

RESEARCH ARTICLE

Open Access

Comparative genomics of methicillin-resistant *Staphylococcus aureus* ST239: distinct geographical variants in Beijing and Hong Kong

Zheng Wang¹, Haokui Zhou¹, Hui Wang², Hongbin Chen², K K Leung³, Stephen Tsui³ and Margaret Ip^{1*}

Abstract

Background: The ST239 lineage is a globally disseminated, multiply drug-resistant hospital-associated methicillin-resistant *Staphylococcus aureus* (HA-MRSA). We performed whole-genome sequencing of representative HA-MRSA isolates of the ST239 lineage from bacteremic patients in hospitals in Hong Kong (HK) and Beijing (BJ) and compared them with three published complete genomes of ST239, namely T0131, TW20 and JKD6008. Orthologous gene group (OGG) analyses of the Hong Kong and Beijing cluster strains were also undertaken.

Results: Homology analysis, based on highest-percentage nucleotide identity, indicated that HK isolates were closely related to TW20, whereas BJ isolates were more closely related to T0131 from Tianjin. Phylogenetic analysis, incorporating a total of 30 isolates from different continents, revealed that strains from HK clustered with TW20 into the 'Asian clade', whereas BJ isolates and T0131 clustered closely with strains of the 'Turkish clade' from Eastern Europe. HK isolates contained the typical ϕ SP β -like prophage with the *SasX* gene similar to TW20. In contrast, BJ isolates contained a unique 15 kb PT1028-like prophage but lacked ϕ SP β -like and ϕ SA1 prophages. Besides distinct mobile genetic elements (MGE) in the two clusters, OGG analyses and whole-genome alignment of these clusters highlighted differences in genes located in the core genome, including the identification of single nucleotide deletions in several genes, resulting in frameshift mutations and the subsequent predicted truncation of encoded proteins involved in metabolism and antimicrobial resistance.

Conclusions: Comparative genomics, based on *de novo* assembly and deep sequencing of HK and BJ strains, revealed different origins of the ST239 lineage in northern and southern China and identified differences between the two clades at single nucleotide polymorphism (SNP), core gene and MGE levels. The results suggest that ST239 strains isolated in Hong Kong since the 1990s belong to the Asian clade, present mainly in southern Asia, whereas those that emerged in northern China were of a distinct origin, reflecting the complexity of dissemination and the dynamic evolution of this ST239 lineage.

Keywords: MRSA, Beijing, Hong Kong, Genome, Orthologous gene groups, ST239

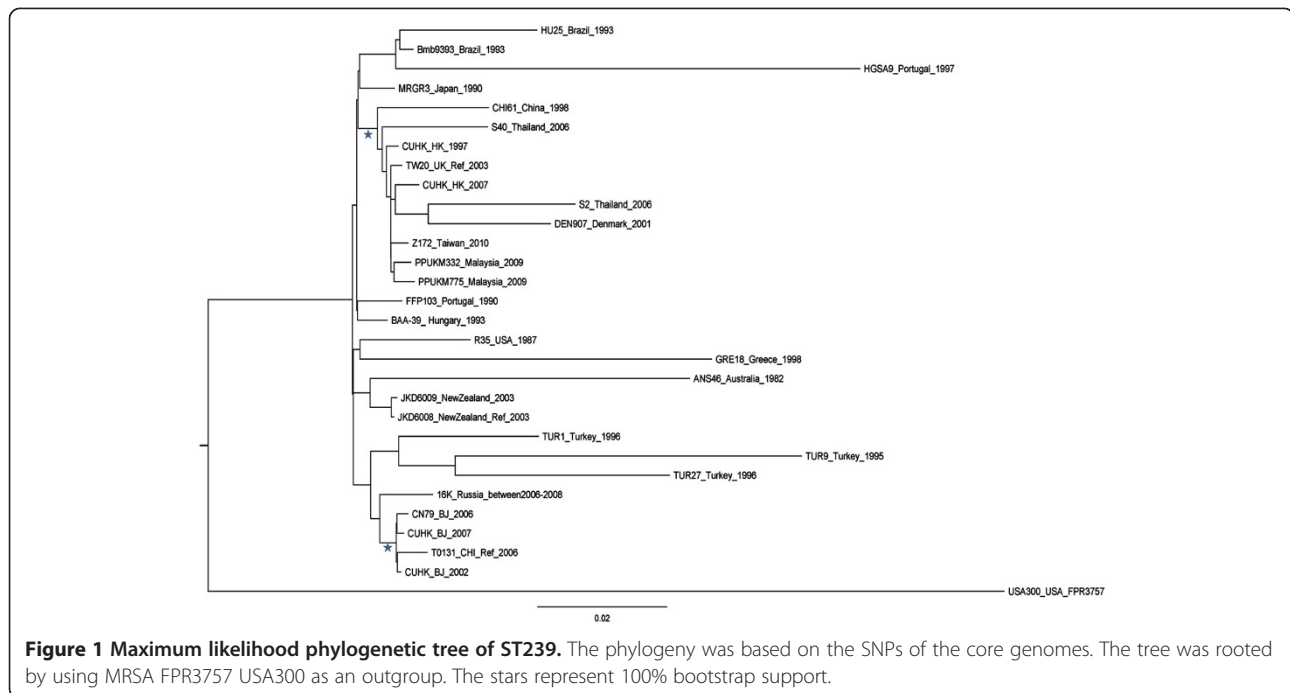
Background

The ST239 lineage of methicillin-resistant *Staphylococcus aureus* (MRSA) is one of the most widely disseminated hospital-associated MRSA (HA-MRSA) [1], which has caused multiple epidemics around the world in recent decades. In China, as in most Asian countries, ST239-SCC*mecIII* has been identified as the predominant clone, accounting for around 75% of observed HA-MRSA

[2,3]. In Hong Kong, ST239-SCC*mecIII* was the most prevalent MRSA clone during the late 1980s and the 1990s [4,5]. However, although it remains one of the most common clones in hospitals in Hong Kong, its prevalence amongst ST239 strains has decreased from the early 2000s [6]. Whole-genome sequencing in clinical microbiology has revolutionized our understanding of MRSA, in areas such as outbreak investigation [7,8], evolutionary and phylogeographic distribution and in recombination studies [9,10]. Harris *et al.* [9] described the comparative genomics, by the reads-mapping method, of globally collected ST239 strains using the TW20 genome

* Correspondence: margaretip@cuhk.edu.hk

¹Department of Microbiology, The Chinese University of Hong Kong, The Prince of Wales Hospital, Ngan Shing Street, Shatin, Hong Kong
Full list of author information is available at the end of the article



T0131. The ϕ SP β -like (TW/20) prophage is considered to be a feature of the ST239 ‘Asian clade’ and it is therefore similarly detected in other ‘Asian clade’ ST239 strains (CHI61, S2, S40, TW20 and DEN907). These strains possess the *SasX* gene, which is located at the 3’ end of the ϕ SP β -like prophage and plays a key role in MRSA colonization, making it a crucial pathogenicity determinant in the spread of ST239 [1]. The *SasX* gene was absent in the BJ cluster.

Another notable difference between the two clusters was the absence of the 43.4 kb prophage ϕ SA1 in the BJ cluster. Both HK genomes contained a 7.29 kb deletion within ϕ SA1 (Figure 4). It has previously been noted that the ϕ SA1 region does not carry any known virulence factors [12]. Both clusters possessed the 44.7 kb prophage ϕ SA3 but some differences were observed, especially in the region between the gene encoding the

transcriptional activator RinB (SAT0131_02106), to the 3’ end (Figure 5). In this region, the BJ cluster contained the additional coding sequences (CDS): *mvaA*, ϕ PVL hypothetical protein, metallo-beta-lactamase superfamily domain-containing protein and *feoB* Ss-1,3-N-acetylglucosaminyl transferase. The proteins encoded by these CDS are involved in antibiotic resistance, virulence and metabolism. BLAST analysis indicated that the *feoB* Ss-1,3-N-acetylglucosaminyl transferase gene was only identified previously in two Australian strains in the NCBI database: JKD6008 (MRSA ST239 clone) and JKD 6159 (MRSA ST93 clone). The metallo-beta-lactamase superfamily domain-containing protein gene has been found in a number of different phages and MRSA clones, such as *Staphylococcus* phages JS01, SP5 and phiNM3; and MRSA clones JKD6008 (MRSA ST239 clone), MRSA252 (MRSA ST36 clone), and 71193 (MRSA ST398 clone).

Table 2 Genome sequencing and contig assembly statistics

Parameters	HK97	HK07	BJ02	BJ07
Total no. of reads	14273328	21087930	12905538	14503760
No. of assembled reads	13807870	20384885	12631772	14335583
No. of contigs	125	81	89	81
Total no. of bases	1284599520	1897913700	1161498420	1305338400
Average contig size (bases)	24068	37011	32465	35090
N50 contig size (bases)	53452	109354	79782	79076
Largest contig size (bases)	167992	206466	310965	258620
Coverage depth	424	629	400	457

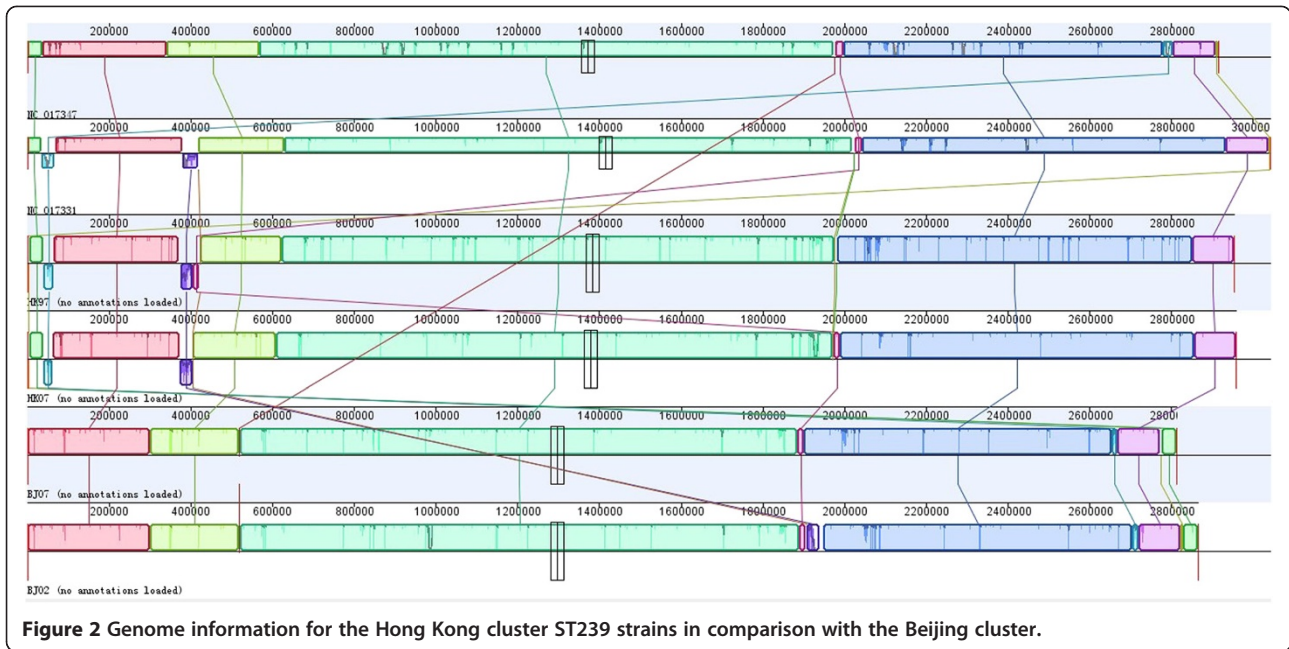


Figure 2 Genome information for the Hong Kong cluster ST239 strains in comparison with the Beijing cluster.

TW20 also harboured a similar gene, but differed by 17 SNPs, and none of the HK strains harboured this gene. For the *mvaA* ϕ PVL hypothetical protein gene, similar genes were found in JKD6008 (ST239 clone), JKD6159 (MRSA ST93 clone), MW2 (ST1 clone) and MSSA476 (MSSA ST1 clone). All six genomes possessed other virulence-associated genes, including the phospholipase C gene, staphylococcal complement inhibitor SCIN, staphylokinase and enterotoxin A [13].

A noteworthy difference in the BJ cluster, absent in the HK strains, was a 15.5 kb region insertion in the gene encoding the 30S ribosomal protein S18, *rpsR* (SATW20_04350). BLAST and PFAST identified that

this region shows high similarity with phage PT1028. This PT1028-like prophage contains important functional genes, such as *sasD*, integrase, *polA*, *deoD1*, SaPI1 and others encoding pathogenicity island proteins (Figure 6).

The analysis also demonstrated other divergences at the MGE level. For example, the pathogenicity island SaPI1 (TW20), which contains enterotoxins K and Q, was absent in all of the BJ cluster strains. Similar to TW20, HK07 carried the SaPI1 island (100% homologous), but it was absent in HK97 (Detailed structures are illustrated in Additional file 1). An 18 kb insertion in T0131 in the chromosome downstream of the *rnr* putative ribonuclease gene, containing an exfoliative toxin A/B gene, was not

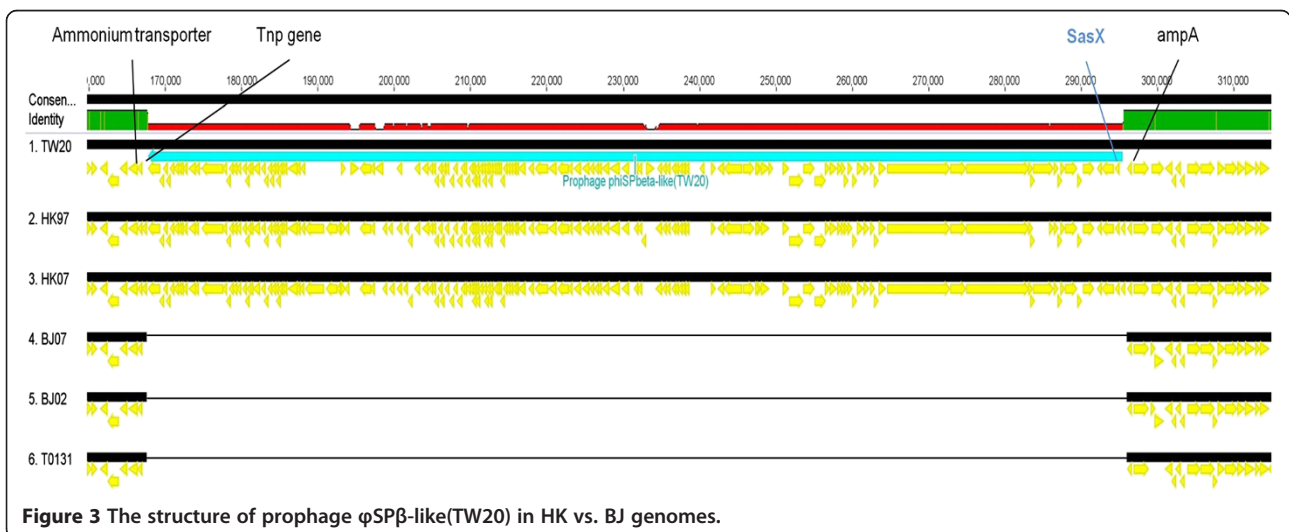
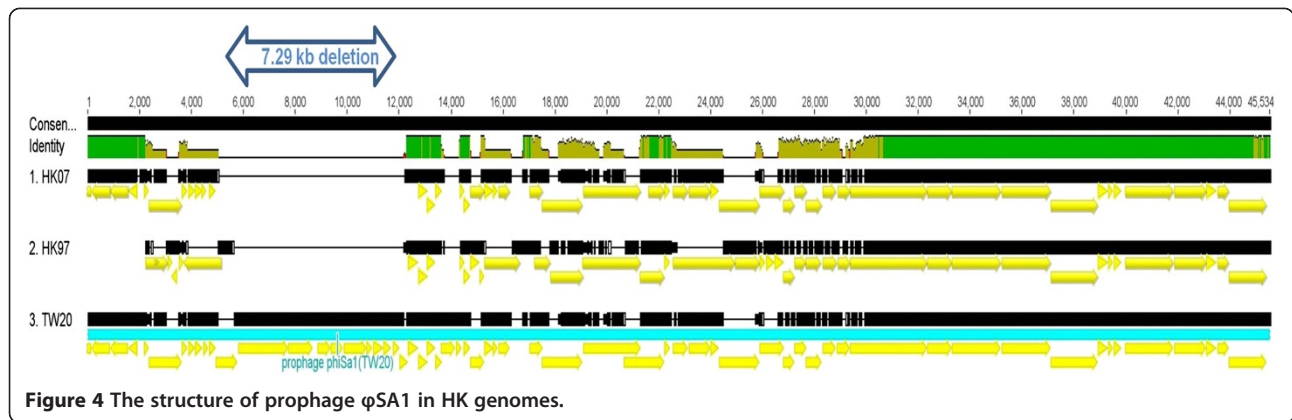


Figure 3 The structure of prophage ϕ SP β -like(TW20) in HK vs. BJ genomes.



present in the HK cluster. Both the BJ strains contained a 6.9 kb fragment of this 18 kb insertion region, but not the virulence genes (Additional file 2).

There are also unique features of individual strains. For example, PFAST indicated a specific 47.5 kb region in BJ02, which showed the greatest similarity with the prophage ϕ NM1 (Newman strain). This ϕ NM1-like prophage (BJ02) harbored two of the three virulence genes, *i.e.*, homologs of SAV0866 and SAV1978 [14], but not homologs of the SAV0862 gene. The detailed structure of this ϕ NM1-like prophage (BJ02) is included in Additional file 3.

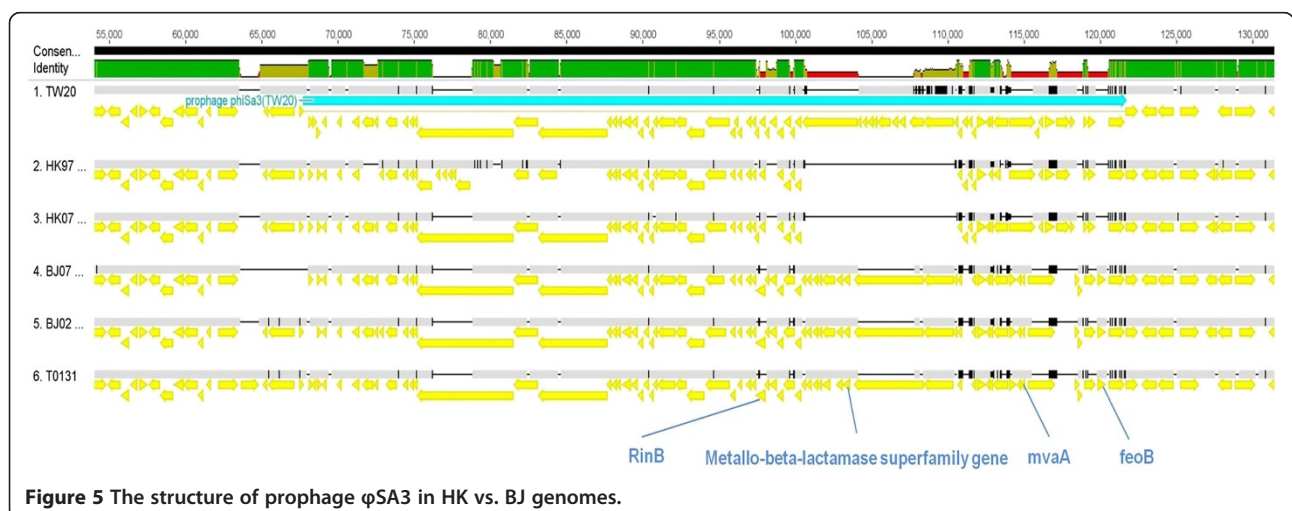
OrthoMCL analysis

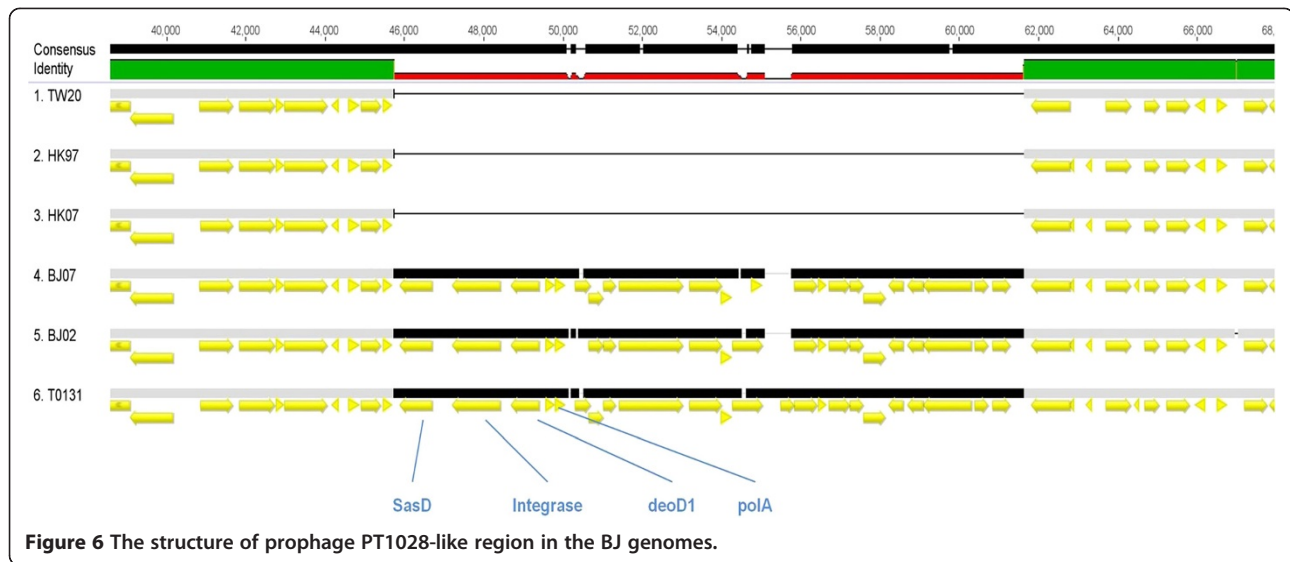
With annotation information and the orthoMCL algorithm, the families of orthologous gene groups were calculated by comparison with TW20 and T0131. The protein sequences of the latter reference genomes were derived from NCBI annotations. The results are illustrated in Venn diagrams (Figure 7). The HK cluster strains shared 2,716 common orthologous gene groups

(OGGs) (Figure 7A). The BJ cluster strains shared 2,588 common OGGs (Figure 7B). BJ02 contained more unique OGGs not present in the other two genomes. Figure 7C shows that 2,483 common OGGs were shared by all six genomes.

Forty-seven OGGs were common in the BJ cluster but were not present in any of the HK cluster genomes. At least 27 of 47 OGGs were related to genomic islands (GIs). Sixteen orthologous groups were composed of hypothetical proteins and pseudogenes, whereas four OGGs consisted of genes orthologous to genes for the following representative proteins from T0131: histidine ammonia-lyase; myosin-cross-reactive streptococcal antigen homolog; methicillin-resistance-related protein *FmhC*; and multidrug resistance transporter B.

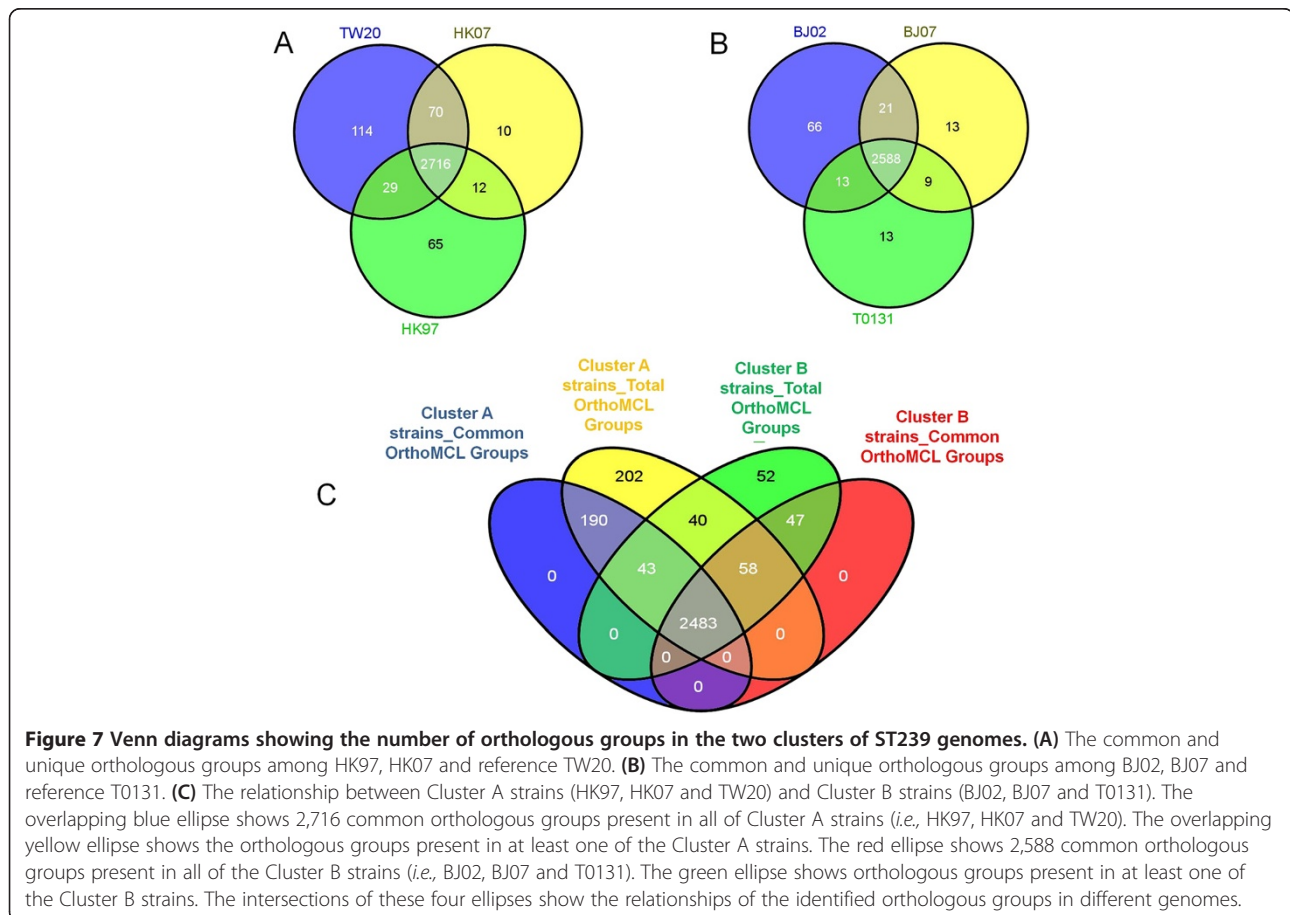
The details of these four OGGs are shown in Additional file 4. Genes in three of the groups (all except the multidrug resistance transporter B group) contained single nucleotide deletions, thus leading to a frameshift mutation and predicted truncation of their encoded proteins. The first group included the gene encoding histidine





ammonia-lyase (SAT0131_00009) and its homologs in BJ02 and BJ07. The 504 aa hutH histidine ammonia-lyase protein (SAR0008) is present in MRSA strains such as MRSA252, JKD6008 and TW20 and plays an important role in histidine metabolism. In the BJ cluster strains, a

'T' deletion occurred at the 1,095th bp of the hutH gene (g.11505delT: NC_017347), causing a frameshift mutation that truncated the hutH into a 382aa length histidine ammonia-lyase protein (SAT0131_00008) and a 120aa protein SAT0131_00009. The putative conserved Lyase-I-like



superfamily domain was detected in this 120 aa protein. Further functional analysis and experimental work are needed to determine whether this has any significant impact on histidine metabolism.

A frameshift mutation (g. 94707delT: NC_017347) was also detected in the BJ cluster strains of the 591 aa myosin-cross-reactive antigen (SAR0111) protein (SAT0131_00089), the representative protein in the second OGG. It is not known what the effects of this mutation might be.

The remaining two groups were related to antimicrobial resistance. Whole-genome alignment of the BJ cluster revealed an 'A' deletion (g.1281606delA: NC_017347) at the 826th bp position of the MW1131 gene (encoding the methicillin resistance-related protein FmhC). This frameshift mutation resulted in a truncated 281 aa protein SAT0131_01298 and a 128 aa length protein SAT0131_01299. The 414 aa length FmhC protein (MW1131) shares identity with FemA and FemB and is classified in the cell wall category [15]. Previously, it has been reported that insertional deactivation of FmhC had no effects on growth, antibiotic susceptibility or the lysostaphin resistance of *S. aureus* strains [16] and its function remains elusive.

In the fourth OGG, multidrug resistance protein B (SAT0131_01869) shared identity with the 393aa multidrug resistance transporter protein B (NWMN_1652), which is present in NEWMAN, COL, USA300_FPR3757, JKD6008 and TW20 MRSA strains. Whether specific local selective pressures by antimicrobials may have contributed to the changes described above, and how these alter antimicrobial resistance, remains to be determined.

The HK cluster genomes contained 190 OGGs that were not present in any of the BJ cluster genomes. Except for 18 groups, which consisted of hypothetical proteins and pseudo genes, all of the other 172 OGGs were associated with genomic islands (GIs). This may reflect the substantial impact of GIs on the evolution of the 'Asian clade' of ST239.

Discussion

The ST239 lineage represents a globally disseminated multi-drug-resistant HA-MRSA. The geographical clustering of strains of the ST239 lineage has been confirmed, both at the continental and national levels [10], and the diversity of European isolates has been discussed in various studies [9,10]. However, Asian isolates were thought to belong to a single, large 'Asian clade', as represented by Thai and Chinese isolates, although some diversity was detected within this clade [9]. Our data revealed geographical clustering – with a diversity of ST239s causing epidemics in hospitals in Beijing and Hong Kong during the late 1990s and the 2000s – and suggested different origins of the ST239s and the

possibility of a more complex distribution in Asia. The HK isolates clustered within the traditional 'Asian clade' and showed a high similarity with TW20, and those present in Taiwan, Thailand and Malaysia. The distinct BJ cluster is more likely to represent the local establishment of an endemic, predominant clone for a number of years, at least from 2002 onwards. According to a recent national surveillance report of MRSA in China, ST239-III-t030s have the same molecular types and characteristics as CUHK_BJ02, and they remain the most predominant HA-MRSA clones in China, with an overall prevalence of 57.1% [2]. The BJ cluster, including the reference genome T0131, showed distinct features, based on their non-synonymous SNPs at the core genome, as well as differences at the MGE level. These observations collectively indicated the possibility that the BJ cluster ST239-III-t030s originated from a relatively recent common ancestor and was disseminated during this period in the ST239 lineage. The CN79 strain, isolated in 2006 and confirmed as a representative ST239 strain in Beijing [17], also clustered within the BJ cluster, and reaffirmed the homogeneity of this cluster. Interestingly, the 'Russia variant' 16K strain [18] also showed close relationship with the BJ cluster and 'Turkish clade', further illustrating the distinct geographical spread of ST239 in the north of China. Further studies are needed, on strains retrieved from older collections and across other parts of China, to estimate the origin and widespread nature of the BJ cluster of ST239.

The reads-mapping assembly method has inherent limitations for MGE detection and non-core genome analysis; this is because MGEs that are absent in the reference genome are not detectable. This disadvantage cannot be overcome, even by using a large number of reference genomes. Thus, in our study, we performed a *de novo* assembly and deep sequencing (more than 400X coverage depth in this study) of each strain. Moreover, three reference genomes were used in the ordering of contigs and the whole-genome alignment process. Recently, an ST239 Russian variant also showed an absence of the ϕ SP β -like (TW20) prophage [18] and this may suggest a potential evolutionary link. However, the unique PT1028-like prophage was not reported in the Russian isolate and the relationship of these strains remains elusive.

The isolates from Hong Kong are representative of the prevalent ST239 clone during the last two decades. ST239 has been the predominant HA-MRSA clone since 1988 and the strains were characterized as multidrug-resistant and prevalent in various hospitals in Hong Kong [4-6]. They are closely related to TW20 of the 'Asian clade' and possessed the ϕ SP β -like prophage and were *SasX*-gene-positive, and they further demonstrated the epidemic wave and dissemination of *sasX*-positive ST239 HA-MRSA in Hong Kong and southern China.

The results of OrthoMCL analysis supported the phylogenetic tree at the protein level. The specific OGGs that were present only in the HK or BJ cluster strains provided further evidence that the horizontal gene transfer of GIs played an important role in ST239 family evolution and geographical clustering. Comparative genomics revealed the common differences between the two clades of ST239 HA-MRSA at SNP, gene and MGE levels. The availability of next generation sequencing on a wider scale will further enhance our understanding of the dynamic evolutionary process in the transmission and spread of globally disseminated multidrug-resistant MRSA.

Conclusions

In summary, comparative genomics, based on *de novo* assembly and deep sequencing, revealed the different origins of the ST239 lineage in northern and southern China and pointed out the common differences between the two ST239 HA-MRSA clades at SNP, gene and protein levels. Besides distinct GIs, which were responsible for the major differences in the two clusters, orthoMCL analyses and whole-genome alignment of the HK and BJ clusters highlighted differences in genes located in the core genome. Single nucleotide deletions, resulting in frameshift mutations, were detected in a number of genes, with the predicted disruption of their encoded proteins, which are known to play an important role in metabolic pathways and antimicrobial resistance. These results reveal the complexity of dissemination and dynamic evolution of the ST239 lineage in China and indicate possible transmission routes. The availability of next generation sequencing technology on a wider scale will further enhance our understanding of the dynamic evolutionary process in the transmission and spread of globally disseminated multidrug-resistant MRSA.

Methods

Bacterial isolates

Four representative ST239 HA-MRSA isolates from bacteremic patients in Hong Kong and Beijing hospitals were selected for whole-genome sequencing. The Hong Kong strains (HK1997 and HK2007) were isolated in 1997 and 2007 at the Dept of Microbiology, Prince of Wales Hospital, Hong Kong, and were representative of strains of indistinguishable PFGE types from a longitudinal surveillance of MRSA strains in Hong Kong in the 1990s to 2007 [4-6]. The Beijing isolates (BJ2002 and BJ2007) were representative of ST239 isolates from bacteremic patients in Beijing in 2002 and 2007, and were strains from the MRSA collection (of Dr H Wang and Dr H Chen) at the Peking Union Medical College Hospital, Beijing. Three complete genomes were used as reference, namely: TW20 [GenBank:FN433596] [19], T0131 [GenBank:CP002643] [11], and JKD6008 [GenBank:

CP002120] [12]. For the phylogenetic analyses, sequence data from another 22 global ST239 genomes (collected between 1982 and 2010) were downloaded from the public databases, NCBI Short Read Archive and Whole Genome Shotgun. The downloaded reads data were mapped to TW20 (using software Geneious 6.1.4, Biomatters Ltd., Auckland, New Zealand) and the downloaded contig data were ordered using TW20 as reference. These strains were 16K [GenBank:BABZ00000000], Bmb9393 [GenBank:CP005288], CN79 [GenBank:ANCI00000000], PPUKM-332-2009 [GenBank:AMRC00000000], PPUKM-775-2009 [GenBank:AMRE00000000], BAA-39 [GenBank:AEEK00000000], MRGR3 [GenBank:AHZL00000000], JKD6009 [GenBank:ABSA00000000], Z172 [GenBank:CP006838] [17,18,20-22] and ANS46, R35, TUR1, TUR9, TUR27, HU25, HGSA9, FFP103, GRE18, S2, CHI61, S40, DEN907 [SRA: ERA000102] [9,10]. The ST8 MRSA genome USA300 FPR3757 [GenBank:CP000255], was used as an outgroup to root the ST239 phylogeny [23]. This genome was aligned with all 29 ST239 strains using the software package Mauve (version 2.3.1), and the SNPs exported were used for phylogenetic analyses. As strains were based on MRSA collections from previous epidemiological studies, no clinical data were obtained nor further ethical submission made.

Molecular typing

The MLST typing was conducted in accordance with the protocol suggested on the MLST website, using seven housekeeping genes (<http://saureus.beta.mlst.net/>) [24]. Staphylococcus protein A (*spa*) typing was carried out by sequencing the PCR product of the *spa* gene, as described, and the *spa* type was confirmed by using the public *spa*-type database and Ridom SpaServer (<http://spa.ridom.de/>; <http://tools.eugenomics.com/>) [25]. SCCmec typing was conducted by multiplex PCR with previously reported primers [26].

Genome sequencing, assembly and annotation

DNA from each MRSA isolate (1 µg for each sample) was extracted with Wizard Genomic DNA Purification Kit (Promega, Madison, USA), following the manufacturer's instructions. Whole-genome sequencing was performed using the HiSeq sequencer (Illumina); unique index-tagged libraries were created for each sample to generate 90 bp paired-end reads. These libraries gave more than 400X coverage (sequencing depth) for each strain on average. *De novo* assembly of the sequencing reads was performed using the Velvet (version 1.2.10) software package [27], coupled with the Velvet Optimiser (version 2.2.4) to select the best kmer length for assembly [28]. The assembly statistics of contigs for each genome are provided in Table 2. The scaffold order of the contigs for each genome was determined by mapping to their

closely related complete reference genomes (TW20, T0131, and JKD6008), using Mauve (version 2.3.1) [29] and Contiguator 2 software [30]. Assembled contigs were submitted to the RAST server (<http://rast.nmpdr.org/>) [31] for genome annotation, and manually inspected using Geneious (version 6.1.4) software (Biomatters Ltd., New Zealand).

Maximum likelihood phylogenetic tree

All genomes of the ST239 representative strains were aligned using the progressive Mauve method with default parameters, and core genome SNPs were retrieved from aligned regions excluding gapped ambiguous columns. The non-core regions were assigned as all sequences that were not present in all 29 ST239 isolates. SNPs were filtered to remove those that were in non-core regions, gaps, and those that included ambiguity codes, and they were finally converted into the phylip format using a PERL script. Maximum likelihood phylogenetic analysis based on core genome SNPs of the ST239 isolates was performed using the RAxML BlackBox [32]. The default CAT model was used, and the ST8 reference genome USA300 FPR3757 [GenBank: CP000255] was included as an outgroup to root the ST239 phylogeny. Supports for nodes were assessed using 100 rapid bootstrap inferences and thereafter by a thorough maximum likelihood search. All free model parameters were estimated by RAxML and likelihood of the final tree was evaluated and optimized under GAMMA.

Mobile genetic element detection and whole genome alignment

The prophage regions were identified by PAST (<http://phast.wishartlab.com>) [33], and the MGE regions in each genome were analyzed by IslandViewer [34]. Whole-genome alignment was performed using Geneious software (Biomatters Ltd., New Zealand) and Mauve (version. 2.3.1) [29] to examine the alignment and distribution of prophage and MGE regions in the genomes.

Gene ortholog analysis

Predicted genes and their translated protein sequences of the four HK and BJ genomes were compared to those of TW20 and T0131, and clustered into ortholog groups using OrthoMCL software [35]. All versus all BLASTP was performed with the default parameter set (an e-value cut-off of 1×10^{-5} , a percent match cut-off of 50%, and an inflation value of 1.5). Common and unique orthologous groups identified among the genomes were analyzed using a Venn diagram [36].

Availability of supporting data

The draft genome sequences of CUHK_HK1997, CUHK_HK2007, CUHK_BJ2002, and CUHK_BJ2007 have been

deposited in the DDBJ/EMBL/GenBank with the accession numbers AZJQ00000000, AZMZ00000000, AZMY00000000, and AZMX00000000, respectively. The versions described in this article are the first versions: AZJQ01000000, AZMZ01000000, AZMY01000000, and AZMX01000000.

Phylogenetic tree data is available in the Dryad Digital Repository (<http://datadryad.org/>), with the following identifier: <http://doi.org/10.5061/dryad.12773>.

Additional files

Additional file 1: Figures S1. The structure of pathogenicity island SaP11 in Beijing Cluster and HongKong Cluster strains.

Additional file 2: Figures S2. The 18 kb insertion structure in Beijing Cluster strains.

Additional file 3: Figures S3. The 47 kb phiNM1-like prophage structure in BJ02 strain.

Additional file 4: Tables S1. The four common functional orthologous gene groups of Beijing cluster strains, which not located on GIs.

Abbreviations

BJ: Beijing; CDS: Coding sequences; GIs: Genomic islands; HA-MRSA: Hospital-associated methicillin-resistant *Staphylococcus aureus*; HK: Hong Kong; MGE: Mobile genetic elements; MRSA: Methicillin-resistant *Staphylococcus aureus*; OGG: Orthologous gene group; SNP: Single nucleotide polymorphism; Spa: *Staphylococcus* protein A.

Competing interests

The authors declare that they have no financial or non-financial competing interests.

Authors' contributions

ZW, HW, HC and MI were involved in the design and execution of the study, and the provision of strains for study. ZW, HZ, KKL and ST performed the laboratory experiments and bioinformatic analyses. ZW, HZ and MI were involved in data analyses. ZW drafted the initial manuscript and all members contributed to the preparation of the final manuscript. All authors read and approved the final manuscript.

Acknowledgments

We would like to thank the Research Fund for the Control of Infectious Diseases (Health and Food Bureau, HKSAR) for financially supporting this study (Commissioned Project No. CU-09-05-01, Principal Investigator: MI). We also acknowledge the use of the MLST website and database, located at Imperial College, London, and funded by The Wellcome Trust.

Author details

¹Department of Microbiology, The Chinese University of Hong Kong, The Prince of Wales Hospital, Ngan Shing Street, Shatin, Hong Kong.

²Department of Clinical Laboratory, Peking University People's Hospital, Beijing 100044, China. ³School of Biomedical Sciences, The Chinese University of Hong Kong, Shatin, Hong Kong.

Received: 23 January 2014 Accepted: 23 June 2014
Published: 26 June 2014

References

- Li M, Du X, Villaruz AE, Diep BA, Wang D, Song Y, Tian Y, Hu J, Yu F, Lu Y, Otto M: MRSA epidemic linked to a quickly spreading colonization and virulence determinant. *Nat Med* 2012, **18**(5):816–819.
- Xiao M, Wang H, Zhao Y, Mao LL, Brown M, Yu YS, O'Sullivan MV, Kong F, Xu YC: National Surveillance of Methicillin-resistant *Staphylococcus aureus* (MRSA) in China Highlights a Still Evolving Epidemiology with Fifteen Novel Emerging Multilocus Sequence Types. *J Clin Microbiol* 2013, **51**(11):3638–3644.

3. Xu BL, Zhang G, Ye HF, Feil EJ, Chen GR, Zhou XM, Zhan XM, Chen SM, Pan WB: **Predominance of the Hungarian clone (ST 239-III) among hospital-acquired methicillin-resistant Staphylococcus aureus isolates recovered throughout mainland China.** *J Hosp Infect* 2009, **71**(3):245–255.
4. Ip M, Lyon DJ, Chio F, Enright MC, Cheng AF: **Characterization of isolates of methicillin-resistant Staphylococcus aureus from Hong Kong by phage typing, pulsed-field gel electrophoresis, and fluorescent amplified-fragment length polymorphism analysis.** *J Clin Microbiol* 2003, **41**(11):4980–4985.
5. Ip M, Lyon DJ, Chio F, Cheng AF: **A longitudinal analysis of methicillin-resistant Staphylococcus aureus in a Hong Kong teaching hospital.** *Infect Control Hosp Epidemiol* 2004, **25**(2):126–129.
6. Ip M, Yung RW, Ng TK, Luk WK, Tse C, Hung P, Enright M, Lyon DJ: **Contemporary methicillin-resistant Staphylococcus aureus clones in Hong Kong.** *J Clin Microbiol* 2005, **43**(10):5069–5073.
7. Harris SR, Cartwright EJ, Torok ME, Holden MT, Brown NM, Ogilvy-Stuart AL, Ellington MJ, Quail MA, Bentley SD, Parkhill J, Peacock SJ: **Whole-genome sequencing for analysis of an outbreak of methicillin-resistant Staphylococcus aureus: a descriptive study.** *Lancet Infect Dis* 2013, **13**(2):130–136.
8. Koser CU, Holden MT, Ellington MJ, Cartwright EJ, Brown NM, Ogilvy-Stuart AL, Hsu LY, Chewapreecha C, Croucher NJ, Harris SR, Sanders M, Enright MC, Dougan G, Bentley SD, Parkhill J, Fraser LJ, Betley JR, Schulz-Trieglaff OB, Smith GP, Peacock SJ: **Rapid whole-genome sequencing for investigation of a neonatal MRSA outbreak.** *N Engl J Med* 2012, **366**(24):2267–2275.
9. Harris SR, Feil EJ, Holden MT, Quail MA, Nickerson EK, Chantratita N, Gardete S, Tavares A, Day N, Lindsay JA, Edgeworth JD, de Lencastre H, Parkhill J, Peacock SJ, Bentley SD: **Evolution of MRSA during hospital transmission and intercontinental spread.** *Science* 2010, **327**(5964):469–474.
10. Castillo-Ramirez S, Corander J, Marttinen P, Aldeljawi M, Hanage WP, Westh H, Boye K, Gulay Z, Bentley SD, Parkhill J, Holden MT, Feil EJ: **Phylogeographic variation in recombination rates within a global clone of Methicillin-resistant Staphylococcus aureus (MRSA).** *Genome Biol* 2012, **13**(12):R126.
11. Li Y, Cao B, Zhang Y, Zhou J, Yang B, Wang L: **Complete genome sequence of Staphylococcus aureus T0131, an ST239-MRSA-SCCmec type III clone isolated in China.** *J Bacteriol* 2011, **193**(13):3411–3412.
12. Howden BP, Seemann T, Harrison PF, McEvoy CR, Stanton JA, Rand CJ, Mason CW, Jensen SO, Firth N, Davies JK, Johnson PD, Stinear TP: **Complete genome sequence of Staphylococcus aureus strain JKD6008, an ST239 clone of methicillin-resistant Staphylococcus aureus with intermediate-level vancomycin resistance.** *J Bacteriol* 2010, **192**(21):5848–5849.
13. Betley MJ, Mekalanos JJ: **Nucleotide sequence of the type A staphylococcal enterotoxin gene.** *J Bacteriol* 1988, **170**(1):34–41.
14. Bae T, Baba T, Hiramatsu K, Schneewind O: **Prophages of Staphylococcus aureus Newman and their contribution to virulence.** *Mol Microbiol* 2006, **62**(4):1035–1047.
15. Ben Zakour NL, Sturdevant DE, Even S, Guinane CM, Barbey C, Alves PD, Cochet MF, Gautier M, Otto M, Fitzgerald JR, Le Loir Y: **Genome-wide analysis of ruminant Staphylococcus aureus reveals diversification of the core genome.** *J Bacteriol* 2008, **190**(19):6302–6317.
16. Aires De Sousa M, Sanches IS, Ferro ML, Vaz MJ, Saraiva Z, Tendeiro T, Serra J, De Lencastre H: **Intercontinental spread of a multidrug-resistant methicillin-resistant Staphylococcus aureus clone.** *J Clin Microbiol* 1998, **36**(9):2590–2596.
17. Chen H, Yang X, Wang Q, Zhao C, Li H, He W, Wang X, Zhang F, Wang Z, Chen M, Zhu B, Wang H: **Insights on evolution of virulence and resistance from the whole-genome analysis of a predominant methicillin-resistant Staphylococcus aureus clone sequence type 239 in China.** *Chin Sci Bull* 2014, **59**(11):1104–1112.
18. Yamamoto T, Takano T, Higuchi W, Iwao Y, Singur O, Reva I, Otsuka Y, Nakayashiki T, Mori H, Reva G, Kuznetsov V, Potapov V: **Comparative genomics and drug resistance of a geographic variant of ST239 methicillin-resistant Staphylococcus aureus emerged in Russia.** *PLoS One* 2012, **7**(1):e29187.
19. Holden MT, Lindsay JA, Corton C, Quail MA, Cockfield JD, Pathak S, Batra R, Parkhill J, Bentley SD, Edgeworth JD: **Genome sequence of a recently emerged, highly transmissible, multi-antibiotic- and antiseptic-resistant variant of methicillin-resistant Staphylococcus aureus, sequence type 239 (TW).** *J Bacteriol* 2010, **192**(3):888–892.
20. Costa MO, Beltrame CO, Ferreira FA, Botelho AM, Lima NC, Souza RC, De Almeida LG, Vasconcelos AT, Nicolas MF, Figueiredo AM: **Complete Genome Sequence of a Variant of the Methicillin-Resistant Staphylococcus aureus ST239 Lineage, Strain BMB9393, Displaying Superior Ability To Accumulate ica-Independent Biofilm.** *Genome Announc* 2013, **1**(4):10.1128.genomeA.00576-13.
21. Neoh HM, Mohamed Hussein ZA, Tan XE, Raja Abd Rahman RM B, Hussin S, Mohamad Zin N, Jamal R: **Draft Genome Sequences of Four Nosocomial Methicillin-Resistant Staphylococcus aureus (MRSA) Strains (PPUKM-261-2009, PPUKM-332-2009, PPUKM-377-2009, and PPUKM-775-2009) Representative of Dominant MRSA Pulosotypes Circulating in a Malaysian University Teaching Hospital.** *Genome Announc* 2013, **1**(1):10.1128.genomeA.00103-12. Epub 2013 Jan 31.
22. Chen FJ, Lauderdale TL, Wang LS, Huang IW: **Complete Genome Sequence of Staphylococcus aureus Z172, a Vancomycin-Intermediate and Daptomycin-Nonsusceptible Methicillin-Resistant Strain Isolated in Taiwan.** *Genome Announc* 2013, **1**(6):10.1128.genomeA.01011-13.
23. Diep BA, Gill SR, Chang RF, Phan TH, Chen JH, Davidson MG, Lin F, Lin J, Carleton HA, Mongodin EF, Sensabaugh GF, Perdreau-Remington F: **Complete genome sequence of USA300, an epidemic clone of community-acquired methicillin-resistant Staphylococcus aureus.** *Lancet* 2006, **367**(9512):731–739.
24. 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**(3):1008–1015.
25. Kreen L, Ramaswamy SV, Graviss EA, Naidich S, Musser JM, Kreiswirth BN: **spa typing method for discriminating among Staphylococcus aureus isolates: implications for use of a single marker to detect genetic micro- and macrovariation.** *J Clin Microbiol* 2004, **42**(2):792–799.
26. Zhang K, McClure JA, Conly JM: **Enhanced multiplex PCR assay for typing of staphylococcal cassette chromosome mec types I to V in methicillin-resistant Staphylococcus aureus.** *Mol Cell Probes* 2012, **26**(5):218–221.
27. Zerbino DR, Birney E: **Velvet: algorithms for de novo short read assembly using de Bruijn graphs.** *Genome Res* 2008, **18**(5):821–829.
28. Gladman SST: **VelvetOptimiser Version 2.2.4.** <http://www.vicbioinformatics.com/software/velvetoptimiser.shtml>.
29. Darling AE, Mau B, Perna NT: **progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement.** *PLoS One* 2010, **5**(6):e11147.
30. Galardini M, Biondi EG, Bazzicalupo M, Mengoni A: **CONTIGuator: a bacterial genomes finishing tool for structural insights on draft genomes.** *Source Code Biol Med* 2011, **6**:11–0473. 6–11.
31. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O: **The RAST Server: rapid annotations using subsystems technology.** *BMC Genomics* 2008, **9**:75–2164. 9–75.
32. Stamatakis A, Hoover P, Rougemont J: **A rapid bootstrap algorithm for the RAxML Web servers.** *Syst Biol* 2008, **57**(5):758–771.
33. Zhou Y, Liang Y, Lynch KH, Dennis JJ, Wishart DS: **PHAST: a fast phage search tool.** *Nucleic Acids Res* 2011, **39**(Web Server issue):W347–W352.
34. Langille MG, Brinkman FS: **IslandViewer: an integrated interface for computational identification and visualization of genomic islands.** *Bioinformatics* 2009, **25**(5):664–665.
35. Fischer S, Brunk BP, Chen F, Gao X, Harb OS, Lodice JB, Shanmugam D, Roos DS, Stoekert CJ Jr: **Using OrthoMCL to assign proteins to OrthoMCL-DB groups or to cluster proteomes into new ortholog groups.** *Curr Protoc Bioinformatics* 2011, **6**:6.12–1–19.
36. Hulsen T, De Vlieg J, Alkema W: **BioVenn - a web application for the comparison and visualization of biological lists using area-proportional Venn diagrams.** *BMC Genomics* 2008, **9**:488–2–164-9. 488.

doi:10.1186/1471-2164-15-529

Cite this article as: Wang et al.: Comparative genomics of methicillin-resistant *Staphylococcus aureus* ST239: distinct geographical variants in Beijing and Hong Kong. *BMC Genomics* 2014 **15**:529.