nature portfolio

Peer Review File

Analysis of Histone Modification Interplay Reveals Two Distinct Domains in Facultative Heterochromatin in *Pyricularia oryzae*

Corresponding Author: Professor Hitoshi Nakayashiki

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)
Comments for the Authors (# COMMSBIO-24-8221)
The manuscript under review presents a thorough investigation into the complex crosstalk of histone modifications in the phytopathogenic fungus Pyricularia oryzae. The authors have skillfully employed bioinformatics analysis of ChIP-seq and RNA-seq data to reveal the impact of specific histone methylation marks on gene expression and PTM interplay. A particularly noteworthy finding is the identification of two distinct facultative heterochromatin compartments, K4-FH and K9-FH, which exhibit differential responses to the loss of specific histone methyltransferases. The study's findings not only advance the understanding of epigenetic regulation in fungi but also highlight the potential role of these histone marks in fungal pathogenicity. Overall, the data presented are of high quality and are interpreted with scientific rigor. I have only a few minor comments that might help further clarify this important finding.

Specific comments:
1. The title and keywords should utilize the species name Pyricularia oryzae.
2. Are MoSet1, MoKmt1, and MoKmt6 the sole enzymes corresponding to the respective modifications? If not, could residual modifications in the mutants potentially impact the results?
3. Could the specific locations of EC, FH, CH, K4-FH, and K9-FH be annotated on the chromosome map?
4. The authors should analyze the A/T levels in the DNA of FH and CH regions to determine if there is a correlation.
5. The authors could discuss whether DNA methylation in regions undergoing FH and CH changes might be altered.
6. Is a 1kb resolution sufficient for analyzing Histone PTMs?
7. What is the relationship and biological significance between "K4-FH and EC" and "K9-FH and CH"? Why are there two distinct trends?
8. Does Fig.2 only analyze chromosome chr1, or all chromosomes? The same question applies to RNA-seq.
9. RNA-seq should undergo statistical analysis to eliminate genes with low significance.
10. Regarding Fig.6c, do genes with low H3K4me2 levels show large FC differences due to low basal expression?
11. Why does Δmoset1 lead to a general increase in global transcription? Are there issues with the internal reference or quantification?
12. In Δmokmt6, there are significant wide alterations in H3K9me3 and RNA levels in the CH region. Could the authors provide a detailed analysis for the changes in gene expression and nearby H3K9me3 correspond? Additionally, why does Fig. 7 primarily analyze changes in H3K4me2 and H3K27me3, but not H3K9me3?
13. Could a functional analysis be conducted on the differentially expressed genes in each mutant?

Reviewer #2

(Remarks to the Author)
Dang et al. investigated the interplay between histone modifications H3K4me2/3, H3K9me3, and H3K27me3 in Pyricularia oryzae. They analyzed the ChIP-seq data of knock-out mutants of histone modification enzymes and found two distinct domains in facultative heterochromatin (K4-FH, adjacent to EC, and K9-FH, adjacent to CH) within the genome of the fungal pathogen. They examined the effect of the loss of a specific PTM on other PTMs. They also analyzed RNA-seq data and

identified a correlation between gene expression and changes in histone modifications. The study is of interest to epigenetic researchers and fungal biologists. The major concern is that this relationship has not been verified through experiments. The writing of this manuscript needs to be improved to enhance the understanding of the interplay between histone modifications. Additionally, they did not analyze the impact of histone crosstalk on fungal pathogenicity. Suggestions are listed below.

1. Verify their findings in selected regions through experimental methods.
2. Analyze known genes critical for fungal pathogenicity using RNA-seq data and investigate the role of histone interplay in fungal pathogenicity.
3. While the definitions of EC-, CH-, and FH-compartments are clear, the location and distribution of K4-FH and K9-FH are not well-defined. A figure or table illustrating the location of each compartment may help readers understand the relationship between the histone modifications in WT or mutants.
4. The manuscript would be easier to understand if the authors described their findings more clearly by improving their writing. For example, the long and complex phrase “These compartments differed in that H3K27me3 levels in K4-FH, but not in K9-FH, were significantly reduced by the loss of MoKmt1, the enzyme catalyzing H3K9me3” could be simplified.
5. Numerous abbreviations in the manuscript also hinder readability and comprehension. It is reasonable to include abbreviations in figures but not in the main text, such as “FC,” “EC,” “FH,” and “CH.”

Reviewer #3

(Remarks to the Author)
In this study, the authors profile three histone post-translational modifications (PTMs) in wild type and histone MTase (KMT) deletion strains of the plant pathogen P. oryzae (strain Br48). The authors examine H3K27me3, H3K9me3, and H3K4me2, classical markers of facultative heterochromatin, constitutive heterochromatin, and euchromatin, respectively. The authors report that loss of individual KMTs leads to altered patterns in PTMs catalyzed by MTases, showing interplay between these various histone marks. The most striking phenotype was observed for H3K27me3 in the mutant lacking the H3K9 MTase. Specifically, H3K27me3 is lost from typical facultative heterochromatin domains and accumulates in constitutive heterochromatin regions, as previously reported for Neurospora and Fusarium. The authors further subdivide constitutive heterochromatin into two sub-types (K4-FH and K9-FH), and report that H3K27me3 is lost only the K4-FH subtype in the Δkmt1 mutant. This is potentially an interesting observation, but it is currently unclear if some major conclusions are supported by the data. There are significant concerns about the analytical strategies used in the study.

Major concerns:

ChIP-seq data are not analyzed by established methods. Normalized Read Counts (RPM) were determined for genome windows (1kb). These windows were then placed into categories based on overlap with called peaks (performed using the HOMER software package).

Additional details of peak calling parameters are needed. What fold-change cut-off was used. What statistical threshold was used?
The locations of called peaks should be provided in supplementary Tables.
Basic features of modification enrichment profiles should be provided based on this peak calling. How many peaks are called for each modification? What is the average peak size for each modification. What is the distribution of peak sizes for each modification? Does H3K4me2 exhibit an enrichment profile in the 5’ end of genes? Do K7me3 peaks typically span multiple genes as in other fungi or do these peaks typically cover single-gene peaks as in plants? Do K9me3 and K27me3 peaks overlap or are these modifications mutually exclusive.
The approach for assigning windows to different chromatin types is flawed. The strategy involves first calling peaks for K9me, K27me, and K4me and then assigning 1kb windows to chromatin types if the window overlaps with called peaks.

Any overlap results in a window being classified based on the overlapping modification. This strategy will artifactually increase the sizes of called peaks, confounding downstream analysis This low overlap threshold will create major problems calculating differential enrichment at windows that span two types of chromatin.
If a 1kb fragment overlapped the boundary between two different modification peaks, the fragment was placed in the category with the highest RPM. This is not appropriate. ChIP-efficiency can vary between antibodies. In this case, the difference in RPM could reflect the affinity of the antibody rather than the level of enrichment. Did RPM calculations account for differences in ChIP-efficiency between antibodies? As described, the approach used could lead to systemic biases in classifying boundary regions. For example, windows on the boundary of K9 and K27me3 peaks could be artifactually classified as FH. This appears to be the case In figure 3, where many K9-FH regions show a decrease in K9 in the Δkmt6 mutants. This suggests that these K9-FH windows should have been classified as CH regions or that the windows span the boundary of the K9 and K27 peaks.
It would be more appropriate to quantify enrichment within called peaks or to restrict analysis to windows that are contained completely within called peaks to avoid artifacts at boundaries.
A second major concern is that peak calling parameters for H3K4me2 were not appropriate based on work from other fungi. In line 620 – The authors write, “since H3K4me2 tends to from broader regions than K9me3 or K27me3, the parameters –size 500 and –minDist 10000 were selected” - This statement is not accurate. H3K4me2 is typically enriched in discrete peaks covering the promoters and 5’ ends of single genes. In yeast, the typical size of an H3K4me2 peak is 200-500 base pairs. By setting the “-minDist” parameter to 10,000, multiple K4me2 peaks are likely being artifactually combined into single

large peaks. The result will be an overestimation of regions classified as EC or K4-FC, leading to a second type of systemic error. The analysis should be performed using appropriate parameters for H3K4me2 peaks.

Throughout the manuscript, statistical thresholds should be applied when plotting differential enrichment values. For example, scatter plots such as those in Fig 2C, are challenging to interpret. The yellow points corresponding to K4-FH appear to be plotted on top of points corresponding to FH-K9. One way to correct this would be to plot points corresponding only to windows with statistically significant differences in enrichment. Alternatively, the spots should be shaded in grey or made transparent. These concerns could be addressed by used validated software for differential enrichment analysis (e.g. diffBind or csaw)

What does this mean? - Line 632 - “The border between K4-FH and K9-FH was determined manually with reference to the FC value of H3K27me3 in the Δmokmt1 relative to WT.

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
My questions were answered.

Reviewer #2

(Remarks to the Author)
The authors have addressed my concerns.

Reviewer #3

(Remarks to the Author)
No further changes requested. Recommend accept.

Response to reviewer;

We greatly appreciate the extensive comments from the reviewers. We found them very helpful for improving our manuscript. We have tried our best to take them into account in the revision of our manuscript. Below are our point-by-point responses to the comments.

Reviewer #1 (Remarks to the Author):

Comments for the Authors (# COMMSBIO-24-8221)
The manuscript under review presents a thorough investigation into the complex crosstalk of histone modifications in the phytopathogenic fungus Pyricularia oryzae. The authors have skillfully employed bioinformatics analysis of ChIP-seq and RNA-seq data to reveal the impact of specific histone methylation marks on gene expression and PTM interplay. A particularly noteworthy finding is the identification of two distinct facultative heterochromatin compartments, K4-FH and K9-FH, which exhibit differential responses to the loss of specific histone methyltransferases. The study's findings not only advance the understanding of epigenetic regulation in fungi but also highlight the potential role of these histone marks in fungal pathogenicity. Overall, the data presented are of high quality and are interpreted with scientific rigor. I have only a few minor comments that might help further clarify this important finding.

Specific comments:
> 1. The title and keywords should utilize the species name Pyricularia oryzae.

We have included the species name in the title as suggested.

> 2. Are MoSet1, MoKmt1, and MoKmt6 the sole enzymes corresponding to the respective modifications? If not, could residual modifications in the mutants potentially impact the results?

Based on the ChIP-seq analysis (Supplementary Fig. S1), almost no peaks of the corresponding histone modification were detectable in the mutant of either MoSet1, MoKmt1, or MoKmt6, indicating that these enzymes are indeed the primary, if not sole, catalysts of the respective modifications.

> 3. Could the specific locations of EC, FH, CH, K4-FH, and K9-FH be annotated on the chromosome map?

We have shown that the locations of EC, FH, CH, K4-FH, and K9-FH (now EC, fHC, cHC, K4-fHC and K9-fHC) by dots with different colors in Fig. 1A and Fig. 7A.

> 4. The authors should analyze the A/T levels in the DNA of FH and CH regions to determine if there is a correlation.

Thank you for the nice suggestion. We have included the data in Fig. 3D.

> 5. The authors could discuss whether DNA methylation in regions undergoing FH and CH changes might be altered.

We have added Fig. 3E to show the average levels of DNA methylation in the genomic compartments including FH and CH, and given some discussion on it in the text.

> 6. Is a 1kb resolution sufficient for analyzing Histone PTMs?

Since sheared chromatin fragments ranged from approximately 100 to 500 bp in our ChIP-seq experiments, we believe that bioinformatic analysis at a 1kb resolution is reasonable.

> 7. What is the relationship and biological significance between "K4-FH and EC" and "K9-FH and CH"? Why are there two distinct trends?

Currently, we do not have the answer. We would like to address this point in the future study.

> 8. Does Fig.2 only analyze chromosome chr1, or all chromosomes? The same question applies to RNA-seq.

In general, data from all chromosomes are included in the figures except Fig.1A and Fig. 7A.

> 9. RNA-seq should undergo statistical analysis to eliminate genes with low significance.

Yes. We have performed statistical analysis on our RNA-seq data and mostly only DEGs are used in the analysis.

> 10. Regarding Fig.6c, do genes with low H3K4me2 levels show large FC differences due to low basal expression?

We believe that the answer is yes. These genes are located in facultative heterochromatic regions, where gene expression is generally repressed.

> 11. Why does Δmoset1 lead to a general increase in global transcription? Are there issues with the internal reference or quantification?

We do not currently have an answer to this question. In fact, when we compared total cellular RNA levels between WT and Δmoset1, we did not observe substantial differences. This supports that transcript levels are generally increased in Δmoset1. We are still working on this issue.

> 12. In Δmokmt6, there are significant wide alterations in H3K9me3 and RNA levels in the CH region. Could the authors provide a detailed analysis for the changes in gene expression and nearby H3K9me3 correspond? Additionally, why does Fig. 7 primarily analyze changes in H3K4me2 and H3K27me3, but not H3K9me3?

Based on Supplementary Fig. 6C, the Δmokmt6 mutation had only a modest impact on gene expression in the CH (now cHC) region. Since the CH region contains only 68 genes and the loss of H3K9me3 appears to have only a minor effect on their expression, we believe that H3K9me3 itself does not play a significant role in regulating cellular genes but may have roles in repressing transposable elements in this fungus.

> 13. Could a functional analysis be conducted on the differentially expressed genes in each mutant?

While this point warrants future investigation, we believe it lies beyond the scope of this study.

Reviewer #2 (Remarks to the Author):

> 1. Verify their findings in selected regions through experimental methods.

We performed ChIP-seq analysis of H3K27 levels under different culture conditions, and found that K27 levels in the K4-FH region are more responsive to environmental changes than those in the K9-FH region (Fig. 3G). These observations provide experimental verification of our major finding that K4-FH and K9-FH are functionally distinct heterochromatic regions in *P. oryzae.*

> 2. Analyze known genes critical for fungal pathogenicity using RNA-seq data and investigate the role of histone interplay in fungal pathogenicity.

We have performed RNA-seq analysis on effector-like genes According to the suggestion by the reviewer, we have analyzed the RNA-seq data focusing on effector-like genes, which are believed to be critical for fungal pathogenicity, during infection (Fig. 3I). The results indicate that approximately half of the upregulated effector-like genes are located in the K4-FH compartment, suggesting that genomic regions shaped by histone interplay contribute to fungal pathogenicity. In addition, our previous results from inoculation tests of Δmoset1, Δmokmt1, and Δmokmt6 (Pham et al., 2015) are included in the text (P131-134) to clarify their contributions to fungal pathogenicity.

> 3. While the definitions of EC-, CH-, and FH-compartments are clear, the location and distribution of K4-FH and K9-FH are not well-defined. A figure or table illustrating the location of each compartment may help readers understand the relationship between the histone modifications in WT or mutants.

We redefined K4-FH and K9-FH based on which of the adjacent EC or CH regions each FH segment is closer to. A schematic diagram illustrating how these compartments are defined is provided in Supplementary Fig. 3D.

> 4. The manuscript would be easier to understand if the authors described their findings more clearly by improving their writing. For example, the long and complex phrase “These compartments differed in that H3K27me3 levels in K4-FH, but not in K9-FH, were significantly reduced by the loss of MoKmt1, the enzyme catalyzing H3K9me3” could be simplified.

We apologize if our English is not as fluent as that of a native speaker. We have revised the manuscript throughout to simplify the language, including the example mentioned by the reviewer.

> 5. Numerous abbreviations in the manuscript also hinder readability and comprehension. It is reasonable to include abbreviations in figures but not in the main text, such as “FC,” “EC,” “FH,” and “CH.”

Thank you for the suggestion. However, without the use of abbreviations, defining K4-FH or K9-FH would be problematic. To improve readability, we have replaced FH, CH and FC with fHC, cHC, and fold change, respectively, throughout the text.

Reviewer #3 (Remarks to the Author):

Major concerns:

> ChIP-seq data are not analyzed by established methods. Normalized Read Counts (RPM) were determined for genome windows (1kb). These windows were then placed into categories based on overlap with called peaks (performed using the HOMER software package).

Rather than focusing on selected called peaks, we would like to include every 1-kb segment of the *P. oryzae* genome to enable a comprehensive analysis. We believe this approach allowed novel insights (see Fig. 1b, for example), particularly into the interplay of histone modifications that may not be captured using standard methods. Called peaks were used solely to define genomic compartments; therefore, detailed characterization of called peaks such as peak numbers or enrichment profiles of individual histone modifications was not central to our analysis.

> Additional details of peak calling parameters are needed. What fold-change cut-off was used. What statistical threshold was used?

We used the default parameter settings of the HOMER software package, except for the parameter "-style histone".

> The locations of called peaks should be provided in supplementary Tables.

The locations of called peaks for H3K4me2, H3K9me3 and H3K27me3 are provided in Supplementary Table 2-4.

> Basic features of modification enrichment profiles should be provided based on this peak calling. How many peaks are called for each modification? What is the average peak size for each modification. What is the distribution of peak sizes for each modification? Does H3K4me2 exhibit an enrichment profile in the 5' end of genes? Do K27me3 peaks typically span multiple genes as in other fungi or do these peaks typically cover single-gene peaks as in plants? Do K9me3 and K27me3 peaks overlap or are these modifications mutually exclusive.

As noted above, our analysis did not specifically focus on called peaks. However, basic features of the called peaks are now provided in Supplementary Table 2-4, or can be calculated from the data, for those interested. Furthermore, some of the information requested by the reviewer is already available in previously published studies (Pham et al., 2015; Zhang et al., 2021; Kobayashi et al., 2023).

> The approach for assigning windows to different chromatin types is flawed. The strategy involves first calling peaks for K9me, K27me, and K4me and then assigning 1kb windows to chromatin types if the window overlaps with called peaks.
Any overlap results in a window being classified based on the overlapping modification. This strategy will artifactually increase the sizes of called peaks, confounding downstream analysis
This low overlap threshold will create major problems calculating differential enrichment at windows that span two types of chromatin.
If a 1kb fragment overlapped the boundary between two different modification peaks, the fragment was placed in the category with the highest RPM. This is not appropriate. ChIP-efficiency can vary between antibodies. In this case, the difference in RPM could reflect the affinity of the antibody rather than the level of enrichment. Did RPM calculations account for differences in ChIP-efficiency between antibodies? As described, the approach used could lead to systemic biases in classifying boundary regions. For example, windows on the boundary of K9 and K27me3 peaks could be artifactually classified as FH. This appears to be the case In figure 3, where many K9-FH regions show a decrease in K9 in the Δkmt6 mutants. This suggests that these K9-FH windows should have been classified as CH regions or that the windows span the boundary of the K9 and K27 peaks.
It would be more appropriate to quantify enrichment within called peaks or to restrict analysis to windows that are contained completely within called peaks to avoid artifacts at boundaries.

Thank you for the thoughtful suggestions. We have re-analyzed all the data in accordance with the reviewer's recommendations (see Materials and Methods). To avoid boundary issues, a

1-kb fragment is assigned to a genomic compartment if more than 80% of its sequence falls within a called peak. In addition, 1-kb fragments that overlap the boundary between two different modification peaks have been reassigned to UA.
The reviewer mentioned “For example, windows on the boundary of K9 and K27me3 peaks could be artifactually classified as FH. This appears to be the case In figure 3, where many K9-FH regions show a decrease in K9 in the Δkmt6 mutants. This suggests that these K9-FH windows should have been classified as CH regions or that the windows span the boundary of the K9 and K27 peaks.”. We believe that this stems from some misunderstanding by the reviewer since the previous figure 3 represented data from the Δmoset1 mutant and the K9 levels in CH regions were rather elevated in the Δkmt6 mutant (previous Fig. 4).
Overall, the main conclusions drawn from the re-analyzed data remain consistent with our previous findings. Therefore, we believe that the potential rate of miscategorization due to boundary issues, if any, was very low and did not affect our main conclusions.

> A second major concern is that peak calling parameters for H3K4me2 were not appropriate based on work from other fungi. In line 620 – The authors write, “since H3K4me2 tends to from broader regions than K9me3 or K27me3, the parameters –size 500 and –minDist 10000 were selected” - This statement is not accurate.

Thank you for pointing this out. What we had originally intended to express was “since euchromatin tends to from broader regions than constitutive heterochromatin or facultative heterochromatin, the parameters –size 500 and –minDist 10000 were selected”. We assume that euchromatin in the *P. oryzae* genome is not limited to H3K4me2 peaks but forms broader regions (see Fig. 1a).

> H3K4me2 is typically enriched in discrete peaks covering the promoters and 5’ ends of single genes. In yeast, the typical size of an H3K4me2 peak is 200-500 base pairs.

Broadly speaking, H3K4me2 is enriched gene body in the *P. oryzae* genome (Pham et al., 2015). In some cases, as reported in yeast, H3K4me2 covers the promoters and 5’ ends of single genes but in other cases, it may be more enriched in 3’ half of single genes.

> By setting the “-minDist” parameter to 10,000, multiple K4me2 peaks are likely being artifactually combined into single large peaks. The result will be an overestimation of regions classified as EC or K4-FC, leading to a second type of systemic error. The analysis should be performed using appropriate parameters for H3K4me2 peaks.

As we showed in the previous Supplementary Fig. 3, most UA segments between EC regions appeared to be regulated similarly to EC segments. Therefore, we assumed that the 'artifactually combined peaks' could belong to the same functional compartment. However, we have decided to follow the reviewer's suggestion, as it makes our conclusion more solid despite increasing the number of UA segments.

> Throughout the manuscript, statistical thresholds should be applied when plotting differential enrichment values. For example, scatter plots such as those in Fig 2C, are challenging to interpret. The yellow points corresponding to K4-FH appear to be plotted on top of points corresponding to FH-K9. One way to correct this would be to plot points corresponding only to windows with statistically significant differences in enrichment. Alternatively, the spots should be shaded in grey or made transparent. These concerns could be addressed by used validated software for differential enrichment analysis (e.g. diffBind or csaw)

We believe that presenting all data within compartments is important for illustrating overall trends, as data within a specific compartment generally follow a common pattern. However, as the reviewer rightly pointed out, overlapping data from different compartments can make interpretation difficult in scatter plots. Therefore, we have included separate plots for each compartment in the supplementary figure when needed.

> What does this mean? - Line 632 - "The border between K4-FH and K9-FH was determined manually with reference to the FC value of H3K27me3 in the Δmokmt1 relative to WT.

This point is also raised by reviewer #2. We redefined K4-FH and K9-FH based on which of the adjacent EC or CH regions each FH segment is closer to. A schematic diagram illustrating how these compartments are defined is provided in Supplementary Fig. 3D.