nature portfolio

Peer Review File

SpotSweeper: spatially-aware quality control for spatial transcriptomics

Corresponding Author: Dr Boyi Guo

This file contains all editorial decision letters in order by version, followed by all author rebuttals in order by version.

Version 0:

Decision Letter:

23rd Jul 2024

Dear Dr Guo,

Your Article, "SpotSweeper: spatially-aware quality control for spatial transcriptomics", has now been seen by three reviewers. As you will see from their comments below, although the reviewers find your work of considerable potential interest, they have raised a number of concerns. We are interested in the possibility of publishing your paper in Nature Methods, but would like to consider your response to these concerns before we reach a final decision on publication.

We therefore invite you to revise your manuscript to address these concerns. As you will see, all three reviewers were concerned about applicability beyond Visium data. We do require that general applicability to other modalities be shown to give the paper full consideration upon resubmission.

The referees also asked about downstream benefits of the technology on specific analysis questions, such as domain detection or cell type deconvolution. While we don't expect an exhaustive survey of all possible downstream analyses, highlighting the importance of this QC step more broadly could strengthen the paper. Similarly, ref 1 wanted you to test more than BayesSpace on downstream applications. We think extensive benchmarking of downstream tools goes beyond the scope of this work, and are happy to discuss how you might modify the paper in response to this concern if you have any questions.

We are committed to providing a fair and constructive peer-review process. 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.

When revising your paper:

* include a point-by-point response to the reviewers and to any editorial suggestions

* please underline/highlight any additions to the text or areas with other significant changes to facilitate review of the revised manuscript

* address the points listed described below to conform to our open science requirements

* ensure it complies with our general format requirements as set out in our guide to authors at www.nature.com/naturemethods

* resubmit all the necessary files electronically by using the link below to access your home page

Link Redacted

Note: This URL links to your confidential home page and associated information about manuscripts you may have submitted, or that you are reviewing for us. If you wish to forward this email to co-authors, please delete the link to your homepage.

We hope to receive your revised paper within three months. If you cannot send it within this time, please let us know. In this event, we will still be happy to reconsider your paper at a later date so long as nothing similar has been accepted for publication at Nature Methods or published elsewhere.

OPEN SCIENCE REQUIREMENTS

REPORTING SUMMARY AND EDITORIAL POLICY CHECKLISTS
When revising your manuscript, please update your reporting summary and editorial policy checklists.

Reporting summary: https://www.nature.com/documents/nr-reporting-summary.zip
Editorial policy checklist: https://www.nature.com/documents/nr-editorial-policy-checklist.zip

If your paper includes custom software, we also ask you to complete a supplemental reporting summary.

Software supplement: https://www.nature.com/documents/nr-software-policy.pdf

Please submit these with your revised manuscript. They will be available to reviewers to aid in their evaluation if the paper is re-reviewed. If you have any questions about the checklist, please see http://www.nature.com/authors/policies/availability.html or contact me.

Please note that these forms are dynamic 'smart pdfs' and must therefore be downloaded and completed in Adobe Reader. We will then flatten them for ease of use by the reviewers. If you would like to reference the guidance text as you complete the template, please access these flattened versions at http://www.nature.com/authors/policies/availability.html.

DATA AVAILABILITY
We strongly encourage you to deposit all new data associated with the paper in a persistent repository where they can be freely and enduringly accessed. We recommend submitting the data to discipline-specific and community-recognized repositories; a list of repositories is provided here: http://www.nature.com/sdata/policies/repositories

All novel DNA and RNA sequencing data, protein sequences, genetic polymorphisms, linked genotype and phenotype data, gene expression data, macromolecular structures, and proteomics data must be deposited in a publicly accessible database, and accession codes and associated hyperlinks must be provided in the "Data Availability" section.

Refer to our data policies here: https://www.nature.com/nature-research/editorial-policies/reporting-standards#availability-of-data

To further increase transparency, we encourage you to provide, in tabular form, the data underlying the graphical representations used in your figures. This is in addition to our data-deposition policy for specific types of experiments and large datasets. For readers, the source data will be made accessible directly from the figure legend. Spreadsheets can be submitted in .xls, .xlsx or .csv formats. Only one (1) file per figure is permitted: thus if there is a multi-paneled figure the source data for each panel should be clearly labeled in the csv/Excel file; alternately the data for a figure can be included in multiple, clearly labeled sheets in an Excel file. File sizes of up to 30 MB are permitted. When submitting source data files with your manuscript please select the Source Data file type and use the Title field in the File Description tab to indicate which figure the source data pertains to.

Please include a "Data availability" subsection in the Online Methods. This section should inform readers about the availability of the data used to support the conclusions of your study, including accession codes to public repositories, references to source data that may be 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", describing which data is available upon request and mentioning any restrictions on availability. If DOIs are provided, please include 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

CODE AVAILABILITY
Please include a "Code Availability" subsection in the Online Methods which details how your custom code is made available. Only in rare cases (where code is not central to the main conclusions of the paper) is the statement "available upon request" allowed (and reasons should be specified).

We request that you deposit code in a DOI-minting repository such as Zenodo, Gigantum or Code Ocean and cite the DOI in the Reference list. We also request that you use code versioning and provide a license.

For more information on our code sharing policy and requirements, please see:
https://www.nature.com/nature-research/editorial-policies/reporting-standards#availability-of-computer-code

MATERIALS AVAILABILITY

As a condition of publication in Nature Methods, authors are required to make unique materials promptly available to others without undue qualifications.

Authors reporting new chemical compounds must provide chemical structure, synthesis and characterization details. Authors reporting mutant strains and cell lines are strongly encouraged to use established public repositories.

More details about our materials availability policy can be found at https://www.nature.com/nature-portfolio/editorial-policies/reporting-standards#availability-of-materials

ORCID
Nature Methods 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.

Please do not hesitate to contact me if you have any questions or would like to discuss these revisions further. We look forward to seeing the revised manuscript and thank you for the opportunity to consider your work.

Sincerely,
Rita

Rita Strack, Ph.D.
Senior Editor
Nature Methods

Reviewers' Comments:

Reviewer #1 (Remarks to the Author):

The current limitations of spatial transcriptomics QC include challenges in handling data noise and technical biases, constraints in spatial resolution impacting accuracy, difficulties in standardization and normalization across samples, and complexities arising from spatial heterogeneity within tissues and cells. Addressing these issues is crucial for advancing the reliability and interpretability of spatial transcriptomics data analyses.
The manuscript presents a novel approach by proposing quality control (QC) metrics and a computational pipeline aimed at identifying low-quality observations and regional artifacts resulting from sample processing errors in Single-cell RNA sequencing (SRT) data. This methodology addresses a significant issue in the field of SRT data analysis, as ensuring data quality is crucial for accurate biological interpretation. The proposed approach is meaningful and holds potential to improve the reliability and reproducibility of SRT data analyses.However, it requires major revisions to improve the clarity, comprehensiveness, and robustness of the proposed methods and their validation.Here are some questions and suggestions for the authors.

1. While the proposed method is promising, the manuscript currently lacks sufficient detail in explaining the QC metrics and the computational pipeline.

A more comprehensive description is needed to enable readers to fully understand and replicate the approach. This includes a step-by-step explanation of the metrics, the computational steps involved, and any assumptions or parameters used.

The manuscript utilizes K-nearest neighbors (KNN) clustering for downstream quality control (QC) purposes. While this approach is described, it would be beneficial to elaborate on the rationale behind choosing KNN clustering over alternative clustering methods commonly used in single-cell RNA sequencing (SRT) data analysis, such as hierarchical clustering. Providing a comparison or justification for selecting KNN clustering would enhance the manuscript's clarity and help readers understand why this method was deemed most suitable for the proposed QC tasks. Additionally, discussing the advantages and potential limitations of KNN clustering in this context would provide a more comprehensive evaluation of the method's appropriateness for the study's objectives.
To further strengthen the article, it would be advantageous to include comprehensive performance evaluations of SpotSweeper. This could encompass comparative analyses with existing QC tools, assessing not only the accuracy and effectiveness of SpotSweeper but also its stability and robustness under diverse experimental conditions and across various categories of organs. Additionally, it would be insightful to investigate the computational efficiency and memory footprint of the QC process. A thorough analysis in this regard could provide valuable insights into the practicality and scalability of SpotSweeper for spatial transcriptome studies.

2. The validation of the proposed QC metrics and computational pipeline appears limited in the current manuscript. It is essential to include more comprehensive benchmarking against existing methods to convincingly demonstrate the effectiveness and advantages of the new approach. In the field of single cell RNA-seq, there are various mature Quality Control assessments. Spotsweeper can compare with these methods or set baseline for benchmark, and highlights the advantages of Spotsweeper than single cells RNAseq data Quality Control.This should involve testing on multiple datasets and scenarios to ensure robustness and generalizability. Additionally, the comparison methods used in the manuscript are few, which limits the scope of the validation.

To substantiate the superiority of the QC method, it is recommended that the authors expand the discussion to include downstream applications. Specifically, the authors could present comparative studies that evaluate the impact of applying the QC method discussed in this paper versus no QC or alternative single-cell QC methods on the outcomes of subsequent analytical steps.

3. The manuscript primarily focuses on the 10X Genomics platform(such as 10x Visium human DLPFC datasets and Mouse Coronal Brain dataset), which is a limitation. To increase the general applicability and robustness of the proposed method, it would be beneficial to test and validate it on a variety of platforms. This will demonstrate the versatility and broader utility of the QC metrics and computational pipeline across different SRT technologies. Besides, many analysis of Spotsweeper only evaluates on single 10X dataset and obtains numerical results, without additional biology discovery. We look forward to new biological discoveries using Spotsweeper in some scenes.
For example,the manuscript could benefit from incorporating targeted case studies that demonstrate the efficacy of SpotSweeper across a variety of SRT data types and organ systems. These case studies should illustrate how SpotSweeper's application leads to tangible improvements in data quality and subsequent analysis.

4. The manuscript only employs BayesSpace method for identifying spatially domin, which is limited. To strengthen the manuscript, it is recommended to explore and compare additional methods for detecting spatially dominant features. This will provide a more comprehensive assessment and validate the effectiveness of the chosen method.

5. The manuscript mentions that the QC metrics can systematically filter out observations with low total Unique Molecular Identifiers (UMIs). However, more detail is needed on how this filtering is achieved and its impact on the overall data quality. Additionally, it would be helpful to include a discussion on the potential consequences of removing low UMI observations, such as the risk of losing biologically relevant information

6. Spotsweeper is limited to Spot Level, and even testing at 10x Visium technology. With the development of spatial sequencing technology, the development of spatial single-cell level technology (Such as Starmap and Baristaseq, etc.) is now developing. Spotsweeper is expected to be more scalability and generalization.

Besides, some minor advice for the authors is as following,
1.It would be better if the Figure.1 add some downstream tasks examples.
2.It would be better if adding an explanation of parameters selection in your algorithm and how to select appropriate parameters.
3.It is recommended to visualize the resultant plots of all slices of DLPFC data in the visualization of sup figure.
4.It is recommended to specify whether the experiment was repeated and how many times it was repeated in order to assess the consistency and reliability of the results.
5.It is recommended to provide an estimate of the computing resources needed to run SpotSweeper, including memory, processor, and time consumption.
Given these points, the manuscript requires major revisions to address the issues mentioned above. I look forward to seeing the authors' revisions and believe that with these improvements, the work will make a valuable contribution to the field.
Thank you very much.

Reviewer #1 (Remarks on code availability):

According to the code provided by the author, we have conducted testing using two datasets. The help file and the annotation in the code are not detailed enough, such as the version information of R. Since I am using the latest version of R, the author's code did not report any errors during installation, but I am unable to call the relevant functions when performing visualizations. The author is requested to provide additional information.

Reviewer #2 (Remarks to the Author):

The authors set out to develop a framework to enable spatially aware quality control. The approach leverages the spatial dependencies imposed in spatial data to identify local and region-based low-quality measurements. The proposed approach, SpotSweeper, is a significant advancement on the current approach to perform quality checks in spatial transcriptomics datasets and holds the potential to form part of every standard spatial transcriptomics analysis. The authors innovatively generalise the analysis such that the tool works with any quality metric and allows multiscale analysis of quality as well. I commend the authors on developing such an impressive quality control tool and on demonstrating its ability extensively in this manuscript. Through analysis of various Visium datasets, the authors identified a set of consistently poorly performing barcode sequences. This result will help improve array barcoding for 10x Visium making SpotSweeper an important tool to discover and avoid problematic barcode sequences in newly developed technologies.

I have no major concerns with the manuscript but do feel that some additional work is needed to generalise the tool to other spot-based technologies. It is an important tool for all spot-based technologies and extending support to the 10x Visium HD and BGI STOmics technologies which have a regular square grid will greatly enhance its potential. Besides this one request from the authors, I have a few minor comments to improve clarity of text in the manuscript.

Major comments:

1. The proposed method can be applied to spot-based sequencing technologies. The authors have reported on a comprehensive analysis of 10x Visium data. However, other spot-based technologies such as the BGI stereo-seq (STOmics) and the 10x Visium HD have not been studied. When the density of spots is increased, as well as their resolution, it would be good to get recommendations on how SpotSweeper should be applied. I suspect there to be more outlier spots in these settings so it would be useful to get guidance on what to expect from these datasets. Publicly available data exist from these technologies so they should be added to the comparisons performed in the paper.

2. Consequently, it would help to have implementations of SpotSweeper where the grid type is not enforced to be a hexagonal tessellation, but a simple square grid.

3. The dataset used to produce figure 4H seems to have tissue folding. This is an artefact that could affect downstream analyses and is visible in the library size. Is it possible to identify these effects using SpotSweeper? If so, it would be worth demonstrating this use case. This is an aspect of quality control, and it would be great if SpotSweeper is able to handle this issue as well. If not, the authors can comment on this issue being present in the older visium samples where CytAssist was not available, and mention that this is an outstanding problem that may need alternative solutions.

Minor comments:

1. The more common approach to defining outliers is using Tukey's approach (>q0.75 + k * IQR and <q0.25 – k * IQR, where k is usually 1.5) that is often used do annotate outliers in boxplots. The authors use an alternative approach based on a modified z-score approach. I would be interested in knowing the reasoning behind the chosen outlier labelling approach.

2. The authors have detailed the computation of z-scores in the first result section, however, they should add a sentence stating the approach to determining outlier thresholds. Normally, this would be some quantile of the z-score, but I see that from the methods section that the proposed approach is to use the z-scores directly. There is nothing wrong with this, but it is better to report it more clearly. I would also advise reporting the quantile matching the z-score cut-off of -3 that is used to demonstrate how strong the outliers are.

3. Add equation numbering

4. Page 3, line 71: Should it state "Then, we calculate a robust z-score using all spots in the neighbourhood:" rather than "for all spots"?

5. I recommend matching the full scaling factor used in the equation for the z-score and that in figure 1 (and possibly annotating it as the q0.75 of a standard normal). It is important to state how this value is obtained. From the referred textbook, I see that the scaling factor is obtained by computing the expectation of the MAD in terms of the standard deviation given a large n which evaluates to the 75 percentile of the standard normal. Writing it as such will enhance clarity.

6. Page 14, line 284 should be < -3 z-scores.

7. Robust linear regression is used to study variance of metrics at multiple scales. I think readers would be interested in understanding the mean variance relationship across scales. Would the authors be able to provide supplementary figures showing examples of this?

8. It would be good to have a function that annotates the 6 problematic barcodes in Visium datasets so that users can discard these with ease and not have to replicate the outlier analysis (the results of which may vary given their threshold selection).

Reviewer #2 (Remarks on code availability):

I have reviewed the code on Bioconductor and it was easy to replicate the analysis. The code itself has already undergone peer-review as part of the submission process. The automatic build and testing systems show that the code is functional.

Reviewer #3 (Remarks to the Author):

This manuscript presents approaches for spatially-aware quality control of spot-based SRT technologies, primarily 10X Visium. They illustrate the method using DLPFC datasets (two studies; Maynard et al and Huuki-Myers et al). In doing so, they also discuss the detection of "dryspot" and "hangnail" artifacts. The approach for using spatial neighborhoods (K-NN) overcomes the issue of confounding of biological information with QC metrics.

However, it is unclear how researchers would select the neighborhood size, how the methods would work at boundary areas of domains (spots in two different domains), and the impact of their QC has on aspects other than domain detection, such as phenotyping/deconvolution methods.

The methods proposed are also not applicable to technologies other than Visium and are not testing in single-cell SRT technologies, which is the future of SRT.

Lastly, it would have been nice to see the QC of more complex/complicated tissues in spatial organization/structure as DLPFC tissues are highly organized (e.g., how would SpotSweeper work on cancer tumor tissue?).

In terms of the package, SpotSweeper is available on Bioconductor and GitHub and has a MIT license. They have a vignette on how to use the package functions and plotting results. They also have a documentation page for the package.

Reviewer #3 (Remarks on code availability):

The package is on GitHub and Bioconductor. They have instructions on how to install the package and examples of how to use the package for SRT QC. The package and code are very usable.

Version 1:

Decision Letter:

6th Feb 2025

Dear Boyi,

Thank you for your letter detailing how you would respond to the reviewer concerns regarding your Article, "SpotSweeper: spatially-aware quality control for spatial transcriptomics". We have decided to invite you to revise your manuscript based on additional referee feedback we got on your proposed responses.

The reviewers think that point 3 raised by reviewer 1 was adequately addressed in your response. Please update the manuscript accordingly. However, they think point 4 needs more attention. They ask that you justify parameter choices for an imaging-based single-cell technology.

We therefore ask you to revise to include these additional data.

Please do not hesitate to contact me if you have any questions or would like to discuss these revisions further.

When revising your paper:

* include a point-by-point response to the reviewers and to any editorial suggestions

* please underline/highlight any additions to the text or areas with other significant changes to facilitate review of the revised manuscript

* address the points listed described below to conform to our open science requirements

* ensure it complies with our general format requirements as set out in our guide to authors at www.nature.com/naturemethods

* resubmit all the necessary files electronically by using the link below to access your home page

Link Redacted

Note: This URL links to your confidential home page and associated information about manuscripts you may have submitted, or that you are reviewing for us. If you wish to forward this email to co-authors, please delete the link to your homepage.

We hope to receive your revised paper within six weeks. If you cannot send it within this time, please let us know. In this event, we will still be happy to reconsider your paper at a later date so long as nothing similar has been accepted for publication at Nature Methods or published elsewhere.

OPEN SCIENCE REQUIREMENTS

REPORTING SUMMARY AND EDITORIAL POLICY CHECKLISTS
When revising your manuscript, please update your reporting summary and editorial policy checklists.

Reporting summary: https://www.nature.com/documents/nr-reporting-summary.zip
Editorial policy checklist: https://www.nature.com/documents/nr-editorial-policy-checklist.zip

If your paper includes custom software, we also ask you to complete a supplemental reporting summary.

Software supplement: https://www.nature.com/documents/nr-software-policy.pdf

Please submit these with your revised manuscript. They will be available to reviewers to aid in their evaluation if the paper is re-reviewed. If you have any questions about the checklist, please see http://www.nature.com/authors/policies/availability.html or contact me.

Please note that these forms are dynamic 'smart pdfs' and must therefore be downloaded and completed in Adobe Reader. We will then flatten them for ease of use by the reviewers. If you would like to reference the guidance text as you complete the template, please access these flattened versions at http://www.nature.com/authors/policies/availability.html.

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.

DATA AVAILABILITY
Please include a "Data availability" subsection in the Online Methods. This section should inform readers about the availability of the data used to support the conclusions of your study, including accession codes to public repositories, references to source data that may be 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", describing which data is available upon request and mentioning any restrictions on availability. If DOIs are provided, please include 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

CODE AVAILABILITY
Please include a "Code Availability" subsection in the Online Methods which details how your custom code is made available. Only in rare cases (where code is not central to the main conclusions of the paper) is the statement "available upon request" allowed (and reasons should be specified).

We request that you deposit code in a DOI-minting repository such as Zenodo, Gigantum or Code Ocean and cite the DOI in the Reference list. We also request that you use code versioning and provide a license.

For more information on our code sharing policy and requirements, please see:
https://www.nature.com/nature-research/editorial-policies/reporting-standards#availability-of-computer-code

MATERIALS AVAILABILITY
As a condition of publication in Nature Methods, authors are required to make unique materials promptly available to others without undue qualifications.

Authors reporting new chemical compounds must provide chemical structure, synthesis and characterization details. Authors reporting mutant strains and cell lines are strongly encouraged to use established public repositories.

More details about our materials availability policy can be found at https://www.nature.com/nature-portfolio/editorial-policies/reporting-standards#availability-of-materials

ORCID
Nature Methods 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.

Please do not hesitate to contact me if you have any questions or would like to discuss these revisions further. We look forward to seeing the revised manuscript and thank you for the opportunity to consider your work.

Sincerely,
Rita

Rita Strack, Ph.D.
Senior Editor
Nature Methods

Reviewers' Comments:

Reviewer #1 (Remarks to the Author):

The revision for the SpotSweeper article mainly added some datasets and conducted more validation experiments. Overall while SpotSweeper presents a technically sound quality control tool with incremental improvements, the revised version of the manuscript right now does not meet the scientific innovation, comprehensive validation, and methodological rigor required for publication in Nature Methods.

My main concerns are as follows.

1.Insufficient Scientific Innovation

The manuscript remains centered on technical validation without demonstrating a truly novel biological discovery. There is no indication of how this tool advances key biological questions or reveals previously unknown insights. Authors emphasize the tool's ability to identify regional artifacts (e.g., dryspots and hangnails) and its improvement over existing QC methods. However, no convincing biological discoveries were presented as evidence of SpotSweeper's practical scientific impact.

2.Lack potential biological applications
Although the authors mention the potential biological applications of SpotSweeper in detecting high-quality spots and artifacts, there is a lack of specific application examples. We strongly recommend that the authors provide one or more concrete application cases to demonstrate how the results identified by SpotSweeper can be used for downstream analysis and elaborate on the biological significance of these analyses. This will not only prove the practicality of SpotSweeper but also enhance readers' understanding of the scope of the method's application.

3.Incomplete Validation

The validation of SpotSweeper remains limited to 10x Visium and partial datasets from VisiumHD and Xenium platforms. The method was not validated on key single-cell spatial technologies such as MERFISH and STARmap , which are essential benchmarks in the field. Authors only added validations on VisiumHD, Xenium, and other 10x datasets. They argued that SpotSweeper generalizes well across spot-based platforms. However, no experiments or benchmarks were conducted for MERFISH or STARmap, and validation remains absent for high-resolution, single-cell data. We explicitly pointed out this limitation, but the authors did not provide data or evidence to address it.

4.Methodological Transparency and Parameter Selection

The justification for parameter choices (e.g., neighborhood size for KNN and multiscale parameters) remains unclear. Systematic comparisons or optimization processes are absent, leaving the method's robustness questionable across varying resolutions and datasets. Authors provided explanations for parameter defaults (e.g., using third-order neighbors) and included a visualization (Figure S7) of parameter choices for hexagonal and square grids. However, they did not provide systematic benchmarks or experimental analyses to justify parameter settings.

Minor comments：
1、The authors mention the selection of parameter k in the manuscript but do not adequately explain its specific impact on the results. We suggest that the authors discuss in detail the criteria for the selection of k, including how the value of k affects the performance and accuracy of SpotSweeper. Additionally, if the value of k has been adjusted and optimized for different datasets, the authors should provide the rationale and a detailed description of the optimization process. This will help readers understand the importance of parameter selection on algorithm performance and increase the transparency and reproducibility of the method.

2、The manuscript uses specific data labels to evaluate the performance of SpotSweeper but does not detail the source and reliability of these labels. We recommend that the authors describe in detail the process of obtaining data labels, including the generation, validation, and quality control measures of the labels. If possible, the authors should also provide additional data or analysis to support the accuracy and reliability of these labels, which is crucial for ensuring the validity of the research results.

Reviewer #2 (Remarks to the Author):

The authors have addressed all my concerns and have made substantial imporvements to the manuscript by including additional datasets, methodology, and software features. I recommend accepting this manuscript for publication.

Reviewer #2 (Remarks on code availability):

As previously stated, the code has been peer reviewed independently and is of high quality.

Reviewer #3 (Remarks to the Author):

The revisions to the manuscript pointed out in my review and the two other reviews were substantial. In particular, they added more datasets and technologies to the benchmarking, including different technology types beyond 10x Visium. They also outlined in more detail the methods used for QC. They also added more methods to testing the downstream results before and after QC (i.e., more domain detection methods). Overall, this is a much-improved manuscript and the tool SpotSweeper will aid researchers in the completion of SRT studies.

Version 2:

Decision Letter:

Our ref: NMETH-A56712B

12th Mar 2025

Dear Boyi,

Thank you for submitting your revised manuscript "SpotSweeper: spatially-aware quality control for spatial transcriptomics" (NMETH-A56712B). It has now been seen by the 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 Methods, 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 within two weeks or so. Please do not upload the final materials and make any revisions until you receive this additional information from us.

TRANSPARENT PEER REVIEW
Nature Methods offers a transparent peer review option for new original research manuscripts submitted from 17th February 2021. We encourage increased transparency in peer review by publishing the reviewer comments, author rebuttal letters and editorial decision letters if the authors agree. Such peer review material is made available as a supplementary peer review file. **Please state in the cover letter 'I wish to participate in transparent peer review' if you want to opt in, or 'I do not wish to participate in transparent peer review' if you don't.** Failure to state your preference will result in delays in accepting your manuscript for publication.

Please note: we allow redactions to authors' rebuttal and reviewer comments in the interest of confidentiality. If you are concerned about the release of confidential data, please let us know specifically what information you would like to have removed. Please note that we cannot incorporate redactions for any other reasons. Reviewer names will be published in the peer review files if the reviewer signed the comments to authors, or if reviewers explicitly agree to release their name. For more information, please refer to our FAQ page.

ORCID
IMPORTANT: Non-corresponding authors do not have to link their ORCIDs but are encouraged to do so. Please note that it will not be possible to add/modify ORCIDs at proof. Thus, please let your co-authors know that if they wish to have their ORCID added to the paper they must follow the procedure described in the following link prior to acceptance: https://www.springernature.com/gp/researchers/orcid/orcid-for-nature-research

Thank you again for your interest in Nature Methods. Please do not hesitate to contact me if you have any questions. We will be in touch again soon.

Sincerely,
Rita

Rita Strack, Ph.D.
Senior Editor
Nature Methods

Reviewer #3 (Remarks to the Author):

The authors have answered adequately the remaining comments.

Version 3:

Decision Letter:

18th Apr 2025

Dear Boyi,

I am pleased to inform you that your Article, "SpotSweeper: spatially-aware quality control for spatial transcriptomics", has now been accepted for publication in Nature Methods. The received and accepted dates will be June 6, 2024 and April 18, 2025. This note is intended to let you know what to expect from us over the next month or so, and to let you know where to address any further questions.

Over the next few weeks, your paper will be copyedited to ensure that it conforms to Nature Methods style. 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. It is extremely important that you let us know now whether you will be difficult to contact over the next month. If this is the case, we ask that you send us the contact information (email, phone and fax) of someone who will be able to check the proofs and deal with any last-minute problems.

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.

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) then you should select the gold OA route, and we will direct you to the compliant route where possible. For authors selecting the subscription publication route, the journal's standard licensing terms will need to be accepted, including self-archiving policies. Those licensing terms will supersede any other terms that the author or any third party may assert apply to any version of the manuscript.

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

If you have posted a preprint on any preprint server, please ensure that the preprint details are updated with a publication reference, including the DOI and a URL to the published version of the article on the journal website.

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

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.

If you are active on Twitter/X or Bluesky, please e-mail me your and your coauthors' handles so that we may tag you when the paper is published.

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.

Please note that you and any of your coauthors will be able to order reprints and single copies of the issue containing your article through Nature Portfolio's reprint website, which is located at http://www.nature.com/reprints/author-reprints.html. If there are any questions about reprints please send an email to author-reprints@nature.com and someone will assist you.

Please feel free to contact me if you have questions about any of these points.

Best regards,
Rita

Rita Strack, Ph.D.
Senior Editor

Totty et al – *Author responses to initial comments*

We extend our sincere thanks to all 3 reviewers for their strong interest in our work, their constructive comments, and their helpful suggestions to improve the manuscript.

> Reviewer #1: "*This methodology addresses a significant issue in the field of SRT data analysis, as ensuring data quality is crucial for accurate biological interpretation. The proposed approach is meaningful and holds potential to improve the reliability and reproducibility of SRT data analyses*."
>
> Reviewer #2: "… *SpotSweeper, is a significant advancement on the current approach to perform quality checks in spatial transcriptomics datasets and holds the potential to form part of every standard spatial transcriptomics analysis. The authors innovatively generalise the analysis such that the tool works with any quality metric and allows multiscale analysis of quality as well. I commend the authors on developing such an impressive quality control tool and on demonstrating its ability extensively in this manuscript*."
>
> Reviewer #3: "*The approach for using spatial neighborhoods (K-NN) overcomes the issue of confounding of biological information with QC metrics*."

We carefully considered and addressed each comment, including analyses on additional datasets as suggested by the reviewers. In our point-by-point response, we cite the reviewers' unaltered comments in regular black typeface, describe our responses and the resulting improvements to the revised manuscript in regular red typeface, and *we highlight the new manuscript text in blue italic typeface*.

Summary of major additions to the paper in revision:

1. We demonstrate the advantages of SpotSweeper in unbiased identification of low-quality spots across a total of 10 datasets of various tissue types and technologies. Results are presented in newly added Section 2.6 and Figure 4, S4-6 and S8. In these analyses, we include the benchmark to an additional QC method that is developed for sc/snRNA , miQC.

2. To improve the clarity of the proposed method, we greatly revise and extend the method sections (Section 2.1 and Section 4.1) to provide in-depth description of local robust z-score, rationale of parameter selection (Figure S7), mean-variance relationship of mitochondria gene proportion (Fig S10), and software functions.

3. We provide additional evidence that outlier spots identified by SpotSweeper are low-quality and demonstrate SpotSweeper's impacts in downstream analyses. We highlight 1) the outliers identified by SpotSweeper receive discordant cluster labels compared to their nearest neighbors across various clustering methods (Fig 2N); 2) removing regional artifacts identified by SpotSweeper greatly improves the identification of spatial domain markers. The regional artifacts can not be correctly identified through other approaches, including multiple spatial clustering algorithms. (Figure S9)

4. Discussions (Section 3) are updated to align with the updated analyses showing SpotSweeper's generalizability. Addition, we argue the importance of the QC step to preserve biological variability in the data to improve the sensitivity for biological discoveries. We further highlight future opportunities for the rapidly advancing SRT technologies.

5. The SpotSweeper software implementation is updated to improve the clarity of the documentation and add newly requested functionalities.

Reviewer #1 (Remarks to the Author):

The current limitations of spatial transcriptomics QC include challenges in handling data noise and technical biases, constraints in spatial resolution impacting accuracy, difficulties in standardization and normalization across samples, and complexities arising from spatial heterogeneity within tissues and cells. Addressing these issues is crucial for advancing the reliability and interpretability of spatial transcriptomics data analyses.

The manuscript presents a novel approach by proposing quality control (QC) metrics and a computational pipeline aimed at identifying low-quality observations and regional artifacts resulting from sample processing errors in Single-cell RNA sequencing (SRT) data. This methodology addresses a significant issue in the field of SRT data analysis, as ensuring data quality is crucial for accurate biological interpretation. The proposed approach is meaningful and holds potential to improve the reliability and reproducibility of SRT data analyses. However, it requires major revisions to improve the clarity, comprehensiveness, and robustness of the proposed methods and their validation.Here are some questions and suggestions for the authors.

Thank you for finding our proposed method meaningful and addressing significant issues in the SRT data analysis. We have now carefully addressed the reviewer's comments. Please find the point-by-point responses below. For a summary of major changes, please see our response to the Editor.

1. While the proposed method is promising, the manuscript currently lacks sufficient detail in explaining the QC metrics and the computational pipeline. A more comprehensive description is needed to enable readers to fully understand and replicate the approach. This includes a step-by-step explanation of the metrics, the computational steps involved, and any assumptions or parameters used.

We have substantially revised and expanded the method sections (see Section 2.1 and 4.1) to describe SpotSweeper's local outlier identification (including local z-score calculation and justification of threshold, neighbor sizes) and regional artifacts detection in great depth.

The manuscript utilizes K-nearest neighbors (KNN) clustering for downstream quality control (QC) purposes. While this approach is described, it would be beneficial to elaborate on the rationale behind choosing KNN clustering over alternative clustering methods commonly used in single-cell RNA sequencing (SRT) data analysis, such as hierarchical clustering. Providing a comparison or justification for selecting KNN clustering would enhance the manuscript's clarity and help readers understand why this method was deemed most suitable for the proposed QC tasks. Additionally, discussing the advantages and potential limitations of KNN clustering in this context would provide a more comprehensive evaluation of the method's appropriateness for the study's objectives.

We have clarified in the manuscript that SpotSweeper does not use *K*-nearest neighbor clustering algorithm. In the local outlier approach, SpotSweeper identifies *K* nearest neighbors based on the euclidean distance in a 2D space to calculate a robust z-score for each spot. For the hangnail artifact detection, SpotSweeper uses the ***k*-means** clustering algorithm with *k*=2 to identify compromised and uncompromised regions. This is advantageous to hierarchical clustering as we know *a priori* the number of clusters we are trying to detect (i.e., compromised and uncompromised tissue). We additionally compared *k*-means clustering to both gaussian mixed model (GMM) and density-based spatial clustering of applications with noise (DBSCAN), the latter of which is commonly used in snRNA-seq analyses. We show that *k*-means is superior to both of these methods in detecting hangnail artifacts (see Figure S12).

To further strengthen the article, it would be advantageous to include comprehensive performance evaluations of SpotSweeper. This could encompass comparative analyses with existing QC tools, assessing not only the accuracy and effectiveness of SpotSweeper but also its stability and robustness under diverse experimental conditions and across various categories of organs.

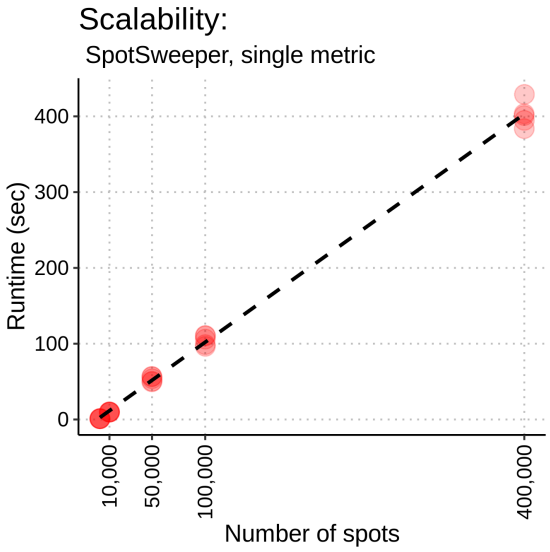
We have now expanded our analyses to evaluate SpotSweeper across a wide range of biological tissue types (human and mouse brain regions, breast and ovarian cancers, and mouse embryo) and spot-based transcriptomic technologies (Slide-seq, Stereo-seq and VisiumHD). Across all of the datasets considered, we demonstrate that the SpotSweeper's localOutlier function greatly reduces the bias in QC metrics across spatial domains. Additionally, we have expanded the number of QC methods to compare with, including global thresholding, 3MAD, and miQC (Hippen et al., 2021). We observed that miQC, which works exceedingly well for snRNA-seq, demonstrates substantial bias across many spatial domains, at times excluding more data points than simple fixed or data-driven thresholds. We have included visualizations of all three major QC metrics across all datasets tested in the main and supplementary figures, as well as visualization for miQC probabilities in main Figure 2 and 4, and Figures S4, S5, S6, and S8.

Overall, through these additional data analysis, we demonstrate that SpotSweeper is accurate, effective, stable and robust

across various types of SRT data.

Additionally, it would be insightful to investigate the computational efficiency and memory footprint of the QC process. A thorough analysis in this regard could provide valuable insights into the practicality and scalability of SpotSweeper for spatial transcriptome studies.

We have now included an analysis testing the scalability of local outlier detection of VisiumHD data containing up to 400,000 spots (Figure S7). We show that SpotSweeper scales linearly with the number of spots taking ~1 second per 1,000 spots to calculate a single QC metric on a single core with 32 GB of RAM. This equates to ~6.67 minutes (400 seconds) to run across 400,000 spots per metric. We have additionally made it so that the localOutlier function now is capable of parallelization implemented with BiocParallel, ensuring that SpotSweeper is efficient and scalable even on large datasets



2. The validation of the proposed QC metrics and computational pipeline appears limited in the current manuscript. It is essential to include more comprehensive benchmarking against existing methods to convincingly demonstrate the effectiveness and advantages of the new approach. In the field of single cell RNA-seq, there are various mature Quality Control assessments. Spotsweeper can compare with these methods or set baseline for benchmark, and highlights the advantages of Spotsweeper than single cells RNAseq data Quality Control.This should involve testing on multiple datasets and scenarios to ensure robustness and generalizability. Additionally, the comparison methods used in the manuscript are few, which limits the scope of the validation.

Thank you for this recommendation. As suggested, we have expanded our analyses to evaluate SpotSweeper across a wide range of biological tissue types and spot-based transcriptomic technologies (main Figure 4 and Figures S4, S5, S6, and S8). Across all of the datasets tested, we show that SpotSweeper's localOutlier function significantly reduces the bias in global QC metrics. In addition, we have expanded our comparisons to include miQC (Hippen et al., 2021), a QC method developed for snRNA-seq that jointly models library and mitochondrial ratio to determine low quality observations. miQC's logic follows that any observation with both low library size and high mitochondrial percentage is indicative of low quality cells/spots. While this method works exceedingly well for snRNA-seq, we found that this method demonstrates substantial bias across many spatial domains, at times excluding more data points than simple fixed or data-driven thresholds.

We would additionally likely highlight that SpotSweeper is the first method developed for identifying regional artifacts in spatial transcriptomics data.

To substantiate the superiority of the QC method, it is recommended that the authors expand the discussion to include downstream applications. Specifically, the authors could present comparative studies that evaluate the impact of applying

the QC method discussed in this paper versus no QC or alternative single-cell QC methods on the outcomes of subsequent analytical steps.

Thank you for your suggestion. We agree that enhanced discussion on downstream impacts would be beneficial to the manuscript. We thus have included the following Discussion around a recent preprint demonstrating the QC (particularly high mito percentage) is confounded with cancer biology:

"Accounting for biological variability in quality control (QC) metrics is essential to preserving the integrity of downstream analyses, as overly stringent filtering can obscure biologically meaningful subpopulations. This issue is particularly critical in contexts such as cancer biology, where malignant cells often exhibit unique transcriptional and metabolic features that deviate from nonmalignant counterparts. For example, a recent study found that high mitochondrial ratio, typically used as a marker for cell stress or death, may instead reflect functional metabolic adaptations in certain malignant cells [27]. By refining QC practices to better accommodate biological variability, researchers can retain these potentially significant cell populations, thereby enhancing the resolution of cell type-specific insights and improving the downstream impact of SRTs data. This underscores the importance of balancing robust outlier detection with the preservation of biologically meaningful variability to avoid misinterpretation of critical subpopulations in both health and disease."

3. The manuscript primarily focuses on the 10X Genomics platform(such as 10x Visium human DLPFC datasets and Mouse Coronal Brain dataset), which is a limitation. To increase the general applicability and robustness of the proposed method, it would be beneficial to test and validate it on a variety of platforms. This will demonstrate the versatility and broader utility of the QC metrics and computational pipeline across different SRT technologies. Besides, many analysis of Spotsweeper only evaluates on single 10X dataset and obtains numerical results, without additional biology discovery. We look forward to new biological discoveries using Spotsweeper in some scenes. For example,the manuscript could benefit from incorporating targeted case studies that demonstrate the efficacy of SpotSweeper across a variety of SRT data types and organ systems. These case studies should illustrate how SpotSweeper's application leads to tangible improvements in data quality and subsequent analysis.

We are happy to share that we have expanded our analyses to include 1) additional tissue types in the 10x Visium platform, 2) three additional spot-based technologies, including VisiumHD at both 8um and 16um resolution, and 3) an imaging-based technology, Xenium. We show that SpotSweeper is highly effective at reducing biases in QC metrics across all spot-based technologies and tissue types. This demonstrates that SpotSweeper is versatile and has broad utility across biological applications. While we also look forward to new biological discoveries made with the assistance of SpotSweeper, however biological discovery is not SpotSweeper's primary use-case. Rather, SpotSweeper's primary use is to reduce bias in quality control in order to ensure the robustness and validity of downstream analyses, which we effectively demonstrate here.

4. The manuscript only employs BayesSpace method for identifying spatially domain, which is limited. To strengthen the manuscript, it is recommended to explore and compare additional methods for detecting spatially dominant features. This will provide a more comprehensive assessment and validate the effectiveness of the chosen method.

We have now additionally included both PRECAST and BANKSY spatial clustering methods on the DLPFC sample with a hangnail artifact (Figure S9). We found that neither were successful in identifying the artifact. In addition to artifact detection, SpotSweeper incorporates a method to automatically annotate artifacts based on the low local variance - a capability beyond any other existing clustering method.

5. The manuscript mentions that the QC metrics can systematically filter out observations with low total Unique Molecular Identifiers (UMIs). However, more detail is needed on how this filtering is achieved and its impact on the overall data quality. Additionally, it would be helpful to include a discussion on the potential consequences of removing low UMI observations, such as the risk of losing biologically relevant information

We have clarified in the manuscript how SpotSweeper identifies low-quality spots using total number of UMIs as a metric. Specifically, we use Equation (1) to calculate a robust z-score for each spot, comparing its total number of UMIs to those

of its *k spatial neighbors*. As the robust z-score follows a standard normal distribution, we use 3 standard deviations (i.e. 3) as the threshold. Overall, Z_i<-3 for each Spot *i* was used to identify those observations with low UMI. Compared to other QC methods, SpotSweeper mitigates the risk of removing biologically relevant information due to globally low UMIs, as shown in Figure 2. As suggested, and previously mentioned above, we have included a paragraph discussing the potential consequences of globally removing low UMI observations, particularly in the context of cancer biology.

6. Spotsweeper is limited to Spot Level, and even testing at 10x Visium technology. With the development of spatial sequencing technology, the development of spatial single-cell level technology (Such as Starmap and Baristaseq, etc.) is now developing. Spotsweeper is expected to be more scalability and generalization.

As we previously mentioned in Item 1 and 3, we have expanded our analyses to include 1) additional tissue types in the 10x Visium platform, 2) three additional spot-based technologies, including VisiumHD at both 8um and 16um resolution, and 3) an imaging-based technology, Xenium. We show that SpotSweeper is highly effective at reducing biases in QC metrics across all spot-based technologies and tissue types. This demonstrates that SpotSweeper is versatile and has broad utility across biological domains and technologies.

While SpotSweeper ran efficiently on the Xenium platform, we show that QC biases are largely absent in imaging-based technologies, likely due to the standardization of the biological unit being sampled (i.e., segmentation of transcripts located in the cell body/nucleus) and the manual selection of a small set of highly expressed genes (typically 100-500). This removes a lot of bias inherent to spot-based technologies. Despite this, SpotSweeper was equally effective at identifying outliers in this Xenium dataset as current methods. We have included a discussion around this point and highlight it as an area for future research, see below.

“While we effectively demonstrate that SpotSweeper outperforms current QC methods in sequencing-based technologies, we observed substantially less variability and bias in library size and unique gene counts across spatial domains in imaging-based methods. We speculate that this reduced variability is due to the standardization of the biological unit being sampled in imaging-based methods, such as nuclei or cell bodies identified through precise cell segmentation. In contrast, sequencing-based technologies sample transcripts indiscriminately from larger, undefined spatial areas, resulting in aggregate signals from cell bodies, neuropil, and extracellular material, which naturally introduces large variability across spatial domains. For example, in the brain datasets analyzed here, we observed consistently lower library sizes and unique gene counts in white matter regions, which contain fewer cell bodies, compared to the cell-body-dense gray matter regions. Moreover, most imaging-based methods are restricted to only a few hundred manually selected genes and typically do not include mitochondrial genes, further reducing variability in genes detected and total counts per cell. While SpotSweeper currently uses multiscale variance of mitochondrial ratio to detect these artifacts, it is possible that a similar approach utilizing the negative control genes normally included in image-based methods may be useful for detecting damaged tissue sections. In addition, rasterization techniques that aggregate mRNA counts into spatial pixels [28] will increase compatibility with the current SpotSweeper workflow, while ensuring the scalability of our approaches for imaging-based platforms. Moreover, the current implementation of SpotSweeper should be amenable to future spot-based technological advancements, such as the VisiumHD platform from 10x Genomics [29], that have substantially increase spatial resolution with complete tissue coverage.”

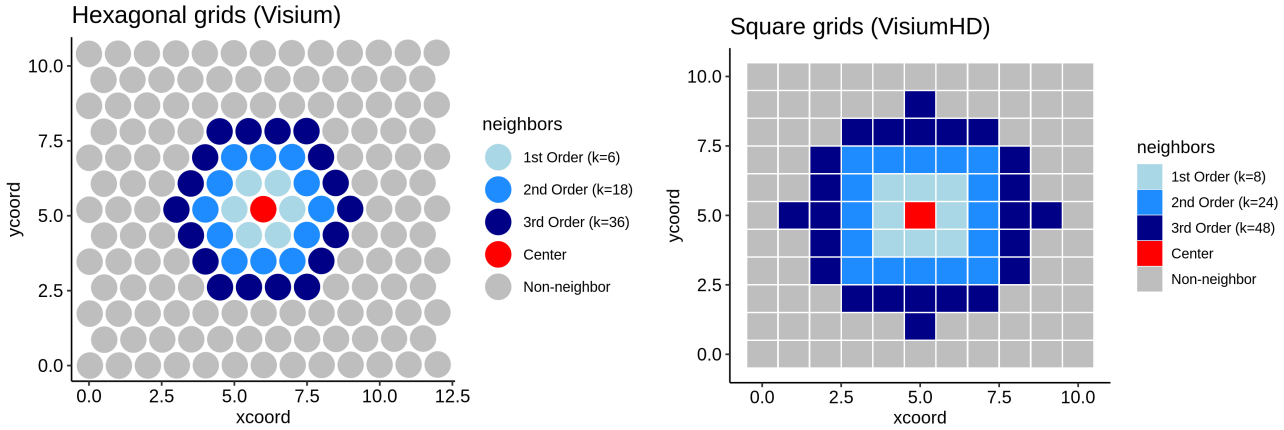
Besides, some minor advice for the authors is as following,
1.It would be better if the Figure.1 add some downstream tasks examples.

We have added a new panel to Figure 2 showing that local outliers display discordant clustering compared to their nearest neighbors. Inaccurate clustering will inherently lead to biases in any downstream results based on clustering.

2.It would be better if adding an explanation of parameters selection in your algorithm and how to select appropriate parameters.

We have now added the following details and figure regarding parameter selection to the methods:

"For SRT technologies that use hexagonal (e.g. Visium) and square (e.g. Stereo-seq and VisiumHD) grid patterns, we find that using third-order neighbors (k=36 and k=48, respectively) for each spot works well. For SRT technologies with non-uniform spacing (e.g., Slide-SeqV2 and Xenium), we find that using the default setting of k=36 works well."



3.It is recommended to visualize the resultant plots of all slices of DLPFC data in the visualization of sup figure.

We have added Figure S1 to visualize the local outliers across all twelve DLPFC Visium samples from Maynard et al. (2021).

4.It is recommended to specify whether the experiment was repeated and how many times it was repeated in order to assess the consistency and reliability of the results.

We thank the reviewer for the comment. The local outlier detection method in SpotSweeper is **static**, meaning the algorithm provides consistent results for the same input data across repetitions. We have clarified this in the discussion section of the manuscript, stating *"Spotsweeper is a static algorithm."*

5.It is recommended to provide an estimate of the computing resources needed to run SpotSweeper, including memory, processor, and time consumption.

We have now included an analysis testing the scalability of local outlier detection of VisiumHD data containing up to 400,000 spots (Figure S7). We show that SpotSweeper scales linearly with the number of spots taking ~1 second per 1,000 spots to calculate a single QC metric on a single core with 32 GB of RAM. This equates to ~6.67 minutes (400 seconds) to run across 400,000 spots per metric. We have additionally made it so that the localOutlier function now is capable of parallelization, ensuring that SpotSweeper is efficient and scalable even on large datasets.

Reviewer #1 (Remarks on code availability):

According to the code provided by the author, we have conducted testing using two datasets. The help file and the annotation in the code are not detailed enough, such as the version information of R. Since I am using the latest version of R, the author's code did not report any errors during installation, but I am unable to call the relevant functions when performing visualizations. The author is requested to provide additional information.

We thank the reviewer for trying out SpotSweeper. We have updated SpotSweeper to development version 1.3.2. We have made improvements regarding functionality and documentation.

Reviewer #2 (Remarks to the Author):

The authors set out to develop a framework to enable spatially aware quality control. The approach leverages the spatial dependencies imposed in spatial data to identify local and region-based low-quality measurements. The proposed approach, SpotSweeper, is a significant advancement on the current approach to perform quality checks in spatial transcriptomics datasets and holds the potential to form part of every standard spatial transcriptomics analysis. The authors innovatively generalise the analysis such that the tool works with any quality metric and allows multiscale analysis of quality as well. I commend the authors on developing such an impressive quality control tool and on demonstrating its ability extensively in this manuscript. Through analysis of various Visium datasets, the authors identified a set of consistently poorly performing barcode sequences. This result will help improve array barcoding for 10x Visium making SpotSweeper an important tool to discover and avoid problematic barcode sequences in newly developed technologies.

I have no major concerns with the manuscript but do feel that some additional work is needed to generalise the tool to other spot-based technologies. It is an important tool for all spot-based technologies and extending support to the 10x Visium HD and BGI STOmics technologies which have a regular square grid will greatly enhance its potential. Besides this one request from the authors, I have a few minor comments to improve clarity of text in the manuscript.

We thank the reviewer for finding our proposed method a significant advancement in quality controlling SRT data. We have carefully addressed all the comments. Please find the point-by-point responses below. For a summary of major changes, please see our response to the Editor.

Major comments:

1. The proposed method can be applied to spot-based sequencing technologies. The authors have reported on a comprehensive analysis of 10x Visium data. However, other spot-based technologies such as the BGI stereo-seq (STOmics) and the 10x Visium HD have not been studied. When the density of spots is increased, as well as their resolution, it would be good to get recommendations on how SpotSweeper should be applied. I suspect there to be more outlier spots in these settings so it would be useful to get guidance on what to expect from these datasets. Publicly available data exist from these technologies so they should be added to the comparisons performed in the paper.

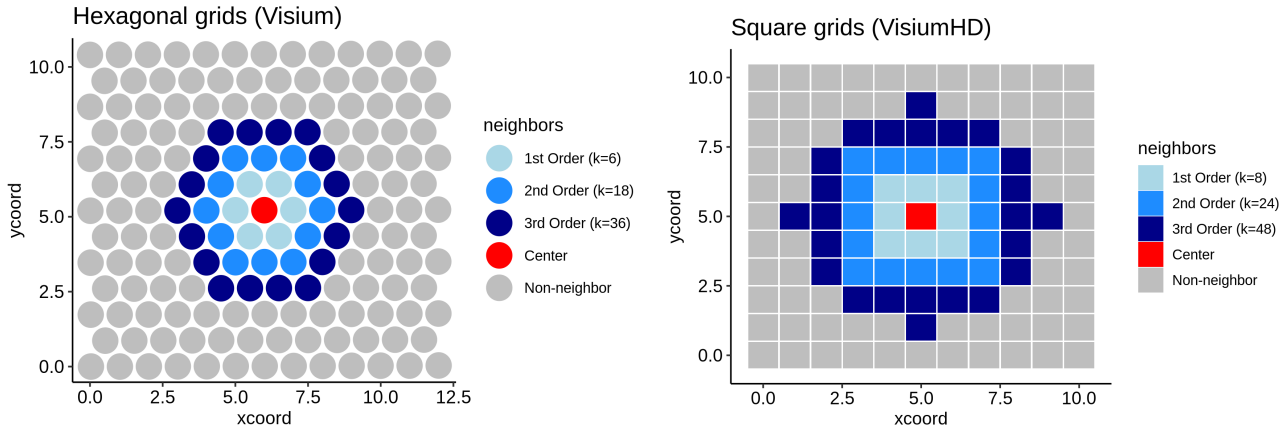
We thank the reviewer for the thoughtful comment! We now include the application of SpotSweeper to a total of 10 datasets that include 1) additional tissue types in the 10x Visium platform (Figure 4), 2) three additional spot-based technologies, including VisiumHD at both 8um and 16um resolution (Figure 4, S6), and 3) an imaging-based technology, Xenium (Fig S8). We show that SpotSweeper is highly effective at reducing biases in QC metrics across all spot-based technologies and tissue types (see Section 2.6). The empirical evidence suggests the default parameters of SpotSweeper generalizes to technologies where spots are more dense and have higher resolution. We are optimistic these additional analyses and their results greatly contribute to developing expectations on quality control practice among rapidly evolving SRT technologies.

2. Consequently, it would help to have implementations of SpotSweeper where the grid type is not enforced to be a hexagonal tessellation, but a simple square grid.

We thank the reviewer for the thoughtful comment! We agree that making SpotSweeper easily accessible to different geometric arrangements would greatly improve its impact. This can be easily achieved by using a different *k* value for each arrangement, when identifying local neighborhoods through *k* nearest neighbor (*k*NN). For example, we suggest using *k*=36 for hexagonal arrangement and *k*=48 for square arrangement as the default in local outlier identification. Additionally, we provide the equations calculating the value of k given the order of neighbors preferred

"For each neighborhood order c, the exact neighborhood size K(c) can be calculated using $K_{hex}(c) = 3c(c + 1)$ *for hexagonal arrangement (e.g., Visium) and* $K_{grids}(c) = 4c(c+1)$ *for square grids."*

and include Figure S7 (see below) to visualize the *k*NN selection to achieve *1st-3rd* order neighbors for both hexagonal (*k*=6/18/36) and square grid arrangements (*k*=8/24/48).



It should be noted that the third-order neighbors do not necessarily conform to a complete square. We argue that acquiring exact *k*NNs is desirable over conforming to a square grid. Biological tissue exhibits complex, non-uniform geometric arrangements, where cells or spots in spatial transcriptomics data that are physically close are more likely to share similar gene expression profiles and/or microenvironments. Thus, unlike a rigid grid system, using exact *k*NN ensures that neighborhood comparisons remain as local as possible. Moreover, in our data analyses, we observe empirically that using exact *k*NNs works well in identifying local outliers. To note, grid arrangements of third-order neighbors were used and validated in all local outlier analyses of VisiumHD and Stereo-seq in the current manuscript.

3. The dataset used to produce figure 4H seems to have tissue folding. This is an artifact that could affect downstream analyses and is visible in the library size. Is it possible to identify these effects using SpotSweeper? If so, it would be worth demonstrating this use case. This is an aspect of quality control, and it would be great if SpotSweeper is able to handle this issue as well. If not, the authors can comment on this issue being present in the older visium samples where CytAssist was not available, and mention that this is an outstanding problem that may need alternative solutions.

Thank you for this thoughtful comment. The downstream impact of tissue folds was previously investigated in Hukki-Myers et al. using the DLPFC data set. The authors concluded that tissue wrinkles and folds do not impact gene-level or spot-level downstream analysis (see Section *Evaluating the impact of histology artifacts on Visium H&E data* in Supplementary Materials). (Huuki-Myers et al., 2024) Additionally, we agree that the CytAssist technology could mitigate tissue folds and wrinkles for future experiments. Therefore, we didn't specifically optimize and experiment SpotSweper to address tissue folds.

Minor comments:

1. The more common approach to defining outliers is using Tukey's approach (>q0.75 + k * IQR and <q0.25 – k * IQR, where k is usually 1.5) that is often used do annotate outliers in boxplots. The authors use an alternative approach based on a modified z-score approach. I would be interested in knowing the reasoning behind the chosen outlier labelling approach.

We thank the reviewer for the thoughtful comment. We use the robust z-score approach for the following reasons. First of all, within each local neighborhood, the number of spots are limited, e.g. 36 for Visium by default. While Tukey's approach with IQR is popular, its utility in data of small-sample-size is not well understood. We respectively argue that calculating quantiles with such small sample sizes could be challenging and unstable. In contrast, the robust z-score approach has been shown to be successful and stable in the geospatial literature (Chen, Lu, Kou, & Chen, 2008). Secondly, the robust z-score provides an improved interpretability as it can be easily mapped to a standard normal distribution and enable a uniform threshold across all spots. In summary, we prefer the robust z-score over IQR.

2. The authors have detailed the computation of z-scores in the first result section, however, they should add a sentence stating the approach to determining outlier thresholds. Normally, this would be some quantile of the z-score, but I see that from the methods section that the proposed approach is to use the z-scores directly. There is nothing wrong with this, but

it is better to report it more clearly. I would also advise reporting the quantile matching the z-score cut-off of -3 that is used to demonstrate how strong the outliers are.

We thank the reviewer for the comment. We greatly revise and expand Section 2.1 and 4.1, describing the local outlier and regional artifacts methods in more depth. Specifically, we explain the rationale of using a robust z-score, that is, it conforms to a standard normal distribution. We further clarify the default threshold choice, i.e. spots with $z_i<3$ total UMI and total unique gene counts and $z_i >3$ for mito ratio are identified as low quality. The default threshold represents 3 standard deviations away from the mean under the standard normal distribution,commonly used as a threshold in statistics literature. Additionally, the choice of threshold can be interpreted that we estimate 1-Phi(3) = 0.13% would be flagged as outlier spots under the normality assumption, where Phi is the cumulative distribution function of a standard normal distribution.

3. Add equation numbering

We have added equation numbering throughout the manuscript.

4. Page 3, line 71: Should it state “Then, we calculate a robust z-score using all spots in the neighbourhood:” rather than “for all spots”?

We thank the reviewer for spotting this typo. We have corrected the sentence in the manuscript.

5. I recommend matching the full scaling factor used in the equation for the z-score and that in figure 1 (and possibly annotating it as the q0.75 of a standard normal). It is important to state how this value is obtained. From the referred textbook, I see that the scaling factor is obtained by computing the expectation of the MAD in terms of the standard deviation given a large n which evaluates to the 75 percentile of the standard normal. Writing it as such will enhance clarity.

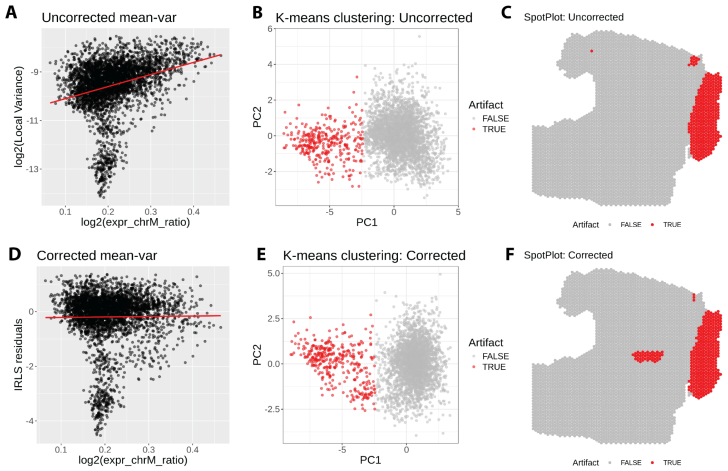
We thank the reviewer for the comment. We have improved the clarity of the Method and Result sections by explaining the rationale of using a robust z-score that is scale to a standard normal distribution. A sentence has been added to explain the scaling factor.

6. Page 14, line 284 should be < -3 z-scores.

We thank the reviewer for spotting this typo. We have corrected the sentence in the manuscript.

7. Robust linear regression is used to study variance of metrics at multiple scales. I think readers would be interested in understanding the mean variance relationship across scales. Would the authors be able to provide supplementary figures showing examples of this?

Thank you for this suggestion. We have now added the below figure to the supplementary information (Figure S10) demonstrating the mean-variance relationship and its correction.



8. It would be good to have a function that annotates the 6 problematic barcodes in Visium datasets so that users can discard these with ease and not have to replicate the outlier analysis (the results of which may vary given their threshold selection).

We thank the reviewer for the thoughtful comment. We have created the *flagVisiumOutliers* function to automatically flag and drop these systematic outlier spots based on the barcodes of those spots.

Reviewer #2 (Remarks on code availability):

I have reviewed the code on Bioconductor and it was easy to replicate the analysis. The code itself has already undergone peer-review as part of the submission process. The automatic build and testing systems show that the code is functional.

We thank the reviewer for the positive feedback regarding the software implementation and documentation. We are pleased to hear that these aspects were clear and helpful, as we aimed to make SpotSweeper both accessible and user-friendly for researchers.

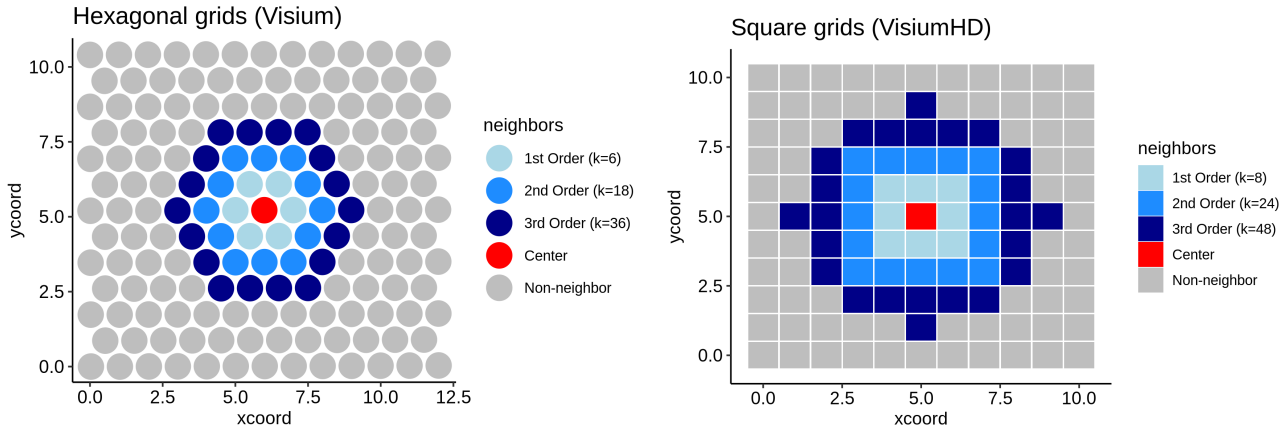
Reviewer #3 (Remarks to the Author):

This manuscript presents approaches for spatially-aware quality control of spot-based SRT technologies, primarily 10X Visium. They illustrate the method using DLPFC datasets (two studies; Maynard et al and Huuki-Myers et al). In doing so, they also discuss the detection of “dryspot” and “hangnail” artifacts. The approach for using spatial neighborhoods (K-NN) overcomes the issue of confounding of biological information with QC metrics.

However, it is unclear how researchers would select the neighborhood size, how the methods would work at boundary areas of domains (spots in two different domains), and the impact of their QC has on aspects other than domain detection, such as phenotyping/deconvolution methods.

We thank the reviewer for the thoughtful comments! We now revise and expand Section 2.1 and 4.1, describing the local outlier and regional artifacts methods in more depth. Please find the revision highlighted in red in the revised manuscript. Specifically, we further clarify how the neighborhood is structured, depending on the spot arrangements of the SRT technologies,

"For SRT technologies that use hexagonal (e.g. Visium) and square (e.g. Stereo-seq and VisiumHD) grid patterns, we find that using third-order neighbors (k=36 and k=48, respectively) for each spot works well. For SRT technologies with non-uniform spacing (e.g., Slide-SeqV2 and Xenium), we find that using the default setting of k=36 works well."



The figure demonstrating the neighbors arrangement for the different spot arrangements is now added to the manuscript as Figure S7. These default neighborhood sizes were examined across various tissue types and SRT technologies and provide empirical evidence for its efficiency in the local outlier identifications. (Figure 2, 4 and S6)

Applying SpotSweeper to various datasets, we provide evidence that SpotSweeper is robust to domain boundaries, specifically not falsely identifying spots at the domain boundaries as outliers which would be a major concern. For example, in Figure 2A-D, we do not see that spots are any more or less likely to be detected as outliers along spatial boundaries. This is true for all additional tissue types and technologies analyzed (Figure 4).

While acknowledging that the reviewer found our demonstration for downstream impact in spatial clustering and domain marker detection useful, we respectfully argue that a comprehensive evaluation of downstream impact is out of the scope of the current manuscript.

The methods proposed are also not applicable to technologies other than Visium and are not testing in single-cell SRT technologies, which is the future of SRT. Lastly, it would have been nice to see the QC of more complex/complicated tissues in spatial organization/structure as DLPFC tissues are highly organized (e.g., how would SpotSweeper work on cancer tumor tissue?).

We thank the reviewer for the thoughtful comments. In the revised manuscript, we apply SpotSweeper to a total of 10 datasets of various tissue types (ranging from mouse brain to human brain to ovarian and breast cancer) as well as technologies (including Visium, Slide-seqV2, Stereo-seq, Visium HD (at both 8um and 16um resolutions)), as well as Xenium). Compared to the current QC practice, SpotSweeper is highly effective at reducing biases in QC metrics across all spot-based technologies and tissue types (Figure 2, 4, S6 and S8). With the substantial evidence demonstrating SpotSweeper's robust performance across tissue types and technologies, we are highly optimistic that SpotSweeper will be highly generalizable to new single-cell SRT technologies, as well as complex tissue structures such as cancer tumor tissues.

In terms of the package, SpotSweeper is available on Bioconductor and GitHub and has a MIT license. They have a vignette on how to use the package functions and plotting results. They also have a documentation page for the package.

We thank the reviewer for the positive feedback regarding the software implementation and documentation.

Reviewer #3 (Remarks on code availability):

The package is on GitHub and Bioconductor. They have instructions on how to install the package and examples of how to use the package for SRT QC. The package and code are very usable.

We thank the reviewer for the positive feedback regarding the software implementation and documentation. We are pleased to hear that these aspects were clear and helpful, as we aimed to make SpotSweeper both accessible and user-friendly for researchers.

Bibliography

Chen, D., Lu, C.-T., Kou, Y., & Chen, F. (2008). On detecting spatial outliers. *GeoInformatica*, *12*(4), 455–475.

Hippen, A. A., Falco, M. M., Weber, L. M., Erkan, E. P., Zhang, K., Doherty, J. A., Vähärautio, A., et al. (2021). miQC: An adaptive probabilistic framework for quality control of single-cell RNA-sequencing data. *PLoS Computational Biology*, *17*(8), e1009290.

Huuki-Myers, L. A., Spangler, A., Eagles, N. J., Montgomery, K. D., Kwon, S. H., Guo, B., Grant-Peters, M., et al. (2024). A data-driven single-cell and spatial transcriptomic map of the human prefrontal cortex. *Science*, *384*(6698), eadh1938.

Totty et al – ***Author responses to reviewers' comments***

We extend our sincere thanks to all 3 reviewers for their strong interest in SpotSweeper and helpful comments. We carefully considered and addressed each comment, including analyses on additional datasets as suggested by the reviewers. In our point-by-point response, we cite the reviewers' unaltered comments in regular black typeface, describe our responses and the resulting improvements to the revised manuscript in regular blue typeface.

Summary of major additions to the paper in revision:

- We extended our validation of SpotSweeper on image-based platforms to include additional MERFISH and STARmap datasets beyond Xenium. We provided strong evidence that SpotSweeper can successfully identify outliers and perform as good, if not better, than the current QC practices in imaging-based platforms.
- We substantially improved transparency of parameter choices for local outlier identification and regional artifact detection for both sequencing-based and imaging-based platforms. We demonstrated that both local outlier and artifact detection are robust across parameter settings of both sequencing-based and imaging-based SRT data, justifying our recommended parameter settings.
- We extended our discussion by highlighting our philosophical viewpoint of data analysis: data analysis for imaging-based technologies, including quality control, would substantially benefit from further research into tailored computational methods due to the inherent differences in data generation, and in turn data properties, from sequencing-based platforms.
- We integrated the above evidence in the manuscript, including five new supplementary figures (two on imaging-based platform validation, and three on impact of parameter choices).

Reviewer #1 (Remarks to the Author):

The revision for the SpotSweeper article mainly added some datasets and conducted more validation experiments. Overall while SpotSweeper presents a technically sound quality control tool with incremental improvements, the revised version of the manuscript right now does not meet the scientific innovation, comprehensive validation, and methodological rigor required for publication in Nature Methods.

We thank the Reviewer for their continual interest in the manuscript. We have now carefully addressed the reviewer's comments. Please find the point-by-point responses below. For a summary of major changes, please see our response to the Editor.

My main concerns are as follows.

1.Insufficient Scientific Innovation

The manuscript remains centered on technical validation without demonstrating a truly novel biological discovery. There is no indication of how this tool advances key biological questions or reveals previously unknown insights. Authors emphasize the tool's ability to identify regional artifacts (e.g., dryspots and hangnails) and its improvement over existing QC methods. However, no convincing biological discoveries were presented as evidence of SpotSweeper's practical scientific impact.

According to the editors of Nature Methods, these points were considered outside the scope of the present work.

2.Lack potential biological applications
Although the authors mention the potential biological applications of SpotSweeper in detecting high-quality spots and artifacts, there is a lack of specific application examples. We strongly recommend that the authors provide one or more concrete application cases to demonstrate how the results identified by SpotSweeper can be used for downstream analysis and elaborate on the biological significance of these analyses. This will not only prove the practicality of SpotSweeper but also enhance readers' understanding of the scope of the method's application.

According to the editors of Nature Methods, these points were considered outside the scope of the present work.

3.Incomplete Validation

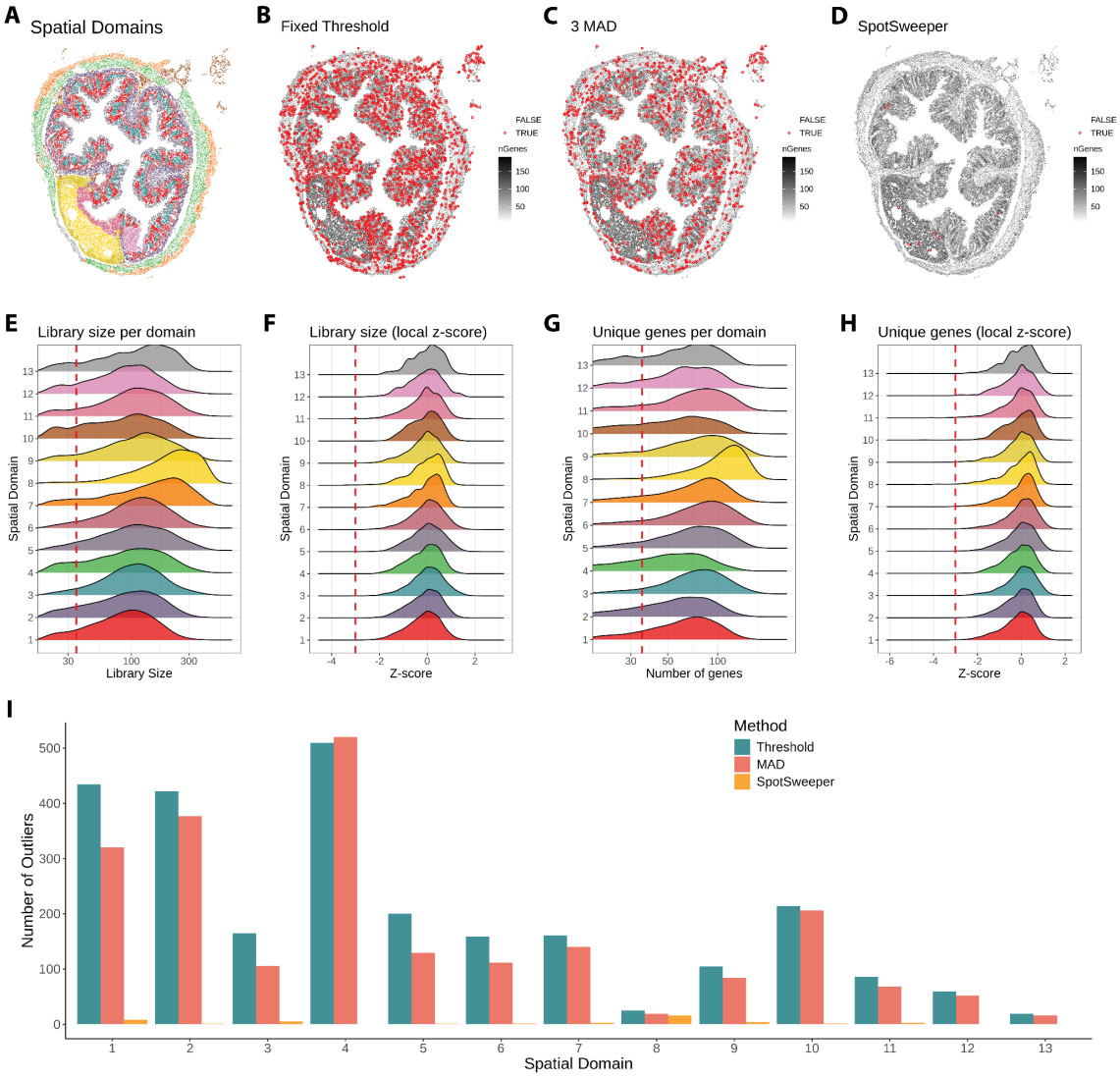
The validation of SpotSweeper remains limited to 10x Visium and partial datasets from VisiumHD and Xenium platforms. The method was not validated on key single-cell spatial technologies such as MERFISH and STARmap , which are essential benchmarks in the field. Authors only added validations on VisiumHD, Xenium, and other 10x datasets. They argued that SpotSweeper generalizes well across spot-based platforms. However, no experiments or benchmarks were conducted for MERFISH or STARmap, and validation remains absent for high-resolution, single-cell data. We explicitly pointed out this limitation, but the authors did not provide data or evidence to address it.

We first would like to respectfully point out that "*validation remains limited to 10x Visium, Visium HD, and Xenium platforms*" is incorrect. We successfully demonstrated that SpotSweeper is highly robust across 9 real-world datasets generated from 5 different technologies across spot, single-cell, and subcellular resolutions in the prior round of revision. This includes Slide-seq and STOmics Stereo-seq technologies. Additionally, the statement that "*validation remains absent for high-resolution, single cell data*" is also inaccurate. In the previous revision, we validated SpotSweeper on two datasets with subcellular resolution; sequencing-based VisiumHD and imaging-based Xenium (Fig. S7, S9).

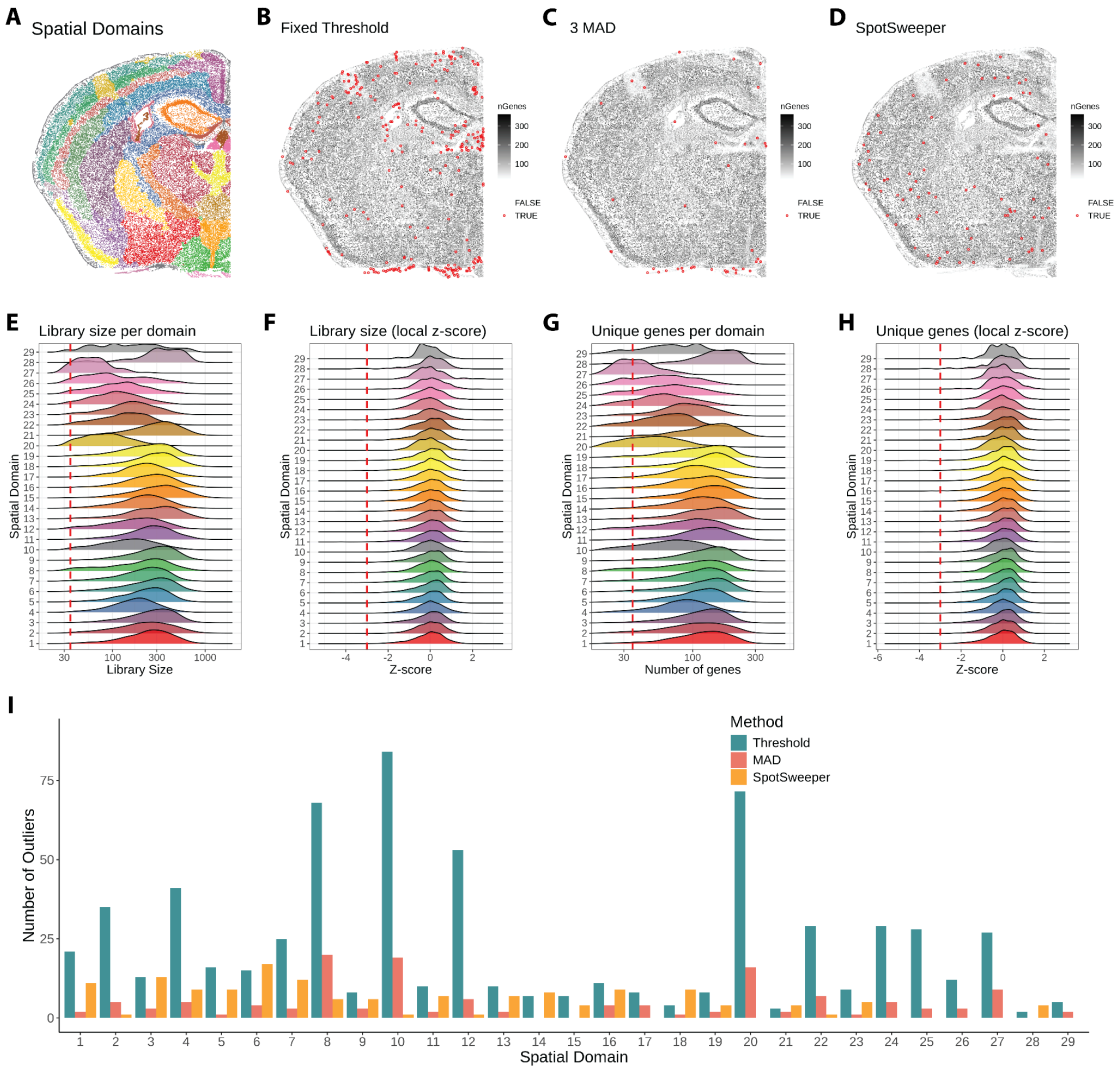
As suggested by all three reviewers in the prior round of reviews, we initially included the Xenium dataset as an example of how SpotSweeper performs on imaging-based technologies. Xenium is a high-resolution, single cell platform that is akin to MERFISH and STARmap. While we did

not originally analyze a MERFISH or STARmap dataset, MERFISH, STARmap, and Xenium are all imaging-based, single molecule *in situ* hybridization approaches. Thus, we believed it was reasonable to assume the validation with Xenium is equally informative.

That being said, ***we have now added additional MERFISH*** (SFig 10) ***and STARmap datasets*** (SFig 11)***.*** Similar to the result of Xenium data, we found that 1) data generated from imaging-based technologies show substantially less biological confounding than those from sequencing-based technologies, and 2) SpotSweeper performs as well as current QC methods on imaging-based platforms. Furthermore, we previously added an extended discussion around these two points and reasoned that the biological unit of individual data observation is standardized due to cell segmentation in imaging-based platforms. In contrast, sequencing-based platforms do not perform segmentation inherently. This results in large variation in the number and type of transcripts captured in each observation due to the sampling of highly heterogeneous tissues underlying each spot. While we acknowledge that we could have been clearer in stating this, ***we found that SpotSweeper successfully identifies outliers in imaging-based platforms***. Philosophically, we believe the data analysis for imaging-based technologies, including quality control, would substantially benefit from further research into tailored computational methods, due to the inherent differences in data generation, and in turn data properties, from sequencing-based platforms. We have added these data to the Results and Supplementary data sections.



Supplementary Figure S10: SpotSweeper performance on a MERFISH mouse colon dataset. (A) Spot plots of a MERFISH dataset of single healthy mouse colon tissue section from Cadinu et al.[24] colored by Banksy spatial domains. (B-D) Spot plots displaying the number of detected genes overlaid with the low quality spots (red) identified using either fixed thresholds (C), three MADs (D), or local outliers as detected by SpotSweeper. (E-H) Ridge plots of the distribution of library size and number of detected genes across spatial domains. Red dotted ines indicate fixed thresholds to identify outliers (< 35 sum genes and < 25 unique genes, or > 3 z-scores). (I) Bar plots of the total number of discarded spots across spatial domains using the four different quality control methods, including SpotSweeper.



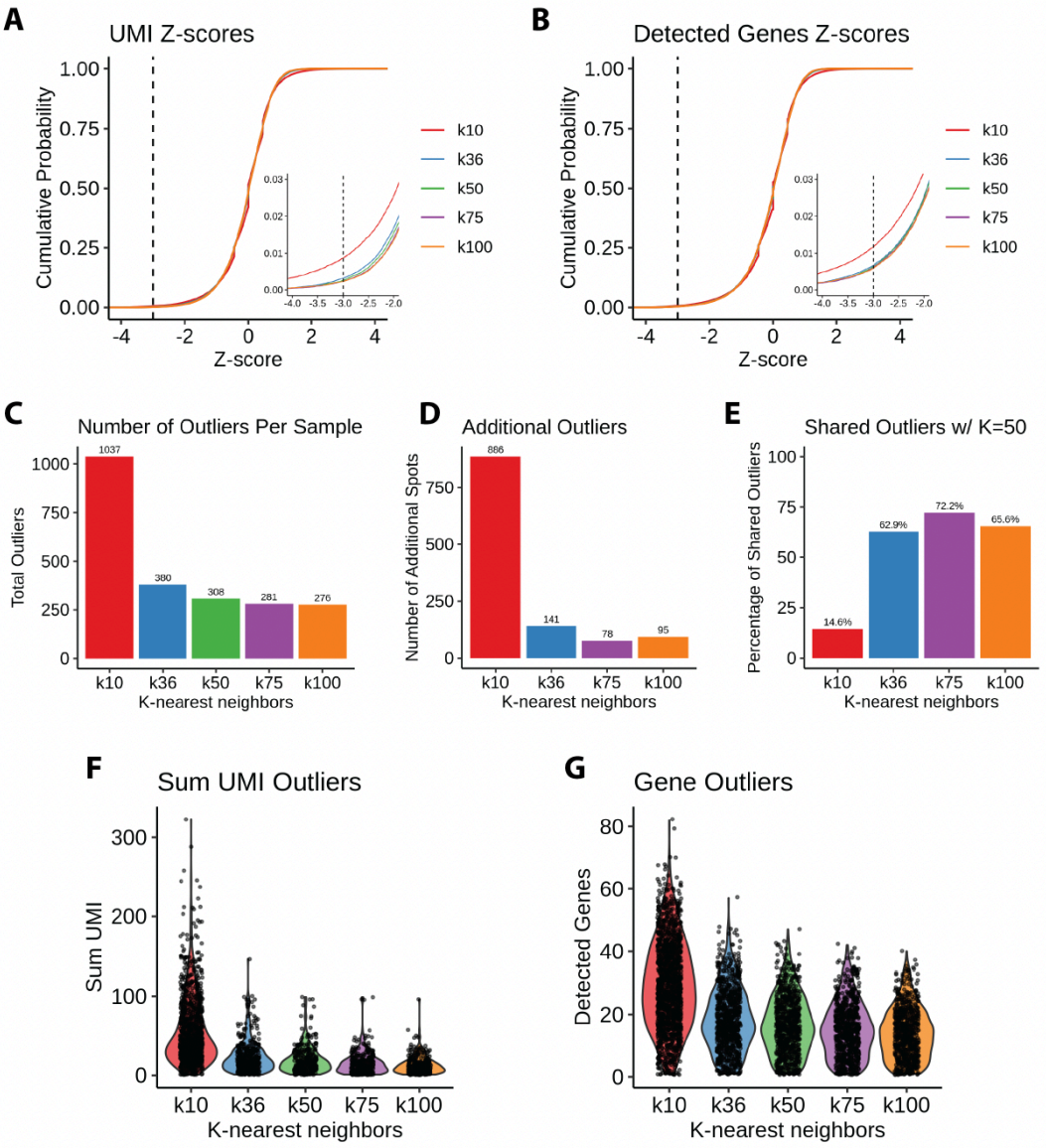
Supplementary Figure S11: SpotSweeper performance on a STARmap mouse brain dataset. (A) Spot plots of a STARmap dataset of one hemisphere coronal section of a mouse brain from Shi et al.[25] colored by Banksy spatial domains. (B-D) Spot plots displaying the number of detected genes overlaid with the low quality spots (red) identified using either fixed thresholds (C), three MADs (D), or local outliers as detected by SpotSweeper. (E-H) Ridge plots of the distribution of library size and number of detected genes across spatial domains. Red dotted ines indicate fixed thresholds to identify outliers (< 35 sum genes and < 25 unique genes, or > 3 z-scores). (I) Bar plots of the total number of discarded spots across spatial domains using the four different quality control methods, including SpotSweeper.

4.Methodological Transparency and Parameter Selection

The justification for parameter choices (e.g., neighborhood size for KNN and multiscale parameters) remains unclear. Systematic comparisons or optimization processes are absent, leaving the method's robustness questionable across varying resolutions and datasets. Authors provided explanations for parameter defaults (e.g., using third-order neighbors) and included a visualization (Figure S7) of parameter choices for hexagonal and square grids. However, they did not provide systematic benchmarks or experimental analyses to justify parameter settings.

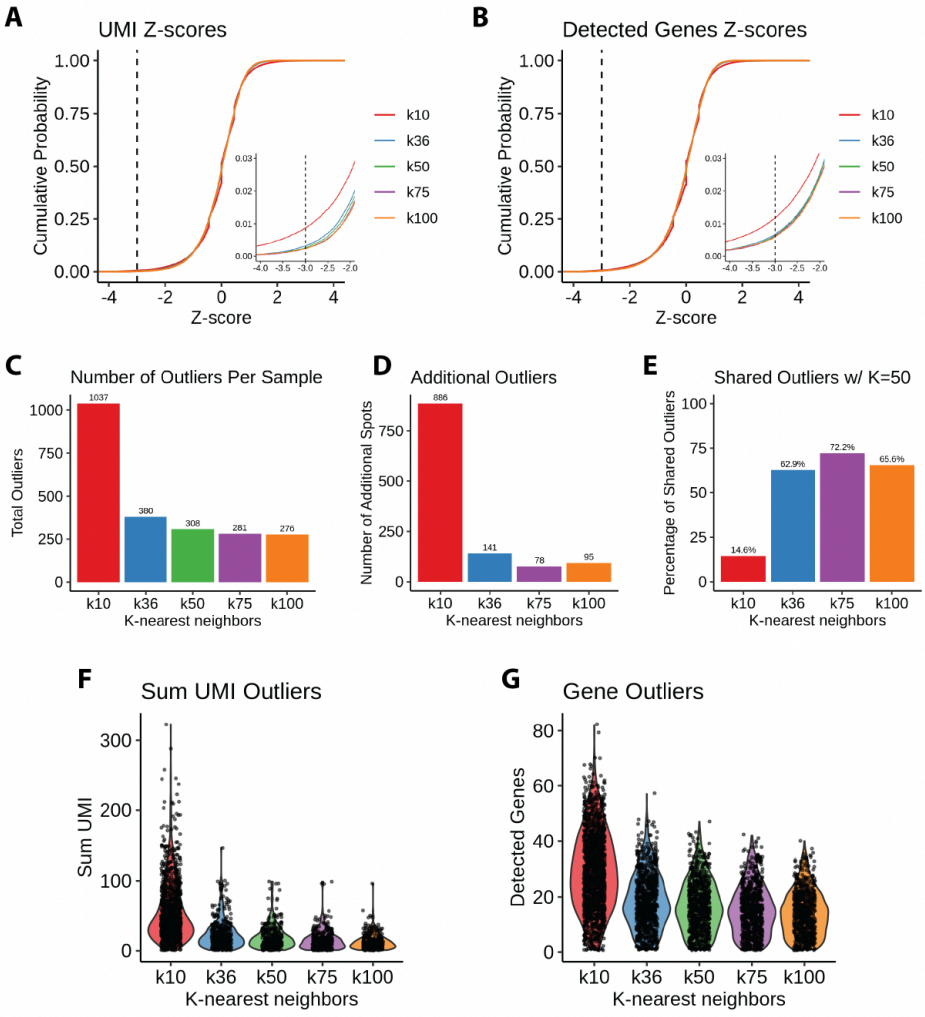
We appreciate that Reviewer #1 acknowledges our revision improved the clarity describing the algorithm of SpotSweeper. In addition to our previous revisions showing that the recommended parameters setting for local outliers are highly robust across tissue types and technologies, we are excited to present new data demonstrating that both local outlier and artifact detection is robust across a range of parameter settings using the Maynard et al. DLPFC Visium and Xenium 5K Coronal Mouse Brain datasets. We respectfully argue that systematically benchmarking parameter settings across many combinations of tissue type and technology is beyond the scope of this manuscript. We have now generated three new supplementary figures providing evidence for robust feature selection across local outlier and artifact detection methods.

For local outliers in sequencing-based platforms, we demonstrate that local outlier detection is robust when using 2nd or greater order of neighbors. Using only 1st order neighbors (k=6 in Visium) results in the spurious identification of otherwise high-quality spots due to the low neighborhood size (i.e., increased variability amplifies sensitivity to small changes). We demonstrate this through several lines of evidence by 1) investigating the distributions of local z-score; 2) profiling the identified outliers. First, the distributions of local z-scores using 2nd-5th order neighbors are remarkably similar; in contrast, the distribution using 1st order neighbors is more fluctuating for all three QC metrics (SFig 2. A-C). The fluctuation suggests that due to the small sample size, the local z-score using 1st order neighborhoods is over-sensitive to local changes of natural biological variation. As a result, there are substantially more local outliers identified using 1st order neighbors than those using 2nd-5th order neighbors; whereas local outliers detected by 2nd-5th order neighbors are highly overlapping (SFig2. D-G). Finally, among the additional outliers detected using 1st order neighbors, many are spurious: these "outliers" contain markedly higher library size and unique genes, and lower mitochondrial ratio, all signs of high quality observations. (SFig2. H-J) In summary, ***this justifies our recommendation to use 3rd order neighbors and demonstrates that SpotSweeper's local outlier detection is highly robust above 1st order neighbors***.



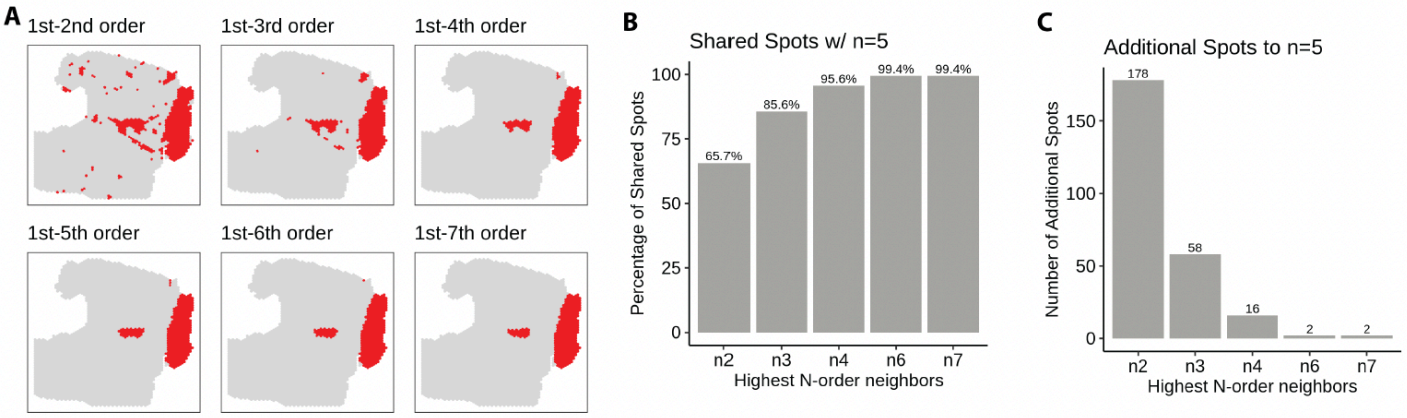
Supplementary Figure S12: Local outlier kNN parameter selection for imaging-based platforms. Cumulative distributions of local z-scores for (A) library size and (B) detected genes across various parameter settings. Dotted lines represent 3 z-score thresholds for local outlier detection and inset plots show the cumulative distribution functions zoomed in at these threshold points. (C) Bar plots showing the number of library size local outliers across parameter settings. (D) Number spots found in addition to the local outliers detected by our recommended setting of k=50 nearest neighbors. (E) Percentage of library size outliers across neighborhood sizes shared with k=50 nearest neighbors. The distributions of (F) library size and (G) detected genes across parameter settings.

For local outliers in imaging-based platforms, we demonstrate that local outlier detection is robust when using 50 or greater nearest neighbors (SFig. 12). Using a small number of neighbors (e.g., k=10 in Xenium) results in the spurious identification of otherwise high-quality spots due to the low neighborhood size (i.e., increased variability amplifies sensitivity to small changes). We demonstrate this using the same metrics as above. In summary, ***this justifies our recommendation to use k=50 nearest neighbors for imaging-based platforms***.



Supplementary Figure S12: Local outlier kNN parameter selection for imaging-based platforms. Cumulative distributions of local z-scores for (A) library size and (B) detected genes across various parameter settings. Dotted lines represent 3 z-score thresholds for local outlier detection and inset plots show the cumulative distribution functions zoomed in at these threshold points. (C) Bar plots showing the number of library size local outliers across parameter settings. (D) Number spots found in addition to the local outliers detected by our recommended setting of k=50 nearest neighbors. (E) Percentage of library size outliers across neighborhood sizes shared with k=50 nearest neighbors. The distributions of (F) library size and (G) detected genes across parameter settings.

For ***hangnail artifact detection***, we demonstrate both qualitatively and quantitatively that artifact detection is robust when using 1st-4th order neighbors or greater for multiscale variance calculations. We show that, although still successfully identifying the artifact, using 1st-2nd and 1st-3rd order neighbors results in noisy artifact detection with many additional spots being detected sparsely throughout the tissue section (Fig 3A). Increasing the number of scales used (i.e., 1st-4th through 1st-7th order neighbors) progressively denoises artifact detection, while accurate detection of artifact itself peaks at the 1st-5th order neighbors (our recommended parameter setting). Using 1st-6th and 1st-7th order neighbors were highly concordant with our recommended setting, resulting in > 99% of identified artifact spots being shared with those of 1st-5th order neighbors (SFig 3B-C). Thus, ***these data demonstrate that hangnail artifact detection is highly robust when using 1st-4th order neighbors or greater, and strongly justifies our recommended multiscale parameter settings.***



Supplementary Figure S17: Impact of multiscale parameter selection of hangnail identification. (A) Spot plots showing detected artifacts (red) across a range of multiscale parameter settings (i.e., 1st-2nd through 1st-7th order neighborhoods). (B) The percentage of annotated spots overlapping with our recommended multiscale parameter setting (1st-5th order neighbors). (C) The number of additional annotated spots found in comparison with our recommendation parameter setting.

Minor comments:

1、The authors mention the selection of parameter k in the manuscript but do not adequately explain its specific impact on the results. We suggest that the authors discuss in detail the criteria for the selection of k, including how the value of k affects the performance and accuracy of SpotSweeper. Additionally, if the value of k has been adjusted and optimized for different datasets, the authors should provide the rationale and a detailed description of the optimization process. This will help readers understand the importance of parameter selection on algorithm performance and increase the transparency and reproducibility of the method.

Please see above responses.

2、The manuscript uses specific data labels to evaluate the performance of SpotSweeper but does not detail the source and reliability of these labels. We recommend that the authors describe in detail the process of obtaining data labels, including the generation, validation, and quality control measures of the labels. If possible, the authors should also provide additional data or analysis to

support the accuracy and reliability of these labels, which is crucial for ensuring the validity of the research results.

The source of all manual and/or ground truth labels are provided in the source data described in the Data Availability section. Otherwise, unsupervised clustering methods (Banksy) were used to find spatial domains, as currently described in the Methods.

Reviewer #2 (Remarks to the Author):

The authors have addressed all my concerns and have made substantial imporvements to the manuscript by including additional datasets, methodology, and software features. I recommend accepting this manuscript for publication.

We would like to thank this reviewer for their valuable contributions to our manuscript!

Reviewer #2 (Remarks on code availability):

As previously stated, the code has been peer reviewed independently and is of high quality.

Reviewer #3 (Remarks to the Author):

The revisions to the manuscript pointed out in my review and the two other reviews were substantial. In particular, they added more datasets and technologies to the benchmarking, including different technology types beyond 10x Visium. They also outlined in more detail the methods used for QC. They also added more methods to testing the downstream results before and after QC (i.e., more domain detection methods). Overall, this is a much-improved manuscript and the tool SpotSweeper will aid researchers in the completion of SRT studies.

We would like to thank this reviewer for their valuable contributions to our manuscript!

Reviewer #3 (Remarks to the Author):

The authors have answered adequately the remaining comments.

We greatly appreciate that the reviewer found our revision addressed the remaining comments.