

A six-member SNP assay on the iPlex MassARRAY platform provides a rapid and affordable alternative for typing major African *Staphylococcus aureus* types

Justin Nyasinga^{1,2}, Cecilia Kyany'a³, Raphael Okoth³, Valerie Oundo³, Daniel Matano³, Simon Wacira³, Willie Sang⁴, Susan Musembi¹ and Lillian Musila^{3,*}

Abstract

Purpose. Data on the clonal distribution of *Staphylococcus aureus* in Africa are scanty, partly due to the high costs and long turnaround times imposed by conventional genotyping methods such as *spa* and multilocus sequence typing (MLST), which means there is a need for alternative typing approaches. This study evaluated the discriminatory power, cost of and time required for genotyping Kenyan staphylococcal isolates using iPlex MassARRAY compared to conventional methods.

Methodology. Fifty-four clinical *S. aureus* isolates from three counties were characterized using iPlex MassARRAY, *spa* and MLST typing methods. Ten single-nucleotide polymorphisms (SNPs) from the *S. aureus* MLST loci were assessed by MassARRAY.

Results. The MassARRAY assay identified 14 unique SNP genotypes, while *spa* typing and MLST revealed 22 *spa* types and 21 sequence types (STs) that displayed unique regional distribution. *spa* type t355 (ST152) was the dominant type overall while t037/t2029 (ST 241) dominated among the methicillin-resistant *S. aureus* (MRSA) isolates. MassARRAY showed 83 % and 82 % accuracy against *spa* typing and MLST, respectively, in isolate classification. Moreover, MassARRAY identified all MRSA strains and a novel *spa* type. MassARRAY had a reduced turnaround time (<12 h) compared to *spa* typing (4 days) and MLST (20 days). The MassARRAY reagent and consumable costs per isolate were approximately \$18 USD compared to *spa* typing (\$30 USD) and MLST (\$126 USD).

Conclusion. This study demonstrated that iPlex MassARRAY can be adapted as a useful surveillance tool to provide a faster, more affordable and fairly accurate method for genotyping African *S. aureus* isolates to identify clinically significant genotypes, MRSA strains and emerging strain types.

INTRODUCTION

Methicillin-resistant *Staphylococcus aureus* (MRSA) is a recognized global threat in the clinical management of infections caused by the pathogen [1, 2], which can cause a wide range of clinical syndromes [3]. Particular strains of MRSA tend to localize in certain geographical regions of the world, for example, USA100 and USA300 in North America [2] and EMRSA15 and EMRSA16 in the United Kingdom [4]. Strain typing can therefore be a useful tool for tracking

high-risk *S. aureus* strains in a region. However, data on the clonal diversity and distribution of MRSA in Africa are scanty and warrant further studies [1, 5, 6]. For example, in Kenya, infections caused by *S. aureus*, including MRSA, are widespread [6–8], yet only two hospital studies have reported on the molecular epidemiology of the pathogen. In the first study, Omuse *et al.* reported the presence of major global MRSA clones such as MLST-CC5, 22, 7 and 30 among a heterogeneous collection of *S. aureus* isolates [6] and the predominance of members of MLST-CC5, such as ST1, ST5, ST8 and

Received 23 November 2018; Accepted 03 March 2019; Published 03 May 2019

Author affiliations: ¹Kenyatta University, Nairobi, Kenya; ²Technical University of Kenya, Nairobi, Kenya; ³United States Army Medical Research Directorate – Africa, Nairobi, Kenya; ⁴Kenya Medical Research Institute, Nairobi, Kenya.

***Correspondence:** Lillian Musila, Lillian.Musila@usamru-k.org

Keywords: *S. aureus*; typing; MassARRAY; *spa*; MLST; Kenya.

Abbreviations: ASPE, Allele-specific primer extension; MALDI-TOF, matrix assisted laser desorption/ionization time of flight mass spectrometry; MLST, multilocus sequence typing; MRSA, methicillin resistant *Staphylococcus aureus*; SAP, Shrimp alkaline phosphatase; SCC*mec*, Staphylococcal cassette chromosome *mec*; SNP, single nucleotide polymorphism; UEPs, unextended primers.

One supplementary table is available with the online version of this article.

000018 © 2019 The Authors

This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

ST241. The second study reported that *spa* type t037 (ST239) was the dominant strain among clinical MRSA isolates [8].

High costs, long turnaround times for results, technical complexity and lack of technical expertise have impeded the adoption of conventional typing methods for surveillance studies, especially in resource-limited settings such as Africa [1, 5, 6]. In recent years, SNP-based genotyping has been proposed as a robust, efficient and low-cost approach to epidemiological characterization of *S. aureus* using methods such as iPlex MassARRAY [9, 10]. Despite its potential, iPlex MassARRAY has not been extensively evaluated against conventional methods such as *spa* typing and MLST, or against *S. aureus* strains from Africa whose epidemiology has been thought to be unique [5].

In this study, the results of a multiplexed SNP assay on the MassARRAY platform for a collection of Kenyan clinical *S. aureus* isolates are reported. In addition, comparisons between iPlex MassARRAY, *spa* and MLST with regard to discriminatory power, assay turnaround times, and reagent and consumable costs are presented.

METHODS

Study population and isolates

This study analysed 54 archived clinical *S. aureus* isolates that had been recovered from samples collected sequentially between 2015 and 2016 from a recruited patient population of 221 persons. The isolates were obtained as part of an ongoing antimicrobial resistance surveillance project covering hospitals in three counties in Kenya where patients ≥ 2 months presenting with blood, urine, throat or skin and soft tissue infections are recruited. Standard microbiological techniques were used for the culture, isolation and identification of *S. aureus*. Isolate identities were further confirmed by conventional PCR for the *femA* gene [11]. Antimicrobial susceptibility testing was performed using both the Siemens MicroScan WalkAway 96 plus (Siemens Healthcare Diagnostics, Inc., New York, USA) and Kirby–Bauer disc diffusion methods. MRSA isolates were tested for the presence of the *mecA* and *SCCmec* types [12].

Of the 54 isolates, 16 were from Nairobi county, while Kisumu and Kericho counties were each represented by 19 isolates. Eighty-seven per cent (47/54) of the isolates were associated with community-acquired infections, defined as infections detected in the patient at presentation or within 48 h of hospitalization. Skin and soft tissue infections (SSTIs) contributed 94 % (51/54) of all isolates, while urinary tract infections accounted for 5.5 % (3/54). Eleven per cent (6/54) of the isolates were classified as MRSA. All MRSA isolates were *mecA*-positive and 4/6 had the *SCCmec* type IV genotype. All of the isolates were preserved as glycerol stocks at -80°C .

DNA extraction

All of the clinical isolates and four reference isolates [*Staphylococcus aureus* ATCC 43300, ATCC 25923, ATCC 25213 and CO-34 (NR-46191) (American Type Culture Collection

and www.beiresources.org] were sub-cultured by streaking a loopful of inoculum on Muller–Hinton agar plates (Sigma-Aldrich, St Louis, Missouri, USA) and incubating it for 18–24 h at 37°C . The ZR/Fungal/Bacterial DNA MiniPrep kit (Zymo Research, Irvine, California, USA) was used for extraction following the manufacturer's instructions, with two modifications: the processing time for the cell disruption step was set at 10 min and DNA was eluted with 60 μl of elution buffer. DNA was quantified using a Qubit DsDNA quantification kit (Thermo Fisher Scientific, Waltham, Massachusetts, USA). For the iPlex MassARRAY assays, DNA concentrations were normalized to 10 $\text{ng } \mu\text{l}^{-1}$. DNA samples were stored at -20°C .

PCR amplification of the *spa* gene

Published primers were used to amplify the X region of the *spa* gene [13]. Each 25 μl PCR reaction mix included 10.5 μl of sterile nuclease-free water, 12.5 μl of Dream *Taq* mix (Thermo Fisher Scientific, Waltham, Massachusetts, USA) and 0.5 μl (10 pmol) of both the forward and reverse primers with 1 μl (approximately 50 ng) of template DNA. Cycling was performed in a GeneAmp 9700 PCR System (Applied Biosystems, Foster City, California, USA) with the following conditions: initial denaturation at 95°C for 5 min, and 35 cycles of denaturation at 95°C for 45 s, primer annealing at 60°C for 30 s, extension at 72°C for 90 s followed by a final extension at 72°C for 10 min. Amplification products were resolved on a 1.5 % agarose in $1\times$ Tris Acetate EDTA buffer (Sigma-Aldrich, St Louis, Missouri, USA) for 60 min at 95V and visualized with EZ Vision DNA stain (Amresco, Inc., Cleveland, Ohio, USA) using a UV transilluminator (UVP LLC, Upland, California, USA).

PCR amplification of MLST loci

Published primers were used to amplify the MLST loci [14]. Each 25 μl reaction comprised 10.5 μl of sterile nuclease-free water, 12.5 μl of Dream *Taq* mix (Thermo Fisher Scientific, Waltham, Massachusetts, USA) and 0.5 μl (10 pmol) of both forward and reverse primers with 1 μl (approximately 50 ng) of template DNA. Thermocycler conditions were set as: initial denaturation at 95°C for 3 min and 35 cycles of strand denaturation at 95°C for 30 s, primer annealing at 55°C for 30 s and extension at 72°C for 60 s with a final extension at 72°C for 10 min. The amplicons were resolved and visualized as described above for the *spa* gene.

Sanger sequencing of *spa* and MLST amplicons

spa and MLST amplicons were purified using the DNA Clean and Concentrator kit (Zymo Research, Irvine, California, USA) following the manufacturer's instructions with a final elution volume of 30 μl . Both forward and reverse strands were sequenced. Each reaction contained 4 μl of sterile distilled water, 2 μl of 5X Big Dye buffer (Applied Biosystems, Foster City, California, USA), 1 μl (4 μM) of the PCR primer, 1 μl of Big Dye terminator mix (Applied Biosystems, Foster City, California, USA) and 4 μl of amplicon DNA. Cycle sequencing was performed on an ABI 9700 thermocycler with cycling conditions set as: 94°C for 5 min followed by 30 cycles of 94°C for 15 s, 55°C for 30 s and 68°C for 2.5 min with a

Table 1. Primary amplification primer sequences for various SNPs and their expected amplicon sizes

SNP ID*	Forward primer sequence (5'–3')**	Reverse primer sequence (5'–3')**	Amplicon length (bp)
arcC162	GTTGGGTTATCAAATCGTGG	ATAGTGATAGAACTGTAGGC	96
arcC210	GAGTCTGGCTGTTCTTTTGG	GATAAAGATGATCCACGATTC	118
aroE132	GGCTTTAATATCACAATTCC	GCACCTGCATTAATCGCTTG	103
aroE252	TCTGGATAAACGCTGTGCAA	AAGATGGCAAGTGGATAGGG	93
gmk129	AACTAGGGATGCGTTTGAAG	GGTGTACCATAATAGTTGCC	104
pta294	GATTAGTAAGTGGTGCAGCG	CCAGATGTTTTTGTAAACGCC	114
tpi36	ATTCCACGAAACAGATGAAG	ACGCTCTTCGTCTGTTTCAC	120
tpi241	ATGAGCCAATCTGGGCAATC	GTTTGACGTACAAATGCACAC	102
tpi243	TGAACCAATCTGGGCAATCG	GTTTGACGTACAAATGCACAC	101
yqil333	CAACCATTGATTGATGTCCC	GGCGGTATGGAGAATATGTC	105

* arcC, carbamate kinase; aroE, shikimate dehydrogenase; gmk, guanylate kinase; pta, phosphate acetyl transferase; tpi, triosephosphate isomerase; yqil, acetyl co-enzyme A acetyl transferase. ** All amplification primers had a 5' 10-mer tag.

final extension of 68 °C for 3 min. Sequencing fragments were purified using Sephadex G50 resin (Sigma-Aldrich, St Louis, Missouri, USA) before loading on the Applied Biosystems 3500 Genetic Analyzer.

Iplex MassARRAY assays for MLST SNPs

The 10 MLST SNPs used in this study have been reported previously [9]. Nine of the 10 MLST SNPs assayed were bi-allelic, while 1 (arcC210) was tri-allelic. The amplification and extension primers were designed using Agena Bioscience Assay Design Suite version 2.0 (Agena Bioscience, Hamburg, Germany). The sequences used to design the primers for primary amplification and extension PCR reactions were downloaded from the staphylococcal MLST database, <http://www.mlst.net> (20th June 2016) [14]. The primer sequences are shown in Table 1. All primary PCR primers had a 10-nucleotide tag (ACGTTGGATG) on the 5' end to exclude them from the 4500–9000 Da mass range of matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) detection. The primary multiplex PCR reactions were run on a GeneAmp 9700 PCR machine (Applied Biosystems, Foster City, California, USA). Each reaction mix contained 1× PCR Buffer, 0.1 μM of forward and reverse primers, 4 mM of MgCl₂, 500 μM of dNTP mix, 0.5 units of *Taq* polymerase and 10 ng of DNA template. The PCR conditions were set as follows: initial denaturation at 95 °C for 2 min, 25 cycles of 95 °C for 30 s, 56 °C for 30 s and 72 °C for 60 s, followed by a final extension at 72 °C for 5 min.

Shrimp alkaline phosphatase (SAP) enzyme was used to dephosphorylate unused dNTPs from the primary PCR reaction before the second allele-specific primer extension (ASPE) PCR. For ASPE reactions, 2 μl of the reconstituted extension primer cocktail was added to each reaction well and the following conditions were set: initial denaturation at 94 °C

for 30 s, 40 cycles of one step at 94 °C for 5 s with 5 sub-cycles of 52 °C for 5 s and 80 °C for 5 s, and a final extension at 72 °C for 3 min. The sequences and masses for unextended primers (UEPs) and extension products for each SNP and SNP allele are shown in Table 2. Extension products were conditioned using a resin, after which 10 nl was dispensed into a 96-well spectroCHIP using the MassARRAY Nanodispenser RS1000 (Agena Bioscience, Hamburg, Germany).

Costing and time comparisons

The reagent and consumable costs per isolate for *spa* typing and MLST were estimated based on local pricing for the items in each step in the processes: DNA extraction, primer sequences, PCR amplification, gel electrophoresis, PCR clean-up, cycle sequencing, fragment purification and fragment analysis. For the iPlex MassARRAY, additional costs for primer sequences, primary and allele-specific PCR amplifications, SAP treatment, sample conditioning and liquid transfer to spectrochips were estimated. Assay time comparisons for the three methods were also made. Technician time was not computed in the cost analysis, but it can be inferred from the assay times.

Data analysis

Raw sequence chromatograms were examined using Chromas version 2.6.2 (Technelysium Pty Ltd, Brisbane, Australia) before consensus sequences were created using BioEdit Sequence Alignment Editor version 7.2.5 [15]. *spa* types were assigned using the online *spa* type finder/identifier software, <http://spatyper.fortinbras.us/> (Fortinbras Research). *spa* types were further confirmed using the Ridom *spa* Server, <https://www.spaserver.ridom.de/> (Ridom GmbH, Würzburg, Germany) [16]. For MLST loci, consensus sequences were aligned using the online alignment tool, MAFFT version 7

Table 2. Allele-specific extension primer sequences and expected masses for various SNP alleles

SNP ID	Alleles	Unextended primer (UEP) sequence (5'-3')*	UEP mass (Da)	Call 1	Mass 1 (Da)	Call 2	Mass 2 (Da)
arcC162	T/A	CTGTAGGCACAATCGT	4881.2	A	5152.4	T	5208.3
arcC210	C/T/A	cTCCACGATTCAATAACCCAAC	6592.3	C	6839.5	T	6919.4
aroE132	A/G	agGATCTAAATACGGTATGATACG	7424.9	G	7672	A	7752
aroE252	T/A	ACAGATGGTATTGGTTATGT	6202	A	6473.3	T	6529.1
gmk129	C/T	CAGCATATTCTATAAAATGGTCATCTTT	8527.6	T	8798.8	C	8814.8
pta294	A/C	TGCTGGACGTACAGTATC	5514.6	C	5801.8	A	5841.7
tpi36	C/T	agCATTCCATGTTTGGAAAATAGC	7046.6	T	7317.8	C	7333.8
tpi241	G/A	AATGCACACATTTTCATTG	5761.8	G	6009	A	6088.9
tpi243	A/G	CAAATGCACACATTTTCATT	5730.8	G	5978	A	6057.9
yqil333	C/T	gctgGTCAACAACAGTCGCTT	6406.2	C	6653.4	T	6733.3

*The bases in lowercase were incorporated into the oligonucleotide sequences during primer design to prevent the extension products for various SNP alleles from being too close to each other in the mass spectrum.

<https://mafft.cbrc.jp/alignment/server/>. Sequence types were assigned on the MLST database, www.mlst.net [14], using the 'Exact or Nearest Match' option. The SpectroAcquire program was used for data acquisition on the MassARRAY Compact Analyzer (Agena Bioscience, Hamburg, Germany) and detection parameters were set at 10 laser shots per raster position with a threshold of 5 good spectra per sample pad. In estimating the accuracy of iPlex MassARRAY, all isolates were listed with their SNP genotypes and corresponding *spa* and MLST sequence types (STs). *Spa* and MLST classifications were used as the reference methods. The accuracy of iPlex MassARRAY was determined as the proportion of isolates whose SNP genotypes corresponded to particular *spa*/MLST types without ambiguity.

Data availability

Raw chromatograms for *spa* and MLST typing as well as data output from the MALDI-TOF MS Compact Analyzer will be made available upon request.

RESULTS

Distribution of *spa* and MLST types

spa and MLST typing yielded congruent results with 22 *spa* types and 21 STs respectively. The isolates displayed considerable heterogeneity with 16/22 *spa* types and 13/21 sequence types being represented by only one isolate. The three counties showed unique genetic fingerprints with minimal overlaps in isolate composition. Only t355 (ST152) was observed across the three counties with 18 *spa* types and 18 STs associating with specific counties. Kericho had 12 *spa* types (ten STs) while Kisumu had 10 *spa* types (eight STs) and Nairobi had five *spa* types (seven STs). Two novel *spa* types t17841 and t17826 were reported and were both associated with MSSA isolates. Despite the heterogeneity, t355 (ST152), t064 (ST8),

t005 (ST22), t2029/t037 (ST241) were the dominant types. The six MRSA isolates were represented by types: [t2029 (ST241) *n*=1, t037 (ST241) *n*=3, t355 (ST152) *n*=1 and t007 (ST39) *n*=1].

Typability and variability of the MLST SNPs

Table 3 summarizes the characteristics of the 10 SNPs analysed. Nine of the 10 SNPs interrogated were highly typeable, with SNP call rates (number of isolates in which a particular SNP

Table 3. Summary of the 10 MLST SNPs analysed by iPlex MassARRAY

SNP	Polymorphism	SNP call rate ^a	Allele frequency ^b
arcC162	T/A	98.3% (57/58)	T 53% (30/57)
arcC210	C/T/A	86% (50/58)	T 60% (30/50)
aroE132	A/G	91.3% (53/58)	A 73.6% (39/53)
aroE252	T/A	86% (50/58)	T 64% (32/50)
gmk129	C/T	84.5% (49/58)	C 67.3% (33/49)
tpi36	C/T	81% (47/58)	C 63.8% (30/47)
tpi241	G/A	86% (50/58)	G 98% (49/50)
tpi243	A/G	94.8% (55/58)	G 60% (33/55)
yqil333	C/T	89.7% (52/58)	T 84.6% (44/52)
pta294	A/C	—*	—*

*pta294 was only identified in one isolate and was excluded when generating SNP genotypes.

a, SNP call rate: The number of isolates in which a given SNP was identified as a proportion of all isolates tested for that SNP.

b, Allele frequency: The number of isolates in which a given SNP allele was identified as a proportion of all isolates in which the SNP was identified.

was identified as a proportion of the total number of isolates tested) that ranged from 81–98.3 % (average 89%). One SNP, pta294, was only identified in one isolate and therefore could not be used to discriminate between isolates and was subsequently excluded from the analyses. Eight of the remaining SNPs were highly variable, with allele frequencies (number of isolates positive for a given SNP allele as a proportion of all isolates in which that SNP was identified) ranging from 53 % (arcC162) to 84.6 % (yqil333).

Iplex MassARRAY SNP genotypes

The combination of nine SNPs for each isolate was used to generate SNP genotypes. To identify the smallest number of SNPs with the highest resolution, different SNP combinations were created by sequentially and progressively reducing the number of SNPs while comparing the resultant classifications with those of the full nine-member SNPs. First, individual SNPs were excluded to create eight-member SNP classifications. Four SNPs (arcC162, arcC210, gmk129 and yqil333) proved to be individually discriminative, while the remaining five (aroE132, aroE252, tpi36, tpi241 and tpi243) contributed to the resolution by acting in combination (Table S1, available in the online version of this article). Next, different SNP pairs from the five SNPs that had combinatorial resolution were excluded to create seven-member SNP classifications. Six different combinations of seven SNPs had a similar resolution to that of the nine-member classification. Lastly, different sets of three SNPs from the five SNPs with combinatorial resolution were excluded to create six-member SNP groupings. In the end, a minimum of six SNPs was required to achieve the resolution of the nine-member SNP profiles. Two unique six-member SNP combinations were identified (arcC162, arcC210, aroE132, gmk129, tpi243 and yqil333 and arcC162, arcC210, gmk129, tpi36, tpi243 and yqil333). In this paper, classifications based on the former combination are presented.

The assay grouped 44/54 isolates into 14 different SNP genotypes. Fig. 1 shows the regional distribution of various SNP genotypes. In Kericho, nine SNP genotypes were observed among the isolates, while the Kisumu and Nairobi isolates showed eight and four SNP genotypes, respectively. The assay successfully typed 11/16, 16/19 and 17/19 isolates from Nairobi, Kericho and Kisumu, respectively, while 10 isolates did not have complete SNP profiles. The classifications achieved by iPlex MassARRAY reflected the heterogeneity and regional distribution patterns revealed by *spa* typing and MLST.

The 14 SNP genotypes were compared with corresponding *spa* and MLST STs. The iPlex MassARRAY showed 83 % (34/41) accuracy against *spa* typing and 82 % (32/39) against MLST. The SNP genotype TTACAC corresponded to a novel *spa* type t17841. iPlex MassARRAY identified all MRSA isolates in this collection. One MRSA was of the ATACGT genotype, which corresponded to t355. t037, which was the *spa* type of three MRSAs, and t2029 (one MRSA) were represented by the ATACAT genotype. ACGTGT corresponded to t007 (one MRSA).

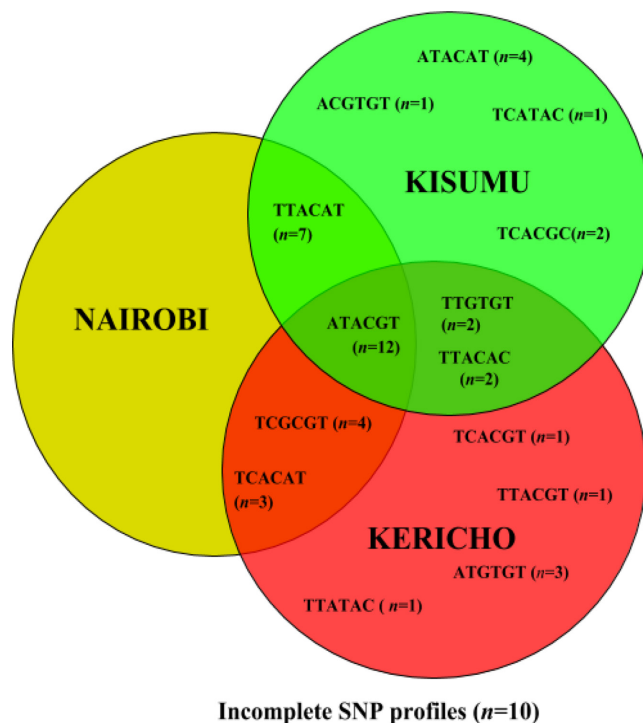


Fig. 1. Distribution of *S. aureus* isolate types in the three counties by MassARRAY SNP genotyping. The six-member SNP genotypes are in the order of arcC162, arcC 210, aroE 132, gmk 129, tpi 243 and yqil 333. Isolates in which at least one of the seven SNPs used for genotype assignment was not identified were considered incomplete.

Three discrepancies in the iPlex MassARRAY assay were observed: SNP genotype TCACGC could not distinguish between t3772 and t13194 (ST 25 and ST80); SNP genotype TTGTGT could not distinguish between t314 and t272 (ST121 and 152); and TCACAT could not distinguish between t084 and t131 (ST 15 and ST1290). Discrepancies involving *spa* types t318/t021 (ST30) and t037/t2029 (ST241) were considered to be minor, as the isolates are closely related and indistinguishable by MLST (Fig. 2).

Turnaround time and reagent and consumable costs per isolate

The iPlex MassARRAY assay took approximately 1 day to complete compared to *spa* typing and MLST, which took approximately 4 and 20 days, respectively (Table 4) to complete. The reagent and consumable costs for analysing an isolate using MassARRAY were approximately \$18 USD compared to \$30 USD/isolate for *spa* typing and \$126 USD/isolate for MLST for the collection of isolates analysed.

DISCUSSION

Six MLST SNPs identified 14 unique genotypes and reflected the heterogeneity and distribution depicted by *spa* typing/MLST of a collection of Kenyan isolates. The isolate types ST152, ST5, ST8, ST15, ST30 and ST241 identified by

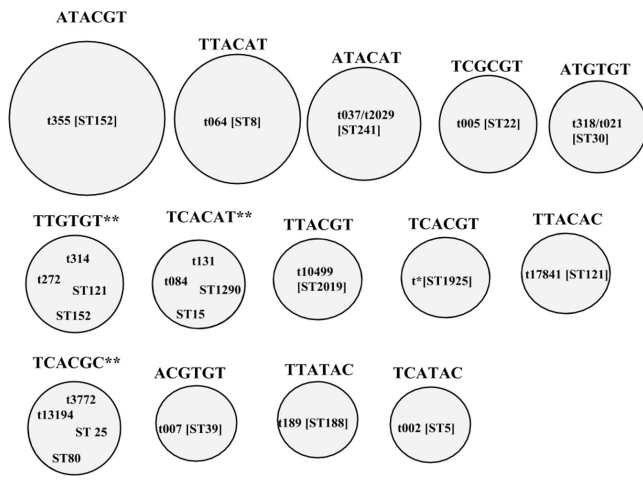


Fig. 2. A comparison of the classifications between iPlex MassARRAY, *spa* typing and MLST. The six-member SNP genotypes indicated above each circle are in the order of arcC162, arcC 210, aroE 132, gmk 129, tpi 243 and yqil 333. The *spa* types are indicated with the prefix t and the MLST types indicated with the prefix ST. The different sizes of the circles represent the frequency of occurrence of various strain types. t*, an unassigned novel *spa* repeat sequence. **, SNP profiles in which discrepant classifications were observed.

MassARRAY have been reported in other African countries, such as Cameroon, Madagascar, Morocco, Niger, Senegal [17] and Ghana [18], highlighting the potential of this assay to be applied in the larger African context.

MassARRAY demonstrated the clear advantages of reduced turnaround time and reagent and consumable costs per isolate based on the time and costs in the Kenyan context, which could vary in different settings. In one study, when the iPlex MassARRAY assay costs were compared to SYBR green real-time PCR, there was a 60 % reduction in reagent costs [9]. Trembizki *et al.* noted an approximately 30 % reduction in costs compared to performing a full MLST analysis [19]. The cost of a MassARRAY platform is higher than that for conventional PCR machines and genetic analysers and therefore prohibitive for a single programme or project to procure, but is best borne as a core institutional capability shared by different projects. The high initial costs for the MassARRAY system are justified by the short turnaround times, multiplexing, automation and throughput capabilities, which can support multiple large-scale studies concurrently.

As the MassARRAY technology is increasingly being adopted, studies utilizing its application in bacterial genotyping are being reported. In one study, a 14-member SNP assay on the MassARRAY platform was developed for genetic characterization of *Neisseria gonorrhoea* in Australia [19] and subsequently applied in a large-scale AMR surveillance study [20]. Even though only 10 SNPs were reported in this study, the technique can be designed to detect more SNPs for increased resolution and even identify binary gene markers to answer clinically important questions about a pathogen, for example regarding virulence and resistance [9].

While the increasing availability of whole-genome sequencing platforms and the concomitant reduction in sequencing costs is recognized, this assay remains superior in terms of the cost and speed of obtaining specific information, such as strain type and presence or absence of resistance or virulence genes, while avoiding the rigours of handling and analysing whole-genome sequencing data. Additionally, the assay would be most relevant for long-term surveillance studies, where it would be impractical to perform *spa* typing, MLST or whole-genome sequencing on hundreds to thousands of isolates.

In some instances, the assay could not distinguish between two or more *spa*/MLST types, an observation noted by others [21, 22]. However, the affected *spa*/MLST types did not belong to major circulating strains, as they were each represented by one isolate, none of which was an MRSA strain. Previously, it has been suggested that increasing the number of SNPs can resolve minor *S. aureus* clones [22]. While the potential of a larger SNP set to improve resolution is recognized, a different set of SNPs in addition to the six identified here may be required.

There were instances, for example with the pta294 SNP, where particular SNPs could not be called. A possible explanation for this could be sequence variation in the target locus for the primer in the primary PCR reactions, an observation that has been noted elsewhere [9, 19]. The quality of DNA

Table 4. Assay time comparisons for *spa*, MLST and MassARRAY typing methods

Step	<i>spa</i> typing		MLST	
	Time	Period	Time	Period
Amplification PCR	3 h	1 day	21 h	2 days
Gel electrophoresis	2 h		7 h	1 day
PCR amplicon clean-up	2 h		14 h	2 days
Sequencing PCR	3 h	1 day	21 h	2 days
Presequencing purification	3 h		21 h	2 days
Sequence fragment analysis	8 h	1 day	56 h	7 days
Data analysis	7 h	1 day	30 h	4 days
Total time	4 days		20 days	
iPlex MassARRAY				
Step	Time	Period		
Amplification PCR	2.5 h			
SAP treatment	0.75 h			
iPlex PCR	3 h			
Sample conditioning	1 h			
Data acquisition	1 h			
Data analysis	3 h			
Total time	1 day			

can potentially affect the success of the assay [19], however, the DNA used for these experiments was of high quality, as extraction was performed using a commercial kit and DNA concentrations were measured and normalized.

Even though the SNPs analysed had generally high identification rates, the combinatorial nature of the method, which requires all SNP results for genotype assignment, meant that 10 isolates could not be classified despite having most of the alleles. This could be solved by further optimization of assay conditions and rescreening of the missing loci to complete the panel for classification.

The lack of a database for MLST SNPs that can be used for matching particular SNP profiles to known STs or *spa* types [19] is a major limitation. This is partly due to the lack of rigorously evaluated MassARRAY SNP data against known *spa* /STs from an international collection of isolates. With increased validation, it should be possible to have an online MLST-style platform where SNPs can be submitted for inter-laboratory comparability of data.

Compelling epidemiological conclusions cannot be drawn from this study due to the modest collection of 54 isolates analysed. A larger collection of isolates from diverse regions and clinical syndromes would not only give a reliable reflection of the epidemiology of *S. aureus* in Kenya but also serve to validate the utility of iPlex MassARRAY for surveillance.

In conclusion, six SNPs derived from the MLST loci provided comparable discriminatory power for resolving a heterogeneous and regionally unique collection of Kenyan clinical *S. aureus* isolates, including MRSA strains. The iPlex MassARRAY demonstrated the advantages of reduced turnaround time and assay costs compared to two conventional typing methods. With increased validation, the assay could serve as a complement to existing typing methods in *S. aureus* surveillance studies to establish the epidemiology of *S. aureus* strains and monitor for the presence or emergence of high-risk strains in a timely fashion.

Funding information

This work was funded by the Armed Forces Health Surveillance Branch (AFHSB) and its Global Emerging Infections Surveillance (GEIS) Section. PROMIS ID: 20160270153 FY17.

Acknowledgement

We acknowledge the use of the *S. aureus* MLST database which is located at Imperial College London and is funded by the Wellcome Trust. The following reagent was provided by the Network on Antimicrobial Resistance in *Staphylococcus aureus* (NARSA) for distribution by BEI Resources, NIAID, NIH: *Staphylococcus aureus*, Strain CO-34, NR-46191. This study made use of the *spa* typing website (<http://www.spaserver.ridom.de/>) that is developed by Ridom GmbH and curated by SeqNet.org (<http://www.SeqNet.org/>). We appreciate technical support from Luíser Ingasia, the MDR Laboratory and Aron Abera. We thank the study participants in the AMR Surveillance Project. This article has been published with permission from the Director, KEMRI. **Disclaimer:** Material has been reviewed by the Walter Reed Army Institute of Research. There is no objection to its presentation and/or publication. The opinions or assertions contained herein are the private views of the author, and are not to be construed as official, or as reflecting true views of the

Department of the Army or the Department of Defense. The investigators have adhered to the policies for protection of human subjects as prescribed in AR 70-25.

Author contributions

Conceptualization, L.M., J.T., Data curation, L.M., C.K., J.T., and R.O.; Formal analysis, J.T., R.O. and C.K.; Funding acquisition, L.M.; Investigation, J.T., C.K., R.O., V.O., D.M. and S.W.; Methodology, J.T., R.O., and C.K.; Supervision, L.M., S.M. and W.S.; Writing- original draft, J.T.; Writing-review & editing, L.M., W.S., S.M., C.K. and R.O.

Conflicts of interest

The authors declare that there are no conflicts of interest.

Ethical statement

This study was reviewed and approved by the Kenya Medical Research Institute (KEMRI) Scientific and Ethics Review Unit (KEMRI/SERU/CCR/0061/3448) and by the WRAIR IRB (WRAIR #2089A). Isolates used in the study were from persons that had provided informed consent to participate in the study.

References

1. Abdulgader SM, Shittu AO, Nicol MP, Kaba M. Molecular epidemiology of Methicillin-resistant *Staphylococcus aureus* in Africa: a systematic review. *Front Microbiol* 2015;6:348.
2. DeLeo FR, Otto M, Kreiswirth BN, Chambers HF. Community-associated methicillin-resistant *Staphylococcus aureus*. *The Lancet* 2010;375:1557–1568.
3. Loughman JA, Fritz SA, Storch GA, Hunstad DA. Virulence gene expression in human community-acquired *Staphylococcus aureus* infection. *J Infect Dis* 2009;199:294–301.
4. Holmes A, McAllister G, McAdam PR, Hsien Choi S, Girvan K *et al.* Genome-wide single nucleotide polymorphism-based assay for high-resolution epidemiological analysis of the methicillin-resistant *Staphylococcus aureus* hospital clone EMRSA-15. *Clin Microbiol Infect* 2014;20:0124–0131.
5. Schaumburg F, Alabi AS, Peters G, Becker K. New epidemiology of *Staphylococcus aureus* infection in Africa. *Clin Microbiol Infect* 2014;20:589–596.
6. Omuse G, Van Zyl KN, Hoek K, Abdulgader S, Kariuki S *et al.* Molecular characterization of *Staphylococcus aureus* isolates from various healthcare institutions in Nairobi, Kenya: a cross sectional study. *Ann Clin Microbiol Antimicrob* 2016;15:51.
7. Akoru C, Kuremu RT, Ndege SK, Obala A, Smith JW *et al.* Prevalence and anti-microbial susceptibility of methicillin *Staphylococcus aureus* at Moi Teaching and Referral Hospital Eldoret. *OJMM* 2016;6:9–16.
8. Aiken AM, Mutuku IM, Sabat AJ, Akkerboom V, Mwangi J *et al.* Carriage of *Staphylococcus aureus* in Thika Level 5 Hospital, Kenya: a cross-sectional study. *Antimicrob Resist Infect Control* 2014;3:22.
9. Syrmis MW, Moser RJ, Whitley DM, Vaska V, Coombs GW *et al.* Comparison of a multiplexed MassARRAY system with real-time allele-specific PCR technology for genotyping of methicillin-resistant *Staphylococcus aureus*. *Clin Microbiol Infect* 2011;17:1804–1810.
10. Struelens MJ, Hawkey PM, French GL, Witte W, Tacconelli E. Laboratory tools and strategies for methicillin-resistant *Staphylococcus aureus* screening, surveillance and typing: state of the art and unmet needs. *Clin Microbiol Infect* 2009;15:112–119.
11. Kobayashi N, Wu H, Kojima K, Taniguchi K, Urasawa S *et al.* Detection of *mecA*, *femA*, and *femB* genes in clinical strains of *Staphylococcus aureus* using polymerase chain reaction. *Epidemiol Infect* 1994;113:259–266.
12. Milheiro C, Oliveira DC, de Lencastre H. Update to the multiplex PCR strategy for assignment of *mec* element types in *Staphylococcus aureus*. *Antimicrob Agents Chemother* 2007;51:3374–3377.
13. Shopsin B, Gomez M, Montgomery SO, Smith DH, Waddington M *et al.* Evaluation of protein A gene polymorphic region DNA sequencing for typing of *Staphylococcus aureus* strains. *J Clin Microbiol* 1999;37:3556–3563.

14. Enright MC, Day NP, Davies CE, Peacock SJ, Spratt BG. Multilocus sequence typing for characterization of methicillin-resistant and methicillin-susceptible clones of *Staphylococcus aureus*. *J Clin Microbiol* 2000;38:1008–1015.
15. Hall TA. BioEdit: a user-friendly biological sequence alignment editor and analysis program for windows 95/98/NT. *Nucleic acids symposium series*. London: Information Retrieval Ltd; 1999. pp. c1979–c2000.
16. Harmsen D, Claus H, Witte W, Rothgänger J, Claus H *et al.* Typing of methicillin-resistant *Staphylococcus aureus* in a university hospital setting by using novel software for *spa* repeat determination and database management. *J Clin Microbiol* 2003;41:5442–5448.
17. Breurec S, Fall C, Pouillot R, Boisier P, Brisse S *et al.* Working group on *Staphylococcus aureus* infections, Garin B, Laurent F. Epidemiology of methicillin-susceptible *Staphylococcus aureus* lineages in five major African towns: high prevalence of Panton-Valentine leukocidin genes. *Clin Microbiol Infect* 2011;17:633–639.
18. Egyir B, Guardabassi L, Sørum M, Nielsen SS, Kolekang A *et al.* Molecular epidemiology and antimicrobial susceptibility of clinical *Staphylococcus aureus* from healthcare institutions in Ghana. *PLoS One* 2014;9:e89716.
19. Trembizki E, Smith H, Lahra MM, Chen M, Donovan B *et al.* High-throughput informative single nucleotide polymorphism-based typing of *Neisseria gonorrhoeae* using the Sequenom MassARRAY iPLEX platform. *J Antimicrob Chemother* 2014;69:1526–1532.
20. Whiley DM, Trembizki E, Buckley C, Freeman K, Baird RW *et al.* Molecular antimicrobial resistance surveillance for *Neisseria gonorrhoeae*, Northern Territory, Australia. *Emerg Infect Dis* 2017;23:1478–1485.
21. Robertson GA, Thiruvengataswamy V, Shilling H, Price EP, Huygens F *et al.* Identification and interrogation of highly informative single nucleotide polymorphism sets defined by bacterial multilocus sequence typing databases. *J Med Microbiol* 2004;53:35–45.
22. Stephens AJ, Huygens F, Inman-Bamber J, Price EP, Nimmo GR *et al.* Methicillin-resistant *Staphylococcus aureus* genotyping using a small set of polymorphisms. *J Med Microbiol* 2006;55:43–51.

Five reasons to publish your next article with a Microbiology Society journal

1. The Microbiology Society is a not-for-profit organization.
2. We offer fast and rigorous peer review – average time to first decision is 4–6 weeks.
3. Our journals have a global readership with subscriptions held in research institutions around the world.
4. 80% of our authors rate our submission process as 'excellent' or 'very good'.
5. Your article will be published on an interactive journal platform with advanced metrics.

Find out more and submit your article at microbiologyresearch.org.