

<https://doi.org/10.1038/s42003-024-05844-z>

# SIMPEL: using stable isotopes to elucidate dynamics of context specific metabolism



Shrikaar Kambhampati<sup>1,7,8</sup>✉, Allen H. Hubbard<sup>1,8</sup>, Somnath Koley<sup>1,8</sup>, Javier D. Gomez<sup>1,2</sup>, Frédéric Marsolais<sup>3,4</sup>, Bradley S. Evans<sup>1</sup>, Jamey D. Young<sup>1,2,5</sup> & Doug K. Allen<sup>1,6</sup>✉

The capacity to leverage high resolution mass spectrometry (HRMS) with transient isotope labeling experiments is an untapped opportunity to derive insights on context-specific metabolism, that is difficult to assess quantitatively. Tools are needed to comprehensively mine isotopologue information in an automated, high-throughput way without errors. We describe a tool, Stable Isotope-assisted Metabolomics for Pathway Elucidation (*SIMPEL*), to simplify analysis and interpretation of isotope-enriched HRMS datasets. The efficacy of *SIMPEL* is demonstrated through examples of central carbon and lipid metabolism. In the first description, a dual-isotope labeling experiment is paired with *SIMPEL* and isotopically nonstationary metabolic flux analysis (INST-MFA) to resolve fluxes in central metabolism that would be otherwise challenging to quantify. In the second example, *SIMPEL* was paired with HRMS-based lipidomics data to describe lipid metabolism based on a single labeling experiment. Available as an R package, *SIMPEL* extends metabolomics analyses to include isotopologue signatures necessary to quantify metabolic flux.

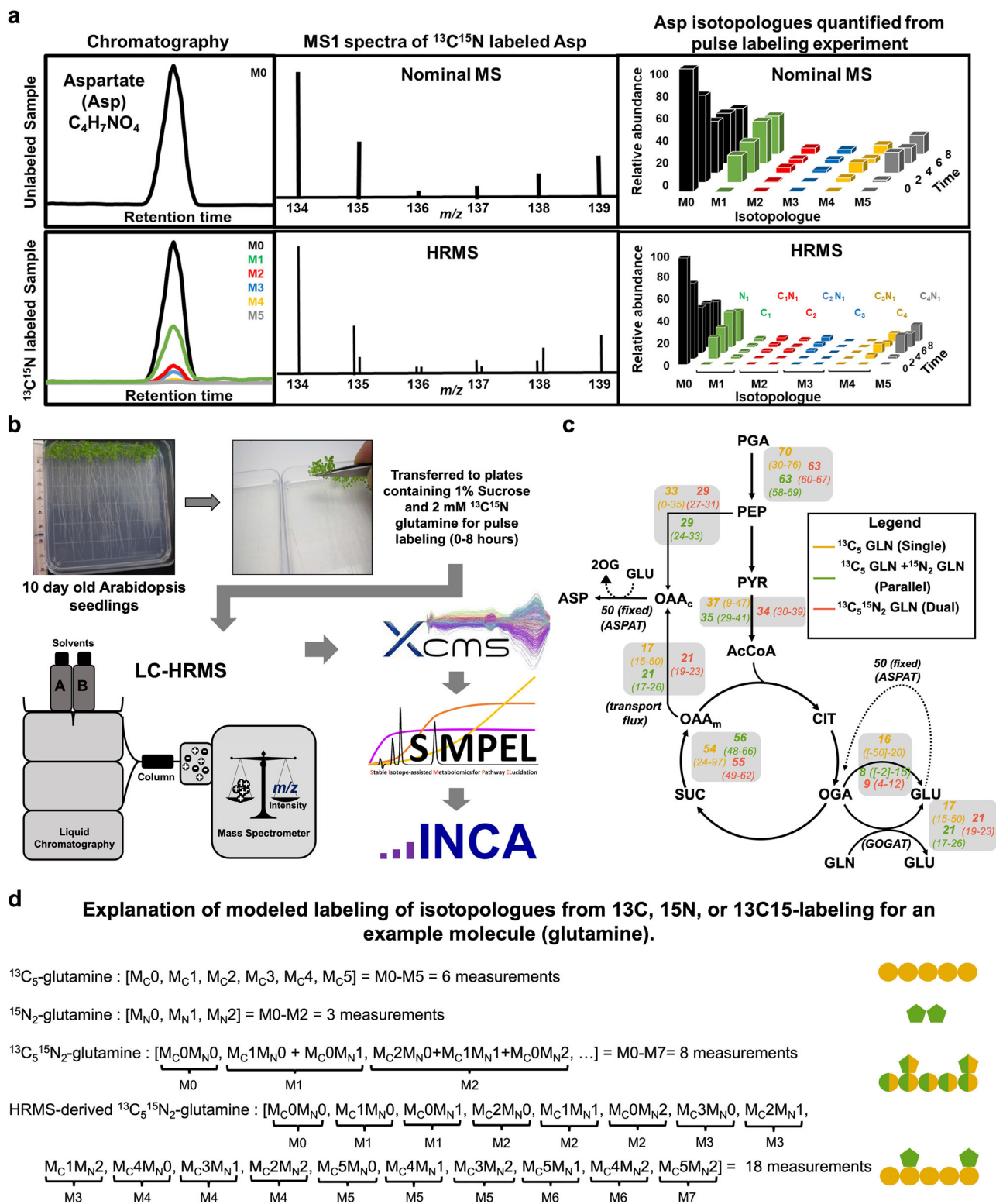
Cellular metabolism is defined by fluxes through biochemical pathways that can be quantified with isotope tracers. Bond rearrangements quantified through a single <sup>13</sup>C tracer provide considerable information for hetero- or mixotrophic plant systems with steady state isotopic labeling<sup>1–8</sup> and, when transient labeling is considered, provide insights on autotrophic<sup>9–17</sup> or heterotrophic metabolism<sup>18–20</sup>. Using targeted approaches, gas chromatography mass spectrometry (GC-MS)<sup>21,22</sup>, and liquid chromatography tandem mass spectrometry (LC-MS/MS) can provide rigorous quantification of compounds and their isotopologues<sup>23–25</sup> to analyze metabolic network function and may consider spatial attributes through imaging-based MS techniques<sup>26–29</sup>. The attention given to developing different methods reflects the complexity of metabolism that is difficult to assess and can benefit from more than one labeling experiment with different <sup>13</sup>C substrates<sup>4,6–8,19,30–34</sup> or <sup>2</sup>H, <sup>15</sup>N or <sup>18</sup>O-based labeling to resolve pathway use<sup>35–43</sup>.

Technological advances now allow monitoring and quantification of large numbers of metabolites through untargeted approaches<sup>44–46</sup> and can link with multi-omics platforms to assess the operational state of a metabolic network, but these new measurement capabilities present several challenges

and directions<sup>47–53</sup> that could partially be addressed through incorporation of tracer studies. Though the acquisition of additional data is not a bottleneck for such studies, manual curation and integration of isotopologues is time consuming and not always compatible with the extensive number of replicates needed for statistically meaningful results. Further, isotopologue monitoring can result in overlapping signals. For example, dual-labeled tracers (e.g., <sup>13</sup>C<sup>15</sup>N-glutamine) produce changes in mass to charge ratio (i.e., *m/z*) in labeled metabolites that indicate the addition of a single neutron from either heavy element and is nominally the same for heavy isotopes from different elements (Fig. 1a). Thus, the capacity to resolve additional fluxes from combining information from multiple heavy isotopes within the same experiment is limited when nominal *m/z* values are recorded.

High resolution mass spectrometry (HRMS) can distinguish metabolites (and isotopologues of heteroatoms) through accurate measurement of *m/z* values that differ by <0.01 Da (Fig. 1a) providing a largely untapped framework to extend isotopic studies and metabolic flux analyses<sup>38,54</sup>. The inclusion of a second heavy isotope in the experimental design increases mass balancing constraints and improves the sensitivity of flux estimates beyond the enhanced data richness associated with additional isotopologue

<sup>1</sup>Donald Danforth Plant Science Center, St. Louis, MO 63132, USA. <sup>2</sup>Chemical and Biomolecular Engineering, Vanderbilt University, Nashville, TN 37235, USA. <sup>3</sup>London Research and Development Center, London, ON N5V 4T3, Canada. <sup>4</sup>Department of Biology, University of Western Ontario, London, ON N6A 5B7, Canada. <sup>5</sup>Department of Molecular Physiology and Biophysics, Vanderbilt University, Nashville, TN 37235, USA. <sup>6</sup>Agricultural Research Service, US Department of Agriculture, St. Louis, MO 63132, USA. <sup>7</sup>Present address: Jack H. Skirball Center for Chemical Biology and Proteomics, The Salk Institute for Biological Studies, La Jolla, CA 92037, USA. <sup>8</sup>These authors contributed equally: Shrikaar Kambhampati, Allen H. Hubbard, Somnath Koley. ✉e-mail: [skambhampati@salk.edu](mailto:skambhampati@salk.edu); [doug.allen@ars.usda.gov](mailto:doug.allen@ars.usda.gov)



measurements. HRMS instruments scan and quantify thousands of  $m/z$  values per second and can easily distinguish isotopologues and compositional changes in mass, for example, the nominal change in  $m/z$  of two Daltons that results from either inclusion of two carbon-13 atoms or alternatively a bond saturation or other chemical modifications<sup>38,54</sup>. Because HRMS can measure  $m/z$  values quickly, many metabolites actively incorporating isotopes can be identified through an unsupervised, unbiased approach with untargeted data acquisition that is blind to the metabolic

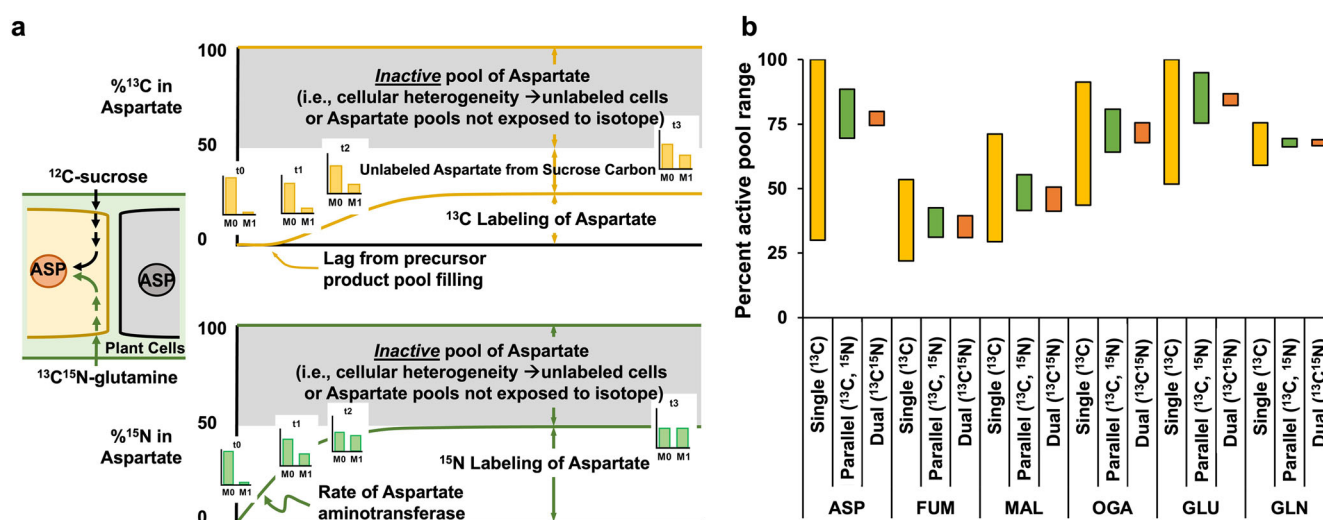
network. The enumerating and quantification of metabolomic features detected in samples enriched with multiple stable isotopes remains an analytical bottleneck<sup>55,56</sup>, as the number of multivariate isotopologues<sup>35</sup> of a given compound ( $I$ ) increases exponentially with the number of labeled atoms ( $N$ ) for each element ( $e$ ) of the molecular formula, following Eq. (1).

$$I = \prod (N_e + 1) \quad (1)$$

### Fig. 1 | Demonstration of the use of high-resolution mass spectrometry (HRMS) to distinguish dual labeled isotopologues and establish fluxes with confidence.

**a** HRMS has the capacity to resolve dual labeled isotopologues that are indistinguishable using nominal resolution MS, a schematic example with expected separation is shown. M0, M1, M2, ... Mn represent the isotopologues that differ by a unit mass, resulting from labeling. **b** Ten-day old Arabidopsis seedlings, transferred to medium contained [ $^{13}\text{C}_5, ^{15}\text{N}_2$ ] glutamine and collected after time course labeling, were used for HRMS data acquisition. Acquired data were first pre-processed using XCMS, and post-processing data analysis was performed using SIMPEL. Data tables exported from SIMPEL were used for INST-MFA with INCA. **c** Network representing the biochemical reactions modeled with INCA to estimate fluxes. Flux values ( $\text{nmol g}^{-1} \text{h}^{-1}$ ) for each reaction are shown next to the arrows along with CIs in parentheses. Color coding is as follows: yellow represents fluxes established from a single  $^{13}\text{C}$  experiment, green reflects a map with parallel  $^{13}\text{C}$  and  $^{15}\text{N}$  experiments, and the orange fluxes resulted from a dual labeled  $^{13}\text{C}^{15}\text{N}$  experiment. Measurements for isotope enrichment within metabolites for all three models were obtained from the same [ $^{13}\text{C}_5, ^{15}\text{N}_2$ ] glutamine labeling experiment, by using different combinations of isotopologues that represent a single ( $^{13}\text{C}$ ), multiple parallel ( $^{13}\text{C}$  and  $^{15}\text{N}$ ) or dual

labeled ( $^{13}\text{C}^{15}\text{N}$ ) experiments (Supplementary data 1, table 1). **d** As an example of the comparative information for modeling, glutamine that has 5 carbons and 2 nitrogen atoms is provided. In the  $^{13}\text{C}$  stand-alone experiment there are 6 isotopologue measurements (M0-M5) (used for modeling the single  $^{13}\text{C}$  labeling experiment). For the  $^{15}\text{N}$  stand-alone experiment there are 3 isotopologue measurements (M0-M2). Thus, as two separate experiments there are 9 measurements (used for modeling the parallel labeling experiment). When a single  $^{13}\text{C}^{15}\text{N}$  experiment is made, a nominal mass instrument would not be able to distinguish whether the increase in  $m/z$  was due to N or C, thus only the additional  $m/z$  values (i.e., M0-M7 isotopologues, a total of 8 measurements) were modeled. Complete HRMS analysis provides fine detail and distinguishes each elemental isotopologue such that 18 measurements can be made and used in principle (data from these measurements were summed to obtain the  $^{13}\text{C}^{15}\text{N}$  isotopologues (M0-M7), the  $^{13}\text{C}$  isotopologues (M0-M5), and the  $^{15}\text{N}$  isotopologues (M0-M2) from a single dual labeling experiment). In general, the number of measurements will be dependent on the number of experiments and tracers, however, for the HRMS case the 18 measurements from soft ESI-MS without fragmentation to the backbone can be generalized using Eq. (1). Abbreviations are listed within Supplementary Data 1, Supplementary Tables 2 and 3.



**Fig. 2 | Comparison of active vs inactive pool determination between single, parallel and dual-labeled MFA. a** Hypothetical inactive pool description for aspartate that is a consequence of unlabeled sucrose in the medium and labeling from  $^{13}\text{C}^{15}\text{N}$  glutamine. **b** 95% confidence intervals (CIs) for active pool estimates

using  $^{13}\text{C}$ -based INST-MFA, compared to parallel  $^{13}\text{C}$  and  $^{15}\text{N}$ -based MFA and a dual labeled  $^{13}\text{C}^{15}\text{N}$ -based MFA for key metabolites in the model, plotted as a range. Abbreviations are listed within Supplementary Data 1, Supplementary Table 3.

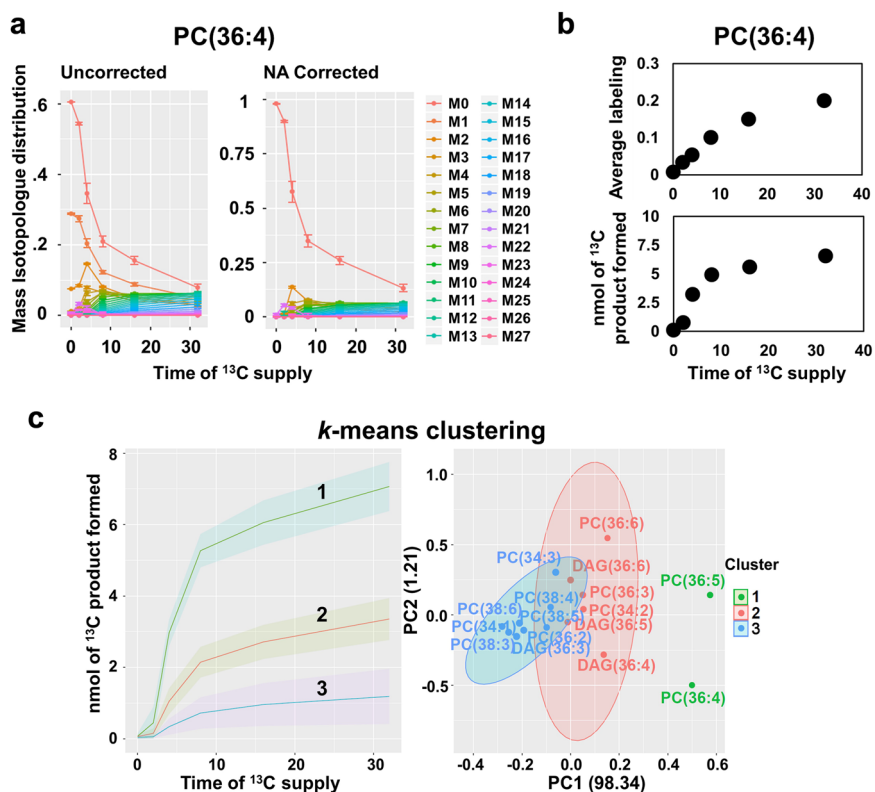
Computational tools compatible with HRMS isotopologue datasets<sup>57–64</sup> offer technical advances in workflows, address changes in instrumental capabilities and technology, and have started to confront bottlenecks in the analyses of isotope enriched HRMS data and their biological implications. Here a new tool is described, Stable Isotope-assisted Metabolomics for Pathway ELucidation (SIMPEL) that capitalizes on HRMS data collected from transient stable isotope labeling experiments. SIMPEL automates post-processing data analyses of isotope enriched metabolomics datasets, generates isotopologue distributions, performs natural abundance corrections along with some global analyses to streamline biological interpretations. Two case studies are presented, where the utility of SIMPEL is highlighted to understand central carbon and lipid metabolism. In the first description, a dual-isotope labeling experiment processed with SIMPEL was used for isotopically nonstationary metabolic flux analysis, to resolve challenging fluxes that traditionally require data from multiple labeling experiments. Providing network-constraining information by tracking two elements in a single labeling experiment enabled more precise flux estimates. In the second example, HRMS-based lipidomics data were processed using SIMPEL to describe dynamic lipid metabolism. Following metabolite precursor-product relationships, the role of PC and DAG in acyl editing was considered within oilseeds.

## Results and discussion

SIMPEL was designed as a data analysis tool to compare labeling and exact mass information enabling determination of the elemental composition of isotopically enriched compounds and application of correlation-based approaches to cluster metabolite peaks with similar patterns of isotopic enrichment from transient labeling experiments. SIMPEL eliminates challenges in data analysis by providing a platform to compatibly link untargeted metabolomics from HRMS with isotopologue labeling from multi-tracer transient labeling experiments. The software combines pre-processed metabolomics feature-lists from contemporary untargeted metabolomics data analysis software such as XCMS<sup>65</sup> or MZmine2<sup>66</sup> with comprehensive lists of compounds of interest including molecular formulae and retention time to search for and identify target metabolites and their isotopologues. Post-processing data-analyses that can be performed with SIMPEL include:

1. Enumeration of isotopologues for a given chemical formula through  $m/z$  tolerances,
2. Identification of isotopologues from libraries of chemical formulae and retention times,
3. Calculation of isotopologue distribution and average labeling per compound from intensities,
4. Correction for natural abundance (NA) with IsoCorrector<sup>67</sup>,

**Fig. 3** |  $^{13}\text{C}$  labeled lipidomic data extraction, plotting, and analyses from pre-processed data using SIMPEL. **a** Isotopologue distribution plot, where the abundances of each of the isotopologues as a proportion of the sum were plotted on the y-axis and the duration of  $^{13}\text{C}$  labeling (in hours) is shown on x-axis. Both the uncorrected and natural abundance corrected plots generated with SIMPEL are shown for PC(36:4) as an example (Supplementary Data 1, Supplementary Table 5). **b** Average labeling calculated using the isotopologue distributions and nmol of  $^{13}\text{C}$  product formed, the latter of which incorporates accurate quantification data for lipids (Supplementary Data 1, Supplementary Table 7). **c**  $k$ -means clustering plot with clusters represented as nmol of  $^{13}\text{C}$  product formed on the y-axis (lines represent the mean enrichment and shaded regions represent standard deviation) and duration of  $^{13}\text{C}$  label incorporation on the x-axis. A 2D representation (1<sup>st</sup> and 2<sup>nd</sup> principal components) of the same clusters with time collapsed to highlight the compounds present in each cluster is shown on the right (data presented in Supplementary Data 1, Supplementary Table 7).



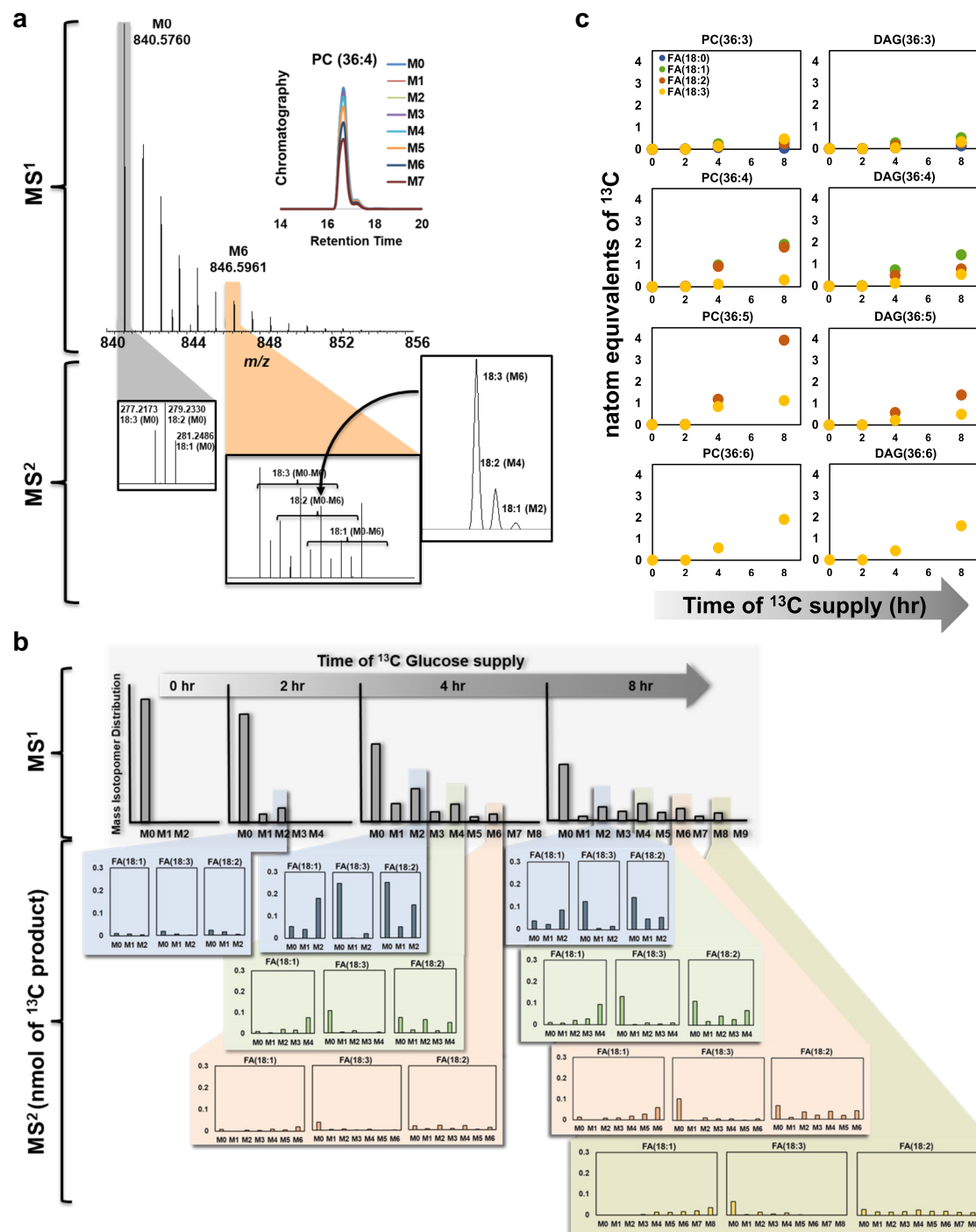
5. Visualization and export of isotopologue distributions and average labeling for all compounds,
6. Global analyses such as principal component analysis (PCA), hierarchical or  $k$ -means clustering of compounds based on label enrichment over time.

Two case studies are provided involving SIMPEL. In the first investigation, ten-day-old *Arabidopsis* roots were grown on a predefined carbon and nitrogen rich medium<sup>68</sup> supplemented with [ $^{13}\text{C}_5^{15}\text{N}_2$ ] glutamine (Fig. 1a) and sampled over time (0–8 h) prior to SIMPEL-based data extraction and metabolic flux analysis. An untargeted metabolomics dataset generated by a qExactive Orbitrap MS and pre-processed with XCMS was imported into SIMPEL (Fig. 1b) resulting in a post-processed data table containing isotopologue distributions for metabolic intermediates that were used to perform isotopically non-stationary metabolic flux analysis (INST-MFA) with INCA<sup>69</sup> (Fig. 1c). Metabolic networks that included either  $^{13}\text{C}$  or  $^{13}\text{C}^{15}\text{N}$  atom transitions were modeled with labeling information from a single  $^{13}\text{C}$  experiment, parallel  $^{13}\text{C}$  and  $^{15}\text{N}$  experiments, or with the addition of  $^{13}\text{C}^{15}\text{N}$  isotopologues that represent information obtained from a single dual-labeling experiment (Fig. 1d). Without HRMS, three independent tracer experiments would be required to obtain these datasets (i.e., one with  $^{13}\text{C}$ , one with  $^{15}\text{N}$ , and one with combined  $^{13}\text{C}/^{15}\text{N}$  labeling). However, all three datasets could be derived from a single  $^{13}\text{C}^{15}\text{N}$ -glutamine labeling experiment by summing isotopologue abundances contained in the comprehensive HRMS dataset (Fig. 1d). As noted in Fig. 1d, combining the three nominal mass resolution  $^{13}\text{C}^{15}\text{N}$  datasets provides fewer measurements than modeling the complete HRMS isotopologue distribution, but INCA does not currently have the capability to model HRMS data directly.

The inclusion of a second isotope (i.e.,  $^{15}\text{N}$ ) enabled active metabolite pool assignments and identification of differences that were a consequence of active pools that incorporated unlabeled carbon from sucrose, versus pools that were inactive or not turned over during the time frame of the experiment (Fig. 2a). The inability to distinguish these descriptions normally confounds labeling interpretations and subsequent flux analyses<sup>70</sup>.

Here, MFA based on parallel or dual labeling with  $^{15}\text{N}$  reduced inactive pool estimates and recast the associated confidence intervals (Fig. 2b). The reactions that use nitrogen differ from those that exclusively use carbon (e.g., aminotransferase reactions relative to carbon shuttling through central metabolism for organic building blocks), thus, pairing isotopes from multiple elements that trace different aspects of metabolism and achieve isotopic equilibrium in metabolites at different rates can help define more realistic metabolic networks. The confidence intervals for the active pool estimates were significantly reduced for the model including dual labeling compared to parallel  $^{13}\text{C}$  and  $^{15}\text{N}$  labeling experiments (Fig. 2b), likely due to the additional paired  $^{13}\text{C}^{15}\text{N}$  isotopologues that describe concomitant use of labeled precursors within a metabolite pool.

Compared to conventional flux studies involving multiple parallel labeling experiments<sup>67,9,33</sup>, a single, dual-labeled experiment can potentially provide additional information<sup>71</sup> from the heavy isotopologue connections that can be quantified through HRMS with SIMPEL resulting in more sensitively determined fluxes and smaller confidence intervals (CIs) (Fig. 1c). Fluxes through organelle transport steps that do not involve metabolite interconversions (e.g., OAA transport between mitochondria and cytosol) are difficult to determine with certainty but could be determined by the software through adjacent reactions and thus can benefit from a second elemental isotope signature. Nitrogen label transfer in addition to the  $^{13}\text{C}$  movement during the conversion of OAA to aspartate enables a better determination of the OAA transfer flux. Bifurcated pathways such as glutamate dehydrogenase (GDH) or glutamine oxoglutarate aminotransferase (GOGAT) could also be resolved based on the nitrogen transfer through coordination with aspartate aminotransferase (ASPART). While two parallel labeling experiments with  $^{13}\text{C}$ -glutamine and  $^{15}\text{N}$ -glutamine, separately, resolved active pool sizes and reliably determined transport fluxes with improved precision compared to a single  $^{13}\text{C}$  experiment, the dual-labeled substrate in combination with HRMS and SIMPEL established tighter CIs for all fluxes likely due to the provision of additional labeling information tracking two elements that were complementary in constraining the flux solution. On average, the  $^{13}\text{C}$  and  $^{15}\text{N}$  parallel labeling-



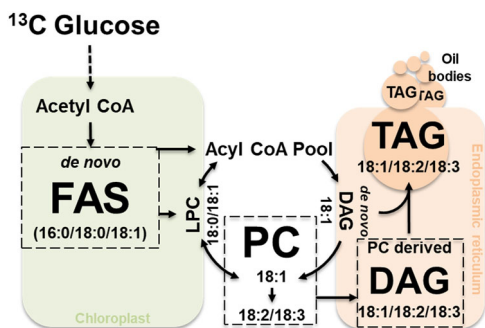
**Fig. 4 | Determination of acyl-chain labeling within lipid molecular species using MS<sup>2</sup> at high resolution.** **a** Tandem MS spectra of PC(36:4) and its isotopologues that co-elute in chromatography. HRMS resolves the different isotopologues, highlighted in gray, M0 (840.5760), and orange, M6 (846.5961), in the MS<sup>1</sup> scan. MS<sup>2</sup> spectra for the unlabeled isotopologue shows 18:1, 18:2 and 18:3 FAs, while the M6 isotopologue shows combinations of labeled FAs that can further be resolved using HRMS. **b** Relative intensities of the isotopologues of PC(36:4) were plotted for each time point (0, 2, 4 and 8 h of <sup>13</sup>C glucose labeling) using the MS<sup>1</sup> data analyzed by SIMPEL.

MS<sup>2</sup> spectra for each of the isotopologues were used to determine the distribution of FA labeling within the acyl chains of PC(36:4). Data at MS<sup>2</sup> represents isotopologue distribution that is adjusted for pool size and the proportions of FAs that constitute PC(36:4) (see methods, Supplementary File 2). **c** <sup>13</sup>C enrichment in acyl chains of PC(36:x) and DAG(36:x) represented as natom equivalents of <sup>13</sup>C (see methods, Supplementary Data 1, Supplementary Tables 8 and 9 for detailed description and calculation) using MS<sup>2</sup> data obtained upon fragmentation of isotopologues.

based MFA and <sup>13</sup>C<sup>15</sup>N dual labeled MFA reduced the CIs by 71 and 84%, respectively, compared to the single <sup>13</sup>C-based MFA (Supplementary Data 1, Supplementary Table 3).

In a second study, SIMPEL was used to probe stable isotope labeled lipidomics data obtained with HRMS. HRMS distinguished the

incorporation of two <sup>13</sup>C molecules ( $\Delta m/z = 2.00671$ ) from a degree of fatty acid unsaturation ( $\Delta m/z = 2.01565$ ) based on  $m/z$  difference of 0.009 Da<sup>38</sup>. Developing *Camelina sativa* seeds, which represent stages corresponding to oil biosynthesis and accumulation, were labeled with [U-<sup>13</sup>C<sub>6</sub>] glucose in a time-course experiment (0–32 h). Embryos were excised from pods, and



**Fig. 5 | Simplified description of seed oil biosynthesis and the central role played by PC.** Seed oil biosynthesis based on  $^{13}\text{C}$  glucose labeling involves the movement of acyl chains onto and off PC for desaturation, known as acyl editing<sup>72–74</sup> and the PC may additionally serve as the shuttling mechanism for the acyl chain export from the chloroplast to the ER<sup>38,75,79</sup>. FAS Fatty acid synthesis, LPC Lyso phosphatidylcholine, PC Phosphatidylcholine, DAG Diacylglycerol, TAG Triacylglycerol.

cotyledons were separated and assessed to gain insights into fatty acid biosynthesis and lipid metabolic pathways in plants. Untargeted lipidomic datasets generated using a Thermo Fusion Lumos Tribrid MS were pre-processed using XCMS and automatically analyzed by SIMPEL (Fig. 3a) resulting in average labeling descriptions for diacylglycerol (DAG) and phosphatidylcholine (PC) that are central to triacylglycerol assembly and polyunsaturation in plants which utilize acyl editing mechanisms<sup>72–74</sup>. Much of what is known about plant lipid biosynthesis has been and continues to be established through inspection of  $^{14}\text{C}$  labeled lipids<sup>72,75–77</sup>; however, HRMS and SIMPEL provide a complementary technique to elucidate the precursor-product relationships and rates of labeling of pathway intermediates in living systems when stable isotope investigations are appropriate.

To enable direct comparison of  $^{13}\text{C}$  enrichment between different lipid species, pool size (nmol) was accounted for (see Supplementary Data 1, Supplementary Tables 4–7) to obtain the quantity of  $^{13}\text{C}$  labeled molecules formed<sup>16,78</sup> (Fig. 3b). A *k*-means clustering analysis of the quantities of  $^{13}\text{C}$  PC and DAG species formed over the time course of labeling revealed three distinct clusters (Fig. 3c). Cluster 1 consisted only of two PCs and was the fastest labeled group, whereas clusters 2 and 3 contained a mixture of other PC and DAG molecular species (Fig. 2c). The results, obtained from a single untargeted assessment of  $^{13}\text{C}$ -labeled lipids, confirmed previous  $^{14}\text{C}$  labeling results that PC is the most rapidly labeled lipid in seed oil production<sup>72,75</sup>. Further, the identification of rapidly labeled PC(36:4) and PC(36:5) species suggested the role of these lipids in connecting fatty acid biosynthesis in the chloroplast with lipid assembly in the endoplasmic reticulum. Membrane lipids such as PC may be a conduit for the direct shuttling of acyl chains from the chloroplast to the ER<sup>38,75,79</sup> through contact sites or to extend channeling mechanisms<sup>80–83</sup>. PC comprises a significant percentage of the lipids present at the chloroplast envelope<sup>84</sup> at around 40%<sup>75</sup>. Tandem MS analysis of  $^{13}\text{C}$  enrichment within the fatty acids of PC and DAG molecular species at 2-, 4- and 8-hour time points (Fig. 4a, b, Supplementary Data 1, Supplementary Tables 8 and 9) indicated highest labeling in 18:1 followed by 18:2 acyl chains at the earliest time point, which represented most of the label enrichment (Fig. 4c). By 4 h, acyl chains within PC(36:5), i.e., FA(18:2) and FA(18:3), were enriched, consistent with desaturation over time. DAGs presented a similar trend, but were delayed relative to PC, with FA(18:1) of DAG(36:4) being the most labeled at both 2 and 4 h while FA(18:2) and FA(18:3) accumulated higher  $^{13}\text{C}$  enrichments by 8 h. The comparable FA enrichment in PC and DAG molecular ions provided a semi-quantitative assessment of the important role of enzymes that interconvert DAG and PC through head group (choline) exchange such as PC:DAG choline phosphotransferase (PDCT)<sup>85</sup> and CDP-choline:DAG choline phosphotransferase (DAG-CPT)<sup>86</sup>. Our contemporary understanding of seed oil

biosynthesis involves de novo synthesis of saturated or monounsaturated fatty acids in the chloroplast, which are then converted to acyl-CoAs that are incorporated into PC for desaturation with acyl chains either released as part of acyl editing, passed to DAG through lipid headgroup exchange (Fig. 5), or used as the third acyl chain in TAG biosynthesis by phospholipid: diacylglycerol acyltransferase (PDAT)<sup>87</sup>. As PC and lysoPC (LPC) can be found on the surface of the outer chloroplast membrane<sup>75,84</sup>, and a significant percentage (30% or more) of lyso-phosphatidylcholine acyltransferase, LPCAT, activity is associated with chloroplasts<sup>88</sup>, the labeled FA(18:1) and FA(18:2) of PC is consistent with PC serving as a carrier of acyl chains that are channeled to the ER<sup>38,75</sup>.

Further, the analyses revealed significant labeling in 16:3/18:3 monogalactosyldiacylglycerol (MGDG) (Supplementary Fig. S1) that suggest chloroplast Kennedy pathway flux of 16:3 in plants not generally considered in green seeds of oil producing crops. Such examples highlight the opportunity for untargeted metabolomics to elucidate undescribed aspects of metabolism with stable isotope labeling and HRMS.

## Conclusions

SIMPEL provides high throughput identification and quantification of isotopologues from untargeted HRMS datasets arising from transient isotope labeling studies, presenting an opportunity to discover and describe novel metabolic activities. With dual labels ( $^{13}\text{C}$ ,  $^{15}\text{N}$ ), SIMPEL enabled more precise flux estimations of central metabolism from a single experiment. Similarly, stable isotope labeling of an oilseed indicated PC and DAG roles in acyl editing that were consistent with channeling of acyl chains through PC to the ER, and MGDG labeling that support the presence of an active chloroplast Kennedy pathway in a green oilseed. SIMPEL is available as an R package for comparative analysis of any pulse labeled stable isotope enrichment datasets generated using traditional metabolomic data acquisition.

## Methods

### Plant growth and experimental conditions

For the metabolomics study using dual-isotope labeling, wildtype *Arabidopsis* ecotype Columbia seeds were grown on vertical plates at 22 °C under continuous light (ca.  $70\ \mu\text{mol m}^{-2}\ \text{s}^{-1}$ ), on a defined nutrient medium previously described<sup>68</sup> for metabolomics with dual-isotopic labeling. The medium consisted of 10 mM potassium phosphate (pH 6.5), 5 mM  $\text{KNO}_3$ , 2 mM  $\text{MgSO}_4$ , 1 mM  $\text{CaCl}_2$ , 0.1 mM  $\text{FeNaEDTA}$ , micronutrients (50 mM  $\text{H}_3\text{BO}_3$ , 12 mM  $\text{MnSO}_4$ , 1 mM  $\text{ZnCl}_2$ , 1 mM  $\text{CuSO}_4$  and 0.2 mM  $\text{Na}_2\text{MoO}_4$ ), 1% sucrose and 1% agar. Ten-day old seedlings were transferred to plates containing the same medium, except the nitrogen source was replaced with 10 mM [ $^{13}\text{C}_5$ ,  $^{15}\text{N}_2$ ]glutamine. Root tissue was excised after exposure to medium containing labeled glutamine for 2, 4, 6 and 8 h to represent time course incorporation of carbon and nitrogen into metabolism. Untreated roots were used as unlabeled (0 h) controls. Each plate yielded ~100 mg of root tissue and served as a single replicate. Four replicates per sample type were collected and flash frozen using liquid  $\text{N}_2$  for total metabolite extraction.

Stable isotope labeled lipidomics was performed with developing embryos of *Camelina sativa*. Plants were grown in greenhouses with day/night temperature maintained at 22/20 °C, 40–50% relative humidity, and 16 h day/8 h night photoperiod. Intact siliques during the seed filling growth stage (15 days after fertilization) were excised and placed in sterile media containing a modified Linsmaier and Skoog medium<sup>89,90</sup> with Gamborg's vitamins (Sigma) and 5 mM MES buffer adjusted to pH 5.8. Fifty mM [ $^{13}\text{C}_6$ ]glucose was used as labeled substrate<sup>91</sup>, and the composition of the remaining carbon and nitrogen sources represented maternal phloem composition to minimize metabolic perturbation and to maintain pseudo in vivo conditions. Silique culturing was performed in a 96-well plate with 0.3 mL of medium and a single silique per well, under continuous light ( $250\ \mu\text{mol m}^{-2}\ \text{s}^{-1}$ ). Tissue was collected and flash frozen immediately after each time point (2, 4, 8, 16 and 32 h). Uncultured siliques excised from the maternal plant were used as unlabeled (0 h) controls. Frozen tissue was

sectioned, on top of dry ice, to excise embryo from the siliques and to separate cotyledons from the embryo axis. Cotyledon samples were extracted and analyzed for lipids in triplicates.

### Total metabolite extraction

Frozen *Arabidopsis* root tissue was homogenized using a tissue lyser, and extraction was carried out using 1 mL of 4:1 methanol: water (v/v) with incubation in an ultra-sonication bath for 30 min followed by shaking for 30 min at 4 °C. The mixture was then centrifuged at 21,000 × g for 10 min at 4 °C; supernatant was transferred into fresh tubes and evaporated to dryness using a speedvac centrifuge at ambient temperature. Dried residue was re-suspended in 200 µL of 1:1 methanol: water (v/v), filtered using 0.2 µm PTFE micro centrifuge filters and transferred to glass vials for HILIC-HRMS runs.

Frozen cotyledon samples from *Camelina* were homogenized using a tissue lyser and the extraction of lipids was carried out using a phase separation method previously described<sup>92</sup>. Briefly, 1 mL 7:3 methanol:chloroform (−20 °C) containing the ultimateSPLASH™ ONE lipid mix (Avanti Polar lipids, Alabaster, AL) as internal standard (1:20 dilution) was added to the samples, vortexed vigorously and incubated on a rotary shaker for 2 h at 4 °C. After incubation, 500 µL of ice-cold water was added to the samples, vortexed and centrifuged at 14,000 rpm at 4 °C for 10 min to achieve phase separation. The upper aqueous phase was carefully removed, 200 µL of methanol was added to the remaining organic phase containing lipids and centrifuged at 14,000 rpm for 5 min to pellet the debris. The organic phase (supernatant) was transferred to a glass tube and dried using a speedvac centrifuge. Samples were re-suspended in 200 µL of 49:49:2 acetonitrile: methanol: chloroform, filtered using 0.2 µm PTFE microcentrifuge filters and transferred to a glass vial for RPLC-HRMS analysis.

### Metabolomics data acquisition using HILIC-HRMS

Chromatographic separation using HILIC was achieved using an Agilent 1290 Infinity II UHPLC system equipped with a SeQuant® ZIC®-HILIC (100 × 2.1 × 3.5 µm) column (EMD Millipore, Burlington, MA). Mobile phases A and B were comprised of 5 mM ammonium acetate (pH 4.0) in water and 90% acetonitrile with 0.1% acetic acid, respectively. A flow rate of 0.3 mL min<sup>−1</sup> was used to elute compounds with the following gradient: 87% B for 5 minutes, decreased to 55% B over the next 8 min and held for 2.5 min before returning to 87% and equilibrating the column for 3 min. The heated electrospray ionization (HESI) conditions used were as follows; spray voltage, 3.9 kV (ESI+), 3.5 kV (ESI−); capillary temperature, 250 °C; probe heater temperature, 450 °C; sheath gas, 30 arbitrary units; auxiliary gas, 8 arbitrary units; and S-Lens RF level, 60%. Full MS data were collected using a Q-Exactive Quadrupole Orbitrap mass spectrometer (Thermo Fisher Scientific) in both positive and negative ionization mode separately from mass ranges 75–1100 *m/z* and 65–900 *m/z*, respectively, at 140,000 resolution. The automatic gain control (AGC) was set to 3 × 10<sup>6</sup> and maximum injection time (IT) used was 524 ms. Top 12 data dependent MS/MS (ddMS2) spectra were also collected for a representative pool of unlabeled samples with MS1 and MS2 data acquired at 35,000 and 17,500 resolution, respectively. The AGC target and maximum IT in the ddMS2 experiment were set to 1 × 10<sup>6</sup>, 128 ms for MS1 spectra and 1 × 10<sup>6</sup>, 64 ms for the MS2 spectra. A 2.0 Da isolation window and normalized collision energy of 25 were used for ddMS2 with the underfill ratio set to 1.0% along with 10 sec dynamic exclusion to reduce redundant spectra.

### Lipidomics data acquisition using RPLC-HRMS

Separations for lipidomics were achieved using the loading pump of a Dionex UltiMate 3000 RSLCnano system (Thermo Fisher Scientific) operating at a flow rate of 40 µL min<sup>−1</sup> equipped with a custom-made C8 column (100 × 0.5 × 5 µm) from Higgins Analytical Inc. (Mountain view, CA) re-packed from a nucleodur C8 Gravity column (Macherey-Nagel, Allentown, PA). Mobile phases comprised of 1% 1 M ammonium acetate, 0.1% acetic acid in water (A) and 1% 1 M ammonium acetate, 0.1% acetic acid in 7:3 (v/v) acetonitrile: isopropanol (B). The following gradient

modified from a previously described method<sup>93</sup> to adapt to micro flow was used; 0–1 min at 55% B, 4 min at 75% B, 12 min at 89% B, 15 min at 99% B, 18 min at 99% B and 20 min at 55% B followed by equilibration up to 30 min. The eluent was sprayed on to the HESI source of an Orbitrap Fusion Lumos Tribrid MS, operated with sheath gas, 25 arbitrary units; auxiliary gas, 5 arbitrary units; ion transfer tube temperature, 300 °C; vaporizer temperature, 100 °C; and S-lens RF level, 60. The spray voltage was 4 kV in both positive and negative modes. Full MS data were collected for mass ranges 450–1200 *m/z* at 240,000 resolution from both positive and negative modes simultaneously, using polarity switch. The AGC target was set to “Standard” and the maximum IT was set to 100 ms.

MS/MS data were collected for representative pooled unlabeled and labeled samples using AcquireX Exclusion-Inclusion lipid characterization template with iterative injections. MS1 data within this experiment were obtained at 120,000 resolution from a mass range of 450–1200 with AGC target set to “standard” and maximum IT set to “auto”. MS2 data were also collected at 120,000 resolution within the orbitrap detector, to achieve enough mass resolution for distinguishing <sup>13</sup>C labeled isotopologues from acyl chain desaturation. A guide to determining the required resolution to sufficiently distinguish *m/z* of interest is provided within supplementary information (Supplementary text T1). The collision energy mode was set to “assisted” with 25, 30 and 35% collision energies with an activation time of 20 ms. The AGC target and maximum IT for MS2 experiments were also set to “standard” and “auto”, respectively. A 0.5 Da isolation window was used and a cycle time of 1.5 s was used within the data dependent mode. The intensity threshold filter, set to 5 × 10<sup>4</sup>, and apex detection were enabled to prioritize ion isolation at their maximum intensity. Data dependent acquisition was also enabled to obtain MS2 on the most intense ions if target ions were not detected within a given scan.

### Data pre-processing and creation of compound database

Full MS data files in Thermo .RAW format obtained in profile mode were first centroided by conversion to mzML format using the MSConvert application within ProteoWizard<sup>94</sup> with peak picking filter applied. For data collected with polarity switch, the positive and negative mode data were separated using the subset filter within MSConvert. Files were imported into R using the XCMS package<sup>65</sup>, and features were detected using the centWave<sup>95</sup> method and 5 ppm mass tolerance. Pre-filter was set to 6 scans with a minimum intensity of 5000; signal to noise ratio and the threshold for noise were set to 5 and 1 × 10<sup>6</sup>, respectively. Retention time alignment was performed using the obiwrap<sup>96</sup> method, grouping included features present in at least 25% of all samples, allowable retention time deviation was 10 s, and *m/z* width was set to 0.01. The “fillPeaks” function was used with default settings, and data was not imputed. All features identified were exported to Excel for use as an input data table for SIMPEL. MS/MS data collected from representative unlabeled samples were used within a commercial software, Compound Discoverer 3.0 (Thermo Scientific), following the “MaxID” workflow. Only the compounds manually verified to have a positive match with the mzCloud MS/MS and/or LIPID MAPS database and a valid chromatographic peak were used to build custom compound databases, for both HILIC and lipidomic methods, which contained chemical formulae and retention times to be used as an input for SIMPEL.

### Post-processing of stable isotope enriched data using SIMPEL

The XCMS pre-processed data tables and the compound databases (described above) for both dual-isotope labeled metabolomics and single-isotope labeled lipidomics datasets were packaged within SIMPEL as example data for users. All the isotopologues and their *m/z* values for compounds within the database were matched against retention time and identified from the pre-processed data using the function “get\_table\_objects\_NA\_corrected”. The function also includes user defined tolerances for *m/z* and *rt* values, which may not be universally applied for all isotope enriched metabolomics datasets, an iterative approach to determine optimal settings may be necessary. The same function also performs NA correction, calculates isotopologue distributions and average labeling for

both NA corrected and uncorrected data and stores them as data objects that are used for plotting based on user preferences. A tutorial describing the usage of each of the functions within *SIMPEL*, and equations that *SIMPEL* uses for calculations, are provided as Supplementary File 3.

### Isotopomer network and determination of fluxes

A reaction network describing glycolysis, TCA cycle, and related amino acid biosynthesis was constructed using a subset of a previously published *Arabidopsis* model<sup>11</sup> and by consulting biochemical literature. The model was adjusted to include carbon and nitrogen atom transitions for the dual-isotope labeling experiment. A list of reactions, and abbreviations can be found in supplementary data 1, table 3. Metabolic fluxes and confidence intervals were estimated using INST-MFA through the MATLAB-based Isotopomer Network Compartmental Analysis (INCA) package, (Version 2.0)<sup>97</sup>. Best-fit values for fluxes were obtained by repeating the evaluation a minimum of 1000 times from random initial values and subjecting the optimized values to a  $\chi^2$  goodness of fit test. 95%-confidence intervals were computed for all estimated parameters by evaluating the sensitivity of the sum-of-squared residuals to parameter variations<sup>98</sup>. The MATLAB model files for the three experimental designs described in the results are also provided as Supplementary Data 2, 3 and 4 for convenience.

### Accurate quantification and label enrichment of lipids and their acyl chains

Unlabeled control samples of *Camelina* cotyledons, that serves as controls, were spiked with nonendogenous, deuterium labeled internal standard mix, i.e. ultimateSPLASH™ ONE and used for accurate quantification of PC and DAG lipid molecular species. Both PC and DAG standards contained a d5 labeled C17 chain in the *sn*-1 position and varying chain lengths from C14 to C22 in the *sn*-2 position. Each standard within the mix was considered a “bucket”, and the concentration of each analyte was calculated using multiple buckets based on its acyl chain composition as determined using MS<sup>2</sup>. For example, PC(34:x) that contains a C16 and a C18 acyl chain was quantified once using d5PC(17:0/16:1) as bucket 1 and again using d5PC(17:0/18:1) as bucket 2. A mean value of the calculated concentration obtained from all possible buckets was treated as the accurate quantity of analyte. Steady-state metabolism was assumed with an expectation that the pool sizes of lipid intermediates such as PCs and DAGs do not change throughout the course of the labeling experiment. PC and DAG quantities obtained in nmol amounts were used to convert NA corrected isotopologue distributions generated for the pulse labeled samples, through *SIMPEL*, into nmol abundances of isotopologues. A sum of all the labeled isotopologues (M1-Mn) in nmol represented the amount of <sup>13</sup>C product formed over the course of the labeling experiment. These data were used to for *k*-means clustering analysis to enable comparison of label enrichment between different pools of PCs and DAGs that are otherwise incomparable due to their differences in pool sizes<sup>30</sup>. MS2 data collected from the labeled samples were used to manually assess and integrate the acyl chain composition and labeling within each of the PC and DAG molecular species. IsoCorrector<sup>67</sup> was used to correct for natural abundance, and the corrected isotopologue distributions were converted to natom equivalents of <sup>13</sup>C, modified and adopted from Arrivault et al.<sup>78</sup>, using the pool size information.

### <sup>13</sup>C label quantification and data interpretation of acyl chains within Lipids using MS<sup>2</sup> fragments

<sup>13</sup>C Glucose labeled cotyledons collected from 2, 4 and 8 h time points along with unlabeled seeds that served as controls were run on a Thermo Orbitrap Fusion Lumos tribrid MS through an untargeted exclusion-inclusion lipid characterization template with iterative injections as described in the methods section and referred to by the manufacturer as AcquireX. The MS<sup>2</sup> ion fragments containing acyl chains and their labeled isotopologues were used to calculate the enrichment in the fatty acids of the lipid molecular species. As an example, negative ion mode fragmentation of *m/z* of 846.5961 that is the formate adduct of the M6 isotopologue of PC(36:4) with (18:1/18:3) and (18:2/18:2) acyl chains contains M0-M6

of 18:1, 18:2 and 18:3 fatty acids (Fig. 4a). The M6 weighting is conserved through summed pairs (i.e., M0/M6, M1/M5, M2/M4, M3/M3, M4/M2, M5/M1, M6/M0) of labeling in acyl chains. Intensities of MS<sup>2</sup> acyl chain labeling distributions were averaged for four scans per precursor *m/z* for M2, M4, M6 and M8 isotopologues. Labeling distributions for each labeling time (2, 4 and 8 h) were corrected for NA, using IsoCorrector, and converted into nmols of <sup>13</sup>C isotopologue by accounting for pool size of the lipid and its isomeric proportions (Fig. 4b) (see Supplementary Data 1, Supplementary Tables 8 and 9 for calculations). As a proxy for labeling in lipids, even labeled isotopologues offset by factors of two (M0, M2, M4, M6 and M8) were integrated and analyzed at each time point representing the total labeling within lipid coming from a uniformly labeled source of glucose. The <sup>13</sup>C enrichment within acyl chains of each lipid were determined by weighting labeled acyl chains by the number of carbons and summing precursor isotopologue values at a given time point to obtain natom equivalents of <sup>13</sup>C (Fig. 4c, Supplementary Data 1, Tables 8 and 9).

### Statistics and reproducibility

Both dual-labeling and lipid labeling experiments were performed independently and reproduced at least twice to ensure robustness of the results. A minimum of three (lipid labeling) and four (dual labeling) independent biological replicates were used. Chi-squared goodness of fit, 95% confidence intervals or *k*-means clustering were used as appropriate, for statistical analyses and were described within the figure legends and methods section.

### Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

### Data availability

Raw and pre-processed data for stable isotope labeled metabolomics using *Arabidopsis* roots and lipidomics using *Camelina* seeds is available at the NIH Common Fund's National Metabolomics Data Repository (NMDR) website, the Metabolomics Workbench, <https://www.metabolomicsworkbench.org/>, where they have been assigned study IDs ST002240 and ST002239 respectively<sup>99</sup>. The data for metabolomics and lipidomics experiments can be accessed directly via their project ID, PR001429 using the <https://doi.org/10.21228/M80X3B>.

### Code availability

The tool *SIMPEL* is freely available for public use at <https://github.com/SIMPELmetabolism/SIMPEL>, as an R package along with test data and source code. A tutorial is also included within GitHub with detailed instructions on *SIMPEL* usage with the test data. The tool is also available at the metabolomics workbench at [https://www.metabolomicsworkbench.org/data/simpel\\_load.php](https://www.metabolomicsworkbench.org/data/simpel_load.php), as a web tool.

Received: 3 August 2022; Accepted: 23 January 2024;

Published online: 12 February 2024

### References

- Lonien, J. & Schwender, J. Analysis of metabolic flux phenotypes for two *Arabidopsis* mutants with severe impairment in seed storage lipid synthesis. *Plant Physiol.* **151**, 1617–1634 (2009).
- Williams, T. C. R. et al. Metabolic network fluxes in heterotrophic *Arabidopsis* cells: Stability of the flux distribution under different oxygenation conditions. *Plant Physiol.* **148**, 704–718 (2008).
- Morley, S. A. et al. Expression of malic enzyme reveals subcellular carbon partitioning for storage reserve production in soybeans. *New Phytologist* <https://doi.org/10.1111/NPH.18835> (2023).
- Masakapalli, S. K., Kruger, N. J. & Ratcliffe, R. G. The metabolic flux phenotype of heterotrophic *Arabidopsis* cells reveals a complex response to changes in nitrogen supply. *Plant J.* **74**, 569–582 (2013).



5. Masakapalli, S. K. et al. Metabolic flux phenotype of tobacco hairy roots engineered for increased geraniol production. *Phytochemistry* **99**, 73–85 (2014).
6. Allen, D. K. & Young, J. D. Carbon and nitrogen provisions alter the metabolic flux in developing soybean embryos. *Plant Physiol.* **161**, 1458–1475 (2013).
7. Carey, L. M. et al. High flux through the oxidative pentose phosphate pathway lowers efficiency in developing camelina seeds. *Plant Physiol.* **182**, 493–506 (2020).
8. Tsogtbaatar, E., Cocuron, J. C., Alonso, A. P. & Lunn, J. Non-conventional pathways enable pennycress (*Thlaspi arvense* L.) embryos to achieve high efficiency of oil biosynthesis. *J. Exp. Bot.* **71**, 3037–3051 (2020).
9. Koley, S. et al. Metabolic synergy in *Camelina* reproductive tissues for seed development. *Sci. Adv.* **8**, 7683 (2022).
10. Chu, K. L. et al. Metabolic flux analysis of the non-transitory starch tradeoff for lipid production in mature tobacco leaves. *Metab. Eng.* **69**, 231–48 (2022).
11. Ma, F., Jazmin, L. J., Young, J. D. & Allen, D. K. Isotopically nonstationary <sup>13</sup>C flux analysis of changes in *Arabidopsis thaliana* leaf metabolism due to high light acclimation. *Proc. Natl Acad. Sci. USA* **111**, 16967–16972 (2014).
12. Fu, X., Gregory, L. M., Weise, S. E. & Walker, B. J. Integrated flux and pool size analysis in plant central metabolism reveals unique roles of glycine and serine during photorespiration. *Nat. Plants* **9**, 169–178 (2022).
13. Xu, Y., Fu, X., Sharkey, T. D., Shachar-Hill, Y. & Walker, B. J. The metabolic origins of non-photorespiratory CO<sub>2</sub> release during photosynthesis: a metabolic flux analysis. *Plant Physiol.* **186**, 297–314 (2021).
14. Medeiros, D. B. et al. <sup>13</sup>CO<sub>2</sub> labeling kinetics in maize reveal impaired efficiency of C<sub>4</sub> photosynthesis under low irradiance. *Plant Physiol.* **190**, 280–304 (2022).
15. Treves, H. et al. Carbon flux through photosynthesis and central carbon metabolism show distinct patterns between algae, C<sub>3</sub> and C<sub>4</sub> plants. *Nat. Plants* **8**, 78–91 (2021).
16. AuBuchon-Elder, T. et al. Sterile spikelets contribute to yield in sorghum and related grasses. *Plant Cell* **32**, 3500–3518 (2020).
17. Weissmann, S. et al. Interactions of C<sub>4</sub> Subtype Metabolic Activities and Transport in Maize Are Revealed through the Characterization of DCT2 Mutants. *Plant Cell* **28**, 466–484 (2016).
18. Kambhampati, S. et al. Temporal changes in metabolism late in seed development affect biomass composition. *Plant Physiol.* **186**, 874–890 (2021).
19. Cocuron, J. C., Koubaa, M., Kimmelfield, R., Ross, Z. & Alonso, A. P. A Combined Metabolomics and Fluxomics Analysis Identifies Steps Limiting Oil Synthesis in Maize Embryos. *Plant Physiol.* **181**, 961–975 (2019).
20. Smith, E. N., Ratcliffe, R. G. & Kruger, N. J. Isotopically non-stationary metabolic flux analysis of heterotrophic *Arabidopsis thaliana* cell cultures. *Front Plant Sci.* **13**, 1049559 (2023).
21. Deller, Y., Filangi, O. & Bouchereau, A. Evaluation of GC/MS-Based <sup>13</sup>C-Positional Approaches for TMS Derivatives of Organic and Amino Acids and Application to Plant <sup>13</sup>C-Labeled Experiments. *Metabolites* **13**, 466 (2023).
22. Deller, Y., Berardocco, S., Berges, C., Filangi, O. & Bouchereau, A. Validation of carbon isotopologue distribution measurements by GC-MS and application to <sup>13</sup>C-metabolic flux analysis of the tricarboxylic acid cycle in *Brassica napus* leaves. *Front. Plant Sci.* **13**, 885051 (2023).
23. Kappelmann, J., Klein, B., Geilenkirchen, P. & Noack, S. Comprehensive and accurate tracking of carbon origin of LC-tandem mass spectrometry collisional fragments for <sup>13</sup>C-MFA. *Anal. Bioanal. Chem.* **409**, 2309–2326 (2017).
24. Koley, S., Chu, K. L., Gill, S. S. & Allen, D. K. An efficient LC-MS method for isomer separation and detection of sugars, phosphorylated sugars, and organic acids. *J. Exp. Bot.* **73**, 2938–2952 (2022).
25. Czajka, J. J., Kambhampati, S., Tang, Y. J., Wang, Y. & Allen, D. K. Application of Stable Isotope Tracing to Elucidate Metabolic Dynamics During *Yarrowia lipolytica*  $\alpha$ -Ionone Fermentation. *iScience* **23**, 100854 (2020).
26. Romsdahl, T. B. et al. Analyzing mass spectrometry imaging data of <sup>13</sup>C-labeled phospholipids in *Camelina sativa* and *Thlaspi arvense* (Pennycress) embryos. *Metabolites* **11**, 148 (2021).
27. Farzana, F. et al. Longitudinal spatial mapping of lipid metabolites reveals pre-symptomatic changes in the hippocampi of Huntington's disease transgenic mice. *Neurobiol. Dis.* **176**, 105933 (2023).
28. Borisjuk, L. et al. Seeing plants as never before. *N. Phytologist* **238**, 1775–1794 (2023).
29. Pareek, V., Tian, H., Winograd, N. & Benkovic, S. J. Metabolomics and mass spectrometry imaging reveal channeled de novo purine synthesis in cells. *Science* **368**, 283–290 (2020).
30. Buescher, J. M. et al. A roadmap for interpreting <sup>13</sup>C metabolite labeling patterns from cells. *Curr. Opin. Biotechnol.* **34**, 189–201 (2015).
31. Allen, D. K., Libourel, I. G. L. & Shachar-Hill, Y. Metabolic flux analysis in plants: Coping with complexity. *Plant Cell Environ.* **32**, 1241–1257 (2009).
32. Crown, S. B., Long, C. P. & Antoniewicz, M. R. Integrated <sup>13</sup>C-metabolic flux analysis of 14 parallel labeling experiments in *Escherichia coli*. *Metab. Eng.* **28**, 151–158 (2015).
33. Schwender, J., Shachar-Hill, Y. & Ohlrogge, J. B. Mitochondrial metabolism in developing embryos of *Brassica napus*. *J. Biol. Chem.* **281**, 34040–34047 (2006).
34. Alonso, A. P., Goffman, F. D., Ohlrogge, J. B. & Shachar-Hill, Y. Carbon conversion efficiency and central metabolic fluxes in developing sunflower (*Helianthus annuus* L.) embryos. *Plant J.* **52**, 296–308 (2007).
35. Slater, K. B. et al. One-shot <sup>13</sup>C/<sup>15</sup>N-metabolic flux analysis for simultaneous quantification of carbon and nitrogen flux. *Mol. Syst. Biol.* **19**, e11099 (2023).
36. Pollard, M. & Ohlrogge, J. Testing Models of Fatty Acid Transfer and Lipid Synthesis in Spinach Leaf Using in vivo Oxygen-18 Labeling. *Plant Physiol.* **121**, 1217–1226 (1999).
37. Allen, D. K. Quantifying plant phenotypes with isotopic labeling & metabolic flux analysis. *Curr. Opin. Biotechnol.* **37**, 45–52 (2016).
38. Allen, D. K. Assessing compartmentalized flux in lipid metabolism with isotopes. *Biochim. Biophys. Acta Mol. Cell Biol. Lipids* **1861**, 1226–1242 (2016).
39. Bednarski, T. K., Rahim, M. & Young, J. D. In vivo <sup>2</sup>H/<sup>13</sup>C flux analysis in metabolomics research. *Curr. Opin. Biotechnol.* **71**, 1–8 (2021).
40. Jacobson, T. B. et al. <sup>2</sup>H and <sup>13</sup>C metabolic flux analysis elucidates in vivo thermodynamics of the ED pathway in *Zymomonas mobilis*. *Metab. Eng.* **54**, 301–316 (2019).
41. Wang, Y., Parsons, L. R. & Su, X. AccuCor2: isotope natural abundance correction for dual-isotope tracer experiments. *Lab. Investig.* **101**, 1403–1410 (2021).
42. Poillet-Perez, L. et al. Autophagy maintains tumour growth through circulating arginine. *Nature* **563**, 569–573 (2018).
43. 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* **5**, 100022 (2020).
44. Schwaiger-Haber, M. et al. A Workflow to Perform Targeted Metabolomics at the Untargeted Scale on a Triple Quadrupole Mass Spectrometer. *ACS Meas. Sci. Au* **1**, 35–45 (2021).
45. Aguilar-Mogas, A., Sales-Pardo, M., Navarro, M., Guimerà, R. & Yanes, O. iMet: a network-based computational tool to assist in the annotation of metabolites from tandem mass spectra. *Anal. Chem.* **89**, 3474–3482 (2017).
46. Aron, A. T. et al. Reproducible molecular networking of untargeted mass spectrometry data using GNPS. *Nat. Protoc.* **15**, 1954–1991 (2020).

47. Zhao, K. & Rhee, S. Y. Omics-guided metabolic pathway discovery in plants: Resources, approaches, and opportunities. *Curr. Opin. Plant Biol.* **67**, 102222 (2022).
48. Zhao, K. & Rhee, S. Y. Interpreting omics data with pathway enrichment analysis. *Trends Genet.* **39**, 308–319 (2023).
49. Fiehn, O. Metabolomics - The link between genotypes and phenotypes. *Plant Mol. Biol.* **48**, 155–171 (2002).
50. Mahajan, P., Fiehn, O. & Barupal, D. IDSL.GOA: Gene Ontology Analysis for Metabolomics. *bioRxiv* <https://doi.org/10.1101/2023.03.25.534225> (2023).
51. Giera, M., Yanes, O. & Siuzdak, G. Metabolite discovery: Biochemistry's scientific driver. *Cell Metab.* **34**, 21–34 (2022).
52. Aseeikh, S. et al. Mass spectrometry-based metabolomics: a guide for annotation, quantification and best reporting practices. *Nat. Methods* **18**, 747–756 (2021).
53. Marshall-Colón, A. & Kliebenstein, D. J. Plant Networks as Traits and Hypotheses: Moving Beyond Description. *Trends Plant Sci.* **24**, 840–852 (2019).
54. Allen, D. K. & Young, J. D. Tracing metabolic flux through time and space with isotope labeling experiments. *Curr. Opin. Biotechnol.* **64**, 92–100 (2020).
55. Misra, B. B. & Mohapatra, S. Tools and resources for metabolomics research community: A 2017–2018 update. *Electrophoresis* **40**, 227–246 (2019).
56. Beller, N. C. & Hummon, A. B. Advances in stable isotope labeling: dynamic labeling for spatial and temporal proteomic analysis. *Mol. Omics* <https://doi.org/10.1039/d2mo00077f> (2022).
57. Hoffmann, F., Jaeger, C., Bhattacharya, A., Schmitt, C. A. & Lisec, J. Nontargeted Identification of Tracer Incorporation in High-Resolution Mass Spectrometry. *Anal. Chem.* **90**, 7253–7260 (2018).
58. Chokkathukalam, A. et al. mzMatch-ISO: an R tool for the annotation and relative quantification of isotope-labelled mass spectrometry data. *Bioinformatics* **29**, 281–283 (2012).
59. Capellades, J. et al. GeoRge: A Computational Tool to Detect the Presence of Stable Isotope Labeling in LC/MS-Based Untargeted Metabolomics. *Anal. Chem.* **88**, 621–628 (2016).
60. Melamud, E., Vastag, L. & Rabinowitz, J. D. Metabolomic Analysis and Visualization Engine for LC-MS Data. *Anal. Chem.* **48**, 9818–9826 (2002).
61. Bueschl, C. et al. MetExtract II: A Software Suite for Stable Isotope-Assisted Untargeted Metabolomics. *Anal. Chem.* **89**, 9518–9526 (2017).
62. Huang, X. et al. X13CMS: Global tracking of isotopic labels in untargeted metabolomics. *Anal. Chem.* **86**, 1632–1639 (2014).
63. Wang, R. et al. Global stable-isotope tracing metabolomics reveals system-wide metabolic alternations in aging *Drosophila*. *Nat. Commun.* **13**, 1–14 (2022).
64. Agrawal, S. et al. EL-MAVEN: A fast, robust, and user-friendly mass spectrometry data processing engine for metabolomics. *Methods Mol. Biol.* **1978**, 301–321 (2019).
65. Smith, C. A., Want, E. J., O'Maille, G., Abagyan, R. & Siuzdak, G. XCMS: Processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification. *Anal. Chem.* **78**, 779–787 (2006).
66. Pluskal, T., Castillo, S., Villar-Briones, A. & Orešič, M. MZmine 2: Modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. *BMC Bioinforma.* **11**, 395 (2010).
67. Heinrich, P. et al. 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.* **8**, 17910 (2018).
68. Kambhampati, S., Pajak, A. & Marsolais, F. Evidence that class I glutamine amidotransferase, GAT1\_2.1, acts as a glutaminase in roots of *Arabidopsis thaliana*. *Plant Sci.* **312**, 111033 (2021).
69. Rahim, M. et al. INCA 2.0: A tool for integrated, dynamic modeling of NMR- and MS-based isotopomer measurements and rigorous metabolic flux analysis. *Metab. Eng.* **69**, 275–285 (2022).
70. Xu, Y., Wieloch, T., Kaste, J. A. M., Shachar-Hill, Y. & Sharkey, T. D. Reimport of carbon from cytosolic and vacuolar sugar pools into the Calvin-Benson cycle explains photosynthesis labeling anomalies. *Proc. Natl Acad. Sci. USA* **119**, e2121531119 (2022).
71. Borah, K. et al. Bayesian multi-model-based <sup>13</sup>C/<sup>15</sup>N-metabolic flux analysis quantifies carbon-nitrogen metabolism in mycobacteria. *bioRxiv* <https://doi.org/10.1101/2022.03.08.483448> (2022).
72. Bates, P. D., Durrett, T. P., Ohlrogge, J. B. & Pollard, M. Analysis of acyl fluxes through multiple pathways of triacylglycerol synthesis in developing soybean embryos. *Plant Physiol.* **150**, 55–72 (2009).
73. Allen, D. K., Bates, P. D. & Tjellström, H. Tracking the metabolic pulse of plant lipid production with isotopic labeling and flux analyses: Past, present and future. *Prog. Lipid Res.* **58**, 97–120 (2015).
74. Bates, P. D. The plant lipid metabolic network for assembly of diverse triacylglycerol molecular species. in *Advances in Botanical Research* (eds Rébeillé, F. & Maréchal, E. B. T.-A. in B. R.) vol. 101 225–252 (Academic Press, 2022).
75. Tjellström, H., Yang, Z., Allen, D. K. & Ohlrogge, J. B. Rapid kinetic labeling of *Arabidopsis* cell suspension cultures: Implications for models of lipid export from plastids. *Plant Physiol.* **158**, 601–611 (2012).
76. Roughan, P. G. & Slack, C. R. Cellular Organization of Glycerolipid Metabolism. *Annu Rev. Plant Physiol.* **33**, 97–132 (1982).
77. Pollard, M. & Shachar-Hill, Y. Kinetic complexities of triacylglycerol accumulation in developing embryos from *Camelina sativa* provide evidence for multiple biosynthetic systems. *J. Biol. Chem.* **298**, 101396 (2022).
78. Arrivault, S. et al. Metabolite pools and carbon flow during C4 photosynthesis in maize: <sup>13</sup>CO<sub>2</sub> labeling kinetics and cell type fractionation. *J. Exp. Bot.* **68**, 283–298 (2017).
79. Karki, N., Johnson, B. S. & Bates, P. D. Metabolically Distinct Pools of Phosphatidylcholine Are Involved in Trafficking of Fatty Acids out of and into the Chloroplast for Membrane Production. *Plant Cell* **31**, 2768–2788 (2019).
80. Andersson, M. X., Goksör, M. & Sandelius, A. S. Optical Manipulation Reveals Strong Attracting Forces at Membrane Contact Sites between Endoplasmic Reticulum and Chloroplasts. *J. Biol. Chem.* **282**, 1170–1174 (2007).
81. Roughan, P. G. Stromal concentrations of coenzyme A and its esters are insufficient to account for rates of chloroplast fatty acid synthesis: evidence for substrate channelling within the chloroplast fatty acid synthase. *Biochem. J.* **327**, 267–273 (1997).
82. Grattan Roughan, P. & Ohlrogge, J. B. Evidence That Isolated Chloroplasts Contain an Integrated Lipid-Synthesizing Assembly That Channels Acetate into Long-Chain Fatty Acids. *Plant Physiol.* **110**, 1239–1247 (1996).
83. Andersson, M. X., Goksör, M. & Sandelius, A. S. Membrane Contact Sites. **2**, 185–187 <https://doi.org/10.4161/psb.2.3.3973> (2007).
84. Dorne, A. J., Joyard, J., Block, M. A. & Douce, R. Localization of phosphatidylcholine in outer envelope membrane of spinach chloroplasts. *J. Cell Biol.* **100**, 1690–1697 (1985).
85. Lu, C., Xin, Z., Ren, Z., Miquel, M. & Browse, J. An enzyme regulating triacylglycerol composition is encoded by the ROD1 gene of *Arabidopsis*. *Proc. Natl Acad. Sci. USA* **106**, 18837–18842 (2009).
86. Slack, C. R., Roughan, P. G., Browse, J. A. & Gardiner, S. E. Some properties of cholinephosphotransferase from developing safflower cotyledons. *Biochim. Biophys. Acta Lipids Lipid Metab.* **833**, 438–448 (1985).
87. Dahlqvist, A. et al. Phospholipid:diacylglycerol acyltransferase: An enzyme that catalyzes the acyl-CoA-independent formation of triacylglycerol in yeast and plants. *Proc. Natl Acad. Sci. USA* **97**, 6487–6492 (2000).

88. Kjellberg, J. M., Trimbom, M., Andersson, M. & Sandelius, A. S. Acyl-CoA dependent acylation of phospholipids in the chloroplast envelope. *Biochim. Biophys. Acta Mol. Cell Biol. Lipids* **1485**, 100–110 (2000).
  89. Thompson, J. F., Madison, J. T. & Muenster, A. maria E. In vitro culture of immature cotyledons of soya bean (*Glycine max* L. Merr.). *Ann. Bot.* **41**, 29–39 (1977).
  90. Hsu, F. C. & Obendorf, R. L. Compositional analysis of in vitro matured soybean seeds. *Plant Sci. Lett.* **27**, 129–135 (1982).
  91. Romsdahl, T. B. et al. Analyzing mass spectrometry imaging data of <sup>13</sup>C-labeled phospholipids in camelina sativa and thlaspi arvense (Pennycress) embryos. *Metabolites* **11**, 1–21 <https://doi.org/10.3390/metabo11030148> (2021).
  92. Kambhampati, S. et al. On the inverse correlation of protein and oil: Examining the effects of altered central carbon metabolism on seed composition using soybean fast neutron mutants. *Metabolites* **10**, 1–15 (2020).
  93. Hummel, J. et al. Ultra performance liquid chromatography and high resolution mass spectrometry for the analysis of plant lipids. *Front. Plant Sci.* **2**, 54 (2011).
  94. Kessner, D., Chambers, M., Burke, R., Agus, D. & Mallick, P. ProteoWizard: Open source software for rapid proteomics tools development. *Bioinformatics* **24**, 2534–2536 (2008).
  95. Tautenhahn, R., Bottcher, C. & Neumann, S. Highly sensitive feature detection for high resolution LC/MS. *BMC Bioinforma.* **9**, 1–16 (2008).
  96. Prince, J. T. & Marcotte, E. M. Chromatographic alignment of ESI-LC-MS proteomics data sets by ordered bijective interpolated warping. *Anal. Chem.* **78**, 6140–6152 (2006).
  97. Young, J. D. INCA: A computational platform for isotopically non-stationary metabolic flux analysis. *Bioinformatics* **30**, 1333–1335 (2014).
  98. Antoniewicz, M. R., Kelleher, J. K. & Stephanopoulos, G. Determination of confidence intervals of metabolic fluxes estimated from stable isotope measurements. *Metab. Eng.* **8**, 324–337 (2006).
  99. Kambhampati, S., Allen, D. K. Mass Spectrometry Datasets – Using stable isotopes and mass spectrometry to elucidate the dynamics of metabolic pathways. *Metabolomics Workbench* <https://doi.org/10.21228/M80X3B> (2022).
- facility and instrument use and assistance with instrument maintenance and plant growth. Special thanks to Ravi Kiran Garimella, Rahul Anand, and Jiahong Zhou for assistance in software design. Thanks also to Dr. Justin Renaud for his assistance with qExactive data acquisition and analysis.

### Author contributions

S.Ka., D.K.A. conceived and designed the experiments, wrote the initial manuscript, S.Ka., A.H.H., S.Ko., J.D.G. performed research and programming and contributed to manuscript writing, B.S.E. provided guidance on instrument setup and data acquisition, F.M., J.D.Y. contributed to experimental design, result interpretation and manuscript writing. All authors reviewed the manuscript.

### Competing interests

The authors declare no competing interests.

### Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s42003-024-05844-z>.

**Correspondence** and requests for materials should be addressed to Shrikaar Kambhampati or Doug K. Allen.

**Peer review information** *Communications Biology* thanks Berkley Walker, Li Chen, and Zhaoyue Zhang for their contribution to the peer review of this work. Primary Handling Editors: Zheng-Jiang Zhu and Joao Valente. A peer review file is available.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

This is a U.S. Government work and not under copyright protection in the US; foreign copyright protection may apply 2024

### Acknowledgements

This work was supported by the National Institute of Health (NIH-U01 CA235508), National Science Foundation (NSF-IOS-1829365), Natural Sciences and Engineering Research Council of Canada (NSERC, RGPN-2018-05478), and the United States Department of Agriculture, National Institute of Food and Agriculture (2023-67017-39419) and Agricultural Research Service (USDA-ARS). Support for the acquisition of instruments used in this work were obtained using NSF (DBI-1827534). The authors would like to acknowledge the Subterranean Influences on Nitrogen and Carbon (SINC) Center, and the Proteomics and Mass Spectrometry Facility, and the Plant Growth Facility at the Donald Danforth Plant Science Center for