

# AccuCor2: Isotope Natural Abundance Correction for Dual-Isotope Tracer Experiments

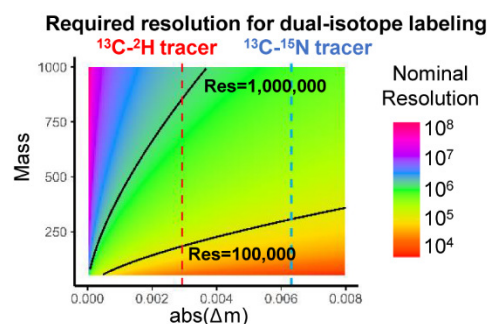
Yujue Wang<sup>1,2</sup>, Lance R. Parsons<sup>3</sup>, Xiaoyang Su<sup>1,2\*</sup>.

<sup>1</sup>Department of Medicine, Rutgers-Robert Wood Johnson Medical School, New Brunswick, NJ, USA

<sup>2</sup>Metabolomics Shared Resource, Rutgers Cancer Institute of New Jersey, New Brunswick, NJ, USA

<sup>3</sup>Lewis Sigler Institute for Integrative Genomics and Department of Chemistry, Princeton University, Princeton, NJ, USA

**ABSTRACT:** Stable isotope labeling techniques have been widely applied in the field of metabolomics and proteomics. Before the measured mass spectrum data can be used for quantitative analysis, it must be accurately corrected for isotope natural abundance and tracer isotopic impurity. Despite the increasing popularity of dual-isotope tracing strategy such as  $^{13}\text{C}$ - $^{15}\text{N}$  or  $^{13}\text{C}$ - $^2\text{H}$ , there is no accurate tool for correcting isotope natural abundance for such experiments. Here, we present AccuCor2 as an R-based tool to perform the correction for  $^{13}\text{C}$ - $^{15}\text{N}$  or  $^{13}\text{C}$ - $^2\text{H}$  labeling experiments. Our results show that the dual-isotope experiments often require a mass resolution that is high enough to resolve  $^{13}\text{C}$  and  $^{15}\text{N}$  or  $^{13}\text{C}$  and  $^2\text{H}$ . Otherwise the labeling pattern is not solvable. However, this mass resolution may not be sufficiently high to resolve other non-tracer elements such as oxygen or sulfur from the tracer elements. Therefore, we design AccuCor2 to perform the correction based on the actual mass resolution of the measurements. Using both simulated and experimental data, we show that AccuCor2 performs accurate and resolution dependent correction for dual-isotope tracer data.



## INTRODUCTION

Stable isotope tracers have been widely used to study the metabolism of carbohydrates, protein, lipids and metabolic co-factors.<sup>1</sup> Tracers can be used to determine the origins of specific metabolites by quantifying the incorporation of the isotope atoms from the labeled substrates.<sup>2-4</sup> Tracers can also be infused *in vivo* in a non-perturbative way to measure the turnover rates of circulating metabolites<sup>5,6</sup>. Another application of stable isotope tracing is metabolic flux analysis (MFA), which provides quantitative evaluation of the rates of biochemical reactions in the entire metabolic network.<sup>7-9</sup> The most commonly used tracing isotopes are  $^{13}\text{C}$ ,  $^2\text{H}$  and  $^{15}\text{N}$ .  $^{13}\text{C}$  tracers are frequently used to study the biosynthesis of metabolites and quantify the carbon flow in the metabolic network<sup>10</sup>. However, there are also many metabolic reactions such as transamination, (de)hydration and redox reactions that do not involve the formation or cleavage of carbon-carbon bonds<sup>11</sup>. In these cases, the  $^{13}\text{C}$  tracers are not informative and tracers that include  $^2\text{H}$  or  $^{15}\text{N}$  should be used instead.  $^2\text{H}$  tracers are commonly used to study isomerase reactions and dehydrogenase reactions that involve NADH or NADPH as co-factors<sup>12,13</sup>. Meanwhile,  $^{15}\text{N}$  tracers are widely used to investigate the metabolism of amino acids and nucleotides<sup>14,15</sup>. Modern mass spectrometry provides high mass resolution that enables unambiguous detection of various stable isotopes simultaneously. As a result, tracers containing dual-isotopes (e.g.  $^{13}\text{C}$ - $^{15}\text{N}$  and  $^{13}\text{C}$ - $^2\text{H}$ ) have been increasingly deployed in biological experiments. For example, L-[ $^{13}\text{C}_6$ ,  $^{15}\text{N}_4$ ] arginine was used to investigate autophagy and melanoma growth<sup>16</sup>. L-[ $^{13}\text{C}$ ,  $^{15}\text{N}$ ] leucine was used to study the whole body amino acid metabolism.<sup>17</sup> Methyl[ $^2\text{H}_3$ ]- $^{13}\text{C}$ -methionine was used to study protein methylation

and breakdown<sup>18</sup>. In these applications, the biological interpretation of the results depends on the accurate determination of the isotopic tracer incorporation in the metabolites.

It should be noted that the measured mass isotopologue distribution does not directly indicate the tracer incorporation. This is because of the natural abundance (NA) of isotopes. For example, the unlabeled metabolite NAD<sup>+</sup> ( $\text{C}_{21}\text{H}_{27}\text{N}_7\text{O}_{14}\text{P}_2$ ) would have 19% being M+1 because of the NA of  $^{13}\text{C}$ . Furthermore, 2.8% of unlabeled NAD<sup>+</sup> could be M+2 because of the NA of  $^{18}\text{O}$ . These natural abundance must be corrected to reveal the actual tracer isotope incorporation, which is done through a process called isotope natural abundance correction (INAC).

Many efforts have been made to advance the methodology of INAC. Biemann *et al.*<sup>19</sup> proposed the first INAC algorithm which subtracts the isotopologue fractions in a stepwise manner (i.e. NA contribution from m+0 were corrected first followed by contributions from m+1, m+2 etc.). Although the stepwise approach is mathematically correct, this method is prone to experimental errors since the measurement error in one mass fraction may be propagated across all other mass fractions and lead to a biased result. This issue was solved by Brauman<sup>20</sup>, Wittmann and Heinzle<sup>21</sup> and Van Winden *et al.*<sup>22</sup> using a least-square approach where all measurement errors were weighted equally. This method first use probability theory to generate a correction matrix that links the isotope labeling pattern and the measured mass fractions affected by NA. Then the least square approach is applied to solve the isotope labeling pattern. Millard *et al.*<sup>23</sup> further improved this method by applying a non-negative least square method to solve the labeling pattern and developed a well-established software IsoCor. IsoCor is an efficient tool that can perform both tracer impurity correction and

$$\mathbf{TM} = \begin{bmatrix} {}^0\text{P}_\text{C} \times {}^0\text{P}_\text{N} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ {}^1\text{P}_\text{C} \times {}^0\text{P}_\text{N} & {}^0\text{P}_\text{C} \times {}^1\text{P}_\text{N} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ {}^2\text{P}_\text{C} \times {}^0\text{P}_\text{N} & {}^1\text{P}_\text{C} \times {}^0\text{P}_\text{N} & {}^0\text{P}_\text{C} \times {}^1\text{P}_\text{N} & 0 & 0 & 0 & 0 & 0 & 0 \\ {}^3\text{P}_\text{C} \times {}^0\text{P}_\text{N} & {}^2\text{P}_\text{C} \times {}^0\text{P}_\text{N} & {}^1\text{P}_\text{C} \times {}^1\text{P}_\text{N} & {}^0\text{P}_\text{C} \times {}^1\text{P}_\text{N} & 0 & 0 & 0 & 0 & 0 \\ {}^0\text{P}_\text{C} \times {}^1\text{P}_\text{N} & 0 & 0 & 0 & {}^0\text{P}_\text{C} \times {}^0\text{P}_\text{N} & 0 & 0 & 0 & 0 \\ {}^1\text{P}_\text{C} \times {}^1\text{P}_\text{N} & {}^0\text{P}_\text{C} \times {}^1\text{P}_\text{N} & 0 & 0 & {}^1\text{P}_\text{C} \times {}^0\text{P}_\text{N} & {}^0\text{P}_\text{C} \times {}^0\text{P}_\text{N} & 0 & 0 & 0 \\ {}^2\text{P}_\text{C} \times {}^1\text{P}_\text{N} & {}^1\text{P}_\text{C} \times {}^1\text{P}_\text{N} & {}^0\text{P}_\text{C} \times {}^1\text{P}_\text{N} & 0 & {}^2\text{P}_\text{C} \times {}^0\text{P}_\text{N} & {}^1\text{P}_\text{C} \times {}^0\text{P}_\text{N} & {}^0\text{P}_\text{C} \times {}^0\text{P}_\text{N} & 0 & 0 \\ {}^3\text{P}_\text{C} \times {}^1\text{P}_\text{N} & {}^2\text{P}_\text{C} \times {}^1\text{P}_\text{N} & {}^1\text{P}_\text{C} \times {}^1\text{P}_\text{N} & {}^0\text{P}_\text{C} \times {}^1\text{P}_\text{N} & {}^3\text{P}_\text{C} \times {}^0\text{P}_\text{N} & {}^2\text{P}_\text{C} \times {}^0\text{P}_\text{N} & {}^1\text{P}_\text{C} \times {}^0\text{P}_\text{N} & {}^0\text{P}_\text{C} \times {}^0\text{P}_\text{N} & 0 \end{bmatrix} \quad (2)$$

INAC for many different tracer elements in large datasets. While IsoCor is accurate for low-resolution mass spectrometry data, it may have an over-correction problem when handling high-resolution data. This is because under high-resolution, some isotopologues may be resolved from the tracer isotopologues and therefore the corresponding fraction should not be subtracted. To address this issue, Su *et al.*<sup>24</sup> developed the first resolution dependent INAC tool called AccuCor. AccuCor uses resolution information to evaluate the impact of all non-tracer isotopes and only corrects the fractions that affect the tracer isotopologue channels. Millard *et al.*<sup>25</sup> and Du *et al.*<sup>26</sup> further refined this method by calculating the correction limits for non-tracer isotope combinations, and developed IsoCor2 and ElemCor respectively. Despite the advancement of INAC methods mentioned above, few tools are available so far to perform INAC for dual-isotope tracer experiments.

In this study, we developed AccuCor2 as an R-based tool to perform a resolution-dependent INAC in data from dual-isotope tracers. We found such correction problems require the two tracer isotopes to be resolved under sufficiently high mass resolution. The non-tracer isotopes are often not fully resolved and AccuCor2 can account for them in a resolution dependent manner. We also compared the performance of AccuCor2 with two other available tools IsoCorrector<sup>27</sup> and PICor using both simulated and experimental data. The result shows both IsoCorrector and PICor under-corrected isotopologues caused by non-tracer isotope natural abundance whereas AccuCor2 correctly solved this problem.

## THEORY

### Theory of isotope natural abundance correction

INAC is the process, whereby, isotope labeling patterns are solved from the measured mass fractions. The isotope labeling pattern (denoted by  $\mathbf{L}$ ) and the measured mass fractions (denoted by  $\mathbf{M}$ ) can be expressed as column vectors. These two vectors are related by the correction matrix (denoted by  $\mathbf{CM}$ ):  $\mathbf{CM} \times \mathbf{L} = \mathbf{M}$ .<sup>21,22</sup> The  $i$ -th row and  $j$ -th column element of the correction matrix  $\mathbf{CM}_{ij}$  represents the probability of the  $j$ -th labeled fraction contributing to the  $i$ -th measured mass fraction due to the isotope NA and the isotopic impurity of the tracer.<sup>21,22</sup> The INAC algorithm that is suitable for high-resolution mass spectrometry data will construct the correction matrix based on the chemical formula of the ion, the mass resolution of the measurements and the isotopic purity of the tracer.<sup>24–26</sup> Once the correction matrix is built, the labeling pattern can be solved by non-negative least squares algorithms such as Hanson-Lawson or L-BFGS-B.<sup>23,24</sup> Here, we will demonstrate the construction of correction matrix for dual-isotope data.

The form of the correction matrix is determined by the forms of the isotope labeling pattern  $\mathbf{L}$  and measured mass fractions  $\mathbf{M}$  vectors. Here, we use serine ( $[\text{M-H}]^+ \text{C}_3\text{H}_6\text{NO}_3$ ) as an example. If serine is only labeled with  $^{13}\text{C}$  in the experiment, the labeling pattern vector  $\mathbf{L} = [{}^{13}\text{C}_0, {}^{13}\text{C}_1, {}^{13}\text{C}_2, {}^{13}\text{C}_3]^T$  and the measured mass fractions vector  $\mathbf{M} = [\text{M} + 0, \text{M} + 1, \text{M} + 2, \text{M} + 3]^T$ . The  $\mathbf{CM}$  connecting these two vectors would be a  $4 \times 4$  matrix. If serine is labeled

with both  $^{13}\text{C}$  and  $^{15}\text{N}$ ,  $\mathbf{L}$  and  $\mathbf{M}$  vectors and their relationship can be expressed as eq. 1.

$$\begin{bmatrix} \text{Correction} \\ \text{Matrix} \end{bmatrix} \times \begin{bmatrix} {}^{13}\text{C}_0 {}^{15}\text{N}_0 \\ {}^{13}\text{C}_1 {}^{15}\text{N}_0 \\ {}^{13}\text{C}_2 {}^{15}\text{N}_0 \\ {}^{13}\text{C}_3 {}^{15}\text{N}_0 \\ {}^{13}\text{C}_0 {}^{15}\text{N}_1 \\ {}^{13}\text{C}_1 {}^{15}\text{N}_1 \\ {}^{13}\text{C}_2 {}^{15}\text{N}_1 \\ {}^{13}\text{C}_3 {}^{15}\text{N}_1 \end{bmatrix} = \begin{bmatrix} \mathbf{M} + \mathbf{C0N0} \\ \mathbf{M} + \mathbf{C1N0} \\ \mathbf{M} + \mathbf{C2N0} \\ \mathbf{M} + \mathbf{C3N0} \\ \mathbf{M} + \mathbf{C0N1} \\ \mathbf{M} + \mathbf{C1N1} \\ \mathbf{M} + \mathbf{C2N1} \\ \mathbf{M} + \mathbf{C3N1} \end{bmatrix} \quad (1)$$

In eq. 1, the element  ${}^{13}\text{C}_m {}^{15}\text{N}_n$  in the  $\mathbf{L}$  vector denotes the fraction with  $m$ - $^{13}\text{C}$  and  $n$ - $^{15}\text{N}$  labeling. The element  $\mathbf{M} + \mathbf{CiNj}$  in the  $\mathbf{M}$  vector denotes the observed mass fraction that corresponds to the incorporation of  $i$   $^{13}\text{C}$  and  $j$   $^{15}\text{N}$  atoms. As we shall see in the result section, in the dual-isotope labeling experiment it is necessary to be able to resolve mass fractions of the same nominal mass such as  $\mathbf{M} + \mathbf{C1N0}$  and  $\mathbf{M} + \mathbf{C0N1}$ . If the measurements are done under unit mass resolution and the  $\mathbf{M} + 1$  ( $m/z$  105 for serine) fraction is obtained instead of the resolved  $\mathbf{M} + \mathbf{C1N0}$  ( $m/z$  105.03867 for serine) and  $\mathbf{M} + \mathbf{C0N1}$  ( $m/z$  105.03235 for serine) fractions, the correction may fail to generate the accurate labeling pattern. For serine which has 3 C and 1 N atoms, the correction matrix  $\mathbf{CM}$  is an  $8 \times 8$  matrix.

### Construction of the correction matrix

$\mathbf{CM}$  can be factorized into three matrices: the tracer isotopic purity matrix  $\mathbf{IPM}$ , the tracer elements matrix  $\mathbf{TM}$  and the non-tracer elements matrix  $\mathbf{NTM}$ :  $\mathbf{CM} = \mathbf{NTM} \times \mathbf{TM} \times \mathbf{IPM}$ . We begin the demonstration with the construction of  $\mathbf{TM}$  because this matrix usually has the greatest impact on  $\mathbf{CM}$ . For convenience, we always use the following convention to arrange the terms in the  $\mathbf{L}$  and  $\mathbf{M}$  vectors: ascending numbers of  $^{13}\text{C}$  followed by ascending numbers of  $^{15}\text{N}$  (eq.1). Therefore, the  $\mathbf{TM}$  for  $^{13}\text{C}$ - $^{15}\text{N}$  labeling of serine can be constructed (eq.2).

As shown in eq. 2, for a metabolite with  $m$  carbon atoms and  $n$  nitrogen atoms,  $\mathbf{TM}$  is a square matrix with  $(m+1) \times (n+1)$  row and columns. In fact,  $\mathbf{TM}$  is composed of  $(n+1) \times (n+1)$  blocks where each block is a  $(m+1) \times (m+1)$  sub-matrix. For example, the element in row 7 column 6 (element [7, 6]) is within block [2, 2] and position [3, 2] in the sub-matrix. This element represents the probability of the  ${}^{13}\text{C}_1 {}^{15}\text{N}_1$ -labeled fraction appearing as  $\mathbf{M} + \mathbf{C2N1}$  due to the carbon and nitrogen natural abundance. Since the  ${}^{13}\text{C}_1 {}^{15}\text{N}_1$ -labeled fraction of serine has 1  $^{13}\text{C}$  from the tracer, there are only 2 carbon atoms that are of NA. Similarly, since the  ${}^{13}\text{C}_1 {}^{15}\text{N}_1$ -labeled fraction of serine has 1  $^{15}\text{N}$  from the tracer, there are no nitrogen atoms that are of NA. We use the term  $\frac{1}{2}\text{P}_\text{C}$  to represent the probability of having 1  $^{13}\text{C}$  out of 2 natural carbon atoms:  $\frac{1}{2}\text{P}_\text{C} = \binom{2}{1} \times p_\text{C}^1 \times (1 - p_\text{C})^{2-1} = 2 \times (0.011)^1 \times (0.989)^1$ , in

$$\mathbf{IPM} = \begin{bmatrix} {}^0\text{P}_{\text{IPC}} \times {}^0\text{P}_{\text{IPN}} & {}^1\text{P}_{\text{IPC}} \times {}^0\text{P}_{\text{IPN}} & {}^2\text{P}_{\text{IPC}} \times {}^0\text{P}_{\text{IPN}} & {}^3\text{P}_{\text{IPC}} \times {}^0\text{P}_{\text{IPN}} & {}^0\text{P}_{\text{IPC}} \times {}^1\text{P}_{\text{IPN}} & {}^1\text{P}_{\text{IPC}} \times {}^1\text{P}_{\text{IPN}} & {}^2\text{P}_{\text{IPC}} \times {}^1\text{P}_{\text{IPN}} & {}^3\text{P}_{\text{IPC}} \times {}^1\text{P}_{\text{IPN}} \\ 0 & {}^0\text{P}_{\text{IPC}} \times {}^0\text{P}_{\text{IPN}} & {}^1\text{P}_{\text{IPC}} \times {}^0\text{P}_{\text{IPN}} & {}^2\text{P}_{\text{IPC}} \times {}^0\text{P}_{\text{IPN}} & 0 & {}^0\text{P}_{\text{IPC}} \times {}^1\text{P}_{\text{IPN}} & {}^1\text{P}_{\text{IPC}} \times {}^1\text{P}_{\text{IPN}} & {}^2\text{P}_{\text{IPC}} \times {}^1\text{P}_{\text{IPN}} \\ 0 & 0 & {}^0\text{P}_{\text{IPC}} \times {}^0\text{P}_{\text{IPN}} & {}^1\text{P}_{\text{IPC}} \times {}^0\text{P}_{\text{IPN}} & 0 & 0 & {}^0\text{P}_{\text{IPC}} \times {}^1\text{P}_{\text{IPN}} & {}^1\text{P}_{\text{IPC}} \times {}^1\text{P}_{\text{IPN}} \\ 0 & 0 & 0 & {}^0\text{P}_{\text{IPC}} \times {}^0\text{P}_{\text{IPN}} & 0 & 0 & 0 & {}^0\text{P}_{\text{IPC}} \times {}^1\text{P}_{\text{IPN}} \\ 0 & 0 & 0 & 0 & {}^0\text{P}_{\text{IPC}} \times {}^0\text{P}_{\text{IPN}} & {}^1\text{P}_{\text{IPC}} \times {}^0\text{P}_{\text{IPN}} & {}^2\text{P}_{\text{IPC}} \times {}^0\text{P}_{\text{IPN}} & {}^3\text{P}_{\text{IPC}} \times {}^0\text{P}_{\text{IPN}} \\ 0 & 0 & 0 & 0 & 0 & {}^0\text{P}_{\text{IPC}} \times {}^1\text{P}_{\text{IPN}} & {}^1\text{P}_{\text{IPC}} \times {}^1\text{P}_{\text{IPN}} & {}^2\text{P}_{\text{IPC}} \times {}^1\text{P}_{\text{IPN}} \\ 0 & 0 & 0 & 0 & 0 & 0 & {}^0\text{P}_{\text{IPC}} \times {}^1\text{P}_{\text{IPN}} & {}^1\text{P}_{\text{IPC}} \times {}^1\text{P}_{\text{IPN}} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & {}^0\text{P}_{\text{IPC}} \times {}^1\text{P}_{\text{IPN}} \end{bmatrix} \quad (3)$$

$$\mathbf{NTM} = \begin{bmatrix} {}^0\text{P}_{\text{H}} \times {}^0\text{P}_{\text{O}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ {}^1\text{P}_{\text{H}} \times {}^0\text{P}_{\text{O}} & {}^0\text{P}_{\text{H}} \times {}^0\text{P}_{\text{O}} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & {}^1\text{P}_{\text{H}} \times {}^0\text{P}_{\text{O}} & {}^0\text{P}_{\text{H}} \times {}^0\text{P}_{\text{O}} & 0 & 0 & 0 & 0 & 0 \\ {}^1\text{P}_{\text{H}} \times {}^1\text{P}_{\text{O}} & 0 & {}^1\text{P}_{\text{H}} \times {}^1\text{P}_{\text{O}} & {}^0\text{P}_{\text{H}} \times {}^1\text{P}_{\text{O}} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & {}^0\text{P}_{\text{H}} \times {}^0\text{P}_{\text{O}} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & {}^1\text{P}_{\text{H}} \times {}^0\text{P}_{\text{O}} & {}^0\text{P}_{\text{H}} \times {}^0\text{P}_{\text{O}} & 0 \\ 0 & 0 & 0 & 0 & 0 & {}^1\text{P}_{\text{H}} \times {}^1\text{P}_{\text{O}} & {}^0\text{P}_{\text{H}} \times {}^1\text{P}_{\text{O}} & 0 \\ 0 & 0 & 0 & 0 & 0 & {}^1\text{P}_{\text{H}} \times {}^1\text{P}_{\text{O}} & {}^0\text{P}_{\text{H}} \times {}^1\text{P}_{\text{O}} & {}^0\text{P}_{\text{H}} \times {}^0\text{P}_{\text{O}} \end{bmatrix} \quad (4)$$

which  $p_{\text{C}}=0.011$  represents the NA of  $^{13}\text{C}$ . Therefore, the element [7, 6] is  ${}^1\text{P}_{\text{C}} \times {}^0\text{P}_{\text{N}}$ . More generally, for the carbon-nitrogen **TM**, the element [(i-1)×(m+1)+f, (j-1)×(m+1)+g] which is in the i-th row and j-th column of the blocks and the f-th row and g-th column of the sub-matrix has the value of  ${}^{f-g}\text{P}_{\text{C}} \times {}^{i-j}\text{P}_{\text{N}}$ . The element [7,6] has the block indices of i=2, j=2 and sub-matrix indices f=3, g=2, and this element has the value of  ${}^{3-2}\text{P}_{\text{C}} \times {}^{2-2}\text{P}_{\text{N}} = {}^1\text{P}_{\text{C}} \times {}^0\text{P}_{\text{N}}$ . Because  ${}^{f-g}\text{P}_{\text{C}} \times {}^{i-j}\text{P}_{\text{N}}=0$  when  $f<g$  or  $i<j$ , **TM** is a lower triangular matrix and each non-zero submatrix is also a lower triangular matrix. Because of this triangularity, **TM** is always well-conditioned.

Similar to **TM**, we can also construct the isotopic purity matrix **IPM** for  $^{13}\text{C}$ - $^{15}\text{N}$  labeled serine (eq. 3). For a metabolite with **m** carbon atoms and **n** nitrogen atoms, **IPM** is a square matrix composed of (n+1) × (n+1) blocks where each block is a (m+1) × (m+1) sub-matrix. The element in row 5 column 7 (element [5, 7]) is within block [2, 2] and position [1, 3] in the sub-matrix. This element represents the probability of the  $^{13}\text{C}_2\text{ }^{15}\text{N}_1$ -labeled fraction appearing as **M + CON1** due to the tracer isotopic impurities. We use the term  ${}^2\text{P}_{\text{IPC}}$  to represent the probability of having 2 natural carbon atoms out of 2 tracer carbon atoms:

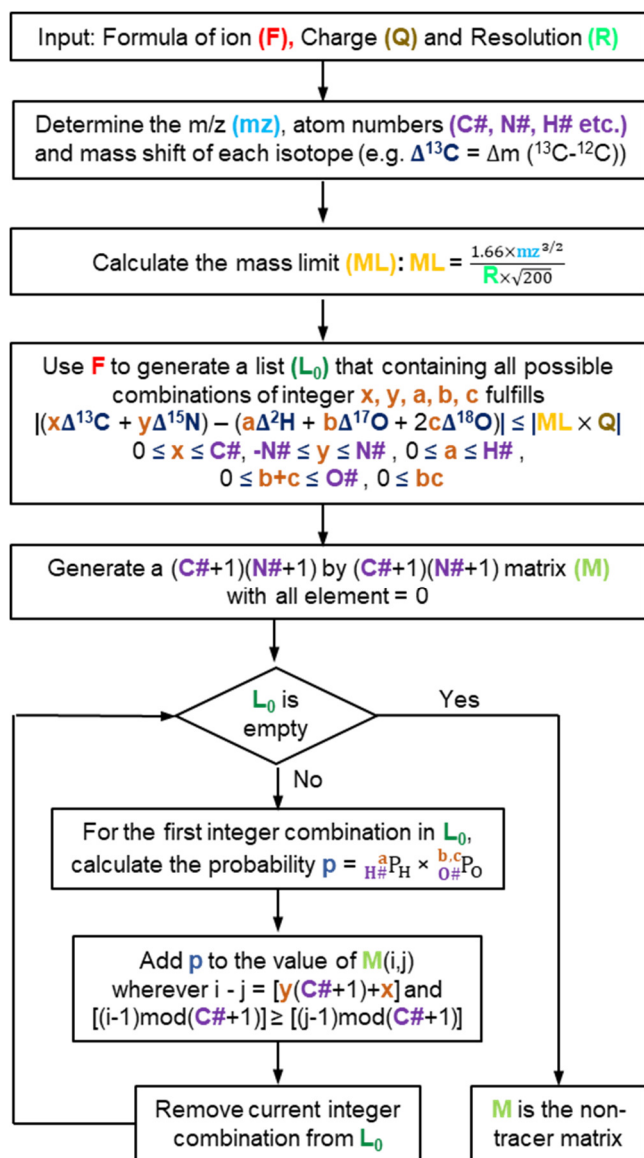
$${}^2\text{P}_{\text{IPC}} = \binom{2}{2} \times (1 - p_{\text{IPC}})^2 \times p_{\text{IPC}}^{2-2} = 1 \times (1 - 0.99)^2 \times (0.99)^0$$

, in which  $p_{\text{IPC}}=0.99$  represents the isotopic purity of  $^{13}\text{C}$  in the tracer. Therefore, the element [5, 7] is  ${}^2\text{P}_{\text{IPC}} \times {}^0\text{P}_{\text{IPN}}$ . More generally, for the carbon-nitrogen **IPM**, the element [(i-1)×(m+1)+f, (j-1)×(m+1)+g] which is in the i-th row and j-th column of the blocks and the f-th row and g-th column of the sub-matrix has the value of  ${}^{g-f}\text{P}_{\text{IPC}} \times {}^{j-i}\text{P}_{\text{IPN}}$ . The element [5,7] has the block indices of i=2, j=2 and sub-matrix indices f=1, g=3, and this element has the value of  ${}^{3-1}\text{P}_{\text{IPC}} \times {}^{2-2}\text{P}_{\text{IPN}} = {}^2\text{P}_{\text{IPC}} \times {}^0\text{P}_{\text{IPN}}$ . Because  ${}^{g-f}\text{P}_{\text{IPC}} \times {}^{j-i}\text{P}_{\text{IPN}}=0$  when  $g<f$  or  $j<i$ , **IPM** is an upper triangular matrix and each non-zero submatrix is also an upper triangular matrix.

Unlike the construction of **TM** and **IPM** described above, the construction of **NTM** is affected by the mass resolution of the measurement<sup>24</sup>. For serine that is  $^{13}\text{C}$ - $^{15}\text{N}$  labeled, it takes a minimal resolution of 50,800 (defined at m/z 200 on Orbitrap analyzer)<sup>24</sup> to resolve  $^{18}\text{O}_1$  and  $^{13}\text{C}_2$  peaks but a minimal resolution of 140,000 to resolve  $^{17}\text{O}_1$  and  $^{13}\text{C}_1$  peaks on the mass spectrum.

Therefore, if data were measured at a mass resolution of 100,000, **NTM** should include the term of  $^{17}\text{O}_1$  but exclude the term of  $^{18}\text{O}_1$  to avoid overcorrection. In our previous work, **NTM** was further factorized. The contribution from each non-tracer element such as oxygen or sulfur was described by separate matrices. These matrices were multiplied together to obtain the full **NTM**. Later, Millard *et al.*<sup>25</sup> pointed out that the combination of certain non-tracer elements may be harder to resolve than individual elements. For example, a minimal resolution of 50,800 and 42,500 are required to resolve  $^{18}\text{O}_1$  versus  $^{13}\text{C}_2$  and  $^2\text{H}_1$  versus  $^{13}\text{C}_1$  of serine respectively. However, it takes a minimal resolution of 260,000 to resolve  $^{18}\text{O}_1+^2\text{H}_1$  versus  $^{13}\text{C}_3$  of serine. As a result, a single **NTM** should be constructed instead of a product of matrices for individual element. The latest version of AccuCor has adopted such an algorithm for improved accuracy. For AccuCor2, we also construct a single **NTM** to account for the NA of the non-tracer elements. The algorithm for constructing **NTM** is shown in Figure 1. First, the algorithm uses the input data to calculate the mass limit (ML) of resolvable peaks. Second, the algorithm generates a list of atom combinations that satisfy the following criteria: 1, for each element, the atom number in the list does not exceed the atom number in the chemical formula; 2, the difference between the mass shift caused by the tracer atoms and the non-tracer atoms is within the mass limit. Lastly, the probability of each atom combination in the list is calculated and added to the corresponding position of **NTM**. The construction completes when all entries in the list are accounted.

As shown in eq. 4, for a metabolite with **m** carbon and **n** nitrogen, **NTM** is a square matrix composed of (n+1)×(n+1) blocks where each block is a (m+1)×(m+1) sub-matrix. The element in row 4 column 1 (element [4,1]) represents the probability of the  $^{13}\text{C}_0\text{ }^{15}\text{N}_0$ -labeled fraction contributing to the M+C3N0 m/z channel due to the non-tracer elements. Under 100,000 resolution, the only non-tracer element combination that make such mass shift is  $^2\text{H}_1+^{18}\text{O}_1$ , which has the probability of  ${}^1\text{P}_{\text{H}} \times {}^0\text{P}_{\text{O}}$ . The term  ${}^1\text{P}_{\text{H}}$  represents the probability of having 1  $^2\text{H}$  atom out of 6 natural hydrogen atoms. The term  ${}^0\text{P}_{\text{O}}$  represents the probability of having 0  $^{17}\text{O}$  and 1  $^{18}\text{O}$  atom out of 3 natural oxygen atoms. Note that **NTM** is not strictly lower-triangular since element [4,5] is non-zero. This represents the probability of  $^{13}\text{C}_0\text{ }^{15}\text{N}_1$  labeled serine contributing to the M+C3N0 channel, which reflects the fact that  $^{15}\text{N}_1+^2\text{H}_2$  and  $^{13}\text{C}_3$  are not resolved for serine under 100,000 resolution.



**Figures 1: NTM construction algorithm for  $^{13}\text{C}$ - $^{15}\text{N}$  labeling experiments.** In this example, carbon and nitrogen are the tracer elements, and we assume, for simplicity, that the non-tracer elements only include hydrogen and oxygen. The algorithm takes the chemical formula of the ion, the charge and the mass resolution as the input information. The output is the resolution dependent **NTM**. Algorithm for  $^{13}\text{C}$ - $^2\text{H}$  labeling experiments follows the same structure.

## EXPERIMENTAL SECTION

### Materials

LCMS-grade methanol (A456), acetonitrile (A955), acetic acid (A35), and water (ACROS 61515) were purchased from Fisher Chemicals (Pittsburgh, PA).  $^{13}\text{C}_3$ - $^{15}\text{N}_1$  L-serine (99% isotopic purity on both  $^{13}\text{C}$  and  $^{15}\text{N}$ , CNLM-474-H-PK) was purchased from Cambridge Isotope Laboratories (Tewksbury, MA). L-serine (S4500) was purchased from MilliporeSigma (Burlington, MA).

### Sample Preparation

100  $\mu\text{M}$  of L-serine,  $^{13}\text{C}_3$ - $^{15}\text{N}_1$  L-serine or 1:1 mixture were dissolve in 40:40:20 (acetonitrile:methanol:water) and centrifuged for 10 min at 4  $^\circ\text{C}$ . The supernatant was collected for LC-MS analysis.

## Liquid chromatography-Mass Spectrometry

The LC-MS method is the same as previously published method. The MS scans were obtained in negative ionization mode with a resolution of 70,000 at  $m/z$  200, in addition to an automatic gain control target of  $3 \times 10^6$  and  $m/z$  scan range of 72 to 1000. Metabolite data was obtained using the MAVEN software package<sup>28</sup> with each labeled isotope fraction (mass accuracy window: 16 ppm).

## Isotope natural abundance correction

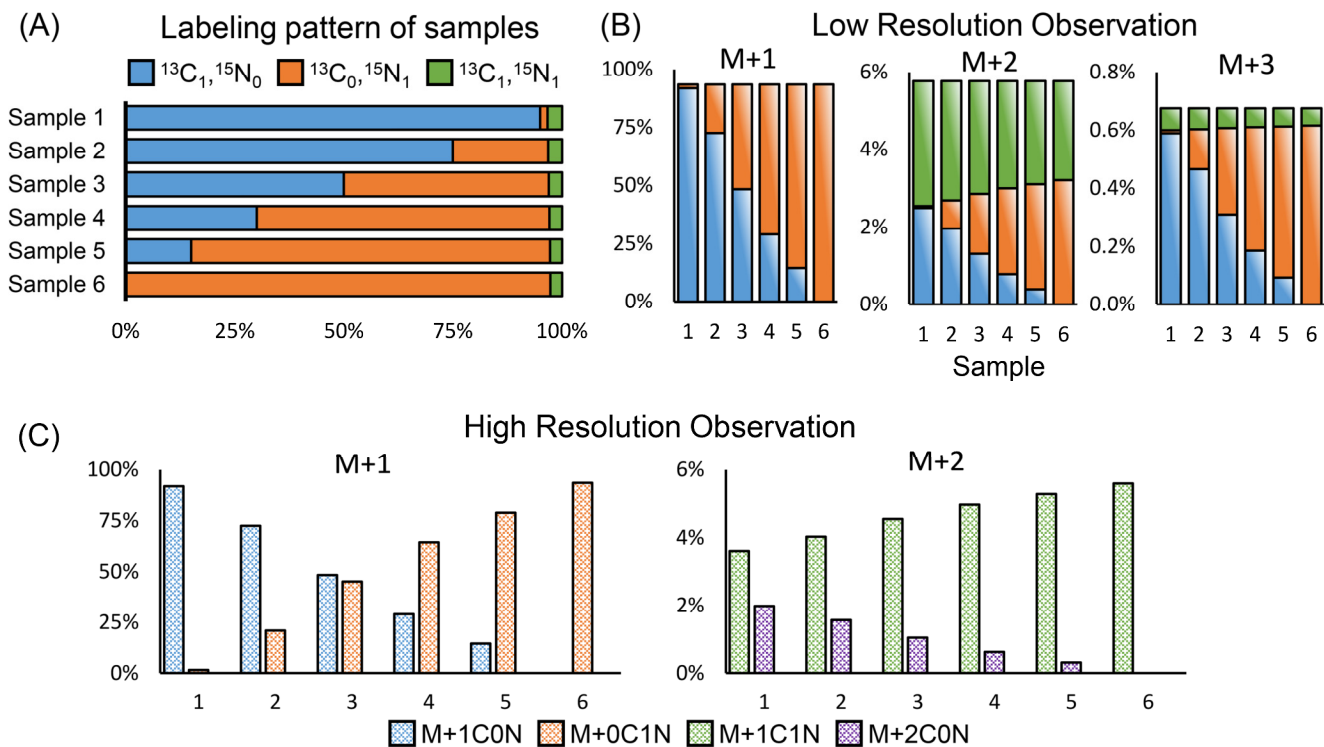
Spectra simulation was done using Thermo Xcalibur Qual Browser. The natural isotope abundance correction code was written in R. The labeling pattern solving was done using the Hanson-Lawson non-negative least squares method, which was implemented in the NNLS package, to avoid negative fractions. The AccuCor2 code is freely available in two versions for  $^{13}\text{C}$ - $^2\text{H}$ , and  $^{13}\text{C}$ - $^{15}\text{N}$  labeling studies (<https://github.com/wangyujue23/AccuCor2>).

## RESULTS AND DISCUSSION

### Requirement on the mass resolution for dual-isotope tracer experiment

Early isotope tracer experiments that use a single isotope were typically done on unit resolution mass spectrometers. The dual-isotope tracer experiments, however, often have a minimum requirement on the mass resolution of the instrument. Take  $^{13}\text{C}$ - $^{15}\text{N}$  serine as an example, there are 8 possible labeled fractions of serine:  $^{13}\text{C}_0$ - $^{15}\text{N}_{0-1}$ . The labeling pattern of serine can be correctly solved if the sample is measured on a high-resolution instrument (see Theory). Under low mass resolution, different labeled fractions with the same nominal mass are combined (e.g.  $\text{M}+\text{C}1\text{N}0$  and  $\text{M}+\text{C}0\text{N}1$  are combined as  $\text{M}+1$ ), resulting in fewer mass fraction measurements than the labeled fractions. Therefore, when measured on a unit resolution mass spectrometer, serine has only 5 mass fractions  $\text{M}+0$  -  $\text{M}+4$ . It is obviously impossible to solve 8 labeling fractions based on the measurements of only 5 mass fractions. Therefore, the labeling fractions of the metabolite cannot be determined.

Even if the choice of the tracer limits the number of possible labeling fractions to be no more than the number of measured mass fractions, the correction may still fail. Suppose serine can be labeled in only 3 forms:  $^{13}\text{C}_1$ - $^{15}\text{N}_0$ ,  $^{13}\text{C}_0$ - $^{15}\text{N}_1$  and  $^{13}\text{C}_1$ - $^{15}\text{N}_1$ . Under unit mass resolution, the  $^{13}\text{C}_1$ - $^{15}\text{N}_0$  and  $^{13}\text{C}_0$ - $^{15}\text{N}_1$  serine are both measured as  $\text{M}+1$  and the  $^{13}\text{C}_1$ - $^{15}\text{N}_1$  fraction is measured as  $\text{M}+2$ . Due to the isotope natural abundance,  $^{13}\text{C}_1$ - $^{15}\text{N}_0$  and  $^{13}\text{C}_0$ - $^{15}\text{N}_1$  serine also contribute to  $\text{M}+2$  and  $\text{M}+3$  fractions and  $^{13}\text{C}_1$ - $^{15}\text{N}_1$  also contributes to the  $\text{M}+3$  fraction. Therefore, it would appear to be possible to determine 3 labeling fractions  $^{13}\text{C}_1$ - $^{15}\text{N}_0$ ,  $^{13}\text{C}_0$ - $^{15}\text{N}_1$  and  $^{13}\text{C}_1$ - $^{15}\text{N}_1$  from the measurement of 3 mass fractions  $\text{M}+1$ ,  $\text{M}+2$  and  $\text{M}+3$ . However, Figure 2 provides a set of examples to demonstrate that such correction cannot be done. We calculated 6 combinations of labeling fractions (Figure 2A, Sample 1-6) which have the  $^{13}\text{C}_1$ - $^{15}\text{N}_0$  fraction ranging from 0% to 95%. These samples have exactly the same  $\text{M}+1$ ,  $\text{M}+2$  and  $\text{M}+3$  fractions when measured under unit mass resolution (Figure 2B, Supplementary Table 1). Therefore, the labeling fractions cannot be calculated from the measured mass fractions. In fact, there are infinitely many combinations that could fit these mass measurements. The mathematical reason behind this phenomenon is that low resolution causes certain columns of the correction matrix, such as those correspond to  $\text{M}+1\text{C}0\text{N}$  and  $\text{M}+0\text{C}1\text{N}$ , to be combined as  $\text{M}+1$ . These combinations make the correction matrix singular. Consequently, the labeling fractions cannot be determined from such measurements.



**Figures 2: Dual-isotope tracer experiment requires high mass resolution.** Serine isotopologue mixtures ( $^{13}\text{C}_1\text{-}^{15}\text{N}_0$ ,  $^{13}\text{C}_0\text{-}^{15}\text{N}_1$  and  $^{13}\text{C}_1\text{-}^{15}\text{N}_1$ ) were measured under low and high resolution MS. Low resolution do not distinguish  $^{13}\text{C}$  and  $^2\text{H}$ , leading to the same observation among all samples. Consequently, the labeling pattern is unsolvable under low resolution. In contrast, high resolution can distinguish  $^{13}\text{C}$  and  $^2\text{H}$  isotopologues, leading to unique solution of labeling pattern.

In contrast, when these samples were measured under a mass resolution that is high enough to resolve  $^{13}\text{C}$  and  $^{15}\text{N}$ ,  $^{13}\text{C}_1\text{-}^{15}\text{N}_0$ ,  $^{13}\text{C}_0\text{-}^{15}\text{N}_1$  and  $^{13}\text{C}_1\text{-}^{15}\text{N}_1$  would be measured as M+1C0N, M+0C1N and M+1C1N respectively (Figure 2C).  $^{13}\text{C}_1\text{-}^{15}\text{N}_0$  would also contribute to M+2C0N due to NA, which can be resolved from M+1C1N under high mass resolution. In this case, all 6 samples give different mass fraction measurements so that their respective labeling patterns can be solved accurately. For serine ( $[\text{M-H}]^-$ ,  $m/z$  104.03532), the nominal mass resolution (for Orbitrap instrument, defined at  $m/z$  200)<sup>24,29</sup> should be  $>19,700$  to resolve  $^{13}\text{C}$  and  $^{15}\text{N}$ . In general, the mass resolution should be high enough to resolve all the labeling fractions of interest. This requirement should be taken into consideration when designing a dual-isotope tracer experiment.

#### Resolution dependent correction for the non-tracer elements

Dual-isotope tracer experiments require a minimum mass resolution under which the tracer isotopes can be resolved. However, this resolution may not be high enough to resolve other non-tracer elements which should also be corrected. As shown by Su *et al.*<sup>24</sup>, the minimum resolution for isotopologue separation depends on both the absolute mass difference ( $\Delta m$ ) and  $m/z$  (represented by  $m$  in the formula).

$$\text{Nominal Resolution} \geq \frac{1.66 \times m^{3/2}}{\Delta m \times \sqrt{200}} \quad (5)$$

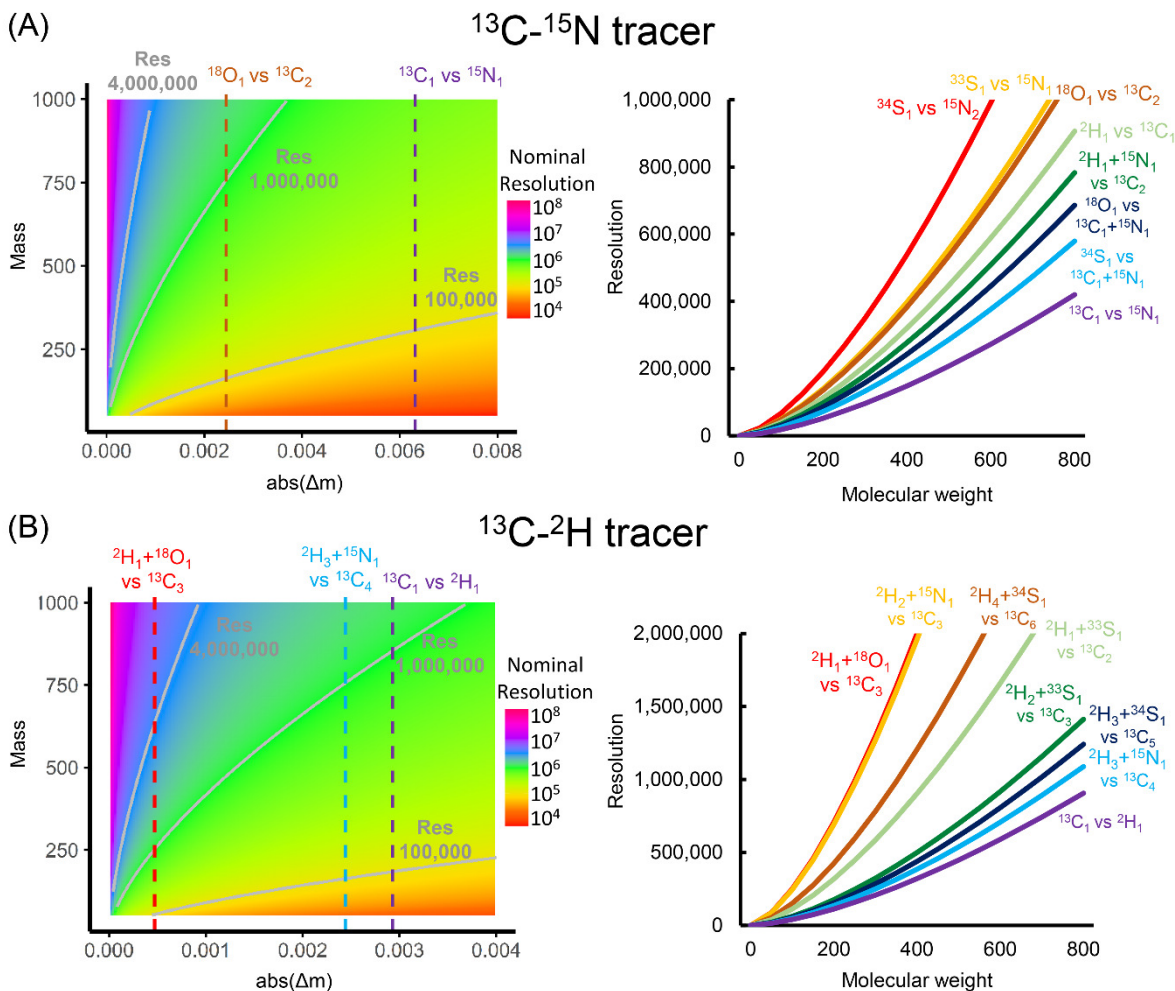
Eq. 5 suggests that as the mass difference  $\Delta m$  goes lower, a higher mass resolution is required to resolve the corresponding isotopologues. Figure 3 provides a visual illustration of the relationship among resolution requirement, mass difference and the mass of ion. For serine ( $[\text{M-H}]^-$ ,  $m/z$  104.03532) that is  $^{13}\text{C}$  and  $^{15}\text{N}$

labeled, it only takes a minimum resolution of 19,700 to resolve the  $^{13}\text{C}_1\text{-}^{15}\text{N}_0$  and  $^{13}\text{C}_0\text{-}^{15}\text{N}_1$  isotopologues. However, it takes a minimum resolution of 50,800 to resolve  $^{13}\text{C}_2$  and  $^{18}\text{O}_1$  ( $\Delta m=0.00245$ ). Similarly, for serine that is  $^{13}\text{C}$  and  $^2\text{H}$  labeled, it takes a minimum resolution of 42,500 to resolve the  $^{13}\text{C}_1\text{-}^2\text{H}_0$  and  $^{13}\text{C}_0\text{-}^2\text{H}_1$  isotopologues. And it takes a minimum resolution of 260,000 to resolve  $^2\text{H}_1\text{-}^{18}\text{O}_1$  and  $^{13}\text{C}_3$  ( $\Delta m=0.00048$ ) and 50,400 to resolve  $^2\text{H}_3\text{-}^{15}\text{N}_1$  and  $^{13}\text{C}_4$  ( $\Delta m=0.00247$ ).

Moreover, as the mass of the ion goes up, the required resolution also goes up (Figure 3). For acetyl-CoA ( $[\text{M-H}]^-, \text{C}_{23}\text{H}_{37}\text{N}_7\text{O}_{17}\text{P}_3\text{S}$ ) that is  $^{13}\text{C}\text{-}^{15}\text{N}$  labeled, it takes a minimum resolution of 427,000 to resolve the  $^{13}\text{C}_1\text{-}^{15}\text{N}_0$  and  $^{13}\text{C}_0\text{-}^{15}\text{N}_1$  isotopologues, which is achievable on modern Orbitrap instruments. However, it takes a minimum resolution of 697,000 to resolve  $^{13}\text{C}_1\text{-}^{15}\text{N}_1$  and  $^{18}\text{O}_1$  ( $\Delta m=0.00387$ ), 1,100,000 to resolve  $^{13}\text{C}_2$  and  $^{18}\text{O}_1$  ( $\Delta m=0.00245$ ) and 1,550,000 to resolve  $^{34}\text{S}_1$  and  $^{15}\text{N}_2$  ( $\Delta m=0.00174$ ). The mass resolution for resolving these isotopologues is not attainable on many current mainstream instruments. Fortunately, as long as the tracer isotopes are resolved, the non-tracer correction matrix is always well-conditioned so that the correction can be done even when the non-tracer isotopes are not resolved.

To accurately account for the contribution from the NA of the non-tracer elements, **NTM** should be constructed based on the actual mass resolution of the experiment (See Theory). Basically, based on the mass resolution, AccuCor2 will calculate whether a specific isotopologue is resolved from any of the tracer isotopologues. If the resolution is not high enough so that the non-tracer NA contributes to a mass channel of a labeling fraction, AccuCor2 will add associated probability to **NTM** and correct for it. If the resolution is high enough to resolve the non-tracer isotopologue from the labeling fraction, AccuCor2 will leave this fraction off from **NTM** to avoid over-correction.





**Figures 3: Effect of resolution on isotopologue separation.** The minimal resolution for separating isotopologues according to  $m/z$  and mass difference ( $\Delta m$ ) for experiments using  $^{13}\text{C}$ - $^{15}\text{N}$  (A) and  $^{13}\text{C}$ - $^2\text{H}$  (B) tracer experiments. The most commonly observed isotopologue pairs are shown in different colors. All isotopologues shown were measured with -1 charge.

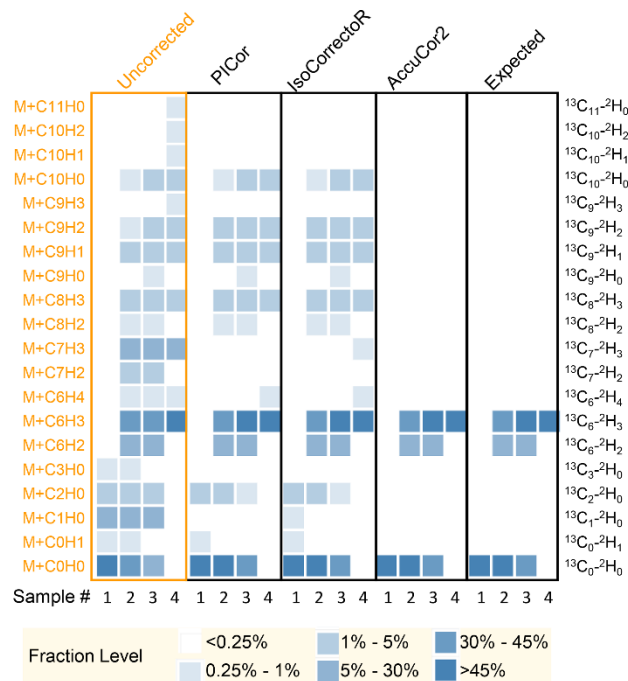
#### AccuCor2 performs accurate natural abundance correction from dual-isotope tracer data

To test the performance of AccuCor2, we first used a simulated dataset of  $^{13}\text{C}$ - $^2\text{H}$  labeled  $\text{NAD}^+$  ( $[\text{M}-\text{H}]^-$ ,  $\text{C}_{21}\text{H}_{26}\text{N}_7\text{O}_{14}\text{P}_2$ ) (Figure 4). The simulation assumed a mass resolution of 750,000 which is sufficient to resolve  $^{13}\text{C}$  and  $^2\text{H}$ . The simulation also assumed the following labeling patterns: sample 1: 100% non-labeled; sample 2: 50% non-labeled + 10%  $^{13}\text{C}_6$ - $^2\text{H}_2$  + 40%  $^{13}\text{C}_6$ - $^2\text{H}_3$ ; sample 3: 36% non-labeled + 14%  $^{13}\text{C}_6$ - $^2\text{H}_2$  + 50%  $^{13}\text{C}_6$ - $^2\text{H}_3$ ; sample 4: 100%  $^{13}\text{C}_6$ - $^2\text{H}_3$ . These isotopologues were chosen because they are possible products when  $^{13}\text{C}_6$ - $^2\text{H}_4$  nicotinamide tracer is used<sup>30</sup>. The simulated dataset shows that besides the expected labeled fractions such as  $\text{M}+0\text{C}0\text{H}$ ,  $\text{M}+6\text{C}2\text{H}$  and  $\text{M}+6\text{C}3\text{H}$ , there are many unexpected isotopologues in the uncorrected data. For example, in sample 4,  $\text{M}+8\text{C}3\text{H}$  is 3.4% of the total intensity and  $\text{M}+9\text{C}1\text{H}$  is 2.1% of the total intensity (Figure 4). Carbon natural abundance alone does not explain the intensity of these signals. The observed signals are in fact mainly contributed by  $^{13}\text{C}_6$ - $^2\text{H}_3$ - $^{18}\text{O}_1$  and  $^{13}\text{C}_6$ - $^2\text{H}_3$ - $^{15}\text{N}_1$  respectively. The mass difference between  $^{13}\text{C}_6$ - $^2\text{H}_3$ - $^{18}\text{O}_1$  and  $^{13}\text{C}_8$ - $^2\text{H}_3$   $\text{NAD}^+$  is 0.00245 so that they are not resolved when the resolution is below 816,000. The mass difference between  $^{13}\text{C}_6$ - $^2\text{H}_3$ - $^{15}\text{N}_1$  and  $^{13}\text{C}_9$ - $^2\text{H}_1$   $\text{NAD}^+$  is 0.00048 so that they are not resolved when the resolution is below 4,200,000. AccuCor2 recognized these fractions that arose due to NA of the non-tracer elements and corrected

them successfully. The corrected results from AccuCor2 is identical to the expected labeling pattern of the samples. PICor and IsoCorrector, on the other hand, failed to correct these fractions. We believe this is because these tools essentially assume infinite mass resolution and ignore non-tracer elements and therefore lead to under-correction of the data (Figure 4).

In addition to the simulated data, we have also tested the performance of AccuCor2 using experimental data (Figure 5). In this experiment, we used [ $^{13}\text{C}_3$ - $^{15}\text{N}_1$ ] serine and unlabeled serine to make the following samples: sample 1: 100% non-labeled; sample 2: 50% non-labeled + 50%  $^{13}\text{C}_3$ - $^{15}\text{N}_1$ ; sample 3: 100%  $^{13}\text{C}_3$ - $^{15}\text{N}_1$ . The measurement was performed at 70,000 resolution. For serine (Figure 5A), this resolution is sufficient to resolve all the non-tracer elements. The only unanticipated measured mass fractions are  $\text{M}+2\text{C}1\text{N}$  and  $\text{M}+3\text{C}0\text{N}$  due to the isotopic impurity of the tracer (99% isotopic purity on both  $^{13}\text{C}$  and  $^{15}\text{N}$ ). AccuCor2 and IsoCorrector both return the expected results, whereas PICor left the  $^{13}\text{C}_2$ - $^{15}\text{N}_1$  and  $^{13}\text{C}_3$ - $^{15}\text{N}_0$  fractions uncorrected because PICor does not handle tracer isotopic impurity (Figure 5A). In this experiment, serine is also observed as a sodium acetate adduct ( $[\text{M}+\text{NaAc}-\text{H}]^-$ ,  $\text{C}_5\text{H}_9\text{NO}_5\text{Na}$ )<sup>31</sup>. We observed signals that correspond to  $\text{M}+5\text{C}1\text{N}$  serine+NaAc (Figure 5B). Since the NaAc is from the LC mobile phase which is unlabeled, this fraction should have been generated

because of isotope natural abundance. In fact, serine+NaAc has the m/z of 186.04, which means a minimum resolution of 120,000 is needed to resolve  $^{13}\text{C}_2$  and  $^{18}\text{O}_1$ . The observed  $^{13}\text{C}_5$ - $^{15}\text{N}_1$  fraction is in fact  $^{13}\text{C}_3$ - $^{15}\text{N}_1$ - $^{18}\text{O}_1$  fraction. The correction results show that AccuCor2 recognized this fraction and corrected it. PICor and IsoCorrectoR failed to correct this fraction. Our tests using simulated and experimental data demonstrate that the resolution dependent correction, as implemented in AccuCor2, is important for the accuracy of isotope natural abundance correction for dual-isotope tracer experiments.



**Figure 4: Performance of PICor, IsoCorrectoR and AccuCor2 in simulated data.** Uncorrected data are generated using NAD<sup>+</sup> MS spectra simulated by Thermo Xcalibur Qual Browser at 100,000 resolution. The original labeling patterns are sample 1: 100% non-labeled; sample 2: 50% non-labeled + 10%  $^{13}\text{C}_6$ - $^2\text{H}_2$  + 40%  $^{13}\text{C}_6$ - $^2\text{H}_3$ ; sample 3: 36% non-labeled + 14%  $^{13}\text{C}_6$ - $^2\text{H}_2$  + 50%  $^{13}\text{C}_6$ - $^2\text{H}_3$ ; sample 4: 100%  $^{13}\text{C}_6$ - $^2\text{H}_3$ .

## CONCLUSION

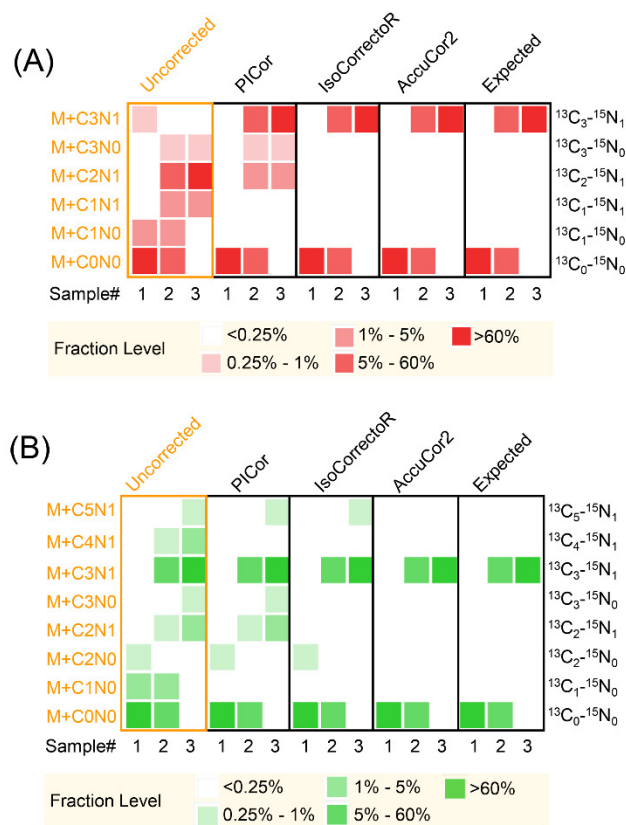
In this study, we developed AccuCor2 as a new tool to perform INAC of data from dual-isotope tracer experiments in a resolution dependent manner. Our results show that the dual-isotope tracer experiment should use a mass resolution that is high enough to resolve all the labeling fractions of interest. Otherwise, the labeling pattern cannot be determined. The non-tracer elements may require an even higher mass resolution to fully resolve, but this can be handled by the resolution dependent correction in AccuCor2. AccuCor2 is freely available in open source format and would enable more accurate calculation of metabolite labeling patterns from MS.

## AUTHOR INFORMATION

### Corresponding Author

\* Xiaoyang Su. Email: xs137@rwjms.rutgers.edu; Tel.: +01-732-235-5447

### Author Contributions



**Figure 5: Performance of PICor, IsoCorrectoR and AccuCor2 in experimental data.** Raw data are serine (A) and sodium acetate adduct of serine (B) ( $\text{C}_5\text{H}_{10}\text{NO}_5\text{Na}$ ) measured under 70,000 resolution. The original labeling patterns are sample 1: 100% non-labeled; sample 2: 50% non-labeled + 50%  $^{13}\text{C}_3$ - $^{15}\text{N}_1$ ; sample 3: 100%  $^{13}\text{C}_3$ - $^{15}\text{N}_1$ .

Y.W. and X.S. wrote the main draft. All authors have reviewed the manuscript. All authors have given approval to the final version of the manuscript.

## ACKNOWLEDGMENT

This research is supported, in part, by NIH grants P30CA072720-5923 (XS).

## REFERENCES

- (1) Kim, I. Y.; Suh, S. H.; Lee, I. K.; Wolfe, R. R. Applications of Stable, Nonradioactive Isotope Tracers in in Vivo Human Metabolic Research. *Experimental and Molecular Medicine*. 2016. <https://doi.org/10.1038/emmm.2015.97>.
- (2) Birsoy, K.; Wang, T.; Chen, W. W.; Freinkman, E.; Abu-Remaileh, M.; Sabatini, D. M. An Essential Role of the Mitochondrial Electron Transport Chain in Cell Proliferation Is to Enable Aspartate Synthesis. *Cell* **2015**, 162 (3), 540–551. <https://doi.org/10.1016/j.cell.2015.07.016>.
- (3) Mills, E. L.; Pierce, K. A.; Jedrychowski, M. P.; Garrity, R.; Winther, S.; Vidoni, S.; Yoneshiro, T.; Spinelli, J. B.; Lu, G. Z.; Kazak, L.; Banks, A. S.; Haigis, M. C.; Kajimura, S.; Murphy, M. P.; Gygi, S. P.; Clish, C. B.; Chouchani, E. T. Accumulation of Succinate Controls Activation of Adipose Tissue Thermogenesis. *Nature* **2018**, 560 (7716), 102–106. <https://doi.org/10.1038/s41586-018-0353-2>.
- (4) TeSlaa, T.; Bartman, C. R.; Jankowski, C. S. R.; Zhang, Z.; Xu, X.; Xing, X.; Wang, L.; Lu, W.; Hui, S.; Rabinowitz, J. D. The Source of Glycolytic Intermediates in Mammalian Tissues. *Cell*

- Metab.* **2021**, *33* (2), 367–378.e5. <https://doi.org/10.1016/j.cmet.2020.12.020>.
- (5) Hui, S.; Ghergurovich, J. M.; Morscher, R. J.; Jang, C.; Teng, X.; Lu, W.; Esparza, L. A.; Reya, T.; Zhan, L.; Yanxiang Guo, J.; White, E.; Rabinowitz, J. D. Glucose Feeds the TCA Cycle via Circulating Lactate. *Nature* **2017**, *551* (7678), 115–118. <https://doi.org/10.1038/nature24057>.
  - (6) Wang, Y.; Kwon, M.; Su, X.; Wondisford, F. E. Glycerol Not Lactate Is the Major Net Carbon Source for Gluconeogenesis in Mice during Both Short and Prolonged Fasting. *Mol. Metab.* **2020**, *31*, 36–44. <https://doi.org/10.1016/j.molmet.2019.11.005>.
  - (7) Wiechert, W.; Möllney, M.; Isermann, N.; Wurzel, M.; De Graaf, A. A. Bidirectional Reaction Steps in Metabolic Networks: III. Explicit Solution and Analysis of Isotopomer Labeling Systems. *Biotechnol. Bioeng.* **1999**, *66* (2), 69–85. [https://doi.org/10.1002/\(SICI\)1097-0290\(1999\)66:2<69::AID-BIT1>3.0.CO;2-6](https://doi.org/10.1002/(SICI)1097-0290(1999)66:2<69::AID-BIT1>3.0.CO;2-6).
  - (8) Antoniewicz, M. R. Methods and Advances in Metabolic Flux Analysis: A Mini-Review. *J. Ind. Microbiol. Biotechnol.* **2015**, *42* (3), 317–325. <https://doi.org/10.1007/s10295-015-1585-x>.
  - (9) Wang, Y.; Wondisford, F. E.; Song, C.; Zhang, T.; Su, X. Metabolic Flux Analysis—Linking Isotope Labeling and Metabolic Fluxes. *Metabolites*. **2020**, pp 1–21. <https://doi.org/10.3390/metabo10110447>.
  - (10) Jang, C.; Chen, L.; Rabinowitz, J. D. Metabolomics and Isotope Tracing. *Cell* **2018**, *173* (4), 822–837. <https://doi.org/10.1016/j.cell.2018.03.055>.
  - (11) Xu, J.; Martien, J.; Gilbertson, C.; Ma, J.; Amador-Noguez, D.; Park, J. O. Metabolic Flux Analysis and Fluxomics-Driven Determination of Reaction Free Energy Using Multiple Isotopes. *Curr. Opin. Biotechnol.* **2020**, *64*, 151–160. <https://doi.org/https://doi.org/10.1016/j.copbio.2020.02.018>.
  - (12) Liu, L.; Shah, S.; Fan, J.; Park, J. O.; Wellen, K. E.; Rabinowitz, J. D. Malic Enzyme Tracers Reveal Hypoxia-Induced Switch in Adipocyte NADPH Pathway Usage. *Nat. Chem. Biol.* **2016**, *12* (5), 345–352. <https://doi.org/10.1038/nchembio.2047>.
  - (13) Chen, L.; Zhang, Z.; Hoshino, A.; Zheng, H. D.; Morley, M.; Arany, Z.; Rabinowitz, J. D. NADPH Production by the Oxidative Pentose-Phosphate Pathway Supports Folate Metabolism. *Nat. Metab.* **2019**, *1* (3), 404–415. <https://doi.org/10.1038/s42255-019-0043-x>.
  - (14) Neinast, M. D.; Jang, C.; Hui, S.; Murashige, D. S.; Chu, Q.; Morscher, R. J.; Li, X.; Zhan, L.; White, E.; Anthony, T. G.; Rabinowitz, J. D.; Arany, Z. Quantitative Analysis of the Whole-Body Metabolic Fate of Branched-Chain Amino Acids. *Cell Metab.* **2019**, *29* (2), 417–429.e4. <https://doi.org/10.1016/j.cmet.2018.10.013>.
  - (15) Spinelli, J. B.; Yoon, H.; Ringel, A. E.; Jeanfavre, S.; Clish, C. B.; Haigis, M. C. Metabolic Recycling of Ammonia via Glutamate Dehydrogenase Supports Breast Cancer Biomass. *Science* (80-. ). **2017**, *358* (6365), 941–946. <https://doi.org/10.1126/science.aam9305>.
  - (16) Poillet-Perez, L.; Xie, X.; Zhan, L.; Yang, Y.; Sharp, D. W.; Hu, Z. S.; Su, X.; Maganti, A.; Jiang, C.; Lu, W.; Zheng, H.; Bosenberg, M. W.; Mehnert, J. M.; Guo, J. Y.; Lattime, E.; Rabinowitz, J. D.; White, E. Autophagy Maintains Tumour Growth through Circulating Arginine. *Nature* **2018**, *563* (7732), 569–573. <https://doi.org/10.1038/s41586-018-0697-7>.
  - (17) Yu, Y. M.; Burke, J. F.; Vogt, J. A.; Chambers, L.; Young, V. R. Splanchnic and Whole Body L-[1-<sup>13</sup>C,<sup>15</sup>N]Leucine Kinetics in Relation to Enteral and Parenteral Amino Acid Supply. *Am. J. Physiol. - Endocrinol. Metab.* **1992**, *262* (5 25-5), <https://doi.org/10.1152/ajpendo.1992.262.5.e687>.
  - (18) Crossland, H.; Smith, K.; Atherton, P. J.; Wilkinson, D. J. A Novel Stable Isotope Tracer Method to Simultaneously Quantify Skeletal Muscle Protein Synthesis and Breakdown. *Metab. Open* **2020**, *5*, 100022. <https://doi.org/10.1016/j.metop.2020.100022>.
  - (19) Biemann, K. *Mass Spectrometry: Organic Chemical Applications*; McGraw-Hill: New York, 1962.
  - (20) Brauman, J. I. Least Squares Analysis and Simplification of Multi-Isotope Mass Spectra. *Anal. Chem.* **1966**, *38* (4), 607–610. <https://doi.org/10.1021/ac60236a020>.
  - (21) Wittmann, C.; Heinze, E. Mass Spectrometry for Metabolic Flux Analysis. *Biotechnol. Bioeng.* **1999**, *62* (6), 739–750. [https://doi.org/10.1002/\(SICI\)1097-0290\(19990320\)62:6<739::AID-BIT13>3.0.CO;2-E](https://doi.org/10.1002/(SICI)1097-0290(19990320)62:6<739::AID-BIT13>3.0.CO;2-E).
  - (22) Van Winden, W. A.; Wittmann, C.; Heinze, E.; Heijnen, J. J. Correcting Mass Isotopomer Distributions for Naturally Occurring Isotopes. *Biotechnol. Bioeng.* **2002**, *80* (4), 477–479. <https://doi.org/10.1002/bit.10393>.
  - (23) Millard, P.; Letisse, F.; Sokol, S.; Portais, J. C. IsoCor: Correcting MS Data in Isotope Labeling Experiments. *Bioinformatics* **2012**, *28* (9), 1294–1296. <https://doi.org/10.1093/bioinformatics/bts127>.
  - (24) Su, X.; Lu, W.; Rabinowitz, J. D. Metabolite Spectral Accuracy on Orbitraps. *Anal. Chem.* **2017**, *89* (11), 5940–5948. <https://doi.org/10.1021/acs.analchem.7b00396>.
  - (25) Millard, P.; Delépine, B.; Guionnet, M.; Heuillet, M.; Bellvert, F.; Létisse, F. IsoCor: Isotope Correction for High-Resolution MS Labeling Experiments. *Bioinformatics* **2019**, *35* (21), 4484–4487. <https://doi.org/10.1093/bioinformatics/btz209>.
  - (26) Du, D.; Tan, L.; Wang, Y.; Peng, B.; Weinstein, J. N.; Wondisford, F. E.; Su, X.; Lorenzi, P. L. ElemCor: Accurate Data Analysis and Enrichment Calculation for High-Resolution LC-MS Stable Isotope Labeling Experiments. *BMC Bioinformatics* **2019**, *20* (1), <https://doi.org/10.1186/s12859-019-2669-9>.
  - (27) Heinrich, P.; Kohler, C.; Ellmann, L.; Kuerner, P.; Spang, R.; Oefner, P. J.; Dettmer, K. Correcting for Natural Isotope Abundance and Tracer Impurity in MS-, MS/MS- and High-Resolution-Multiple-Tracer-Data from Stable Isotope Labeling Experiments with IsoCorrector. *Sci. Rep.* **2018**, *8* (1), <https://doi.org/10.1038/s41598-018-36293-4>.
  - (28) Melamud, E.; Vastag, L.; Rabinowitz, J. D. Metabolomic Analysis and Visualization Engine for LC - MS Data. *Anal. Chem.* **2010**, *82* (23), 9818–9826. <https://doi.org/10.1021/ac1021166>.
  - (29) Zubarev, R. A.; Makarov, A. Orbitrap Mass Spectrometry. *Anal. Chem.* **2013**, *85* (11), 5288–5296. <https://doi.org/10.1021/ac4001223>.
  - (30) Liu, L.; Su, X.; Quinn, W. J.; Hui, S.; Krukenberg, K.; Frederick, D. W.; Redpath, P.; Zhan, L.; Chellappa, K.; White, E.; Migaud, M.; Mitchison, T. J.; Baur, J. A.; Rabinowitz, J. D. Quantitative Analysis of NAD Synthesis-Breakdown Fluxes. *Cell Metab.* **2018**, *27* (5), 1067–1080.e5. <https://doi.org/10.1016/j.cmet.2018.03.018>.
  - (31) Su, X.; Chiles, E.; Maimouni, S.; Wondisford, F. E.; Zong, W.-X.; Song, C. In-Source CID Ramping and Covariant Ion Analysis of Hydrophilic Interaction Chromatography Metabolomics. *Anal. Chem.* **2020**, *92* (7), 4829–4837. <https://doi.org/10.1021/acs.analchem.9b04181>.