nature portfolio

Peer Review File

Staphylococcus aureus ST764-SCCmecII high-risk clone in bloodstream infections revealed through national genomic surveillance integrating clinical data

Corresponding Author: Dr Motoyuki Sugai

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
Thank you for the opportunity to review Staphylococcus aureus ST764-SCCmecII high-risk clone in bloodstream infections revealed through national genomic surveillance integrating clinical data by Hisatsune and colleagues. In their study, they performed a genomic epidemiology study of 580 S. aureus isolates collected during 2019–2020 from national surveillance of MRSA bloodstream infections and compared this sample to isolates collected from 1994–2000. Most notably, they identified a MRSA ST764-SCCmecII clone that demonstrated a high rate of morbidity and mortality. Overall, their study has a lot of merit. I appreciated the detailed genomic, phylogenetic, and coalescent analysis the authors performed, which they carried out using the appropriate methods. The findings are of interest to the microbial genomics research community, and published genomic data with metadata from this region of the world are a welcome addition. Last, the supplemental and main figures are all nicely constructed. I have a number of suggestions below, but my primary comment is that the text could be more concise. The detailed description of all differences in virulence and antibiotic resistance genes is often hard to follow and it is difficult to identify the notable observations. Many of these results are already in tables and figures. Please see my additional comments below:

- For the results describing the population structure, the authors should add significance tests and then focus on the significant differences in lineage distribution. The list of proportions can be hard to track in the text and may not be needed because most of these results are in tables

- For the comparison dataset of isolates from 1994-2000, there should be a more detailed description of their source in the results section. Did their collection mirror the more recent sample? As the 2019-2020 data show regional differences in population structure, this would be important information to include as a limitation.

- I find it confusing that BAPS was run independently on the comparison sample from 1994-2000, resulting in the same SC number and different associated MLST/CCs. My approach would have been to run BAPS on entire dataset of 580+183 isolates. This way, the SCs would have stayed the same, making comparison easier. I suspect this is why the authors switch to using CCs to describe changes in antimicrobial resistance patterns, which negates the original purpose of using BAPS to assess population structure.

- In the results section describing the differences in virulence and antimicrobial resistance determinants, it is unclear if these observations are statistically significant as there are not p-values for tests. Additionally, it can be a bit tedious to read through all of these comparisons. I think that keeping them as a table/figure and detailing a couple notable differences is sufficient. The evolutionary origins are much more interesting.

- Results 312-315 – must likely the observations from the effective populations size estimates are the result of the sampling bias. The gap in sampling may have led to Ne being artificially inflated. In the root-to-tip correlations, you can see how the uneven sampling has driven the temporal signal. I wouldn't focus too much on the Ne results.

- Methods lines 607-610 – How did the authors determine that a skyline model with strict clock were the best fit? Were other

models tested? For example, did you try a skygrid model with relaxed molecular clock? Technically, all demographic and clock models should have been compared with Bayes factors to determine the best fitting model.

Minor comments
In the abstract, you should expressly state the location of the national surveillance

Line 446 – LTCF not previously defined

Reviewer #2

(Remarks to the Author)
In their study, the authors conducted exemplary national genomic surveillance of S. aureus from BSIs collected between 2019 and 2020. 580 S. aureus isolates from 55 hospitals were included and clinical metadata and resistance data were also integrated into the analyses. The focus of the analyses was on three dominant clones (CC1, CC5 and CC8). The results were compared with genome data from surveillance studies from 1994 to 2000. This comparison made it possible to analyse the genomic evolution of S. aureus over the past 20 years in Japan. The authors were able to trace the emergence and spread of a 'high-risk clone' that evolved from the ST5-SCCmecII clone, which was widespread in the 1990s, through the acquisition of various toxin- and resistance gene-carrying mobile genetic elements.

In my opinion, the study is a very good example of the development of findings from a combination of genotypic, phenotypic surveillance data and patient-related metadata. The text is largely easy to understand and is accompanied by a large number of excellent illustrations. I personally find some passages in the results section somewhat unstructured and lengthy. It could be helpful here to shorten the content a little, especially as it is presented very clearly in the illustrations.

I am not a trained bioinformatician, but from my point of view, state of the art techniques and tools were used to analyse genome data.

Specific comments:
l- 75 – 78
Check sentence and context.
Fig. 1c
Is there really an additional value for this part of the figure? Check for redundancy.
l. 175 ff.
Check for redundancies.
Supp. Fig. 2 STs
Please check legend (grey color).
Supp. Fig. 3 MRSA 1994-2000
Please check legend (blue color?).
l. 207 "resistance to cefazolin (CEZ), cefmetazole (CMZ) decreased in the total sample of MRSA", although all MRSA are positive for mecA. Do you have any explanation? See also Supp. Fig. 4c . l. 220, Fig. 2c (CC1 and CC8). Is there any known association between SCCmecIV and unexpected susceptibility towards cephalosporins I am not aware of? How is resistance defined here? Which breakpoints are used?
In general, the result section describing the presence of antibiotic resistance genes and phenotypic resistance patterns seems difficult to read. Perhaps this part could be tightened up a bit, in especially because the content is perfectly presented in figures; one question I miss is to what extent genotypic and phenotypic data are concordant.
As mentioned already for resistance genes above, I would also tighten up the passage dealing with the description of virulence genes, as it is also nicely presented in figures.
Fig. 4
Check for redundancy: Presence of VFRs and ARGs (Fig 2 and Fig 3)
l. 373 OXA most probably abbreviates Oxacillin, which was previously abbreviated MPIPC? Please check and correct.
l. 374 "A comparison of the resistance profiles and genome data of BSI-derived isolates in a single university hospital between 2011 and 2019 revealed that increased resistance to the three drugs corresponded to CC8-SCCmecIV"
Citation?

l. 513 "We randomly selected…."
I assume that the selection process has ensured a representative sample....?
l. 518 "infected nest"
Please explain.
l. 576 "antibiotic resistance genes encoded by chromosomal mutations"
Actually, these are housekeeping genes carrying resistance associated mutations.
l. 627 ff.
Which Guideline was used for which substance?

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
Thank you for the opportunity to review the revised manuscript. The authors have addressed my comments in detail, making several revisions to the analysis, figures, and text. It also appears that they have made significant changes in response to Reviewer 2's comments. Overall, I believe this has greatly improved the manuscript. In addition, the limitations are now clearly delineated. I have no additional comments at this time.

REVIEWER COMMENTS

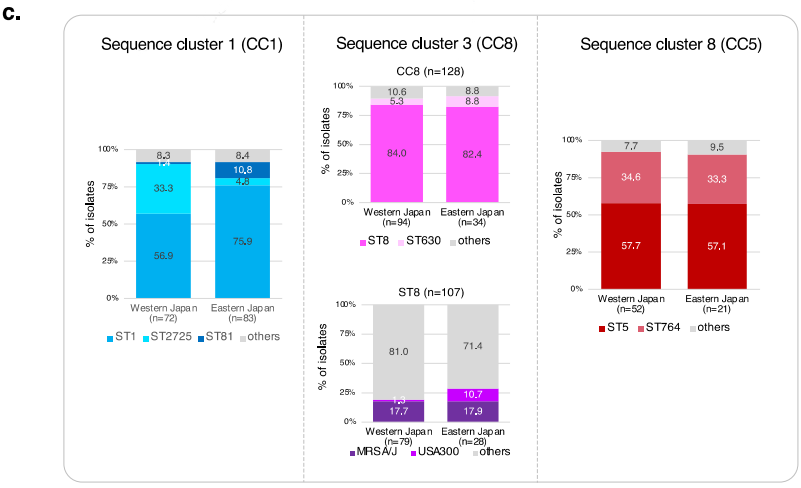
Reviewer #1 (Remarks to the Author):
Thank you for the opportunity to review Staphylococcus aureus ST764-SCCmecII high-risk clone in bloodstream infections revealed through national genomic surveillance integrating clinical data by Hisatsune and colleagues. In their study, they performed a genomic epidemiology study of 580 S. aureus isolates collected during 2019–2020 from national surveillance of MRSA bloodstream infections and compared this sample to isolates collected from 1994–2000. Most notably, they identified a MRSA ST764-SCCmecII clone that demonstrated a high rate of morbidity and mortality. Overall, their study has a lot of merit. I appreciated the detailed genomic, phylogenetic, and coalescent analysis the authors performed, which they carried out using the appropriate methods. The findings are of interest to the microbial genomics research community, and published genomic data with metadata from this region of the world are a welcome addition. Last, the supplemental and main figures are all nicely constructed. I have a number of suggestions below, but my primary comment is that the text could be more concise. The detailed description of all differences in virulence and antibiotic resistance genes is often hard to follow and it is difficult to identify the notable observations. Many of these results are already in tables and figures. Please see my additional comments below:

- For the results describing the population structure, the authors should add significance tests and then focus on the significant differences in lineage distribution. The list of proportions can be hard to track in the text and may not be needed because most of these results are in tables

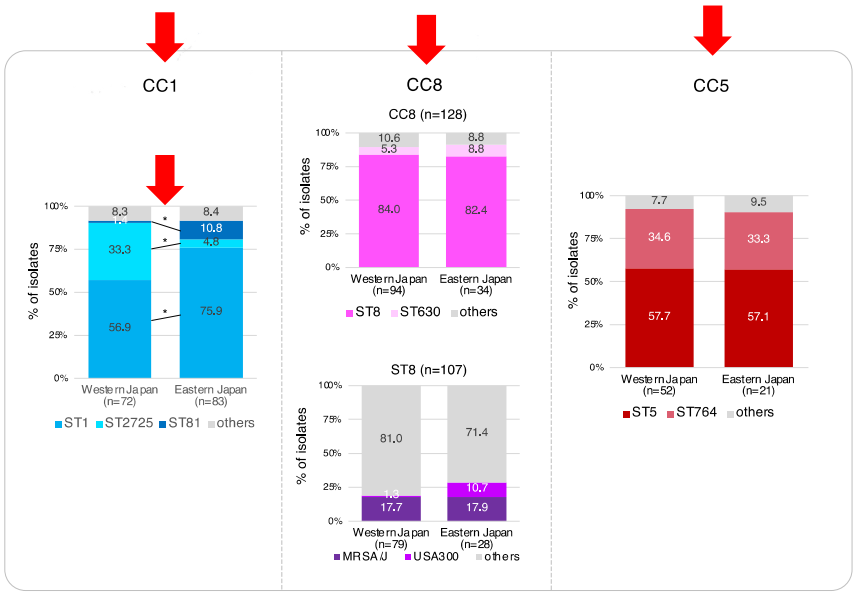
Response:
Thank you for your valuable comments.
Results (lines 130-134, revised text): We have added a test for significant differences in STs between eastern and western Japan and edited the text for clarity. We have also included information on significant differences in Figure 1C. The added or revised parts are marked with red arrows in the figure below.

Fig.1

Before



After



- For the comparison dataset of isolates from 1994-2000, there should be a more detailed description of their source in the results section. Did their collection mirror the more recent sample? As the 2019-2020 data show regional differences in population structure, this would be important information to include as a limitation.

Response:

Thank you for pointing this out. We have described the source information for the isolates from 1994-2000 in lines 503-505 (revised text): "For comparative analysis of BSI-derived *S. aureus* isolates from decades ago, we used a nationwide collection (also known as the YK-Collection) of BSI-derived *S. aureus* isolates from 1994 to 2000, described in a previous study[21]." Additionally, in lines 167-169 (revised text), we have stated: "We

collected the BSI-derived *S. aureus* isolates from post-marketing nationwide surveillance conducted during 1994–2000 after levofloxacin (LVFX) was released in Japan[21].” More detailed information on the 183 strains used for the comparison dataset can be found in Supplementary Table 2 or reference #21.

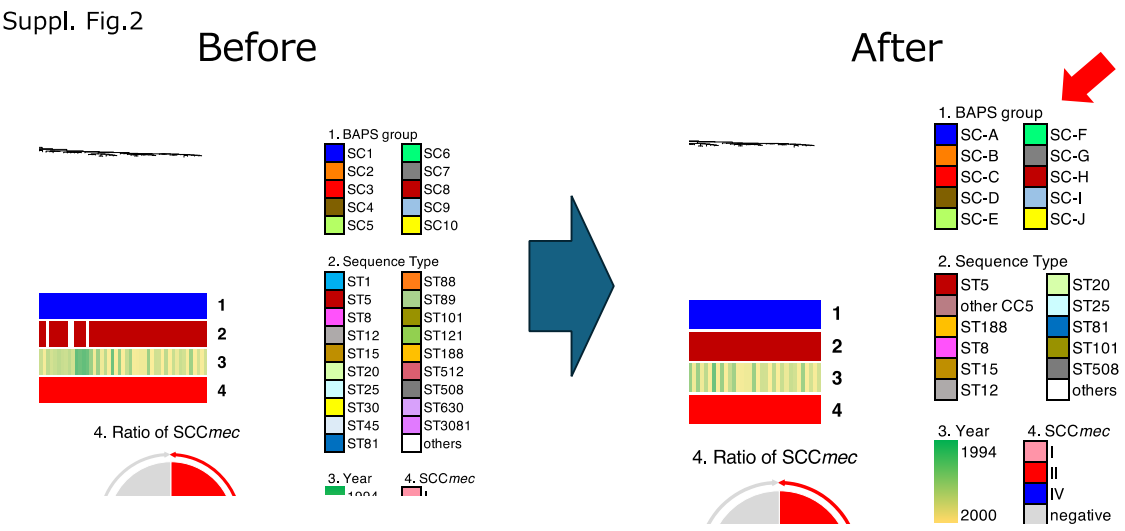
In the 1994–2000 dataset, geographical information was missing, so it was not possible to compare the geographical distribution with the 2019–2020 data. In lines 465-467 (revised text), we have updated the following: "Second, geographic and clinical data for isolates from the period 1994–2000 were not available. Therefore, we were unable to compare the geographical distribution of isolates between the two datasets (2019-2020 and 1994-2000)…"

- I find it confusing that BAPS was run independently on the comparison sample from 1994-2000, resulting in the same SC number and different associated MLST/CCs. My approach would have been to run BAPS on entire dataset of 580+183 isolates. This way, the SCs would have stayed the same, making comparison easier. I suspect this is why the authors switch to using CCs to describe changes in antimicrobial resistance patterns, which negates the original purpose of using BAPS to assess population structure.

Response:

Thank you for your valuable suggestion. The aim of this study was to clarify the genomic characteristics of MRSA isolated from bloodstream infections (BSI) in 2019-2020. Since the dominant lineage of BSI-derived MRSA 20 years ago, which we used as a comparative dataset, is ST5-MRSA (New York/Japan clone), we analyzed the two periods separately. As you pointed out, using the same BAPS/SC numbers in the two datasets cause confusion. Therefore, we have changed the identification code for the BAPS/SC in the 1994-2000 collection from “SC1” to “SC-A” in Supplementary Fig. 2 to clarify that the BAPS/SC numbers are not comparable between the two datasets. We have also rechecked and corrected the ST color.

Suppl. Fig.2



- In the results section describing the differences in virulence and antimicrobial resistance determinants, it is unclear if these observations are statistically significant as there are not p-values for tests. Additionally, it can be a bit tedious to read through all of these comparisons. I think that keeping them as a table/figure and detailing a couple notable differences is sufficient. The evolutionary origins are much more interesting.

Response:

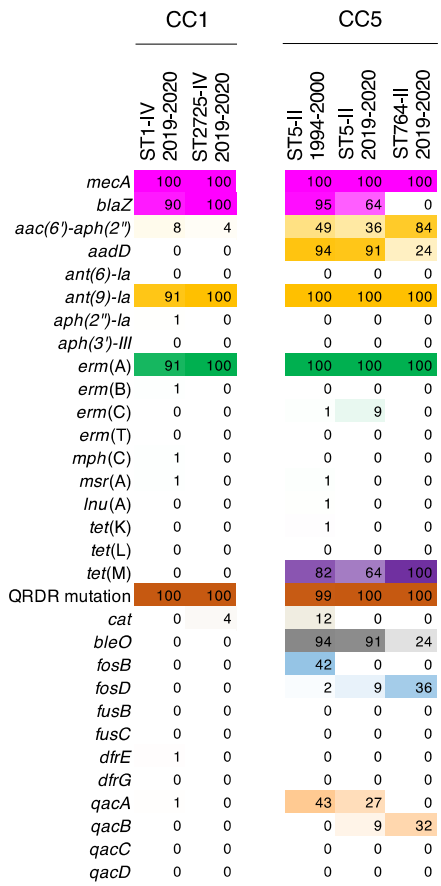
Thank you for raising this point. We conducted statistical significance tests on the data comparing virulence factor and antimicrobial resistance gene determinants and incorporated the results into Fig. 2b, Fig. 3, and Supplementary Fig. 4b-d. The added parts are indicated by red arrows in the figures below. We have added "*p<0.05; ST5-II_1994-2000 vs. ST5-II_2019-2020. #p<0.02; ST5-II_2019-2020 vs. ST764-II_2019-2020. " to the legend of Figure 2b. Furthermore, we have added "*p<0.02; ST5-II_1994-2000 vs. ST5-II_2019-2020." to the legend of Figure 3, and "*p<0.05; 1994-2000 vs. 2019-2020." has been included in the legend of Supplementary Figure 4b-d. Additionally, we have revised the text to clarify several differences that should be compared.

Before

After

Fig.2

b.



b.

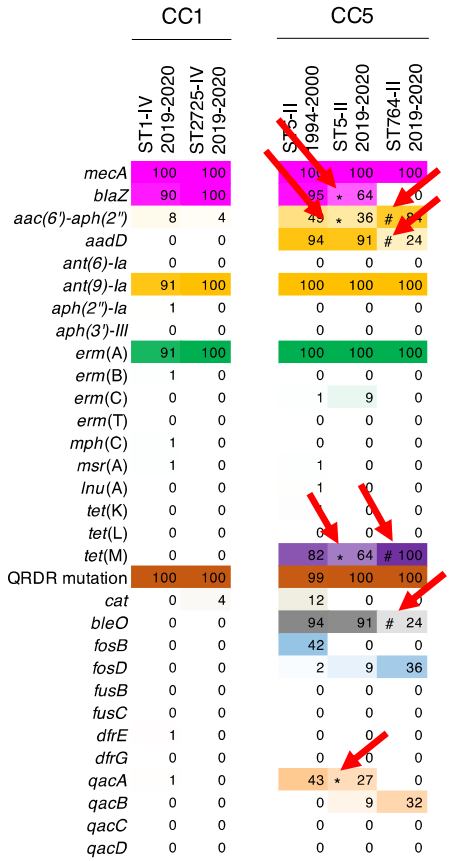
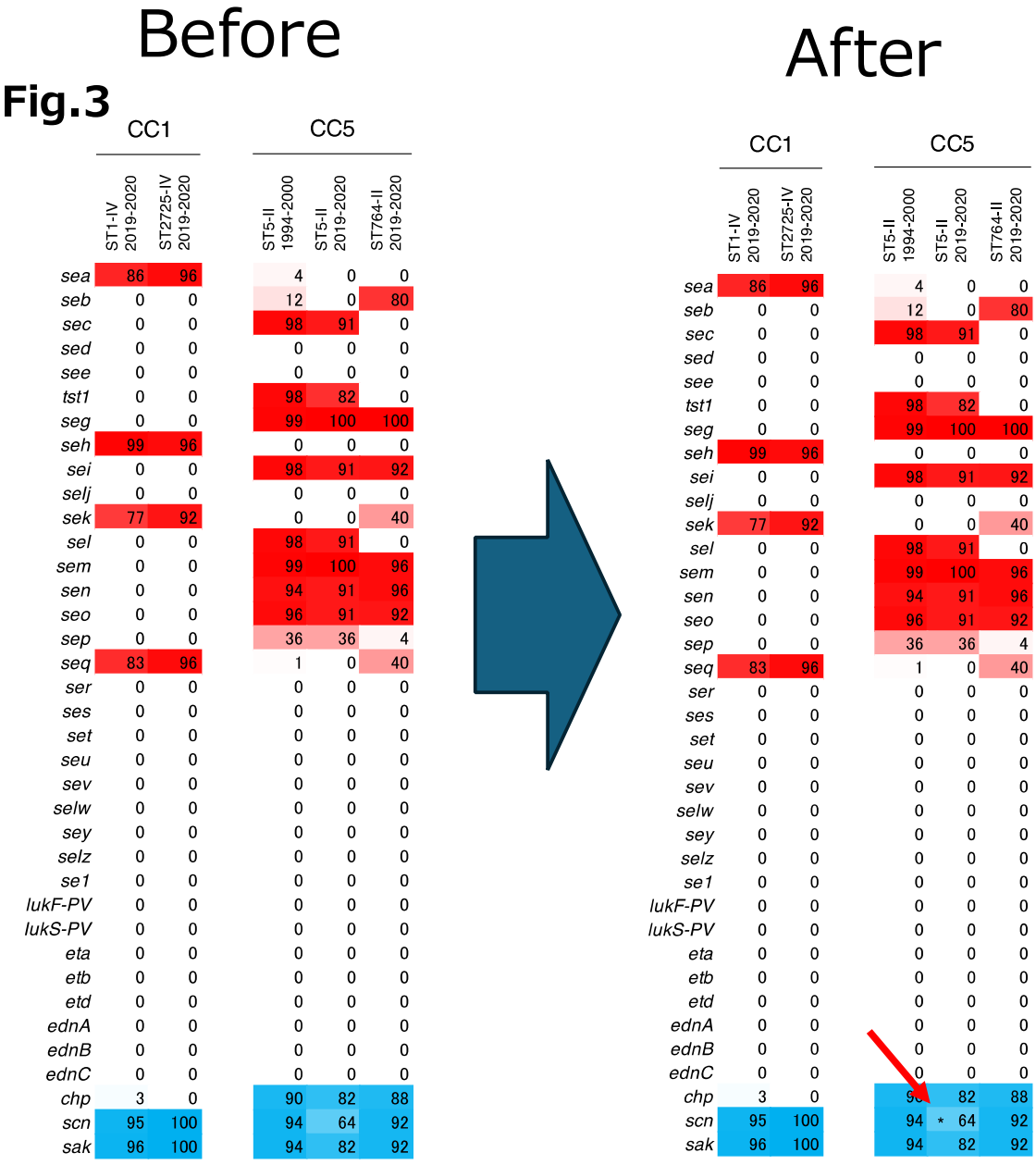
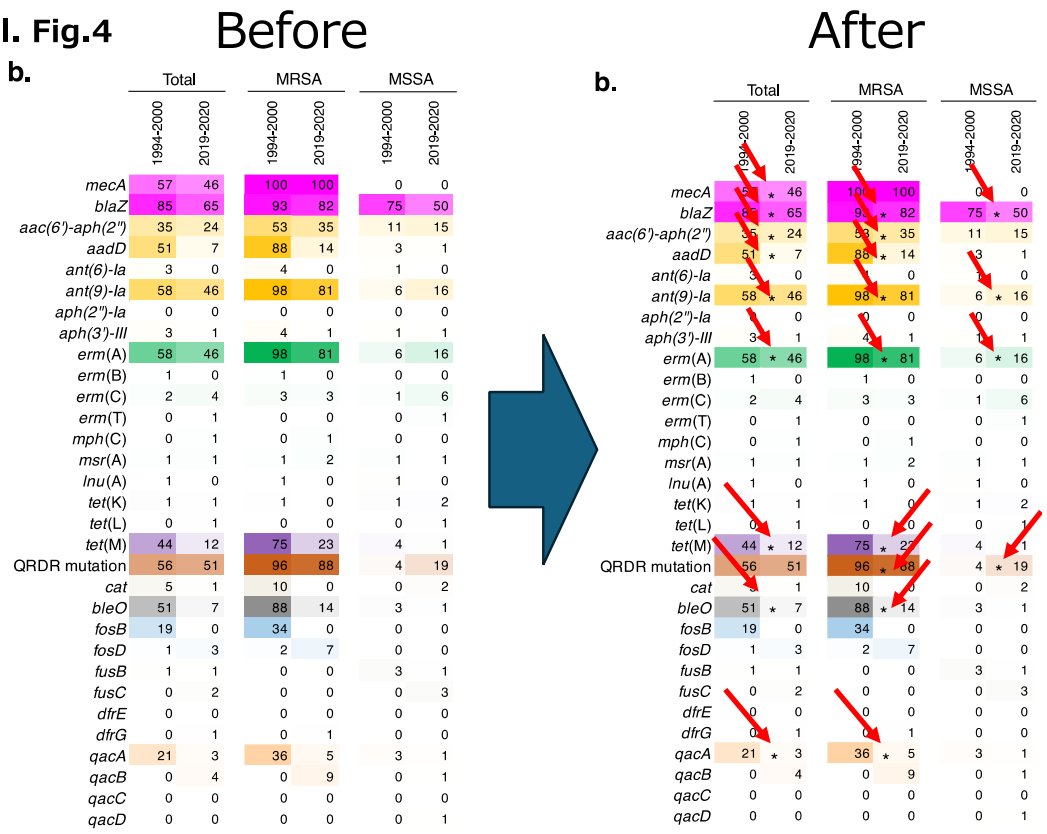


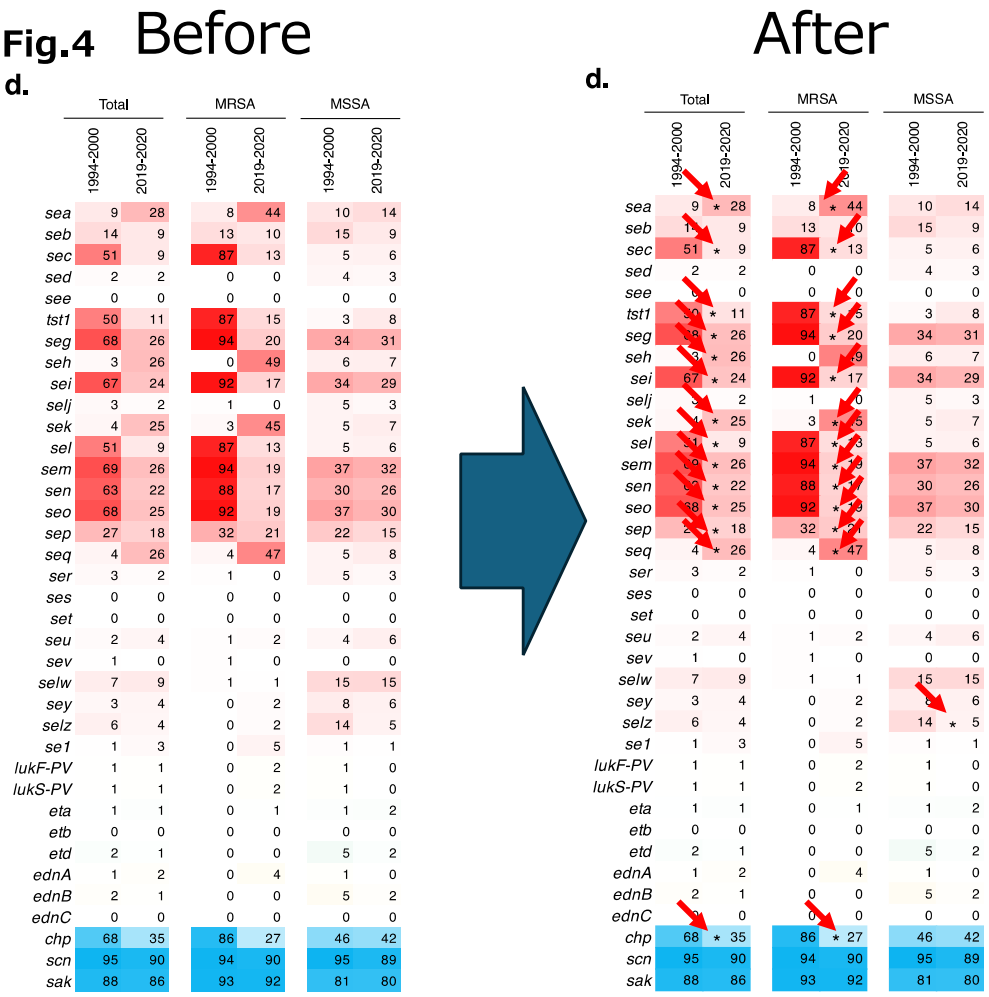
Fig.3


Before
After
CC1
CC5
ST1-IV 2019-2020
ST2725-IV 2019-2020
ST5-II 1994-2000
ST5-II 2019-2020
ST764-II 2019-2020
sea 86 96 4 0 0
seb 0 0 12 0 80
sec 0 0 98 91 0
sed 0 0 0 0 0
see 0 0 0 0 0
tst1 0 0 98 82 0
seg 0 0 99 100 100
seh 99 96 0 0 0
sei 0 0 98 91 92
selj 0 0 0 0 0
sek 77 92 0 0 40
sel 0 0 98 91 0
sem 0 0 99 100 96
sen 0 0 94 91 96
seo 0 0 96 91 92
sep 0 0 36 36 4
seq 83 96 1 0 40
ser 0 0 0 0 0
ses 0 0 0 0 0
set 0 0 0 0 0
seu 0 0 0 0 0
sev 0 0 0 0 0
selw 0 0 0 0 0
sey 0 0 0 0 0
selz 0 0 0 0 0
se1 0 0 0 0 0
lukF-PV 0 0 0 0 0
lukS-PV 0 0 0 0 0
eta 0 0 0 0 0
etb 0 0 0 0 0
etd 0 0 0 0 0
ednA 0 0 0 0 0
ednB 0 0 0 0 0
ednC 0 0 0 0 0
chp 3 0 90 82 88
scn 95 100 94 64 92
sak 96 100 94 82 92
CC1
CC5
ST1-IV 2019-2020
ST2725-IV 2019-2020
ST5-II 1994-2000
ST5-II 2019-2020
ST764-II 2019-2020
sea 86 96 4 0 0
seb 0 0 12 0 80
sec 0 0 98 91 0
sed 0 0 0 0 0
see 0 0 0 0 0
tst1 0 0 98 82 0
seg 0 0 99 100 100
seh 99 96 0 0 0
sei 0 0 98 91 92
selj 0 0 0 0 0
sek 77 92 0 0 40
sel 0 0 98 91 0
sem 0 0 99 100 96
sen 0 0 94 91 96
seo 0 0 96 91 92
sep 0 0 36 36 4
seq 83 96 1 0 40
ser 0 0 0 0 0
ses 0 0 0 0 0
set 0 0 0 0 0
seu 0 0 0 0 0
sev 0 0 0 0 0
selw 0 0 0 0 0
sey 0 0 0 0 0
selz 0 0 0 0 0
se1 0 0 0 0 0
lukF-PV 0 0 0 0 0
lukS-PV 0 0 0 0 0
eta 0 0 0 0 0
etb 0 0 0 0 0
etd 0 0 0 0 0
ednA 0 0 0 0 0
ednB 0 0 0 0 0
ednC 0 0 0 0 0
chp 3 0 82 88
scn 95 100 94 * 64 92
sak 96 100 94 82 92

Suppl. Fig.4


Before
After
b.
b.
Total
MRSA
MSSA
1994-2000
2019-2020
mecA 57 46 100 100 0 0
blaZ 85 65 93 82 75 50
aac(6')-aph(2") 35 24 53 35 11 15
aadD 51 7 88 14 3 1
ant(6)-Ia 3 0 4 0 1 0
ant(9)-Ia 58 46 98 81 6 16
aph(2")-Ia 0 0 0 0 0 0
aph(3')-III 3 1 4 1 1 1
erm(A) 58 46 98 81 6 16
erm(B) 1 0 1 0 0 0
erm(C) 2 4 3 3 1 6
erm(T) 0 1 0 0 0 1
mph(C) 0 1 0 1 0 0
msr(A) 1 1 1 2 1 1
lnu(A) 1 0 1 0 1 0
tet(K) 1 1 1 0 1 2
tet(L) 0 1 0 0 0 1
tet(M) 44 12 75 23 4 1
QRDR mutation 56 51 96 88 4 19
cat 5 1 10 0 0 2
bleO 51 7 88 14 3 1
fosB 19 0 34 0 0 0
fosD 1 3 2 7 0 0
fusB 1 1 0 0 3 1
fusC 0 2 0 0 0 3
dfrE 0 0 0 0 0 0
dfrG 0 1 0 1 0 0
qacA 21 3 36 5 3 1
qacB 0 4 0 9 0 1
qacC 0 0 0 0 0 0
qacD 0 0 0 0 0 1

Suppl. Fig.4



- Results 312-315 – must likely the observations from the effective populations size estimates are the result of the sampling bias. The gap in sampling may have led to Ne being artificially inflated. In the root-to-tip correlations, you can see how the uneven sampling has driven the temporal signal. I wouldn't focus too much on the Ne results.

Response:

Thank you for your insightful suggestion.

We have removed this sentence from lines 312-315 of the original text.

- Methods lines 607-610 – How did the authors determine that a skyline model with strict

clock were the best fit? Were other models tested? For example, did you try a skygrid model with relaxed molecular clock? Technically, all demographic and clock models should have been compared with Bayes factors to determine the best fitting model.

Response:

We sincerely appreciate your guidance on this technical point. We had not tested the SkyGrid model previously; however, we conducted an analysis comparing the following four models: 1) skyline model with a strict clock, 2) skyline model with a relaxed clock, 3) SkyGrid model with a strict clock, and 4) SkyGrid model with a relaxed clock. The models with a relaxed clock failed to achieve convergence in the MCMC because more than half of the model parameters had an ESS <200. A comparison of models with a strict clock, based on the Bayes factor, strongly supported the skyline model over the SkyGrid model (Bayes factor >20). We have added this explanation to lines 599-607 in the Methods section.

Minor comments

In the abstract, you should expressly state the location of the national surveillance

Response:

Thank you for bringing this to our attention. We have specified the locations compared for geographic distribution in this study in line 39 of the abstract (revised text).

Line 446 – LTCF not previously defined

Response:

We appreciate your attention to this matter. In line 427, we have defined LTCFs.

Reviewer #2 (Remarks to the Author):

In their study, the authors conducted exemplary national genomic surveillance of S. aureus from BSIs collected between 2019 and 2020. 580 S. aureus isolates from 55 hospitals were included and clinical metadata and resistance data were also integrated into the analyses. The focus of the analyses was on three dominant clones (CC1, CC5 and CC8). The results were compared with genome data from surveillance studies from 1994 to 2000. This comparison made it possible to analyze the genomic evolution of S. aureus over the past 20 years in Japan. The authors were able to trace the emergence and spread of a ‘high-risk clone’ that evolved from the ST5-SCCmecII clone, which was widespread in the 1990s, through the acquisition of various toxin- and resistance gene-carrying mobile genetic elements.

In my opinion, the study is a very good example of the development of findings from a combination of genotypic, phenotypic surveillance data and patient-related metadata. The text is largely easy to understand and is accompanied by a large number of excellent illustrations. I personally find some passages in the results section somewhat unstructured and lengthy. It could be helpful here to shorten the content a little, especially as it is presented very clearly in the illustrations. I am not a trained bioinformatician, but from my point of view, state of the art techniques and tools were used to analyse genome data.

Response:

Thank you for your review. We have revised the section (lines 188-244, revised text) to improve clarity and explicitly stated their correlations. Similarly, we have condensed and refined the section on virulence factor genotypes (lines 246-271) to make the key points more accessible.

Specific comments:

l- 75 – 78

Check sentence and context.

Response:

Thank you for your insightful comment. We have removed "of MRSA BSIs" from line 77 of the original text.

Fig. 1c

Is there really an additional value for this part of the figure? Check for redundancy.

Response:

Thank you for your perceptive comment. Fig. 1b compares the composition of all ST types between eastern and western Japan, whereas Fig. 1c focuses on the composition of ST types within major CCs. Through these comparisons, we aim to emphasize the regional differences in the distribution of STs belonging to CC1. Additionally, we believe this analysis provides valuable insights, such as the lack of regional variation in invasive clones like ST764, MRSA/J, and USA300.

l. 175 ff.

Check for redundancies.

Response:

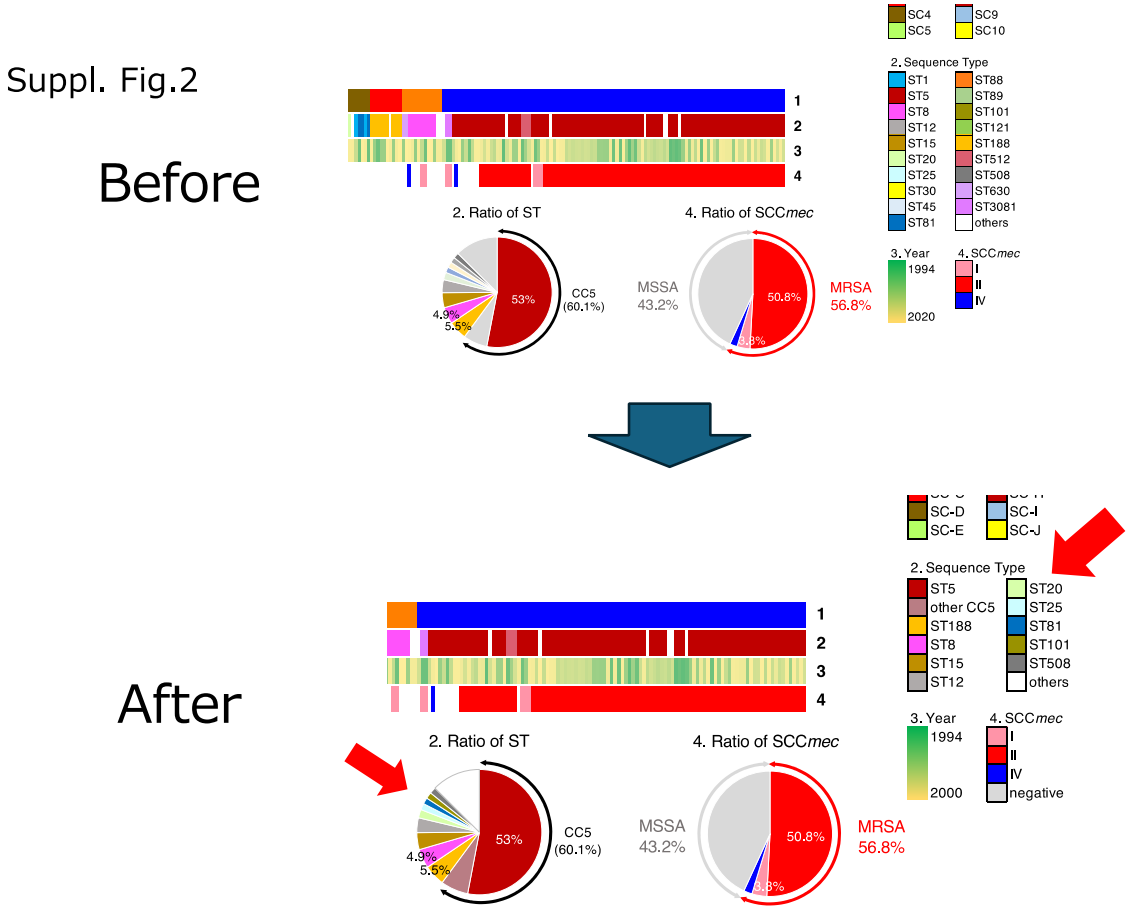
Thank you for your thoughtful comment. Indeed, the sentence in the Results, lines 174-176 (original text), was duplicated, so we have deleted it.

Supp. Fig. 2 STs

Please check legend (grey color).

Response:

Thank you for your important comment. We have revised the ST index colors in Supplementary Fig. 2. The revised parts are indicated by red arrows in the figure below.



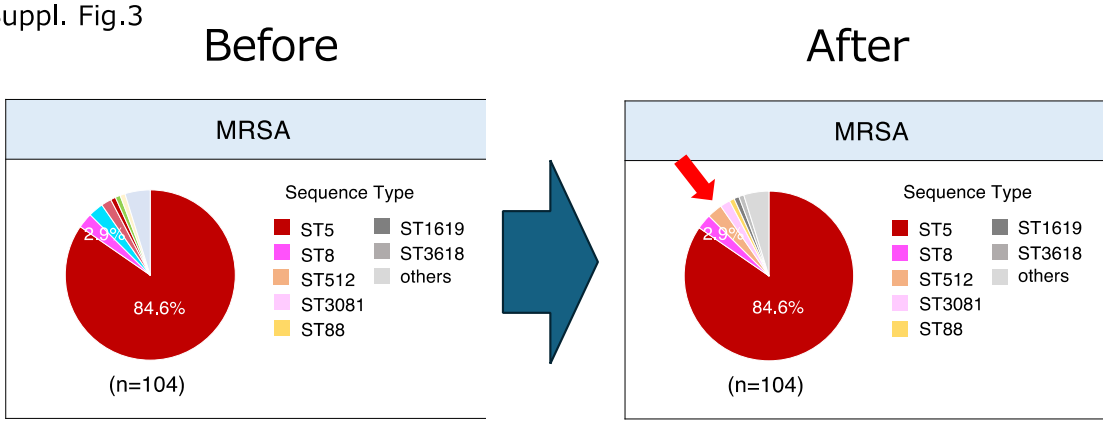
Supp. Fig. 3 MRSA 1994-2000

Please check legend (blue color?).

Response:

Thank you for your meaningful comment. We have revised the ST index colors for MRSA in 1994-2000.

Suppl. Fig.3



l. 207 “resistance to cefazolin (CEZ), cefmetazole (CMZ) decreased in the total sample of MRSA”, although all MRSA are positive for mecA. Do you have any explanation? See also Supp. Fig. 4c . l. 220, Fig. 2c (CC1 and CC8). Is there any known association between SCCmecIV and unexpected susceptibility towards cephalosporins I am not aware of? How is resistance defined here? Which breakpoints are used?

Response:

Thank you very much for pointing this out. Regarding the decrease in cefazolin- and cefmetazole-resistant strains among BSI-derived MRSA isolates in 2019–2020, we believe this trend reflects a shift characterized by a reduction in SCC*mec*II strains and an increase in SCC*mec*IV strains; however, the detailed mechanisms remain unclear. We have added this explanation to the Discussion section (lines 375-381, revised text): “Another notable feature is the significant decrease in resistance rates to CEZ and CMZ among BSI-derived MRSA isolates in 2019–2020. Similar findings have been reported by Yamaguchi et al.[28]. This reduction in MICs for CEZ and CMZ in BSI-derived MRSA isolates is likely attributable to the decline in multidrug-resistant MRSA-SCC*mec*II, which was predominant in the past in Japan, and the corresponding increase in MRSA-SCC*mec*IV. However, the detailed mechanism remains unknown.” We have also added one reference (#28), so the numbers for subsequent references have been adjusted accordingly.

In lines 619-623 (revised text), we have added the following: “AMR to MPIPC, CEZ, CMZ, GM, EM, CLDM, MINO, and LVFX was determined according to the breakpoints defined in the CLSI guidelines (Thirty-First Edition: M100). Since the CLSI guidelines do not provide interpretive criteria for ABK MICs, GM was used as a reference.” The breakpoints for each antimicrobial agent, based on the aforementioned guidelines, are as follows: MPIPC ≥4, CEZ ≥32, CMZ ≥32, GM ≥16, EM ≥8, CLDM ≥4, MINO ≥16, LVFX

≥4, and ABK ≥16.

In general, the result section describing the presence of antibiotic resistance genes and phenotypic resistance patterns seems difficult to read. Perhaps this part could be tightened up a bit, in especially because the content is perfectly presented in figures;

Response:

Your meaningful comment is duly noted. We have refined the Results section describing the presence of antibiotic resistance genes and phenotypic resistance patterns as follows:

First, we have removed lines 198-203 (original text): "In contrast, the prevalence of aminoglycoside resistant gene *ant(9)-Ia*, macrolide resistant genes *erm*(A) and *erm*(C), FOM resistance gene *fosD*, chlorhexidine resistant gene *qacB*, and the quinolone resistance-determining region (QRDR) mutations (GrlA/S80F/E84G and GryA/S84L/E88G) contributing to quinolone resistance in chromosomal *grlA* and *gyrA* were not significantly affected in the total samples or MRSA."

Second, we have condensed lines 224-234 (original text): "The proportions of ARGs in ST5-SCC*mec*II during the two periods were high for *aadD*, *ant(9)-Ia*, *erm*(A), *bleO*, and QRDR mutations (GrlA/S80F/E84K and GyrA/S84L/S85P) (Fig. 2b, CC5 and Supplementary Table 3). In contrast, the ARGs with a decreasing trend in ST5-SCC*mec*II in 2019–2020 were *blaZ*, *aac(6')-aph(2")*, *tet*(M), *cat*, *fosB*, and *qacA*. The ARGs showing an increasing trend in ST5-SCC*mec*II in 2019–2020 were *erm*(C), *fosD*, and *qacB*. ARGs with high proportions in the ST764-SCC*mec*II strain included *ant(9)-Ia*, *erm*(A), and QRDR mutations (GrlA/S80Y/E84K and GyrA/S84L/E88G); particularly, ARGs with a higher proportion than those in ST5-SCC*mec*II were *aac(6')-aph(2")*, *tet*(M), *fosD*, and *qacB*. The ARGs of ST764-SCC*mec*II *aadD*, *erm*(C), and *bleO* were lower in abundance than those of ST5-SCC*mec*II."

To lines 220-223 (revised text): "In ST5-SCC*mec*II, *blaZ*, *aac(6')-aph(2")*, *tet*(M), *cat*, *fosB*, and *qacA* showed a decreasing trend in 2019-2020 compared to 1994-2000 ($p<0.05$) (Fig. 2b, CC5; Supplementary Table 3). In ST764-SCC*mec*II, *aac(6')-aph(2")*, *tet*(M), *fosD*, and *qacB* were detected at a higher prevalence compared to ST5-SCC*mec*II ($p<0.02$)."

Third, we have removed lines 241-245 (original text): "The average number of ARGs in ST8-SCC*mec*IV was lower than that in ST8-SCC*mec*I; however, the difference was statistically significant between ST8-SCC*mec*IVj and ST8-SCC*mec*IVa. Most ST8-

SCC*mec*I possessed *blaZ*, *aac(6')-aph(2")*, *ant(9)-Ia*, *erm*(A), *tet*(M), and QRDR mutations (GrlA/S80F and GyrA/S84L/S85P) (Fig. 2b, CC8 and Supplementary Table 3)."

one question I miss is to what extent genotypic and phenotypic data are concordant.

Response:

Thank you for your valuable question. Regarding the concordant between genotype and phenotype data, we calculated Kendall's coefficient of concordance (W) as described in our previous study and presented the results in the new Supplementary Table 4. We also added the following sentences to lines 240-244 (revised text): "We calculated Kendall's concordance coefficient (W) as described in our previous study[22] and found a certain degree of concordance between genotype and phenotype in CC1, CC5, and CC8 (Supplementary Table 4). High concordance (Kendall's W >0.8) was observed for EM, GM, CLDM, MINO, and LVFX, along with their corresponding resistance determinants."

The numbering following Supplementary Table 5 has been corrected.

We have also added a reference (#22), and the numbering of subsequent references has been updated accordingly.

As mentioned already for resistance genes above, I would also tighten up the passage dealing with the description of virulence genes, as it is also nicely presented in figures.

Response:

We appreciate your thoughtful recommendation. First, we have refined lines 261-263 (original text): "EGC (*seg, sei, sem, sen, and seo*) in the total samples or MRSA of 2019–2020 was drastically reduced compared to that in 1994–2000; however, it remained approximately 30% in MSSA (Supplementary Fig. 4d)."

As lines 249-252 (revised text): "Additionally, the proportion of the staphylococcal enterotoxin gene cluster EGC (*seg*, *sei*, *sem*, *sen*, and *seo*) drastically decreased in total samples and MRSA isolates from 2019–2020 compared to 1994–2000 (p<0.001), but remained approximately 30% in MSSA isolates."

Second, we have condensed lines 278-280 (original text): "The proportions of *seb*, *seg*, *sei*, *sem*, *sen*, and *seo* in ST764-SCC*mec*II were high, whereas the proportions of *sek* and *seq* were low (40%). ST764-SCC*mec*II did not possess *sec*, *tst-1*, or *sel*."

As lines 266-268 (revised text): "ST764-SCC*mec*II exhibited high proportions of *seb*, *seg*, *sei*, *sem*, *sen*, and *seo* but lacked *sec*, *tst-1*, and *sel*."

Third, we have tightened lines 280-286 (original text): “In CC8, ST8-SCC*mec*I and ST8-SCC*mec*IVj possessed 50 and 75% of *sep*, respectively, with no other SE genes present (Fig. 3, CC8). The proportion of *chp* in the ST8-SCC*mec*IVj and IVl clones was significantly lower than that in the other clones. ST8-SCC*mec*IVl contained high proportions of *sec*, *tst-1*, and *sel* (90%). In addition, ST8-SCC*mec*IVl possessed 90% *sel* and 40% *se1* and *ednA*. ST8-SCC*mec*IVa had *sek*, *seq*, and PVL genes (*lukF-PV* and *lukS-PV*) belonging to the USA300 lineage.”

As lines 268-271 (revised text): “In CC8, *sep* was prevalent in ST8-SCC*mec*I, IVj, and IVl (Fig. 3, CC8). ST8-SCC*mec*IVl showed high proportions of *sec*, *tst-1*, and *sel*, along with additional *se1* and *ednA*. ST8-SCC*mec*IVa carried *sek*, *seq,* and PVL genes (*lukF/S-PV*), which are characteristic of the USA300 lineage.”

Fig. 4

Check for redundancy: Presence of VFRs and ARGs (Fig 2 and Fig 3)

Response:

Your helpful suggestion is much appreciated. We have rechecked the redundancy in the presence of VFGs and ARGs in Fig. 2-4, and confirmed that there are no issues. The similarity in notation between “sel” and “se1” may have caused some confusion. These are distinct genes.

Additionally, the colors of the VRGs and ARGs in Fig. 4 have been modified to match the legend of Fig. 4.

l. 373 OXA most probably abbreviates Oxacillin, which was previously abbreviated MPIPC? Please check and correct.

Response:

We are grateful for your considerate advice. We have replaced OXA with MPIPC.

l. 374 “A comparison of the resistance profiles and genome data of BSI-derived isolates in a single university hospital between 2011 and 2019 revealed that increased resistance to the three drugs corresponded to CC8-SCCmecIV“

Citation?

Response:

Your insightful comment is recognized. In line 359 (revised text), we have added reference #27.

l. 513 “We randomly selected….”

I assume that the selection process has ensured a representative sample....?

Response:

Your insightful comment is appreciated. Yes, we selected it assuming that a representative sample was secured. In line 505 (revised text), we have deleted “randomly” to simplify the text.

l. 518 “infected nest”

Please explain.

Response:

We are thankful for your helpful guidance. The “infected nest” information for each strain is clearly stated in Supplementary Table 1. A tag has been added to line 503 (revised text). Additionally, we have moved the sentence from lines 516-518 (original text) to lines 500-503 (revised text).

l. 576 “antibiotic resistance genes encoded by chromosomal mutations”

Actually, these are housekeeping genes carrying resistance associated mutations.

Response:

Thank you kindly for your helpful suggestion. In lines 565-567 (revised text), we have revised the sentence to: “PointFinder v4.1.11[50] was used to detect QRDR mutations in the DNA topoisomerase IV subunit A (GrlA) and DNA gyrase A (GyrA) on the chromosome, also with 90% identity and 90% query coverage cutoffs.”

l. 627 ff.

Which Guideline was used for which substance?

Response:

Thank you for your important question. In lines 619-623 (revised text), we have added the following: “AMR to MPIPC, CEZ, CMZ, GM, EM, CLDM, MINO, and LVFX was determined according to the breakpoints defined in the CLSI guidelines (Thirty-First Edition: M100). Since the CLSI guidelines do not provide interpretive criteria for ABK MICs, GM was used as a reference.”