nature portfolio

Peer Review File

Altered Genomic Methylation Promotes Staphylococcus aureus Persistence in Hospital Environment

REVIEWER COMMENTS

Reviewer #1 (Remarks to the Author):

In this manuscript, the authors investigate the ability of S. aureus to adapt to the hospital environment by studying isolates collected during a hospital outbreak. While the initial clone of the epidemic possesses a functional Agr system, several "sub-clones" express a "flexible" dysfunction of the Agr system associated with various phenotypes. The "flexible" dysfunction of the Agr system is associated with decreased 5mC methylation and altered MraW expression. An mraW mutant recapitulates the "flexible" dysfunction of the Agr system. Thus, a "flexible" dysfunction of the Agr system resulting from an altered methylation profile could favor hospital survival.

The role of genomic methylation has not been extensively studied in S. aureus, and this study based on isolates collected during an epidemic presents some interesting hypotheses.

Major comments:

1. Unfortunately, the regulation of Agr by methylation is not sufficiently explored in the current version of the manuscript.

2. Furthermore, the main problem with the manuscript is the way it is written. The description of the "subclones" is hard to follow, overly descriptive, with no logical transitions to introduce the different assays (e.g. line 198, line 210, line 302). The nomenclature used to name the "sub-clones" is not informative for the reader and longer, more informative names should be given (e.g. indicating their agr status in upper script...).

3. There is a lack of mechanistic explanation of the phenotypes reported (e.g. lines 208-209: persister formation, plasmid acquisition, survival within MP...).

4. line 196 "by overcoming RM systems". by which mechanisms?

Minor:

- A table summarizing the phenotypes of agr+/agr-/EA-agr including the list of “subclones” belonging to each category could help the reader.
- There is a lack of proper definitions of “lineage clones”, clone, “subclones” (lines 137-…)
- The choice of the term “flexible” could be more explained (line 167)
- line 105 : “The majority” can you precise (n=…)
- Lines 169-174: could be part of the introduction
- Line 175: which growth phase ?
- Lines 178, 191, 193 : “higher” can you precise values
- line 201: by gentamycin protection assay (instead of by treatment with antibiotics)
-line 210, 21: Agr positive, Agr expressing replace by Agr functional or something else but with

consistency.
-line 211: under 3 culture conditions: why this choice ?
-line 213: “agr regulated genes”
-line 217: what is exactly “agr suppression conditions” ?
-line 218: not italic
-line 220: these data do not “suggest that altered gene expression…contributed to…” but rather “prompted us to investigate…”
-lines 253, 349: compared to ?
-line 284: “explained” is not appropriate here
-fig4d: the parental strain should be shown and compared
-line 312 313: n=?
-line 321: can you show the data supporting the 2.9%
-lines 365-366: maybe not useful affirmation
-line 385: can you detail aeration conditions

Reviewer #2 (Remarks to the Author):

The study by Yamazaki et al. document a case where new S. aureus lineages emerged with agr defects (2 isolates) as well as environmentally adapted (AE) accessory gene regulator (Agr) expression during a NICU outbreak. The study benefits from the detailed characterization of multiple isolates but in aggregate the results does not provide sufficient support for the main premise that altered genomic methylation through disregulation of mraW promotes S. aureus persistence.

Major comments
• CN35 differs from the other isolates by more than 170 core genome SNVs, which is much larger than what would be expected by the S. aureus mutation rate over a period of ~2 years. This makes it unlikely that this isolate is clonally related to the rest of the outbreak cluster.
• The authors should present growth curves for each isolate to assess whether the Agr phenotype observed under low-aeration conditions could be due to a growth defect under these conditions.
• It is unclear which subclones of which isolates are shown in Figures 2a-e, and wether more than one isolate was tested in each category. For example there are two Agr-defective mutants in the study (CN31 and CN08) but 20 data points are shown for the agr mutant bar in Figure 2b. Were multiple subclones grown and tested for each isolate, or was only one isolate selected and tested multiple times? The interpretation of this figure depends on what data is shown here.
• The RNA-Seq assay also does not have sufficient details to interpret the results. Agr positive (CN02 and CN06), agr mutant (CN08 and CN31), and EA-Agr (CN09 and CN17) isolates were profiled under three

culture conditions, with Figure 2f showing differentially expressed genes between Agr-expression and Agr-suppresion groups. Which groups are compared in this figure and under what culture conditions? Does the Agr-suppression group include both agr mutant and EA-Agr isolates? Does a positive value mean more expression in the Agr-expressing group? How does expression vary between culture conditions and when comparing the agr-mutant and EA-Agr isolates? Is there variability between isolates within each group? A heatmap of top expression changes across the isolates and growth conditions should be included as a supplemental figure to aid interpretation of the results. In particular, the normalized expression levels of the hsdM/S should also be shown in detail across isolates and culture conditions.

• Multiple genes appear upregulated and downregulated between Agr-suppression and Agr-expressing groups, including DNA repair enzymes and exonucleases that could conceivable also impact acquirement of foreign resistance genes.

• Different EA-Agr subclones are used for different experiments. For example, CN09 and CN17 were used for the RNA-Seq experiments, while CN06, CN31, and CN07 were used for mouse infections. What was the rationale for selecting different isolates for each of the experiments in the study?

• It is difficult to interpret the methylation pattern results presented in extended data tables 5-7 and the main results should be summarized in a figure. It should also be clarified whether the observed difference in e.g. the 6mA methylation profile of CN31 can be explained by a decreased sequencing depth and/or reduced sensitivity to detect modified motifs.

• It is unclear which Agr positive and EA-Agr isolates were profiled for extended data table 9 and how expression levels were compared. Notably the expression of mraW is not significantly different between groups after correcting for multiple testing (p 0.06). The method used for multiple testing correction needs to be indicated.

• The statement that the “CN02 mraW-deficient mutant, CN02ΔmraW+pTXΔ16, showed Agr suppression that was explained by differences in cell density under low aeration conditions” needs further clarification. Did this mutant display a growth defect under low aeration conditions that could impact Agr activity indirectly? The authors should also examine the mutant clone to exclude potential mutations in the Agr locus. Likewise, the growth properties of the CN07 subclone under high aeration conditions should be clarified.

Minor comments

• ML-phylogenetic tree of isolates. More details needed. Size of the core genome assessed in the alignment. Background of isolates in the same region would help to assess whether this is a true outbreak or multiple independent introductions.

• Minimum spanning tree of core genome SNVs separating the isolates of the outbreak cluster for just the CC1 isolates provides a better perception of changes and could be introduced earlier when discussing the outbreak cluster.

• The authors should clarify on what basis CN02 was determined to be the founder clone.

• Since long-read sequencing was only done for a few genomes, the presence/absence of Tn4001 and arsenate elements should be verified in all clones by aligning reads against the CN33 and CN35 reference plasmids that contain these element.

• Consider moderating the statement “and became extinct”.

• Please add a legend for the isolate names in Figure 1b and the node colors in Figure 1c.

• Line 136, CN04 should be omitted from the range of CN01-CN06 as is a different clonal complex/ST and

not part of the outbreak cluster.

Reviewer #3 (Remarks to the Author):

The manuscript by Yamazaki et al., “Altered Genomic Methylation Promotes Staphylococcus aureus Persistence in Hospital Environment” presents several original data about lineages displaying an altered Agr expression, including an increased transformation efficiency and an increased mice colonization. Using SMRT sequencing, the authors analyzed the DNA methylation patterns of the genomes (methylomes) of several of these lineages. They also used illumina sequencing of cDNA to analyze the gene expression.
In my opinion, additional data should be presented before clearly drawing the proposed conclusions.

Major concern:

- The current title is not illustrative of the presented data. The authors claim that alteration of genomic methylation promotes persistence. However, this concerns only a limited number of DNA-modifications detected by SMRT-seq. While SMRT-seq is a powerful tool to study m6A and m4C, the well-known limit of SMRT-sequencing for m5C detection should be taken in consideration. Here, the observed 10% diminution of m5C detection could be biased by SMRT-seq technical limitations. To have an exhaustive DNA methylome analysis, and especially when focusing on m5C, additional methods (BSF-seq or EM-seq) should have been used.

- Bacterial DNA-MTases are known to methylate DNA after recognition of specific DNA motifs. Therefore, methylome analysis of a DNA-MTase deletion mutant in bacteria usually leads to differentially methylated sites located in a given motif.
Here in the Gram-positive S. aureus, the identified differentially methylated sites on the DNA are randomly distributed in the genome, not located in a particular motif, and this might be linked to MraW. MraW has been primarily described to have rRNA methylation function. Its putative role in DNA methylation modification has been shown in one study in the Gram-negative bacterium E. coli. Whether a similar function exists in S. aureus remains to be shown.

- The stability of the 16S RNA should be studied, since it is possible that the structure of 16S RNA is modified in the mraW deletion mutant. This could impact various physiological processes in the cells.

Other comments:
- line 261-262: there are no mention of m5C in S. aureus in the cited reference. Another ref should be used here.

- Are the differentially methylated sites on the DNA related to genes with differential expression?

-Table7: it is unclear what are the values indicated in columns D-I. Does this table present predicted

DNA-methylated sites or just DNA-modification?

- Table 8: Exact values of the samples should be shown. Please do not present an average of only 2 samples.

- Fig4C. CN07 is mapping very close to CN02 and CN06. What is the proposed explanation for that?

Reviewer #4 (Remarks to the Author):

I commend the research team for their great efforts in this infant cohort-based study on a MRSA outbreak and the comprehensive genome sequencing and comparative genomics analysis. This work is particularly relevant given the rising concern over MRSA as a critical pathogen contributing to increased mortality and antibiotic resistance. It is also exciting to see the manuscript implicate the possible role of epigenetic regulation in MRSA adaption to hospital environment because bacterial epigenetics represents a new direction to better understand bacterial adaption. However, several major issues in the manuscript warrant further investigation and clarification before re-evaluation.

1. The genetic diversity among the studied clones is substantial. Even though there are no consensus differences as emphasized by the authors, the collective impact of multiple genetic variations on the phenotype cannot be ruled out. Comparing several isolates and clones with numerous genetic differences falls short in convincingly attributing the observed traits solely to factors like Agr and the RM system. This is a general challenge that is worth careful consideration by the authors.

2. The use of culture media to study Agr phenotypes, while informative, may not adequately mimic the environmental selective pressures the study aims to elucidate. If epigenetics is indeed mediating the reversible suppression of Agr, its effects should be susceptible to changes in culture conditions (from MRSA real environmental conditions). Unless some underlying upstream genetic factors stabilize this difference, which would then imply that epigenetics is only an intermediary in this case.

3. The RNAseq data showed an upregulation of hsdM/S, the methyltransferase and specificity unit, but a downregulation of Type I Restriction Enzyme. Typically, Type I RMS systems are regulated in concert, so this divergence in gene expression trends warrants careful scrutiny to rule out any potential artifacts in the RNAseq data analysis.

4. Regarding the detection of 5mC methylation, I'm concerned about the reliability of the PacBio sequencing data. PacBio is great at detecting 6mA and 5mC, but the signal for 5mC is generally modest and can be influenced by sequencing depth (deeper sequencing tends to call higher 5mC %), leading to potentially misleading interpretations of methylation percentages. For a claim as critical as the epigenetic regulation of 5mC, more robust methods like bisulfite sequencing or EM-seq (or even better: NEBNext® Enzymatic Methyl-seq Kit) should be employed to validate these findings.

5. While the genetic investigation of the 5mC methyltransferase MraW is a step in the right direction,

natural regulation of 5mC methyltransferase and complete deletion mutations are not equivalent. Confirming the extent of 5mC changes remains pivotal. Additionally, understanding the prevalence of the MraW gene across different MRSA strains (is MraW common or rare) could help understand how broadly applicable are the findings in this study.

6. The phenotypic disparities between CN08 and CN31, and the unique RM system motifs in CN31, add complexity to the interpretation of the underlying genetic and epigenetic drivers.

In summary, while the study presents some promising avenues of research, its current form leaves room for ambiguity and requires additional experiments and analyses to solidify its conclusions.

We thank the referees for their constructive and encouraging comments, which helped us to improve the quality and content of our manuscript. We have conducted additional experiments and revised the Figures and the text to address the points raised by the referees. The new results are shown in 21 new and revised panels/tables (Figure 1b, 1c, 2b, 5a, 5c, Extended data figure 2b, 3a, 3e, 4a, 4b, 4c, 7, 8a, 8b, 8c Extended data table 4, 8, 9, 10, 11,12). The revised sentences are indicated with underlining for clarity.

REVIEWER COMMENTS

Reviewer #1 (Remarks to the Author):

In this manuscript, the authors investigate the ability of S. aureus to adapt to the hospital environment by studying isolates collected during a hospital outbreak. While the initial clone of the epidemic possesses a functional Agr system, several "sub-clones" express a "flexible" dysfunction of the Agr system associated with various phenotypes. The "flexible" dysfunction of the Agr system is associated with decreased 5mC methylation and altered MraW expression. An mraW mutant recapitulates the "flexible" dysfunction of the Agr system. Thus, a "flexible" dysfunction of the Agr system resulting from an altered methylation profile could favor hospital survival.

The role of genomic methylation has not been extensively studied in S. aureus, and this study

based on isolates collected during an epidemic presents some interesting hypotheses.

Reply> *Thank you very much for your positive and insightful comments on our manuscript. We greatly appreciate your recognition of our work.*

Major comments:

1. Unfortunately, the regulation of Agr by methylation is not sufficiently explored in the current version of the manuscript.

Reply> *We appreciate this feedback and agree that further investigation was necessary. In response to Reviewer 3 and 4's suggestions, we additionally performed Enzymatic Methyl-seq for 5mC methylation on these clinical isolates. Our new data reveal that altered 5mC methylation on transcriptional regulator genes (rpsD and pcrA) is conserved in hospital-associated EA-Agr isolates. These findings further support our hypothesis that altered methylation profiles play a crucial role in the flexible dysfunction of the Agr system, thereby enhancing the adaptability and survival of S. aureus in hospital environment. New data are shown in Figure 5a, 5b and Extended table 10 and described in the main text on pages 14-15, lines 335-354.*

2. Furthermore, the main problem with the manuscript is the way it is written. The description of the "subclones" is hard to follow, overly descriptive, with no logical transitions to introduce the different assays (e.g. line 198, line 210, line 302). The nomenclature used to name the "sub-clones" is not informative for the reader and longer, more informative names should be given (e.g. indicating their agr status in upper script...).

Reply> *We appreciate this feedback and have made several revisions to address these concerns. Specifically, we have added the Agr status (Agr positive, agr mutant, and EA-Agr) in upper script with the strain names throughout the manuscript. Additionally, we have included the following statement to clarify this change: "To make the characteristics of the subclones clearer, we have described the names of the outbreak subclones to include their Agr status in upper script (Agr positive: agr^{+}, Agr mutant: agr^{-}, EA-Agr: EA)" (page 6, lines 139-141).*

3. There is a lack of mechanistic explanation of the phenotypes reported (e.g. lines 208-209: persister formation, plasmid acquisition, survival within MP...).

Reply> *As the reviewer pointed out, our initial manuscript lacked detailed mechanistic explanations. In response, we have now included more detailed descriptions of how Agr expression is related to these phenotypes. Specifically, we have expanded on the following*

paragraphs: page8 lines 190-193, page9 lines 220－221, page10, lines 230-235, page11, lines 250-257).

4. line 196 "by overcoming RM systems". by which mechanisms?

Reply> *We performed additional qPCR analysis and confirmed that hsdR expression is suppressed by Agr suppression condition, leading to the conclusion, that the decreased hsdR expression in the restriction-modification system and increased expressions of DNA repair enzymes and exonucleases in subclones with suppressed Agr contributed to the acquirement of foreign resistance genes in agr mutant and EA-Agr subclones. We revised the text based on these new data (page11, line249-257).*

Minor:

- A table summarizing the phenotypes of agr+/agr-/EA-agr including the list of "subclones" belonging to each category could help the reader.

Reply> *We added the Agr categories (Agr positive/agr mutant/ EA-Agr) of each isolate in Extended Table 4.*

- There is a lack of proper definitions of "lineage clones", clone, "subclones" (lines 137-…)

Reply> *We revised the text. "Clone" is now only used for CN02, while the other outbreak CC1*

lineage isolates are defined as "subclones".

- The choice of the term “flexible” could be more explained (line 167)

Reply> *We added the following sentence to explain the term "flexible". "These subclones demonstrated flexibility in Agr expression, adapting their Agr activation according to the aeration levels, which differed from the typical Agr behavior." on page 7, lines154-155.*

- line 105 : “The majority” can you precise (n=…)

Reply> *We added the number (n=14, 70%) of patients who were only asymptomatically colonized by MRSA and patients infected by MRSA (n=6, 30%) on page 5, lines 114-115.*

- Lines 169-174: could be part of the introduction

Reply> *We moved these sentences to the page 3, lines 68-72.*

- Line 175: which growth phase ?

Reply> *We revised the sentence "late growth phage (4 hours culture)…" on page 8,line 193.*

- Lines 178, 191, 193 : “higher” can you precise values

Reply> *We added p-value as the reviewer suggested on page 8, line 196, page 9, line 212, page 9, line 213.*

- line 201: by gentamycin protection assay (instead of by treatment with antibiotics)

Reply> *We changed the sentence to "by gentamycin/vancomycin protection assay" on page 9, line 223.*

-line 210, 21: Agr positive, Agr expressing replace by Agr functional or something else but with consistency. & -line 211: under 3 culture conditions: why this choice ? & -line 217: what is exactly “agr suppression conditions” ?

Reply> *We apologize for the confusion caused by the lack of explanation. The meanings of "Agr positive" and "Agr expressing" are different. We added additional explanation on page 10, lines 235-242 and an Extended Fig. 4 to clarify why and how the Agr expressing and Agr-suppressed groups were set up and the purpose of creating these groups. We categorized Agr expressing condition (Agr positive in high and low aeration conditions and EA-Agr in high aeration condition) and Agr-supression condition (EA-Agr in low aeration conditions and agr mutant in high aeration condition) for maximizing the detection sensitivity of Agr-dependent gene expression.*

-line 213: “agr regulated genes”

Reply> *We revised the sentence on page 10, lines 243-244.*

-line 218: not italic

Reply> *We revised the sentence on page 11, lines 249.*

-line 220: these data do not “suggest that altered gene expression…contributed to…” but rather “prompted us to investigate…”

Reply> *We additionally performed qPCR for hsdM and hsdR gene expressions and added*

some new data as Reviewer 2&4 suggested and toned down this conclusion on page 11, line 249-257.

-lines 253, 349: compared to ?

Reply> *We revised the sentence on page 12, line 289, page 17, lines 408-409.*

-line 284: "explained" is not appropriate here

Reply> *We revised the sentence on page 15, line 360*

-fig4d: the parental strain should be shown and compared

Reply> *We added the parental strains in new Figure 5e.*

-line 312 313: n=?

Reply> *We added the numbers of isolates on page13, lines 294-295.*

-line 321: can you show the data supporting the 2.9%

Reply> *We already showed how to calculate the percentage of EA-Agr phenotype in all sequenced strain (4 strains from 137 isolates) in skin-adapted strains, as already stated on page 13, lines 303-304.*

-lines 365-366: maybe not useful affirmation

Reply> *As reviewer suggested, we removed the last sentence on page 18.*

-line 385: can you detail aeration conditions

Reply> *Following the description of aeration conditions, we added a sentence " We defined*

this Monod culture apparatus as high aeration because the use of the Monod culture apparatus relatively increases aeration compared to the plastic tube culture conditions. " on page 19, lines 451-453.

Reviewer #2 (Remarks to the Author):

The study by Yamazaki et al. document a case where new S. aureus lineages emerged with agr defects (2 isolates) as well as environmentally adapted (AE) accessory gene regulator (Agr) expression during a NICU outbreak. The study benefits from the detailed characterization of multiple isolates but in aggregate the results does not provide sufficient support for the main premise that altered genomic methylation through disregulation of mraW promotes S. aureus persistence.

Reply> *Thank you for highlighting the novelty of our study. As the reviewer pointed out, we created a mutant of the methyltransferase mraW as a model to understand how cytosine methylation might influence EA-Agr phenotype. However, it is important to note that mutations in mraW were not detected in the clinical isolates of EA-Agr. The reviewer's comment is indeed accurate, and in response, we have revised both the abstract and the main text to either remove or tone down the statements regarding the direct contribution of mraW. This revision ensures that our manuscript accurately reflects the findings and their*

implications.

Major comments

• CN35 differs from the other isolates by more than 170 core genome SNVs, which is much larger than what would be expected by the S. aureus mutation rate over a period of ~2 years. This makes it unlikely that this isolate is clonally related to the rest of the outbreak cluster.

Reply> *Thank you for your insightful comments. As the reviewer's suggested, we removed CN35 from Fig1c and mentioned "CN35 differs from the other isolates by more than 170 core genome SNVs, which is much larger than what would be expected by the S. aureus mutation rate over a period of approximately 2 years." on page 5 line 111-1113. Additionally, we noted that the reviewer appropriately used "SNV" instead of "SNP," and we have updated our manuscript to use "SNV" throughout.*

• The authors should present growth curves for each isolate to assess whether the Agr phenotype observed under low-aeration conditions could be due to a growth defect under these conditions.

Reply> *We added the growth curves for each isolate of Agr positive and EA-Agr under low-aeration condition in the new Extended data figure 2b in addition to Extended data figure 2a.*

EA phenotype did not affect their growth low-aeration conditions.

• It is unclear which subclones of which isolates are shown in Figures 2a-e, and wether more than one isolate was tested in each category. For example there are two Agr-defective mutants in the study (CN31 and CN08) but 20 data points are shown for the agr mutant bar in Figure 2b. Were multiple subclones grown and tested for each isolate, or was only one isolate selected and tested multiple times? The interpretation of this figure depends on what data is shown here.

Reply> *We apologize for the confusion caused by the lack of explanation. We revised the figure legend of Figure 2 and added information on that how many experimental replicates from each subclones are shown in each figure panel.*

• The RNA-Seq assay also does not have sufficient details to interpret the results. Agr positive (CN02 and CN06), agr mutant (CN08 and CN31), and EA-Agr (CN09 and CN17) isolates were profiled under three culture conditions, with Figure 2f showing differentially expressed genes between Agr-expression and Agr-suppresion groups. Which groups are compared in this figure and under what culture conditions? Does the Agr-suppression group include both agr mutant and EA-Agr isolates? Does a positive value mean more expression in the Agr-

expressing group? How does expression vary between culture conditions and when comparing the agr-mutant and EA-Agr isolates? Is there variability between isolates within each group? A heatmap of top expression changes across the isolates and growth conditions should be included as a supplemental figure to aid interpretation of the results. In particular, the normalized expression levels of the hsdM/S should also be shown in detail across isolates and culture conditions.

Reply> *We apologize for the confusion caused by the lack of explanation. The meanings of "Agr positive" and "Agr expressing" are different. We have added additional sentences on page 10, lines 235-242 and a new Extended data figure 4 to clarify why and how the Agr expressing and Agr-suppressed groups were set up and the purpose of creating these groups.We categorized Agr expressing condition (Agr positive in high and low aeration conditions and EA-Agr in high aeration condition) and Agr-supression condition (EA-Agr in low aeration conditions and agr mutant in high aeration condition) for maximizing the detection sensitivity of Agr-dependent gene expression. Because most gene expressions were dependent on culture conditions based on PCA analysis shown in Extended data figure 4a. A heatmap of top expression changes across the isolates and growth conditions was shown in New Extended data figure 4b. HsdR and hsdM* gene expressions normalized by housekeeping gene are shown *in New Extended data figure 4c. We also revised the sentences*

on page 10-11, line 235-257.

• Multiple genes appear upregulated and downregulated between Agr-suppression and Agr-expressing groups, including DNA repair enzymes and exonucleases that could conceivable also impact acquirement of foreign resistance genes.

Reply> *Thank you for your insightful comments. We mentioned this possibility on page 11, line 254-257.*

• Different EA-Agr subclones are used for different experiments. For example, CN09 and CN17 were used for the RNA-Seq experiments, while CN06, CN31, and CN07 were used for mouse infections. What was the rationale for selecting different isolates for each of the experiments in the study?

Reply> *We had already used CN02, CN08, and CN17 for in vivo experiments in the Main Figure, and CN06, CN31, and CN07 in Extended Data Figure 5 to verify the phenotypic consistency across multiple subclones. For the strain examination, the selection varied for each experiment when performed without bias by the respective experimenter. As the reviewer pointed out, confirming the EA-Agr phenotype in multiple subclones in vitro is also important. Therefore, we have conducted additional in vitro experiments using different*

subclones. We have replaced Figure 2b and added the results to the New Extended Data Figure 3a and 3e to ensure the consistency of the EA-Agr phenotype across various subclones.

• It is difficult to interpret the methylation pattern results presented in extended data tables 5-7 and the main results should be summarized in a figure. It should also be clarified whether the observed difference in e.g. the 6mA methylation profile of CN31 can be explained by a decreased sequencing depth and/or reduced sensitivity to detect modified motifs.

Reply> *We additionally analyzed 5mC by Enzymatic methyl-seq (EM-seq) as suggested by Reviewer 3 and Reviewer 4. We also repeated the analysis of 6mA and 4mC using PacBio Sequel IIe and reanalyzed the data. As recommended by Reviewer 3 and Reviewer 4, we removed all 5mC data from PacBio and consequently removed the old Extended Data Tables 7 and 8. Our additional experiments using EA-seq showed that two gene expression regulators (pcrA and rpsD) are differentially methylated between Agr positive and EA-Agr. The new data are shown in the new Figure 5a, 5b, the new Extended data table 9 and new Extended data Table 10. Of note, in the initial PacBio sequencing, the coverage of sequence for all analyzed strains was standardized to 100, meaning that the sequencing depth did not affect the 6mA methylation profile of CN31. We have included this coverage information in the new Extended data table 9.*

• It is unclear which Agr positive and EA-Agr isolates were profiled for extended data table 9 and how expression levels were compared. Notably the expression of mraW is not significantly different between groups after correcting for multiple testing (p 0.06). The method used for multiple testing correction needs to be indicated.

Reply> *We apologize for not making the description of the results clearer. We additionally mentioned which Agr positive and EA-Agr isolates were profiled for the revised extended data table 11 in the maintaxt (page15, line 346-347). We applied the Benjamini-Hochberg method for multiple comparison testing (Default for DESeq2) as now indicated on page 28, lines 639-640. Given our tolerance for false positives in order to discover novel mechanisms, we considered an adjusted p-value/False Discovery Rate (FDR) of < 0.1 to be significant. Interestingly, compared to the low aeration condition, decreased expressions of mraW were observed in the high aeration condition in the Agr-positive group. In contrast, elevated expressions were observed in the high aeration condition in EA-Agr subclones. Therefore, we expected that Agr expression in EA-Agr subclones would have a high dependency on cytosine methylation for Agr expression. Based on these findings, we considered the gene to be important for Agr expression, despite not showing significant differences at FDR < 0.05. We have added the explanation for this to page 15, lines 349-354.*

• The statement that the “CN02 mraW-deficient mutant, CN02ΔmraW+pTXΔ16, showed Agr suppression that was explained by differences in cell density under low aeration conditions” needs further clarification. Did this mutant display a growth defect under low aeration conditions that could impact Agr activity indirectly? The authors should also examine the mutant clone to exclude potential mutations in the Agr locus. Likewise, the growth properties of the CN07 subclone under high aeration conditions should be clarified.

Reply> *Thank you for pointing out the error. CN02ΔmraW+pTXΔ16 does not cause growth defects. In contrast, CN07ΔmraW+pTXΔ16 showed growth defect in high aeration condition. We have removed the sentence "that was explained by differences in cell density." and mentioned about the growth of CN07ΔmraW+pTXΔ16 on page16, lines 365-366.*

These specific mraW mutants and complemented strains do not have agr mutations which was confirmed by HiFi sequencing. We added this in the method section (page 20, line 473-474).

Minor comments

• ML-phylogenetic tree of isolates. More details needed. Size of the core genome assessed in the alignment. Background of isolates in the same region would help to assess whether

this is a true outbreak or multiple independent introductions.

Reply> *As suggested by the reviewers, we performed a background analysis of 164 S. aureus isolates from the same region (the same hospital and prefecture in Japan). Using Roary with default parameters, we identified 1,769 core genes, 140 soft core genes, 1,188 shell genes, and 3,784 cloud genes within the background of isolates. Additionally, we re-analyzed the core genome size after removing reference genomes, CC8 strains, and CN35 using Roary, and identified 2,523 core genes, 143 shell genes, and 248 cloud genes within the CC1 outbreak lineage. The results are described in the Methods section (page 22, lines 502-514). By providing the genetic background of the isolates from the same region, we concluded the observed cases represent a true outbreak.*

• Minimum spanning tree of core genome SNVs separating the isolates of the outbreak cluster for just the CC1 isolates provides a better perception of changes and could be introduced earlier when discussing the outbreak cluster.

Reply> *We revised Figure1c based on core genome SNVs separating the isolates of the outbreak cluster for just the CC1 isolates.*

• The authors should clarify on what basis CN02 was determined to be the founder clone.

Reply> *Since CN02 agr+ was genetically the closest to CN07 EA among CN01-03 agr+ and was identified the earliest, we referred to CN02 agr+ as the founder clone in this study. We*

mentioned this on page 7, line 173-175 and in revised Figure 1c.

• Since long-read sequencing was only done for a few genomes, the presence/absence of Tn4001 and arsenate elements should be verified in all clones by aligning reads against the CN33 and CN35 reference plasmids that contain these element.

Reply> *We confirmed the Tn4001 element in all CC1 outbreak subclones except for CN35 because CN35 was already excluded from the lineage analysis. We found that Tn4001 only exists in CN33 but not in other subclones, which is consistent with the observation that gentamicin resistance was only acquired in by CN33.*

• Consider moderating the statement "and became extinct".

Reply> *We removed the sentence as the reviewer suggested.*

• Please add a legend for the isolate names in Figure 1b and the node colors in Figure 1c.

Reply> *We added the isolate names and the node colors in the legends of Figure 1b and Figure 1c*.

• Line 136, CN04 should be omitted from the range of CN01-CN06 as is a different clonal complex/ST and not part of the outbreak cluster.

Reply> *We omitted CN04 from CN01-CN06 on page 6, line145.*

Reviewer #3 (Remarks to the Author):

The manuscript by Yamazaki et al., "Altered Genomic Methylation Promotes Staphylococcus aureus Persistence in Hospital Environment" presents several original data about lineages displaying an altered Agr expression, including an increased transformation efficiency and an increased mice colonization. Using SMRT sequencing, the authors analyzed the DNA methylation patterns of the genomes (methylomes) of several of these lineages. They also used illumina sequencing of cDNA to analyze the gene expression.

In my opinion, additional data should be presented before clearly drawing the proposed conclusions.

Reply> *Thank you for identifying the important points of our research. Below, we describe the additional analysis results as pointed out by the reviewer.*

Major concern:

- The current title is not illustrative of the presented data. The authors claim that alteration of genomic methylation promotes persistence. However, this concerns only a limited number of DNA-modifications detected by SMRT-seq. While SMRT-seq is a powerful tool to study m6A and m4C, the well-known limit of SMRT-sequencing for m5C detection should be taken in consideration. Here, the observed 10% diminution of m5C detection could be biased by SMRT-seq technical limitations. To have an exhaustive DNA methylome analysis, and

especially when focusing on m5C, additional methods (BSF-seq or EM-seq) should have been used.

Reply> *We additionally analyzed 5mC by Enzymatic methyl-seq (EM-seq) as suggested by the Reviewers (This was also mentioned by Reviewer 4). Our additional experiments using EM-seq showed that two gene expression regulators (pcrA and rpsD) are differentially methylated between Agr positive and EA-Agr. The new data is shown in the new Figure 5a, 5b and new Extended data table 10. We consequently removed the old Extended data tables 7 and 8. The new data are described on pages14-15, lines 335-344 and page18 lines 419-428.*

- Bacterial DNA-MTases are known to methylate DNA after recognition of specific DNA motifs. Therefore, methylome analysis of a DNA-MTase deletion mutant in bacteria usually leads to differentially methylated sites located in a given motif. Here in the Gram-positive S. aureus, the identified differentially methylated sites on the DNA are randomly distributed in the genome, not located in a particular motif, and this might be linked to MraW. MraW has been primarily described to have rRNA methylation function. Its putative role in DNA methylation modification has been shown in one study in the Gram-negative bacterium E. coli. Whether a similar function exists in S. aureus remains to be shown.

Reply> *Thank you for your insightful comments. We conducted additional analyses on rRNA*

methylation. There was no change in the state of rRNA cytosine methylation between Agr positive and EA-Agr CC1 lineage subclones. However, in the mraW mutant, rRNA methylation was severely impaired in the knockout strain and partially rescued in the reconstituted strain. These results suggest that the phenomena observed in the mraW mutant do not fully reflect the characteristics of EA-Agr, but our additional results at least indicate the importance of DNA cytosine methylation in EA-Agr. We have included the results and limitations of these experiments in the main text (page16, line370-377) and added new Extended data figure7.

- The stability of the 16S RNA should be studied, since it is possible that the structure of 16S RNA is modified in the mraW deletion mutant. This could impact various physiological processes in the cells.

Reply> *16SrRNA/23rRNA ratio detected by Bioanalyzer was not changed between CN02ΔmraW+pTX16 and CN02ΔmraW+pTXmraW. This data is shown in the new Extended data figure 7a.*

Other comments:

- line 261-262: there are no mention of m5C in S. aureus in the cited reference. Another ref should be used here.

Reply> *We replaced the reference which mentioned of the m5C in S. aureus (ref.30, page14, line 318).*

- Are the differentially methylated sites on the DNA related to genes with differential expression?

Reply> *According to the EM-seq result, the genes pcrA (ATP-dependent DNA helicase) and rpsD (30S ribosomal protein S4), which are involved in gene expression regulation, exhibited increased 5mC methylation in Agr positive compared to EA-Agr subclones. Quantitative RT-PCR analysis revealed that expressions of pcrA and rpsD were significantly increased in EA-Agr subclones compared to Agr positive subclones in the late growth phase (4 hours culture) of low aeration condition. New data were shown in Fig. 5b and described on page 15, lines 342-344.*

-Table7: it is unclear what are the values indicated in columns D-I. Does this table present predicted DNA-methylated sites or just DNA-modification?

Reply> *We consequently removed the old Extended data tables 7 and 8.*

- Table 8: Exact values of the samples should be shown. Please do not present an average of only 2 samples.

Reply> *We consequently removed the old Extended data tables 7 and 8.*

- Fig4C. CN07 is mapping very close to CN02 and CN06. What is the proposed explanation for that?

Reply> *We consequently removed the old Figure4 C.*

Reviewer #4 (Remarks to the Author):

I commend the research team for their great efforts in this infant cohort-based study on a MRSA outbreak and the comprehensive genome sequencing and comparative genomics analysis. This work is particularly relevant given the rising concern over MRSA as a critical pathogen contributing to increased mortality and antibiotic resistance. It is also exciting to see the manuscript implicate the possible role of epigenetic regulation in MRSA adaption to hospital environment because bacterial epigenetics represents a new direction to better understand bacterial adaption. However, several major issues in the manuscript warrant further investigation and clarification before re-evaluation.

Reply> *Thank you very much for your positive and insightful comments on our manuscript. We greatly appreciate your recognition of our work.*

1. The genetic diversity among the studied clones is substantial. Even though there are no consensus differences as emphasized by the authors, the collective impact of multiple genetic

variations on the phenotype cannot be ruled out. Comparing several isolates and clones with numerous genetic differences falls short in convincingly attributing the observed traits solely to factors like Agr and the RM system. This is a general challenge that is worth careful consideration by the authors.

Reply> *Thank you for your valuable comment. We additionally compared the HiFi sequencing (PacBio Sequel IIe) data of CN02^{agr+} and CN07^{EA} corrected with Illumina sequencing data, the full length of the coding sequence was completely identical between these two subclones and only one intragenic mutation was detected in these 2 subclones. Therefore, we believe the collective impact of multiple genetic variations on the phenotype can be ruled out. The new analysis results are described on page 7, lines 170-173 and the method is described on pages 23-24, lines 546-549. Based on this analysis, the genetic tree of Figure 1c was also revised.*

2. The use of culture media to study Agr phenotypes, while informative, may not adequately mimic the environmental selective pressures the study aims to elucidate. If epigenetics is indeed mediating the reversible suppression of Agr, its effects should be susceptible to changes in culture conditions (from MRSA real environmental conditions). Unless some underlying upstream genetic factors stabilize this difference, which would then imply that

epigenetics is only an intermediary in this case.

Reply> *We acknowledge the reviewer's point that further investigation into the upstream mechanisms is necessary. Therefore, we have discussed the limitations of our study regarding this matter in the discussion section (page18, lines 419-428).*

3. The RNAseq data showed an upregulation of hsdM/S, the methyltransferase and specificity unit, but a downregulation of Type I Restriction Enzyme. Typically, Type I RMS systems are regulated in concert, so this divergence in gene expression trends warrants careful scrutiny to rule out any potential artifacts in the RNAseq data analysis.

Reply> *We agree with the Reviewer. We additionally performed qPCR analysis of hsdR and hsdM and confirmed that hsdR expression is suppressed by Agr suppression condition while no difference was detected in hsdM. Therefore, we concluded that the decreased hsdR expression in the restriction-modification system may contribute to the acquirement of foreign resistance genes in agr mutant and EA-Agr subclones. We revised the text based on this new data (page11, lines 251-257) and added a new Extended data figure 4c.*

4. Regarding the detection of 5mC methylation, I'm concerned about the reliability of the PacBio sequencing data. PacBio is great at detecting 6mA and 5mC, but the signal for 5mC is

generally modest and can be influenced by sequencing depth (deeper sequencing tends to call higher 5mC %), leading to potentially misleading interpretations of methylation percentages. For a claim as critical as the epigenetic regulation of 5mC, more robust methods like bisulfite sequencing or EM-seq (or even better: NEBNext® Enzymatic Methyl-seq Kit) should be employed to validate these findings.

Reply> *We additionally analyzed 5mC by enzymatic methyl-seq (EM-seq) as suggested by the Reviewers (This was also mentioned by Reviewer 3). Our additional experiments using EM-seq showed that two gene expression regulators (pcrA and rpsD) are differentially methylated between Agr positive and EA-Agr. The new data is shown in the revised new Figure 5a, 5b and new Extended data table 10. We consequently removed the old Extended data tables 7 and 8. The new data are described on pages 14-15, lines 335-344 and page18 lines 419-424.*

5. While the genetic investigation of the 5mC methyltransferase MraW is a step in the right direction, natural regulation of 5mC methyltransferase and complete deletion mutations are not equivalent. Confirming the extent of 5mC changes remains pivotal. Additionally, understanding the prevalence of the MraW gene across different MRSA strains (is MraW common or rare) could help understand how broadly applicable are the findings in this study.

Reply> *Thank you for your insightful comments. We have confirmed that mraW is conserved in the genomes of 860 S. aureus strains in the database (pages 20-21, lines 475-479). However, as shown in the rRNA methylation results conducted following Reviewer 3's suggestion, mraW-depletion also affects rRNA methylation. These results suggest that the phenomena observed in the mraW mutant do not fully reflect the characteristics of EA-Agr, but our additional results at least indicate the importance of DNA cytosine methylation in EA-Agr. We have included the results and limitations of these experiments in the main text (page 16, lines 370-377 and page 18, lines 424-428) and added new Extended data figure 7.*

6. The phenotypic disparities between CN08 and CN31, and the unique RM system motifs in CN31, add complexity to the interpretation of the underlying genetic and epigenetic drivers.

Reply> *Although CN31 showed different 6mA motifs from CN08, these strains showed the same phenotype in vivo and in vitro. Additionally, CovSA strains (ST15 and ST45) which showed different 6mA motifs from CC1 outbreak subclones also developed EA-Agr subclones (Revised Figure 4 and new Extended data table 9). While the 6mA methylation pattern may occasionally change downstream of agr mutation, we believe it is not directly related to Agr phenotype.*

In summary, while the study presents some promising avenues of research, its current form leaves room for ambiguity and requires additional experiments and analyses to solidify its conclusions.

REVIEWERS' COMMENTS

Reviewer #1 (Remarks to the Author):
The manuscript by Yamazaki et al was considerably improved during the revision process, gaining in clarity and precision. It is now much easier to understand the links between the different isolates of the outbreak and their phenotypes. I therefore have few additional comments for the authors. My main comments on the revised version are as follows:
- A graphical representation of the main molecular mechanisms identified in the study linking the contribution of PcrA, RpsD, HsdM,S,R and MraW would be useful.

Minor comments:
- line 56: Abstract: I think some words are missing (In this work, we discovered how alteration...)
- line 96: Can you briefly present the background to this screening (following cases of severe infection)?
- line 98: Is 17 the total or a selection?
- line 142: 4h is sufficient if in a “rich culture medium”.
- line 172-174: identical coding sequences, but what about regulatory sequences? We can only conclude that it is “more likely” that this is not a mutation.
- line 194: a word seems to be missing from this sentence
- line 227 and line 379: did you test the control strains LACwt and LACdelta-agr (as in Extended data figure 3 a) in the gentamycin protection assay in macrophages? LACdelta-agr generally survives better in cells, including macrophages, than WT (cf 10.1111/mmi.15184). Since you've observed the opposite with clinical strains, can you comment on this?
- line 236-7: this sentence is vague. Can you be more precise?
- line 253: some words are missing
- line 292, line 304: since the skin isolates were taken from 6-month-old children, I'd prefer “skin-isolated” or “collected from skin in non-hospitalized children” instead of “skin-adapted”.
- line 384: “methylation controls the EA-Agr...” a statement too strong here.

Comments on behalf of Reviewer #2:
*The qPCR analysis in response to Reviewer #4.
The reviewer 4 concern regarding RNAseq data showing a decreased expression of hsdR but not of hsdM/S between Agr expressing/suppressed clones in the initial version of the manuscript was legitimate.
In the current version of the manuscript, the authors have confirmed this differential expression of hsdR using RT-qPCR. It is slight, but significant (extended fig 4c).
Therefore, the technical limitation of RNAseq alone can now be ruled out.
Although hsdMSR can be transcribed as an operon, the existence of a specific regulation of hdsR (by a specific promoter upstream from hsdR, or by putative small interfering RNAs), or differences in mRNA half-life may exist. Anyway, the slight difference observed here seems genuine and therefore the proposed conclusion of the author (softened by the word “suggest”) is fine, in my opinion.

Reviewer #3 (Remarks to the Author):

I am satisfied by the additional information to address the issues that were raised.

Reviewer #1 (Remarks to the Author):

The manuscript by Yamazaki et al was considerably improved during the revision process, gaining in clarity and precision. It is now much easier to understand the links between the different isolates of the outbreak and their phenotypes. I therefore have few additional comments for the authors. My main comments on the revised version are as follows:

Reply> *Thank you very much for your positive feedback on our revised manuscript. We appreciate the reviewer's comments acknowledging the improvements in clarity and precision, as well as the reviewer's recognition of the strengthened connections between the outbreak isolates and their phenotypes. The reviewer's insights have greatly contributed to enhancing the quality of our work.*

In response to the reviewer's latest comments, we have carefully prepared replies to each of the points the reviewer raised. Please find our detailed responses below, addressing the reviewer's comments on the revised version.

- A graphical representation of the main molecular mechanisms identified in the study linking the contribution of PcrA, RpsD, HsdM,S,R and MraW would be useful.

Reply> *As per the reviewer's suggestion, we created a graphical representation and added it to Yamazaki_Supplementary Figure 8.*

Minor comments:

- line 56: Abstract: I think some words are missing (In this work, we discovered how alteration...)

Reply> *We added some words as the reviewer suggested in line 56.*

- line 96: Can you briefly present the background to this screening (following cases of severe infection)?

Reply> *This screening is performed as a routine measure since before the outbreak. We added this sentence on page 19, line 435-436.*

- line 98: Is 17 the total or a selection?

Reply> *We added the sentence "All isolates from Period A were included in the analysis without exclusion" on page 19, line 439-440.*

- line 142: 4h is sufficient if in a "rich culture medium".

Reply> *We added the sentence "in rich culture medium" on page 6 line 143.*

- line 172-174: identical coding sequences, but what about regulatory sequences? We can only conclude that it is “more likely” that this is not a mutation.

Reply> *We also analyzed the regulatory sequences using the same method. Therefore, we revised the sentence to read, “the full length of the coding and regulatory sequences was completely identical between these two subclones,” on page 7, lines 173–175.*

- line 194: a word seems to be missing from this sentence.

Reply> *Thank you for pointing out the error. We added the sentence “and when the bacteria was” on page 8 line 194.*

- line 227 and line 379: did you test the control strains LACwt and LACdelta-agr (as in Extended data figure 3 a) in the gentamycin protection assay in macrophages? LACdelta-agr generally survives better in cells, including macrophages, than WT (cf 10.1111/mmi.15184). Since you've observed the opposite with clinical strains, can you comment on this?

Reply> *As we mentioned in the paper using three references, Agr-positive S. aureus is believed to survive more effectively within macrophages, likely due to factors such as PSMs that are produced in an Agr-dependent manner (Ref. 25, 26, 27, 28). On the other hand, as noted in the paper referenced by the reviewer (doi: 10.1111/mmi.15184), there is research indicating that in the LAC strain, Agr-deleted mutant has a survival advantage within phagocytes (THP-1 cells). In this study, lysostaphin was used in addition to gentamicin for initial S. aureus killing. Gentamicin acts by binding to bacterial ribosomes and inhibiting protein synthesis, thereby preventing their growth. In contrast, lysostaphin kills S. aureus by cleaving the pentaglycine cross-bridges between stem peptides in the peptidoglycan. Lysostaphin-killed S. aureus (lysibody) has been reported to enhance the natural functions of phagocytes (doi: 10.1128/AAC.01056-18) and this may contribute to variations in the role of Agr in phagocytic cell survival reported across different studies.*

- line 236-7: this sentence is vague. Can you be more precise?

Reply> *We revised the sentence “Comparison of RNA-sequencing results from samples collected under three different culture conditions revealed that the majority of genes exhibit changes dependent on the culture conditions (Extended data fig. 4a, top).” on page 10, line 236-8.*

- line 253: some words are missing

Reply> *Thank you for pointing out the error. We added the sentence "the involvement of" on page 11 line 253.*

- line 292, line 304: since the skin isolates were taken from 6-month-old children, I'd prefer "skin-isolated" or "collected from skin in non-hospitalized children" instead of "skin-adapted".
Reply> We revised the sentences (line 292 and 304) "skin-isolated S. aureus strains".

- line 384: "methylation controls the EA-Agr..." a statement too strong here.
Reply> *We revised the sentence, "genomic cytosine methylation may potentially control the EA-Agr phenotype" on page 16, line 384-5.*

Comments on behalf of Reviewer #2:
*The qPCR analysis in response to Reviewer #4.
The reviewer 4 concern regarding RNAseq data showing a decreased expression of hsdR but not of hsdM/S between Agr expressing/suppressed clones in the initial version of the manuscript was legitimate.
In the current version of the manuscript, the authors have confirmed this differential expression of hsdR using RT-qPCR. It is slight, but significant (extended fig 4c).
Therefore, the technical limitation of RNAseq alone can now be ruled out.
Although hsdMSR can be transcribed as an operon, the existence of a specific regulation of hdsR (by a specific promoter upstream from hsdR, or by putative small interfering RNAs), or differences in mRNA half-life may exist. Anyway, the slight difference observed here seems genuine and therefore the proposed conclusion of the author (softened by the word "suggest") is fine, in my opinion.
Reply> *Thank you for your positive feedback and for acknowledging the revised RT-qPCR analysis in response to Reviewer #4's comments. We appreciate your recognition of our efforts to confirm the differential expression of hsdR between Agr-expressing and Agr-suppressed clones, which indeed aligns with the RNA-seq data.*

Reviewer #3 (Remarks to the Author):
I am satisfied by the additional information to address the issues that were raised.
Reply> *Thank you for your positive feedback.*