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

4,300
Open access books available

116,000
International authors and editors

125M
Downloads

154
Countries delivered to

TOP 1%
Our authors are among the most cited scientists

12.2%
Contributors from top 500 universities

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

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

Numbers displayed above are based on latest data collected. For more information visit www.intechopen.com
Genotyping Approaches for Identification and Characterization of *Staphylococcus aureus*

Mazen M. Jamil Al-Obaidi, Zarizal Suhaili and Mohd Nasir Mohd Desa

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/intechopen.75969

Abstract

Genotyping methods are vital epidemiological tools for discriminating different bacterial isolates within same species, which in turn provide useful data in tracing source of infection and disease management. There have been a revolutionary efforts in ways to distinguish between bacterial types and subtypes at molecular level utilizing DNA in the genomes. Notably, the growth of various DNA typing methods has provided innovative apparatuses for improved surveillance and outbreak investigation. Thus, early identification and genotyping are indispensable as resources for managing therapeutic treatment and the control of rapid expansion of clinically important bacteria. Methicillin-resistant *Staphylococcus aureus* (MRSA) has been in a great attention due to its contagious nature and subjected to various typing analyses. Thus, in this chapter, we aimed to review the contribution of various genotyping methods of commonly used as well as those unique to *staphylococci* in understanding its epidemiology, infection and dissemination pattern, and to provide an impression of specific advantages and disadvantages of each tool.

Keywords: genotyping, MLST, RFLP, *Staphylococcus aureus*, WGS

1. Introduction

Typing is a process to characterize the species and properties of organisms, in particular the discrimination at the strain level both phenotypically and genetically. Conventional typing such as serotypes, biotype, and phage type has been in practice for many years. Nevertheless, typing at molecular level is nowadays very essential due to its specificity, which is often used to support the associated phenotypic characteristics. For example, one species may comprise
many subtypes or a subpopulation which one might be more pathogenic than the others. Thus, genotyping plays an important role to identify potential differences at genetic level as well as for epidemiological traceability of all the presented isolates [1].

A good typing method must have the discriminatory power to differentiate all unrelated isolates epidemiologically to facilitate any outbreak investigation. This will allow investigation to demonstrate person-to-person strain transmission, subsequently, allowing preventive measures to be designed to inhibit further dispersion of the pathogens. Additionally, genotyping method must be inexpensive, rapid, easy to interpret and highly reproducible [1]. For a continuous surveillance, genotyping methods must produce results with a sufficient stability over time. Also, it should produce portable data and can be easily accessed through open source web-based database or a client-server database connected via the internet, facilitating global comparison of the isolates.

Genotyping methods are basically based on phenotyping, PCR/sequence typing and genome typing approaches. Remarkably, a great effort has been put up in epidemiological investigations of *Staphylococcus aureus* due to its role as a leading nosocomial, community and livestock-acquired bacterial pathogen. Globally, dynamic spread of methicillin-resistant *Staphylococcus aureus* (MRSA) strains stimulates the increasing rates of these strains in several regions rapidly. Additionally, the global emergence of MRSA had influenced significantly the health care systems all around the worlds over the past 50 years. The epidemiological changes of *S. aureus* infection in human beings and animals are being focused for two main reasons: (i) to understand the evolution and dissemination pattern of the species and (ii) to find a proper antimicrobial treatment strategy and effective infection management. Therefore, epidemiological studies utilizing various typing techniques are continuously on the go in various parts of the world especially those with increasing rates of MRSA infections. In this chapter, we aim to review the contribution of various genotyping methods commonly in use as well as those unique to *Staphylococal* particularly MRSA, that might assist to detect the outbreak infections, conduct epidemiological surveillance by means of rapid typing and to provide an impression of specific advantages and disadvantages of each typing tool.

2. Phenotypic detection and identification of MRSA

Numerous conventional or molecular methods can be applied for detection and identification of *S. aureus* including colony morphology, production of coagulase activity and by various enzyme activity. Also, commercial latex agglutination tests and the API Staph system (bioMérieux) are examples of assays available for identification of *S. aureus*, which remains the methods of choice due to their feasibility and low cost. Additionally, there are other phenotypic methods such as biotyping and immunoblotting, serotyping, phage-typing and multilocus enzyme electrophoresis (MLEE) that have been frequently applied previously.

Upon identification of *S. aureus*, antimicrobial susceptibility profile is always performed so that the choice of antimicrobial treatment can be formulated. The antimicrobial testing procedures to a broad range of antimicrobial agents have been standardized and improved for
the accuracy of reporting, following guidelines either by the Clinical & Laboratory Standards Institute (CLSI), The European Committee on Antimicrobial Susceptibility Testing (EUCAST) and British Society for Antimicrobial Chemotherapy (BSAC) guidelines. The methods are usually carried out to determine which antibiotic shall be the most effective in treating bacterial infection in vivo. This simple and useful method offer relatively inexpensive and is usually applied for clinical investigation of pathogen isolated from various types of infection. Examples of the antimicrobial susceptibility tests are disc diffusion, agar/Tube agar dilution, CHROMagar oxacillin resistance screening as well as agar-based and Epsilometer test (E-Test).

Currently, these tests are considered as the most popular methods of choice that can support genotyping data [2]. Nevertheless, previous reports suggested that phenotypic methods of identification have drawbacks due to the variability in expression of phenotypic characterization by isolates belonging to the same species and their reliance on subjective interpretation of test results [2]. As a result, phenotypic detection and identification was heavily burdened with several issues; including low reproducibility, reliability, sensitivity and specificity as well as lack of resolution in epidemiology investigation. Therefore, several reports have suggested that genotypic identification and detection methods offered a higher discriminatory power, reliability, reproducibility and typeability [3]. Genotypic identification can be done with the phenotypic approach together for a better comparative analysis. Furthermore, the introduction of molecular screening for MRSA detection as well as identification directly from clinical specimens has been developed to enhance and identify common Staphylococcal spp., as well as to speed up the detection methods especially in clinical research [4].

3. Genotyping of MRSA

The ultimate goals for bacterial typing are to further clarify the population dynamicity and also to track the spread of the microorganisms. As mentioned earlier, traditional bacterial typing of phenotypic-based alone, does not provide the prudent resolution for identifying and tracking an infection-causing pathogen, and also does not clearly describe the transmission pattern of an outbreak. However, molecular typing has been an invaluable tool for molecular epidemiologist as well as clinical researchers for tracing the spread of particular strains, discovering the route of dissemination and the potential reservoirs. Usually, the outcomes of epidemiological investigations are often used to guide and assist the clinical treatment of the patients by selecting the appropriate antimicrobial agents. Furthermore, molecular typing contributes to the comprehensive understanding of the epidemiology of infection and facilitates infection control measures as well as management [5].

It is well known that S. aureus is frequently associated with clonal spread as reported by many studies utilizing various typing methods on huge numbers of S. aureus strains. For example, molecular strain typing of MRSA is implemented in order to elucidate genetic variation to guide in outbreak investigation as well as to characterize genetic macroevolution for spatial-temporal and evolutionary studies [6]. In those studies, PCR-based methods are commonly used for typing as they are easy and fast technique. Other methods such as pulsed-field gel electrophoresis (PFGE), coagulase gene PCR-restriction fragment length polymorphism
(RFLP) and *Staphylococcal* cassette chromosome mec typing (SCCmec) also play similar important roles in molecular typing of both MSSA and MRSA [7]. Additionally, sequence-based techniques also play an vital role in genotyping, including *spa* gene-typing and multilocus sequence typing (MLST), that have been considered as a very useful tool for epidemiological studies, particularly MRSA [8]. Several reports suggested two methods known as PFGE and MLST that are considered as ‘gold standard’ in typing of both MSSA and MRSA, although these typing methods are often time-consuming, costly and laborious [8]. In the subsequent sections, various genotyping methods are presented to elaborate each extent to establish molecular epidemiology studies.

3.1. Polymerase chain reaction (PCR)-based identification

Polymerase chain reaction (PCR) is known as enzymatic method used to exponentially amplify a specific preselected fragment of DNA. It is well known that PCR uses a thermostable DNA polymerase *in-vitro* to multiply copies of a specific nucleic acid region exponentially. The procedures require DNA template from the organisms being typed, thermostable DNA polymerase, two synthetic oligonucleotide primers and four standard deoxyribonucleoside triphosphate that are incorporated into newly synthesized DNA. There are numerous PCR-based amplification methods that have been applied widely in the subtyping of various microorganisms, including *S. aureus* especially MRSA as stated in this chapter. Various phenotypic tests have been used for identification of MRSA from other *Staphylococcal* spp., such as screening for production of protein A, cell-bound clumping factor, extracellular coagulase and heat-stable nuclease [9]. However, a good package of rapid molecular detection is also required for screening and identification of certain antibiotic resistance determinants as well as virulence factors of *S. aureus* [3, 10]. To date, most of the molecular approaches for the identification of MRSA have been a PCR-based method with a range of primers designed to amplify specific targeted markers encompassing species-specific, antibiotics resistance as well as virulence determinant [11]. Other PCR sequencing-based methods have been developed for the identification of *S. aureus* from other coagulase-negative *staphylococci* targeting 16S rRNA, RNA polymerase B (*rpoB*), *femA*, *tuf* and *gap* genes. However, these approaches have their own limitations as it is not sufficiently discriminatory to differentiate closely related *staphylococcal* spp., where database of these genes only include a limited number of *Staphylococcal* spp. [12].

3.1.1. Amplified fragment length polymorphism (AFLP)

AFLP is a PCR-based method applied in DNA fingerprinting and genetic research. In this method, restriction enzymes (e.g. endonuclease) are used to cut the genomic DNA of the typed species, subsequently, double-stranded oligonucleotide adaptors which are comprised of a core sequence and an enzyme-specific sequence, are bound to one of the sticky ends of the restricted fragments. After that, a PCR thermocycler is used to amplify those restricted fragments ending with the adapter selectively, using primers complementary to the adapter sequence. Then, the restriction site sequence and a number of additional nucleotides from the end of the unknown DNA are designed. Usually, the restriction fragments (50–100) are amplified using fluorescent dye-labeled PCR primers, to detect those separated fragments by size using automated DNA sequencer. Likewise, gel electrophoresis can also be used to visualize
and analyze the amplified fragments of typed DNA. Upon analysis, a high-resolution banding is generated via computer reflecting the genetic relatedness among bacterial isolates [13]. This kind of technique has a higher discriminatory power in comparison to PFGE, where it was shown in a study that AFLP analysis provided greater genetic resolution and was less sensitive to DNA quality during genetic typing of bacterial pathogens E.coli O157:H7 in epidemiological investigation [14]. Additionally, like other DNA banding pattern-based method, AFLP can be automated and has portable results, as well as reproducible approach to facilitate the analysis [14]. Previous study has been conducted by Fossum and Bukholm [15] reported MRSA population was revolutionized from hospital-acquired MRSA (HA-MRSA) to community-acquired MRSA CA-MRSA in the south-eastern part of Norway through increase in MRSA clones harboring SCCmec IV as shown by AFLP, MLST and spa typing methods. AFLP analysis grouped the MRSA isolates into clusters according to the clonal complexes (CCs), but did not discriminate among the different sequence types (STs) or spa types inside each CC [15]. In the United Kingdom, there are 16 phage types of epidemic MRSA (eMRSA) strains that have been identified, of which eMRSA-3, -15 and -16 now predominate [16]. Through this approach specifically fluorescent AFLP (FAFLP), it was able to classify eMRSA phage type from 1 to 16 by identifying eMRSA phage type of S. aureus (eMRSA-15) from UK [17] and into 9 clone clusters in European isolates [18]. Thus, AFLP is considered as a tool with highly discriminatory power against these strains of MRSA. As a result, this technique is considered suitable for MRSA epidemics surveillance at national and international levels as well as reproducible approach. Additionally, it is found that AFLP approach is more reproducible than PFGE and MLST, and it is more suitable for inter-laboratory data exchange using sequence-based data [19]. The main drawbacks of this method are labor-intensive and expensive.

3.1.2. 16s ribosomal RNA (16s rRNA)

16S rRNA comprises ~1500 pair nucleotide sequence coding for catalytic RNA that is part of the 30S ribosomal subunit. 16S rRNA gene is comprised of nine variable regions (V1–V9/30–100 base pairs long), that show sequence diversity among different bacterial species, subsequently enable for identification purposes. V1–2–3 regions are located at the 50 end of the 16s RNA gene which is shown to be appropriate and more sensitive than other regions for identification of different types of bacteria [20]. This gene is constant in function, promising a valid molecular chronometer, where it exists in all prokaryotic cells. Therefore, it is used to elucidate both close and distant phylogenetic relationships at the genus and at the species level [21] based on the differences in the nucleotide sequence of 16s RNA gene. Additionally, dedicated 16S databases [22] that include near full-length sequences for a large number of strains and their taxonomic placements are available. The sequence from an unknown strain can be compared against these available sequences in the database. This approach is considered as a common substitute for traditional methods using (rRNA) gene sequencing [23]. It is less time-consuming and labor intensive, where DNA sequencing can offer more absolute taxonomic classification than culture-based approaches for numerous organisms [23]. However, there are also limitations with this approach associated with the short read lengths, variances ascending from the diverse regions selected, sequencing errors and difficulties in evaluating operational taxonomic units (OTUs) [23, 24]. Additionally, single marker (16S rRNA) usage
is considered challenging to assess the bacterial diversity, subsequently difficult to identify bacterial species [25], as well as the resolution of 16S rRNA that is very limited among closely related species. A previous study has shown that 16S rRNA combined with meca and muci using multiplex PCR, is considered as useful tool for rapid characterization of MRSA [26]. Thus, this multi-gene technique is considered a better discrimination tool among unrelated isolates, particularly in S. aureus [27].

3.1.3. Staphylococcal cassette chromosome mec typing (SCCmec)

SCCmec complex is a mobile genetic element that confers the methicillin resistance profile in S. aureus. MSSA may emerge to become MRSA upon acquiring this genetic complex. SCCmec contains essential elements which can be detected by regular PCR; (i) ccr genes which are constituted by ccrA and ccrB and (ii) meca gene complex which is composed of meca gene and its regulatory genes, mecI and mecRI. Currently, 11 major types of SCCmec elements (I–XI) have been identified based on the organization of the meca gene complex, ccr gene complex and integrated plasmids (http://www.SCCmec.org/). To date, there are four allotypes (types 1, 2, 3 and 5) of ccr complex and three classes (A, B and C) of meca complex [28]. Different combinations of these complex classes and allotypes generate various SCCmec types. SCCmec elements are currently classified into types I–V based on the nature of the meca gene complex and ccr allotypes [26]. At present time, multiplex PCR is used for the characterization of SCCmec types. For example, Okuma et al. [29] developed primers that were specific for SCCmec IVa and SCCmec type IVb, meanwhile, Hisata et al. [30] developed multiplex PCR for the specific identifications of SCCmec type IIa, IIb, IVc and IVd.

Also, two different multiplex PCR methods were developed by Zhang et al. [28] and Milheirico et al. [31] for specific characterization of SCCmec type I to SCCmec type V. Likewise, 9 pairs of primers were used by Zhang et al. [28] for identification of SCCmec type I, II, III, IV (a, b, d) and V, whereas 10 pairs of primers were used as described by Milheirico et al. [31]. Interestingly, Boye et al. [32] developed an easy screening of MRSA SCCmec typing only by using multiplex PCR with a combination of four pairs of primers, where clear and easily discriminated band pattern was obtained for all major five types of SCCmec. These characterization methods could be used to distinguish HA-MRSA and CA-MRSA typing, where SCCmec types I, II, III and VIII are usually acquired by HA-MRSA, while SCCmec types IV, V, VI and VII are acquired in CA-MRSA [33]. Thus, it is very useful and important molecular tool in understanding the potential epidemiological background of the strains.

3.1.4. Multiple-locus variable-number tandem repeat assay (MLVA)

Multiple-locus variable-number tandem repeat analysis (MLVA), was previously known as a variable-number tandem repeat (VNTR) [34] by making use the VNTR polymorphism. In 2008, after the introduction of spa typing as a standard molecular typing method in the Germany MRSA surveillance, MLVA was added as a supportive typing technique. This method involves PCR amplification of five specific loci (sdr, elfA, elfB, ssp and spa) of S. aureus which is composed of seven individual genes (sdrCDE, elfA, elfB, sspA, spa, meca and fnbP) [35] using multiplex PCR mixture followed by separation of the amplified bands on agarose gel and comparison of the band patterns between strains to identify genetic clusters or clones [36]. This genotyping
method showed a successful typing of MRSA isolates in many studies [37], in term of determining the genetic diversity and evolutionary lineage with discriminatory power.

It was shown that this approach has a reproducibility as good as PFGE technique [34]. The main drawback of this approach is that in highly conserved genomes, there may not be sufficient DNA polymorphisms in these limited sequence targets to exhibit alleles. Another limitation was, small deletions and insertion in the regions flanking the repeat units may lead to misinterpretations, making the MLVA results slightly more ambiguous than sequenced-based methods. However, to overcome this limitation, the DNA sequence of each new allele is determined to confirm the deduced number of repeats [38]. However, the level of discrimination can be increased by adding more loci and repeating the assay with different restriction enzymes [39].

3.1.5. Repetitive element polymerase chain reaction (rep-PCR)

rep-PCR is a DNA-based technique that discriminates microbes at subspecies or strain level by observing genomic DNA fingerprint patterns [40]. In this approach, the hybridization of primers to noncoding intergenic repetitive sequences takes place across the genome. The amplicons are produced during DNA amplification of the repetitive elements. Depending on the distribution of the repeat elements across the genome, the genetic relatedness between the bacterial isolates can be inferred by comparing the banding pattern of the amplicons. Enterobacterial repetitive intergenic consensus’ (ERIC 124–127 bp), ‘the repetitive extragenic palindromic’ (REP 35–40 bp), and the ‘BOX 154 bp’ sequences are examples of conserved repeat sequences that have been used successfully in rep-PCR typing [40]. This kind of approach is considered as highly discriminatory tool for different bacterial organisms such as S. aureus and Campylobacter jejuni [41, 42]. However, there was one drawback for this method which was low rate of reproducibility, due to the uses of traditional agarose gels for electrophoresis, which might result in a discrepancy in relation to the use of different reagents and gel electrophoresis systems.

Alternatively, rep-PCR approach is developed and used by a semi-automated method using DiversiLab system (bioMérieux, Marcy l’Etoile, France), where clinically important organisms can be detected by commercial PCR kits [43]. Then, high-resolution chip-based microfluidic capillary electrophoresis is used to separate amplified genomic DNA within repetitive elements, where chip-based microfluidic capillary electrophoresis can increase the determination and reproducibility of the rep-PCR method compared to traditional gel. In the next step, DiversiLab software is used to normalize and analyze the data automatically. Several reports have evaluated the usefulness of this approach (DiversiLab) in outbreak-related and epidemiology unrelated bacterial isolates [44]. It was shown that this approach is rapid, reproducible and easy for typing microorganisms. Hospital outbreaks of MRSA have been identified using this useful DiversiLab tool by Fluit et al., [45]. In contrast, other study found that this tool is not highly discriminative tool for MRSA typing particularly in outbreak setting [46]. The main limitation of this approach is the DiversiLab databases are stored only on manufacturer server, resulting in some users not allowed to use this typing system due to security purposes.

3.1.6. Restriction fragment length polymorphism (RFLP)

In PCR-RFLP approach, restricted enzymes are used to detect the variations in homologous DNA fragments. Then, the DNA fragments are amplified using regular PCR, subsequently
these fragments are separated by gel electrophoresis based on length of the fragments. For example, coagulase gene (coa) and Staphylococcal protein A (spa) gene RFLP amplified fragment of DNA could be identified through this technique. Previous studies have shown PCR-RFLP typing of coa gene as useful tool to discriminate S. aureus strains on the basis of sequence variation within the 3’ end coding region of the gene [47]. The amplification discriminatory power of coa gene depends on the heterogeneity of the region containing 81 bp tandem repeats at the 3’ coding region of the coa, where this region is different in the number of tandem repeats and the location of AluI and HaeIII restriction sites among different isolates [48]. AluI is better than HaeIII in S. aureus typing, but both can be used to be more reliable and sufficient power in discrimination issues. It is found that coa-RFLP typing has discriminatory power for S. aureus strains particularly in MRSA strains [49]. On the other hand, the repeated part of spa is located at 3’end and identified as X region; the repetitive part of region X comprises of up to 12 elements each with a length of 24 nucleotides. High polymorphic is defined by this 24-nucleotide region with respect to the number and sequence of repeats. Variety of X region causes protein A variation [50]. Thus, the potential dissemination of MRSA can be detected by the number of repeats in the region X of spa [51]. As a result, the PCR-RFLP assays (coa and spa RFLPs) are useful molecular markers for a rapid, and initial study of MRSA outbreaks [51]. Wichelhaus et al. [52] reported that this method is proven as to have a good discriminatory power, type-ability and reproducibility in MRSA typing. Moreover, this technique can be used in routine infection control program in health care systems as well as epidemiological investigations [48].

3.2. Sequence typing method for bacterial identification

3.2.1. Staphylococcal protein A (spa) typing

Staphylococcal protein A (spa) typing is a sequence-based method that targets VNTR of the spa gene region encoding protein A [53]. The spa gene region is polymorphic as a result of spontaneous mutations and loss or gain of repeat. Besides, spa gene is reported to be a highly effective tool in subtyping both MSSA and MRSA [54]. As mentioned above, the region X of spa gene consists of 24 bp repeats sequences, and the diversity of the strains is recognized by duplications and deletions of the sequence in this region of the gene. The variation in the sequences is used to assign repeats numbers [55]. Sequenced data can be analyzed using free accessible offline bioinformatics tool Ridom Bioinformatics (Ridom, GmBH, Germany) (http://spaserver.ridom.de) [56]. This spa server database also provides global frequencies information related to the mapping of the spa with the MLST S. aureus database. To date, 748 diverse repeats with more than 17,416 spa types have been described from 131 countries with total strains 384,806 (http://spaserver.ridom.de). Sequences of perfect quality are synchronized with spa server (Ridom server) [57] specifically for spa typing, providing a typical worldwide nomenclature together with integral quality control.

Subsequently, Based Upon Repeat Pattern (BURP) algorithm is used to analyze the diverse spa types associated to each other. The analysis shows a good consistency with MLST-CCs, where ST that shares at least five of seven identical alleles are grouped into a single CC [58]. The advantages of this typing method are the results generated are easy to interpret, less time-consuming, highly reproducible, less laborious and highly comparable between laboratories...
via ridom.spa.server compared with PFGE. Besides, spa typing is impressive for its ease of interpretation and suitability for international comparison. It is also able to detect both slowly and rapidly accumulated molecular variations as well as to investigate outbreaks in epidemiological studies and molecular evolutions of population structure [59]. However, non-typeable (NT) isolates are increasingly found in the Dutch MRSA surveillance as well as globally. Thus, to overcome the issue of NT strains, other typing method should be concurrently used to be a supportive method [60]. Malachowa et al. [61] found that spa typing was more approximate to MLST approach upon comparing four genotyping methods (PFGE, MLST, MLVA and spa typing) in 59 S. aureus strains. Additionally, HA-MRSA, CA-MRSA and livestock MRSA (LA-MRSA) dissemination can be monitored by a combination of these analyses together with spa typing in epidemiological studies at a global level [62].

3.2.2. Multilocus sequences typing (MLST)

MLST has been invented to overcome the poor or insufficient portability of traditional and older molecular typing application. The main idea of this tool is based on MLEE [63] which depends on the differences in electrophoretic mobility of various enzymes exist in a bacterial species. Neisseria meningitidis was the first species subjected to MLST analysis in 1998 [64]. After that, this tool was developed to detect other type of bacterial species, where it became a widely accepted tool for molecular epidemiological studies as well as evolutionary studies of pathogen at the molecular level [65]. This molecular subtyping method was developed for bacterial characterization to facilitate rapid and global comparisons among species [66]. In term of MRSA, seven housekeeping genes are amplified and sequenced for internal sequences [67]. In the subsequent analysis, MRSA isolates are grouped within a single CC when five out of seven housekeeping genes (400–500 bases) in that particular MRSA isolates having identical sequences and isolates with the seven same allelic profiles may be descended from a common ancestor [66, 67]. If there are various alleles at each of the seven loci, the isolates are unlikely to have the identical allelic profiles by chance, while isolates that have similar allelic profile can be considered members of the same clone [66]. The variations found among these genes are mostly synonymous and neutral. Since these genes accumulate variations in a slow manner, they are considered to be reliable indicator of evolutionary history [68]. The main advantage of this tool is that whole produced data are obvious due to standardized nomenclature internationally and reproducible. Additionally, ST profile as well as alleles sequences are available in huge central databases (http://pubmlst.org and www.mlst.net) [69] that are freely accessible online. Moreover, the genetic relatedness between bacterial strains within a species can also be identified via the databases.

Thus, it is a useful tool to compare the data with other laboratories via web-based electronic data. Furthermore, it allows the exchange of data collected over internet through the MLST database. BURST software package can analyze the evolutionary events within S. aureus population [67]. For instance, MRSA-ST239 was found to disperse in different countries although carrying a similar ST [70]. The drawback of the technique is the high cost, time-consuming, labor-intensive, and also has no discrimination power for cases related to short-term outbreak. For the later, this technique may not discriminate well the epidemic spread of bacterial strains within a limited time frame [19]. Nevertheless, MLST is still considered as the rapid
method for subtyping for MRSA in clinical research, and has been shown to be useful in global epidemiological studies of *S. aureus* [67].

### 3.3. Genomics-based typing tool

#### 3.3.1. Pulsed-field gel electrophoresis (PFGE)

PFGE is an approach used to detect the dispersion of large segments of DNA using gel with high electrical fields that facilitate changing in DNA direction periodically [71]. In brief, molecular sieve of gel is used to transform DNA from cathode to anode using common electrophoresis method. Two electric fields are used in PFGE technique, where it allows to change the directions of the DNA as mentioned above. Subsequently, ethidium bromide dye is used to differentiate the DNA band spectrum as a typing result. Clinically, various types of bacteria can be genotyped by PFGE which is considered as the “gold standard” genotyping method. It is assumed as an epidemiological tool for most bacterial species since 1990s [71]. Currently, PFGE is used worldwide to identify and characterize isolates of bacteria in outbreak investigations [19, 71, 72]. It is also considered as prototype tool to analyze center to center transmission events [73].

In term of *S. aureus*, isolation of intact bacterial chromosomes are required prior to PFGE procedures, where these isolated chromosomes subsequently is broken down into large DNA fragments using cutting restriction endonuclease such as SmaI. Subsequently, the restriction fragments can be separated via agarose gel “pulse-field” electrophoresis, where those separated DNA fragment could be monitored as a banding pattern in the gel. For easier analysis, large restriction fragments (30 kb–1 Mb) are separated based on their size in a dependent manner, yielding few bands on the gel [74]. It is well known that traditional electrophoresis is able to separate DNA fragments up to 20–50 kb only. Thus, this method has been invented to overcome this weakness through modifying the direction of the electrical field to mobilize DNA fragments of up to ~2 Mb [75]. Subsequently, the gels are dyed and captured by an imaging system and analyzed using BioNumerics software programs with the Dice coefficient and un-weight pair group matching analysis (UPGMA) setting according to the criteria as described by Fred et al. [72]. After that, graphical dendrogram may be generated by DNA fingerprinting software.

PFGE has been found to show a higher discriminatory power than PCR-RFLP of coa gene and other PCR-based fingerprinting methods as it enables the entire chromosome to be analyzed, whereas the PCR-based fingerprinting methods explore only selected (random) portions of it [76]. Previous studies stated the reproducibility of PFGE is considered high due to the standardization of protocols [77], allowing national and international surveillance systems [78], and standard interpretation guidelines to investigate the emergence of bacterial species particularly *S. aureus*. Previous study had been done to compare different tools such as MLST, PFGE and AFLP for genetic typing of *S. aureus*. It was found that PFGE is less reproducible, and less useful for long-term epidemiological investigations or phylogenetic relationships evolution in *S. aureus* strains [19]. Thus, this method is found extremely helpful
in the short-term investigation and identification of MRSA outbreaks in hospital, community and livestock-associated [79]. The solid advantage of this application is the ability to address a large number of an investigated genome (>90%). However, there are certain disadvantages of this application such as time-consuming and labor-intensive, as well as insufficient resolution power to differentiate bands of identical size. It also requires highly skilled operators and there are no standardized reagents with technically laborious and lack of centralized criteria for interpreting the banding patterns [80].

3.3.2. DNA microarray

DNA microarray typing method uses a collection of DNA probes that are attached to a solid surface in ordered manner. Ideally, complementary nucleotide sequences for specific bacterial isolates are detected by DNA probes. This approach is specific tools to identify several genes for specific bacterial strains. It can also be used to identify allelic variations of a gene which exists in all strains for particular species. Usually, target DNA could be labeled by chemical, enzymatic reaction and DNA microarray hybridization. Then, labeled target DNA and an immobilized probe create signal due to successful hybridization, giving measurement automatically using scanner. Currently, this approach is extensively used to analyze genomic mutations such as single-nucleotide polymorphisms (SNPs). It is also found that this approach is an excellent application to identify exceptional antibiotic resistance and virulence genes simultaneously to represent epidemiological markers of certain isolates of interest [81]. Whole genome microarray approach is the alternative tool for whole genome sequencing (WGS) for saving time, expenses and efforts, where it has ability to investigate genetic features of isolates involved in outbreak. For example, 31 chromosomes and 46 plasmids were identified from a various set of E. coli isolates, subsequently, the presence or absence of genes were detected in very recently emerged E. coli O104:H4 using microarray system [82]. Interestingly, more than 3000 clinical and veterinary isolates of MRSA were characterized epidemiologically through Alere StaphyType DNA microarray system, covering 334 target sequences, including 170 distinct genes and their allelic variants [83], showing a high level of biodiversity among MRSA, especially among strains harboring SCCmec IV and V elements. Overall, this technology is highly accurate, but the reproducibility data needs to be established to the broad application to be shared globally. Additionally, this approach is considered not practical if the target of typing is SNPs of highly clonal species. Another disadvantage of this approach that the detection is limited only to sequences that is included in the array.

3.4. Whole genome sequencing (WGS)

To investigate genome variations, cost-effective way has been invented for genetic investigations, which is second generation sequencing (NGS) or high-throughput sequencing. This technique is named second generation to differentiate it from first generation sequencing based on the Sanger method. The main advantage of this approach over several traditional sequencing methods is the ability to create millions of reads (35–700 bp length) in one shot,
which also leads to a reduction in cost. The nucleotide sequence of the genome is constructed by gathering numerous short sequences reads from overlapping regions, or comparison with previous reference sequences genomes (re-sequencing).

Currently, WGS is considered as a high attractive tool for epidemiological studies [84], and it is believed that this method has the potential as routine tool for bacterial identification and characterization in the near future. Nevertheless, the main challenge for this approach is the interpretation and computation of the huge set of data. This approach is currently used to determine the genetic relatedness between bacterial isolates based on sequence analysis of the whole genome. Additionally, WGS has the ability to distinguish various genomes within an SNP, which cannot be achieved in conventional molecular typing approaches. Thus, characterization of transmission events and outbreaks will be accurate. However, extensive studies must be conducted to translate this prospective tool into a routine practice. It is well known that the methods based on SNPs permit detailed and targeted analysis of variations among related organisms. Thus, WGS using SNPs analysis can identify the isolates related to an outbreak from non-outbreak isolates. Moreover, various phenotypic characteristics such as virulence and antibiotic resistance of particular pathogen can also be inferred by WGS technique. Finally, this approach enables the search for genetic markers, such as the presence or absence of a gene or an amino acid substitution in a protein, facilitating the linkage with the occurrence, severity and virulence of the disease.

Clinically, SNPs analysis on MRSA isolates recovered from an outbreak in a unit care for neonates using WGS sequencing approach was able to offer relevant data within a time frame that can stimulate patient care [85]. Additionally, through WGS, data can also reliably predict antibiotic susceptibility phenotype of MRSA from an outbreak scene [86], leading to development of hospital infection management and patient outcomes in routine clinical practice. Some previous studies took benefits from WGS by investigating CA-MRSA in USA including USA300-0114 [86], where genetic variation was found. Considering the fact that the isolates were recovered and originated from a confined geographical area, the WGS analysis suggested the continuous evolution of this clone within the limited region. These results offer additional support for the use of WGS as a first-line screening method, which is comparable with those gained by phenotypic methods [87]. Furthermore, additional genes may be added to the panel to increase the coverage and sensitivity, where sequenced isolates can be screened to recognize new resistance genes.

As a result, WGS is considered as a rapid prediction of resistance which contributes to effective clinical management, particularly for S. aureus. Subsequently, this approach permits the characterization of transmission routes to improve infection control strategies and manage the outbreak. Once the genetic basis of virulence is understood, WGS could permit determination of emerging infectious strains (and new virulence genes) locally and globally. Furthermore, if genetic diversity is characterized over time, it will provide new knowledge of the S. aureus population structure, subsequently leading us to obtain extra information and understand the genetic basis of the disseminating strains. As a whole, it is suggested that WGS is considered as a very useful tool in epidemiological investigations to discriminate MRSA, and it may assist to trace person-to-person transmission in health care systems [88].
4. Conclusions

Higher rates of morbidity and excessive healthcare costs are the two main reasons that can be caused by the growing number of HA, CA and LA-MRSA infections. The management of these infections must be conducted through the screening of individuals as well as infection control program. Currently, MRSA can be reliably detected within hours using rapid screening methods. However, the continuous evolution of SCCmec MRSA strains requires frequent monitoring of the strains. Therefore, genotyping techniques must be sufficient with internal and external quality control and standardized internationally for MRSA diagnostics. For that reason, to reduce the severe clinical and economical effects of MRSA, rapid and accurate typing is required especially for epidemiological investigations. Currently, conventional and molecular methods are used in combination for MRSA typing. Nevertheless, controversy is still on-going to choose which molecular typing methods will suit every requirement to ascertain molecular epidemiology studies.

For instance, AFLP has a higher discriminatory power in comparison to PFGE, where it provided greater genetic resolution and is less sensitive to DNA quality during genetic typing of bacterial pathogens in epidemiological investigation. Additionally, AFLP can be automated and has portable results, as well as reproducible approach to facilitate the analysis. Moreover, AFLP it is more suitable for inter-laboratory data exchange using sequence-based data. 16s rRNA analysis is considered a good discrimination approach among unrelated isolates, particularly in *S. aureus* only if it is combined with other gene identification such as *nuc* and *meca*. SCCmec is also very useful and important molecular tool in understanding the epidemiology of methicillin resistance as well as supporting the clonal strain relatedness. DiversiLab rep-PCR tool is very useful to identify MRSA in the hospital outbreaks. In contrast, it is reported that rep-PCR is not highly discriminative tool for MRSA typing particularly in outbreak setting. RFLP can be used in routine infection control program in health care systems as well as epidemiological investigation. It has a good discriminatory power, typeability and reproducibility in MRSA typing. Spa typing is a based sequence typing method, where its results are easy to interpret, less time-consuming, highly reproducible, less laborious and highly comparable between laboratories via ridom.spa.server. MLST is still considered as the rapid method for subtyping for MRSA in clinical research, and has been shown to be useful in global epidemiological studies of *S. aureus*, and the results are comparable between laboratories using MLST server for interpretation. PFGE is found extremely helpful in the short-term investigation and identification of MRSA outbreaks in hospital, community and livestock-associated; however, this method has insufficient resolution power to differentiate bands in identical size as the main drawback of this method. DNA microarray technology is highly accurate, but the reproducibility data needs to be established to the broad application of this technology, and it is not practical if the target of typing is SNPs of highly clonal species. Also, this approach is difficult to identify sequences not included in the array.

Finally, WGS is considered as a very useful tool in epidemiological investigations to discriminate MRSA, and it may assist to trace person-to-person transmission in health care
systems. Additionally, this technique is considered suitable for MRSA epidemics surveillance at national and international levels as well as reproducible approach, which is essential as baseline resources for managing therapeutic treatment and the control of rapid expansion of these strains.

Author details

Mazen M. Jamil Al-Obaidi, Zarizal Suhaili and Mohd Nasir Mohd Desa*

*Address all correspondence to: mnasir@upm.edu.my

1 Faculty of Medicine and Health Sciences, Universiti Putra Malaysia, Selangor, Malaysia
2 Faculty of Bioresources and Food Industry, Universiti Sultan Zainal Abidin, Terengganu, Malaysia

References


[60] Bosch T, Pluister GN, van Luit M, Landman F, et al. Multiple-locus variable number tandem repeat analysis is superior to spa typing and sufficient to characterize MRSA for surveillance purposes. Future Microbiology. 2015;10:1155-1162


[70] Neetu TJP, Murugan S. Genotyping of methicillin resistant *Staphylococcus aureus* from tertiary care hospitals in Coimbatore, South India. Journal of Global Infectious Diseases. 2016;8:68-74


