

Exploiting Bacterial Genomes to Develop Biomarkers for Identification

Ravi Kumar, Shikha Koul, and Vipin Chandra Kalia

Abstract Bacteria have unique abilities to adjust itself to diverse environmental conditions. Under adverse conditions, their genetic reservoir provides necessary help. Although bacteria have been perceived as pathogens to most living beings, the most critical are the ones which infect human being. Bacteria also harbour the human gut and skin and have been shown to be helpful. The pathogenic bacteria cause diseases and contribute to ill health. The need is to identify them rapidly, diagnose the disease and initiate the treatment. Most bacteria can be easily identified on the basis of their 16S rRNA (*rrs*) gene. However, in case where multiple copies of *rrs* are present within a bacterial genome, it is difficult to identify them, since they show great homology with other species of a genus. Here, novel approaches have been reviewed, which rely upon certain genes which are common to a large number of species of *Clostridium*, *Lactobacillus*, *Staphylococcus*, *Streptococcus*, *Vibrio* and *Yersinia* and show unique digestion patterns on treatment with restriction endonucleases.

1 Introduction

Bacteria have been bestowed with inherent abilities to withstand extreme environmental conditions and undergo genetic changes to evolve rapidly (Kalia 2010). Their interactions with human beings are extremely varied. The benign type of bacteria persists within the human gut while quite a few exist and survive well on the human body (Yu et al. 2014; Arasu et al. 2015). The pathogenic bacteria are a

R. Kumar

Microbial Biotechnology and Genomics, CSIR – Institute of Genomics and Integrative Biology (IGIB), Delhi University Campus, Mall Road, Delhi 110007, India

S. Koul • V.C. Kalia (✉)

Microbial Biotechnology and Genomics, CSIR – Institute of Genomics and Integrative Biology (IGIB), Delhi University Campus, Mall Road, Delhi 110007, India

Academy of Scientific & Innovative Research (AcSIR), 2, Rafi Marg, Anusandhan Bhawan, New Delhi 110001, India

e-mail: vckalia@igib.res.in; vc_kalia@yahoo.co.in

major threat to human health (Gautam et al. 2014; Wang et al. 2014). Bacteria invade the human body and then cause extensive damage, especially in the cases of infectious diseases. The bacterial pathogens may live and survive inside the body for a long term, e.g. tuberculosis, or for a short term as in diarrheas. In these two extreme scenarios, the former are slow growers with an excellent ability to evade the human immune system. In the latter case, the growth is rapid and the expression turns violent. Among the pathogens, which are a worry for Health Departments, notable ones are *Streptococcus*, *Shigella*, *Helicobacter*, *Vibrio*, *Clostridium*, *Yersinia*, *Salmonella*, etc. (Mahale et al. 2014; Kalia et al. 2015, 2016; Kalia and Kumar 2015; Kekre et al. 2015; Koul et al. 2015a, b; Koul and Kalia 2016; Kumar et al. 2016; Meza-Lucas et al. 2016; Puri et al. 2016; Yagnik et al. 2016). The discovery of antibiotics turned out to be a blessing for human beings. However, this arsenal invariably used to get rid of these pathogens didn't live up to its expectations. Bacteria could develop resistance to most of the antibiotics. Few research groups are still showing courage to invest time and money to develop novel antibiotics (Agarwala et al. 2014; Saxena et al. 2014; Prakasham et al. 2014; Alipiah et al. 2015; Sajid et al. 2015). Pharmaceutical companies have completely lost interest in investing in this area. It is now for the governments to look into this societal issue. A mechanism which is adding fuel to the fire is the bacterial ability to become "resistant" to antibiotics, without undergoing any genetic changes. Here, they multiply silently without any evident signs of their presence. Once they reach a threshold population density, they start expressing genes which never get expressed while the bacteria are in low numbers. This phenomenon of quorum sensing is responsible for the expression of pathogenic factors (Kalia and Purohit 2011; Kalia et al. 2011b; Gui et al. 2014; Hema et al. 2015; Kaur et al. 2015; Koul et al. 2015b; Arya and Princy 2016). Among these biofilm formation is one of the most dangerous activities. It provides a shield to the bacteria which thus show "resistance" to antibiotics (Shang et al. 2014). There seems to be a battle which is always going on between bacteria and antibiotics (Kalia 2013a, 2014, 2015b). Under all these circumstances, the need is to rapidly identify the pathogen and diagnose the disease in an unambiguous manner before it becomes unmanageable.

2 The Conventional Ways

A wide range of bacterial identification methods is employed depending upon the need and available facilities: amplified fragment length polymorphism (AFLP), microarray, restriction endonuclease digestion (RE), PCR-ribotyping, multilocus sequence analysis, randomly amplified polymorphic DNA and DNA-DNA re-association (Prakash et al. 2014). One of the most frequently employed molecular markers to identify microbes is the 16S rRNA (*rrs*) gene sequence. The evidence for its popularity among researchers can be gauged from the Ribosomal

Database Project (RDP), which contains around 3.2 million *rrs* entries (<https://rdp.cme.msu.edu/>). The credit goes to the innovative thinking of Prof. Carl R. Woese (Kalia 2013b; Prakash et al. 2013). In certain cases, the *rrs* gene sequence does not prove helpful, especially if the organisms belong to closely related taxa. As an alternative to *rrs* gene, people resort to the usage of a few housekeeping genes (HKGs), which are also highly conserved throughout the bacterial world—*recA*, *gyrA*, *gyrB*, *rpoB*, etc. These HKGs are used in various combinations and have proved effective in distinguishing closely related organisms (Porwal et al. 2009; Bhushan et al. 2015). Here, invariably up to eight genes are required to achieve meaningful result, which obviously amounts to higher investments of time, resources and money.

3 Exploring the Hidden Potential of *rrs*

At times, full length comparisons of *rrs* genes from different species or strains don't prove effective enough to distinguish and identify the bacteria unambiguously. Variable regions of *rrs* and their combinations have also been employed for establishing their identity and phylogenetic relations. More recent efforts have revealed the presence of a few features, which have not been exploited for bacterial identification. Molecular markers like unique nucleotide signatures and specific endonuclease restriction digestion patterns were developed for *Bacillus*, *Clostridium* and *Pseudomonas*. Another interesting parameter developed in these studies was the phylogenetic frame work, which could define the phylogenetic limits of a species (Porwal et al. 2009; Kalia et al. 2011a; Bhushan et al. 2013, 2015; Kalia 2015a). It could be extrapolated to identify those bacteria which are yet to be seen by the databases. These genomic tools have been extended to identify members of the genus *Streptococcus* (Lal et al. 2011) and *Helicobacter* (Puri et al. 2016).

4 The Trouble with *rrs* Genes

In the organisms like *Clostridium*, *Lactobacillus*, *Staphylococcus*, *Streptococcus*, *Vibrio* and *Yersinia*, each genome possesses multiple copies (4–13) of the *rrs* gene (Table 1). (Klappenbach et al. 2001; Kalia et al. 2015, 2016; Kalia and Kumar 2015; Kekre et al. 2015; Koul et al. 2015a; Koul and Kalia 2016; Kumar et al. 2016). This is expected to result in overestimation of bacterial populations. The other limitation is a high level of similarity observed between *rrs* copies of genomes from different species, which may lead to mislabelling of the organism.

Table 1 Characteristics of sequenced genomes of certain Gram-positive bacteria: *Lactobacillus*, *Clostridium* and *Staphylococcus* species (www.ncbi.nlm.nih.gov)

Organism	Genome					References
	Size (Mb)	%GC	No. of genes	No. of proteins	<i>rrs</i> copies	
<i>Clostridium</i>	2.54–6.00	27.40–32.02	3911–4057	2315–5020	8–13	Kekre et al. (2015)
<i>Lactobacillus</i>	1.37–3.36	33.01–51.50	1434–3148	1298–3004	4–9	Koul and Kalia (2016)
<i>Staphylococcus</i>	2.56–3.07	32.05–37.60	2386–3163	2275–3041	5–6	Kumar et al. (2016)
<i>Streptococcus</i>	1.75–2.38	35.60–41.60	1693–2408	1762–2270	4–7	Kalia et al. (2016)
<i>Vibrio</i>	4.03–6.32	38.37–47.49	3656–5807	3406–5574	7–11	Kalia et al. (2015)
<i>Yersinia</i>	3.60–4.96	46.95–48.05	3219–5596	3079–5498	6–8	Kalia and Kumar (2015)

Table 2 Genes commonly used for identifying different bacterial species

Genus	Genes	Reference
<i>Lactobacillus</i>	<i>recA</i> , <i>pheS</i> , <i>pyrG</i> , <i>tuf</i> , <i>sphI</i> , <i>mub</i> , <i>fbp</i> , <i>bsh</i>	Naser et al. (2007), Sarmiento-Rubiano et al. (2010), Nguyen et al. (2013), Yu et al. (2015), Koul and Kalia (2016)
<i>Staphylococcus</i>	<i>coa</i> , <i>femA</i> , <i>femB</i> , <i>gyrA</i> , <i>spa</i> , <i>mecA</i> , <i>atlE</i> , <i>tuf</i> , <i>ileS</i>	Brown et al. (2005), Pichon et al. (2012), Roberts (2014), Kumar et al. (2016)
<i>Streptococcus</i>	<i>cpsA</i> , <i>gdh</i> , <i>groESL</i> , <i>lytA</i> , <i>psaA</i> , <i>pspA</i> , <i>recA</i> , <i>recN</i> , <i>rpoA</i> , <i>rpoB</i> , <i>sodA</i>	Carvalho et al. (2007), Bishop et al. (2009), Abdeldaim et al. (2010), Kalia et al. (2016)
<i>Vibrio</i>	<i>rpoB</i> , <i>hsb60</i> , <i>sodB</i> , <i>flaE</i>	Tarr et al. (2007), Taneja et al. (2012), Bhattacharyya and Hou (2013), Kalia and Kumar (2015)

5 The Potential Alternatives to *rrs* Gene

Molecular tools used for distinguishing *Lactobacillus*, *Staphylococcus*, *Streptococcus* and *Vibrio* species involve a wide range of genes (Table 2). The molecular tools like (1) loop-mediated isothermal amplification (LAMP), (2) LAMP combined with lateral flow dipstick (Surasilp et al. 2011; Plaon et al. 2015; Thongkao et al. 2015), (3) a silicon-based optical thin-film biosensor chip (Bai et al. 2010), (4) Fourier transform infrared spectroscopy and (5) matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (Kuhm et al. 2009; Ayyadurai et al. 2010; Stephan et al. 2011) are highly sensitive but costly as well.

Most of these genes had been effective in identifying bacteria up to the species level (Table 2) (Brown et al. 2005; Carvalho et al. 2007; Naser et al. 2007; Tarr et al. 2007; Bishop et al. 2009; Abdeldaim et al. 2010; Kingston et al. 2010;

Sarmiento-Rubiano et al. 2010; Pichon et al. 2012; Taneja et al. 2012; Bhattacharyya and Hou 2013; Liu et al. 2013; Nguyen et al. 2013; Roberts 2014; Illegheems et al. 2015; Kalia and Kumar 2015; Kalia et al. 2015, 2016; Yu et al. 2015; Koul and Kalia 2016; Kumar et al. 2016). A major limitation with the previous studies has been the use of genes which were not present in all the species of a given genus, e.g. *Yersinia* and *Lactobacillus* (Bhagat and Viridi 2007; Kishore et al. 2012; Raftis et al. 2014; Illegheems et al. 2015; Moroeanu et al. 2015; Petrova et al. 2015). In addition, multiple gene analyses have also been employed in many cases. The use of different genes in different studies reflects that in spite of their usage over a long period, no consensus gene has been identified as yet.

6 Screening Genomes for Biomarkers

An innovative approach for searching novel markers in organisms possessing multiple copies of *rrs* has been developed recently (Table 1) (Kalia et al. 2015, 2016; Kalia and Kumar 2015; Kekre et al. 2015; Koul et al. 2015a; Koul and Kalia 2016; Kumar et al. 2016). It involves finding genes, which are common to almost all the species of a genus. From this pool of common genes, 8–34 representative genes (200–4000 nucleotides) were selected (Table 3) (Kalia et al. 2015, 2016). In silico digestion of each of these genes with ten type II restriction endonucleases (REs) (4–6 base cutters) revealed interesting size and sequence of fragments. Unique RE digestion patterns are selected out and can be used for identification of particular organism with high precision (Table 4). Biomarkers for identifying strains have been deduced by using combination of REs (*AluI*, *BfaI*, *BfuCI*, *CviAII*, *HpyCH4V*, *RsaI*, *TaqI* and *Tru9I*) and common genes (Table 4): (a) *Clostridium*—(i) *recN*, *dnaJ*, *secA*, *mutS* and *grpB* (Kekre et al. 2015)(b) *Lactobacillus*—*dnaA*, *dnaJ*, *gyrB*, *pusA*, *recA* and *ruvB* (Koul and Kalia 2016)(c) *Staphylococcus*—*argH*, *argR*, *cysS*, *gyrB*, *purH*, *pyrE* and *recA* (Kumar et al. 2016)(d) *Streptococcus*—*dnaA*, *dnaK*, *gabG*, *mraY*, *purH*, *purK* and *pyrH* (Kalia et al. 2016)(e) *Vibrio*—*dapF*, *hisD*, *ilvH*, *lpxC*, *recF*, *recR*, *rph* and *ruvB* (Kalia et al. 2015)(f) *Yersinia*—*aceE*, *cysJ*, *fadJ*, *gltB*, *gyrB*, *leuD*, *ligA*, *mukB*, *rpoB* and *secA* (Kalia and Kumar 2015) This approach allows detection of bacteria even in a mixed population. It also offers a choice to select a single gene or a combination of genes. Since it does not involve any costly and sophisticated equipment, which demands time, money and highly skilled manpower, the chances of its being exploited on a large scale are high. Thus, the rapidity with which the bacteria can be identified will allow the treatment to initiate quickly.

It was interesting to learn that in the case of the genes which are common to different genera, this approach still has the potential to prove effective in distinguishing them unambiguously (Tables 5 and 6):(1) The *dnaJ* gene in *Clostridium* and *Lactobacillus* and the *gyrB* of *Lactobacillus* and *Yersinia* species were found to provide unique digestion patterns with RE—*AluI* (Table 5).(2) *purH* gene in *Staphylococcus* and *Streptococcus* showed unique patterns on digestion with REs

Table 3 List of genes common to most species within a genus: digested with different restriction endonucleases^a

Gene	Genus												
	<i>Clostridium</i>	<i>Lactobacillus</i>	<i>Staphylococcus</i>	<i>Streptococcus</i>	<i>Vibrio</i>	<i>Yersinia^b</i>	Gene	<i>Clostridium</i>	<i>Lactobacillus</i>	<i>Staphylococcus</i>	<i>Streptococcus</i>	<i>Vibrio</i>	<i>Yersinia</i>
<i>aceE</i>	-	-	-	-	+	+	<i>gyrB</i>	-	+	+	+	+	+
<i>argG</i>	-	+	+	-	-	-	<i>ileS</i>	-	-	-	-	+	+
<i>argH</i>	-	-	+	-	-	-	<i>ligA</i>	-	-	-	-	+	+
<i>argR</i>	-	-	+	+	-	-	<i>mraY</i>	-	+	+	+	-	-
<i>argS</i>	-	+	+	+	-	-	<i>murC</i>	-	+	+	+	-	-
<i>aroB</i>	-	-	+	-	-	-	<i>nusA</i>	-	-	+	+	+	+
<i>aroD</i>	-	-	-	+	-	-	<i>phcS</i>	-	+	-	-	-	-
<i>asnB</i>	-	-	-	-	+	-	<i>pheT</i>	-	-	+	+	+	+
<i>carB</i>	-	+	-	-	+	+	<i>purH</i>	-	-	+	+	-	-
<i>clpB</i>	+	-	-	-	-	-	<i>purK</i>	-	-	-	+	-	-
<i>clpX</i>	+	-	-	-	-	-	<i>purR</i>	-	-	-	+	-	-
<i>cpxA</i>	-	-	-	-	+	-	<i>pyrB</i>	-	+	-	-	-	-
<i>cpxP</i>	-	-	-	-	-	+	<i>pyrE</i>	-	-	+	-	-	-
<i>cysJ</i>	-	-	-	-	-	+	<i>pyrG</i>	-	+	-	+	-	-
<i>cysS</i>	-	+	+	+	-	+	<i>recA</i>	+	+	+	+	-	-
<i>dnaA</i>	+	+	+	+	-	-	<i>recR</i>	+	-	-	-	+	-
<i>dnaE</i>	-	-	-	-	+	-	<i>recU</i>	-	+	+	+	-	-
<i>dnaI</i>	+	+	-	-	-	-	<i>rhlL</i>	-	-	-	-	+	-
<i>dnaK</i>	-	+	+	+	-	+	<i>rpoB</i>	-	-	-	-	-	+
<i>fabG</i>	-	-	-	+	-	-	<i>ruvB</i>	+	+	+	+	+	+
<i>glyA</i>	-	-	-	+	-	-	<i>secA</i>	+	-	+	+	+	+
<i>guaA</i>	-	+	+	-	-	-	<i>trmD</i>	-	+	-	+	-	-

References: Kekre et al. (2015), Kalia and Kumar (2015), Kalia et al. (2015, 2016), Koul and Kalia (2016), Kumar et al. (2016)

^a*AluI*, *BfuCI*, *CviAI*, *HaeI*, *HinI*, *HpyCH4V*, *RsaI*, *TaqI* and *Ttr9I*^bThe following genes were studied only in the case of *Yersinia*: *aceK*, *fadI*, *feoB*, *glnU*, *glpQ*, *gltV*, *gltX*, *hisG*, *lecZ*, *lecZ*, *leuD*, *malE*, *malE*, *metH*, *mltA*, *mukB*, *nagB*, *ribD*, *ripA*, *ripW* and *thiP* (Kalia and Kumar 2015)

Table 4 Gene–restriction endonuclease combinations with unique digestion patterns

Organisms	Genes	Restriction endonucleases										References	
		<i>AluI</i> ^a	<i>Bfal</i>	<i>BfuCI</i>	<i>CviAII</i>	<i>HpyCH4V</i>	<i>RsaI</i>	<i>TaqI</i>	<i>Tru9I</i>				
<i>Clostridium</i>	<i>recN</i>	+	+	– ^b	–	–	–	–	–	–	–	–	Kekre et al. (2015)
	<i>dnaJ</i>	+	–	–	–	–	–	–	–	–	–	–	
	<i>secA</i>	+	–	–	–	–	–	–	–	–	–	–	
	<i>musS</i>	–	+	–	–	–	–	–	–	–	–	+	
	<i>grpB</i>	–	–	–	–	–	–	–	–	–	–	+	
	<i>dnaA</i>	+	–	–	–	–	–	–	–	–	–	–	
<i>Lactobacillus</i>	<i>dnaJ</i>	+	–	+	–	–	–	–	–	–	–	–	Koul and Kalia (2016)
	<i>gyrB</i>	+	–	–	–	–	–	–	–	–	–	–	
	<i>pusA</i>	+	–	–	–	–	–	–	–	–	+	–	
	<i>recA</i>	–	–	+	–	–	–	–	–	–	–	–	
	<i>ruvB</i>	+	–	+	–	–	–	–	–	–	+	–	
	<i>argH</i>	–	–	+	–	–	–	–	–	–	–	–	
	<i>argR</i>	–	–	–	–	–	–	–	–	–	–	+	
	<i>cysS</i>	+	–	–	–	–	–	–	–	–	–	–	
	<i>gyrB</i>	–	–	–	–	–	–	–	–	–	–	–	
	<i>purH</i>	–	–	+	–	–	–	–	–	–	–	–	
<i>Staphylococcus</i>	<i>pyrE</i>	–	–	–	–	–	–	–	–	–	–	–	Kumar et al. (2016)
	<i>recA</i>	+	–	–	+	–	–	–	–	–	–	–	
	<i>dnaA</i>	–	–	–	–	–	–	–	–	–	–	–	
	<i>dnaK</i>	–	–	–	–	–	–	–	–	–	–	+	
	<i>fabG</i>	–	–	–	–	–	–	–	–	–	–	+	
	<i>mraY</i>	+	–	–	–	–	–	–	–	–	–	–	
<i>Streptococcus</i>	<i>purH</i>	–	–	+	–	–	–	–	–	–	–	–	Kalia et al. (2016)
	<i>purK</i>	+	–	–	–	–	–	–	–	–	–	–	
	<i>pyrH</i>	+	–	–	–	–	–	–	–	–	–	–	
	<i>purK</i>	+	+	–	–	–	–	–	–	–	–	–	
	<i>pyrH</i>	+	–	–	–	–	–	–	–	–	–	–	
	<i>pyrH</i>	+	–	–	–	–	–	–	–	–	–	–	

(continued)

Table 4 (continued)

Organisms	Genes	Restriction endonucleases										References
		<i>AluI</i>	<i>BfaI</i>	<i>BfuCI</i>	<i>CviAII</i>	<i>HpyCH4V</i>	<i>RsaI</i>	<i>TaqI</i>	<i>Tru9I</i>			
<i>Vibrio</i>	<i>dapF</i>	-	-	+	+	+	-	+	-	-	-	Kalia et al. (2015)
	<i>hisD</i>	+	-	+	-	+	+	+	-	-	-	
	<i>ihvH</i>	+	-	-	-	+	-	+	-	-	-	
	<i>lpxC</i>	+	-	-	-	-	+	+	+	+	+	
	<i>recF</i>	+	-	-	-	+	+	+	+	-	-	
	<i>recR</i>	+	-	-	-	+	+	+	+	-	-	
	<i>rph</i>	-	-	+	+	+	-	-	-	-	+	
	<i>ravB</i>	-	-	+	+	+	-	+	-	-	-	
	<i>aceE</i>	-	-	+	-	-	-	-	-	-	-	
	<i>cysI</i>	-	-	-	-	+	-	-	-	-	-	
	<i>fadI</i>	+	-	+	+	+	-	-	-	-	-	
<i>Yersinia</i>	<i>gltB</i>	+	-	+	-	-	+	-	-	-	-	Kalia and Kumar (2015)
	<i>gyrB</i>	+	-	-	-	-	-	-	-	-		
	<i>leuD</i>	-	-	-	-	+	-	-	-	+		
	<i>ligA</i>	+	-	-	-	-	-	-	-	-		
	<i>mukB</i>	+	-	+	+	+	-	+	-	-		
	<i>rpoB</i>	-	-	-	-	-	-	-	-	+		
	<i>secA</i>	+	-	+	-	-	-	-	-	+		

^aUnique RE digestion pattern^bNot applicable

Table 5 Unique restriction endonuclease digestion patterns in genes common to different genera^a

Genus	Genes	Restriction endonuclease	References
		<i>AluI</i>	
<i>Clostridium</i>	<i>dnaI</i>	74•33•26•190•491•175•47	Kekre et al. (2015)
<i>Lactobacillus</i>	<i>dnaI</i>	84•268•66•330•109•287•20	Koul and Kalia (2016)
<i>Lactobacillus</i>	<i>gyrB</i>	1041•391•280•238	Koul and Kalia (2016)
<i>Yersinia</i>	<i>gyrB</i>	884•266•12•775•216•262	Kalia and Kumar (2015)
<i>Staphylococcus</i>	<i>gyrB</i>	–	Kumar et al. (2016)
		<i>TaqI</i>	
<i>Lactobacillus</i>	<i>ruvB</i>	344•271•184•171•50•27	Koul and Kalia (2016)
<i>Vibrio</i>	<i>ruvB</i>	19•121•218•117•30•253•319	Kalia and Kumar (2015)

^aThese patterns are representative of the genus.

Table 6 Unique restriction endonuclease digestion patterns in genes common to different genera^a

Genus	Genes	Restriction endonucleases		
		<i>CviAII</i>	<i>BfuCI</i>	<i>HpyCH4V</i>
<i>Staphylococcus</i>	<i>purH</i>	208•207•33•222•98•8•44•477•63•72•47	234•284•784•108•78	+
<i>Streptococcus</i>	<i>purH</i>	15•211•36•21•144•413•8•448•267	117•318•438•261•414	–
<i>Lactobacillus</i>	<i>recA</i>	299•232•232•231•80•24	+	296•242•228•202•52•33
<i>Staphylococcus</i>	<i>recA</i>	133•57•447•47•204•445•32	–	145•237•153•509

References: Kalia et al. (2016), Koul and Kalia (2016), Kumar et al. (2016)

^aThese patterns are representative of the genus

(*CviAII* and *BfuCI*) and *recA* gene in *Lactobacillus* and *Staphylococcus* with REs—*CviAII* and *HpyCH4V* (Table 6).(3) The *ruvB* gene–*TaqI* combinations in *Lactobacillus* and *Vibrio* species (Table 5).

7 Opinion

Very closely related species and strains possess genes which differ only in a few nucleotides. Efforts to exploit these differences on a global scale don't provide any clues for their being distinct. The search for nucleotide signatures unique to a species along with a unique RE digestion pattern allowed clear-cut distinction to a large extent. However, this approach is likely to fail in case of organisms possessing multiple copies of *rrs*. In either scenario where *rrs* is not able to distinguish organisms, an obvious approach is to go in for exploiting other HKGs. In fact, most studies which used HKGs were successful only on employing information from six to eight genes. Here again, the selection of genes may be limited by the fact that the same genes may not be present in all the strains or

species of a genus. Thus, the need is to bank upon genes which are present in most if not all the species of a genus. The approach used in these works relied primarily on genes common to all the sequence genomes. And in case two strains seem quite close in their nucleotide sequences, an additional gene–RE combination can be searched using in silico approach and primers can be developed accordingly. This approach has the potential to be extended to other genes of an organism.

Acknowledgements We are thankful to the Director of CSIR-Institute of Genomics and Integrative Biology (IGIB) and CSIR project GENESIS (BSC0121) for providing the necessary funds, facilities and moral support. Authors are also thankful to Academy of Scientific and Innovative Research (AcSIR), New Delhi.

References

- Abdeldaim G, Herrmann B, Mölling P, Holmberg H, Blomberg J, Olcén P, Strålin K (2010) Usefulness of real-time PCR for *lytA*, *ply*, and *Spn9802* on plasma samples for the diagnosis of pneumococcal pneumonia. *Clin Microbiol Infect* 16:1135–1141. doi:10.1111/j.1469-0691.2009.03069.x
- Agarwala M, Choudhury B, Yadav RNS (2014) Comparative study of antibiofilm activity of copper oxide and iron oxide nanoparticles against multidrug resistant biofilm forming uropathogens. *Indian J Microbiol* 54:365–368. doi:10.1007/s12088-014-0462-z
- Alipiah NM, Shamsudin MN, Yusoff FM, Arshad A (2015) Membrane biosynthesis gene disruption in methicillin-resistant *Staphylococcus aureus* (MRSA) as potential mechanism for reducing antibiotic resistance. *Indian J Microbiol* 54:41–49. doi:10.1007/s12088-014-0488-2
- Arasu MV, Al-Dhabi NA, Rejiniemon TS, Lee KD, Huxley VAJ, Kim DH, Duraipandiyani V, Karuppiah P, Choi KC (2015) Identification and characterization of *Lactobacillus brevis* P68 with antifungal, antioxidant and probiotic functional properties. *Indian J Microbiol* 55:19–28. doi:10.1007/s12088-014-0495-3
- Arya R, Princy SA (2016) Exploration of modulated genetic circuits governing virulence determinants in *Staphylococcus aureus*. *Indian J Microbiol* 56:19–27. doi:10.1007/s12088-015-0555-3
- Ayyadurai S, Flaudrops C, Raoult D, Drancourt M (2010) Rapid identification and typing of *Yersinia pestis* and other *Yersinia* species by matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometry. *BMC Microbiol* 10:285. doi:10.1186/1471-2180/10/285
- Bai S, Zhao J, Zhang Y, Huang W, Xu S, Chen H, Fan L-M, Chen Y, Deng XW (2010) Rapid and reliable detection of 11 foodborne pathogens using thin-film biosensor chips. *Appl Microbiol Biotechnol* 86:983–990. doi:10.1007/s00253-009-2417-6
- Bhagat N, Virdi JS (2007) Distribution of virulence associated genes in *Yersinia enterocolitica* biovar 1A correlates with clonal groups and not the source of isolation. *FEMS Microbiol Lett* 266:177–183. doi:10.1111/j.1574-6968.2006.00524.x
- Bhattacharyya N, Hou A (2013) A pentaplex PCR assay for detection and characterization of *Vibrio vulnificus* and *Vibrio parahaemolyticus* isolates. *Lett Appl Microbiol* 57:233–240. doi:10.1111/lam.12101
- Bhushan A, Joshi J, Shankar P, Kushwah J, Raju SC, Purohit HJ, Kalia VC (2013) Development of genomic tools for the identification of certain *Pseudomonas* up to species level. *Indian J Microbiol* 53:253–263. doi:10.1007/s12088-013-0412-1

- Bhushan A, Mukherjee T, Joshi J, Shankar P, Kalia VC (2015) Insights into the origin of *Clostridium botulinum* strains: evolution of distinct restriction endonuclease sites in *rrs* (16S rRNA gene). *Indian J Microbiol* 55:140–150. doi:10.1007/s12088-015-0514-z
- Bishop CJ, Aanensen DM, Jordan GE, Kilian M, Hanage WP, Spratt BG (2009) Assigning strains to bacterial species via the Internet. *BMC Biol* 7:3. doi:10.1186/1741-7007-7-3
- Brown DF, Edwards DI, Hawkey PM, Morrison D, Ridgway GL, Towner KJ, Wren MW, Joint Working Party of the British Society for Antimicrobial Chemotherapy, Hospital Infection Society, Infection Control Nurses Association (2005) Guidelines for the laboratory diagnosis and susceptibility testing of methicillin-resistant *Staphylococcus aureus* (MRSA). *J Antimicrob Chemother* 56:1000–1018. doi:10.1093/jac/dki372
- Carvalho MG, Tondella ML, McCaustland K, Weidlich L, McGee L, Mayer LW, Steigerwalt A, Whaley M, Facklam RR, Fields B, Carlone G, Ades GW, Dagan R, Sampson JS (2007) Evaluation and improvement of real-time PCR assays targeting *lytA*, *ply*, and *psaA* genes for detection of pneumococcal DNA. *J Clin Microbiol* 45:2460–2466. doi:10.1128/JCM.02498-06
- Gautam N, Sharma N, Ahlawat OP (2014) Purification and characterization of bacteriocin produced by *Lactobacillus brevis* UN isolated from Dhulliachar: a traditional food product of North East India. *Indian J Microbiol* 54:185–189. doi:10.1007/s12088-013-0427-7
- Gui Z, Wang H, Ding T, Zhu W, Zhuang X, Chu W (2014) Azithromycin reduces the production of α -hemolysin and biofilm formation in *Staphylococcus aureus*. *Indian J Microbiol* 54:114–117. doi:10.1007/s12088-013-0438-4
- Hema M, Balasubramanian S, Princy SA (2015) Meddling *Vibrio cholerae* murmurs: a neoteric advancement in cholera research. *Indian J Microbiol* 55:121–130. doi:10.1007/s12088-015-0520-1
- Illegheems K, De Vuyst L, Weckx S (2015) Comparative genome analysis of the candidate functional starter culture strains *Lactobacillus fermentum* 222 and *Lactobacillus plantarum* 80 for controlled cocoa bean fermentation processes. *BMC Genomics* 16:766. doi:10.1186/s12864-015-1927-0
- Kalia VC (2010) Extending genomic limits through metagenomic exploration. *J Cosmol* 13:3625–3627
- Kalia VC (2013a) Quorum sensing inhibitors: an overview. *Biotechnol Adv* 31:224–245. doi:10.1016/j.biotechadv.2012.10.004
- Kalia VC (2013b) The visionary: Prof. Carl R. Woese. *Indian J Microbiol* 53:245–246. doi:10.1007/s12088-013-0417-9
- Kalia VC (2014) Microbes, antimicrobials and resistance: the battle goes on. *Indian J Microbiol* 54:1–2. doi:10.1007/s12088-013-0443-7
- Kalia VC (2015a) Let's explore the latent features of genes to identify bacteria. *J Mol Genet Med* 9, e105. doi:10.4172/1747-0862.1000E105
- Kalia VC (2015b) Microbes: the most friendly beings? In: Kalia VC (ed) Quorum sensing vs quorum quenching: a battle with no end in sight. Springer India, pp 1–5. ISBN 978-81-322-1981-1. doi:10.1007/978-81-322-1982-8_1
- Kalia VC, Kumar P (2015) Genome wide search for biomarkers to diagnose *Yersinia* infections. *Indian J Microbiol* 55:366–374. doi:10.1007/s12088-015-0552-6
- Kalia VC, Purohit HJ (2011) Quenching the quorum sensing system: potential antibacterial drug targets. *Crit Rev Microbiol* 37:121–140. doi:10.3109/1040841X.2010.532479
- Kalia VC, Mukherjee T, Bhushan A, Joshi J, Shankar P, Huma N (2011a) Analysis of the unexplored features of *rrs* (16S rDNA) of the genus *Clostridium*. *BMC Genomics* 12:18. doi:10.1186/1471-2164-12-18
- Kalia VC, Raju SC, Purohit HJ (2011b) Genomic analysis reveals versatile organisms for quorum quenching enzymes: acyl-homoserine lactone-acylase and -lactonase. *Open Microbiol J* 5:1–13. doi:10.2174/187428580110501000
- Kalia VC, Kumar P, Kumar R, Mishra A, Koul S (2015) Genome wide analysis for rapid identification of *Vibrio* species. *Indian J Microbiol* 55:375–383. doi:10.1007/s12088-015-0553-5

- Kalia VC, Kumar R, Kumar P, Koul S (2016) A genome-wide profiling strategy as an aid for searching unique identification biomarkers for *Streptococcus*. *Indian J Microbiol* 56:46–58. doi:[10.1007/s12088-015-0561-5](https://doi.org/10.1007/s12088-015-0561-5)
- Kaur G, Rajesh S, Princy SA (2015) Plausible drug targets in the *Streptococcus mutans* quorum sensing pathways to combat dental biofilms and associated risks. *Indian J Microbiol* 55:349–357. doi:[10.1007/s12088-015-0534-8](https://doi.org/10.1007/s12088-015-0534-8)
- Kekre A, Bhushan A, Kumar P, Kalia VC (2015) Genome wide analysis for searching novel markers to rapidly identify *Clostridium* strains. *Indian J Microbiol* 55:250–257. doi:[10.1007/s12088-015-0535-7](https://doi.org/10.1007/s12088-015-0535-7)
- Kingston JJ, Radhika M, Roshini PT, Raksha MA, Murali HS, Batra HV (2010) Molecular characterization of lactic acid bacteria recovered from natural fermentation of beet root and carrot Kanji. *Indian J Microbiol* 50:292–298. doi:[10.1007/s12088-010-0022-0](https://doi.org/10.1007/s12088-010-0022-0)
- Kishore P, Lalitha KV, Joseph TC, Thampuran N (2012) Biotyping and antibiotic resistance profile of *Yersinia enterocolitica* associated with seafoods from South-west coast of India. *Fish Technol* 49:64–71
- Klappenbach JA, Saxman PR, Cole JR, Schmidt TM (2001) rrndb: the ribosomal RNA operon copy number database. *Nucleic Acids Res* 29:181–184. doi:[10.1093/nar/29.1.181](https://doi.org/10.1093/nar/29.1.181)
- Koul S, Kalia VC (2016) Comparative genomics reveals biomarkers to identify *Lactobacillus* species. *Indian J Microbiol* 56:253–263. doi:[10.1007/s12088-016-0605-5](https://doi.org/10.1007/s12088-016-0605-5)
- Koul S, Kumar P, Kalia VC (2015a) A unique genome wide approach to search novel markers for rapid identification of bacterial pathogens. *J Mol Genet Med* 9:194. doi:[10.4172/1747-0862.1000194](https://doi.org/10.4172/1747-0862.1000194)
- Koul S, Prakash J, Mishra A, Kalia VC (2015b) Potential emergence of multi-quorum sensing inhibitor resistant (MQSIR) bacteria. *Indian J Microbiol*. doi:[10.1007/s12088-015-0558-0](https://doi.org/10.1007/s12088-015-0558-0)
- Kuhm AE, Suter D, Felleisen R, Rau J (2009) Identification of *Yersinia enterocolitica* at the species and subspecies levels by Fourier transform infrared spectroscopy. *Appl Environ Microbiol* 75:5809–5813. doi:[10.1128/AEM.00206-09](https://doi.org/10.1128/AEM.00206-09)
- Kumar R, Koul S, Kumar P, Kalia VC (2016) Searching biomarkers in the sequenced genomes of *Staphylococcus* for their rapid identification. *Indian J Microbiol* 56:64–71. doi:[10.1007/s12088-016-0565-9](https://doi.org/10.1007/s12088-016-0565-9)
- Lal D, Verma M, Lal R (2011) Exploring internal features of 16S rRNA gene for identification of clinically relevant species of the genus *Streptococcus*. *Ann Clin Microbiol Antimicrob* 10:28. doi:[10.1186/1476-0711-10-28](https://doi.org/10.1186/1476-0711-10-28)
- Liu Q, Wang S, Zhi J-F, Ming H, Teng D (2013) Efficient production of lactic acid from sweet sorghum juice by a newly isolated *Lactobacillus salivarius* CGMCC 7.75. *Indian J Microbiol* 53:332–336. doi:[10.1007/s12088-013-0377-0](https://doi.org/10.1007/s12088-013-0377-0)
- Mahale KN, Paranjape PS, Marathe NP, Dhotre DP, Chowdhury S, Shetty SA, Sharma A, Sharma K, Tuteja U, Batra HV, Shouche YS (2014) Draft genome sequences of *Yersinia pestis* strains from the 1994 plague epidemic of Surat and 2002 Shimla outbreak in India. *Indian J Microbiol* 54:480–482. doi:[10.1007/s12088-014-0475-7](https://doi.org/10.1007/s12088-014-0475-7)
- Meza-Lucas A, Pérez-Villagómez M, Martínez-López JP, García-Rodea R, Martínez-Castelán MG, Escobar-Gutiérrez A, de-la-Rosa-Arana JL, Villanueva-Zamudio A (2016) Comparison of DOT-ELISA and Standard-ELISA for detection of the *Vibrio cholerae* toxin in culture supernatants of bacteria isolated from human and environmental samples. *Indian J Microbiol* 56:379–382. doi:[10.1007/s12088-016-0596-2](https://doi.org/10.1007/s12088-016-0596-2)
- Moroeanu VI, Vamanu E, Paun G, Neagu E, Ungureanu OR, Eremia SAV, Radu GL, Ionescu R, Pelinescu DR (2015) Probiotic strains influence on infant microbiota in the in vitro colonic fermentation model GIS1. *Indian J Microbiol* 55:423–429. doi:[10.1007/s12088-015-0542-8](https://doi.org/10.1007/s12088-015-0542-8)
- Naser SM, Dawyndt P, Hoste B, Gevers D, Vandemeulebroecke K, Cleenwerck I, Vancanneyt M, Swings J (2007) Identification of lactobacilli by *pheS* and *rpoA* gene sequence analysis. *Int J Syst Evol Microbiol* 57:2777–2789. doi:[10.1099/ij.s.0.64711-0](https://doi.org/10.1099/ij.s.0.64711-0)
- Nguyen THK, Doan VTT, Ha LD, Nguyen HN (2013) Molecular cloning, expression of *minD* gene from *Lactobacillus acidophilus* VTCC-B-871 and analyses to identify *Lactobacillus*

- rhamnosus* PN04 from Vietnam *Hottuyntia cordata* Thunb. Indian J Microbiol 53:385–390. doi:[10.1007/s12088-013-0384-1](https://doi.org/10.1007/s12088-013-0384-1)
- Petrova MI, Lievens E, Malik S, Imholz N, Lebeer S (2015) *Lactobacillus* species as biomarkers and agents that can promote various aspects of vaginal health. Front Physiol 6:81. doi:[10.3389/fphys.2015.00081](https://doi.org/10.3389/fphys.2015.00081)
- Pichon B, Hill R, Laurent F, Larsen AR, Skov RL, Holmes M, Edwards GF, Teale C, Kearns AM (2012) Development of a real-time quadruplex PCR assay for simultaneous detection of *nuc*, Panton–Valentine leucocidin (PVL), *mecA* and homologue *mecA*_{LGA251}. J Antimicrob Chemother 67:2338–2341. doi:[10.1093/jac/dks221](https://doi.org/10.1093/jac/dks221)
- Plaon S, Longyant S, Sithigorngul P, Chaivisuthangkura P (2015) Rapid and sensitive detection of *Vibrio alginolyticus* by loop mediated isothermal amplification combined with a lateral flow dipstick targeted to the *rpoX* gene. J Aquat Anim Health 27:156–163. doi:[10.1080/08997659.2015.1037468](https://doi.org/10.1080/08997659.2015.1037468)
- Porwal S, Lal S, Cheema S, Kalia VC (2009) Phylogeny in aid of the present and novel microbial lineages: diversity in *Bacillus*. PLoS One 4, e4438. doi:[10.1371/journal.pone.0004438](https://doi.org/10.1371/journal.pone.0004438)
- Prakash O, Jangid K, Shouche YS (2013) Carl Woese: from biophysics to evolutionary microbiology. Indian J Microbiol 53:247–252. doi:[10.1007/s12088-013-0401-4](https://doi.org/10.1007/s12088-013-0401-4)
- Prakash O, Pandey PK, Kulkarni GJ, Mahale KN, Shouche YS (2014) Technicalities and glitches of terminal restriction fragment length polymorphism (T-RFLP). Indian J Microbiol 54:255–261. doi:[10.1007/s12088-014-0461-0](https://doi.org/10.1007/s12088-014-0461-0)
- Prakasham RS, Kumar BS, Kumar YS, Kumar KP (2014) Production and characterization of protein encapsulated silver nanoparticles by marine isolate *Streptomyces parvulus* SSNP11. Indian J Microbiol 54:329–336. doi:[10.1007/s12088-014-0452-1](https://doi.org/10.1007/s12088-014-0452-1)
- Puri A, Rai A, Dhanaraj PS, Lal R, Patel DD, Kaicker A, Verma M (2016) An *in silico* approach for identification of the pathogenic species, *Helicobacter pylori* and its relatives. Indian J Microbiol 56:277–286. doi:[10.1007/s12088-016-0575-7](https://doi.org/10.1007/s12088-016-0575-7)
- Raftis EJ, Forde BM, Claesson MJ, O'Toole PW (2014) Unusual genome complexity in *Lactobacillus salivarius* JCM1046. BMC Genomics 15:771. doi:[10.1186/1471-2164-15-771](https://doi.org/10.1186/1471-2164-15-771)
- Roberts AL (2014) Identification of *Staphylococcus epidermidis* in the clinical microbiology laboratory by molecular methods. Methods Mol Biol 1106:33–53. doi:[10.1007/978-1-62703-736-5_3](https://doi.org/10.1007/978-1-62703-736-5_3)
- Sajid A, Arora G, Singhal A, Kalia VC, Singh Y (2015) Protein phosphatases of pathogenic bacteria: role in physiology and virulence. Annu Rev Microbiol 69:527–547. doi:[10.1146/annurev-micro-020415-111342](https://doi.org/10.1146/annurev-micro-020415-111342)
- Sarmiento-Rubiano LA, Berger B, Moine D, Zúñiga M, Pérez-Martínez G, Yebra MJ (2010) Characterization of a novel *Lactobacillus* species closely related to *Lactobacillus johnsonii* using a combination of molecular and comparative genomics methods. BMC Genomics 11:504. doi:[10.1186/1471-2164-11-504](https://doi.org/10.1186/1471-2164-11-504)
- Saxena A, Mukherjee M, Kumari R, Singh P, Lal R (2014) Synthetic biology in action: developing a drug against MDR-TB. Indian J Microbiol 54:369–375. doi:[10.1007/s12088-014-0498-0](https://doi.org/10.1007/s12088-014-0498-0)
- Shang Z, Wang H, Zhou S, Chu W (2014) Characterization of N-acyl-homoserine lactones (AHLs)-deficient clinical isolates of *Pseudomonas aeruginosa*. Indian J Microbiol 54:158–162. doi:[10.1007/s12088-014-0449-9](https://doi.org/10.1007/s12088-014-0449-9)
- Stephan R, Cernela N, Ziegler D, Pflueger V, Tonolla M, Ravasi D, Fredriksson-Ahomaa M, Haechler H (2011) Rapid species specific identification and subtyping of *Yersinia enterocolitica* by MALDI-TOF mass spectrometry. J Microbiol Methods 87:150–153. doi:[10.1016/j.mimet.2011.08.016](https://doi.org/10.1016/j.mimet.2011.08.016)
- Surasilp T, Longyant S, Rukpratanporn S, Sridulyakul P, Sithigorngul P, Chaivisuthangkura P (2011) Rapid and sensitive detection of *Vibrio vulnificus* by loop-mediated isothermal amplification combined with lateral flow dipstick targeted to *rpoS* gene. Mol Cell Probes 25:158–163. doi:[10.1016/j.mcp.2011.04.001](https://doi.org/10.1016/j.mcp.2011.04.001)

- Taneja N, Sangar G, Chowdhury G, Ramamurthy T, Mishra A, Singh M, Sharma M (2012) Molecular epidemiology of *Vibrio cholerae* causing outbreaks & sporadic cholera in northern India. *Indian J Med Res* 136:656–663
- Tarr CL, Patel JS, Puhr ND, Sowers EG, Bopp CA, Strockbine NA (2007) Identification of *Vibrio* isolates by a multiplex PCR assay and *rpoB* sequence determination. *J Clin Microbiol* 45:134–140. doi:[10.1128/JCM.01544-06](https://doi.org/10.1128/JCM.01544-06)
- Thongkao K, Longyant S, Silprasit K, Sithigorngul P, Chaivisuthangkura P (2015) Rapid and sensitive detection of *Vibrio harveyi* by loop-mediated isothermal amplification combined with lateral flow dipstick targeted to vhhP2 gene. *Aquac Res* 46:1122–1131. doi:[10.1111/are.12266](https://doi.org/10.1111/are.12266)
- Wang R, Fang S, Xiang S, Ling S, Yuan J, Wang S (2014) Generation and characterization of a scFv antibody against T3SS needle of *Vibrio parahaemolyticus*. *Indian J Microbiol* 54:143–150. doi:[10.1007/s12088-013-0428-6](https://doi.org/10.1007/s12088-013-0428-6)
- Yagnik B, Patel S, Dave M, Sharma D, Padh H, Desai P (2016) Factors affecting inducible expression of outer membrane protein A (OmpA) of *Shigella dysenteriae* Type-1 in *Lactococcus lactis* using nisin inducible controlled expression (NICE). *Indian J Microbiol* 56:80–87. doi:[10.1007/s12088-015-0556-2](https://doi.org/10.1007/s12088-015-0556-2)
- Yu S, Peng Y, Chen W, Deng Y, Zheng Y (2014) Comparative genomic analysis of two-component signal transduction systems in probiotic *Lactobacillus casei*. *Indian J Microbiol* 54:293–301. doi:[10.1007/s12088-014-0456-x](https://doi.org/10.1007/s12088-014-0456-x)
- Yu S, Peng Y, Zheng Y, Chen W (2015) Comparative genome analysis of *Lactobacillus casei*: insights into genomic diversification for niche expansion. *Indian J Microbiol* 55:102–107. doi:[10.1007/s12088-014-0496-2](https://doi.org/10.1007/s12088-014-0496-2)