Indian J Microbiol (Jan–Mar 2016) 56(1):64–71
DOI 10.1007/s12088-016-0565-9

ORIGINAL ARTICLE

Searching Biomarkers in the Sequenced Genomes of *Staphylococcus* for their Rapid Identification

Ravi Kumar[1] · Shikha Koul[1,2] · Prasun Kumar[1] · Vipin Chandra Kalia[1,2]

Received: 31 December 2015 / Accepted: 11 January 2016 / Published online: 20 January 2016

Abstract Bacterial identification using *rrs* (16S rRNA) gene is widely reported. Bacteria possessing multiple copies of *rrs* lead to overestimation of its diversity. *Staphylococcus* genomes carries 5–6 copies of *rrs* showing high similarity in their nucleotide sequences, which lead to ambiguous results. The genomes of 31 strains of *Staphylococcus* representing 7 species were searched for the presence of common genes. In silico digestion of 34 common genes using 10 restriction endonucleases (REs) lead to select gene-RE combinations, which could be used as biomarkers. RE digestion of *recA* allowed unambiguous identification of 13 genomes representing all the 7 species. In addition, a few more genes *(argH, argR, cysS, gyrB, purH,* and *pyrE*) and RE combinations permitted further identification of 12 strains. By employing additional RE and genes unique to a particular strain, it was possible to identify the rest 6 *Staphylococcus aureus* strains. This approach has the potential to be utilized for rapid detection of *Staphylococcus* strains.

Keywords Biomarkers · Diagnosis · Genome · In silico · Restriction endonuclease · *Staphylococcus*

Electronic supplementary material The online version of this article (doi:10.1007/s12088-016-0565-9) contains supplementary material, which is available to authorized users.

✉ Vipin Chandra Kalia
vckalia@igib.res.in; vc_kalia@yahoo.co.in

[1] Microbial Biotechnology and Genomics, CSIR - Institute of Genomics and Integrative Biology (IGIB), Delhi University Campus, Mall Road, New Delhi 110007, India

[2] Academy for Scientific and Innovative Research (AcSIR), 2 Rafi Marg, New Delhi 110001, India

Introduction

Staphylococcus species are Gram-positive pathogens, which colonize and infect the skin and mucous membranes of human beings. These infections in humans are often associated with exposure to livestock [1]. Bacteria demonstrate a unique ability to rapidly acquire genetic material, which increases its pathogenicity and confers resistance to antibiotics. *Staphylococcus aureus* has evolved as an organism responsible for epidemics, which are difficult to control. *Staphylococcus epidermidis* causes a wide range of nosocomial infections, which include nosocomial bloodstream, eye, ear, nose, throat and cardiovascular system [2]. In fact, *S. aureus* has developed the competence to withstand the threats posed by the human immune system [3]. *Staphylococcus aureus* cause infections in open wounds through mucosal surfaces or skin [3, 4]. *Staphylococcus* is also reported to cause abscesses, bacteremia endocarditis, gastroenteritis, food intoxications and septicemia [5]. Children and diabetic patients with HIV are highly susceptible to colonization by *S. aureus* [3]. The most effective mechanism by which *S. aureus* expresses its virulence is regulated by a quorum sensing (QS) mediated accessory gene regulator system [6]. QS regulated biofilm formation is considered as the main cause of infections in this organism. These QS systems have been rigorously studied as potential therapeutic targets [7–12].

Bacterial Identification

Biochemical Tests

Reference methods for identification of *Staphylococcus* species include: (1) enzyme assays—alkaline phosphatase,

coagulase, amino acid decarboxylases, urease, (2) nitrate reduction, and acid production from a wide range of sugars, and (3) hemolysis [5, 13]. A few other ancillary tests to identify *Staphylococcus* include anaerobic utilization of glucose and mannitol, lysostaphin sensitivity, and thermostable nuclease production. *S. epidermidis* can be distinguished from *S. aureus*, as it lacks abilities such as coagulase production, thermonuclease production, and mannitol utilization [14].

Molecular Assays

Guidelines for the laboratory level detection of *S. aureus* recommend the molecular targets such as *coa*, *femA, femB*, *gyrA*, *nuc*, *spa*, *Sa442*, and 16S rRNA (*rrs*) genes [15]. The triplex qPCR assay based on *mecA*, and *femA* (from *S. aureus* and *S. epidermidis*) could be completed within a short span of 6 h [16]. Autolysin encoded by *atlE* which influences primary attachment, is one of the most studied genes [2]. Real-time PCR for amplifying *altE*, *lukS*-PV, *mecA,* homologue $mecA_{LGA251}$, *mvaA, nuc,* and *tuf* genes from blood allows rapid identification of *S. aureus* strains within 2–3 h [17–20]. PCR-amplification of the *nuc, mecA, ileS, lukS-lukF* genes are used to identify *S. aureus* [21]. Here, *S. aureus* ATCC25923 act as a positive control, whereas *S. epidermidis* ATCC12228 is used as a negative control [5]. Identification of genetically diverse isolates of methicillin-resistant *S. aureus* (MRSA) has been successfully done using *orfX*-staphylococcal cassette chromosome *mec*- (SCC*mec*) based assays carrying *blaZ*, *ccr, mecA*, *mecI*, and *mecR1* genes [1, 22–25]. A few more genes used for identifying *S. epidermidis* include: *sodA*, *rpoB*, *tuf*, *gap*, *dnaJ*, *hsp40*, and tRNA intergenic spacer [20]. The MRSA PCR detection kit (Multiplex) targets the following 4 genes: *rrs*, *mecA*, *lukPVL* and *femA* (http://himedialabs.com/TD/MBPCR020.pdf). Although commercially available kits are effective in identifying the *mecA* gene, however, these methods need pure cultures [15]. In spite of the availability and usage of a large number of genes for identifying *Staphylococcus*, there seems to be no consensus so far.

The most widely used gene for identifying bacteria is *rrs.* Recent works have proved helpful in further enhancing their value by revealing their unique latent features [26–29]. However, the major limitation in the use of *rrs* gene is encountered in bacteria having multiple copies, which is responsible for overestimation of bacterial species. Secondly, the multiple copies of *rrs* show very high similarity with those of other species as well [30–35]. We need to resort to other conserved genes for better bacterial identification. Recent works have used a set of genes which are common to all the species of a genus. These genes were digested in silico with different Restriction Endonucleases (REs). Unique RE digestion patterns obtained with a specific gene were shown to be potentially useful for rapid bacterial identification. Since *Staphylococcus* genomes have multiple copies of *rrs*, we have adopted the strategies proposed earlier for searching novel markers in pathogenic bacteria [30–32, 34], for identifying this bacteria as well.

Materials and Methods

Comparative Analysis of Sequenced Genomes

Sequenced genomes of the *Staphylococcus*: 7 species—31 strains were used here (http://www.ncbi.nlm.nih.gov/): *S. aureus* (24 strains), *S. carnosus, S. epidermidis* (2 strains), *S. haemolyticus, S. lugdunensis, S. pseudintermedius,* and *S. saprophyticus* (Table S1). Certain features of these *Staphylococcus* genomes have been presented in Table S1. Comparative analysis of *Staphylococcus* genomes, allowed us to select 53 genes common to all of them. These common genes, varied from 179 to 4316 nucleotides (nts) (Tables S1 and S2). In addition, *rrs* was also used in this analysis. Orientation (5′–3′) of sequences was checked with the help of BioEdit [36].

In silico Digestion of Common Genes with Restriction Endonucleases

Ten Type II REs: (1) *AluI*, *BfaI*, *BfuCI*, *CviAII*, *HpyCH4 V*, *RsaI*, *TaqI*, *Tru9I* (4 base cutters) and (2) *HaeI* and *Hin1I* (6 base cutters) were used for in silico digestion of common genes [32]. RE digestion patterns of these genes were obtained through Cleaver (http://cleaver.sourceforge.net/) (Table S2). REs which resulted in 5–15 fragments were employed for comparative analysis of the gene sequences [32]. A genome wide search was performed in the following genomes—*S. aureus* MW2, *S. aureus* MSSA476, *S. aureus* Mu3 and *S. aureus* N315. Genes that were not common to all the 31 genomes were treated as unique.

Results

In silico RE Digestion of *rrs*

In each of 31 genomes of *Staphylococcus* strains there were 5–6 copies of *rrs*. The 162 copies of *rrs* could be segregated into 11 groups on the basis of multiple sequence alignment (MSA). Each group was represented by 5–68 copies, which were quite similar to each other. Twenty five genomes were represented by 5 groups, with 6–68 copies in each. It allowed unambiguous segregation of only 6 genomes representing 32 copies: *S. aureus* subsp. *aureus* T0131, *S. aureus* subsp.

aureus VC40, *S. epidermidis* ATCC12228, *S. epidermidis* RP62A, *S. lugdunensis* N920143 and *S. saprophyticus* subsp. *saprophyticus* ATCC15305.

Unique in silico RE digestion patterns were observed in the following strains: (1) *S. aureus*—5 strains, (2) *S. carnosus,* (3) *S. epidermidis*—2 strains, (4) *S. haemolyticus*, (5) *S. lugdunensis*, (6) *S. pseudintermedius*, and (7) *S. saprophyticus* (Table 1). Here, unique digestion patterns in the *rrs* gene were recorded with 8 REs (Table 1). RE-*Hin1I* did not provide unique digestion pattern in any of the strains. There was a reasonable amount of variation in RE digestion patterns even within a genome: (1) RE-*AluI* and *S. carnosus* subsp. *carnosus* TM300; *S. epidermidis* ATCC12228, (2) RE-*BfaI* and *S. haemolyticus* JCSC1435, (3) RE-*BfuCI* and *S. aureus* subsp. *aureus* TW20; *S. epidermidis* ATCC12228, *S. epidermidis* RP62A (4) RE-*CviAII* and *S. aureus* subsp. *aureus* 6850; *S. aureus* subsp. *aureus* MRSA252; *S. saprophyticus* subsp. *saprophyticus* ATCC15305, (5) RE-*HpyCH4 V* and *S. aureus* RF122; *S. aureus* subsp. *aureus* MRSA252; *S. carnosus* subsp. *carnosus* TM300, (6) RE-*TaqI* and *S. lugdunensis* N920143, and (7) RE-*Tru9I* and *S. aureus* RF122 (Table 1). In *S. aureus* subsp. *aureus* TCH60 and *S. pseudintermedius* ED99, RE digestion pattern was observed to be exactly similar in all the *rrs* copies in their respective genomes. In silico digestion of 162 *rrs* gene sequences with 10 REs, was quite effective in identifying 12 *Staphylococcus* strains representing 7 species. These 12 strains included only 4 out of 6 strains (*S. epidermidis* ATCC12228, *S. epidermidis* RP62A, *S. lugdunensis* N920143 and *S. saprophyticus* subsp. *saprophyticus* ATCC15305), which could be segregated through MSA. On the basis of these analyses, we may conclude that we need to resort to other genes for identifying the rest 19 genomes having 99 *rrs* copies.

In silico RE Digestion of Common Genes

In 31 *Staphylococcus* genomes, 53 genes (in addition to *rrs*) were found to be common to all of them. The size of these 53 genes varied from 179 to 4316 nts (Table S2). In silico digestion of 34 out of 53 genes with different REs showed certain unique features, which can be used to distinguish different *Staphylococcus* genomes. Interestingly, there were 4 *S. aureus* genomes (Table 4), which did not show any unique RE digestion pattern with any of the REs. The genes which showed unique digestion patterns with a few REs were: *argH, argR, cysS, gyrB, purH, pyrE* and *recA* (Tables 2a, b, 3).

recA

Digestion of *recA* gene (1349 nts) was recorded with 8 REs. REs-*AluI*, *HpyCH4V* and *Tru9I* proved effective in generating unique RE digestion patterns with *recA* in 7–8 genomes. However, if all the combinations are taken into consideration, these gene-RE combinations can decipher 11 strains belonging to 7 *Staphylococcus* species (Table 2a). In addition, another genome of *S. aureus* strain RF122 can also be identified because of unique RE digestion patterns generated with *CviAII, TaqI* and *BfuCI* (Table 2b).

argH, argR, cysS, gyrB, purH, and pyrE

As, *recA* was helpful in being used as a marker for identifying 13 out of 31 strains, we resorted to other genes for the rest of the strains. Six more genes (*argH, argR, cysS, gyrB, purH,* and *pyrE*) in combinations with REs—*AluI*, *BfuCI*, *CviAII*, *HpyCH4 V*, *RsaI*, and *Tru9I* provided markers for identifying 12 more *S. aureus* strains (Table 3). The details of unique gene-RE combinations and the genomes so identified were as follows (Table 3): (a) *purH*: (1) *BfuCI* for *S. aureus* TCH60 and *S. aureus* T0131, (2) *CviAII* for *S. aureus* ED98; (3) *HpyCH4V* for *S. aureus* 11819-97; (b) *argH-BfuCI* for *S. aureus* 04-02981; (c) *argR-Tru9I* for *S. aureus* Newman and *S. aureus* JKD6008; (d) *gyrB-RsaI* for *S. aureus* COL and *S. aureus* JKD6159; (e) *cysS-AluI* for *S. aureus* Mu50; (f) *pyre-Tru9I* for *S. aureus* VC40; and (g) *recR-Tru9I* for *S. aureus* MRSA252.

However, among all the common genes, none of the gene-RE combination could distinguish the 6 genomes of *S. aureus*. For identifying *S. aureus* NCTC8325 and *S. aureus* USA300_FPR3757, two different gene-RE combinations were used (Table 3). Incidentally, the strains of *S. aureus* MW2 and *S. aureus* MSSA476 were found to have a similar RE pattern with *argH-BfuCI*. These two genomes could only be identified by using another gene *seh*, which is only present in *S. aureus* MW2. Gene *seh* can be digested only with *AluI, HpyCH4 V* and *Tru9I* (Table 4). On the other hand, rest two genomes of *S. aureus* Mu3 and *S. aureus* N315 were quite similar and hard to distinguish on the basis of various common genes used in the study. A genome wide search led to the identification of quite a few unique genes that may be used to distinctly identify them (Table 4).

Thus, using these approaches of common and unique genes in combination with 10 different REs, we could identify quite a few biomarkers. *recA* alone was effective in distinguishing 13 out of 31 genomes. An additional 14 genomes could be identified on the basis of specific RE digestion patterns of genes such as: *argH, argR, cysS, gyrB, purH,* and *pyrE.* Four genomes of *S. aureus* could be distinguished by taking help of genes unique to a given strain. These biomarker genes enabled us to segregate all the 31 genomes distinctly.

Table 1 In silico restriction endonuclease (RE) digestion pattern (5′-3′) of *rrs* gene of *Staphylococcus* strains

Staphylococcus spp.	GenBank ID	Copies of *rrs*	Unique RE digestion pattern[a]
RE-AluI			
S. aureus subsp. *aureus* MRSA252	BX571856	5	82•86•87•615•210•51•214•124•87
S. aureus subsp. *aureus* TW20	FN433596	5	74•86•87•824•51•214•124•80
S. carnosus subsp. *carnosus* TM300	AM295250	2/5	82•173•615•209•51•214•124•87
		3/5	82•173•615•209•51•170•44•124•87
S. epidermidis ATCC12228	AE015929	2/5	87•123•214•51•209•478•137•87•86•82
		3/5	87•123•214•51•209•615•87•86•82
S. haemolyticus JCSC1435	AP006716	5	82•86•87•615•209•265•123•87
S. pseudintermedius ED99	CP002478	5	82•173•186•429•209•265•124•87
S. saprophyticus subsp. *saprophyticus* ATCC15305	AP008934	6	81•86•87•615•209•51•138•76•125•87
RE-BfaI			
S. aureus subsp. *aureus* MRSA252	BX571856	5	256•763•336•115•86
S. aureus subsp. *aureus* TCH60	CP002110	6	248•762•335•115•79
S. epidermidis ATCC 12228	AE015929	5	84•114•335•763•258
S. haemolyticus JCSC1435	AP006716	1/5	1019•335•114•86
		4/5[b]	256•763•335•114•86
S. lugdunensis N920143	FR870271	5/5	252•763•335•113•63
S. saprophyticus subsp. *saprophyticus* ATCC15305	AP008934	6	255•763•140•195•116•86
RE-BfuCI			
S. aureus subsp. *aureus* TCH60	CP002110	6	7•223•74•119•351•582•175•8
S. aureus subsp. *aureus* TW20	FN433596	1/5	7•223•74•119•352•582•100•75•8
		4/5[b]	7•223•74•119•352•582•175•8
S. carnosus subsp. *carnosus* TM300	AM295250	5	15•416•351•582•175•15
S. epidermidis ATCC12228	AE015929	2/5	11•174•354•228•352•119•74•223•19
		3/5	11•174•582•352•119•74•223•19
S. epidermidis RP62A	CP000029	1/6	15•223•74•119•352•228•354•174•15
		5/6[b]	15•223•74•119•352•582•174•15
S. haemolyticus JCSC1435	AP006716	5	15•416•352•582•174•15
S. lugdunensis N920143	FR870271	5	15•297•119•352•582•162
S. saprophyticus subsp. *saprophyticus* ATCC15305	AP008934	6	14•297•471•582•176•15
RE-CviAII			
S. aureus subsp. *aureus* 6850	CP006706	1/5	47•140•493•269•108•148•35•90•34•176
		4/5[b]	47•140•493•269•108•148•125•34•176
S. aureus subsp. *aureus* MRSA252	BX571856	1/5	26•29•139•493•269•108•148•125•34•183
		4/5[b]	55•140•493•269•108•148•125•34•173
S. aureus subsp. *aureus* TCH60	CP002110	6	47•140•492•269•108•148•125•34•176
S. epidermidis ATCC12228	AE015929	5	180•159•148•108•269•493•140•57
S. epidermidis RP62A	CP000029	6	55•140•493•269•108•148•159•182
S. pseudintermedius ED99	CP002478	5	55•128•12•493•269•108•148•125•34•183
S. saprophyticus subsp. *saprophyticus* ATCC15305	AP008934	1/6	54•633•269•108•148•70•55•34•184
		5/6	54•140•493•269•108•148•70•55•34•184
RE-HpyCH4V			
S. aureus RF122	AJ938182	2/5	58•191•43•188•186•40•161•197•262•229
		3/5[b]	58•191•43•188•186•201•197•262•229
S. aureus subsp. *aureus* MRSA252	BX571856	1/5	248•43•188•186•201•197•262•229
		4/5[b]	58•191•43•188•186•201•197•262•229

Table 1 continued

Staphylococcus spp.	GenBank ID	Copies of *rrs*	Unique RE digestion pattern[a]
S. carnosus subsp. *carnosus* TM300	AM295250	1/5	58•128•8•10•462•88•113•197•262•229
		1/5	58•128•18•462•88•112•197•262•229
		3/5	58•128•18•462•88•113•197•262•229
S. epidermidis ATCC12228	AE015929	5	228•262•197•201•186•422•58
S. lugdunensis N920143	FR870271	5	58•422•186•201•197•262•201
S. pseudintermedius ED99	CP002478	5	58•422•186•88•113•197•262•229
S. saprophyticus subsp. *saprophyticus* ATCC15305	AP008934	6	57•422•186•201•197•222•40•230
RE-RasI			
S. carnosus subsp. *carnosus* TM300	AM295250	5	494•405•357•146•152
S. epidermidis ATCC12228	AE015929	5	151•503•406•494
RE-TaqI			
S. epidermidis ATCC12228	AE015929	5	77•144•361•907•65
S. haemolyticus JCSC1435	AP006716	5	63•124•14•769•361•144•79
S. lugdunensis N920143	FR870271	1/5	59•907•361•143•57
		4/5[b]	63•907•361•144•42
S. saprophyticus subsp. *saprophyticus* ATCC15305	AP008934	6	62•907•361•146•79
RE-Tru9I			
S. aureus RF122	AJ938182	1/5	497•104•278•86•136•25•19•410
		4/5[b]	601•278•86•136•25•19•410
S. carnosus subsp. *carnosus* TM300	AM295250	5	601•277•86•136•25•19•410
S. epidermidis ATCC12228	AE015929	5	407•19•25•136•86•278•603
S. lugdunensis N920143	FR870271	5	601•278•86•136•25•401
S. pseudintermedius ED99	CP002478	5	601•278•86•136•454

Symbol (•) indicates RE site in the gene sequences

[a] Values represent restriction fragments (nucleotides)

[b] This pattern is not unique. It has been presented to indicate the RE digestion pattern of the rest of the *rrs* copies

Discussion

Identification of pathogenic bacteria is the first prerequisite for diagnosis and treatment. Quite a few metabolic characteristics of these bacteria give a very clear cut indication of the potential organisms. However, a confirmatory statement can be made on the basis of their molecular features. Most identification methods use *rrs* gene sequences [7, 12, 26–28]. In spite of its successful usage, there are situations where this gene becomes ambiguous. Such cases involve bacteria possessing multiple copies of *rrs,* which also show high similarity among them [12, 26, 33–35]. In general, other highly conserved genes become helpful, which requires additional time and money. In most of the molecular approaches used for identifying *Staphylococcus* species, 27 genes have been widely used either individually or in combination [1, 17–20, 22–24]. Of these 27 genes, only *gyrA* is among those which are common to all the genomes. It implies that the rest 26 genes may not prove effective in all the *Staphylococcus* strains. In our study, the most important feature is the usage of genes which were common to all the *Staphylococcus* strains. And *recA* alone was effective in identifying strains which represented all the 7 species. Analysis of *gyrA* with 10 different REs showed very poor resolution, such that a very large number of fragments were generated, which are not very convenient to be used for this purpose.

For employing the biomarkers deduced in this study for identifying *Staphylococcus* isolates, the following procedure is suggested. Standard molecular procedures, including DNA extraction from the infected sample, designing primers and optimizing conditions are required for gene amplification. The amplified gene product can be digested with selected REs or the sequence can be digested in silico with designated RE. Such a strategy for searching biomarkers from genes which are common to all the species has been shown for pathogenic bacteria such as:

Table 2 Unique in silico restriction endonuclease digestion pattern (5′-3′) of *recA* gene of *Staphylococcus* genomes

Staphylococcus spp.	Digestion patterns (nucleotides) with restriction endonucleases		
	Tru9I	*HpyCH4V*	*AluI*
a			
S. aureus USA300_TCH1516	140•35•10•26•378•160•59•102•37•44•90•21•31•50•182	328•72•32•187•264•240•104•10•128	442•100•203•321•64•14•221
S. aureus CA-347	–[a]	637•42•336•29	–
S. aureus RF122	–	–	–
S. aureus 6850	38•404•63•135•49•39•9•174•133	–	–
S. aureus LGA251	38•404•63•135•49•39•9•204•103	406•231•42•365	–
S. aureus TW20	38•404•63•135•49•39•9•174•30•103	–	–
S. carnosus TM300	38•404•63•223•9•80•147•104	70•195•15•9•573•206	223•216•15•315•252•47
S. *epidermidis* ATCC12228	–	–	22•201•165•66•108•254•234
S. epidermidis RP62A	–	–	22•201•165•66•108•254•234
S. haemolyticus JCSC1435	442•63•67•68•48•40•213•106	238•30•39•231•99•114•111•156•29	22•201•165•111•63•485
S. lugdunensis N920143	38•404•199•47•49•24•286	223•9•33•42•129•102•99•42•90•278	268•42•96•33•15•45•29•484•35
S. pseudintermedius ED99	38•467•67•156•9•307	145•237•153•509	454•108•75•16•391
S. saprophyticus ATCC15305	253•42•307•36•276•311•37•47•38•3	250•207•651•242	331•17•178•60•114•30•257•61•302

Staphylococcus spp.	Digestion patterns (nucleotides) with restriction endonucleases				
	CviAII	*RsaI*	*TaqI*	*BfaI*	*BfuCI*
b					
S. aureus USA300_TCH1516	133•57•447•47•204•445•32	167•105•81•226•617•135•34	511•221•56•199•161•217	676•451•116•122	262•23•165•528•387
S. aureus CA-347	–[a]	–	–	–	–
S. aureus RF122	286•577•181	–	275•54•715	–	9•285•60•600•90
S. aureus 6850	–	–	–	–	–
S. aureus LGA251	–	–	–	–	–
S. aureus TW20	–	–	–	–	–
S. carnosus TM300	229•57•72•126•379•205	245•313•113•49•348	275•54•67•29•396•165•82	25•57•210•776	–
S. *epidermidis* ATCC12228	–	–	–	–	–
S. epidermidis RP62A	–	–	–	–	–
S. haemolyticus JCSC1435	229•129•126•379•184	245•623•179	275•121•489•68•33•61	82•894•71	9•285•528•60•72•93
S. lugdunensis N920143	229•57•72•126•354•209	395•264•209•179	–	25•85•937	9•115•170•60•600•93
S. pseudintermedius ED99	45•97•87•255•299•261	245•475•113•211	182•111•751	25•111•908	96•198•750
S. saprophyticus ATCC15305	261•22•38•133•236•141•399•120	141•345•225•526•113	317•18•42•40•400•533	–	336•513•501

Symbol (•) indicates RE site in the gene sequences

[a] No unique pattern was observed

Table 3 *Staphylococcus* genomes distinguished based on unique Gene: restriction endonuclease approach

Staphylococcus spp.	Gene	RE	RE digestion patterns[a]	
S. aureus TCH60	*purH*	*BfuCI*	234•284•784•108•78	
S.aureus T0131			61•266•401•370	
S.aureus ED98		*CviAII*	208•207•33•222•98•8•44•477•63•72•47	
S.aureus 11819-97		*HpyCH4V*	325•72•42•93•120•54•45•114•41•28•255•162•90•38	
S.aureus 04-02981	*argH*	*BfuCI*	123•123•114•1047	
S.aureus Newman	*argR*	*Tru9I*	41•156•194•6•46•7	
S.aureus JKD6008			41•105•51•36•90•20•48•6•46•7	
S.aureus COL	*gyrB*	*RsaI*	425•45•46•190•164•270•318•477	
S.aureus JKD6159			83•342•45•46•354•270•318•477	
S.aureus Mu50	*cysS*	*AluI*	675•442•15•399•160•137	
S.aureus VC40	*pyrE*	*Tru9I*	86•35•39•4•5•4•83•108•42•105•18•83	
S.aureus MRSA252	*recR*	*Tru9I*	32•117•11•205•50•64•119	
S.aureus NCTC8325	*argH*	*BfuCI*	81•15•123•114•1047[b]	*BfaI:* 679•267•165•136•133
S.aureus USA300_FPR3757				*BfaI:* 458•221•267•165•136•133

Symbol (•) indicates RE site in the gene sequences

[a] Values represent restriction fragments (nucleotides)

[b] These two strains can be distinguished on the basis of additional RE*(BfaI)* digestion pattern

Table 4 Distinction of *Staphylococcus* genomes based on unique genes

Staphylococcus spp.	Gene-RE: digestion pattern[a]		
S.aureus MW2	*argH-BfuCI*	Unique gene[b]	*AluI:* 1096•144•467
	81•138•114•1047	*seh*	*HpyCH4V:* 121•123•233•249
			Tru9I: 5•9•83•52•12•32•76•176•168•54•59
S.aureus MSSA476		–[b]	
S.aureus Mu3	*argH-BfuCI*	Unique genes[c]: *bacA, dltX, dnlJ, fmt, graA, grab, graC, graD, graE, graF, graR, graS, mgrA, narQ, Pfk, phdB, saeR, saeS, sak, sarA, sarH, sarR, sepA, tetM, tufa, yhcR, yhcS, yjbM, yrhB*	
	96•135•118•1047		
S.aureus N315		Unique genes[d]: *atl, clfA, fab, fdhD, fmtA, isdA, isdB, isdC, isdD,isdE, isdF, isdG, lig, lpl4, lysP, metB, msrA, orfX, pdhB, pfkA, prfC, rnr, srtB, tuf, uppP*	

Symbol (•) indicates RE site in the gene sequences

[a] Values represent restriction fragments (nucleotides)

[b] *S. aureus* strains MW2 and MSSA476 can be distinguished by the presence of *seh* gene in the former strain

[c,d] Genes unique to these two strains respectively

Clostridium, *Vibrio, Streptococcus*, and *Yersinia* [30–32, 34]. It must be emphasized that since *recA* is present in many organisms, including *Streptococcus*, a close relative of *Staphylococcus*, however, the RE digestion patterns do not match with those of *Staphylococcus* species (Data not shown). This approach has provided genes which can be utilized as biomarkers for rapidly detecting *Staphylococcus* strains [10, 12].

Acknowledgments We are thankful to the Director of CSIR-Institute of Genomics and Integrative Biology (IGIB), and CSIR projects—GENESIS (BSC0121) and INDEPTH (BSC0111) for providing the necessary funds, facilities and moral support. Authors are also thankful to the Academy of Scientific & Innovative Research (AcSIR), New Delhi.

References

1. Iwg SCC (2009) Classification of staphylococcal cassette chromosome *mec* (SCCmec): guidelines for reporting novel SCCmec elements. Antimicrob Agents Chemother 53:4961–4967. doi:10.1128/AAC.00579-09
2. Vuong C, Gerke C, Somerville GA, Fischer ER, Otto M (2003) Quorum-sensing control of biofilm factors in *Staphylococcus epidermidis*. J Infect Dis 188:706–718. doi:10.1086/377239
3. Liu GY (2009) Molecular pathogenesis of *Staphylococcus aureus* infection. Pediatr Res 65:71R–77R. doi:10.1203/PDR.0b013e31819dc44d
4. Plata K, Rosato AE, Węgrzyn G (2009) *Staphylococcus aureus* as an infectious agent: overview of biochemistry and molecular genetics of its pathogenicity. Acta Biochim Pol 56:597–612
5. Kateete DP, Kimani CN, Katabazi FA, Okeng A, Okee MS, Nanteza A, Joloba ML, Najjuka FC (2010) Identification of

Staphylococcus aureus: DNase and Mannitol salt agar improve the efficiency of the tube coagulase test. Ann Clin Microbiol Antimicrob 9:23. doi:10.1186/1476-0711-9-23
6. Gray B, Hall P, Gresham H (2013) Targeting *agr*- and *agr*-like quorum sensing systems for development of common therapeutics to treat multiple gram-positive bacterial infections. Sensors 13:5130–5166. doi:10.3390/s130405130
7. Kalia VC (2013) Quorum sensing inhibitors: an overview. Biotechnol Adv 31:224–245. doi:10.1016/j.biotechadv.2012.10.004
8. Kalia VC (2014) Microbes, antimicrobials and resistance: the battle goes on. Indian J Microbiol 54:1–2. doi:10.1007/s12088-013-0443-7
9. Kalia VC (2015) 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
10. 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
11. 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
12. Koul S, Prakash J, Mishra A, Kalia VC (2015) Potential emergence of multi-quorum sensing inhibitor resistant (MQSIR) bacteria. Indian J Microbiol. doi:10.1007/s12088-015-0558-0
13. Iorio NL, Ferreira RB, Schuenck RP, Malvar KL, Brilhante AP, Nunes AP, Bastos CC, Dos Santos KR (2007) Simplified and reliable scheme for species-level identification of *Staphylococcus* clinical isolates. J Clin Microbiol 45:2564–2569. doi:10.1128/JCM.00679-07
14. Bennett RW, Lancette GA (2001) Bacteriological analytical manual. Chapter 12 *Staphylococcus aureus.* BAM: *Staphylococcusaureus*—Food and Drug Administration
15. 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
16. Francois P, Pittet D, Bento M, Pepey B, Vaudaux P, Lew D, Schrenzel J (2003) Rapid detection of methicillin-resistant *Staphylococcus aureus* directly from sterile or nonsterile clinical samples by a new molecular assay. J Clin Microbiol 41:254–260. doi:10.1128/JCM.41.1.254-260.2003
17. Gröbner S, Kempf VA (2007) Rapid detection of methicillin-resistant staphylococci by real-time PCR directly from positive blood culture bottles. Eur J Clin Microbiol Infect Dis 26:751. doi:10.1007/s10096-007-0368-3
18. Schuenck RP, Em Pereira, Iorio NLP, dos Santos KRN (2008) Multiplex PCR assay to identify methicillin-resistant *Staphylococcus heamolyticus*. FEMS Immunol Med Microbiol 52:431–435. doi:10.1111/j.1574-695X.2008.00387.x
19. 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 mecALGA251. J Antimicrob Chemother 67:2338–2341. doi:10.1093/jac/dks221
20. 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
21. Thomas LC, Gidding HF, Ginn AN, Olma T, Iredell J (2007) Development of a real-time *Staphylococcus aureus* and MRSA (SAM-) PCR for routine blood culture. J Microbiol Methods 68:296–302. doi:10.1016/j.mimet.2006.09.003
22. Francois P, Bento M, Renzi G, Harbarth S, Pittet D, Schrenzel J (2007) Evaluation of three molecular assays for rapid identification of methicillin-resistant *Staphylococcus aureus*. J Clin Microbiol 45:2011–2013. doi:10.1128/JCM.00232-07
23. Malhotra-Kumar S, Haccuria K, Michiels M, Ieven M, Poyart C, Hryniewicz W, Goossens H (2008) Current trends in rapid diagnostics for methicillin-resistant *Staphylococcus aureus* and glycopeptide-resistant *Enterococcus* species. J Clin Microbiol 46:1577–1587. doi:10.1128/JCM.00326-08
24. Malhotra-Kumar S, Van Heirstraeten L, Lee A, Abrahantes JC, Lammens C, Vanhommerig E, Molenberghs G, Aerts M, Harbarth S, Goossens H, MOSAR WP2 Study Team (2010) Evaluation of molecular assays for rapid detection of methicillin-resistant *Staphylococcus aureus*. J Clin Microbiol 48:4598–4601. doi:10.1128/JCM.00004-10
25. Shore AC, Deasy EC, Slickers P, Brennan G, O'Connell B, Monecke S, Ehricht R, Coleman DC (2011) Detection of staphylococcal cassette chromosome *mec* type XI carrying highly divergent *mecA*, *mecI*, *mecR1*, *blaZ*, and *ccr* genes in human clinical isolates of clonal complex methicillin-resistant *Staphylococcus aureus*. Antimicrob Agents Chemother 55:3765. doi:10.1128/AAC.00187-11
26. 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
27. Kalia VC, Mukherjee T, Bhushan A, Joshi J, Shankar P, Huma N (2011) Analysis of the unexplored features of rrs (16S rDNA) of the genus *Clostridium*. BMC Genom 12:18. doi:10.1186/1471-2164-12-18
28. 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
29. Kalia VC (2015) Let's explore the latent features of genes to identify bacteria. J Mol Genet Med 9:e105. doi:10.4172/1747-0862.1000E105
30. 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
31. 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
32. 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
33. 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
34. Kalia VC, Kumar R, Kumar P, Koul S (2015) A genome-wide profiling strategy as an aid for searching unique identification biomarkers for *Streptococcus*. Indian J Microbiol. doi:10.1007/s12088-015-0561-5
35. Koul S, Kumar P, Kalia VC (2016) A unique genome wide approach to search novel markers for rapid identification of bacterial pathogens. J Mol Genet Med (in Press)
36. Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp Ser 41:95–98