nature portfolio

Peer Review File

Complex multicellularity is linked with expanded specialized metabolite production in microbes

Corresponding Author: Professor Cameron Currie

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
The paper by Salamzade et al. describes a very interesting study linking multicellular lifestyles of bacteria and fungi to expanded chemical arsenals in microbes. This is a carefully conducted study with a clear storyline, rigorous analyses and clear figures. While it is not the very first paper to link biosynthetic potential with multicellularity, it is the first one to do it this explicitly, comprehensively and quantitatively. The code for all analyses is also well documented. Overall, it will make a nice contribution to scientific literature. Still, I have a number of remarks that I hope will be useful to the authors:
- Line 93-94: The link between multicellularity and secondary metabolic complexity has been made before. It would be good to acknowledge/cite reviews that have discussed this (e.g., doi.org/10.1111/brv.12605, doi:10.1128/jb.00153-23, doi:10.1039/C5NP00013K). Also, at least one large-scale study of microbial biosynthetic diversity has identified a connection between multicellularity and biosynthetic potential (doi:10.1016/j.cell.2014.06.034, Table S2).
- Line 92: 'facilitated' and Line 186: 'due to'. In these sentences, directional causality is implied, while there does not seem to be data that clearly supports this. The evidence is associational.
- In the analysis of biosynthetic repertoires, BGC richness among taxa is analysed extensively. Adding also diversity metrics (how diverse are BGC repertoires of related species/strains) would be very useful to provide a more complete picture. The suggestion that multicellular organisms might be prime sources for novel metabolite discovery (line 472, closing sentence of the paper) could also be substantiated by such an analysis.
- Line 113-115: interesting hypotheses are provided for the links between secondary metabolite production and multicellular lifestyles. However, this does not necessarily explain the expansion of numbers of BGCs per genome; even with a single BGC, a bacterium might invest a lot of energy making a lot of that metabolite. The fact that a bacterium has more BGCs in its genome does not necessarily mean that it invests more energy in secondary metabolism throughout its life cycle.
- Line 122: the summed lengths of predicted BGC regions were used. Was any effort made to trim/correct greedily identified borders? Otherwise, the total size could easily be heavily overestimated.
- Line 273: if vertical evolution is inferred, can this be substantiated by checking that gene phylogeny follows species phylogeny? And/or through a parsimony analysis to show that this is the most parsimonious model for the data?

Minor:
- Figure 1c: 'Number of distinct metabolites' . I guess this refer to 'known metabolites'. Please specify.
- Line 321: 'two independent origins' -> Figure 3A seems to suggest three?

(Remarks on code availability)
The GitHub repository is well structured and the code is well documented too. There is also a README file that clearly outlines what can be found where.

Reviewer #2

(Remarks to the Author)
This paper provides a broad phylogenetic analysis of the coincidence of specialized metabolite production and multicellularity in bacteria and fungi. The phylogenetic analysis appears to be quite rigorous, and the statistical associations are not terribly surprising, but nice to see evaluated in a formal fashion and to have all of these analyses in one place. My comments are rather limited.

1) The authors need to provide an objective definition of a BGC at the outset. Bugs synthesize all kinds of metabolites, from amino acids to nucleotides to cofactors, up to the more advanced compounds that the authors seem to be focusing on, but where is the line drawn?

2) This is not entirely on the authors, but more a comment on the extremes to which journals now seem to be encouraging things. The font sizes and tiny graphs included here are approaching a rather ridiculous level. Readers should not have to expand a pdf

to the entire width of a screen in hopes of seeing things; they ought to be visible on a 8 x 11 hardcopy.

3) The authors favor the hypothesis that multicellularity paved the way for the evolution of biosynthetic pathways, and there certainly appears to be an association. I'm not sure how it can be teased out, but how do we know that the order isn't reversed? Some discussion of this is needed.

4) The issue of division of labor is raised in the paper, and this does indeed seem to be an issue in the evolution of complex multicellularity. But no evidence in provided on this, so this should also be elaborated on a bit.

5) If the authors' hypothesis is correct and general, then other lineages of multicellulars would be expected to produce complex external metabolites. It would be helpful to summarize some data on this, which likely exists in some of the photosynthetic eukaryote lineages.

(Remarks on code availability)

Reviewer #3

(Remarks to the Author)
Summary: The manuscript "Complex multicellularity linked with expanded chemical arsenals in microbes" analyzes existing genomic data to address the question of how the evolution of multicellularity is related to an organism or lineages genomic diversity biosynthetic gene clusters. Overall, I found the manuscript to be well-written and engaging to read, and the complexity and scale of the analyses across different microbial groups are impressive. The results certainly provide strong evidence for the association between multicellularity and BGC diversity in both Bacteria and Fungi and I mainly have suggestions for clarification of some of the methods or statements.

Questions/Suggestions:

My major concern about the methods is how the taxon selection might skew the results. For example, it appears that the fungal genomes that were used are quite skewed towards certain groups (e.g., chytrids), while other lineages (e.g., "zygos" such as Basidiobolus) are underrepresented. I understand that not all genomes can be analyzed, but it seems that if you had to choose 312 genomes you would want to make sure that all clades are well represented, to the extent possible with existing data. For example, Supplemental Table S7 lists the fungal genomes that were included in the analyses. In looking at these, it seems that Chytrid genomes are over-represented compared to other non-Dikarya lineages (i.e., 23% of genomes), such as Mucoromycota (5% of genomes).

Along those same lines, why did the authors restrict genomes to those that did not have "fam_Incertae_sedis" or "unidentified" in their taxonomic designations in the UFCG database when the types of analyses that they performed did not rely heavily on these lower taxonomic levels? A huge diversity of Fungi are not yet fully identified or named. By excluding these genomes my concern is that the genomes are not fully representative of the genomic diversity of fungi.

In the methods, the authors state "Representative genomes were selected based on BGC-ome size." More information would be useful. i.e., were rep genomes chosen to represent a diversity of BGC-ome sizes, or the maximum sizes, etc?

Is there a table that lists the Pfams with NACHT or HET domains that were used in the analyses? In the supplementary methods, the authors state "Domain architectures for HI-associated proteins were assessed based on comprehensive Pfam annotations using
custom scripts provided in the associated code repository of this study." It would be useful to have more information in the Supplementary methods as to how this was done rather than having to go hunt down the github code.

You may want to define dikaryon/Dikarya for non-fungal biologists?

It is really interesting that the authors identified HI domains as being associated with BGC diversity. However, I do wonder whether all of the putative HI genes are functioning in that capacity? For example, Neurospora crassa has 8-10 het loci, but a search of the genome for het annotations results in many more hits. Did the authors check to see whether the proteins that were identified as having HI domains were validated in species where this is known? You mention that these proteins have homology to antiphage proteins in bacteria, but are there other known functions for proteins with these domains in fungi?

For the sentence "Pezizomycotina and Agaricomycetes likely originated around the same time as the emergence of land plants" the authors provide a reference to a review paper by Berbee (ref 46). Rather than only citing a review paper I would also directly cite the research underlying this statement, such as Lutzoni et al., 2018 https://www.nature.com/articles/s41467-018-07849-9

Lines 417-420. The statement "Indeed, our finding that BGC expansions in both fungi and bacteria coincides with significant enrichment in carbohydrate utilization enzymes suggest that catabolic processes help these microbes obtain the energy required to produce specialized metabolites at ecologically relevant scales." is consistent with earlier work by Franco et al., (2022) New Phytologist that demonstrated that expansion of BGCs in a clade of xylarialean fungi was associated with increased numbers of CAZymes (as well as other saptrotrophic/pathogenicity-related genes) in non-symbiont genomes, yet not in fungi isolated as symbiotic endophytes. This suggests that while this trend may be generally true across fungi, there may be exceptions when

more closely related taxa are analyzed?

The authors discuss how Pezizmycotina has expanded BGC-omes compared to Agaricomycotina, but there is a very large amount of variation in BGC diversity among the Pezizmycotina that the authors do not address. How does that variation fit into the model they propose? I think it they authors are suggesting that fungi with large BGC-omes may be undergoing sexual reproduction less frequently? However, I don't think clonal vs. sexual reproduction frequency fully explains the variation in BGC-omes across the Pezizomycotina.

Is there a relationship between multicellularity, BGC-ome size, and HGT frequency? Previous studies have demonstrated (at least in Fungi) that BGCs are frequently obtained via HGT, but the authors do not address this.

The authors state: "Large-scale annotation efforts have also previously suggested that BGC carriage is limited in many bacterial species18 and databases cataloging the taxonomic origins of known natural products have shown that they are disproportionately sourced from just six microbial phyla19,20 (Fig. 1b)." However, I wonder to what extent the database is biased towards particular more well-studied taxa?

(Remarks on code availability)

Decision Letter:

26th September 2025

*Please ensure you delete the link to your author homepage in this e-mail if you wish to forward it to your co-authors.

Dear Cameron,

Thank you for your patience while your manuscript "Complex multicellularity linked with expanded chemical arsenals in microbes" was under peer-review at Nature Microbiology. It has now been seen by 3 referees, whose expertise and comments you will find at the end of this email. Although they find your work of some potential interest, they have raised a number of concerns that will need to be addressed before we can consider publication of the work in Nature Microbiology.

In particular, referee #1 asks to better integrate previous literature, and to add diversity metrics to the analysis (how diverse are BGC repertoires of related species/strains). The referee also has some concerns regarding overestimation of the size of predicted BGC regions, and asks to further validate vertical evolution. Referee #2 says that an objective definition of a BGC needs to be provided at the outset, and that division of labor needs to be further elaborated on. Referee #3 has important concerns about how the taxon selection might skew the results. The referee also has concerns about the exclusion of certain genomes. Further, referee #3 asks if there is a relationship between multicellularity, BGC-ome size, and HGT frequency. Editorially, we will require these and the other referee concerns to be addressed in full.

Should further experimental data allow you to address these criticisms, we would be happy to look at a revised manuscript.

We are committed to providing a fair and constructive peer-review process. Please do not hesitate to contact us if there are specific requests from the reviewers that you believe are technically impossible or unlikely to yield a meaningful outcome.

We strongly support public availability of data. Please place the data used in your paper into a public data repository, if one exists, or alternatively, present the data as Source Data or Supplementary Information. If data can only be shared on request, please explain why in your Data Availability Statement, and also in the correspondence with your editor. For some data types, deposition in a public repository is mandatory - more information on our data deposition policies and available repositories can be found at https://www.nature.com/nature-research/editorial-policies/reporting-standards#availability-of-data.

Please include a data availability statement as a separate section after Methods but before references, under the heading "Data Availability". This section should inform readers about the availability of the data used to support the conclusions of your study. This information includes accession codes to public repositories (data banks for protein, DNA or RNA sequences, microarray, proteomics data etc…), references to source data published alongside the paper, unique identifiers such as URLs to data repository entries, or data set DOIs, and any other statement about data availability. At a minimum, you should include the following statement: "The data that support the findings of this study are available from the corresponding author upon request", mentioning any restrictions on availability. If DOIs are provided, we also strongly encourage including these in the Reference list (authors, title, publisher (repository name), identifier, year). For more guidance on how to write this section please see: http://www.nature.com/authors/policies/data/data-availability-statements-data-citations.pdf

When revising your manuscript:

* Include a "Response to referees" document detailing, point-by-point, how you addressed each referee comment. If no action was taken to address a point, you must provide a compelling argument. This response will be sent back to the referees along with the revised manuscript.

* If you have not done so already we suggest that you begin to revise your manuscript so that it conforms to our Article format

instructions at http://www.nature.com/nmicrobiol/info/final-submission. Refer also to any guidelines provided in this letter.

* Include a revised version of any required reporting checklist. It will be available to referees (and, potentially, statisticians) to aid in their evaluation if the manuscript goes back for peer review. A revised checklist is essential for re-review of the paper.

When submitting the revised version of your manuscript, please pay close attention to our href="https://www.nature.com/nature-portfolio/editorial-policies/image-integrity">Digital Image Integrity Guidelines. and to the following points below:

-- that unprocessed scans are clearly labelled and match the gels and western blots presented in figures.
-- that control panels for gels and western blots are appropriately described as loading on sample processing controls
-- all images in the paper are checked for duplication of panels and for splicing of gel lanes.

EXTENDED DATA FIGURES

When re-submitting your manuscript, please ensure that any supplementary figures and tables that are crucial to the manuscript's conclusions are converted into Extended Data figures and tables to increase visibility of these data. Extended Data figures and tables are online-only (present in the online PDF and full-text HTML versions of the paper), peer-reviewed display items that provide essential background to the article but are not included in the main article due to space constraints. A maximum of ten Extended Data display items (figures and tables) is permitted.

Finally, please ensure that you retain unprocessed data and metadata files after publication, ideally archiving data in perpetuity, as these may be requested during the peer review and production process or after publication if any issues arise.

Please use the link below to submit a revised paper:

Link Redacted

Note: This url links to your confidential homepage and associated information about manuscripts you may have submitted or be reviewing for us. If you wish to forward this e-mail to co-authors, please delete this link to your homepage first.

Nature Microbiology is committed to improving transparency in authorship. As part of our efforts in this direction, we are now requesting that all authors identified as 'corresponding author' on published papers create and link their Open Researcher and Contributor Identifier (ORCID) with their account on the Manuscript Tracking System (MTS), prior to acceptance. This applies to primary research papers only. ORCID helps the scientific community achieve unambiguous attribution of all scholarly contributions. You can create and link your ORCID from the home page of the MTS by clicking on 'Modify my Springer Nature account'. For more information please visit please visit www.springernature.com/orcid.

If you wish to submit a suitably revised manuscript we would hope to receive it within 6 months. If you cannot send it within this time, please let us know.

In the meantime we hope that you find our referees' comments helpful.

Yours sincerely,

****
Reviewer Expertise:

Referee #1: Biosynthetic gene clusters, natural products, Bacterial and fungal BGCs, computational biology
Referee #2: Population genetics
Referee #3: Fungal evolution, phylogenetics

Reviewer Comments:

Reviewer #1 (Remarks to the Author):

The paper by Salamzade et al. describes a very interesting study linking multicellular lifestyles of bacteria and fungi to expanded chemical arsenals in microbes. This is a carefully conducted study with a clear storyline, rigorous analyses and clear figures. While it is not the very first paper to link biosynthetic potential with multicellularity, it is the first one to do it this explicitly, comprehensively and quantitatively. The code for all analyses is also well documented. Overall, it will make a nice contribution to scientific literature. Still, I have a number of remarks that I hope will be useful to the authors:
- Line 93-94: The link between multicellularity and secondary metabolic complexity has been made before. It would be good to acknowledge/cite reviews that have discussed this (e.g., doi.org/10.1111/brv.12605, doi:10.1128/jb.00153-23, doi:10.1039/C5NP00013K). Also, at least one large-scale study of microbial biosynthetic diversity has identified a connection between multicellularity and biosynthetic potential (doi:10.1016/j.cell.2014.06.034, Table S2).
- Line 92: 'facilitated' and Line 186: 'due to'. In these sentences, directional causality is implied, while there does not seem to be data that clearly supports this. The evidence is associational.
- In the analysis of biosynthetic repertoires, BGC richness among taxa is analysed extensively. Adding also diversity metrics

(how diverse are BGC repertoires of related species/strains) would be very useful to provide a more complete picture. The suggestion that multicellular organisms might be prime sources for novel metabolite discovery (line 472, closing sentence of the paper) could also be substantiated by such an analysis.
- Line 113-115: interesting hypotheses are provided for the links between secondary metabolite production and multicellular lifestyles. However, this does not necessarily explain the expansion of numbers of BGCs per genome; even with a single BGC, a bacterium might invest a lot of energy making a lot of that metabolite. The fact that a bacterium has more BGCs in its genome does not necessarily mean that it invests more energy in secondary metabolism throughout its life cycle.
- Line 122: the summed lengths of predicted BGC regions were used. Was any effort made to trim/correct greedily identified borders? Otherwise, the total size could easily be heavily overestimated.
- Line 273: if vertical evolution is inferred, can this be substantiated by checking that gene phylogeny follows species phylogeny? And/or through a parsimony analysis to show that this is the most parsimonious model for the data?

Minor:
- Figure 1c: 'Number of distinct metabolites' . I guess this refer to 'known metabolites'. Please specify.
- Line 321: 'two independent origins' -> Figure 3A seems to suggest three?

Reviewer #1 (Remarks on code availability):

The GitHub repository is well structured and the code is well documented too. There is also a README file that clearly outlines what can be found where.

Reviewer #2 (Remarks to the Author):

This paper provides a broad phylogenetic analysis of the coincidence of specialized metabolite production and multicellularity in bacteria and fungi. The phylogenetic analysis appears to be quite rigorous, and the statistical associations are not terribly surprising, but nice to see evaluated in a formal fashion and to have all of these analyses in one place. My comments are rather limited.

1) The authors need to provide an objective definition of a BGC at the outset. Bugs synthesize all kinds of metabolites, from amino acids to nucleotides to cofactors, up to the more advanced compounds that the authors seem to be focusing on, but where is the line drawn?

2) This is not entirely on the authors, but more a comment on the extremes to which journals now seem to be encouraging things. The font sizes and tiny graphs included here are approaching a rather ridiculous level. Readers should not have to expand a pdf to the entire width of a screen in hopes of seeing things; they ought to be visible on a 8 x 11 hardcopy.

3) The authors favor the hypothesis that multicellularity paved the way for the evolution of biosynthetic pathways, and there certainly appears to be an association. I'm not sure how it can be teased out, but how do we know that the order isn't reversed? Some discussion of this is needed.

4) The issue of division of labor is raised in the paper, and this does indeed seem to be an issue in the evolution of complex multicellularity. But no evidence in provided on this, so this should also be elaborated on a bit.

5) If the authors' hypothesis is correct and general, then other lineages of multicellulars would be expected to produce complex external metabolites. It would be helpful to summarize some data on this, which likely exists in some of the photosynthetic eukaryote lineages.

Reviewer #3 (Remarks to the Author):

Summary: The manuscript "Complex multicellularity linked with expanded chemical arsenals in microbes" analyzes existing genomic data to address the question of how the evolution of multicellularity is related to an organism or lineages genomic diversity biosynthetic gene clusters. Overall, I found the manuscript to be well-written and engaging to read, and the complexity and scale of the analyses across different microbial groups are impressive. The results certainly provide strong evidence for the association between multicellularity and BGC diversity in both Bacteria and Fungi and I mainly have suggestions for clarification of some of the methods or statements.

Questions/Suggestions:

My major concern about the methods is how the taxon selection might skew the results. For example, it appears that the fungal genomes that were used are quite skewed towards certain groups (e.g., chytrids), while other lineages (e.g., "zygos" such as Basidiobolus) are underrepresented. I understand that not all genomes can be analyzed, but it seems that if you had to choose 312 genomes you would want to make sure that all clades are well represented, to the extent possible with existing data. For example, Supplemental Table S7 lists the fungal genomes that were included in the analyses. In looking at these, it seems that Chytrid genomes are over-represented compared to other non-Dikarya lineages (i.e., 23% of genomes), such as Mucoromycota (5% of genomes).

Along those same lines, why did the authors restrict genomes to those that did not have “fam_Incertae_sedis” or “unidentified” in their taxonomic designations in the UFCG database when the types of analyses that they performed did not rely heavily on these lower taxonomic levels? A huge diversity of Fungi are not yet fully identified or named. By excluding these genomes my concern is that the genomes are not fully representative of the genomic diversity of fungi.

In the methods, the authors state “Representative genomes were selected based on BGC-ome size.” More information would be useful. i.e., were rep genomes chosen to represent a diversity of BGC-ome sizes, or the maximum sizes, etc?

Is there a table that lists the Pfams with NACHT or HET domains that were used in the analyses? In the supplementary methods, the authors state “Domain architectures for HI-associated proteins were assessed based on comprehensive Pfam annotations using
custom scripts provided in the associated code repository of this study.” It would be useful to have more information in the Supplementary methods as to how this was done rather than having to go hunt down the github code.

You may want to define dikaryon/Dikarya for non-fungal biologists?

It is really interesting that the authors identified HI domains as being associated with BGC diversity. However, I do wonder whether all of the putative HI genes are functioning in that capacity? For example, Neurospora crassa has 8-10 het loci, but a search of the genome for het annotations results in many more hits. Did the authors check to see whether the proteins that were identified as having HI domains were validated in species where this is known? You mention that these proteins have homology to antiphage proteins in bacteria, but are there other known functions for proteins with these domains in fungi?

For the sentence “Pezizomycotina and Agaricomycetes likely originated around the same time as the emergence of land plants” the authors provide a reference to a review paper by Berbee (ref 46). Rather than only citing a review paper I would also directly cite the research underlying this statement, such as Lutzoni et al., 2018 https://www.nature.com/articles/s41467-018-07849-9

Lines 417-420. The statement “Indeed, our finding that BGC expansions in both fungi and bacteria coincides with significant enrichment in carbohydrate utilization enzymes suggest that catabolic processes help these microbes obtain the energy required to produce specialized metabolites at ecologically relevant scales.” is consistent with earlier work by Franco et al., (2022) New Phytologist that demonstrated that expansion of BGCs in a clade of xylarialean fungi was associated with increased numbers of CAZymes (as well as other saptrotrophic/pathogenicity-related genes) in non-symbiont genomes, yet not in fungi isolated as symbiotic endophytes. This suggests that while this trend may be generally true across fungi, there may be exceptions when more closely related taxa are analyzed?

The authors discuss how Pezizmycotina has expanded BGC-omes compared to Agaricomycotina, but there is a very large amount of variation in BGC diversity among the Pezizmycotina that the authors do not address. How does that variation fit into the model they propose? I think it they authors are suggesting that fungi with large BGC-omes may be undergoing sexual reproduction less frequently? However, I don’t think clonal vs. sexual reproduction frequency fully explains the variation in BGC-omes across the Pezizomycotina.

Is there a relationship between multicellularity, BGC-ome size, and HGT frequency? Previous studies have demonstrated (at least in Fungi) that BGCs are frequently obtained via HGT, but the authors do not address this.

The authors state: “Large-scale annotation efforts have also previously suggested that BGC carriage is limited in many bacterial species18 and databases cataloging the taxonomic origins of known natural products have shown that they are disproportionately sourced from just six microbial phyla19,20 (Fig. 1b).” However, I wonder to what extent the database is biased towards particular more well-studied taxa?

****

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
I am satisfied with the changes made. The article looks excellent in its current state.

(Remarks on code availability)

Reviewer #3

(Remarks to the Author)
Overview: The authors have done a substantial number of new analyses to address my concerns/questions as the those of the other reviewers, which is commendable. I enjoyed reading their response. Overall, I feel the authors provide some very interesting results and new avenues for research thinking that are very intriguing. I also appreciate that the authors made the code and data easily available for future use, which is not always done.

Below are some minor comments/suggestions.

Intro, lines 93-94. The wording could be rephrased given that plants also produce specialized metabolites and it's not just a microbial thing. Perhaps you can change specialized to 'bioactive' since you want to focus on this aspect of specialized metabolism in microbes? Also, fyi I think a lot of people associate 'microbe' with only Bacteria or Archaea and not Fungi; you might want to explicitly say both microbial prokaryotes and eukaryotes (although I know prokaryote is problematic).

Lines 102-104. This sentence is missing references. “The natural products derived from microbial specialized metabolites represent a critical source of bioactive compounds with immense potential for drug discovery, having yielded numerous life-saving pharmaceuticals like antibiotics, immunosuppressants, and anticancer agents.”

Line 142 and 172 (and elsewhere in paper). The authors should not italicize “Ascomycetes” as this is an informal term and not a valid taxonomic name. Change to Ascomycota or ascomycetes.

Line 158. Change “productions” to “products”? “Building on the intriguing pattern of key sources of natural products being multicellular…”

Line 160. The authors added this sentence to the paper “…multicellularity facilitated expansions in specialized metabolite production” but I would suggest changing “facilitated” to something that doesn’t imply causation. Perhaps “associated with”?

Line 179. I would suggest that you cite the paper by Richards & Tablot if you are taking about public goods. Doi:10.1038/nrmicro3108

Line 275. This sentence is a little unclear, perhaps change as follows. “The only family found to be associated with BGC-ome size in four or more of the five phyla was glycosyl hydrolase family 19 that encode chitinases, which are rare in microbes and hypothesized to have been horizontally acquired in Streptomyces from plants26.”

Line 413. I would suggest rephrasing this sentence. Ascomycota aren’t the “hallmark”, the products they produce are, right? Perhaps this instead? “A trove of fungal natural products are produced by Ascomycota, including critically important drugs such as the antibiotic penicillin, the immunosuppressant cyclosporine, and the cholesterol-lowering statins.”

This sentence is also missing a reference.

Lines 415-418. I think the text below clarifies what you mean by this, but this sentence is problematic for me because your premise is that multicellularity is associated with expanded BGCs, yet multicellularity arose before the Ascomycota. Thus, why are BGCs only expanded in certain lineages? You get to this later, but it takes a while.

“Like the other major microbial source of natural product drugs, the Actinomycetota, the Ascomycota contain genera enriched in BGCs and have the ability to form complex cooperative units in the form of mycelia, suggesting that enrichment in biosynthetic potential in fungi is associated with multicellularity.”

I see why you want to start with Ascomycota since they are well-known for producing bioactive molecules, but I think it would make more sense just to start this section with Fungi “To systematically investigate BGC enrichment across Kingdom Fungi, which contain many taxa known for their production of bioactive metabolites, …” and then narrow the focus down to phyla and why only certain lineages are expanded.

Line 419 and elsewhere. Fungi, not fungi (since you are referring to the Kingdom) rather than using it as a common noun or informal term (i.e., fungi often = Fungi + Oomycota).

Line 441. Ditto. “Fonticula, a close outgroup of Fungi.”

Line 442. Will non-phylogenetic biologists know what symplesiomorphic means? :-)

Line 454. What subkingdom are you referring to? Dikarya? Please clarify.

Line 462. Specify what classes are within the “monophyletic subdivision” you are referring to? It is hard to see in the Figure.

Line 502. I feel like the wording of this may be confusing to people as it makes it sound like only Basidiomycota form dikaryons. This goes back to my suggestion that you define Dikarya for non-fungal readers and you can state that the primary synapomorphy of the subkingdom Dikarya (Ascomycota and Basidiomycota) is the formation of dikaryons, cells with two nuclei containing separate genomes that are physically associated and co-replicate.

Line 574. If you are referring to the Pezizomycotina and Agaricomycotina, use ‘subphyla’ instead of “classes”. Adding a comma after subphyla could help clarify the sentence’s meaning too.

(Remarks on code availability)

Decision Letter:

Our ref: NMICROBIOL-25072637A

19th March 2026

Dear Cameron,

Thank you for submitting your revised manuscript "Complex multicellularity is linked with expanded chemical arsenals in microbes" (NMICROBIOL-25072637A). It has now been seen by two of the three original referees and their comments are below. The reviewers find that the paper has improved in revision, and therefore we'll be happy in principle to publish it in Nature Microbiology, pending minor revisions to satisfy the referees' final requests and to comply with our editorial and formatting guidelines.

We are now performing detailed checks on your paper and will send you a checklist detailing our editorial and formatting requirements in about a week. Please do not upload the final materials and make any revisions until you receive this additional information from us.

Thank you again for your interest in Nature Microbiology. Please do not hesitate to contact me if you have any questions.

Sincerely,

Reviewer #1 (Remarks to the Author):

I am satisfied with the changes made. The article looks excellent in its current state.

Reviewer #3 (Remarks to the Author):

Overview: The authors have done a substantial number of new analyses to address my concerns/questions as the those of the other reviewers, which is commendable. I enjoyed reading their response. Overall, I feel the authors provide some very interesting results and new avenues for research thinking that are very intriguing. I also appreciate that the authors made the code and data easily available for future use, which is not always done.

Below are some minor comments/suggestions.

Intro, lines 93-94. The wording could be rephrased given that plants also produce specialized metabolites and it's not just a microbial thing. Perhaps you can change specialized to 'bioactive' since you want to focus on this aspect of specialized metabolism in microbes? Also, fyi I think a lot of people associate 'microbe' with only Bacteria or Archaea and not Fungi; you might want to explicitly say both microbial prokaryotes and eukaryotes (although I know prokaryote is problematic).

Lines 102-104. This sentence is missing references. “The natural products derived from microbial specialized metabolites represent a critical source of bioactive compounds with immense potential for drug discovery, having yielded numerous life-saving pharmaceuticals like antibiotics, immunosuppressants, and anticancer agents.”

Line 142 and 172 (and elsewhere in paper). The authors should not italicize “Ascomycetes” as this is an informal term and not a valid taxonomic name. Change to Ascomycota or ascomycetes.

Line 158. Change “productions” to “products”? “Building on the intriguing pattern of key sources of natural products being multicellular…”

Line 160. The authors added this sentence to the paper “…multicellularity facilitated expansions in specialized metabolite production” but I would suggest changing “facilitated” to something that doesn’t imply causation. Perhaps “associated with”?

Line 179. I would suggest that you cite the paper by Richards & Tablot if you are taking about public goods. Doi:10.1038/nrmicro3108

Line 275. This sentence is a little unclear, perhaps change as follows. “The only family found to be associated with BGC-ome size in four or more of the five phyla was glycosyl hydrolase family 19 that encode chitinases, which are rare in microbes and hypothesized to have been horizontally acquired in Streptomyces from plants26.”

Line 413. I would suggest rephrasing this sentence. Ascomycota aren’t the “hallmark”, the products they produce are, right? Perhaps this instead? “A trove of fungal natural products are produced by Ascomycota, including critically important drugs such as the antibiotic penicillin, the immunosuppressant cyclosporine, and the cholesterol-lowering statins.”

This sentence is also missing a reference.

Lines 415-418. I think the text below clarifies what you mean by this, but this sentence is problematic for me because your premise is that multicellularity is associated with expanded BGCs, yet multicellularity arose before the Ascomycota. Thus, why are BGCs only expanded in certain lineages? You get to this later, but it takes a while.

“Like the other major microbial source of natural product drugs, the Actinomycetota, the Ascomycota contain genera enriched in BGCs and have the ability to form complex cooperative units in the form of mycelia, suggesting that enrichment in biosynthetic potential in fungi is associated with multicellularity.”

I see why you want to start with Ascomycota since they are well-known for producing bioactive molecules, but I think it would make more sense just to start this section with Fungi “To systematically investigate BGC enrichment across Kingdom Fungi, which contain many taxa known for their production of bioactive metabolites, …” and then narrow the focus down to phyla and why only certain lineages are expanded.

Line 419 and elsewhere. Fungi, not fungi (since you are referring to the Kingdom) rather than using it as a common noun or informal term (i.e., fungi often = Fungi + Oomycota).

Line 441. Ditto. “Fonticula, a close outgroup of Fungi.”

Line 442. Will non-phylogenetic biologists know what symplesiomorphic means? :-)

Line 454. What subkingdom are you referring to? Dikarya? Please clarify.

Line 462. Specify what classes are within the “monophyletic subdivision” you are referring to? It is hard to see in the Figure.

Line 502. I feel like the wording of this may be confusing to people as it makes it sound like only Basidiomycota form dikaryons. This goes back to my suggestion that you define Dikarya for non-fungal readers and you can state that the primary synapomorphy of the subkingdom Dikarya (Ascomycota and Basidiomycota) is the formation of dikaryons, cells with two nuclei containing separate genomes that are physically associated and co-replicate.

Line 574. If you are referring to the Pezizomycotina and Agaricomycotina, use ‘subphyla’ instead of “classes”. Adding a comma after subphyla could help clarify the sentence’s meaning too.

Version 2:

Decision Letter:

12th May 2026

Dear Cameron,

I am pleased to accept your Article "Complex multicellularity is linked with expanded specialized metabolite production in microbes" for publication in Nature Microbiology. Thank you for having chosen to submit your work to us and many congratulations.

Over the next few weeks, your paper will be copyedited to ensure that it conforms to Nature Microbiology style. We look particularly carefully at the titles of all papers to ensure that they are relatively brief and understandable.

Once your paper is typeset, you will receive an email with a link to choose the appropriate publishing options for your paper and our Author Services team will be in touch regarding any additional information that may be required. Once your paper has been scheduled for online publication, the Nature press office will be in touch to confirm the details.

You may wish to make your media relations office aware of your accepted publication, in case they consider it appropriate to organize some internal or external publicity. Once your paper has been scheduled you will receive an email confirming the publication details. This is normally 3-4 working days in advance of publication. If you need additional notice of the date and time of publication, please let the production team know when you receive the proof of your article to ensure there is sufficient time to coordinate. Further information on our embargo policies can be found here:
https://www.nature.com/authors/policies/embargo.html

After the grant of rights is completed, you will receive a link to your electronic proof via email with a request to make any corrections within 48 hours. If, when you receive your proof, you cannot meet this deadline, please inform us at rjsproduction@springernature.com immediately. You will not receive your proofs until the publishing agreement has been received through our system

Due to the importance of these deadlines, we ask you please us know now whether you will be difficult to contact over the next month. If this is the case, we ask you provide us with the contact information (email, phone and fax) of someone who will be able to check the proofs on your behalf, and who will be available to address any last-minute problems.

Acceptance of your manuscript is conditional on all authors' agreement with our publication policies (see https://www.nature.com/nmicrobiol/editorial-policies). In particular your manuscript must not be published elsewhere.

Authors may need to take specific actions to achieve compliance with funder and institutional open access mandates. If your research is supported by a funder that requires immediate open access (e.g. according to Plan S principles or the NIH public access policy) then you should select the gold OA route, and we will direct you to the compliant route where possible. Because authors warrant under our subscription licensing terms that they haven't committed to licensing any version of their article under a licence inconsistent with the terms of our agreement – including the applicable embargo period – publication under the subscription model isn't suitable for authors whose funders require no embargo.

If you have any questions about our publishing options, costs, Open Access requirements, or our legal forms, please contact ASJournals@springernature.com

An online order form for reprints of your paper is available at https://www.nature.com/reprints/author-reprints.html. All co-authors, authors' institutions and authors' funding agencies can order reprints using the form appropriate to their geographical region.

We welcome the submission of potential cover material (including a short caption of around 40 words) related to your manuscript; suggestions should be sent to Nature Microbiology as electronic files (the image should be 300 dpi at 210 x 297 mm in either TIFF or JPEG format). Please note that such pictures should be selected more for their aesthetic appeal than for their scientific content, and that colour images work better than black and white or grayscale images. Please do not try to design a cover with the Nature Microbiology logo etc., and please do not submit composites of images related to your work. I am sure you will understand that we cannot make any promise as to whether any of your suggestions might be selected for the cover of the journal.

You can now use a single sign-on for all your accounts, view the status of all your manuscript submissions and reviews, access usage statistics for your published articles and download a record of your refereeing activity for the Nature journals.

To assist our authors in disseminating their research to the broader community, our SharedIt initiative provides you with a unique shareable link that will allow anyone (with or without a subscription) to read the published article. Recipients of the link with a subscription will also be able to download and print the PDF.

As soon as your article is published, you will receive an automated email with your shareable link.

Congratulations once again and I look forward to seeing the article published.

With kind regards,

P.S. Click on the following link if you would like to recommend Nature Microbiology to your librarian http://www.nature.com/subscriptions/recommend.html#forms

** Visit the Springer Nature Editorial and Publishing website at www.springernature.com/editorial-and-publishing-jobs for more information about our career opportunities. If you have any questions please click here.**

In response to the reviewers' constructive and insightful suggestions, we conducted several new analyses that have strengthened our study. Below is a brief overview of this additional work, which we detail further in our responses to specific reviewer comments.

- Performed a more thorough investigation of the origin and evolution of ansamycin BGCs in *Actinomycetes* (**Extended Data Fig. 7**).
- Updated our investigation of heterokaryon incompatibility in fungi to comprehensively assess protein domains from characterized proteins associated with the cellular function (**Extended Data Fig. 9**).
- Performed multi-approach, systematic investigations of recent horizontal gene transfer in bacteria and fungi (**Extended Data Fig. 3**).
- Evaluated the diversity of distinct Pfam domains mapping to key biosynthetic genes as a function of ribosomal diversification for each bacterial order (**new panel Fig. 1c**).
- Assessed literature for support on associations between biosynthetic traits and multicellularity in other eukaryotic lineages besides fungi, including animals, slime molds, and plants.

Reviewer #1 (Remarks to the Author):

The paper by Salamzade et al. describes a very interesting study linking multicellular lifestyles of bacteria and fungi to expanded chemical arsenals in microbes. This is a carefully conducted study with a clear storyline, rigorous analyses and clear figures. While it is not the very first paper to link biosynthetic potential with multicellularity, it is the first one to do it this explicitly, comprehensively and quantitatively. The code for all analyses is also well documented. Overall, it will make a nice contribution to scientific literature. Still, I have a number of remarks that I hope will be useful to the authors:

Thank you for your feedback and assessment of our study. We have updated our text, figures, and code to reflect your suggestions for improvement, as detailed below.

- Line 93-94: The link between multicellularity and secondary metabolic complexity has been made before. It would be good to acknowledge/cite reviews that have discussed this (e.g., doi.org/10.1111/brv.12605, doi:10.1128/jb.00153-23, doi:10.1039/C5NP00013K). Also, at least one large-scale study of microbial biosynthetic diversity has identified a connection between multicellularity and biosynthetic potential (doi:10.1016/j.cell.2014.06.034, Table S2).

We appreciate the reviewer for their comments and for pointing out these 4 papers. We have adjusted the wording of these lines to clarify that this has not been ‘systematically investigated’ (*lines 82-86*). We have also carefully considered these suggested citations. Here are our thoughts on each of them:

doi.org/10.1111/brv.12605 (Naranjo-Ortiz and Gabaldón 2020, *Biological Reviews*). This excellent review of fungi was already cited and we have added reference to it here as well.

doi:10.1128/jb.00153-23 (Schlimpert and Elliot 2023, *JBact*). This is an excellent review paper on model species of *Streptomyces.* We unfortunately are over the limit in the number of citation, given it does not address the evolution of biosynthetic enrichment we felt it was less relevant to our work than the other papers we cite.

doi:10.1039/C5NP00013K (Traxler and Kolter 2015, *Nat Prod Rep*). This is another excellent review paper, which includes discussion of natural products and multicellularity. We have now added it to the paper.

doi:10.1016/j.cell.2014.06.034, Table S2 (Cimermancic et al. 2014, *Cell*). We have added this citation to our paper. This is a key study in the field, which we and many others have taken inspiration from. We appreciate the reviewer pointing out that supplemental Table S2 features a plot showing the relationship between multicellularity and biosynthetic potential that we had missed, in large part cause the words 'multicellular/multicellularity do not appear in the main text of the paper. We do want to note that it is only referenced as a single sentence in the supplemental/extended text and methodological details pertaining to the analysis, including those important for understanding the chart in Table S2, are unavailable.

- Line 92: 'facilitated' and Line 186: 'due to'. In these sentences, directional causality is implied, while there does not seem to be data that clearly supports this. The evidence is associational.

Thank you for catching these instances. We agree that it is important to be careful about stating directionality in the introduction and result sections. We have adjusted the wording to "closely associated with" from "facilitated" (*lines 85-86*) and to "coinciding with" from "due to" (*lines 171-173*).

- In the analysis of biosynthetic repertoires, BGC richness among taxa is analysed extensively. Adding also diversity metrics (how diverse are BGC repertoires of related species/strains) would be very useful to provide a more complete picture. The suggestion that multicellular organisms might be prime sources for novel metabolite discovery (line 472, closing sentence of the paper) could also be substantiated by such an analysis.

Thank you for your suggestion, this is a great point. We avoided investigating intra-taxon BGC diversity mostly because there was a recent study that explored this in depth (Gavriilidou, Kautsar *et al.* 2022, doi: 0.1038/s41564-022-01110-2).

While we were already citing this work, we now directly reference it in the first Results section (*line 144*) and have added a small analysis to complement results from that study. While their study primarily determined intra-taxon biosynthetic diversity by clustering analogous BGCs into gene cluster families (GCFs) - we looked at the number of distinct Pfam domains mapping to key biosynthetic genes as a function of ribosomal diversification for each order. Results from this analysis are now shown in **Fig. 1c**.

- Line 113-115: interesting hypotheses are provided for the links between secondary metabolite production and multicellular lifestyles. However, this does not necessarily explain the expansion of numbers of BGCs per genome; even with a single BGC, a bacterium might invest a lot of energy making a lot of that metabolite. The fact that a bacterium has more BGCs in its genome does not necessarily mean that it invests more energy in secondary metabolism throughout its life cycle.

Thank you for your feedback. We agree that unicellular organisms could invest substantial amounts of energy into secondary metabolism of even a single BGC. We have reworded these lines to now read: "Finally, we discuss the association between multicellularity and the genomic enrichment in both specialized metabolism and carbohydrate-active enzymes, including through

the lens that in extracellular environments these molecules act as public goods whose production and regulation may be shaped by tragedy of the commons dynamics." (*lines 105-109*).

- Line 122: the summed lengths of predicted BGC regions were used. Was any effort made to trim/correct greedily identified borders? Otherwise, the total size could easily be heavily overestimated.

Thank you for pointing this out. No, we did not refine BGC borders. We now provide an additional panel in **Extended Data Fig. 1** showing that there is strong correlation between the protocore region size of BGCs with the total BGC region size and have added a sentence in the Results clarifying this (*lines 121-124*):

> "Both the median number of BGCs as well as the median BGC-ome size, the summed length of predicted BGC regions in a single genome, were assessed across 4,732 genus-representative genomes from the orders (**Fig. 1a; Supplementary Table S1**). A strong correlation was observed between the two metrics and while complete genomes were preferentially used where available, BGC-ome size was largely assessed throughout this study to alleviate over-assessment of biosynthetic capacity arising from assembly fragmentation issues. In addition, although BGC-ome sizes reported might be inclusive of flanking genomic regions unrelated to specialized metabolism, trends in BGC-ome size were correlated with other measures of biosynthetic capacity, including those applying stricter criteria for defining BGCs and their proximities (**Extended Data Fig. 1**)."

- Line 273: if vertical evolution is inferred, can this be substantiated by checking that gene phylogeny follows species phylogeny? And/or through a parsimony analysis to show that this is the most parsimonious model for the data?

Thank you for your interest and suggestions relating to the evolution of the ansamycin and AHBA biosynthesis machinery. We agree that further investigation was warranted and have greatly improved and expanded our evolutionary analysis relating to the origins of ansamycin BGCs. Due to space limits in the main text, much of this analysis is detailed in the Supplementary Notes section: "**Evolutionary analysis of anscamycin synthesizing BGCs and rifamycin resistance across *Actinomycetota"***. The text there discusses the content of Extended Data Figure 7 (formerly 6), which has also been substantially updated.

Briefly, we added a phylogenetic analysis of AHBA synthase, improved the composite phylogenetic analysis of ansamycin BGCs to be at the domain-resolution, and investigated overlap in domain-resolution ortholog groups between BGC instances.

Ultimately, we now clarify that both vertical and horizontal transfer appear to have contributed to the evolution of ansamycin BGCs. Our main conclusion that ansamycin biosynthesis is an innovation originating in *Actinomycetes* remains and is further supported by these new analyses.

Minor:

- Figure 1c: 'Number of distinct metabolites'. I guess this refer to 'known metabolites'. Please specify.

Thank you for your comment. Yes, this is known/characterized metabolites. We agree it is good to clarify this to avoid confusion and have updated the figure label accordingly.

- Line 321: 'two independent origins' -> Figure 3A seems to suggest three?

Thank you for noting this, we now recognize that our wording was misleading. In this section we are specifically addressing expansions in *Dikarya*. We have attempted to clarify the wording, as follows: "Next, in our evaluation of the evolution of biosynthesis in fungi, we find support for two independent origins of BGC enrichment occurring separately within each of the two phyla of *Dikarya*, each coinciding with developments in physiological complexity and the formation of macroscopic sexual fruiting bodies, such as cup-shaped ascocarps and mushroom basidiocarp structures." (*lines 282-286*)

We should note that the reason we do not highlight the *Neocallimastigomycota* more generally in the paper is because they are relatively constrained in biodiversity and their biosynthetic enrichment is from lateral acquisition of BGCs from bacteria within rumens as shown by Swift et al. 2021 (doi: 10.1073/pnas.2019855118).

Reviewer #1 (Remarks on code availability):

The GitHub repository is well structured and the code is well documented too. There is also a README file that clearly outlines what can be found where.

Thank you for checking the associated code repositories, we appreciate your feedback.

Reviewer #2 (Remarks to the Author):

This paper provides a broad phylogenetic analysis of the coincidence of specialized metabolite production and multicellularity in bacteria and fungi. The phylogenetic analysis appears to be quite rigorous, and the statistical associations are not terribly surprising, but nice to see evaluated in a formal fashion and to have all of these analyses in one place. My comments are rather limited.

Thank you for your feedback and suggestions to improve our manuscript.

1) The authors need to provide an objective definition of a BGC at the outset. Bugs synthesize all kinds of metabolites, from amino acids to nucleotides to cofactors, up to the more advanced compounds that the authors seem to be focusing on, but where is the line drawn?

Thank you for your comment. We have changed the opening few sentences of the Introduction to improve our description of what we regard as specialized metabolites and aid reader comprehension (*lines 34-39*).

We also now directly mention that we applied antiSMASH, a commonly used software in microbial genomics for BGC annotation that allows for the comprehensive detection of multiple types of BGCs encoding a variety of specialized metabolites. Specifically, we added the following statement to the final introductory paragraph: "To evaluate the breadth of known microbial specialized metabolite classes — including polyketides, non-ribosomal peptides, and terpenes — we comprehensively annotated for over 80 types of BGCs using the software antiSMASH[21]." (*lines 90-92*).

2) This is not entirely on the authors, but more a comment on the extremes to which journals now seem to be encouraging things. The font sizes and tiny graphs included here are approaching a rather ridiculous level. Readers should not have to expand a pdf to the entire width of a screen in hopes of seeing things; they ought to be visible on a 8 x 11 hardcopy.

Thank you for your feedback and we apologize for any difficulties you experienced in assessing the figures. We agree that it is important to have figures legible if printed and have attempted to adjust the main figures accordingly. Some of the extended data figures are dense with panels still; however, they are provided in high-resolution to allow readers to zoom-in on individual panels.

3) The authors favor the hypothesis that multicellularity paved the way for the evolution of biosynthetic pathways, and there certainly appears to be an association. I'm not sure how it can be teased out, but how do we know that the order isn't reversed? Some discussion of this is needed.

We greatly appreciate the reviewer's comment; it prompted us to give more consideration to this potential. Below we outline our thinking as well as changes to the manuscript.

Yes, it is possible that specialized metabolites facilitated the origin of multicellularity. This was something we hinted at in the discussion of our original manuscript and have now slightly expanded on in our revision. It is important to note that our hypothesis is that the origin of multicellularity facilitated substantial enrichment in BGCs. We think it is unlikely that the transition to multicellularity required a large suite of specialized metabolites, rather it is much more likely that one or a few specialize metabolite might have played a role. Given that virtually all bacterial orders have at least some genera with at least a few BGCs (**Fig. 1d**), it is highly likely that each multicellular lineage we investigated evolved from a unicellular ancestor had the capacity to produce one or a few specialized metabolites. Finally, it is hard to envision unicellular taxa evolving an enrichment in specialized metabolism in the absence of both cellular differentiation and reduction in the costs associated with public goods (*see points below*).

To address this in our paper we have added the following text in supplemental material. We note that we do not have space to expand the discussion in the main text.

'Our work and hypothesis do not focus on whether multicellularity is associated with the ability to produce a few specialized metabolites, but specifically whether it is associated with a substantial enrichment of BGCs. While it is possible that substantial enrichment of BGC occurred prior to the transition from unicellular to multicellular growth, we think this is less likely due to the following:

a) Carriage of a larger number of BGCs has a cost, even just for replication of the resulting larger genome. The benefits from carriage of a large numbers of BGCs is clear for

multicellular organisms (i.e., reduced conflict associated with public goods).

b) Unlike unicellular organisms, multicellular organisms have an expanded division of labor of cells. This is predicted to increase the benefits conferred from a higher diversity of BGCs, such as parallel production of different antibiotics where different cells of a multicellular organism specialize in the production of individual ones.

c) It is increasingly clear that for microorganisms specialized metabolites play important roles mediating ecological interactions. For each of *Cyanobacteriota*, *Myxococcota,* and *Actinomycetota* we show positive associations between counts for known protein markers of multicellularity and sporulation with BGC-ome size (**Fig. 1e-g**). This supports a view that an important benefit associated with an increased degree of cooperation is an increase in diversity in specialized metabolites.

4) The issue of division of labor is raised in the paper, and this does indeed seem to be an issue in the evolution of complex multicellularity. But no evidence in provided on this, so this should also be elaborated on a bit.

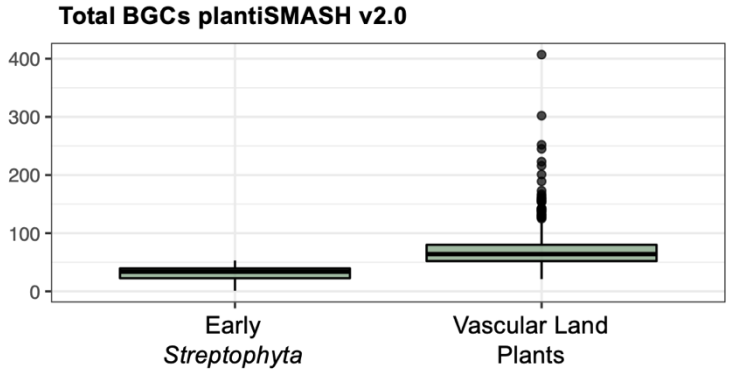
Thank you for your comment. It led us to two recent and important studies pertaining to microbial division of labor dynamics. Zhang et al. 2020 (doi: 10.1126/sciadv.aay5781) found that genotypic heterogeneity led to increased antibiotic production and division of labor dynamics in *Streptomyces* (multicellular actinomycetes that has historically contributed the most natural products from bacteria). Furthermore, while *Streptomyces* is one of the most complex bacterial taxa, other *Actinomycetes* have been described that exhibit more simple and transient division of labor arrangements that fit our findings (doi: 10.1016/j.mib.2022.102148). We have added citations to these papers and expanded on it in the Discussion (*lines 362-367*).

5) If the authors' hypothesis is correct and general, then other lineages of multicellulars would be expected to produce complex external metabolites. It would be helpful to summarize some data on this, which likely exists in some of the photosynthetic eukaryote lineages.

Thank you for your encouragement to dig deeper into whether developments of multicellularity parallel biosynthetic expansions in other branches of life beyond bacteria and fungi. We have added a small **Supplementary Notes** section ("Cursory assessment of associations between multicellularity and biosynthetic potential in non-fungal eukaryotes") briefly touching on existing knowledge of the biosynthetic potential of other key multicellular organisms (e.g. slime molds, animals, and plants) and their phylogenetic neighbors that exhibit simpler morphologies. This section is referenced now in the second paragraph in the Discussion section (*lines 351-357*).

We think the author's comment specifically alludes to plants, a well-known source of natural products. To our knowledge, differences in the biosynthetic potential between morphologically simpler green algae and more advanced land plants has not been investigated. However, recently some of the authors of antiSMASH (which we used for annotation of BGCs in this study and is the most popular software for the task), also released a new version of their software plantiSMASH - including pre-computed BGC predictions for various plants in the phylum of *Streptophyta.* Using data from their preprint's supplementary data, we performed a quick analysis and found that vascular land plants, which exhibit more complex morphologies and life cycles, have higher biosynthetic potential than morphologically simpler early-diverging plants in

the phylum. Code for the analysis can be found in the GitHub repo associated with this study along with a citation pointer to the plantiSMASH study.



Reviewer #3 (Remarks to the Author):

Summary: The manuscript "Complex multicellularity linked with expanded chemical arsenals in microbes" analyzes existing genomic data to address the question of how the evolution of multicellularity is related to an organism or lineages genomic diversity biosynthetic gene clusters. Overall, I found the manuscript to be well-written and engaging to read, and the complexity and scale of the analyses across different microbial groups are impressive. The results certainly provide strong evidence for the association between multicellularity and BGC diversity in both Bacteria and Fungi and I mainly have suggestions for clarification of some of the methods or statements.

Thank you for your feedback and questions. We have attempted to address them and think they have contributed to a stronger manuscript.

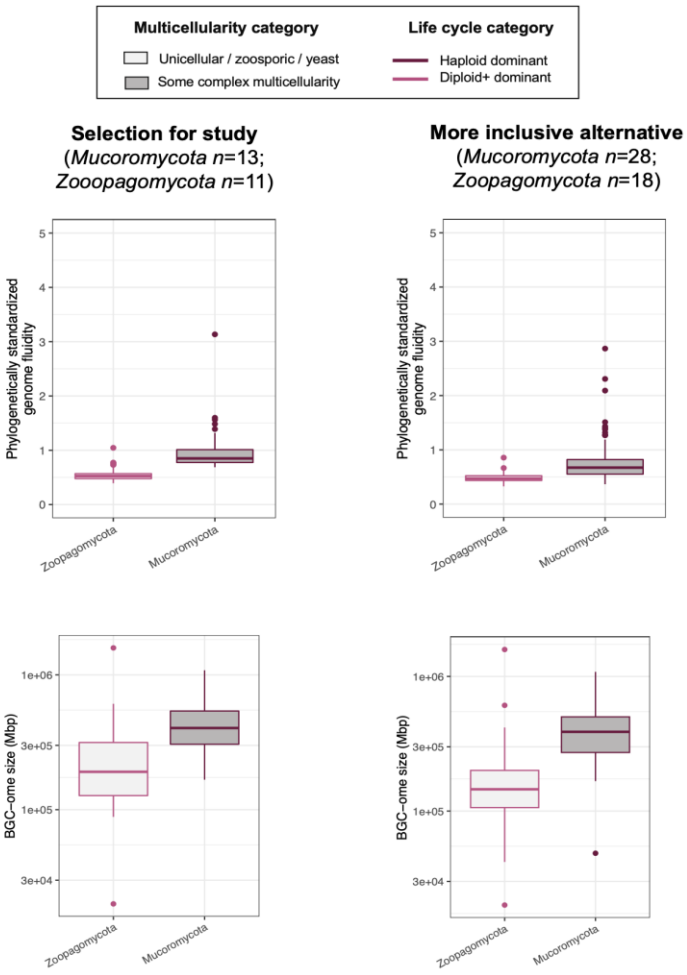
Questions/Suggestions:

My major concern about the methods is how the taxon selection might skew the results. For example, it appears that the fungal genomes that were used are quite skewed towards certain groups (e.g., chytrids), while other lineages (e.g., "zygos" such as Basidiobolus) are underrepresented. I understand that not all genomes can be analyzed, but it seems that if you had to choose 312 genomes you would want to make sure that all clades are well represented, to the extent possible with existing data. For example, Supplemental Table S7 lists the fungal genomes that were included in the analyses. In looking at these, it seems that Chytrid genomes are over-represented compared to other non-Dikarya lineages (i.e., 23% of genomes), such as Mucoromycota (5% of genomes). Along those same lines, why did the authors restrict genomes to those that did not have "fam_Incertae_sedis" or "unidentified" in their taxonomic designations in the UFCG database when the types of analyses that they performed did not rely heavily on these lower taxonomic levels? A huge diversity of Fungi are not yet fully identified or named. By excluding these genomes my concern is that the genomes are not fully representative of the genomic diversity of fungi.

We understand and appreciate these concerns. We have now extensively assessed the impact of including additional genomes for taxa that were of concern for being unsampled or under-sampled.

As noted in our paper, we included 312 fungal genomes, selected to be as broadly representative of the genomic diversity of fungi as possible. Nevertheless, as the reviewer noted, some lineages have less representation in our analyses than others. This is almost entirely due to fewer genomes being available for these groups, which is in part a reflection of these fungal groups likely not being as diverse (i.e., it is estimated that there are many more dikaryotic fungal species than non-dikaryotic fungal species, see for example doi: 10.1111/brv.12418). As we will detail below, we included virtually all high quality non-*Dikarya* genomes from UFCG. Additional non-*Dikarya* genomes from Amses et al. 2022 were also included, which was a comprehensive analysis of the fungal kingdom similar to our study. This led to a total of 11 *Zoopamycota* (11 distinct genera), 16 *Mucoromycota* (13 distinct genera), 72 *Chytridiomycota* (including *Neocallimastigomycota*), 6 *Blastocladiomycota*, and 3 early-diverging lineages (including *Rozellomycota*) being used in our study.

The UFCG, a comprehensive database of fungal genomes, was established in 2022, and contains genomes for 1,587 species. Of these, one species' genome has been suppressed from being downloaded by NCBI and another 20 genomes have too low of an assembly quality (N50 < 10,000) to investigate BGC content. Of the remaining 1,566 species-representative genomes, 1,147 are *Ascomycota* and 341 are *Basidiomycota*. This leaves only 78 non-dikaryotic species in consideration. Only two of these were unidentified and dropped because they are not recognized in NCBI Taxonomy, described as yeast-like symbionts in UFCG. Five were further flagged as "*fam_Incertae_sedis*", however, because of our inclusion of genomes from Amses et al., we actually included all five of these in our study.



We do note that there are two groups of non-*Dikarya* fungi where we could expand the inclusion of available genomes: the *Mucoromycota* and *Zoopagomycota*. We also want to point out that our analyses were further designed to account for under sampling and diversity differences between phyla. For instance, the psaps method specifically standardizes the commonly used genomic fluidity metric for core genome phylogenetic breadth.

To assess whether under sampling of these 2 phyla would have an impact on conclusions in our study, we investigated

the distributions of BGC-ome sizes and phylogenomically standardized genome fluidity measures from psaps between our current selection and if we had included all genera with available genomes belonging to *Zoopagomycota* and *Mucoromycota* from UFCG. Ultimately, trends for both measures were robust across the two genome sets - *Mucoromycota* tended to possess greater BGC-ome sizes and phylogenetically standardized genomic fluidity compared to *Zoopagomycota*. Figure included here in our response for your information.

In the methods, the authors state "Representative genomes were selected based on BGC-ome size." More information would be useful. i.e., were rep genomes chosen to represent a diversity of BGC-ome sizes, or the maximum sizes, etc?

Thank you for pointing out the lack of detail here. We have now expanded our explanation on why we used BGC-ome size to select genus-level representative genomes for the 20 genera with multiple genomes (*lines 558-560*):

> For comparative analyses of biosynthetic genes or gene clusters across taxonomic groups, only a single representative genome per genus was regarded, the one with the largest BGC-ome size. Most of the genera with multiple genomes were non-*Dikarya* (18 of 20) and this approach would allow us to more conservatively assess whether they are depleted in BGCs relative to the higher fungi.

Is there a table that lists the Pfams with NACHT or HET domains that were used in the analyses? In the supplementary methods, the authors state "Domain architectures for HI-associated proteins were assessed based on comprehensive Pfam annotations using custom scripts provided in the associated code repository of this study." It would be useful to have more information in the Supplementary methods as to how this was done rather than having to go hunt down the github code.

Thank you for your feedback. We have added a *new* **Supplementary Table S11** with domain architectures of proteins with the HET domain and expanded on the methods description. The full Pfam-A database was used for determining protein architectures. Note, we de-emphasized investigations into NACHT after re-investigating HI, please see below for our response to your other comment on HI.

You may want to define dikaryon/Dikarya for non-fungal biologists?

Thank you for this suggestion. We agree that it is valuable to describe *Dikarya* and have also added definition of dikaryon to make the study accessible to folks with different backgrounds (*lines 284-286* and *lines 298-301*).

It is really interesting that the authors identified HI domains as being associated with BGC diversity. However, I do wonder whether all of the putative HI genes are functioning in that capacity? For example, Neurospora crassa has 8-10 het loci, but a search of the genome for het annotations results in many more hits. Did the authors check to see whether the proteins that were identified as having HI domains were validated in species where this is known? You mention that these proteins have homology to antiphage proteins in bacteria, but are there other known functions for proteins with these domains in fungi?

Thank you very much for your interest, encouragement to dig deeper, and suggestions here. We have now significantly improved upon our HI-related analyses and substantially updated **Extended Data Fig. 9** (formerly Extended Data Fig. 8). We have also updated our discussion around the potential role of proteins with HI-associated domains beyond HI.

As a recap, we investigated HI because HET-containing ortholog groups were detected as associated with increased BGC-ome size by GWAS. We then systematically identified HI-associated proteins that featured at least one of two protein domains, HET or NACHT, analogous to the study by Van der Nest *et al.* 2014.

In the study by Zhao *et al.* 2015 from the lab of Professor Louise Glass (doi: 10.1093/molbev/msv125), a leading and pioneering expert on this topic, they state “The HET domain contains approximately 150 amino acids and has no identified cellular or biochemical function other than its association with vegetative incompatibility.” In this same paper, the authors identify 73 proteins in *Neurospora crassa* with homology to a HET consensus sequence that they putatively regard as involved in HI-interactions and find that 34 of them exhibit signatures of balancing selection, which they explain would be expected for allorecognition proteins. For this study, we included two *Neurospora* genomes from different species and identified 28 and 32 proteins with HET domains, respectively.

Then, in the study by Kibby et al. 2023 by the Whitely lab (doi: 10.1016/j.cell.2023.04.015), the authors showed that the NACHT domains in fungi were likely acquired from bacteria and originally functioned in viral defense. Based on phylogenetic analysis of the NACHT domain, they identified a specific clade of fungal NACHT domains that corresponded to proteins functioning in self-recognition (clade 18 from Figure 2 of their study). As such, while we think the presence of the HET domain very likely relates to heterokaryon incompatibility, proteins with only the NACHT domain without HET might function in something besides HI, perhaps viral defense similar to NACHT’s function in bacteria. We have thus deemphasized mention of NACHT in the study.

To perform a more thorough investigation, we manually put together a multi-FASTA database of 47 characterized HI-associated proteins from *Ascomycota* and determined the domain architecture of them using Pfam-A (*new* **Extended Data Fig. 9a**; *lines 641-671 of Supplementary Notes and Methods document*). We then searched for these HI-associated domain architectures in our genome set and reported the average copy-count and range of copy-counts for each domain architecture for the BGC-enriched *Pezizomycotina*, *Agaricomycetes*, and other fungi. Ultimately, our investigation revealed that the main architecture of interest are proteins with only HET domain(s) – what drove our GWAS signal. Such proteins have further been validated as related to heterokaryon incompatibility in three separate filamentous *Ascomycota* genera.

Also, your inquiry actually led us to reassess whether it was appropriate to search for domains using hmmsearch without the “--domtblout” option (or equivalent in PyHMMER; due to reasons discussed here: https://github.com/althonos/pyhmmer/issues/65). While we were already using this option for some HI-related analyses, we have now updated all other analyses (throughout the paper) that involved annotating domains with hmmsearch but had not used this option. This resulted in slight adjustments to **Fig. 3b** as well as **Fig. 4a/Extended Data Fig. 9**. The conclusions of the study were not affected.

Pertaining to our thoughts on alternate functions of proteins with HI-associated domains, we have updated the Discussion text (*lines 387-390*) to focus more on the HET domain and

speculate about its potential links to exciting new findings of uneven chromosome distributions across nuclei in some *Pezizomycotina* by Xu et al. 2025, *Science* (doi: 10.1126/science.abn7811):

"The ability to self-distinguish is likely important for fungi when competing against other fungal species for access to potentially limited carbon sources within their natural habitats; however, HI could have other roles in fungal biology, including viral defense[61] and promoting haplontic life cycles. Recent research has also shown that multiple species of *Pezizomycotina* have uneven distributions of chromosomes across nuclei when in a haploid state[62] and HI-associated proteins might relate to this atypical biological phenomenon."

For the sentence "Pezizomycotina and Agaricomycetes likely originated around the same time as the emergence of land plants" the authors provide a reference to a review paper by Berbee (ref 46). Rather than only citing a review paper I would also directly cite the research underlying this statement, such as Lutzoni et al., 2018 https://www.nature.com/articles/s41467-018-07849-9

Thank you for your suggestion, we have added a citation to Lutzoni *et al.* 2018.

Lines 417-420. The statement "Indeed, our finding that BGC expansions in both fungi and bacteria coincides with significant enrichment in carbohydrate utilization enzymes suggest that catabolic processes help these microbes obtain the energy required to produce specialized metabolites at ecologically relevant scales." is consistent with earlier work by Franco et al., (2022) New Phytologist that demonstrated that expansion of BGCs in a clade of xylarialean fungi was associated with increased numbers of CAZymes (as well as other saptrotrophic/pathogenicity-related genes) in non-symbiont genomes, yet not in fungi isolated as symbiotic endophytes. This suggests that while this trend may be generally true across fungi, there may be exceptions when more closely related taxa are analyzed?

Thank you for your perspective and sharing that interesting study. Yes, we agree and now address the likely impact of alternate factors shaping variability in CAZyme and BGC distributions at finer phylogenetic resolutions. We also now cite the study you reference as part of the changes we made to the Results text pertaining to your next comment (see below).

For the Discussion sentence referenced in this comment, we think it is appropriately worded. We are simply suggesting that CAZymes are co-associated with biosynthetic potential and providing a possible explanation. Based on our interpretation of Figure 1 from the study you shared, although the endosymbionts in the possible clade referenced possess high biosynthetic potential while not having more CAZymes, biosynthetic associations might have occurred ancestrally for the clade before the fungi developed a symbiotic association with plants. This could occur for instance if endosymbionts lost unneeded CAZymes as their new lifestyle allowed them to specialize and need to process only a limited set of carbohydrates produced by the plants they associate with. Alternatively, if the endosymbionts in this clade were generalists then perhaps they lost CAZymes because they acquire simpler nutrients from their host plants that need less processing.

The authors discuss how Pezizmycotina has expanded BGC-omes compared to Agaricomycotina, but there is a very large amount of variation in BGC diversity among the Pezizmycotina that the authors do not address. How does that variation fit into the model they propose? I think it they authors are suggesting that fungi with large BGC-

omes may be undergoing sexual reproduction less frequently? However, I don't think clonal vs. sexual reproduction frequency fully explains the variation in BGC-omes across the Pezizomycotina.

We agree that there are very interesting patterns to explore relating to the diversity of BGC-ome sizes within and between the *Agaricomycotina* and *Pezizomycotina*. It is important to note that our sampling and study design limits our ability to address the intra-taxon diversity within *Pezizomycotina* or other fungal clades in this study. Pertaining to life cycles, we are only suggesting that *Pezizomycotina* are generally more enriched in BGCs than *Agaricomycotina,* potentially in part because they appear to be in a haploid state more often. We believe future studies designed to follow-up on this question would be exciting.

With regards to intra-taxon/clade variability of biosynthetic content, including *Pezizomycotina* and *Agaricomycotina*, we think differences in lifestyles and inter-species interactions likely underlie variability in biosynthetic capacity. As the study you referenced in your previous comment showed, generalist fungi might have more diverse BGCs and more pressure to potentially acquire or develop new ones via HGT or duplication. We have updated the final sentence of the Results (*lines 330-334*) to now elaborate that lifestyles also contribute to biosynthetic capacity referencing the study by Franco *et al.* 2022 in addition to a recent preprint that systematically assessed the impacts of lifestyle across the fungal kingdom (doi: 10.1101/2024.09.14.613087) and a perspective on the biosynthetic potential of lichen-forming fungi (doi: 10.1111/nph.70003).

Is there a relationship between multicellularity, BGC-ome size, and HGT frequency? Previous studies have demonstrated (at least in Fungi) that BGCs are frequently obtained via HGT, but the authors do not address this.

Thank you for your inquiry. We agree that horizontal transfer is a relevant factor shaping the biosynthetic landscape of both bacteria and fungi. However, the frequency of horizontal transfer is difficult to measure, especially at the deeper phylogenetic levels we focus on in our study here. To assess only the impact of recent horizontal transfer on BGC evolution, we have updated and applied our program codoff - which, in the latest version, reports a "Discordance Percentile" metric. Lower values of this metric indicate discordant codon usage of a BGC to its genomic context, thus indicating possible recent acquisition via horizontal transfer. We tested codoff on BGCs across >1,000 complete *Streptomyces* genomes (https://github.com/Kalan-Lab/codoff/wiki/Large-scale-application-to-BGCs-from-Streptomyces) and found that BGCs on the tail end of chromosomes exhibited more discordant codon usage with respect to their genomic contexts. This matches expectations since tail regions have been shown to more frequently undergo HGT (dois: 10.1099/mgen.0.000525, 10.1007/978-3-642-41566-1_14, 10.1128/mbio.01533-19). Codoff also detects many genomically integrated phages as exhibiting discordant codon usage to the genome-wide profile, regardless of genomic location.

Interestingly, most fungal BGCs align with the codon usage profiles of their genomic context (*new* **Extended Data Fig. 2d**). Just 129 of 6,647 (1.9%) BGCs were in the top fifth percentile for discordant codon usage suggesting they were recently acquired. However, codon usage discordance can only detect recent horizontal transfer events. This is likely why we did not detect anything unusual about the BGCs of *Neocallimastigomycota*, which are believed to be acquired from bacteria ancestrally. Swift *et al.* 2021 showed that although 63% of BGCs were estimated to be from bacterial sources, their sequence similarity to bacterial homologs were

quite low - indicating that the BGCs had already assimilated their codon usage to their new genomic contexts.

Mobile genetic elements called “Starships” have recently been characterized and reported to mediate horizontal transfer in *Pezizomycotina* (doi: 10.1093/molbev/msac109). We further assessed the kingdom-wide distribution and prevalence of tyrosine recombinases, the central “captain” gene underlying the mobilization of these elements. We found they were mostly rare outside of the *Pezizomycotina* (*new* **Extended Data Fig. 2e**).

The authors state: “Large-scale annotation efforts have also previously suggested that BGC carriage is limited in many bacterial species18 and databases cataloging the taxonomic origins of known natural products have shown that they are disproportionately sourced from just six microbial phyla19,20 (Fig. 1b).” However, I wonder to what extent the database is biased towards particular more well-studied taxa?

Thank you for your inquiry. Yes, we agree that most natural products are isolated from just six microbial phyla partly because these contain more well-studied species. Another bias is the ability to readily culture these microbes from the environment as well as the ability to conduct microbial fermentation at scale for investigating natural products. Nevertheless, it is interesting that at the bacterial order level the median BGC-ome size is closely associated with the number of characterized natural products (*see* **Fig. 1D**, first column is the number of specialized metabolites). But, as the reviewer might be noting, those orders not enriched in BGCs but have been sources for many characterized natural products fall into well studied bacterial groups (e.g., *Pseudomonadales*, *Enterobacterales*). We would be very interested in adding more details on this to our paper, but are already well over the word limit, so we plan on addressing this in subsequent papers.

Reviewer #1 (Remarks to the Author):

I am satisfied with the changes made. The article looks excellent in its current state.

Reviewer #3 (Remarks to the Author):

Overview: The authors have done a substantial number of new analyses to address my concerns/questions as the those of the other reviewers, which is commendable. I enjoyed reading their response. Overall, I feel the authors provide some very interesting results and new avenues for research thinking that are very intriguing. I also appreciate that the authors made the code and data easily available for future use, which is not always done.

Below are some minor comments/suggestions.

Thank you so much for your initial and additional feedback. Your comments have significantly improved the quality of our study for which we are grateful. We have further revised our manuscipt to address your remaining concerns and suggestions.

Note, we experienced difficulties matching up the line numbers you reference in your comments - this is a known issue with Word documents where line numbers are dynamic and could depend on how you are personally viewing the file (the degree of zoom). In our replies here, we thus include excerpts of the text as they appear in the manuscript to enable more efficient reviewer assessment.

Intro, lines 93-94. The wording could be rephrased given that plants also produce specialized metabolites and it's not just a microbial thing. Perhaps you can change specialized to 'bioactive' since you want to focus on this aspect of specialized metabolism in microbes? Also, fyi I think a lot of people associate 'microbe' with only Bacteria or Archaea and not Fungi; you might want to explicitly say both microbial prokaryotes and eukaryotes (although I know prokaryote is problematic).

Thank you for your feedback. After careful consideration, we have chosen to retain the wording of this sentence. First, pertaining to replacing "microbial life", with "*Bacteria* and *Fungi*", we think that in this specific sentence, using the replacement can be interpreted incorrectly as stating that BGCs evolved independently in these taxa, which we are not attempting to claim here. We ultimately do not think it is confusing that plants, as well as other microbes like slime molds, can make specialized metabolites, because our abstract should clearly communicate that the focus for this study is on *Bacteria* and *Fungi*. During the initial round of reviews, another reviewer had also asked about the trend in other taxa, including plants and animals. To address their request,

we had added some Discussion text based on literature review around expectations for the link between multicellularity and BGC expansions also being observed for these taxa.

We further think that specialized (or perhaps secondary) metabolite is the appropriate and standard term to use. We think this terminology will ensure the broadest comprehension of the category of metabolites that we are referring to and is used by recent reviews/studies, such as Cox 2024, *Nature Reviews Chemistry* (Title: *Engineered and total biosynthesis of fungal specialized metabolites*; DOI: 10.1038/s41570-023-00564-0) and Gavriilidou and Kautsar *et al* 2022; *Nature Microbiology* (Title: *Compendium of specialized metabolite biosynthetic diversity encoded in bacterial genomes*; DOI: 10.1038/s41564-022-01110-2). In addition to literature review, in response to feedback from Reviewer 2, we had performed a cursory analysis referenced in the Discussion (and described in greater detail in the Supplementary Notes) investigating the link of complex physiology and expanded BGCs per genome within the plant phylum of *Streptophyta*:

> Other multicellular organisms, especially plants, are known to produce various specialized metabolites. Interestingly, a re-analysis of recent biosynthetic characterizations across the plant phylum *Streptophyta* suggests that vascular land-plants also experienced an expansion in BGCs that parallels increased morphological complexity (see **Supplementary Notes**). Given this, as well as reports that other multicellular *Eukarya*, like slime molds and animals, also contain machinery for the biosynthesis of specialized metabolites, future systematic studies in the *Eukarya* might reveal our findings apply across the tree of life.

Lines 102-104. This sentence is missing references. "The natural products derived from microbial specialized metabolites represent a critical source of bioactive compounds with immense potential for drug discovery, having yielded numerous life-saving pharmaceuticals like antibiotics, immunosuppressants, and anticancer agents."

We have now added reference to the following review article at the end of this sentence:

Demain and Sanchez 2009, *The Journal of Antibiotics* (Title: Microbial drug discovery: 80 years of progress; DOI: 10.1038/ja.2008.16).

Line 142 and 172 (and elsewhere in paper). The authors should not italicize "Ascomycetes" as this is an informal term and not a valid taxonomic name. Change to Ascomycota or ascomycetes.

Thank you again for the correction, we now use "*Ascomycota"* in place of "*Ascomycetes"* as suggested. We were unable to find a second instance of ascomycetes in the manuscript; however, think that you might be referring to our use of *Actinomycetes*. We want to note that this is a recognized taxonomic class in the Genome Taxonomy Database (GTDB) so we have kept it italicized with the first letter capitalized.

Line 158. Change "productions" to "products"? "Building on the intriguing pattern of key sources of natural products being multicellular…"

Thank you for catching this typo. We have now corrected “productions” to “products” as you noted.

Line 160. The authors added this sentence to the paper “…multicellularity facilitated expansions in specialized metabolite production” but I would suggest changing “facilitated” to something that doesn’t imply causation. Perhaps “associated with”?

Thank you for your feedback. We have avoided stating causation throughout our paper, but here this text is simply stating our hypothesis. Although we avoid stating causation throughout our manuscript, we would like restate the reasons why we think multicellularity facilitated biosynthetic expansions (as noted to reviewer 2 in our prior response):

a) Carriage of a larger number of BGCs has a cost, even just for replication of the resulting larger genome. The benefits from carriage of a large numbers of BGCs is clear for multicellular organisms (i.e., reduced conflict associated with public goods).

b) Unlike unicellular organisms, multicellular organisms have an expanded division of labor of cells. This is predicted to increase the benefits conferred from a higher diversity of BGCs, such as parallel production of different antibiotics where different cells of a multicellular organism specialize in the production of individual ones.

c) It is increasingly clear that specialized metabolites play important roles mediating ecological interactions for microorganisms. For each phylum of *Cyanobacteriota*, *Myxococcota,* and *Actinomycetota* we show positive associations between counts for known protein markers of multicellularity and sporulation with BGC-ome size (**Fig. 1e-g**). This supports a view that an important benefit associated with an increased degree of cooperation is an increase in diversity in specialized metabolites and potentially ecological breadth.

Line 179. I would suggest that you cite the paper by Richards & Tablot if you are taking about public goods. Doi:10.1038/nrmicro3108

Thank you for sharing that reference, it is a very engaging Review article and certainly relevant to our study. However, we think that its primary focus on HGT of public good encoding genes is not appropriate to cite when we discuss secondary metabolites as public goods. We were previously citing a Review article that discussed how some specialized metabolites correspond to public goods; however, it was largely focused on the bacterial genus *Salinospora*. We have thus added reference to another study on fungal siderophores by:

Haas, Eisendle, and Turgeon 2008, *Annual Review of Phytopathology* (Tilte: Siderophores in Fungal Physiology and Virulence, DOI: 10.1146/annurev.phyto.45.062806.094338)

This is because fungal siderophores, secreted secondary metabolites used to scavenge for iron, are often thought of as quintessential “public goods” of microbial populations and to make up for

the lack of regard for a reference focused on a fungal public good as we believe you were suggesting.

Line 275. This sentence is a little unclear, perhaps change as follows. "The only family found to be associated with BGC-ome size in four or more of the five phyla was glycosyl hydrolase family 19 that encode chitinases, which are rare in microbes and hypothesized to have been horizontally acquired in Streptomyces from plants26."

Thank you for drawing our attention to this sentence. We agree that it was a long and poorly written, so we have now split it up and rewritten the second part to improve clarity:

> The only family found to be associated with BGC-ome size in four or more of the five phyla was glycosyl hydrolase family 19, chitinases that are rare in microbes. Within *Streptomyces*, homologs of this enzyme family are hypothesized to have been acquired via HGT from plants.

Line 413. I would suggest rephrasing this sentence. Ascomycota aren't the "hallmark", the products they produce are, right? Perhaps this instead? "A trove of fungal natural products are produced by Ascomycota, including critically important drugs such as the antibiotic penicillin, the immunosuppressant cyclosporine, and the cholesterol-lowering statins."

This sentence is also missing a reference.

Thank you for your feedback. We have now corrected for this grammatical issue and substantially updated the opening to this section to open with a focus on the full kingdom as you suggest in the next comment.

Lines 415-418. I think the text below clarifies what you mean by this, but this sentence is problematic for me because your premise is that multicellularity is associated with expanded BGCs, yet multicellularity arose before the Ascomycota. Thus, why are BGCs only expanded in certain lineages? You get to this later, but it takes a while.

"Like the other major microbial source of natural product drugs, the Actinomycetota, the *Ascomycota* contain genera enriched in BGCs and have the ability to form complex cooperative units in the form of mycelia, suggesting that enrichment in biosynthetic potential in fungi is associated with multicellularity."

I see why you want to start with *Ascomycota* since they are well-known for producing bioactive molecules, but I think it would make more sense just to start this section with Fungi "To systematically investigate BGC enrichment across Kingdom Fungi, which contain many taxa known for their production of bioactive metabolites, …" and then narrow the focus down to phyla and why only certain lineages are expanded.

Thank you for this feedback, we completely agree with you - it is better to start off with reference to the full kingdom to avoid confusion on the section's scope and focus. Part of the reason why we referenced *Ascomycota* is because of the analysis of characterized BGC records in MIBiG shown in Fig. 1b. We have updated the area of the text you reference as such:

> In addition to the domain *Bacteria*, the eukaryotic kingdom *Fungi* has historically served as another key reservoir of natural products and the source of several important drugs including the antibiotic penicillin, the immunosuppressant cyclosporine, and the cholesterol-lowering statins. The kingdom also contains diverse species capable of forming complex cooperative units, such as some *Ascomycota*, which can form mycelia resembling structures produced by some bacterial *Actinomycetes*. Similar to filamentous *Actinomycetes*, filamentous *Ascomycota* are also renowned for their diverse chemical repertoire, suggesting that enrichment in biosynthetic potential in *Fungi* is also associated with multicellularity.

Line 419 and elsewhere. Fungi, not fungi (since you are referring to the Kingdom) rather than using it as a common noun or informal term (i.e., fungi often = Fungi + Oomycota).

Thank you for the correction. We now capitalize the first letter of *Fungi* and italicize the word in this and most other instances, but not all of them. Given that we use *Fungi* to describe the premise of our analysis early on in the fungal Results section, we think that it should now be clear to readers that when we use the informal term "fungi" in this manuscript, we use it to refer to only the kingdom *Fungi*, not inclusive of *Oomycota*.

Line 441. Ditto. "Fonticula, a close outgroup of Fungi."

Thank you again. Yes, we agree this is a critical instance where it makes sense to refer to the formal kingdom by name, *Fungi*. We now do so, capitalizing the first letter and italicizing the word.

Line 442. Will non-phylogenetic biologists know what symplesiomorphic means? :-)

Thank you for your feedback, we agree that it is important for the text to use accessible language designed for a broader audience range. We have updated the first instance of the term's usage to include a brief definition:

> These trends and the lack of BGCs in *Fonticula*, a close outgroup for *Fungi*, support the conclusion that limited biosynthetic potential is symplesiomorphic in *Fungi* - that is, limited biosynthetic potential is likely an ancestral characteristic of the kingdom.

Line 454. What subkingdom are you referring to? Dikarya? Please clarify.

Thank you for catching this typo. We have now corrected it:

> Although we found limited genomic indications of recent HGT of BGCs in fungi, over the course of fungal evolution, HGT has likely contributed extensively to the biosynthetic diversity of the ~~subkingdom~~**fungal kingdom**.

Line 462. Specify what classes are within the "monophyletic subdivision" you are referring to? It is hard to see in the Figure.

Thank you for the great advice. This will indeed allow readers familiar with fungal taxonomy to better understand the boundaries of the BGC-enriched clade within *Pezizomycotina*. We are limited in the main text by word limits but now describe which classes belong to this clade in the main Methods text and Fig. 3's legend. In brief, the BGC-enriched clade within *Pezizomycotina* consists of at least seven classes: *Sordariomycetes*, *Dothideomycetes*, *Eurotiomycetes*, *Leotiomycetes*, *Lecanoromycetes*, *Xylonomycetes*, and *Arthoniomycetes*.

Line 502. I feel like the wording of this may be confusing to people as it makes it sound like only Basidiomycota form dikaryons. This goes back to my suggestion that you define Dikarya for non-fungal readers and you can state that the primary synapomorphy of the subkingdom Dikarya (Ascomycota and Basidiomycota) is the formation of dikaryons, cells with two nuclei containing separate genomes that are physically associated and co-replicate.

Thank you for the feedback, we have updated the wording of this paragraph to improve clarity as you suggest.

First, we simplified this sentence:

> As with the *Pezizomycotina*, the *Agaricomycetes* represent the other major fungal group well known for possessing complex multicellularity.

In the beginning of the paragraph, we also do a better job at introducing the subkingdom of Dikarya, now providing a definition of dikaryon earlier on for readers:

> Next, in our evaluation of the evolution of biosynthesis in *Fungi*, we find support for two independent origins of BGC enrichment occurring separately within the subkingdom of *Dikarya*, which are able to form dikaryons, cells with two nuclei containing separate genomes that are physically associated and co-replicate. Many members of this subkingdom exhibit physiological complexity and form macroscopic sexual fruiting bodies, such as cup-shaped ascocarps and mushroom basidiocarp structures.

Line 574. If you are referring to the Pezizomycotina and Agaricomycotina, use 'subphyla' instead of "classes". Adding a comma after subphyla could help clarify the sentence's meaning too.

Thank you for catching this typo. We updated the use of "classes" in the following sentence to "subphyla":

A monophyletic clade within *Pezizomycotina* enriched for BGCs was identified through visual examination and validated to include taxonomic ~~classes~~**subphyla** described as BGC enriched in prior studies.

Additional changes:

We made a couple changes to the main text or figures that were not related to reviewer or editor requests:

1.) Removal of text in Discussion around findings from Xu *et al* 2025: We removed the following sentence from our Discussion speculating the relationship between HI-associated proteins and chromosome partitioning referencing a recent study published in Science last year.

> 'Recent research has also shown that multiple species of *Pezizomycotina* have uneven distributions of chromosomes across nuclei when in a haploid state and HI-associated proteins might relate to this atypical biological phenomenon.'

The validity and interpretation of the findings by Xu *et al* 2025, *Science* (Title: Distribution of haploid chromosomes into separate nuclei in two pathogenic fungi; DOI: 10.1126/science.abn7811) has recently been called into question by Zheng, van Kan, and Auxier 2026 in a pre-print (Title: Multiple nuclei means multiple chromosome sets in *Botrytis cinerea* and *Neurospora crassa*; DOI: 10.64898/2026.03.14.711691v1), so we felt it best to avoid speculating around this in our paper.

2.) Removal of panel b in Fig. 1: After further consideration, we thought it was best to remove Fig. 1b, allowing us to enlarge the other panels of the figure, and these results are largely congruent with the MIBiG study, which we reference here.