nature portfolio

Peer Review File

Charting Unknown Metabolic Reactions by Mass Spectrometry-Resolved Stable-isotope Tracing Metabolomics

Corresponding Author: Professor Zheng-Jiang Zhu

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
Gao et al developed a new tool (IsoNet) for the identification of unknown metabolites in LCMS datasets with the help of stable-isotope labeling. Usually, Mass Isotopomer Distribution (MID) patterns between two metabolites of the same reaction are similar. Built on this premise the authors developed an isotopologue similarity networking algorithm that identifies related metabolites and further elucidates similarities of previously annotated and unknown metabolites based on mass differences. To confirm this technique, the authors analyzed a demo dataset, which is unfortunately not part of the publication. With the help of this dataset they identified multiple closely related unknown metabolites of glutathione, validated them with measurements of synthesized standards, and even found a new transsulfuration reaction which is on its own of importance. The manuscript and the developed tool are of high relevance for the mass-spec and metabolism community.

Although the manuscript in most parts clearly describes the methods, it is not clear how the authors calculated reaction-pair distances. In agreement with others, they propose that reaction-paired metabolites tend to share similar isotopologue patterns. To demonstrate this principle they group annotated metabolites and pairwise calculate their respective “reaction steps” based on a curated KEGG reaction network using “a distance function” (see Methods lines 599-601 and 639-643). However, it is unclear what specific curated reaction network was used, what exactly is meant with “reaction steps” and how the distance was calculated.

The proposed algorithm takes advantage of a complex condition based distance calculations. I am wondering if and how these complex calculations outperform more general concepts such as dynamic programming (e.g. alignment) that do not require a prior classification and nevertheless produce normalized and comparable distances.

Furthermore, while the authors present fully analyzed XCMS, MetDNA, and MetTracer data, to test their algorithm MS2 data is required but was not provided and the analyzed data does not match their IsoNet demo data structure (their Github page does not specify required data structures or preprocessing steps). As mentioned above, the MS raw data is missing and thus presented results cannot be reproduced. Upon publication all these data should be made publicly available and all data structures should be defined in detail to allow users to use the tool with their own preprocessing pipeline.

Regarding the application of IsoNet, the authors report the discovery of a novel transsulfuration reaction converting glutathione into γ-glutamyl-seryl-glycine. This is an intriguing finding that significantly extends our understanding of glutathione metabolism. Using a combination of stable-isotope tracing and high-resolution mass spectrometry, the authors propose a compelling mechanism for how glutathione is converted into γ-glutamyl-seryl-glycine.

The experiments convincingly demonstrate the presence of this new metabolite and suggest it is synthesized directly from glutathione rather than being formed de novo from serine, glycine, and glutamate. However, further investigation is needed to elucidate the biochemistry and physiological relevance of this pathway.

Homologous Enzyme Search: The proposed reaction (R03923) is catalyzed by a bacterial enzyme. To strengthen their findings, the authors should perform a sequence similarity analysis to identify whether a homologous enzyme exists in the human genome.

Validation in a Glutathione-Deficient Model: To confirm that γ-glutamyl-seryl-glycine is synthesized directly from glutathione, the authors could utilize cells deficient in glutathione biosynthesis, such as those with a GCLC knockout. This model would also provide valuable insights into the physiological relevance of γ-glutamyl-seryl-glycine by allowing the study of its effects in the absence of endogenous production.

While the presence of γ-glutamyl-seryl-glycine is convincingly demonstrated, addressing these points would provide a more comprehensive understanding of the biochemistry and enhance the manuscript's impact.

(Remarks on code availability)

Reviewer #2

(Remarks to the Author)
I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

(Remarks on code availability)

Reviewer #3

(Remarks to the Author)
The authors describe the use of IsoNet for the discovering of unknown metabolic reactions using SIRM MS setups. An example for its application was shown with the discovery of gamma-glutamyl-serine-glycine in cell cultures and animal models.

In general, the IsoNet seems to be very attractive for the discovery of new metabolites, one weakness to the system might be due to the similarity networking, promoting the discovery of 'large' and 'composite' molecules. With composite meaning that they consists of multiple "smaller" molecules. Can the authors comment on this? For instance, I would expect that the networking approach of metabolites like glycolytic intermediates, Krebs cycle, amino acids ... will most likely not discover novel compounds and networking is probably much more difficult to link metabolites with each other.

I also believe that looking at lipids would be beneficial (and impactful) for the (novel) discovery of related (isotope labeling) species. Did the authors look at lipid species?

How do the authors take into account that from a network point of view and the respective MS method (positive versus negative ionisation mode), key molecules might be overlooked. Does IsoNet take into account that ionisation mode limits the detectability of certain species in the network when looking for similarities in isotope patterns?

Were there any other novel metabolites discovered? Or was the y-Glu-Ser-Gly the only example?

(Remarks on code availability)
NA

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
Dear Authors,
Thank you for addressing my previous comments. I have carefully reviewed the revised manuscript and appreciate the thorough revisions made by the authors. I am pleased to see that all my concerns have been appropriately addressed, resulting in a significant improvement in the manuscript.
Regarding the biological relevance of IsoNet, the newly included gene-silencing experiments add substantial value to the study. These findings further support the conclusion that γ-glutamyl-seryl-glycine is indeed derived from glutathione.

(Remarks on code availability)

Reviewer #2

(Remarks to the Author)

I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

(Remarks on code availability)
The IsoNet software provides a thorough README with multiple demo datasets. The installation with docker was easy. The IsoNet usage worked with the demo datasets and the results were reproducible.

Reviewer #3

(Remarks to the Author)
the authors have addressed my comments adequately

(Remarks on code availability)

Reviewer #4

(Remarks to the Author)
This manuscript by Gao et al describe an approach to the discovery of novel metabolites and novel metabolic reactions using an technique they call isotopologue similarity networking (IsoNet), to identify unknown metabolic reactions through mass spectrometry-resolved stable-isotope tracing metabolomics. Using IsoNet, they constructed the isotopologue similarity networks for labeled metabolites both in vitro and in vivo. Using this approach, they discovered unknown metabolic reactions. Focusing on glutathione metabolism, they describe a heretofore uncharacterized transsulfuration reaction of glutathione, resulting in γ-Glu-Ser-Gly. This work expands the techniques for identifying as yet unexplored metabolic pathways that govern cellular metabolism, energetics, and signaling.

The bulk of the experiments were conducted in vitro using 293T cells, and these data form the core of the data in the manuscript. The authors assess endogenous levels of novel metabolites and labeling patterns after incubating cells with ([U-13C]-glutamine, [U-13C]-glucose, and [U-13C]-acetate The authors have thoroughly addressed the limitations and caveats of IsoNet, and they provide valuable comparisons to other technologies for the identification of novel metabolites

For in vivo experiments, the authors assessed γ-Glu-Ser-Gly in a variety of tissues of adult mice and found the spleen to exhibit the highest levels of the metabolite. They then administered [U-13C]-glucose in the diet to adult mice and harvested the liver 24 hours later, indicating that γ-Glu-Ser-Gly was synthesized endogenously and suggesting a direct product of glutathione metabolism. This conclusion was supported by data using the BSO, an inhibitor of GCS. These studies were further supported by in vitro studies in 293T cells in which GCS was knocked down and resulted in a reduction in γ-Glu-Ser-Gly levels.

Overall, I find this study to be carefully controlled and the interpretations to be cautious. The interpretations are balanced, and the authors have placed the results in the context of other technologies that are complementary. This is a valuable addition to the literature.

(Remarks on code availability)

Response to the reviewers:

The authors would like to thank the reviewers for the helpful comments. We feel these comments have strengthened the manuscript considerably.

Reviewer #1

General comments: *"Gao et al developed a new tool (IsoNet) for the identification of unknown metabolites in LCMS datasets with the help of stable-isotope labeling. Usually, Mass Isotopomer Distribution (MID) patterns between two metabolites of the same reaction are similar. Built on this premise the authors developed an isotopologue similarity networking algorithm that identifies related metabolites and further elucidates similarities of previously annotated and unknown metabolites based on mass differences. To confirm this technique, the authors analyzed a demo dataset, which is unfortunately not part of the publication. With the help of this dataset they identified multiple closely related unknown metabolites of glutathione, validated them with measurements of synthesized standards, and even found a new transsulfuration reaction which is on its own of importance. The manuscript and the developed tool are of high relevance for the mass-spec and metabolism community."*

Ans: We sincerely appreciate the reviewer's positive feedback. In the revised manuscript, we have made the demo datasets available, including stable-isotope tracing metabolomics data for **293T cells** (**Figure 2**) and **iBMDM cells** (**Supplementary Figures 4** and **15)**. Specifically, the raw LC-MS data files generated in this study have been deposited in the National Omics Data Encyclopedia (https://www.biosino.org/node) under the accession code OEP00004597. Additionally, the IsoNet R package, instructions of data preparation, input data files for demo datasets, and detailed data processing workflow are available on GitHub (https://github.com/ZhuMetLab/IsoNet). IsoNet requires three input files: (i) a table with isotopologue patterns (.csv); (ii) a table with identification results (.csv); and (iii) MS2 data files (.msp). The relevant input files associated with the demo datasets reported in our study have also been deposited on GitHub.

We have added the availability of raw LC-MS data files and the demo datasets and in the revised manuscript as following:

"The raw LC-MS data files generated in this study have been deposited in the National Omics Data Encyclopedia (https://www.biosino.org/node) under the accession code OEP00004597. The IsoNet R package, instructions of data preparation, input data files for demo datasets, and detailed data processing workflow are available on GitHub (https://github.com/ZhuMetLab/IsoNet)."

Comment #1: *"Although the manuscript in most parts clearly describes the methods, it is not clear how the authors calculated reaction-pair distances. In agreement with others, they propose that reaction-paired metabolites tend to share similar isotopologue patterns. To demonstrate this principle they group annotated metabolites and pairwise calculate their respective "reaction steps" based on a curated KEGG reaction network using "a distance function" (see Methods lines 599-601 and 639-643). However, it is unclear what specific curated reaction network was used, what exactly is meant with "reaction steps" and how the distance was calculated."*

Ans: We thank a lot for the reviewer's comment. The calculation of reaction steps is based on the knowledge-based metabolic reaction network (MRN) curated from the KEGG database reported in our previous work (Zhou *et al.*, ***Nature Communications***, 2022, https://doi.org/10.1038/s41467-022-34537-6). This reaction network comprises 6,397 metabolites and 8,129 reactions (**Supplementary Fig. 2e**). The reaction step represents the minimum number of metabolic reactions occurred between two metabolites in the metabolic reaction network. Calculation of reaction distance between two metabolites was achieved by using a distance calculation function in R package "igraph" (v1.2.9), which determines the number of connections between any two nodes within the network. By inputting KEGG IDs, the number of reaction steps between two metabolites can be derived from the MRN network. For example, the reaction step between UDP-glucose (C00029) and UDP-glucuronate (C00167) is one step, involving the reaction R00286 (**Figure R1a**). Similarly, the reaction step between malate (C00149) and aspartate (C00049) is two steps, corresponding to two reactions, R00343 and R00357 (**Figure R1b**).

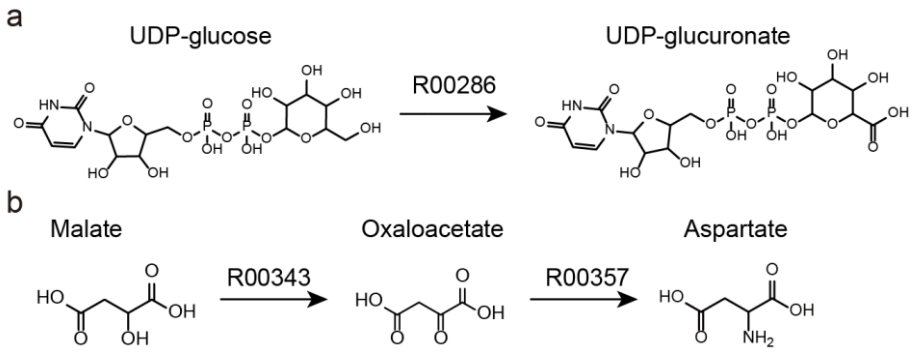


Figure R1. Illustration of the reaction steps between UDP-glucose and UDP-glucuronate (a), and between malate and aspartate (b).

We have revised the manuscript to clarify as following:

"*For known reactions, their reaction steps were calculated based on the knowledge-based metabolic reaction network curated from the KEGG database reported in our previous work of MetDNA. This reaction network comprises 6,397 metabolites and 8,129 reactions (**Supplementary Fig. 2e**). The reaction step represents the minimum number of metabolic reactions occurred between*

two metabolites in the metabolic reaction network. Calculation of reaction distance between two metabolites was realized by using a distance calculation function in R package "igraph" (v1.2.9), which determines the number of connections between any two nodes within the network. By inputting KEGG IDs, the number of reaction steps between two metabolites can be derived from the MRN network."

Comment #2: *"The proposed algorithm takes advantage of a complex condition based distance calculations. I am wondering if and how these complex calculations outperform more general concepts such as dynamic programming (e.g. alignment) that do not require a prior classification and nevertheless produce normalized and comparable distances."*

Ans: We thank the reviewer for the excellent idea. After a careful comparison, we believe that the principles of distance calculation for dynamic programming and IsoNet developed in this study are highly similar. For example, for the isotopologue pattern similarity of ADP-ribose and ATP, the dynamic programming algorithm scored 0.77 and IsoNet scored 0.78 (**Figure R2a**). Similarly, when calculating the isotopologue pattern similarity of glutathione and S-lactoylglutathione, the dynamic programming algorithm scored 0.77 and IsoNet scored 0.79 (**Figure R2b**).

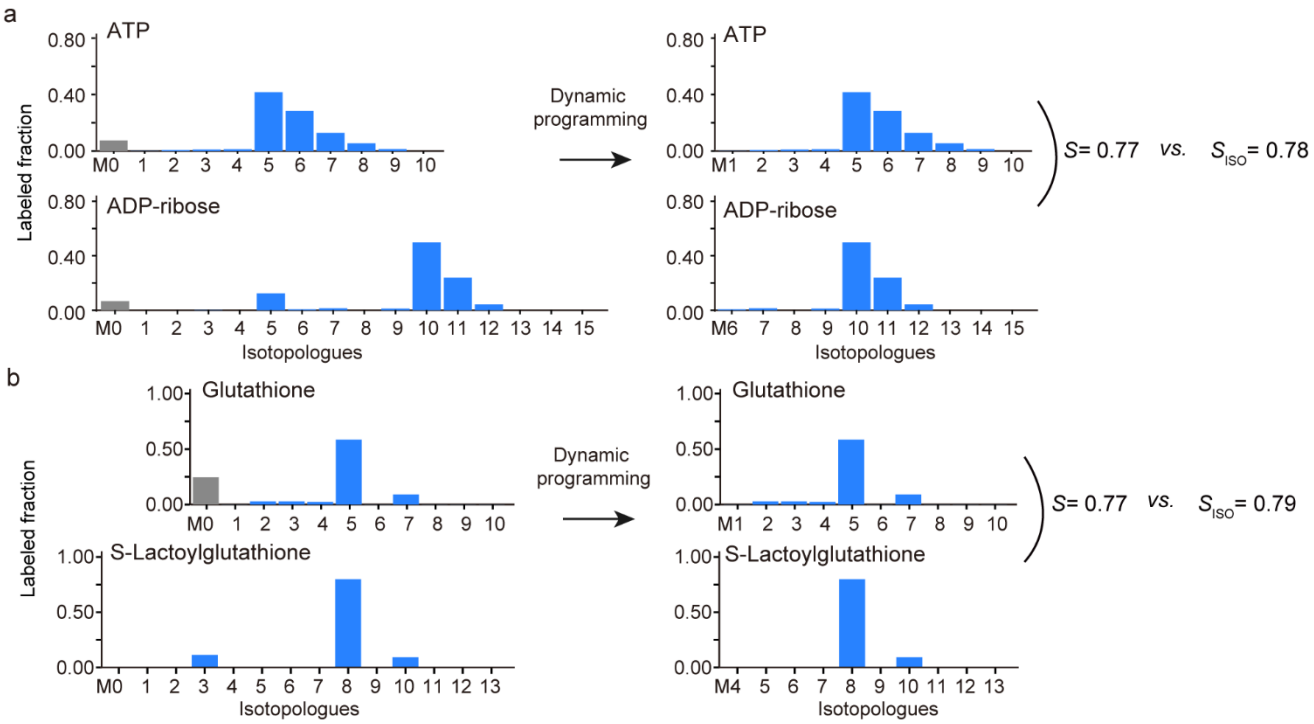


Figure R2. Calculation of isotopologue pattern similarity for metabolite pairs using the dynamic programming.

Specifically, the dynamic programming uses a point-to-point alignment approach to obtain the optimal alignment segments for similarity computations. In comparison, IsoNet employs the

biochemical principles to directly align isotopologue pattern segments (defined as "motifs"; see details in **Methods**). The motif-based alignment offers better interpretability since the isotopologue pattern often represents a specific substructure of metabolites from metabolic reactions. For example, purine nucleotides display distinct stepwise isotopologue patterns (e.g., isotopologues *M*+6 to *M*+10 of ATP; **Figure R2a** and **Supplementary Fig. 3b**), which corresponds to the *de novo* synthesis of purine using glucose. Therefore, we incorporated the concept of motifs into the IsoNet algorithm to improve the interpretation of substructure information embedded in isotopologue patterns.

We have revised the discussion section in the manuscript as following:

"*IsoNet calculates isotopologue pattern similarity based on biochemical principles which directly aligns isotopologue pattern segments (defined as "motifs"). The motif-based alignment offers better interpretability compared with point-to-point alignment approaches such as dynamic programing, since the isotopologue pattern represents a specific substructure of metabolites from metabolic reactions. However, in principle, isotopologue pattern similarity calculation using IsoNet is highly similar to conventional dynamic programming.*"

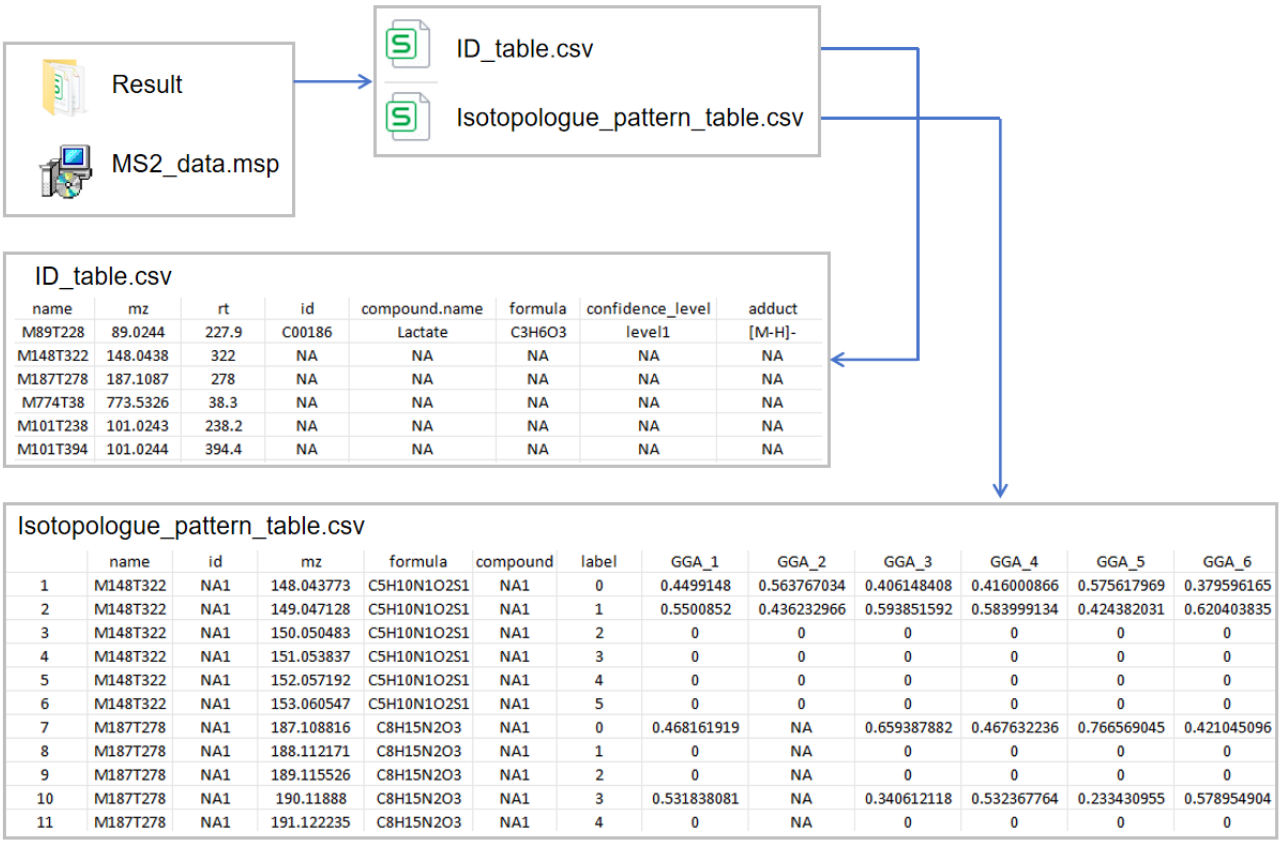
Comment #3: *"Furthermore, while the authors present fully analyzed XCMS, MetDNA, and MetTracer data, to test their algorithm MS2 data is required but was not provided and the analyzed data does not match their IsoNet demo data structure (their Github page does not specify required data structures or preprocessing steps). As mentioned above, the MS raw data is missing and thus presented results cannot be reproduced. Upon publication all these data should be made publicly available and all data structures should be defined in detail to allow users to use the tool with their own preprocessing pipeline."*

Ans: We thank a lot for the reviewer's comments. In the revised manuscript, we have added the data availability. The raw LC-MS data files generated in this study have been deposited in the National Omics Data Encyclopedia (https://www.biosino.org/node) under the accession code OEP00004597. The IsoNet R package, instructions of data preparation, input data files for demo datasets, and detailed data processing workflow are available on GitHub (https://github.com/ZhuMetLab/IsoNet).

Specifically, IsoNet requires three input files: i) a table with isotopologue patterns (.csv) ; ii) a table with identification results (.csv); iii) MS2 data files (.msp) (see **Figure R3**). The input data files for demo datasets reported in the study are now available on GitHub. The data structures and formats of input files should follow:

1. The isotopologue patterns of known and unknown labeled metabolites in the isotopologue pattern table are used to calculate isotopologue pattern similarity. The first column should be an index, and the remaining column names must be: "name", "id", "mz", "formula", "compound", "label", and sample names.

2. The identification results are used for annotating unknown reactions. The column names must be: "name", "mz", "rt", "id", "compound.name", "formula", "confidence_level", and "adduct". The definitions of confidence level and adduct forms are the same as those used in MetDNA2. level 1: metabolites annotated using in-house metabolite standards with three orthogonal properties (i.e., MS1 + RT + MS/MS); level 2: metabolites annotated using two orthogonal properties from the standard MS/MS libraries without RT available (i.e., MS1 + MS/MS). "level1" or "level2" are considered known metabolites. The forms of adducts should be: "[M-H]-", "[M-H2O-H]-", "[M+CH3COO]-", etc. in negative mode and "[M+H]+", "[M-H2O+H]+", "[M+Na]+", "[M+NH4]+", etc. in positive mode. Multiple annotation is separate them with `;`.



ID_table.csv

name	mz	rt	id	compound.name	formula	confidence_level	adduct
M89T228	89.0244	227.9	C00186	Lactate	C3H6O3	level1	[M-H]-
M148T322	148.0438	322	NA	NA	NA	NA	NA
M187T278	187.1087	278	NA	NA	NA	NA	NA
M774T38	773.5326	38.3	NA	NA	NA	NA	NA
M101T238	101.0243	238.2	NA	NA	NA	NA	NA
M101T394	101.0244	394.4	NA	NA	NA	NA	NA

Isotopologue_pattern_table.csv

	name	id	mz	formula	compound	label	GGA_1	GGA_2	GGA_3	GGA_4	GGA_5	GGA_6
1	M148T322	NA1	148.043773	C5H10N1O2S1	NA1	0	0.4499148	0.563767034	0.406148408	0.416000866	0.575617969	0.379596165
2	M148T322	NA1	149.047128	C5H10N1O2S1	NA1	1	0.5500852	0.436232966	0.593851592	0.583999134	0.424382031	0.620403835
3	M148T322	NA1	150.050483	C5H10N1O2S1	NA1	2	0	0	0	0	0	0
4	M148T322	NA1	151.053837	C5H10N1O2S1	NA1	3	0	0	0	0	0	0
5	M148T322	NA1	152.057192	C5H10N1O2S1	NA1	4	0	0	0	0	0	0
6	M148T322	NA1	153.060547	C5H10N1O2S1	NA1	5	0	0	0	0	0	0
7	M187T278	NA1	187.108816	C8H15N2O3	NA1	0	0.468161919	NA	0.659387882	0.467632236	0.766569045	0.421045096
8	M187T278	NA1	188.112171	C8H15N2O3	NA1	1	0	NA	0	0	0	0
9	M187T278	NA1	189.115526	C8H15N2O3	NA1	2	0	NA	0	0	0	0
10	M187T278	NA1	190.11888	C8H15N2O3	NA1	3	0.531838081	NA	0.340612118	0.532367764	0.233430955	0.578954904
11	M187T278	NA1	191.122235	C8H15N2O3	NA1	4	0	NA	0	0	0	0

Figure R3. Overview of data preparation for data analysis using IsoNet.

Here, we also provide demo code to run IsoNet using the input files for the 293T cell demo dataset (Figure 2). Again, the input files can be downloaded from the GitHub (https://github.com/ZhuMetLab/IsoNet).

```
library(IsoNet)
wd <- '.'
setwd(wd)
constructMIDNet(tracer_table_file = "./Result/Isotopologue_pattern_table.csv",
```

```
ms2_file = "ms2_data.msp",
id_table = "./Result/ID_table.csv",
sample_names = c("GGA_1","GGA_2","GGA_3","GGA_4","GGA_5","GGA_6"),
dir_path = '.',
mid_cutoff = 0.7,
mid_fc = 20,
mid_isoDegree = 0.1,
mid_min_motifLen = 0.5,
ms2_score_cutoff = 0.5,
mid_max_motif = 1000L,
ignore_max = TRUE,
scoring_approach = 'gnps',
mass_diff_freq_cut_off = 4L)
```

Comment #4: *"Regarding the application of IsoNet, the authors report the discovery of a novel transsulfuration reaction converting glutathione into γ-glutamyl-seryl-glycine. This is an intriguing finding that significantly extends our understanding of glutathione metabolism. Using a combination of stable-isotope tracing and high-resolution mass spectrometry, the authors propose a compelling mechanism for how glutathione is converted into γ-glutamyl-seryl-glycine. The experiments convincingly demonstrate the presence of this new metabolite and suggest it is synthesized directly from glutathione rather than being formed de novo from serine, glycine, and glutamate. However, further investigation is needed to elucidate the biochemistry and physiological relevance of this pathway. Homologous Enzyme Search: The proposed reaction (R03923) is catalyzed by a bacterial enzyme. To strengthen their findings, the authors should perform a sequence similarity analysis to identify whether a homologous enzyme exists in the human genome."*

Ans: We thank for the reviewer's positive comments regarding the application of IsoNet. We have taken the reviewer's comment and validated the biosynthesis of γ-Glu-Ser-Gly in a glutathione-deficient cell model to elucidate the biochemistry and physiological relevance of this pathway (please see **Reponses to Comment #5**).

We have also considered the reviewer's comment regarding the homologous enzyme search. The reaction R03923 occurs in bacteria and is catalyzed by the tRNA thiolation enzyme ThiI, which is involved in the sulfur modification of tRNA. However, a sequence similarity analysis did not reveal any particularly convincing candidates in the human genome.

Nevertheless, in human, there is a set of tRNA thiolation enzymes such as mitochondrial tRNA-specific 2-thiouridylase 1 (MTU1) and cytoplasmic tRNA 2-thiolation protein 1/2 (CTU1/2), catalyzing

the same reaction as in bacteria for sulfur modification. More importantly, the human tRNA-modification genes remain largely unexplored. Approximately 23% of the human tRNA modification genes (cytosolic and mitochondrial combined) remain unknown (Crécy-Lagard ***et al.***, ***Nucleic Acids Research***, 2019, https://doi.org/10.1093/nar/gkz011). Given that glutathione is a critical cellular sulfur donor, it is highly plausible that these unidentified genes and enzymes may also be involved in tRNA thiolation and the transsulfuration reaction of glutathione. To uncover these undiscovered tRNA thiolation enzymes related to glutathione, we believe new proteomics technologies are required, such as peptide-centric local stability assay (Li *et al.*, ***Nature Methods***, 2024, https://doi.org/10.1038/s41592-024-02553-7).

We have taken the reviewer's comments and have revised the discussion section in the manuscript as following:

"*In human, there is a set of tRNA thiolation enzymes such as mitochondrial tRNA-specific 2-thiouridylase 1 (MTU1) and cytoplasmic tRNA 2-thiolation protein 1/2 (CTU1/2), catalyzing the sulfur modification on tRNAs. More importantly, the human tRNA-modification genes remain largely unexplored. Approximately 23% of the tRNA modification genes in human remain unknown. Given that glutathione is a critical cellular sulfur donor, it is highly plausible that these unidentified genes and enzymes may also be involved in tRNA thiolation and the transsulfuration reaction of glutathione.*"

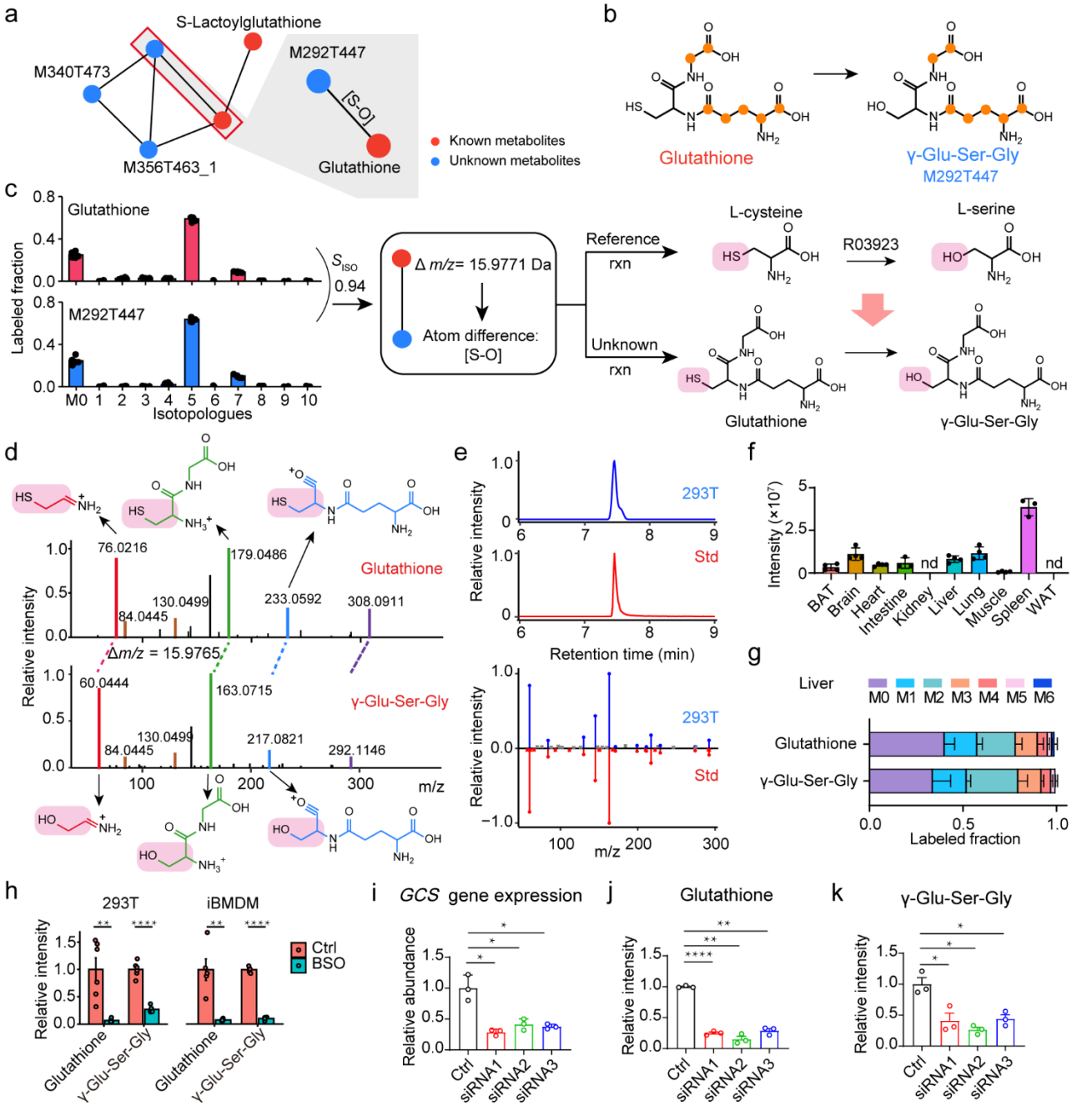
Comment #5: *"Validation in a Glutathione-Deficient Model: To confirm that γ-glutamyl-seryl-glycine is synthesized directly from glutathione, the authors could utilize cells deficient in glutathione biosynthesis, such as those with a GCLC knockout. This model would also provide valuable insights into the physiological relevance of γ-glutamyl-seryl-glycine by allowing the study of its effects in the absence of endogenous production. While the presence of γ-glutamyl-seryl-glycine is convincingly demonstrated, addressing these points would provide a more comprehensive understanding of the biochemistry and enhance the manuscript's impact."*

Ans: We greatly appreciate the reviewer's constructive comment. In the revised manuscript, we have addressed the reviewer's suggestion and generated gamma-glutamylcysteine synthetase (GCS, also known as glutamate-cysteine ligase, GCL) knockdown 293T cells deficient in glutathione biosynthesis using RNA interference (revised **Fig. 3i**). The LC-MS analyses showed that GCS knockdown dramatically decreased intracellular levels of glutathione (revised **Fig. 3j**). Accordingly, the intracellular levels of γ-Glu-Ser-Gly were also significantly reduced (revised **Fig. 3k**). Altogether, these results confirm that γ-Glu-Ser-Gly is synthesized directly from glutathione.

We have added the results and method description in the revised manuscript as following:

"*To confirm that γ-Glu-Ser-Gly is synthesized directly from glutathione, we further generated GCS knockdown 293T cells deficient in glutathione biosynthesis using RNA interference (**Fig. 3i**). As results*

*showed in **Fig. 3j** and **3k**, GCS knockdown dramatically decreased cellular levels of not only glutathione but also γ-Glu-Ser-Gly, demonstrating the direct link between γ-Glu-Ser-Gly biosynthesis and endogenous glutathione level."*



Revised Figure 3. Glutathione-associated isotopologue similarity subnetwork reveals new metabolic reactions. (**a**) The glutathione-associated isotopologue similarity subnetwork in positive ionization mode. (**b**) The unknown metabolic reaction between glutathione and M292T447. Orange points represent the 13C-labeled atoms. (**c**) Elucidation of the metabolic reaction between glutathione and M292T447 (γ-Glu-Ser-Gly). (**d**) Interpretation of MS/MS spectra between glutathione and γ-Glu-Ser-Gly. (**e**) Validation of γ-Glu-Ser-Gly using chemical standard (Std): retention time match (top) and

MS2 spectral match (bottom). (**f**) The relative levels of γ-Glu-Ser-Gly in various mouse tissues (n=3-4) quantified by LC−MS. (**g**) Isotopologue pattern similarity between glutathione and γ-Glu-Ser-Gly in mouse liver tissue. Mice were administered with [U-13C]-glucose for 24 h (n=4). (**h**) Intracellular γ-Glu-Ser-Gly levels in 293T and iBMDM cells after the treatment of buthionine sulfoximine (BSO; 100 μM) for 12 h or 10 h, respectively (n=5-6). (**i**) Quantitative real-time PCR analyses of mRNA levels of GCS gene expression in control group and siRNA-transfected 293T cells (n=3). (**j,k**) Intracellular levels of glutathione (**j**) and γ-Glu-Ser-Gly (**k**) in control group and the siRNA-transfected 293T cells (n=3). P-values were determined by a two-tailed Student's t-test. **, p-value < 0.01; ****, p-value < 0.0001. Values represent means ± SD.

Reviewer #2

General comments: *"I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts."*

Ans: We thank for the reviewer's input during the peer review of the manuscript.

Reviewer #3

Comment #1: *"The authors describe the use of IsoNet for the discovering of unknown metabolic reactions using SIRM MS setups. An example for its application was shown with the discovery of gamma-glutamyl-serine-glycine in cell cultures and animal models. In general, the IsoNet seems to be very attractive for the discovery of new metabolites, one weakness to the system might be due to the similarity networking, promoting the discovery of 'large' and 'composite' molecules. With composite meaning that they consists of multiple "smaller" molecules. Can the authors comment on this? For instance, I would expect that the networking approach of metabolites like glycolytic intermediates, Krebs cycle, amino acids ... will most likely not discover novel compounds and networking is probably much more difficult to link metabolites with each other."*

Ans: We greatly appreciate the reviewer's positive feedback on the IsoNet approach. We agree that IsoNet is particularly effective at discovering "large" and "composite" unknown metabolites. This is because large, composite metabolites with multiple substructures tend to exhibit diverse isotopologue patterns in stable-isotope labeling. Consequently, the motif-based isotopologue pattern similarity networking is more likely to connect these metabolites. Additionally, the diverse isotopologue patterns facilitate the discovery and annotation of new metabolites and related biochemical reactions. The discovery of glutathione-associated new metabolites and reactions serves as a prominent example.

For small metabolites, such as glycolytic intermediates and amino acids, IsoNet can also connect them through isotopologue pattern similarity networking (see examples in **Fig. 1c** and **1d**). However, due to the lack of structural motifs, the annotation of unknown metabolites within these subnetworks remain a challenge. We anticipate that the annotation functions within IsoNet can be further improved to continue filling gaps in the metabolic network. We have added the following discussion in the revised manuscript.

"In addition, IsoNet tends to discover large and composite unknown metabolites. This is because that large, composite metabolites with multiple substructures tend to display more diverse isotopologue patterns in the stable-isotope labeling. Consequently, the motif-based isotopologue pattern similarity networking is more likely to connect these metabolites. Additionally, the diverse isotopologue patterns facilitate the discovery and annotation of new metabolites and related biochemical reactions. The discovery of glutathione-associated new metabolites and reactions serves as a prominent example. For small metabolites, such as glycolytic intermediates and amino acids, IsoNet can also connect them through isotopologue pattern similarity networking (see examples in Fig. 1c and 1d). However, due to the limited structural information that can be obtained from the isotopologue patterns or MS/MS spectra of small metabolites, the annotation of unknown metabolites within these subnetworks remains a challenge. We anticipate that the annotation functions within

IsoNet can be further improved to continue filling gaps in the metabolic network."

Comment #2: *"I also believe that looking at lipids would be beneficial (and impactful) for the (novel) discovery of related (isotope labeling) species. Did the authors look at lipid species?"*

Ans: We greatly appreciate the reviewer's comment. We agree that investigating lipids would be beneficial for the discovery of related lipid species. In our study, our original focus was primarily on metabolomics. As such, untargeted metabolomics protocols, including sample preparation and data acquisition, were applied. Under the experimental conditions, some lipids, such as phosphatidylcholines (PCs) and phosphatidylethanolamines (PEs), could be measured, but other major lipids, such as triglycerides and ceramides, were not covered as they are typically addressed by lipidomics methods.

To address the reviewer's comment, we employed IsoNet to discover unknown lipid species in the dataset of 293T cells (**Figure R4**). Classes of PC and PE lipids were isotopically labeled and clustered into two large isotopologue similarity subnetworks by IsoNet (**Figure R4a**). Within the PC lipid subnetwork, we annotated an unknown feature (M777T592) as an oxidized phospholipid (OxPL) (**Figure R4b**). To annotate this unknown lipid species, we examined the metabolite pair of PC(34:2) and M777T592, which has an isotopologue similarity score (S_{ISO}) of 0.70 and a mass difference ($\Delta m/z$) of 18.0124 Da. Matching the mass difference to the delta mass library suggests that the atom difference between the two metabolites is [2H+O]. Based on this atom difference, we retrieved the related reference reaction (R01082) from the reaction library, which involves the conversion of fumarate to malate. Inspired by this reference reaction, we annotated the unknown lipid species M777T592 as PC(34:1;O). The MS/MS spectrum of M777T592 also has the characteristic fragment ion with *m/z* 184.0734 Da, corresponding to the choline headgroup substructure (**Figure R4c**).

It is important to note that major lipid classes, such as triglycerides and ceramides, were not detected or labeled under our experimental conditions. We believe that with suitable experimental conditions designed for lipidomic analysis, IsoNet can also uncover more unknown lipid species. Furthermore, the reference reaction library within IsoNet was primarily designed for metabolites and may not fully account for lipid metabolism and reactions.

In the revised manuscript, we have added the following discussion in the revised manuscript.

"In our study, our original focus was primarily on metabolomics. We believe that with suitable experimental conditions designed for lipidomic analysis (e.g., optimized measurement settings and labeling protocols), IsoNet can also be applicable for the discovery of unknown lipid species. Furthermore, the reference reaction library within IsoNet was primarily designed for metabolites and should be further developed to account for lipid metabolism and reactions."

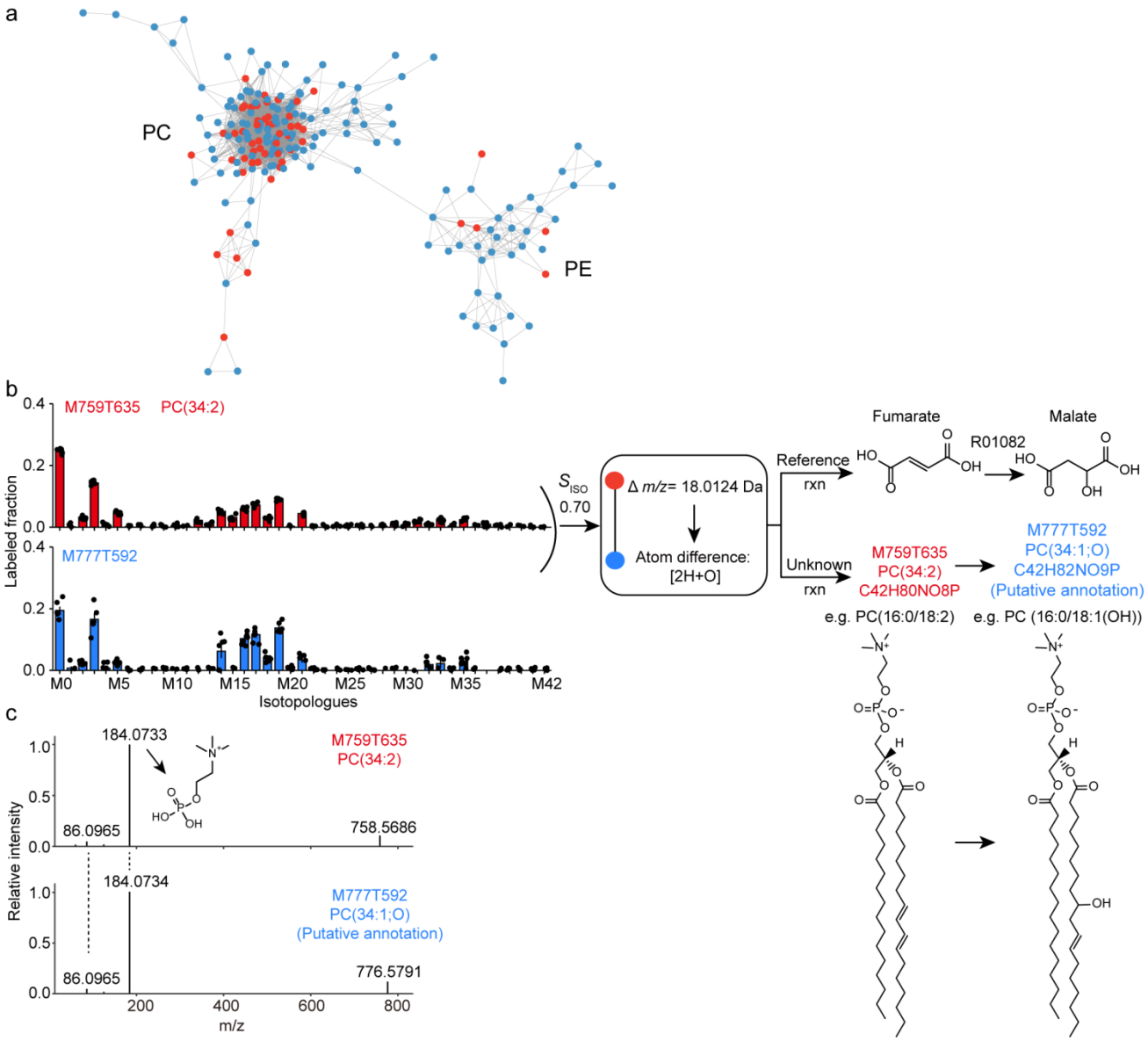


Figure R4. Lipid-associated isotopologue similarity subnetwork reveals potential unknown lipid species. (**a**) The lipid-associated isotopologue similarity subnetwork in positive ionization mode in 293T cells. (**b**) Elucidation of potential unknown lipid species M777T592. (**c**) Interpretation of MS/MS spectra of PC(34:2) and PC(34:1;O) (putative annotation).

Comment #3: *"How do the authors take into account that from a network point of view and the respective MS method (positive versus negative ionisation mode), key molecules might be overlooked. Does IsoNet take into account that ionisation mode limits the detectability of certain species in the network when looking for similarities in isotope patterns?"*

Ans: We greatly appreciate the reviewer's comment. We agree with the reviewer that the coverage of metabolites being measured is influenced by the MS ionization mode. To address this, we acquired data using both positive and negative ionization modes to improve the coverage of detected

metabolites. Then, IsoNet was employed to construct isotopologue similarity networks under both positive and negative modes, and subsequently merged the two polarities together. For examples, the isotopologue similarity networks of 293T cells (**Figure 2b** and **Supplemental Data 2**) and iBMDM cells labeled with multiple tracers (**Supplemental Figure 5** and **Supplemental Data 5**) were constructed by integrating datasets of both positive and negative ionization modes. To further expand coverage of metabolites in the isotopologue similarity network, methods such as chemical derivatization and multidimensional chromatography separation can also be employed.

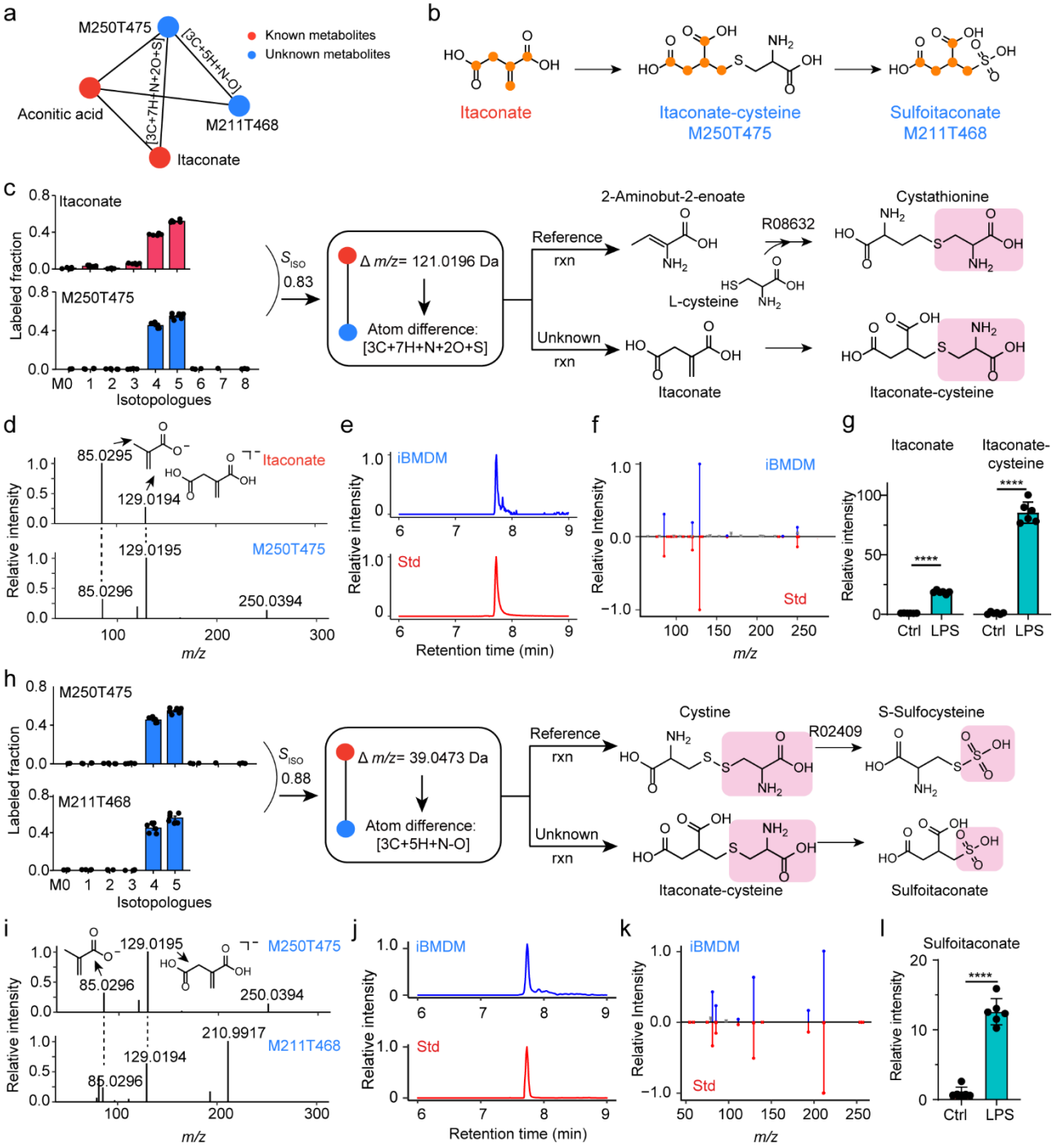
We have taken the reviewer's comment and have added the following discussion in the revised manuscript.

"*The coverage of metabolites and reactions in the isotopologue similarity network is influenced by the MS ionization mode. To address this, we therefore acquired data using both positive and negative ionization modes to improve the coverage of detected metabolites. Then, we used IsoNet and constructed an integrated isotopologue similarity network which combines data from individual ionization modes. To further expand coverage of metabolites in the isotopologue similarity network, methods such as chemical derivatization and multidimensional chromatography separation can also be employed.* "

Comment #4: *"Were there any other novel metabolites discovered? Or was the γ-Glu-Ser-Gly the only example?"*

Ans: Thanks a lot for the reviewer's valuable comment. To clarify, in the glutathione-associated subnetwork, seven new metabolites absent in the Human Metabolome Database (HMDB) were successfully discovered and validated (**Supplementary Table 1**). Among them, γ-Glu-3-sulfamoyl-Ala-Gly is a new structure that has not been reported in PubChem. In our study, we mainly focused on the validation of the biosynthesis pathway of γ-Glu-Ser-Gly because it represents a new transsulfuration reaction, which is important in its own right. The initial discovery of a transsulfuration reaction involving serine and cysteine dates back to the 1960s. This marks the identification of the second transsulfuration reaction known in mammals.

To further address the reviewer's comment, in the revised manuscript, we added two new metabolite examples related to itaconate (**Supplementary Fig. 15** and **Supplementary Data 8**). In LPS-stimulated iBMDM cells, we employed IsoNet to discover and annotate itaconate-cysteine and sulfoitaconate from the itaconate-associated isotopologue similarity subnetwork. **Both metabolites have not been reported in HMDB**.



Supplementary Figure 15. Itaconate-associated isotopologue similarity subnetwork disclose newly identified metabolites. (**a**) The itaconate-associated isotopologue similarity subnetwork in negative ionization mode of iBMDM cells with LPS stimulation. (**b**) The unknown metabolic reactions between itaconate, M250T475, and M211T468. Orange points represent the ^{13}C-labeled atoms. (**c**) Elucidation of the metabolic reaction between itaconate and M250T475 (itaconate-cysteine, putative structure). (**d**) Interpretation of MS/MS spectra of itaconate and M250T475. (**e-f**) Validation of itaconate-cysteine using an in-house synthesized chemical standard (Std) by mixing itaconate and cysteine standards: retention time match (**e**) and MS/MS spectral match (**f**). (**g**) Intracellular levels of itaconate and M250T475 (itaconate-cysteine, putative structure) in iBMDM cells after the treatment

of LPS (100 ng/ml) for 10 h, respectively (n=6). (**h**) Elucidation of the metabolic reaction between itaconate-cysteine (putative structure) and M211T468 (sulfoitaconate). (**i**) Interpretation of MS/MS spectra of itaconate-cysteine (putative structure) and sulfoitaconate. (**j-k**) Validation of sulfoitaconate using chemical standard (Std): retention time match (**j**) and MS/MS spectral match (**k**). (**l**) Intracellular levels of sulfoitaconate in iBMDM cells after the treatment of LPS (100 ng/ml) for 10 h (n=6). P-values were determined by a two-tailed Student's t-test. ****, p-value < 0.0001. Values represent means ± SD.

We have added the results in **Supplementary Fig. 15** and **Supplementary Data 8** in the revised manuscript as following:

*"**Itaconate-associated isotopologue similarity subnetwork disclose new metabolites.** Itaconate is a crucial anti-inflammatory metabolite, which modulates immune responses upon infections and the pathogenesis of inflammatory diseases. Next, we constructed an isotopologue similarity network in the lipopolysaccharide (LPS)-treated iBMDM cells, uncovering previously unknown reactions and metabolites related to inflammation (**Supplementary Data 8**). In the itaconate-associated isotopologue similarity subnetwork, we identified two new metabolites and associated reactions (**Supplementary Fig. 15a** and **15b**). The metabolites produced by these reactions were identified as itaconate-cysteine (**Supplementary Fig. 15c-g**) and sulfoitaconate (**Supplementary Fig. 15h-l**) by isotopologue similarity comparison, relevance to reference reactions, interpretation of MS/MS spectrum, and validation using chemical standards. Neither itaconate-cysteine nor sulfoitaconate is recoded in the KEGG or HMDB databases. Moreover, both itaconate-cysteine and sulfoitaconate were found to increase dramatically by LPS stimulation in iBMDM cells, which were even more profound than the elevation of itaconate (**Supplementary Fig. 15f** and **15k**), suggesting their relevance with inflammation. In conclusion, IsoNet facilitates the discovery of previously unknown metabolites linked to inflammation, offering valuable potentials for understanding anti-inflammatory processes."*

Response to the reviewers:

We appreciate the reviewers' positive comments toward publication.

Reviewer #1

General comments:

"Dear Authors,

Thank you for addressing my previous comments. I have carefully reviewed the revised manuscript and appreciate the thorough revisions made by the authors. I am pleased to see that all my concerns have been appropriately addressed, resulting in a significant improvement in the manuscript.

Regarding the biological relevance of IsoNet, the newly included gene-silencing experiments add substantial value to the study. These findings further support the conclusion that γ-glutamyl-seryl-glycine is indeed derived from glutathione."

Ans: Thanks a lot for the reviewer's positive comments towards publication.

Reviewer #2

General comments: *"I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts. The IsoNet software provides a thorough README with multiple demo datasets. The installation with docker was easy. The IsoNet usage worked with the demo datasets and the results were reproducible."*

Ans: We thank for the reviewer's input during the peer review of the manuscript. Thanks a lot for the reviewer's positive comments towards publication.

Reviewer #3

General comments: *"the authors have addressed my comments adequately."*

Ans: Thanks a lot for the reviewer's positive comments towards publication.

Reviewer #4

General comments: *"This manuscript by Gao et al describe an approach to the discovery of novel metabolites and novel metabolic reactions using an technique they call isotopologue similarity networking (IsoNet), to identify unknown metabolic reactions through mass spectrometry-resolved stable-isotope tracing metabolomics. Using IsoNet, they constructed the isotopologue similarity networks for labeled metabolites both in vitro and in vivo. Using this approach, they discovered unknown metabolic reactions. Focusing on glutathione metabolism, they describe a heretofore uncharacterized transsulfuration reaction of glutathione, resulting in γ-Glu-Ser-Gly. This work expands the techniques for identifying as yet unexplored metabolic pathways that govern cellular metabolism, energetics, and signaling.*

The bulk of the experiments were conducted in vitro using 293T cells, and these data form the core of the data in the manuscript. The authors assess endogenous levels of novel metabolites and labeling patterns after incubating cells with ([U-13C]-glutamine, [U-13C]-glucose, and [U-13C]-acetate The authors have thoroughly addressed the limitations and caveats of IsoNet, and they provide valuable comparisons to other technologies for the identification of novel metabolites

For in vivo experiments, the authors assessed γ-Glu-Ser-Gly in a variety of tissues of adult mice and found the spleen to exhibit the highest levels of the metabolite. They then administered [U-13C]-glucose in the diet to adult mice and harvested the liver 24 hours later, indicating that γ-Glu-Ser-Gly was synthesized endogenously and suggesting a direct product of glutathione metabolism. This conclusion was supported by data using the BSO, an inhibitor of GCS. These studies were further supported by in vitro studies in 293T cells in which GCS was knocked down and resulted in a reduction in γ-Glu-Ser-Gly levels.

Overall, I find this study to be carefully controlled and the interpretations to be cautious. The interpretations are balanced, and the authors have placed the results in the context of other technologies that are complementary. This is a valuable addition to the literature."

Ans: Thanks a lot for the reviewer's positive comments towards publication.