DU Journal of Undergraduate Research and Innovation
Volume 2, Issue 1 pp 149-165, 2016

In silico Analysis of Genus Proteus for Identification of Redundant Sequences in the Database

Samarth Kulshrestha[1], Arnab Kapuria [2], Dev Dutt Patel[2],
P.S. Dhanaraj[2], Mansi Verma [2,4]*
*mansiverma20@gmail.com
[1] Department of Biological Sciences, [2] Department of Zoology, Sri Venkateswara College,
Benito Juarez Marg, Dhaula Kuan, New Delhi -110021

ABSTRACT

Proteus is a clinically important genus belonging to the family *Enterobacteriaceae* and exhibit pathogenic relationship with the human gastrointestinal tract. Conventional methods like Indole test, swarming pattern and serotyping are some of the misleading tests for the identification of interspecies within this genus due to its high evolving rate. In the present study, 16S rRNA sequences of 423 strains belonging to four species of genus *Proteus* were evaluated. Segregation of species was done using framework analysis and supported by restriction pattern analysis. The 16S rRNA gene analysis proved able for intra-species discrimination from the heterogeneous phylogenetic framework tree indicating the genetic variability among the species. This is further supported by species-specific restriction enzyme digestion pattern. Using these approaches, the resolution of interspecies heterogeneity was found to increase which may help in reducing redundancy in database and provide an effective tool for proper species identification.

Keywords: *Proteus spp., phylogenetic framework, in silico restriction enzyme analysis, uncharacterized species identification*

INTRODUCTION

The bacteria belonging to *Enterobacteriaceae* family are of special microbiological interest because of their pathogenic relationship with the human gastrointestinal tract (1). Currently, there are more than 35 genera described in this family, *Proteus* being one of them (2).The genus was originally described by Hauser in 1885 (3). Since then, it has endured significant taxonomic emendations and is currently divided into distinct groups. These include *Proteus mirabilis, P. vulgaris, P. penneri and P. hauseri* and three unnamed genomo-species, *Proteus* 4, 5 and 6 respectively (3,4).

P. mirabilis are cells vacillating amidst vegetative swimmers and hyper-flagellated swarmers (3, 5). This species can be visualized in water and soil as a free- living microbe (6, 7). Almost 90% of the infections caused by *Proteus* are evoked by the forerunning species. As a fact, 48% of *P. mirabilis* strains are multi-drug resistant and thus a strenuous target for treatment (8, 9). ***Proteus vulgaris*** is a chemo-heterotrophic, rod-shaped bacterium cornering on peritrichous flagella (3). Soil, polluted water, raw meat, gastrointestinal tract of animals and dust are its primary domiciles (10, 11). . In 1982, Hickman et al., cataloged three bio-groups within the *Proteus vulgaris* species and on the groundwork of biochemical parameters, bio-group 1 was elucidated as a new species tagged as ***P. penneri*** (3, 12, 13). The latter species fruitages

negative test for indole production, exploits urea, and yields hydrogen sulphide and gas from glucose (14, 15). ***P. hauseri*** corresponds to *Proteus* genomo-species 3 (12).
Routine biochemical tests have also been utilized for species classification such as fermentation of maltose and mannitol, hydrogen sulphide and indole production, ability to utilize citrate, liquefaction of gelatine and positive ornithine decarboxylase activity (16, 17 and 18). But some of them lead to misidentification of the species due to overlapping characteristics of the discrete species. (19).
Genus *Proteus* embodies the typical flora of the intestinal tract of animals and humans and can be isolated from distinctive environments. Some of the stately species of genus *Proteus* causes primary and secondary infections surrogating as an opportunistic pathogen (20). In auxiliary, these species can also be entangled with bacteremia, neonatal meningoencephalitis, empyema, osteomyelitis and subcutaneous abscesses (2, 20).
In recent times, molecular methods comprising of Polymerase Chain Reaction (PCR), plasmid profiling, Outer-membrane protein profiles, Randomly Amplified Polymorphic DNA-PCR (RAPD-PCR), Restriction Fragment Length Polymorphism (RFLP) have been used for the identification of many species associated with various pathological conditions belonging to the genus *Proteus* (21, 22, 23, 24, 25). However, routine implementation of such processes is expensive. From a practical perspective, a simplified approach for preliminary identification of species is required.
Since 1992, 16S rRNA gene sequencing has proved to be a vital tool for the classification and identification of microorganisms (26, 27).Therefore, molecular tools based on the internal features of 16S rRNA can be developed and used for identification against clinically significant microorganisms (28, 29, 30, 31).
Thus, the current study aims to explore the inherent characteristics of 16S rRNA gene, to devise ways to prevent the redundancy and to aid in preliminary identification of species during diagnostic treatments.

METHODOLOGY

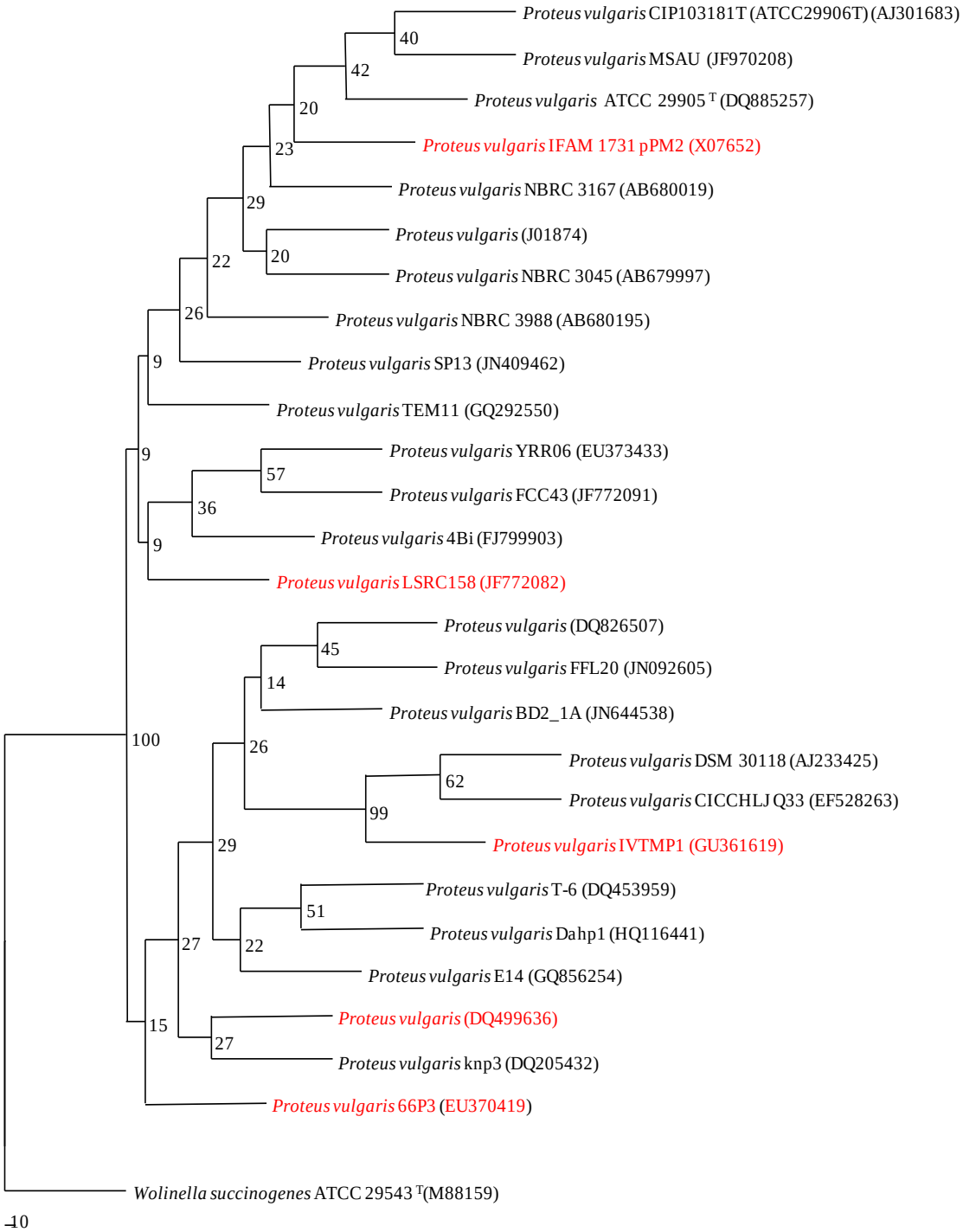
16S rRNA gene sequences related to the genus *Proteus* were analysed through the taxonomy browser at National Centre for Biotechnology Information (NCBI) (www.ncbi.nic.in). Out of 7 species reported in NCBI taxonomy browser including 72 uncharacterized sequences, sequences of 4 significant species with higher persistence value were used as reference species. In the present study, 423 clinical strains (> 1200 nts) were downloaded and investigated from Ribosomal Database Project (RDP) database (http://rdp.cme.msu.edu/) belonging to the genus (32). These include sequences from the isolates of *P. mirabilis* (90 sequences), *P .vulgaris* (56 sequences), *P. hauseri (*13 sequences) and P. *penneri* (9 sequences).
The downloaded sequences for each reference species set with an outgroup sequence were aligned using CLUSTAL X (33). Using DNADIST of the PHYLIP 3.6 package, evolutionary distances between the sequences were calculated with Kimura correction (34). The program NEIGHBOR was used to generate phylogenetic tree and statistical analysis was performed using the programmes SEQBOOT and CONSENSE by generating 100 replicates of the data set. The trees were viewed using MEGA 4.0 visualising software (35). From each phylogenetic tree, sequences that formed clusters were aligned and a consensus sequence was acquired using JALVIEW sequence editor (36). Using BLASTN, the sequence closest to the consensus sequence among the cluster was opted as a representative sequence of that clade. Similarly, a testimonial set of total 26 representative sequences were selected to denote the genetic diversification among the species under the genus *Proteus.* The phylogenetic framework generated was validated by using species specific 16S rRNA sequences from 4 *Proteus* sp. respectively.

BIOPHP (www.biophp.org.in) and NEB CUTTER (37) were used to obtain the restriction pattern of the 4 master data sets. To determine the unique restriction digestion pattern of each species, all the 423 sequences were assayed with 192 different restriction enzymes one at a time. When a restriction site found common to all sequences belonging to single species but astray from other species domain, it was considered as unique restriction site.
The 72 16S rRNA sequences categorised only up to genus level were tested for identification upto species level. The above validated framework sequences were used to categorise the uncharacterized sequences by constructing phylogenetic tree. Further, *in silico* restriction digestion pattern unique to each species was applied to support the hypothesis of reclassification of uncharacterized sequences.

RESULTS

In the present study, 16S rRNA gene sequences downloaded from the RDP database under the genus *Proteus* underwent phylogenetic analysis and restriction enzyme analysis. Such methodology is used to devise a quick way of identification for the clinically relevant microbes and the degree to which sequence related redundancy is present in the database.
The phylogenetic trees were constructed for each *Proteus* species with the aid of PHYLIP 3.6 package (Figure I, II, III and IV). In total, 26 master sequences were first lined for generating framework tree (Table I). With construction of framework tree, diversification within the genus *Proteus* was evaluated. Out of 26 master sequences, 18 sequences belonging to the species *P. mirabilis* and *P. hauseri* can be clearly distinguished intimating about their homogeneity within the genus *Proteus.* Sequences belonging to the species *P. vulgaris* and *P. penneri* showed high degree of similarity thus, making the identification of these two species a difficult task (Fig V). Based on the framework sequences with other sequences of the isolates of single species, a new phylogenetic tree was constructed. Similar validated trees were formed for each species under the genera (Fig VI, VII, VIII and IX). Each of the master sequences was seen to form a distinct clade in each validated phylogenetic tree. The presence of homogeneity within the phylogenetic trees indicates the precise validation of the framework tree generated above. Though some of the species sequences were found clustered with other species sequences and thus reported here under the need of reinvestigation (Table II).
In-silico restriction enzyme analysis was used to generate species specific markers and thus, further supporting the (i) study of heterogeneous species sequences in the genus *Proteus* and (ii) classification of uncharacterized species. When sequences of *Proteus vulgaris* were inspected under the process of *in-silico* restriction digestion, 3 different restriction digestion pattern were discovered unique to the species caused by *Scrf*I, *Msp*R92I and *Bme*1390I cutting at different locations. The restriction enzymes *Vsp*I, *Ase*I, *Psh*BI configured a unique restriction pattern in *Proteus vulgaris*. *Aat*II, *Bsa*HI, and *Hsp*92I generate restriction patterns which were common to all *Proteus* species except for *Proteus mirabilis*. Similarly, *Bsp*HI, *Pag*I, *Rca*I restriction sites were found to be present in 3 *Proteus* species except for *Proteus penneri*. Therefore, their absence can be used as proof of relatedness to a particular species. The results of the restriction sites and length of the nucleotide cuts have been tabulated in Table III.
With similar approach for validation of framework tree, uncharacterized sequences of *Proteus* sp. were checked for preliminary classification. Each of the 72 uncharacterized sequences was aligned with the reference species sequences. The result provided a rough picture of similarity chart of each of the uncharacterized sequence. To consolidate, a

Figure-I: Phylogenetic tree of *Proteus vulgaris***.** The tree was constructed by neighbor-joining method with Jukes and Cantor correction. The numbers at node represent bootstrap values (based on 100 re-sampling) and sequences in red font were exercised as signature sequences.



phylogenetic tree consisting of master sequences and uncharacterized sequences was constructed (Fig X). Out of 72 such sequences, 41 were chosen to be enquired for restriction pattern analysis. About 23 uncharacterized sequences of genus *Proteus* showed similar phylogenetic clustering and restriction pattern as of *P. mirabilis*. Similarly, 4, 4 and 2 uncharacterized sequences confirmed their inclination towards *P. vulgaris*, *P. penneri* and *P. hauseri*, respectively (Table IV).

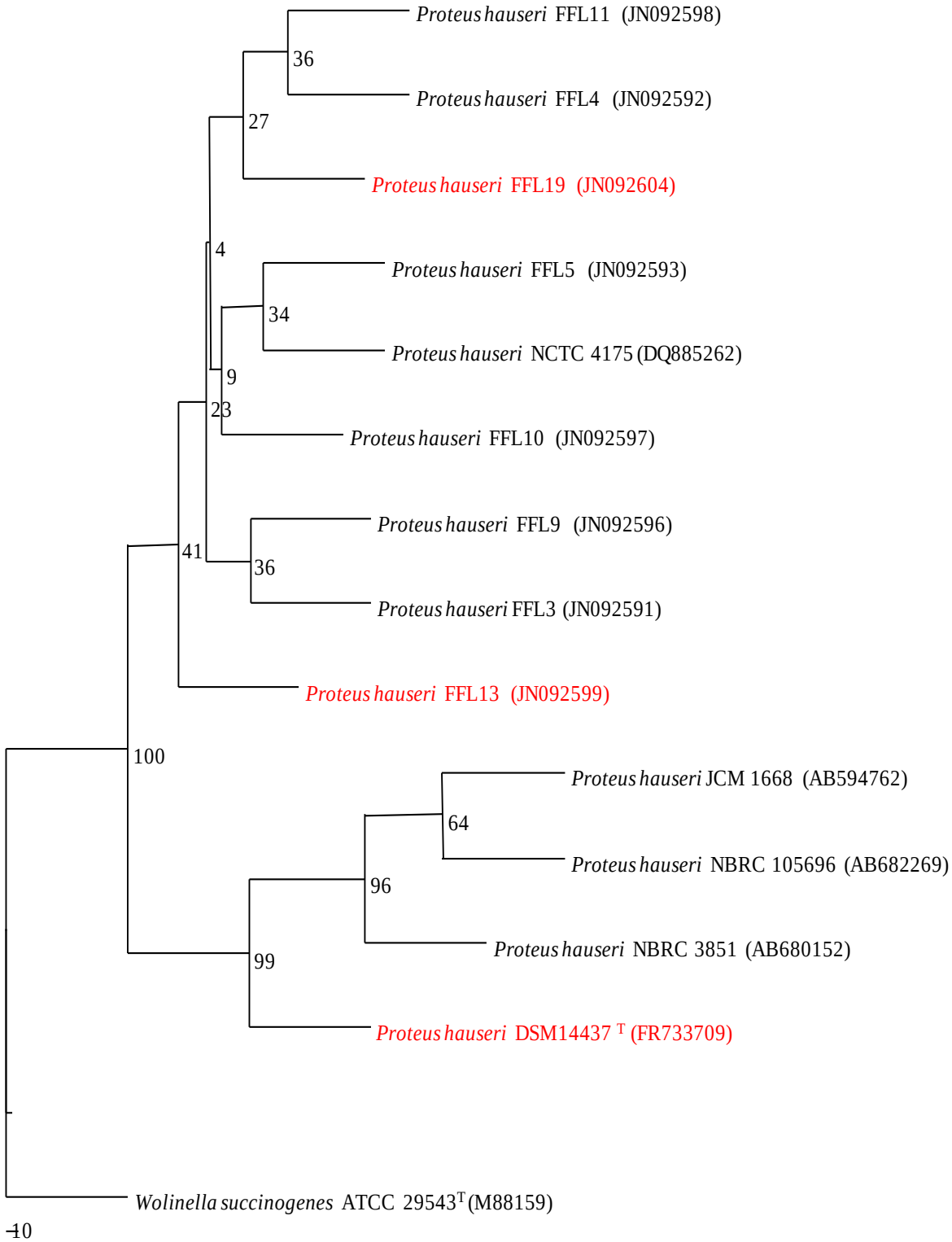


Figure-II: Phylogenetic tree of *Proteus hauseri***.** The tree was constructed by neighbor-joining method with Jukes and Cantor correction. The numbers at node represent bootstrap values (based on 100 re-sampling) and sequences in red font were exercised as signature sequences.

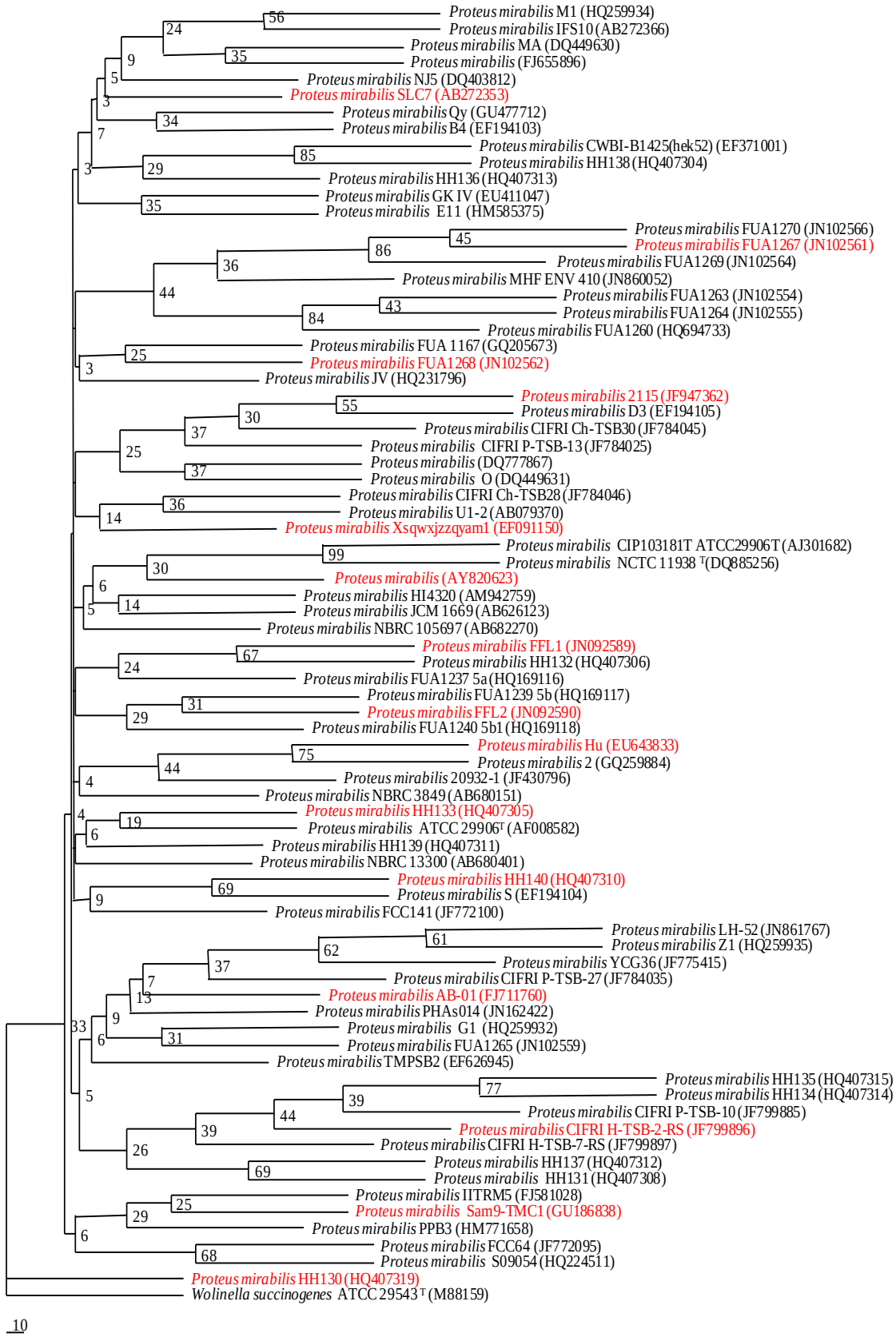


Figure-III: Phylogenetic tree of *Proteus mirabilis*. The tree was constructed by neighbor-joining method with Jukes and Cantor correction. The numbers at node represent bootstrap values (based on 100 re-sampling) and sequences in red font were exercised as signature sequences.

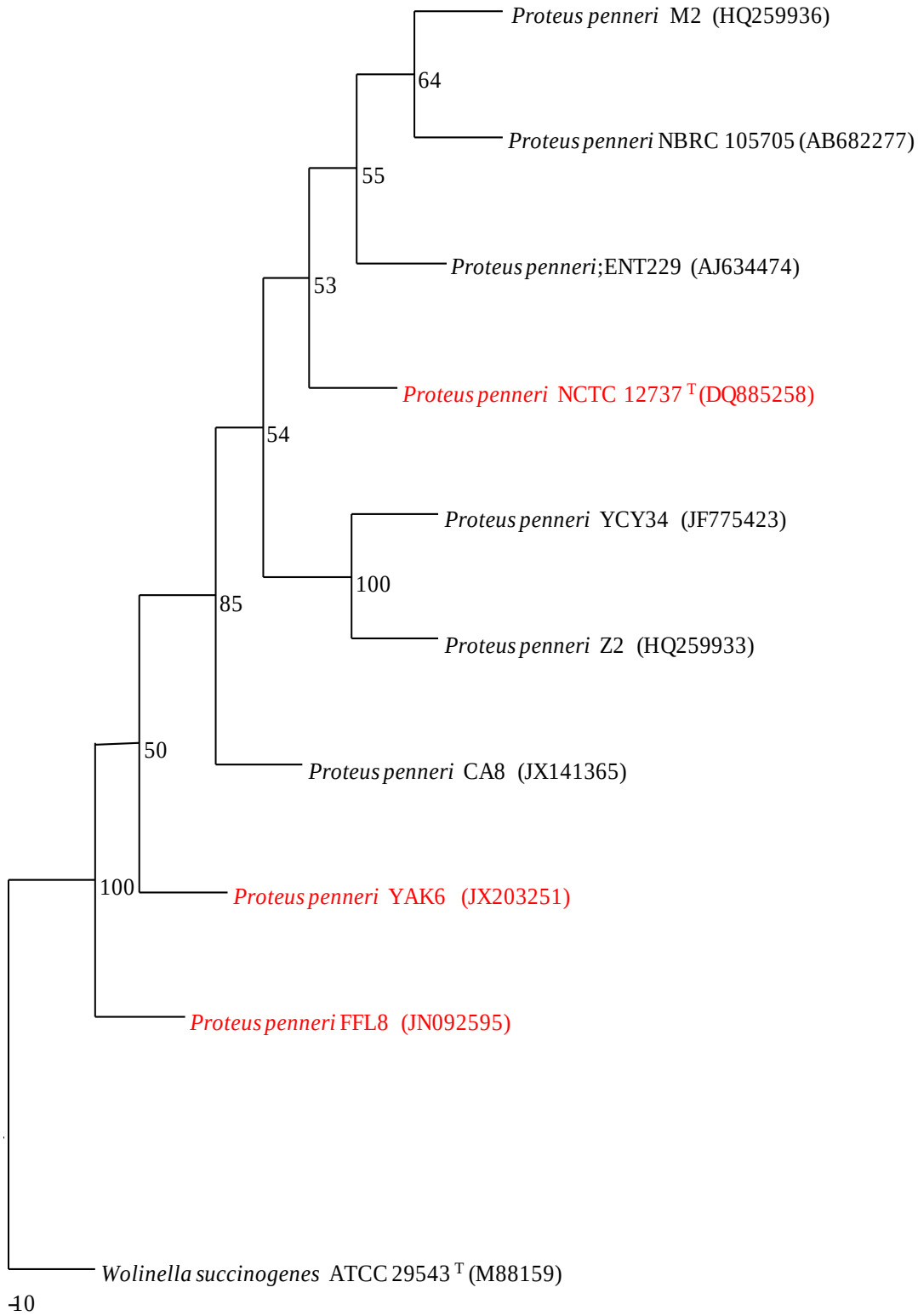


Figure-IV: Phylogenetic tree of *Proteus penneri***.** The tree was constructed by neighbor-joining method with Jukes and Cantor correction. The numbers at node represent bootstrap values (based on 100 re-sampling) and sequences in red font were exercised as signature sequences.

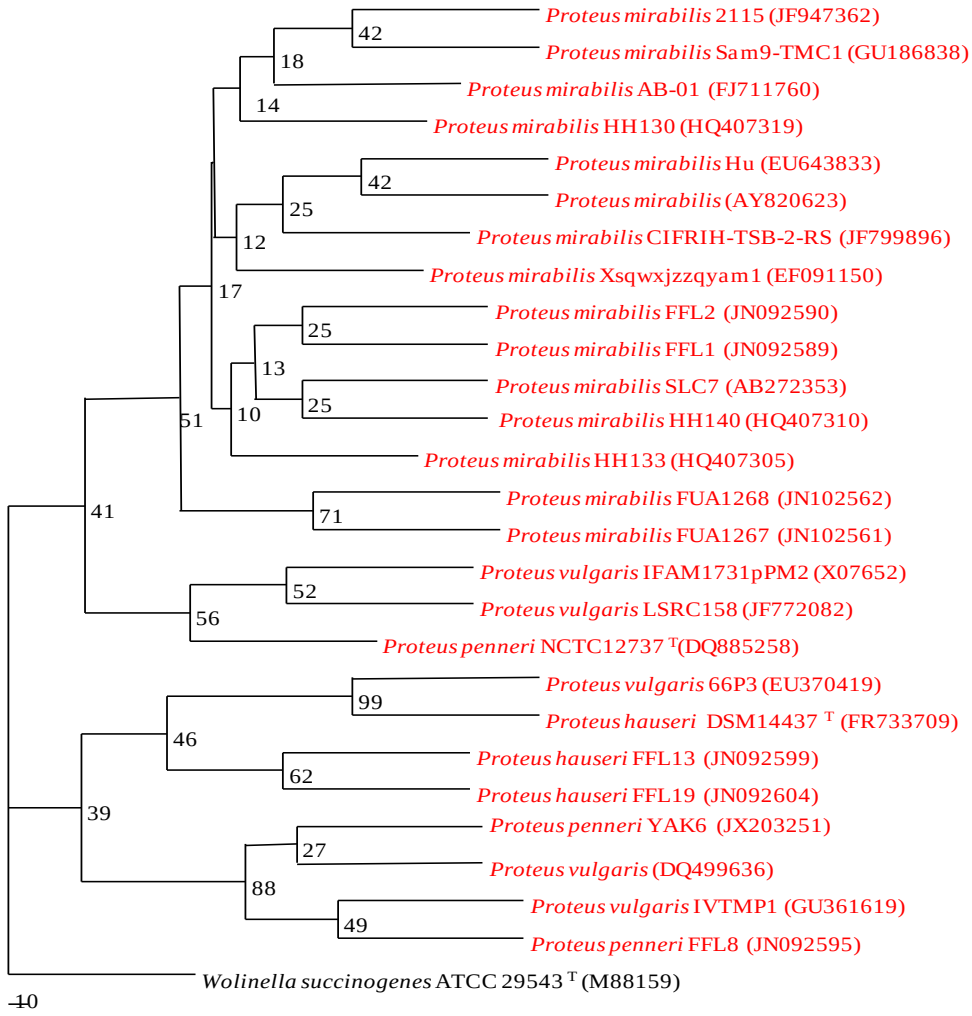


Figure-V: Phylogenetic tree of 26 framework sequences (Red color) of Proteus spp**.** The tree was constructed by neighbor-joining method with Jukes and Cantor correction. The numbers at node represent bootstrap values (based on 100 re-sampling).

Table-I Sequences used for generating framework

Species	No. of Sequences used	No. of clusters obtained	No. of representative sequences	Accession no. of signature sequences
P. mirabilis	90	15	15	JN102562, JN102561, EF091150, HQ407310, JF947362, AY820623, EU64383, JN092589, JN092590, HQ467305, AB272353, JF799896, FT11760, GU186838, HQ407319
P. vulgaris	56	5	5	X07652, JF772082, JN644538, EU370419, DQ205432
P. hauseri	13	3	3	FR733709, JN092599, JN092604
P. penneri	9	3		JN092595, DQ885258, JX203251

S.No.	Species name	Accession number	Modifies species	Restriction pattern analysis
1.	*Proteus mirabilis*	JF772095	*Proteus vulgaris*	+
2.	*Proteus mirabilis*	HQ224511	*Proteus vulgaris*	+
3.	*Proteus vulgaris*	JF772082	*Proteus mirabilis*	+
4.	*Proteus penneri*	JN092595	*Proteus vulgaris*	+
5.	*Proteus vulgaris*	GU361619	*Proteus penneri*	+
6.	*Proteus hauseri*	DQ885262	*Proteus vulgaris*	+
7.	*Proteus penneri*	DQ885258	*Proteus vulgaris*	+
8.	*Proteus vulgaris*	JN644538	*Proteus hauseri*	+
9.	*Proteus penneri*	JF775423	*Proteus mirabilis*	+
10.	*Proteus penneri*	HQ259933	*Proteus mirabilis*	+
11.	*Proteus vulgaris*	EU370419	*Proteus hauseri*	+
12.	*Proteus penneri*	HQ259936	*Proteus vulgaris*	+
13.	*Proteus penneri*	AB682277	*Proteus vulgaris*	+
14.	*Proteus penneri*	AT634474	*Proteus vulgaris*	+

Table-II List of species need to be reconsider for reclassification (above)

Table-III *In-silico* restriction enzyme analysis

Restriction enzyme	Restriction site	Cut site	*P. penneri* $^{\delta}$	*P. hauseri**	*P. vulgaris*^{o}	*P. mirabilis*$^{\times}$
Aat II	1158	G˅ACGTC	+	+	+	-
Hsp92I	1165	GR˅CGYC	+	+	+	-
BspHI	1473	T˅CATG	-	+	+	+
Bme1390I	1382	CC˅NG	-	-	+	-
VspI	460	ATTA˅A	-	+	-	-
BsaHI	1192	GR˅CGYC	+	+	+	-
PagI	1530	T˅CATG	-	+	+	+
MspR91	1389	CC˅NG	-	-	+	-
AseI	460	ATTA˅A	-	+	-	-
RcaI	1462	T˅CATG	-	+	+	+
ScrfI	1392	CC˅NG	-	-	+	-
PshBI	460	ATTA˅A	-	+	-	-

$^{\delta}$ Reference sequence *Proteus penneri* ENT229; AJ634474
* Reference sequence *Proteus hauseri* FFL9; JN092604
^{o} Reference sequence *Proteus vulgaris* TEM11; GQ292550
$^{\times}$ Reference sequence *Proteus mirabilis* FFL2; JN092590

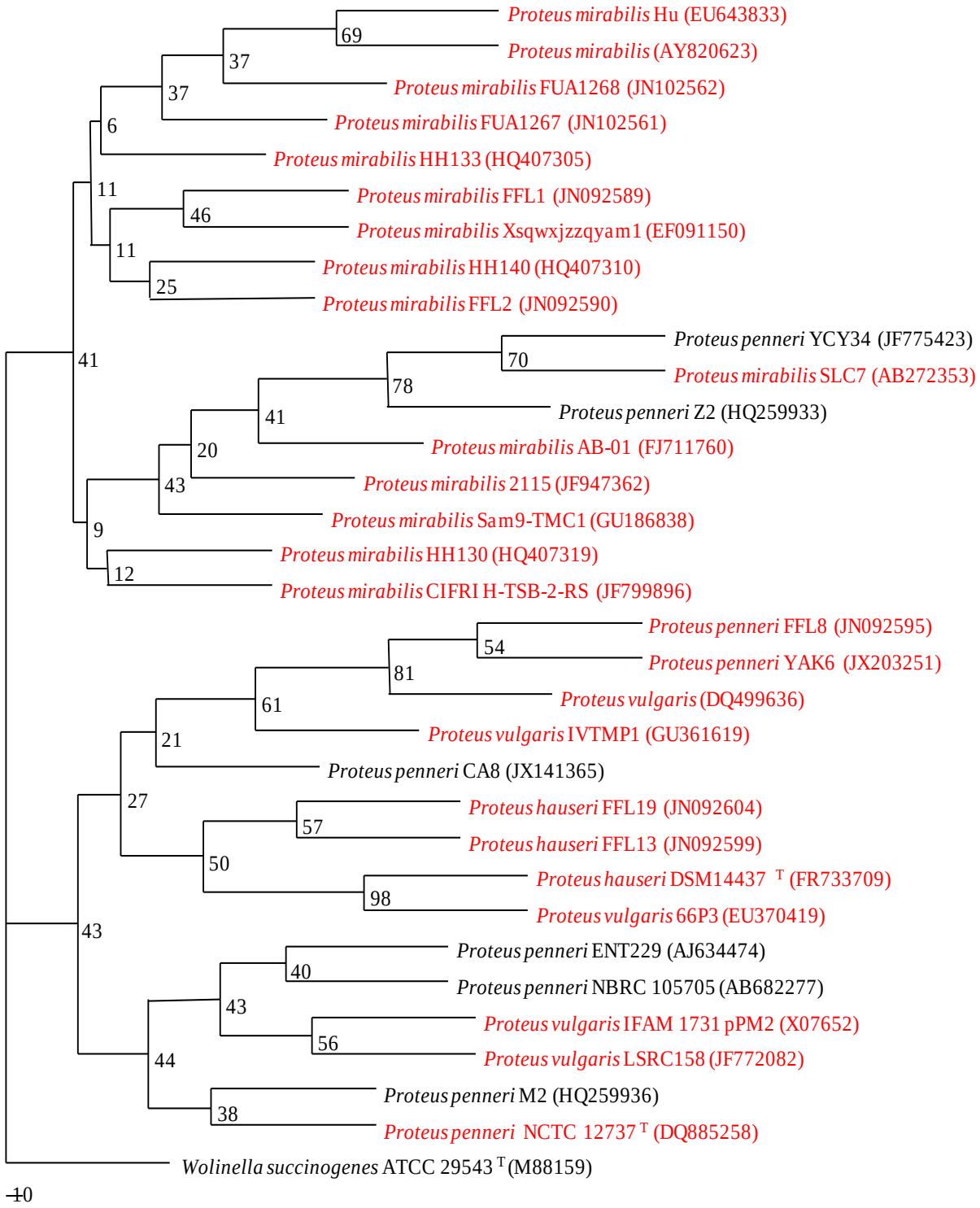


Figure-VII: Validation tree of 26 framework sequences (Red Color) and characterized *Proteus penneri*. The tree was constructed by neighbor-joining method with Jukes and Cantor correction. The numbers at node represent bootstrap values (based on 100 re-sampling).

Figure-VIII: Validation tree of 26 framework sequences (Red Color) and characterized *Proteus mirabilis*. The tree was constructed by neighbor-joining method with Jukes and Cantor correction. The numbers at node represent bootstrap values (based on 100 re-sampling).

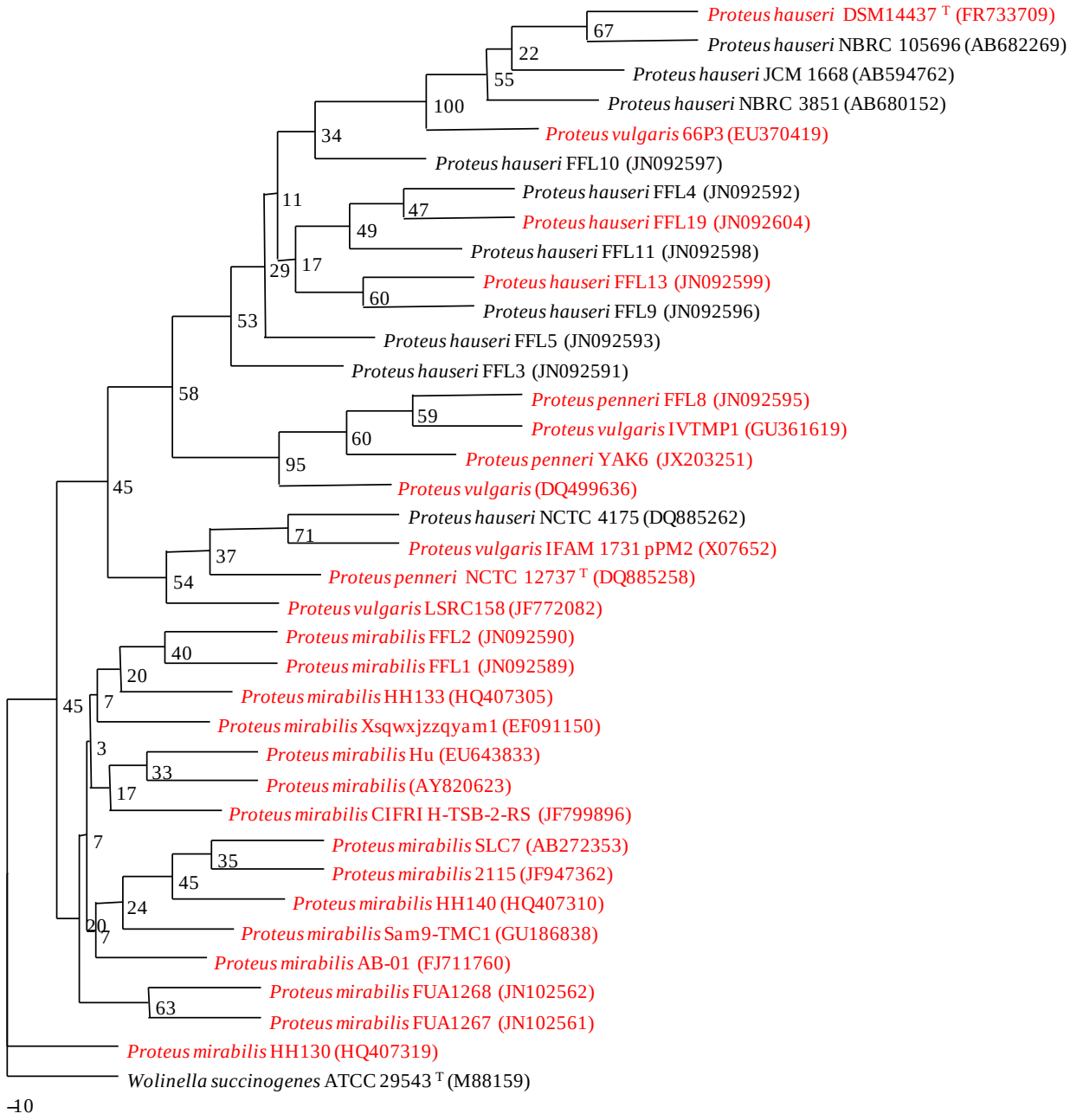


Figure IX: Validation tree of 26 framework sequences (Red Color) and characterized *Proteus hauseri*. The tree was constructed by neighbor-joining method with Jukes and Cantor correction. The numbers at node represent bootstrap values (based on 100 re-sampling).

Table-IV Characterization of uncharacterized species

S.NO.	Accession number	Strain classified into Species	Restriction analysis
1.	JF946807	*Proteus mirabilis*	+
2.	JF946804	*Proteus mirabilis*	+
3.	JF946778	*Proteus mirabilis*	+
4.	JQ695940	*Proteus mirabilis*	+
5.	JX104035	*Proteus mirabilis*	+
6.	HQ009354	*Proteus mirabilis*	+
7.	JF946779	*Proteus mirabilis*	+
8.	JF946780	*Proteus mirabilis*	+
9.	EU382215	*Proteus mirabilis*	+
10.	JF946785	*Proteus mirabilis*	+
11.	AB754811	*Proteus mirabilis*	+
12.	JF946784	*Proteus mirabilis*	+
13.	EF426446	*Proteus vulgaris*	+
14.	GU812899	*Proteus vulgaris*	+
15.	GQ383895	*Proteus vulgaris*	+
16.	JN106439	*Proteus vulgaris*	+
17.	JQ322832	*Proteus penneri*	+
18.	JQ322826	*Proteus penneri*	+
19.	JQ322825	*Proteus penneri*	+
20.	JQ322828	*Proteus penneri*	+
21.	EU710747	*Proteus hauseri*	+
22.	AB538871	*Proteus hauseri*	+

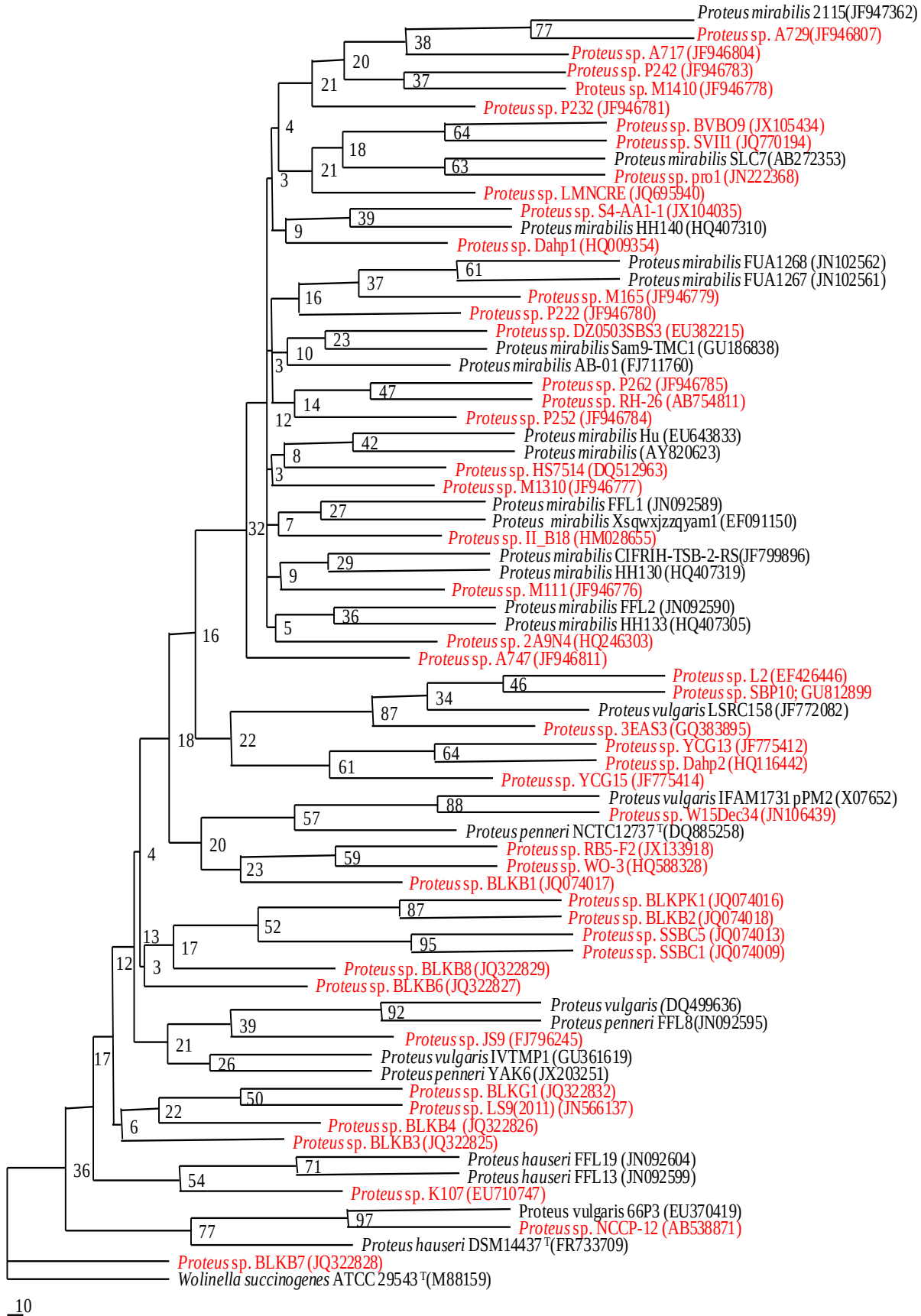


Figure-X: Phylogenetic tree of 26 framework sequences (Red Color) and uncharacterized *Proteus* **spp.** The tree was constructed by neighbor-joining method with Jukes and Cantor correction. The numbers at node represent bootstrap values (based on 100 re-sampling).

DISCUSSION

Considering the harmful effects of pathogenic microorganisms, much of the work focuses on understanding the physiological, biochemical, taxonomical, ecological and phylogenetic behavior of the pathogenic members. Hence, there has been a flood of genomic and other sorts of molecular data entering into the databases. Moreover, the rapidly evolving strains can form another major hurdle for identification process due to their distinctive molecular features but similar phenotypic characters. Therefore, a void of knowledge arose in correct identification of species/strain which contributes to distortion in our medical treatment. The present study involves development of a measure to identify appropriate species sequences. The study was conducted by exploring the internal features of 16S rRNA gene sequences of different *Proteus* species.

From the above framework analysis, it can be concluded that *P .mirabilis* and *P. hauseri* had patterns identical or very similar to those of their respective strain types, so that they could be grouped in two distinct ribo-groups corresponding to two different species.

There were inconsistencies raised within the framework tree. Sequences from *Proteus penneri* were found to be grouped with *Proteus vulgaris* clade. This observation has suggested either about the heterogeneity between two species or their improper classification. To investigate such discrepancies, the sequences were checked for similarities using BLASTN as well as species specific restriction digestion patterns. The result of present study indicated the grouping of *P. penneri*; genetically close to yet distinct from *P. vulgaris*. It can be considered as a most recently differentiated species within the genus *Proteus.*

The validity of the phylogenetic framework sequences was established by categorical segregation of members of given *Proteus* species against a total of 26 reference species sequences. Yet, disparities were raised where some of the sequences of a species were clading with the other species sequences. These discrepant sequences were critically analyzed to establish the heterogeneity among the *Proteus* species. Some of the *P. mirabilis* sequences (JF772095, HQ224511) were found to be more close to *P. vulgaris* sequences. Also, few of *P. vulgaris* sequences (JF772082, GU361619) showed more resemblance with *P. hauseri* and *P. penneri* sequences. A sequence of *P. vulgaris* (EU 370419) was found to form a distinct clade with *P. hauseri* in almost every tree generated. Interestingly, the query sequence showed more similarity and higher bootstrap value towards the *P. hauseri* sequences (99% similarity) rather than *P. vulgaris* sequences (98% similarity). The query sequence also showed similar restriction enzyme digestion pattern of *P. hauseri.* This suggested the need for further study of reclassification of sequences. The result of the above study is tabulated in table III.

The applied nature of the above approach has been used for characterization of the uncharacterized *Proteus* sequences. For 72 uncharacterized sequences identified up to genus level, 33 sequences were found clading with framework sequences. These were also supported by higher bootstrap values. The bootstrap values indicate the probability of a character belonging to the same taxa. Out of 33 concordances, 23 segregated with *P. mirabilis* strains, 4 clustered with *P. vulgaris* strains, 4 branched with *P. penneri* group and 2 with *P. hauseri* species. The characterization was validated using restriction enzyme digestion pattern analysis thus providing a cue about the identification of uncharacterized *Proteus* sequences. The consideration can further be verified using motif analysis and other house-keeping genes to obtain more clear vision.

On account of rapid and precise identification which allows instantaneous diagnosis and in turn, its treatment, it is inclined to model in subtraction of economic catastrophe and degree of mortality (37). This procedure was able to clarify intra-species, inter-species and uncharacterized sequence correlations. This can record the repetitive data in the database and

hence decrease the presence of recurrent sequences to allow more precise identification of pathogenic bacterial sequences.

CONCLUSIONS

The present study aims to identify signatures (inner secrets of 16S rRNA gene) and employing them to identify repetitive species sequences which are present in the database. The phylogenetic framework construction and *in-silico* restriction enzyme analysis aims to define a range of genetic variability within the species which later can be exploited for explanation of redundancy in different species sequences. The result of the present research confirms the grouping of *Proteus mirabilis, Proteus hauseri, Proteus penneri* and *Proteus vulgaris* into distinct species. The intra-species discrimination has been resolved usin g above approaches with a proposal to the identification of some uncharacterized *Proteus* species sequences and reclassification for some of them. Factors like being a conserved sequence, minimal human error and also less biochemical variability mold the above procedure into a quick and cost effective approach. This will reduce the database redundancy to allow accurate and precise identification of organisms based on the genomic data.

ACKNOWLEDGMENTS

Samarth Kulshrestha and Arnab Kapuria acknowledge University of Delhi Innovative project for providing the research fellowship.

REFERENCES

(1) Tumbarello, M., Trecarichi, E.M., Flori, B., hostel, A.R., Fadda, G. and Spanu, T. Multi-Drug Resistant *Proteus mirabilis* bloodstream infections: Risk factors and outcomes. *Antimicrobial agents and Chemotherapy*, 2012, 56(6), 3224-3231.

(2) Dauga, C. Evolution of the gyrB gene and molecular phylogeny of *Enterobacteriaceae*: a model molecule for molecular systematic studies. *International Journal of Systematic and Evolutionary Microbiology*, 2002, 52, 531-547.

(3) Manas, J. and Belas, R. The Genera *Proteus, Providencia and Morganella. Prokaryotes,* 2006, 6, 254-269.

(4) Kishore, J. Isolation, identification and characterization of *Proteus penneri*- a missed rare pathogen. *Indian Journal of Medical Research.* 2002, 135, 341-355.

(5) Sabbuba, N.A., Mahenthiralingam, A., and Stickler, D.J. Molecular Epidemiology of *Proteus mirabilis* infections of the catheterized UT. *Journal of Clinical Microbiology*. 2003, 41(11), 4961-4965.

(6) Burall, L. S., Harro, J.M., Li, X., Lockatell, C.V., Himpsl, S.D., Hebel, J.R., Johnson, D.E. and Mobley, H.L.T. *Proteus mirabilis* that contribute to pathogenesis of Urinary Tract infections: Identification of 25 signature-tagged mutants attenuated at least 100-fold. *Infection and Immunity*, 2004, 72(5), 2922-2938.

(7) Jacobsen, S.M. and Shirtiff, M.E. *Proteus mirabilis* biofilms and catheter associated urinary tract infections. *Virulence*, 2011, 2(5), 460-465.

(8) Nakamura, S., Maeda, N., Miron, I.M., Yoh, M., Izutsu, K., Kataoka, C., Honda, T., Yasunaga, T., Nakaya, T., Kawai, J., Hayashizaki, Y., Horii, T. and Iida, T. Metagenomic analysis of bacterial identification. *Emerging infectious disease*, 2008, 14, 11, 1784-1786.

(9) Mordi, R.M. and Momoh, M.I. Incidence of *Proteus* species in wound infections and their sensitivity pattern in the University of Benin Teaching Hospital. *African Journal of Biotechnology*. 2009, 8(5), 725-730.

(10) Feglo, P.K., Gbedema, S.Y., Quay, S.N.A., Adu-Sarkodie, Y., Opoku-Okrah, C. Occurrence, species distribution and antibiotic resistance of *Proteus* isolation: A case study at the Komfa Anokye Teaching Hospital (KATH) in Ghana. *International Journal of Pharma Science and Research*. 2010, 1(9), 347-352.

(11) Nwankwo, E. Isolation of pathogenic bacteria from somites in the operating rooms of a specialist hospital in Kano, North-western Nigeria. *Pan African Medical Journal*. 2012, 12(90), 1-10.

(12) O'Hara, C.M., Brenner, F.W., Steigerwalt, A.G., Hill, B.C., Miller, J.M. and Brenner, D.J. Classification of *P. vulgaris* biogroup 3 with recognition of *Proteus hauseri* sp. Nov, nom. Rev. and unnamed *Proteus* genomo-species. 4, 5 and 6. *Intenational Journal of Systematic and Evolutionay Microbiology*, 2000, 50, 1869-1875.

(13) Janda, J.M., Abbott, S.L., Khashe, S. and Probert,W. Biochemical identification and characterization of DNA groups within the *Proteus vulgaris* complex. *Journal of Chemical Microbiology*, 2001, 39(4), 1231-1234.

(14) Engler, H.D., Troy, K., Bottone, E.J. Bacteremia and subcutaneous abscess caused by *Proteus penneri* in a neutropenic host. *Journal of Clinical Microbiology*, 1990, 28 (7), 1645–1646.
(15) Hickman, F.W., Steigerwalt, A.O., Farmer, J.J, 3rd and Brenner. Identification of *Proteus penneri* sp. nov., formerly known as Proteus vulgaris indole negative or as *P. vulgaris* biogroup 1. *Journal Clinical Microbiology*. 1982, 15(6), 1097-1102.
(16) Knirel, Y.A., Perepelov, A.V., Kondakova, A.N., Sidorczyk, Z., Kaca, W. Structure and serology of O-antigens as the basis for classification of *Proteus* strains. *Innate Immunity*. 2011, 17(1), 70-96.
(17) Huang, C.T. Multi-test media for rapid identification of *Proteus* species with notes on biochemical reactions of strains isolated from urine and pus. *Journal of Clinical Pathology*, 1966, 19, 438.
(18) O'Hara, C.M., Brenner, F.W. and Miller, J.M. Classification, identification and clinical significance of *Proteus, Providencia and Morganella*. *Clinical Microbiology Review*, 2000, 13(4), 534-546.
(19) Lal,D., Verma, M., and Lal, R. Exploring internal features of 16sr RNA gene for identification of clinical relevant species of the genus *Streptococcus*. *Annals of Microbiology and Antimicrobials*, 2011, 10(28), 1-11.
(20) Giammanco, G., Pignato, S., Grimont, P., Grimont, F., Giammanco, G. Genotypy of the genus *Proteus* by rpoB sequence analysis. *Italian Journal of Public Health*, 2006, 19-22.
(21) Krajden, S., Fuksa, M., Petrea, C., Crisp, L.J., Penner, J.L. Expanded clinical spectrum of infections caused by *Proteus penneri*. *Journal of Clinical Microbiology*, 1987, 25(3), 578–579.
(22) Aquilini, E., Azevedo, E., Jimenez, N., Bouamama, L., Tomas, J.M. and Regue, M. Functional identification of the *Proteus mirabilis* Core Lipopolysaccharide Biosynthesis Genes. *Journal of Biotechnology*, 2010, 4413-4424.
(23) (27) Widjojoatmodjo, M.N., Fluit, A.C. and Verhoef, J. Rapid identification of bacteria by PCR-single-strand conformation polymorphism. *Journal of Clinical Microbiology*, 1994, 32(12), 3002-3007.
(24) Himpsl, S.D., Lockatell, C.V., Hebel, J.R., Johnson, D.E., Mobley, H.T. Identification of virulence determinants in uropathogenic *Proteus mirabilis* using signature-tagged mutagenesis. *Journal of Medical Microbiology*, 2008, 57, 1068-1078.
(25) Goswami, N. N., Trivedi, H.R., Goswami, A.P., Patel, T. and Tripathi, C.B. Antibiotic sensitivity profile of bacterial pathogens in postoperative wound infections at a tertiary care hospital in Gujarat, India. *Journal of Pharmacology Pharmacotherapy*, 2011, 2(3), 158-164.
(26) Janda, J.N. and Abbott, S.L. 16S rRNA Gene Sequencing for Bacterial Identification in the Diagnostic Laboratory: Pluses, Perils, and Pitfalls. *Journal of Clinical Microbiology*, 2007, 45(9), 2761–2764.
(27) Woo, P. C. Y., Lau, S. K. P., Teng, J. L. L., Tse H. and Yuen, K.-Y. Then and now: use of 16S rDNA gene sequencing for bacterial identification and discovery of novel bacteria in clinical microbiology laboratories. *Clinical Microbiology and Infections*, 2008, 14, 908–934.
(28) Woo, P.C.Y., Kenneth H. L. Ng, Lau, S.K.P., Yip, K.T., Fung, A.M.Y., Leung, K.T., Tam, D.M.W., Que, T.L. and Yuen, K.Y. Usefulness of the MicroSeq 500 16S Ribosomal DNA-Based Bacterial Identification System for Identification of Clinically Significant Bacterial Isolates with Ambiguous Biochemical Profiles. *Journal of Clinical Microbiology*, 2003, 41(5), 1996–2001.
(29) Kalia, V.C., Mukherjee, T., Bhushan, A., Joshi, J., Pratap and Huma, N. Analysis of the unexplored features of rrs (16S rDNA) of the Genus *Clostridium*. *BMC Genomics*, 2011, 12(18), 1-35.
(30) Dempsey, K.E., Riggio, M.P., Lennon, A., Hannah, V.E., Ramage, G., Allan, D. and Bagg, J. Identification of bacteria on the surface of clinically infected and non-infected prosthetic hip-joints removed during revision arthroplasties by 16S rRNA gene sequencing and by microbiological culture. *Arthritis Research and Therapy*, 2007, 9(3), 1-11.
(31) Puri, A., Rai, A., Dhanaraj, P.S., Lal, R., Patel, D.D., Kaicker, A., Verma, M. An in silico approach for identification of the pathogenic species, *Helicobacter pylori* and its relatives. *Indian J Microbiol*, 2016, doi:10.1007/s12088-016-0575-7.
(32) Maidak, B.L., Cole, J.R., Lilburn, T.G., Parker, C.T., Saxman, P.R., Stredwick, J.M., Garrity, G.M., Li, B., Olsen, G.J., Pramanik, S., Schmidt, T.M., Tiedje, J.M. The RDP (Ribosomal Database Project) continues. *Nucleic Acids Research*, 2000, 28, 173-174.
(33) Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., Higgins, D.G. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research*, 1997, 25, 4876-4882.
(34) Hendy, M.D., Penny, D. and Steel, M.A. A discrete Fourier analysis for evolutionary tress. *Evolution*, 1994, 91, 3339-3343.
(35) Tamura, K., Dudley, J., Nei, M., Kumar, S. MEGA 4: Molecular Evolutionary Genetics Analysis Software version 4.0. *Molecular Biology and Evolution*, 2007, 24(8), 1596-1599.
(36) Clamp, M., Cuff, J., Searle, S.M., Barton, G.J. "The Jalview Java Alignment Editor". *Bioinformatics*, 2004, 20, 426-427.
(37) Vincze, T., Posfai, J. and Roberts, R.J. NEBcutter: a program to cleave DNA with restriction enzymes. *Nucleic Acids Research*, 2003, 31(13), 3688–3691.