, K., Swings, J., 1995. Reclassification of *Xanthomonas*. Int. J. Syst. Evol. Microbiol. 45, 472–489.

Vauterin, L., Rademaker, J., Swings, J., 2000. Synopsis on the taxonomy of the genus *Xanthomonas*. Phytopathology 90, 677–682.

Veith, B., Herzberg, C., Steckel, S., Feesche, J., Maurer, K.H., Ehrenreich, P., Baumer, S., Henne, A., Liesegang, H., Merkl, R., Ehrenreich, A., Gottschalk, G., 2004. The complete genome sequence of *Bacillus licheniformis* DSM13, an organism with great industrial potential. J. Mol. Microbiol. Biotechnol. 7, 204–211.

Ventura, M., Canchaya, C., Bernini, V., Del Casale, A., Dellaglio, F., Neviani, E., Fitzgerald, G.F., van Sinderen, D., 2005. Genetic characterization of the *Bifidobacteriumbreve* UCC 2003 *hrcA* locus. Appl. Environ. Microbiol. 71, 8998–9007.

Ventura, M., Canchaya, C., Casale, A.D., Dellaglio, F., Neviani, E., Fitzgerald, G.F., van Sinderen, D., 2006. Analysis of bifidobacterial evolution using a multilocus approach. Int. J. Syst. Evol. Microbiol. 56, 2783–2792.

Versalovic, J., Schneider, M., de Bruijn, F.J., Lupski, J.R., 1994. Genomic fingerprinting of bacteria using repetitive sequence based polymerase chain reaction. Methods Mol. Cell. Biol. 5, 25–40.

Vives-Rigo, J., Lebaron, P., Nebe-von Caron, G., 2000. Current and future applications of flow cytometry in aquatic microbiology. FEMS Microbiol. Rev. 24, 429–448.

Walder, M., Sandstedt, K., Ursing, J., 1983. Phenotypic characteristics of thermotolerant *Campylobacter* from human and animal sources. Curr. Microbiol. 9, 291–295.

Wang, L.T., Lee, F.L., Tai, C.J., Kasai, H., 2007. Comparison of gyrB gene sequences, 16S rRNA gene sequences and DNA–DNA hybridization in the Bacillus subtilis group. Int. J. Syst. Evol. Microbiol. 57, 1846–1850.

Wauters, G., Baere, T.D., Willems, A., Falsen, E., Vaneechoutte, M., 2003. Description of *Comamonas aquatica* comb.nov. and *Comamonas kerstersii* sp. nov. for two subgroups of *Comamonas terrigena* and emended description of *Comamonas terrigena*. Int. J. Syst. Evol. Microbiol. 53, 859–862.

Weid, I.V., Duarte, G.F., van Elsas, J.D., Seldin, L., 2002. Paenibacillusbrasilensis sp. nov., a novel nitrogen-fixing species isolated from the maize rhizosphere in Brazil. Int. J. Syst. Evol. Microbiol. 52, 2147–2153.

Weidner, S., Arnold, W., Puhler, A., 1996. Diversity of uncultured microorganisms associated with the seagrass *Halophila stipulacea* estimated by restriction fragment length polymorphism analysis of PCR-amplified 16S rRNA genes. Appl. Environ. Microbiol. 62, 766–771.

Wen, A., Fegan, M., Hayward, C., Chakraborty, S., Sly, L.I., 1999. Phylogenetic relationships among members of the Comamonadaceae, and description of *Delftia acidovorans* (den Dooren de Jong 1926 and Tamaoka et al. 1987) gen. nov., comb. nov. Int. J. Syst. Bacteriol. 49, 567–576.

Whittaker, P., Mossoba, M.M., Al-Khaldi, S., Fry, F.S., Dunkel, V.C., Tall, B.D., Yurawecz, M.P., 2003. Identification of foodborne bacteria by infrared spectroscopy using cellular fatty acid methyl esters. J. Microbiol. Methods 55, 709–716.

Willems, A., Falsen, E., Pot, B., Jantzen, E., Hoste, B., Vandamme, P., Gillis, M., Kersters, K., Ley, J.D., 1990. Acidovorax, a new genus for *Pseudomonas facilis*, *Pseudomonas delafieldii*, E. Falsen (EF) group 13, EF group 16, and several clinical isolates, with the species *Acidovorax facilis* comb. nov., *Acidovorax delafieldii* comb. nov., and *Acidovorax temperans* sp. nov. Int. J. Syst. Bacteriol. 40, 384–398.

Willems, A., Pot, B., Falsen, E., Vandamme, P., Gillis, M., Kersters, K., Ley, J.D., 1991. Polyphasic taxonomic study of the emended genus *Comamonas*: relationships to *Aquaspirillum aquaticum*, E. Falsen group 10, and other clinical isolates. Int. J. Syst. Bacteriol. 41, 427–444.

Wilson, L.A., Sharp, P.M., 2006. Enterobacterial repetitive intergenic consensus (ERIC) sequences in *Escherichia coli*: evolution and implications for ERIC-PCR. Mol. Biol. Evol. 23, 1156–1168.

Winson, M.K., Swift, S., Hill, P.J., Sims, C.M., Griesmayr, G., Bycroft, B.W., Williams, P., Stewart, G.S.A.B., 1998. Engineering the luxCDABE genes from Photorhabdus luminescens to provide a bioluminescent reporter for constitutive and promoter probe plasmids and mini-Tn5 constructs. FEMS Microbiol. Lett. 163, 193–202.

Xu, M., Voorhees, K.J., Hadfield, T.L., 2003. Repeatability and pattern recognition of bacterial fatty acid profiles generated by direct mass spectrometric analysis of in-situ thermal hydrolysis/methylation of whole cells. Talanta 59, 577–589.

Yang, P., Vauterin, L., Vancanneyt, M., Swings, J., Kersters, K., 1993. Application of fatty acid methyl esters for the taxonomic analysis of the genus *Xanthomonas*. Syst. Appl. Microbiol. 16, 47–71.

Yeo, I.C., Lee, N.K., Hahm, Y.T., 2012. Genome sequencing of *Bacillus subtilis* SC-8, antagonistic to the *Bacillus cereus* group, isolated from traditional Korean fermented-soybean food. J. Bacteriol. 194, 536–537.

Yuen, G.Y., Alvarez, A.M., Benedict, A.A., Trotter, K.J., 1986. Use of monoclonal antibodies to monitor the dissemination of *Xanthomonas campestris* pv.*campestris*. Phytopathology 77, 366–370.

Zhang, H., Ishige, K., Kornberg, A., 2002. A polyphosphate kinase (PPK2) widely conserved in bacteria. Proc. Natl. Acad. Sci. U. S. A. 99, 16678–16683.