

Corisin Induces Proteostasis Stress to Drive Epithelial Injury and Pulmonary Fibrosis

Corresponding Author: Professor Esteban Gabazza

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)

The manuscript by Fujimoto et al, explores the role of corisin in epithelial injury and fibrosis. The manuscript is well written, easy to follow and provides compelling data that synthetic or endogenously expressed corisin can induce a number of profibrotic activities in pulmonary epithelial cells including induction of apoptosis, SASP secretion, and EMT phenotypes. The work is comprehensive exploring both cellular phenotypes in vitro with synthetic peptide, or expression driven by transfected plasmids. One caveat of the work however, is that corisin is part of a larger protein made by bacteria, and the larger protein does not have activity, which I don't think was made clear enough in the introduction. The authors Nat. Comm. paper in 2020 did show that corisin is at least identified in bacterial supernatants which proves that the corisin peptide can be generated by the microbiome. While much of the paper explores effects using A549 cells which are a cancer line that does not always mimic, normal AT2 cell behavior, the authors have repeated some key measures in primary human alveolar epithelial cells which adds important context and impact.

Major:

Please provide information in the introduction to make it clear that corisin is a peptide derived from a non-pathogenic larger bacterial protein, a transglycosylase. I think the introduction of this paper needs to clarify that the peptide has been found to be generated by the bacteria in measurable amounts in the lung to set the stage for the current work.

Because most of the data are with synthetic corisin or genetically expressed corisin, the data in Supplemental Fig. 12C seems to be highly relevant since the corisin was immunodepleted using the anti-corsin Ab. Does this Ab only immunodeplete the corisin peptide, or does it immunodeplete the whole transglycosylase enzyme? Is it known? The reason is because the authors tested the depleted BALF, but the pull down of corisin was not tested and would have been a nice addition if it pulled down active peptide. It seems that supplemental Figure 12 C should be in the main text and the pull down of corisin added if performed.

Minor:

line 789 has a reference not formatted correctly

Do the authors have hydroxyproline data for the fibrosis measurements to supplement the Ashcroft scoring?

Reviewer #2

(Remarks to the Author)

Fujimoto et al. presents the manuscript "Corisin induces proteostasis stress to drive epithelial injury and pulmonary fibrosis".

In this manuscript the authors present a comprehensive investigation into a microbiota-derived factor, corisin, and its role in epithelial injury and pulmonary fibrosis. They integrate molecular, cellular, and in vivo approaches to shed light on this peptide, and from their stepwise assessment they present corisin as a novel, potent peptide present in human bronchoalveolar lavage samples, exerting fibrosis driving activities in patients with idiopathic pulmonary fibrosis, but not in healthy controls.

The technical and methodological details are comprehensive and my review will focus on aspects of the study regarding the study participant data, the microbiota data, and the strength of certain interpretations. I here provide feedback regarding aspects that could be improved.

Human participant data:

The manuscript is framed in the context of human disease, progressive pulmonary fibrosis, and particularly idiopathic pulmonary fibrosis (IPF). Yet, only a limited part of the data is derived directly from participant bronchoalveolar lavage samples (BAL) (Figure 1, Supplementary Table 1 (Characteristics), and BAL experiments in Supplementary Fig. 12). The majority of the study relies on in vitro systems and mouse models. This is appropriate for mechanistic work, and the distinction between human-derived observations and experimental modeling is stated as a limitation. Unfortunately the information provided about the participants is not clear, and lacking important elements when studying the microbiota:

Supplementary Table 1 reports 14 IPF patients, and no healthy controls. Figure 1 includes only a subset (IPF n=8/14 control n=2/3), without explanation of this selection given in the Methods section or elsewhere in the manuscript.

When Figure 1 shows control h2 and h4, Supplementary Fig. 12: h1, h2, h3. and the source material providing no information about participants in Figure 1 it is very hard to understand the use of participant samples.

As the authors further state that corisin-encoding bacteria are present in all IPF patients, yet no comprehensive overview of all samples is provided, the statement should be adjusted to fit with what is shown. Corisin-related sequences are detected in both IPF patients and healthy controls. This raises important questions regarding disease specificity, like whether differences are quantitative rather than qualitative, and how presence in healthy individuals aligns with the proposed pathogenic role. This conceptual point would benefit from clearer discussion.

No clinical information for healthy controls is reported, despite their inclusion in analyses. Supplementary Table 1 should contain information about all 17 participants, and also whose samples were used. Why only some samples are analysed need to be addressed.

I also miss information about: Antibiotic use before sampling, exacerbation frequency the last year, time from diagnosis to sampling, degree of fibrosis in IPF patients, BAL yield.

Clarifying these points would improve transparency and allow assessment of representativeness.

Bronchoscopy and BAL samples:

There is very little information given about the BAL sampling procedure. During which time span were participants sampled? Were stable state collected before or after exacerbations? What were the safety criteria for performing the procedure both during stable and exacerbated states? How much liquid was instilled and were samples collected from the same anatomical location for all participants? Were any measures taken to lower the risk of contamination during sampling? Why were there not a processing (negative) control sample taken for each participant sample?

To improve the presentation of participant data it would also be very interesting to see the effect of acute exacerbations per participant. If possible I would encourage the team to address potential differences within individuals comparing their exacerbated and stable BAL.

Identification of bacteria and methodological limitations:

Bacterial identification is based on targeted PCR amplification of a short (~150 bp) fragment of a corisin-related gene, followed by sequencing and bioinformatic annotation. The authors should give information about the needed number of PCR cycles to achieve signal as BAL fluid is very sparse on microbiota DNA, and a high number of PCR cycles increase the chimera risk.

While amplicon sequencing is suitable for detecting specific gene families there are inherent limitations. The short sequence length and use of a non-standard taxonomic marker limit confidence in species-level identification. From reading the manuscript it seems still as if species level information was obtained from this amplicon sequencing. As a result, while the presence of corisin-like sequences is supported, the identification of specific bacterial species is more uncertain and should be interpreted cautiously.

Also page 37 line 991 says Shotgun sequencing. This is not described, and should be corrected to amplicon sequencing.

No overview of detected taxa across all samples is provided. Neither is the number of reads per sample for the taxa in Figure 1. With sparse data this should be conveyed.

For the microbiome assessment contamination should be considered as there seem to be only 1 negative control sample collected for the entire study. Is this the case decontamination tools like decontam in R should be used to make sure corisin producing bacteria is not results of laboratory contamination.

3) Statistical considerations

Several experiments (e.g. Figure 2, Figure 3) are based on relatively small sample sizes (n≈4–6). The methods description indicate that parametric or non-parametric tests were used depending on data distribution, but it is unclear which tests were applied to specific datasets. Also normality evaluations for as few data points as 4 is not well powered. The authors have used graphpad for their analyses and this tool provides the following on this topic:

https://www.graphpad.com/guides/prism/latest/statistics/choosing_parametric_vs_nonpar.htm

Given the small sample sizes, clearer reporting and acknowledgment of this limitation would improve the statistical assessments, as would running non-parametric tests for all data to check for robustness of the results.

Figure 5c: On page 11 the authors state that UMAP visualizations show clear separation between corisin-treated/non-treated cells, although samples are plotted separately and do not allow direct comparison.

Context of corisin research

It would be helpful to more clearly acknowledge that this research group appears to be the primary source of work on corisin. While this does not diminish the findings, it is important context for readers in evaluating novelty and independent validation.

Source document

Could the authors double check sheet names and content in the source material. Some columns are not labeled and some Figures seem not to be supported by content on corresponding sheet.

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)

The authors have adequately addressed my comments and the manuscript is significantly improved.

Reviewer #2

(Remarks to the Author)

I thank the authors for their careful and thorough revision. They have addressed the points raised in my review in a satisfactory manner.

In particular, I appreciate the clarification and expansion of the participant information in Supplementary Table 1. Regarding antibiotic use, if not possible to retrieve information at latest time for AB use, the authors can leave as is, but it would be interesting to know if AB use occurred within the last 90 days before sampling both stable and exacerbated state. Time from diagnosis to sampling is now perfectly clear, if possible time to first exacerbation should be added. Degree of fibrosis, and the details of the BAL sampling procedure are both well described.

The revised description of sample usage explaining that PCR was attempted on all available samples but sequencing was restricted to those meeting predefined amplicon-quality criteria resolves my earlier confusion about the subset of samples. Thank you for clarifying.

I am also pleased that the authors now report the PCR cycling conditions and acknowledge the chimera/artifact risk associated with low-biomass specimens, and that the taxonomic claims have been appropriately tempered to reflect sequence-similarity assignments rather than definitive species identification. The correction of "shotgun" to "amplicon" sequencing, the added sequencing-depth information, and the explicit description of the amplicon pooling strategy all improve transparency.

The added longitudinal paired stable-versus-exacerbation analyses (Supplementary Fig. 1) are a valuable addition, and the expanded discussion of disease specificity framing corisin as a context-dependent effector requiring "second hits" rather than a simple presence/absence determinant addresses my conceptual concern well. The uniform adoption of non-parametric tests and the acknowledgment that much of the corisin literature originates from this group are also appropriate and improve the balance of the manuscript.

I have no further concerns within the scope of my review.

Open Access This Peer Review File is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

In cases where reviewers are anonymous, credit should be given to 'Anonymous Referee' and the source.

The images or other third party material in this Peer Review File are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>

Response to Comments from Reviewer #1:

Comment 1

The manuscript by Fujimoto et al, explores the role of corisin in epithelial injury and fibrosis. The manuscript is well written, easy to follow and provides compelling data that synthetic or endogenously expressed corisin can induce a number of profibrotic activities in pulmonary epithelial cells including induction of apoptosis, SASP secretion, and EMT phenotypes. The work is comprehensive exploring both cellular phenotypes in vitro with synthetic peptide, or expression driven by transfected plasmids. One caveat of the work however, is that corisin is part of a larger protein made by bacteria, and the larger protein does not have activity, which I don't think was made clear enough in the introduction. The authors Nat. Comm. paper in 2020 did show that corisin is at least identified in bacterial supernatants which proves that the corisin peptide can be generated by the microbiome. While much of the paper explores effects using A549 cells which are a cancer line that does not always mimic, normal AT2 cell behavior, the authors have repeated some key measures in primary human alveolar epithelial cells which adds important context and impact.

Response

We sincerely thank the reviewer for the thoughtful, constructive, and encouraging comments. We are grateful for the reviewer's positive evaluation of the manuscript and for recognizing the comprehensive nature of the work, including the *in vitro* studies using synthetic and endogenously expressed corisin, as well as the validation of key findings in primary human alveolar epithelial cells. We also appreciate the important comment regarding the biological origin of corisin as a peptide released from a larger bacterial transglycosylase precursor protein. In response, we have revised the Introduction to clarify that corisin is generated by proteolytic release from this precursor protein, that the intact precursor protein itself does not reproduce the biological activity of corisin, and that corisin has previously been detected in bacterial culture supernatants and in measurable amounts in BALF. **Please refer to pages 5, lines 108-129.**

Comment 2

Major:

Please provide information in the introduction to make it clear that corisin is a peptide derived from a non-pathogenic larger bacterial protein, a transglycosylase. I think the introduction of this paper needs to clarify that the peptide has been found to be generated by the bacteria in measurable amounts in the lung to set the stage for the current work.

Response

We thank the reviewer for this important comment and agree that additional background was needed to better contextualize the biological origin of corisin and its relevance to the lung environment. In response, we revised the Introduction to clarify that corisin is a 19-amino-acid peptide generated by proteolytic processing of a larger IsaA-like transglycosylase precursor protein involved in bacterial cell-wall remodeling. We also added information indicating that the full-length precursor protein itself does not reproduce the biological activity of corisin, emphasizing that biological activity emerges following peptide release rather than being an intrinsic property of the intact bacterial protein.

Please refer to page 5, lines 108-118, and the transcribed text below.

"...Corisin is a 19-amino acid microbiota-derived peptide originally identified in the culture supernatant of Staphylococcus nepalensis isolated from fibrotic lung tissue.⁹ The corisin sequence is embedded within a larger IsaA-like transglycosylase precursor protein involved in bacterial cell-wall remodeling, and biological activity is acquired only after proteolytic release of the peptide from this precursor protein.⁹ Importantly, the intact precursor protein itself does not reproduce the biological activity of corisin, indicating that pathogenic activity resides in the released peptide rather than in the

full-length bacterial protein.⁹ Subsequent studies demonstrated that corisin and corisin-like peptides can be generated from transglycosylases of several commensal and pathogenic bacterial species, including *S. cohnii*, *S. hominis*, *Mycobacteroides abscessus*, and *Listeria monocytogenes*, supporting the concept that microbiota-derived peptide effectors may arise from diverse bacterial communities.¹¹ ...”

In addition, to address the reviewer’s point regarding physiological relevance, we expanded the Introduction to describe previous evidence demonstrating that corisin was originally identified in bacterial culture supernatants and that corisin and corisin-like peptides have been detected in measurable amounts in bronchoalveolar lavage fluid (BALF), with concentrations increased in patients with IPF and further elevated during acute exacerbation.

Please refer to pages 5 and 6, lines 118-129, and the transcribed text below.

“... Moreover, corisin and corisin-like peptides were detected in measurable amounts in bronchoalveolar lavage fluid (BALF), demonstrating that these bioactive peptides are generated within the pulmonary microenvironment in vivo.^{9, 11} Corisin concentrations were significantly increased in BALF from patients with IPF compared with healthy controls and further increased during acute exacerbation.^{9, 11} Experimental studies have shown that corisin promotes mitochondrial dysfunction, reactive oxygen species production, epithelial injury, and fibrotic progression in murine models of lung fibrosis.^{9, 11-13} Conversely, neutralization of corisin attenuates acute exacerbation and fibrosis severity in experimental models of lung fibrosis.¹¹ Beyond the lung, corisin has been implicated in infection-associated acute lung injury, chronic kidney disease, acute cholangitis, and adverse pregnancy outcomes, suggesting broader systemic profibrotic potential, enabled by its interaction serum albumin.¹³⁻¹⁷ ...”

Comment 3

Because most of the data are with synthetic corisin or genetically expressed corisin, the data in Supplemental Fig. 12C seems to be highly relevant since the corisin was immunodepleted using the anti-corsin Ab. Does this Ab only immunodeplete the corisin peptide, or does it immunodeplete the whole transglycosylase enzyme? Is it known? The reason is because the authors tested the depleted BALF, but the pull down of corisin was not tested and would have been a nice addition if it pulled down active peptide. It seems that supplemental Figure 12 C should be in the main text and the pull down of corisin added if performed.

Response

We thank the reviewer for this insightful comment. The anti-corsin monoclonal antibody used for immunodepletion was generated against the corisin peptide sequence and recognizes the PESSGNP epitope (<https://pmc.ncbi.nlm.nih.gov/articles/PMC8943153/>). In the present study, we did not directly evaluate whether immunodepletion removed only the released corisin peptide or also captured larger transglycosylase precursor proteins or precursor-derived fragments present in BALF. Therefore, we cannot exclude the possibility that additional corisin-containing molecular species were co-depleted. We agree that direct biochemical characterization of the immunoprecipitated fraction (for example by immunoprecipitation followed by Western blotting or mass spectrometry) would strengthen interpretation and have added this point as a limitation and future direction in the revised manuscript. We also transferred **Supplementary Figure 12C** into the main figures (**Figure 8**, in the revised manuscript) to better highlight the relevance of the native BALF depletion experiments.

Please see Figure 8 (formerly Supplementary Figure 12) and the Discussion/limitation section on pages 23 and 24, lines 617-620, and the transcribed text below.

“...Another limitation is that we did not determine whether anti-corisin antibody-mediated immunodepletion selectively removed the released bioactive corisin peptide or also depleted larger corisin-containing precursor proteins or precursor-derived fragments present in BALF....”

Comment 4

Minor:

Line 789 has a reference not formatted correctly

Response

We have corrected the format.

Please refer to page 34, line 897.

Comment 5

Do the authors have hydroxyproline data for the fibrosis measurements to supplement the Ashcroft scoring?

Response

We thank the reviewer for this helpful suggestion. In response, we have now included hydroxyproline quantification as an additional biochemical measure of pulmonary fibrosis in corisin transgenic mice (**Figure 9f**), complementing the histopathological Ashcroft scoring. Hydroxyproline content could not be measured in the bleomycin experiments because all lung tissues from those experiments had already been used for histopathological analysis. The new results have been incorporated into the revised Results section and the corresponding figure panel, and the Methods section has been updated accordingly.

Please see Figure 9f, the Results section on page 18, lines 469–470, and the Methods section on page 37, lines 995–996, and the transcribed texts below.

Results section on page 18, lines 469–470:

*“...The resulting transgenic lines were viable and fertile, producing normal litter sizes. Strikingly, however, these mice developed spontaneous lung fibrosis beginning at approximately 12 weeks of age, as evidenced by micro-CT, increased Ashcroft fibrosis scores, and extensive collagen deposition on Masson’s trichrome staining, and increased lung hydroxyproline content (**Fig. 9a–f**)....”*

Methods section on page 37, lines 996–997:

“...Lung hydroxyproline content was quantified using a colorimetric assay kit (BioVision Inc., San Francisco, CA, USA)....”

Response to Comments from Reviewer #2 :

Comment 1

Fujimoto et al. presents the manuscript “Corisin induces proteostasis stress to drive epithelial injury and pulmonary fibrosis”.

In this manuscript the authors present a comprehensive investigation into a microbiota-derived factor, corisin, and its role in epithelial injury and pulmonary fibrosis. They integrate molecular, cellular, and in vivo approaches to shed light on this peptide, and from their stepwise assessment they present corisin as a novel, potent peptide present in human bronchoalveolar lavage samples, exerting fibrosis driving activities in patients with idiopathic pulmonary fibrosis, but not in healthy controls.

The technical and methodological details are comprehensive and my review will focus on aspects of the study regarding the study participant data, the microbiota data, and the strength of certain interpretations. I here provide feedback regarding aspects that could be improved.

Response

We are very grateful to the reviewer for the insightful and constructive evaluation of our manuscript. We particularly appreciate the reviewer's careful assessment of the study participant information, microbiota-related analyses, and interpretation of the findings. These comments have helped us improve the clarity, balance, and rigor of the revised manuscript.

Comment 2

Human participant data:

The manuscript is framed in the context of human disease, progressive pulmonary fibrosis, and particularly idiopathic pulmonary fibrosis (IPF). Yet, only a limited part of the data is derived directly from participant bronchoalveolar lavage samples (BAL) (Figure 1, Supplementary Table 1 (Characteristics), and BAL experiments in Supplementary Fig. 12). The majority of the study relies on in vitro systems and mouse models. This is appropriate for mechanistic work, and the distinction between human-derived observations and experimental modeling is stated as a limitation. Unfortunately the information provided about the participants is not clear, and lacking important elements when studying the microbiota:

Supplementary Table 1 reports 14 IPF patients, and no healthy controls. Figure 1 includes only a subset (IPF n=8/14 control n=2/3), without explanation of this selection given in the Methods section or elsewhere in the manuscript.

When Figure 1 shows control h2 and h4, Supplementary Fig. 12: h1, h2, h3. and the source material providing no information about participants in Figure 1 it is very hard to understand the use of participant samples.

As the authors further state that corisin-encoding bacteria are present in all IPF patients, yet no comprehensive overview of all samples is provided, the statement should be adjusted to fit with what is shown.

Response

We apologize for the lack of clarity. PCR amplification was attempted using all available BALF DNA samples, including 14 IPF samples obtained during stable disease, 14 corresponding samples obtained during acute exacerbation, and 3 healthy control samples. However, sequencing was performed only for PCR products that showed a clearly resolved band of sufficient intensity at the expected amplicon size (~150 bp), because weak or poorly resolved bands were considered unlikely to provide reliable sequence information.

Based on this technical criterion, sequencing was performed for 2 stable IPF samples, 6 acute-exacerbation IPF samples, and 2 healthy control samples. We have now clarified this point in the Methods and Results sections and revised the text to indicate that these data demonstrate detection in technically analyzable samples, rather than prevalence across the entire cohort.

Please refer to the revised Results section on page 7, lines 162–201, the Methods section on page 41, lines 1094–1111, and the transcribed texts below. The limitations of this technical approach are also addressed in the Discussion/Limitations section; please refer to page 23, lines 601–608, and the transcribed text below.

Results section on page 7, lines 162–201:

“....PCR amplification targeting corisin-related sequences was attempted in all available BALF-derived DNA samples. However, sequencing was restricted to samples yielding amplicons of sufficient intensity and resolution at the expected size (~150 bp), as described in the Methods.¹⁷

Based on this technical quality criterion, downstream amplicon sequencing and bioinformatic analysis were performed using amplicons from 2 stable IPF samples, 6 acute exacerbation IPF samples, and 2 healthy control samples.

PCR products of the expected size were observed in the selected samples, whereas no amplification was detected in the no-template negative control (**Fig. 1a**). Nucleotide sequencing of DNA extracted from the excised gel (corresponding to the ~150 bp amplicons) identified sequences containing regions encoding canonical corisin and corisin-like peptides. Sequence alignment demonstrated highest similarity to transglycosylase-derived sequences annotated in GenBank from bacteria, including *Staphylococcus nepalensis* and *Staphylococcus cohnii*, both encoding peptides with 100% amino acid identity to canonical corisin, across all sequenced IPF and healthy control samples. In contrast, corisin-like peptide-encoding sequences showing highest similarity to annotated sequences from *Staphylococcus haemolyticus* and *Staphylococcus taiwanensis* were detected only among the sequenced IPF samples (**Fig. 1b,c**).

To further explore the sequence diversity and potential microbial origin of the detected amplicons, amplicon sequence variant (ASV)-based analysis followed by BLASTN alignment against reference databases was additionally performed. This exploratory analysis identified multiple ASVs showing the highest sequence similarity to members of the genus *Staphylococcus*, including sequences annotated as *S. nepalensis*, *S. haemolyticus*, *S. hominis*, *S. xylosus*, and related taxa (**Supplementary Dataset 1**). Collectively, these findings support the presence of genetically diverse corisin- and corisin-like peptide-encoding sequences within BALF samples from patients with IPF.

Because the amplified fragment (~150 bp) was specifically designed to recover the corisin-encoding region instead of use as a phylogenetic marker, these bacterial annotations represent sequence similarity to reference database entries and should not be interpreted as definitive species identification. Moreover, given that homologous transglycosylase genes carrying corisin/corisin-like sequences may occur across multiple bacterial taxa, inference of bacterial origin based solely on this short amplicon remains limited.

These findings provide the first sequence-level evidence that corisin- and corisin-like peptide-encoding bacterial sequences are detectable in BALF from patients with IPF. However, these data should be interpreted as evidence of sequence detectability rather than definitive evidence for enrichment of specific bacterial species associated with disease. Because sequencing was performed only on samples that met predefined amplicon-quality criteria, these findings should also be interpreted as evidence of detectability among technically analyzable samples rather than as a comprehensive prevalence analysis across the entire cohort....”

[Methods section page 41, lines 1094–1111:](#)

“....To ensure reliable sequence identification and reduce potential misinterpretation associated with low-template amplification, sequencing analysis was restricted to samples yielding amplicons of sufficient intensity and a clearly resolved band at the expected size (~150 bp). Samples showing weak, ambiguous, or poorly resolved amplification products were excluded from sequencing for technical quality reasons and were not interpreted as negative for the presence of corisin-related sequences.

PCR amplification targeting corisin-related sequences was performed using BALF-derived DNA obtained from all 14 patients with IPF during the stable phase, the corresponding BALF samples collected during acute exacerbation, and three healthy controls. Based on these predefined quality criteria, sufficient amplicons for sequencing were obtained from two stable IPF samples, six acute exacerbation IPF samples, and two healthy control samples. The selected amplicons were subsequently extracted from excised gels, sequenced and subjected to downstream bioinformatic analysis. Negative control reactions performed in the absence of template DNA showed no detectable amplification....”

[The Discussion/limitation section on page 23, lines 601–608:](#)

“...An additional limitation concerns the sequencing approach used to detect corisin- and corisin-like peptide-encoding DNA sequences. The targeted ~150 bp amplicon was designed to recover the corisin-encoding region and therefore does not provide the phylogenetic resolution of established taxonomic markers or metagenomic approaches. In addition, horizontal dissemination of corisin-containing transglycosylase genes across multiple bacterial taxa may further limit inference of organismal origin. Consequently, bacterial annotations presented in this study should be interpreted as sequence-similarity assignments rather than definitive species identification....”

Comment 3

Corisin-related sequences are detected in both IPF patients and healthy controls. This raises important questions regarding disease specificity, like whether differences are quantitative rather than qualitative, and how presence in healthy individuals aligns with the proposed pathogenic role. This conceptual point would benefit from clearer discussion.

Response

We thank the reviewer for this important conceptual comment and agree that detection of corisin-related sequences in both IPF patients and healthy controls requires additional discussion regarding disease specificity and mechanism. In response, we expanded the Discussion to clarify that our findings do not support a simple presence–absence model of corisin-producing bacteria. Instead, we now discuss the possibility that corisin acts as a context-dependent pathogenic effector whose biological consequences depend on quantitative and host-related factors rather than mere detectability. Specifically, we propose a “second-hit” framework in which epithelial vulnerability, altered microbiome composition, impaired proteostatic resilience, chronic inflammatory signaling, increased local corisin accumulation, and/or prolonged exposure contribute to pathogenic activation.

To support this interpretation, we incorporated a discussion of prior (and current) observations showing that corisin concentrations are increased in BALF from patients with IPF and are further elevated during acute exacerbation, suggesting that quantitative differences in exposure may influence disease outcomes. We also integrated our *in vivo* findings demonstrating that sustained corisin overexpression induces spontaneous pulmonary fibrosis and that corisin-overexpressing transgenic mice exhibit worsened fibrosis and reduced survival following bleomycin-mediated injury. Together, these observations support the concept that corisin may function as an amplifier of epithelial injury and fibrotic remodeling in susceptible lungs rather than as a disease-specific microbial product that is pathogenic solely by its presence.

We believe these additions provide a clearer conceptual framework that reconciles the detection of corisin-related sequences in healthy individuals with their proposed pathogenic role in IPF.

Please refer to the revised Discussion section on page 22, lines 571–596, and the transcribed text below.

“...The marked potency of native corisin observed in vitro prompted us to investigate its functional consequences in vivo. Consistent with its ability to induce apoptosis, senescence, EMT, and proteostasis disruption at disease-associated picogram concentrations, constitutive expression of native corisin in TG mice was sufficient to induce spontaneous pulmonary fibrosis in the absence of external injury, demonstrating that sustained corisin exposure alone can initiate and maintain chronic epithelial dysfunction and fibrotic remodeling. However, the pathological impact of corisin became even more pronounced under conditions of epithelial stress. Following bleomycin-mediated injury, corisin TG mice exhibited significantly reduced survival and developed more extensive fibrosis than wild-type littermates, indicating that corisin lowers the threshold for severe lung injury and amplifies fibrotic responses. Notably, saline-treated TG animals did not exhibit overt abnormalities, suggesting that corisin overexpression is not acutely toxic but instead promotes disease progression in a context-dependent manner. This concept is further supported by the observation that corisin-related

sequences were detected not only in IPF samples but also in healthy controls, indicating that the mere presence of corisin-producing bacteria may not be sufficient to drive disease. Rather, these findings support a model in which corisin acts as a context-dependent pathogenic effector requiring additional permissive conditions ("second hits"), including epithelial vulnerability, impaired proteostatic resilience associated with aging, altered microbiome composition, chronic inflammatory signaling, impaired barrier integrity, increased local accumulation of corisin, or prolonged exposure to the peptide. Consistent with this interpretation, corisin concentrations in BALF are increased in patients with IPF and further elevated during acute exacerbation,⁹ suggesting that quantitative differences in exposure, together with altered host responsiveness, determine pathological outcome. Collectively, these findings support a model in which corisin functions not simply as a disease-specific microbial product but as an amplifier of epithelial injury and fibrotic remodeling in susceptible lungs, thereby contributing to the progressive and episodic nature of IPF...."

Comment 4

No clinical information for healthy controls is reported, despite their inclusion in analyses. Supplementary Table 1 should contain information about all 17 participants, and also whose samples were used. Why only some samples are analysed need to be addressed.

I also miss information about: Antibiotic use before sampling, exacerbation frequency the last year, time from diagnosis to sampling, degree of fibrosis in IPF patients, BAL yield.

Clarifying these points would improve transparency and allow assessment of representativeness.

Response

We thank the reviewer for this important comment. In response to the reviewer's suggestions, we have substantially expanded **Supplementary Table 1** to include clinical information for all participants, including both patients with IPF and healthy controls. Specifically, we added information on antibiotic use before sampling, exacerbation frequency during the previous year, time from diagnosis to sampling, and the degree of fibrosis in patients with IPF. BALF yield has also been added to the Methods section.

Please refer to the revised Supplementary Table 1 and the updated description of the study participants and BALF collection procedures in the Methods section on page 25, lines 647–671, and the transcribed text below.

To further clarify sample usage, as noted above, we state in the Methods section that PCR amplification was initially attempted with all available BALF-derived DNA samples. Sequencing analysis was then restricted to samples that yielded amplicons of sufficient quality and intensity at the expected size for reliable sequence interpretation.

Please refer to the revised Methods section on page 41, lines 1094–1111, and the transcribed texts below.

Methods section on page 25, lines 647–671:

"...Subjects. This study enrolled 14 Japanese patients with idiopathic pulmonary fibrosis (IPF) (mean age, 69.0 ± 7.2 years; 13 males and 1 female) and three healthy Japanese male volunteers (mean age, 49.3 ± 16.2 years). The clinical characteristics of the participants are summarized in **Supplementary Table 1**. The diagnosis of IPF was established according to internationally accepted criteria.^{43, 44} The degree of pulmonary fibrosis was assessed using pulmonary function parameters, including forced vital capacity (FVC).^{45, 46} Bronchoscopic procedures were performed in accordance with the American Thoracic Society guidelines, and bronchoalveolar lavage fluid (BALF) samples were obtained from all 14 patients with IPF and from the three healthy controls.⁴³ BALF samples were obtained from the same patients during both stable and acute exacerbation phases. Stable-phase BALF was performed at the time of diagnosis, before the occurrence of acute exacerbation, whereas exacerbation-phase BALF was collected on the day of emergency hospitalization for acute exacerbation. All BALF procedures were performed in accordance with institutional safety criteria. In

patients with acute exacerbation, BALF was performed only when the PaO₂/FiO₂ (P/F) ratio was >150, either under continuous positive airway pressure (CPAP) support or during short-term mechanical ventilation after intubation. The procedure was discontinued if oxygen saturation decreased below 90%. BALF was performed in the right middle lobe, either from segment B4 or B5 depending on the bronchial segment most suitable for wedge positioning. A total of 150 mL sterile saline was instilled (50 mL x 3 aliquots). To minimize contamination, sterile saline, sterile tubes, and sterile gloves were used during sample collection and processing. Negative processing control samples were not collected because BALF specimens were obtained as part of routine clinical practice rather than specifically for research purposes. None of the patients had received antibiotic treatment before BALF sampling. Among the 14 patients included in the study, only one had experienced an acute exacerbation during the previous year (**Supplementary Table 1**).

Methods section on page 41, lines 1094–1111:

“...To ensure reliable sequence identification and reduce potential misinterpretation associated with low-template amplification, sequencing analysis was restricted to samples yielding amplicons of sufficient intensity and a clearly resolved band at the expected size (~150 bp). Samples showing weak, ambiguous, or poorly resolved amplification products were excluded from sequencing for technical quality reasons and were not interpreted as negative for the presence of corisin-related sequences.

PCR amplification targeting corisin-related sequences was performed using BALF-derived DNA obtained from all 14 patients with IPF during the stable phase, the corresponding BALF samples collected during acute exacerbation, and three healthy controls. Based on these predefined quality criteria, sufficient amplicons for sequencing were obtained from two stable IPF samples, six acute exacerbation IPF samples, and two healthy control samples. The selected amplicons were subsequently sequenced and subjected to downstream bioinformatic analysis. Negative control reactions performed in the absence of template DNA showed no detectable amplification....”

Comment 5

Bronchoscopy and BAL samples:

There is very little information given about the BAL sampling procedure. During which time span were participants sampled? Were stable state collected before or after exacerbations? What were the safety criteria for performing the procedure both during stable and exacerbated states? How much liquid was instilled and were samples collected from the same anatomical location for all participants? Were any measures taken to lower the risk of contamination during sampling? Why were there not a processing (negative) control sample taken for each participant sample?

Response

Following the reviewer's suggestions, we expanded the Methods section. Specifically, we clarified the timing of sample collection and now state that BALF samples obtained during stable disease and acute exacerbation were collected longitudinally from the same patients, with stable-phase BALF obtained at diagnosis, before the occurrence of an acute exacerbation, and exacerbation-phase BALF collected on the day of emergency hospitalization. We also added details on the safety criteria used for BALF, particularly during acute exacerbation, including P/F ratio thresholds, use of CPAP or short-term mechanical ventilation support when required, and predefined oxygen saturation criteria for procedure termination.

In addition, we included methodological details regarding the sampling procedure itself, including the anatomical location used for BALF collection (right middle lobe, B4 or B5), lavage volume (150 mL sterile saline administered as three 50 mL aliquots), and measures taken to minimize contamination during collection and processing (use of sterile saline, tubes, and gloves).

To address cohort-related concerns, we added information regarding antibiotic exposure before BALF sampling and exacerbation history during the preceding year. Finally, we clarified that negative processing controls were not collected because BALF specimens were obtained as part of routine clinical practice rather than under a protocol specifically designed for microbiome analysis. We believe these revisions provide a clearer description of the sampling procedures and improve the interpretation of the study findings.

Please refer to the revised Supplementary Table 1, the revised Method section on page 25, lines 647–672, and the transcribed text below.

“...Subjects. This study enrolled 14 Japanese patients with idiopathic pulmonary fibrosis (IPF) (mean age, 69.0 ± 7.2 years; 13 males and 1 female) and three healthy Japanese male volunteers (mean age, 49.3 ± 16.2 years). The clinical characteristics of the participants are summarized in **Supplementary Table 1**. The diagnosis of IPF was established according to internationally accepted criteria.^{43, 44} The degree of pulmonary fibrosis was assessed using pulmonary function parameters, including forced vital capacity (FVC).^{45, 46} Bronchoscopic procedures were performed in accordance with the American Thoracic Society guidelines, and bronchoalveolar lavage fluid (BALF) samples were obtained from all 14 patients with IPF and from the three healthy controls.⁴³ BALF samples were obtained from the same patients during both stable and acute exacerbation phases. Stable-phase BALF was performed at the time of diagnosis, before the occurrence of acute exacerbation, whereas exacerbation-phase BALF was collected on the day of emergency hospitalization for acute exacerbation. All BALF procedures were performed in accordance with institutional safety criteria. In patients with acute exacerbation, BALF was performed only when the PaO₂/FiO₂ (P/F) ratio was >150, either under continuous positive airway pressure (CPAP) support or during short-term mechanical ventilation after intubation. The procedure was discontinued if oxygen saturation decreased below 90%. BALF was performed in the right middle lobe, either from segment B4 or B5 depending on the bronchial segment most suitable for wedge positioning. A total of 150 mL sterile saline was instilled (50 mL x 3 aliquots). To minimize contamination, sterile saline, sterile tubes, and sterile gloves were used during sample collection and processing. Negative processing control samples were not collected because BALF specimens were obtained as part of routine clinical practice rather than specifically for research purposes. None of the patients had received antibiotic treatment before BALF sampling. Among the 14 patients included in the study, only one had experienced an acute exacerbation during the previous year (**Supplementary Table 1**). Unprocessed BALF aliquots were collected into sterile tubes and stored at -80°C until further analysis....”

Comment 6

To improve the presentation of participant data it would also be very interesting to see the effect of acute exacerbations per participant. If possible I would encourage the team to address potential differences within individuals comparing their exacerbated and stable BAL.

Response

We thank the reviewer for this important suggestion. In response, we performed additional longitudinal analyses comparing paired BALF samples obtained from the same patients during stable disease and acute exacerbation. Specifically, we quantified corisin and inflammatory/fibrotic mediators in paired BALF samples and evaluated within-individual changes between the stable and exacerbated states. These analyses showed increased BALF levels of corisin and associated inflammatory mediators during acute exacerbation, together with significant correlations between BALF corisin levels and selected cytokines. The corresponding data have now been incorporated into the revised manuscript and figure set.

Please note that BALF corisin concentrations in the present study were measured using a more sensitive assay than that employed in our previous report. Consequently, the absolute concentrations

reported here are higher than those previously reported in BALF samples. (<https://pmc.ncbi.nlm.nih.gov/articles/PMC7093394/>) ..

Please refer to the newly added **Supplementary Fig. 1a-b** and the corresponding description in the **Results** section on pages 8–9, lines 201–217, and the **Methods** section on pages 37–38, lines 989–1006, and the transcribed texts below.

Results section on pages 8–9, lines 202–218:

“....Longitudinal bronchoalveolar lavage fluid analysis links corisin to inflammatory activation during acute exacerbation of idiopathic pulmonary fibrosis. Having established the presence of corisin- and corisin-like peptide–encoding DNA sequences in bronchoalveolar lavage fluid (BALF), we next investigated whether BALF corisin levels were associated with inflammatory, epithelial-injury, and fibrinolytic-remodeling responses during disease exacerbation. To address this question, we quantified corisin together with a panel of inflammatory and epithelial injury–associated mediators, including interleukin-8 (IL-8), monocyte chemoattractant protein-1 (MCP-1), tumor necrosis factor- α (TNF- α), Fas ligand (FasL), high-mobility group box 1 (HMGB1), and plasminogen activator inhibitor-1 (PAI-1), in paired BALF samples collected from the same patients during stable disease and subsequent acute exacerbation. Corisin levels and the concentrations of all measured mediators were significantly increased during acute exacerbation compared with the stable disease state (**Supplementary Fig. 1a**). Notably, corisin levels showed significant positive correlations with these mediators in BALF obtained during acute exacerbation, whereas no such associations were observed in samples collected during stable disease (**Supplementary Fig. 1b**). Together, these findings indicate that increased BALF corisin levels are closely linked to the inflammatory and epithelial-injury milieu characteristic of acute exacerbation of idiopathic pulmonary fibrosis....”

Methods section on pages 37–38, lines 990–1007:

“...Biochemical analysis. The concentrations of interleukin-8 (IL-8) and tumor necrosis factor- α (TNF- α) were quantified using BD OptEIA enzyme-linked immunosorbent assay (ELISA) kits from BD Biosciences (San Diego, CA, USA). Fas ligand, plasminogen activator inhibitor-1 (PAI-1), and monocyte chemoattractant protein-1 (MCP-1) were measured using commercially available ELISA kits from R&D Systems, Inc. (Minneapolis, MN, USA). High-mobility group box 1 (HMGB1) was quantified using an ELISA kit from Shino-Test Corporation (Tokyo, Japan). Lung hydroxyproline content was quantified using a colorimetric assay kit (BioVision Inc., San Francisco, CA, USA). All assays were performed according to the manufacturer’s instructions. Corisin levels were measured using an in-house ELISA as previously described.¹¹ Briefly, 96-well plates were coated with polyclonal anti-transglycosylase 351 antibody at 2 μ g/mL in phosphate-buffered saline and incubated overnight at 4°C. After blocking with 1% bovine serum albumin in phosphate-buffered saline and thorough washing with phosphate-buffered saline containing Tween, the wells were incubated overnight at 4°C with serially diluted corisin standards and plasma samples. After washing, horseradish peroxidase-conjugated streptavidin from R&D Systems, Inc. (Minneapolis, MN, USA) was added. Following further washing and incubation, the substrate solution was applied to develop color, and absorbance was measured at 450 nm using an iMark™ Microplate Absorbance Reader from Bio-Rad Laboratories, Inc. (Hercules, CA, USA)....”

Comment 7

Identification of bacteria and methodological limitations:

Bacterial identification is based on targeted PCR amplification of a short (~150 bp) fragment of a corisin-related gene, followed by sequencing and bioinformatic annotation. The authors should give information about the needed number of PCR cycles to achieve signal as BAL fluid is very sparse on microbiota DNA, and a high number of PCR cycles increase the chimera risk.

Response

We thank the reviewer for this important comment. We agree that the number of PCR amplification cycles is an important consideration when analyzing low-biomass specimens such as BALF, as extensive amplification may increase the risk of PCR artifacts, including chimera formation. We have

therefore expanded the Methods section to explicitly describe the PCR conditions used. PCR amplification was performed with an initial denaturation at 93°C for 2 min, followed by 40 cycles of 93°C for 10 s, 50°C for 20 s, and 73°C for 40 s, with a final extension at 73°C for 1 min.

A total of 40 amplification cycles were used due to the low abundance of bacterial DNA typically present in BALF samples. To improve the reliability of the analysis and minimize the interpretation of potential amplification artifacts, only samples yielding a clearly resolved amplification product of the expected size (~150 bp) and sufficient intensity were advanced for sequencing. Weak or ambiguous amplification products were excluded from further analysis. In addition, no amplification was observed in template-free negative controls. These details have now been added to the revised Methods section.

Please refer to the revised Methods section on page 41, lines 1089–1111, and the transcribed text below.

“...PCR amplification was performed using an initial denaturation step at 93°C for 2 min, followed by 40 cycles of denaturation at 93°C for 10 s, annealing at 50°C for 20 s, and extension at 73°C for 40 s, with a final extension at 73°C for 1 min. Because bronchoalveolar lavage fluid (BALF) is a low-biomass specimen with limited microbial DNA yield, 40 amplification cycles were used to achieve sufficient sensitivity for downstream sequencing and taxonomic annotation. To ensure reliable sequence identification and reduce potential misinterpretation associated with low-template amplification, sequencing analysis was restricted to samples yielding amplicons of sufficient intensity and a clearly resolved band at the expected size (~150 bp). Samples showing weak, ambiguous, or poorly resolved amplification products were excluded from sequencing for technical quality reasons and were not interpreted as negative for the presence of corisin-related sequences.

PCR amplification targeting corisin-related sequences was performed using BALF-derived DNA obtained from all 14 patients with IPF during the stable phase, the corresponding BALF samples collected during acute exacerbation, and three healthy controls. Based on these predefined quality criteria, sufficient amplicons for sequencing were obtained from two stable IPF samples, six acute exacerbation IPF samples, and two healthy control samples. The selected amplicons were subsequently sequenced and subjected to downstream bioinformatic analysis. Negative control reactions performed in the absence of template DNA showed no detectable amplification....”

Comment 8

While amplicon sequencing is suitable for detecting specific gene families there are inherent limitations. The short sequence length and use of a non-standard taxonomic marker limit confidence in species-level identification. From reading the manuscript it seems still as if species level information was obtained from this amplicon sequencing. As a result, while the presence of corisin-like sequences is supported, the identification of specific bacterial species is more uncertain and should be interpreted cautiously.

Also page 37 line 991 says Shotgun sequencing. This is not described, and should be corrected to amplicon sequencing.

No overview of detected taxa across all samples is provided. Neither is the number of reads per sample for the taxa in Figure 1. With sparse data this should be conveyed.

Response

We thank the reviewer for this thoughtful comment and agree that our targeted amplicon sequencing approach has inherent limitations regarding taxonomic resolution.

First, we acknowledge that the ~150 bp PCR amplicon analyzed in this study is not a phylogenetic marker and was not designed for definitive bacterial species identification. Consequently, species-level assignments based on sequence similarity should be interpreted with caution. We have revised

the manuscript to state explicitly that the bacterial names mentioned in the text reflect sequence matches to annotated GenBank entries and do not constitute evidence that the detected sequence originated unequivocally from those specific species.

Please refer to the revised Results section on page 8, lines 188–201, and the transcribed text below.

Importantly, our objective was not microbiome profiling or taxonomic classification, but rather the detection of DNA fragments encoding corisin and corisin-like peptides. Corisin is a 19-amino-acid proapoptotic peptide embedded within the C-terminal region of a transglycosylase-like protein that is distributed across multiple bacterial taxa. Previous studies have identified homologous corisin-containing sequences in several *Staphylococcus* species, including *S. nepalensis*, *S. haemolyticus*, *S. hominis*, and *S. cohnii* (<https://pmc.ncbi.nlm.nih.gov/articles/PMC7093394/>). However, the corresponding transglycosylase gene has also been reported in non-staphylococcal organisms, including *Mycobacteroides abscessus*, *Listeria monocytogenes*, and *Burkholderia multivorans* (<https://pmc.ncbi.nlm.nih.gov/articles/PMC8943153/>), suggesting horizontal gene transfer and a broader taxonomic distribution.

To further explore the potential microbial origin of the detected sequences, we performed amplicon sequence variant (ASV)-based sequence analysis followed by BLASTN alignment against reference databases. This exploratory analysis identified the highest-similarity matches predominantly among members of the genus *Staphylococcus*, including sequences annotated as *S. nepalensis*, *S. haemolyticus*, *S. hominis*, and related taxa. However, because these assignments were derived from a short targeted amplicon designed specifically to recover the corisin-encoding region rather than taxonomic markers, we interpreted these results (in the revised manuscript) conservatively as sequence-similarity relationships rather than definitive species identification.

Please refer to the revised Results section on page 8, lines 180–193, and the newly added Supplementary Dataset 1, and the transcribed text below.

Given this multi-genus distribution, inferring the bacterial origin of the detected sequences solely from this short amplicon is inherently uncertain. Accordingly, we avoided making definitive taxonomic conclusions throughout the revised manuscript. Instead, our sequencing strategy was designed to recover the peptide-encoding region and infer the presence of corisin or corisin-like sequences, while treating organismal assignment as exploratory rather than confirmatory.

We have added a statement regarding this technical limitation to the Discussion/Limitations section. Please refer to the revised Discussion/Limitations section on page 23, lines 601–608, and the transcribed text below.

We also appreciate the reviewer's observation regarding terminology and have corrected "Shotgun DNA library preparation and sequencing" to "Amplicon sequencing" throughout the manuscript.

Finally, in response to the reviewer's suggestion, we have expanded the Results and Figure presentation to include sequencing depth information (number of reads per sample), ASV-based sequence similarity analysis, and an overview of "the potential" taxa detected across samples, thereby improving transparency regarding data sparsity and sequence interpretation. In addition, we have clarified in the Methods section that, for samples with insufficient individual amplicon yield, selected DNA amplicon samples were pooled prior to sequencing to increase amplicon intensity and enable reliable sequence identification. Specifically, extracted amplicons from healthy controls H2 and H4 were pooled into one sequencing sample, extracted amplicons from patients 1, 4, and 7 (all 3 derived

from patients under acute exacerbation-IPF) were pooled into a second sequencing sample, and extracted amplicons from patients 15 and 16 (derived from patients under stable disease-IPF) were pooled into a third sequencing sample.

Please refer to the revised method section on page 41, lines 1099–1103, and the legend of Supplementary Dataset 1.

Results section on page 8, lines 188–201:

“...Because the amplified fragment (~150 bp) was specifically designed to recover the corisin-encoding region rather than a phylogenetic marker, these bacterial annotations represent sequence similarity to reference database entries and should not be interpreted as definitive species identification. Moreover, given that homologous transglycosylase genes carrying corisin/corisin-like sequences may occur across multiple bacterial taxa, inference of bacterial origin based solely on this short amplicon remains limited.

These findings provide the first sequence-level evidence that corisin- and corisin-like peptide-encoding bacterial sequences are detectable in BALF from patients with IPF. However, these data should be interpreted as evidence of sequence detectability rather than definitive evidence for enrichment of specific bacterial species associated with disease. Because sequencing was performed only on samples that met predefined amplicon-quality criteria, these findings should also be interpreted as evidence of detectability among technically analyzable samples rather than as a comprehensive prevalence analysis across the entire cohort....”

Results section on page 8, lines 180–193:

*“...To further explore the sequence diversity and potential microbial origin of the detected amplicons, amplicon sequence variant (ASV)-based analysis followed by BLASTN alignment against reference databases was additionally performed. This exploratory analysis identified multiple ASVs showing the highest sequence similarity to members of the genus *Staphylococcus*, including sequences annotated as *S. nepalensis*, *S. haemolyticus*, *S. hominis*, *S. xylosus*, and related taxa (**Supplementary Dataset 1**). Collectively, these findings support the presence of genetically diverse corisin- and corisin-like peptide-encoding sequences within BALF samples from patients with IPF.*

Because the amplified fragment (~150 bp) was specifically designed to recover the corisin-encoding region rather than a phylogenetic marker, these bacterial annotations represent sequence similarity to reference database entries and should not be interpreted as definitive species identification. Moreover, given that homologous transglycosylase genes carrying corisin/corisin-like sequences may occur across multiple bacterial taxa, inference of bacterial origin based solely on this short amplicon remains limited....”

Discussion/Limitations section on page 23, lines 601–608:

“.....An additional limitation concerns the sequencing approach used to detect corisin- and corisin-like peptide-encoding DNA sequences. The targeted ~150 bp amplicon was designed to recover the corisin-encoding region and therefore does not provide the phylogenetic resolution of established taxonomic markers. In addition, the earlier reported horizontal dissemination of corisin-containing transglycosylase genes across multiple bacterial taxa may further limit inference of organismal origin.⁹ Consequently, bacterial annotations presented in this study should be interpreted as sequence-similarity assignments rather than definitive species identification....”

Results section on page 41, lines 1099–1103:

“...For samples with insufficient individual amplicon yield, DNA amplicons from healthy controls H2 and H4 were pooled into a first sequencing sample; amplicons from patients 1, 4, and 7 into a second; and amplicons from patients 15 and 16 into a third, to increase total amplicon intensity and ensure reliable sequence identification...”

For the microbiome assessment contamination should be considered as there seem to be only 1 negative control sample collected for the entire study. Is this the case decontamination tools like decontam in R should be used to make sure corisin producing bacteria is not results of laboratory contamination.

Response

We appreciate the reviewer's concerns with contaminants and have discussed this issue with our bioinformatics experts and came to the conclusion that applying decontam in this case may not be feasible. This is because the data were sequences obtained from gel-purified bands amplified from samples, and the negative control sample had no bands to purify and therefore were not pursued further. Please, note that our method here is quite different from the standard 16S amplicon sequencing approach, where from PCR to sequencing there are no intermediate purification steps like gel purification. Thus, the use of R 'decontam' may not be applicable in this particular case, since this is a tool designed for the purpose of removing or mitigating the abundance of detected taxa from blank/buffer negative controls (https://benjjneb.github.io/decontam/vignettes/decontam_intro.html).

To address the reviewer's concern, we have clarified this point in the revised Methods and added a limitation stating that, although the absence of a target-sized band in the negative control argues against reagent or laboratory contamination as the major source of the detected sequences, we cannot completely exclude low-level contamination. We have therefore interpreted the sequencing data cautiously and avoided making definitive taxonomic claims regarding the bacterial origin of corisin-producing organisms.

Please refer to the revised Discussion/limitation section on page 23, lines 608–611, and the transcribed text below.

“...Although the absence of a detectable target-sized amplicon in negative-control reactions argues against reagent or laboratory contamination as a major source of the detected sequences, the possibility of low-level contamination cannot be completely excluded.....”

We hope this explanation is satisfactory to the reviewer and clarifies why we did not use the decontam package in the first place.

Comment 10

3) Statistical considerations

Several experiments (e.g. Figure 2, Figure 3) are based on relatively small sample sizes ($n \approx 4-6$). The methods description indicate that parametric or non-parametric tests were used depending on data distribution, but it is unclear which tests were applied to specific datasets. Also normality evaluations for as few data points as 4 is not well powered. The authors have used graphpad for their analyses and this tool provides the following on this topic:

https://www.graphpad.com/guides/prism/latest/statistics/choosing_parametric_vs_nonpar.htm

Given the small sample sizes, clearer reporting and acknowledgment of this limitation would improve the statistical assessments, as would running non-parametric tests for all data to check for robustness of the results.

Response

We thank the reviewer for this insightful and constructive comment regarding the statistical methodology. We agree that formal normality testing has limited statistical power when applied to small sample sizes (e.g., $n = 4-6$), and that selecting statistical methods solely on the basis of normality testing under these conditions may not be sufficiently robust.

In response to this concern, we comprehensively re-evaluated the statistical analyses throughout the manuscript. Specifically, all comparisons presented in all figures and panels were reanalyzed using non-parametric methods irrespective of the outcome of normality testing. Comparisons between two unpaired groups were performed using the two-sided Mann–Whitney U test; between two paired groups, using the two-sided Wilcoxon rank test; and among three or more groups, using the Kruskal–Wallis test followed by Dunn’s post hoc analysis.

Normality assessment (Shapiro–Wilk or Kolmogorov–Smirnov test) was retained; however, its role was revised and is now used exclusively to guide the descriptive presentation of the data (mean $\pm$ SD, mean $\pm$ SEM, or median with interquartile range [IQR]) rather than to determine the inferential statistical test applied. This modification was implemented to avoid overinterpreting distribution tests in small datasets while maintaining appropriate descriptive reporting.

Importantly, following reanalysis using the uniform non-parametric approach, the study's overall statistical outcomes and biological conclusions remained unchanged, supporting the robustness and reproducibility of the findings. To improve transparency, the Statistical Analysis section has been revised to explicitly describe the statistical test used for each comparison and clarify the rationale for adopting non-parametric methods throughout the study.

Please refer to the revised Method section on pages 43-44, lines 1154–1162, the figure legends of all figures, and the transcribed text below.

*“...**Statistical Analysis.** Data distribution was assessed using the Shapiro–Wilk or Kolmogorov–Smirnov test to guide the selection of the most appropriate descriptive statistics for each dataset. Data are presented as mean $\pm$ standard deviation (SD), mean $\pm$ standard error of the mean (SEM), or median with interquartile range (IQR). Given the relatively small sample sizes in several experiments, group comparisons were performed using non-parametric statistical tests. Comparisons between two independent groups were performed using the two-sided Mann–Whitney U test, whereas comparisons between two paired groups were performed using the two-sided Wilcoxon signed-rank test. Comparisons among three or more groups were performed using the Kruskal–Wallis test followed by Dunn’s post hoc test.....”*

Comment 11

Figure 5c: On page 11 the authors state that UMAP visualizations show clear separation between corisin-treated/non-treated cells, although samples are plotted separately and do not allow direct comparison.

Response

Based on the reviewer's suggestion, we have revised Figure 5c to include an integrated UMAP visualization that displays both conditions within the same embedding space, enabling direct comparison of cellular distributions and transcriptional states between groups.

Please refer to the revised Figure 5c.

Comment 12

Context of corisin research

It would be helpful to more clearly acknowledge that this research group appears to be the primary source of work on corisin. While this does not diminish the findings, it is important context for readers in evaluating novelty and independent validation.

Response

We thank the reviewer for this thoughtful comment and agree that it is important to provide readers with appropriate context regarding the current stage of corisin research. Corisin is a recently identified microbiota-derived peptide, and to date, much of the published work characterizing its biological functions and disease associations has originated from our research group. We acknowledge that broader independent replication will be important to further establish the generalizability and translational significance of these findings.

In response to this comment, we have revised the Discussion/Limitations section to explicitly acknowledge this point and to emphasize the need for validation across independent cohorts, experimental systems, and research groups. At the same time, we note that the present study incorporated collaborative analyses across multiple institutions and included analytical procedures designed to reduce interpretive bias, such as blinded or independent evaluation of selected datasets, including single-cell transcriptomic analyses, screening of intracellular corisin-binding proteins, and assessment of corisin penetration into cells. We believe these additions provide readers with clearer context regarding both the novelty and current developmental stage of the corisin field.

Please refer to the revised Discussion/limitation section on page 24, lines 622–629, and the transcribed text below.

“.....Finally, corisin is a recently identified microbiota-derived peptide, and much of the currently available literature characterizing its biological roles has originated from our research group. Therefore, independent replication across additional research groups, diverse experimental systems, and larger clinical cohorts will be important to establish the reproducibility, biological relevance, and translational significance of these observations. Although the present study incorporated collaborative analyses across multiple institutions and included blinded evaluation of selected datasets, further external validation remains warranted”

Comment 13

Source document

Could the authors double check sheet names and content in the source material. Some columns are not labeled and some Figures seem not to be supported by content on corresponding sheet.

Response

We thank the reviewer for carefully checking the source data files and for pointing out these inconsistencies. We have now thoroughly reviewed and revised the source document to ensure that all sheet names correctly correspond to the associated figures and panels. In addition, unlabeled columns have been appropriately annotated, and missing or mismatched datasets have been corrected to ensure that each figure is fully supported by the corresponding source data sheet. We apologize for the oversight and appreciate the reviewer’s attention to detail, which helped improve the clarity and transparency of the data presentation.

Point-by-point Response to Comments from the Reviewers

Response to Comments from Reviewers #1:

Comment 1

The authors have adequately addressed my comments and the manuscript is significantly improved.

Response

We sincerely thank the reviewer for their positive assessment of our revised manuscript and for their valuable comments, which have significantly improved the quality of our work.

Response to Comments from Reviewer #2:

Comment 1

I thank the authors for their careful and thorough revision. They have addressed the points raised in my review in a satisfactory manner.

Response

We sincerely thank the reviewer for their positive assessment of our revised manuscript. We greatly appreciate the reviewer's time and careful evaluation.

Comment 2

In particular, I appreciate the clarification and expansion of the participant information in Supplementary Table 1. Regarding antibiotic use, if not possible to retrieve information at latest time for AB use, the authors can leave as is, but it would be interesting to know if AB use occurred within the last 90 days before sampling both stable and exacerbated state. Time from diagnosis to sampling is now perfectly clear, if possible time to first exacerbation should be added. Degree of fibrosis, and the details of the BAL sampling procedure are both well described.

Response

We sincerely thank the reviewer for this valuable comment. Based on the available medical records, none of the patients received antibiotics at our institution within 90 days before BALF sampling at either the stable or acute exacerbation state. However, because information on treatment received at other medical institutions was not always available, antibiotic use elsewhere during this period could not be entirely ruled out.

Regarding the time to first exacerbation, the mean interval from the initial BALF sampling during the stable state to the first acute exacerbation was $1,037 \pm 812.5$ days, with a median of 679.5 days. We have added this information to the Methods section of the revised manuscript. The mean interval from the initial BALF sampling during the stable state to the first acute exacerbation has also been added to Supplementary Table 1, and the corresponding interval for each individual patient has been provided in the Source Data file.

Please refer to page 15, lines 666-671, in the methods section, and to the updated Supplementary Table 1, and the transcribed text described below:

"None of the patients received antibiotics at our institution within 90 days before BALF sampling at either the stable or acute exacerbation state. However, antibiotic use at other medical institutions

during this period could not be completely excluded. The mean time from the initial BALF sampling at the stable state to the first acute exacerbation was $1,037 \pm 812.5$ days, with a median of 679.5 days.”

Comment 3

The revised description of sample usage explaining that PCR was attempted on all available samples but sequencing was restricted to those meeting predefined amplicon-quality criteria resolves my earlier confusion about the subset of samples. Thank you for clarifying.

I am also pleased that the authors now report the PCR cycling conditions and acknowledge the chimera/artifact risk associated with low-biomass specimens, and that the taxonomic claims have been appropriately tempered to reflect sequence-similarity assignments rather than definitive species identification. The correction of “shotgun” to “amplicon” sequencing, the added sequencing-depth information, and the explicit description of the amplicon pooling strategy all improve transparency.

The added longitudinal paired stable-versus-exacerbation analyses (Supplementary Fig. 1) are a valuable addition, and the expanded discussion of disease specificity framing corisin as a context-dependent effector requiring “second hits” rather than a simple presence/absence determinant addresses my conceptual concern well. The uniform adoption of non-parametric tests and the acknowledgment that much of the corisin literature originates from this group are also appropriate and improve the balance of the manuscript.

I have no further concerns within the scope of my review.

Response

We sincerely thank the reviewer again for the positive assessment of our revisions and for the thoughtful and constructive comments, which have helped us improve the clarity, transparency, and overall quality of our manuscript.