Chapter 12

Isotopically Nonstationary MFA (INST-MFA) of Autotrophic Metabolism

Lara J. Jazmin, John P. O'Grady, Fangfang Ma, Doug K. Allen, John A. Morgan, and Jamey D. Young

Abstract

Metabolic flux analysis (MFA) is a powerful approach for quantifying plant central carbon metabolism based upon a combination of extracellular flux measurements and intracellular isotope labeling measurements. In this chapter, we present the method of isotopically nonstationary ^{13}C MFA (INST-MFA), which is applicable to autotrophic systems that are at metabolic steady state but are sampled during the transient period prior to achieving isotopic steady state following the introduction of $^{13}CO_2$. We describe protocols for performing the necessary isotope labeling experiments, sample collection and quenching, nonaqueous fractionation and extraction of intracellular metabolites, and mass spectrometry (MS) analysis of metabolite labeling. We also outline the steps required to perform computational flux estimation using INST-MFA. By combining several recently developed experimental and computational techniques, INST-MFA provides an important new platform for mapping carbon fluxes that is especially applicable to autotrophic organisms, which are not amenable to steady-state ^{13}C MFA experiments.

Key words Metabolic flux analysis, Isotopomer analysis, Nonaqueous fractionation, Mass spectrometry, Elementary metabolite unit, Isotopomer modeling, Autotrophic metabolism

1 Introduction

Metabolic engineering of plants has great potential for the production of novel compounds and chemical feedstocks at low costs. However, in order to take full advantage of this potential, a better understanding of plant central carbon metabolism is needed. The ability to quantitatively map intracellular carbon fluxes using isotope tracers and metabolic flux analysis (MFA) is critical for identifying pathway bottlenecks and elucidating network regulation in plants, especially transgenic varieties that have been engineered to alter their native metabolic capacities [1, 2]. Typically, MFA relies on the assumption of both metabolic and isotopic steady state. Achieving this situation experimentally involves (1) equilibrating the system in a stable metabolic state, (2) introducing an

Martine Dieuaide-Noubhani and Ana Paula Alonso (eds.), *Plant Metabolic Flux Analysis: Methods and Protocols*, Methods in Molecular Biology, vol. 1090, DOI 10.1007/978-1-62703-688-7_12,

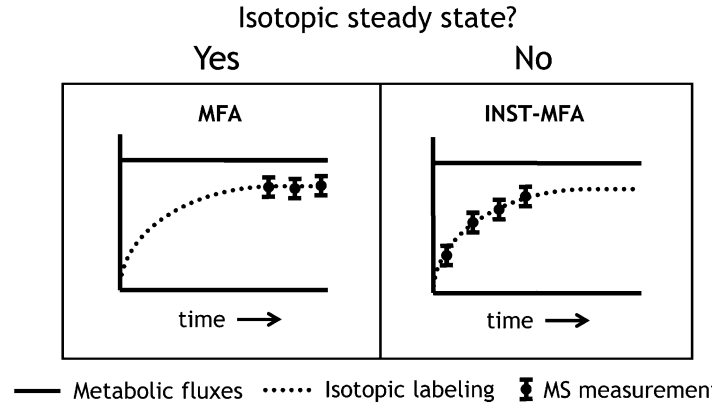

Fig. 1 Comparison between steady-state and nonstationary MFA methodologies. The relative speed of metabolic and isotopic dynamics will influence the type of MFA study performed. The *left panel* shows the conventional MFA approach under both metabolic and isotopic steady state. The *right panel* shows INST-MFA at metabolic steady state but not isotopic steady state

isotopically labeled substrate without perturbing the metabolic steady state, (3) allowing the system to establish a new isotopic steady state that is dictated by the underlying metabolic fluxes, and (4) measuring isotopic labeling in the fully equilibrated system as shown in Fig. 1. Depending on the relative speed of metabolic versus isotopic dynamics, however, other experimental scenarios can be envisioned. If labeling occurs slowly but metabolism is maintained in a fixed state, isotopically nonstationary MFA (INST-MFA; Fig. 1) can be applied to determine fluxes from transient isotope labeling measurements [3]. This requires isotopomer measurements to be collected at multiple time points during the pre-steady-state labeling period followed by least-squares fitting of dynamic isotopomer balance equations to identify best-fit estimates of flux parameters.

Although ^{13}C is the preferred isotope tracer for quantifying central carbon metabolism in heterotrophic systems, autotrophic organisms assimilate carbon solely from CO_2 and therefore produce a uniform steady-state ^{13}C-labeling pattern that is insensitive to fluxes (Fig. 2). This makes conventional steady-state ^{13}C-MFA ineffective for quantifying autotrophic metabolism [4]. However, transient measurements of isotope incorporation following a step change from natural CO_2 to $^{13}CO_2$ can be used to determine fluxes by application of INST-MFA. Furthermore, INST-MFA has the ability to quantify metabolite pool sizes based solely on their labeling dynamics [5, 6], thus providing a potential framework for integrating metabolomic datasets with MFA. Despite its advantages, however, the increased complexity of INST-MFA introduces

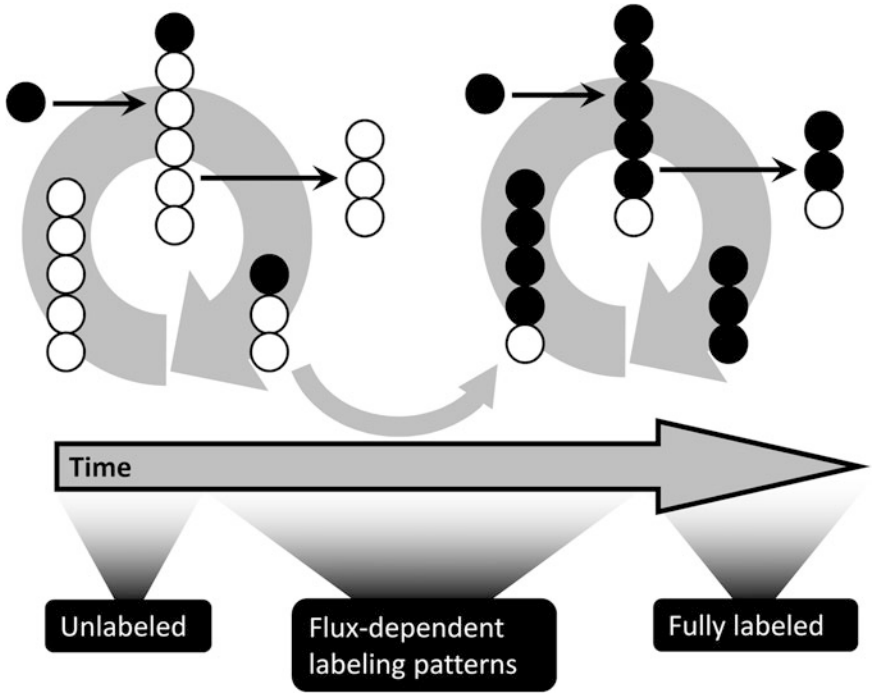

Fig. 2 Example of carbon labeling in an autotrophic system. Following the introduction of $^{13}CO_2$ to the Calvin cycle, intracellular metabolites become gradually labeled over time. Once steady-state labeling is achieved, all metabolites are uniformly ^{13}C enriched irrespective of fluxes and intracellular pool sizes. Labeling patterns observed during the isotopically transient period, however, can be computationally analyzed to determine fluxes

additional difficulties at both the computational and experimental levels that must be addressed. First, the solution of large-scale ordinary differential equation (ODE) models poses a substantial challenge to efficiently simulate transient isotope labeling experiments. The application of EMU decomposition to INST-MFA has greatly reduced this computational burden and has enabled determination of fluxes and accurate confidence intervals in biologically relevant networks [7, 8]. Second, the requirement for isotopically nonstationary measurements adds further complexity to experimental design, including selection of sampling time points and metabolite concentration measurements. Finally, rapid sampling and metabolite quenching must be applied in order to obtain meaningful isotopomer data from rapidly labeled intracellular metabolites. The field of metabolomics has witnessed considerable progress in this area, and some of these measurement techniques have already been successfully adapted for isotopomer studies in plants [9–12].

In this contribution, we present up-to-date protocols for performing INST-MFA on photosynthetic tissues under conditions

of metabolic steady state, which is enabled by recent advancements in both experimental techniques for obtaining transient isotopomer measurements and computational tools for nonstationary flux analysis. Specifically, we describe experimental protocols for performing transient isotope labeling experiments, sample collection and quenching, extraction of intracellular metabolites, and mass spectrometry (MS) analysis of metabolite labeling, We also describe the important technique of nonaqueous fractionation (NAF), which can be used to obtain compartment-specific information on isotopic enrichment [13]. We then discuss the computational steps required to simultaneously integrate mass isotopomer distribution (MID) measurements from multiple pathways and compartments to reconstruct comprehensive flux maps of autotrophic metabolism. We anticipate that these methods will become part of a growing MFA toolbox, which has already been rapidly adopted by the plant biology and metabolic engineering communities due to its ability to quantify network-wide metabolic phenotypes in seeds and other non-photosynthetic tissues.

2 Materials

2.1 ^{13}C Labeling Experiment and Quenching

1. Plants at metabolic steady state (*see* **Note 1**).
2. Glove bag (Glas-Col, 17″ × 17″ × 11″) or other labeling chamber.
3. ^{13}C labeled CO_2 gas; $^{13}CO_2/N_2/O_2$ gas mix, ratio 0.033/78/21.967 (Sigma).
4. Liquid nitrogen.

2.2 CO_2 Consumption and Production Measurement

1. Li-COR 6400XT portable photosynthesis system (Lincoln, NE, USA).

2.3 Starch Production Measurement

1. 75 % (v/v) methanol.
2. α-amylase.
3. Amyloglucosidase.
4. Megazyme Kit (K-TSTA).
5. UV–Vis spectrophotometer.

2.4 Nonaqueous Fractionation

1. Tetrachloroethylene.
2. Heptane.
3. Centrifuge (capable of at least 13,000 × *g*).
4. Inert centrifuge tubes, preferably transparent.

2.5 Enzymatic Assays

2.5.1 Phosphoenolpyruvate Carboxylase Assay Buffer

1. 90 μL of 110 mM Tris sulfate adjusted to pH 8.5 at 25 °C with NaOH (all volumes are per sample).
2. 5 μL of 300 mM $MgSO_4 \cdot 7H_2O$.
3. 15 μL of 100 mM $NaHCO_3$.
4. 5 μL of 6 mM β-NADH.
5. 15 μL of dioxane.
6. 5 μL of 300 mM dithioerythritol (DTE).
7. 0.5 μL of 600 units/mL malic dehydrogenase (MDH).

2.5.2 Prephenate Aminotransferase Storage Buffer

1. 10 wt% glycerol.
2. 5 wt% of 1 M sodium phosphate, adjusted to pH 8.0 at 25 °C with NaOH.
3. 1 wt% of 10 mM pyridoxal 5-phosphate (PLP).

2.5.3 Prephenate Aminotransferase Assay Solution

1. 1.25 μL of 1 M sodium phosphate, adjusted to pH 8.0 at 25 °C with NaOH (all volumes are per sample).
2. 0.5 μL of 10 mM PLP.
3. 5 μL of 100 mM aspartic acid.
4. 2.5 μL of 10 mM prephenic acid.

2.5.4 Prephenate Aminotransferase Assay Derivatizing Solution

1. 5 mg Phthaldialdehyde (all volumes are per sample).
2. 900 μL of 0.5 M boric acid, adjusted to pH 10.2 at 20 °C with NaOH.
3. 125 μL Methanol.
4. 5 μL β-mercaptoethanol.

2.5.5 α-Mannosidase Assay

1. 50 mM sodium citrate buffer adjusted to a pH 4.5 at 25 °C.
2. 5 mM *p*-nitrophenol-α-D-mannoside.
3. 0.2 M boric buffer adjusted to pH 9.8 at 25 °C.
4. PD-10 protein desalting columns.
5. Microplate reader.
6. HPLC system equipped with a UV–Vis detector.

2.6 Metabolite Extraction

1. Cellulose acetate spin-X columns.
2. Methanol.
3. Chloroform.
4. DI water.
5. Bench-top centrifuge (capable of at least 8,000 × *g*).

2.7 Mass Spectrometry Analysis and Data Processing

1. GC-MS and/or LC-MS/MS.
2. Derivatization agents, vials, heating blocks, and nitrogen evaporator for GC-MS sample preparation (*see* **Note 2**).
3. Vials, columns, gases, buffers, solvents, and other consumables for GC-MS or LC-MS/MS.
4. Computer equipped with either (1) freeware MS analysis software (*see* **Note 3**) or (2) commercial software for searching and integrating mass spectra.
5. Mass spectral library for compound verification, such as the NIST/EPA/NIH Mass Spectral Database [14], Golm Metabolome Database [15], FiehnLib [16], Metlin [17], HMDB [18], or MassBank [19].

2.8 Isotopically Nonstationary Metabolic Flux Analysis

1. Computer equipped with research code or publically available software capable of performing INST-MFA, such as Isotopomer Network Compartmental Analysis (INCA; http://mfa.vueinnovations.com/), which runs through the computing environment of MATLAB.

3 Methods

3.1 ^{13}C Labeling Experiment and Quenching

The ^{13}C labeling experiment should be initialized once the plants are producing maximum biomass without production of reproductive organs (i.e., flowers and seeds) and approaching a metabolic (pseudo) steady state (*see* **Note 1**).

1. To maintain steady state metabolism, phototrophs should be isotopically labeled in the chamber near midday or 3–4 h after exposure to light.
2. Remove samples from the chamber at multiple time points (~5–15) prior to reaching isotopic steady state and quench immediately.
3. Initiate quenching by immediately pouring liquid nitrogen onto plant tissue in a way that minimizes shading.
4. Grind using a mortar and pestle.

3.2 CO_2 Consumption and Production Measurement

1. Measure leaf CO_2 assimilation in an incubator via a Li-COR 6400XT portable photosynthesis system and 2 × 3 cm standard clear top or red/blue light source leaf chamber with the following settings:
   (a) Photon flux density (PFD): 200 μmol/(m^2 s).
   (b) Air temperature: 24 ± 1 °C.
   (c) Reference CO_2 concentration: ambient level.

3.3 Starch Production Measurement

1. Collect approximately 100 mg of fresh tissue.
2. Add tissue to 1.4 mL of 75 % (v/v) methanol at 70 °C for 30 min.
3. Add 750 μL chloroform (−20C) and 1500 μL DI water (4C) and vortex
4. Spin down the sample to fraction the nonpolar metabolites into the chloroform.

   Remove the methanol/water phase for further analysis of water-soluble metabolites.
5. Re-extract the remaining residue with 1.5 mL 80 % (v/v) ethanol at 80 °C three times to remove all the remaining water-soluble metabolites.
6. Dry down the remaining residue.
7. Digest the pellet with α-amylase and amyloglucosidase to hydrolyze starch into maltodextrins and then into glucose.
8. Determine the starch concentration spectrophotometrically as glucose equivalents using the Megazyme Kit.

3.4 Nonaqueous Fractionation

NAF is an optional, yet invaluable, technique for separating metabolites from different subcellular compartments in an ordered manner so that they can be characterized and their relative quantities at each locus can be identified. This procedure, first outlined for plants by Gerhardt and Heldt [20], begins with the rapid freezing and lyophilizing of the tissue. This arrests metabolism, and the removal of water prevents polar metabolites from being able to diffuse—instead they remain with the membrane of the compartments in which they were localized while all enzymes are retained compartmentally but inactive. At this point the cellular material can be stratified by compartment using a nonaqueous density gradient. This separation is quantified by extracting the enzymes of each compartment and analyzing their relative activity in each fraction of the density gradient. Using information on how the compartments have been separated, coupled with a thorough analysis of the metabolite profile of each fraction of the density gradient, the metabolic profiles of each compartment can be deconvoluted and the resulting data employed for further analysis (*see* **Note 5**).

1. Grind the leaf tissue. It is crucial that frequency and time be optimized to get small particles.
2. Filter the leaf powder by size before suspension in the loading fraction of the gradient.
3. Prepare the steps in the gradient by mixing together various ratios of tetrachloroethylene and heptanes. The NAF gradient is prepared to achieve the maximum separation of biomass (*see* **Note 6**).

4. Carefully layer the steps in the gradient on top of each other in the centrifuge tube, allowing only minimal mixing of the layers. A loading layer and density well are prepared for the top and bottom layers, respectively.
5. Place the loading layer with the sample suspended in it on top of the gradient, once it is created.
6. Centrifuge the tube for 2 h at 13,000 × *g* to ensure full separation of the cellular fragments throughout the gradient.
7. Remove the tube after centrifugation and draw out the sample in fractions. The sample is drawn from the bottom of the tube using a syringe of glass or other nonreactive material. Fraction volumes are selected to help normalize the amount of biomass in each fraction. Generally, six or seven fractions are drawn.
8. While still suspended, further separate the fractions in order to accommodate the number of assays that will be performed (i.e., if three compartmental markers will be analyzed and each requires suspension of cellular material in a separate buffer, then the fraction will be separated into four parts—three for the different assays and one for metabolite analysis).
9. Dry down the fractions under a stream of N_2 to remove the nonaqueous solvent before subsequent analysis.

3.5 Enzymatic Assays

Enzymatic assays are used to determine the distribution of each subcellular compartment throughout the fractions. Because the cellular fragments frequently contain parts of multiple compartments, this information is necessary to deconvolute the labeling patterns observed in the metabolite profiles of the fractions. The enzymes assayed are selected because they are easily and reliably quantifiable and they are known to be present in only one compartment. The assays outlined here are for the determination of cytosolic (phospho*enol*pyruvate carboxylase), plastidial (prephenate aminotransferase), and vacuolar (α-mannosidase) distribution.

Mitochondrial activity has been determined by others using the method of Bergmeyer [21], although, due to the large standard deviation of the assay, the compartment has not been unambiguously delineated for any plant species [22]. Activity measurements for each protein assay are normalized by their protein content. Protein content is measured using a Pierce BCA protein assay kit from Thermo Fisher Scientific (Rockford, IL).

3.5.1 Phosphoenolpyruvate Carboxylase Assay

1. Resuspend samples in 0.5 mL of Tris sulfate buffer and allow to incubate for 30–60 min at 4 °C to effectively extract the protein.
2. Spin down the samples at 17,000 × *g* for 30 s.
3. Aliquot 10 μL of supernatant from each sample into the wells of a 96-well plate.

4. Following the procedure prescribed by Sigma-Aldrich [23], add to the extract phospho*enol*pyruvate carboxylase (PEPc) assay buffer, as described in Subheading 2.5.
5. Allow the wells to equilibrate at 25 °C for 5 min.
6. Add 5 μL of 30 mM PEP solution to each well, except for the last one, which will serve as a blank. Mix thoroughly for 1 min.
7. Measure the OD_{340} of each well periodically over the next 10 min to determine relative enzyme activity.

3.5.2 Prephenate Aminotransferase Assay

A modified form of the method presented by Maeda et al. [24] is used to measure prephenate aminotransferase (PAT) activity.

1. Resuspend the samples in 0.4 mL of PAT storage buffer. Allow the samples to incubate for 30–60 min at 4 °C to effectively extract the protein.
2. Spin down at 18,000 × *g* for 30 s.
3. Load the supernatant onto PD-10 columns, which have been prewashed with the PAT storage buffer solution.
4. Elute fractions from the PD-10 columns 0.25 mL at a time.
5. Measure the OD_{280} of each PD-10 column fraction to determine the relative protein content. The set of subfractions with the highest absorbance will be used in the PAT assay.
6. Perform the assay by adding PAT assay buffer to 15 μL of the desalted enzyme solution.
7. Incubate the samples at 37 °C for 30–60 min.
8. Quench the samples with 50 μL of methanol and store at −20 °C overnight to precipitate the protein.
9. Centrifuge the samples at 18,000 × *g* for 5 min to remove all precipitated protein. The supernatant is analyzed by high-performance liquid chromatography (HPLC).
10. Derivatize the samples immediately prior to injection using a PAT assay derivatizing solution.
11. 10 μL samples are derivatized with 10 μL of derivatizing agent in the injection needle for 2.5 min prior to sample separation. This derivatizing solution should be changed every 12 h.
12. Run sample on an Agilent Eclipse XDB-C8 column (4.6 × 150 mm) maintained at 35 °C. The mobile phase is a ratio of 0.1 % ammonium acetate:methanol that varies according to Table 1 with a constant flow rate of 0.4 mL/min. Derivative concentrations were measured using a UV/Vis detector observing absorbance at 336 nm with a reference of 390 nm. Figure 3 shows an example of the chromatogram.

3.5.3 α-Mannosidase Assay

1. Extract protein samples in 50 mM sodium citrate buffer.
2. Allow samples to incubate for 30–60 min at 4 °C to effectively extract the protein.

Table 1
HPLC mobile-phase gradient for PAT assay

Time (min)	% MeOH
0	15
15	65
25	65
27	15
32	15

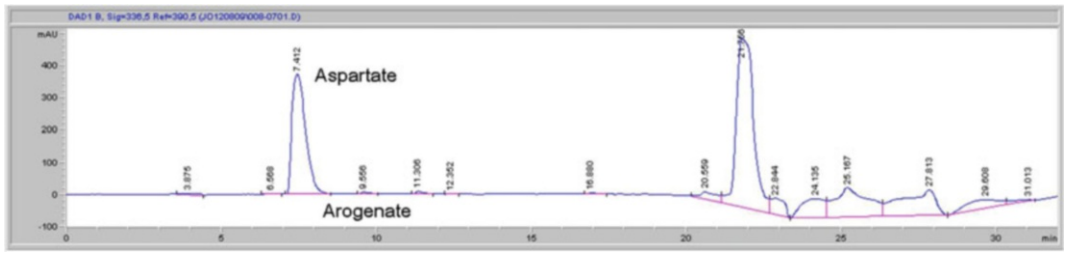

Fig. 3 Sample HPLC chromatogram of the PAT assay using phthaldialdehyde as a derivatizing agent

3. Following the method of Li and Primate [25], add the extracted enzyme to 5 mM *p*-nitrophenol-α-D-mannoside solution at 25 °C in a 10:1 ratio.
4. Terminate the reaction after 10 min with a volume of 0.2 M borate buffer at pH 9.8 equal to twice the reaction mixture.
5. Measure the OD_{400} of each sample to determine enzyme activity.

3.6 Metabolite Extraction

1. Weigh 200 mg tissue powder (kept in liquid N_2) to a plastic 2 mL tube.
2. Add 3 mL 7:3 methanol–chloroform (−20 °C); after vigorous shaking, the mixture is incubated at −20 °C for 2 h with occasional shaking.
3. Water-soluble metabolites are extracted from the chloroform phase by adding 2.4 mL H_2O; after vigorous shaking, centrifuge for 5 min at 420 × *g*.
4. Remove the upper, aqueous methanol–H_2O phase to a glass reaction tube and repeat **step 3**; mix the second upper phase with the first one; the final volume of the extract should be about 6.5 mL.
5. Dry the extract with N_2 at room temperature.

6. Redissolve the dry extract with 200 μL H_2O; remove the viscous, high-molecular-mass components from the samples by applying extracts to 0.45 μm cellulose acetate centrifuge tube filters.

3.7 Mass Spectrometry Analysis

Divide the extracted sample from each fraction into three different vials. One vial will be used for LC-MS/MS analysis and the remaining two vials will be dried down the rest of the way for amino acid and organic acid analyses using GC-MS.

3.7.1 GC-MS Analysis of Amino Acids

GC-MS analysis of amino acids is most readily achieved following conversion to *tert*-butyldimethylsilyl (TBDMS) derivatives [26]. TBDMS derivatives are approximately 1,000 times more stable than trimethylsilyl (TMS) derivatives and produce characteristic $[M\text{-}57]^+$, $[M\text{-}85]^+$, and $[M\text{-}157]^+$ fragment ions that facilitate identification. Ion masses and relative retention times are detailed by Wittman et al. [27].

1. Dissolve dried sample in 100 μL anhydrous pyridine in a fume hood.
2. Add 70 μL MTBSTFA + 1 % TBDMCS.
3. Sonicate for 5 min.
4. Incubate for 2 h at 45 °C on dry block heater.
5. Transfer supernatant to injection vial.
6. Run sample on single quadrupole GC-MS with the following settings:
   (a) Injection volume: 1 μL.
   (b) Split injection mode, 10:1 ratio (*see* **Note 7**).
   (c) Column: DB-5MS (or similar), 30 m × 0.25 mm ID, 0.25 μm film.
   (d) Column flow: He at 1 mL/min (*see* **Note 8**).
   (e) Inlet temp: 270 °C.
   (f) Interface temp: 300 °C.
   (g) Temp profile: 135 °C for 3 min, ramp at 3 °C/min to 280 °C, hold at 280 °C for 2 min (53.3-min run time).
   (h) MS settings: Electron impact (EI) ionization, full-scan detection (100–500 m/z), 5-min solvent delay.

3.7.2 GC-MS Analysis of Organic Acids

Organic acids can be converted to their TBDMS derivatives and analyzed in much the same way as amino acids. However, it is recommended to first convert all ketone and aldehyde functional groups to their associated methyloxime derivatives. This prevents keto-enol tautomerization, which would otherwise result in multiple TBDMS derivatives. Ion masses are detailed by Roessner et al. [28].

1. Dissolve dried sample in 100 µL pyridine with 20 mg/mL methoxyamine hydrochloride MOX reagent in a fume hood.
2. Sonicate for 5 min.
3. Incubate for 90 min at 30 °C on dry block heater.
4. Add 100 µL MTBSTFA + 1 % TBDMCS.
5. Incubate for 60 min at 37 °C on dry block heater.
6. Remove from heating block and incubate overnight at room temperature (*see* **Note 9**).
7. Centrifuge for 5 min at 18,000 × *g* to remove solid debris.
8. Transfer liquid to injection vial.
9. Run sample on single quadrupole GC-MS with the following settings:
   (a) Injection volume: 1 µL.
   (b) Purged splitless mode, set to activate at 1 min (*see* **Note 7**).
   (c) Column: DB-5MS (or similar), 30 m × 0.25 mm ID, 0.25 µm film.
   (d) Column flow: He at 1 mL/min (*see* **Note 8**).
   (e) Inlet temp: 270 °C.
   (f) Interface temp: 300 °C.
   (g) Temp profile: 70 °C for 3 min, ramp at 4 °C/min to 280 °C, hold at 280 °C for 2 min (75-min run time).
   (h) MS settings: Scan mode (100–550 *m/z*), 5-min solvent delay.

3.7.3 LC-MS/MS Analysis of Sugar Phosphates

Although analysis of labeling in macromolecule components such as protein, starch, and lipids has dominated the MFA literature due to their high abundance in tissues, macromolecules label too slowly for INST-MFA and recent advances in MS technology have facilitated direct analysis of free intracellular metabolites. This approach enables quantification of MIDs in sugar phosphate intermediates that participate directly in glycolysis and Calvin cycle reactions, providing comprehensive and dynamic information on flux through these important pathways. However, for intermediates found in multiple subcellular compartments (e.g., glucose-6-phosphate in both the cytosol and chloroplast), this will only provide the *average* labeling among the various intracellular pools, unless NAF is applied to generate enriched subfractions first.

The vast majority of sugar phosphate analysis has been performed using LC-MS/MS to avoid thermal degradation of these nonvolatile analytes [29]. The LC-MS/MS conditions provided are based on the method developed by Luo et al. [30], which has been subsequently modified and applied to analyze labeling in cyanobacteria extracts by Shastri et al. [31].

Table 2
LC-MS/MS ion transitions and method parameters

Analyte	Retention time (min)	MS1	MS2	DP	EP	CE	CXP	CUR	IS	TEM	GS1	GS2	CAD
Glycolate	6.2	75	43	−52	−10	−15	−7	40	−4,500	550	85	90	High
G6P/F6P	12.4/13.3	259	97	−50	−2	−23	−10	40	−4,500	550	85	90	High
GAP/DHAP	12.2/16.9	169	97	−50	−3	−12	−8.5						
S7P	12.9	289	97	−26	−11	−14	−4.5						
R5P/X5P/RU5P	12.7/14.2/14.5	229	97	−79	−6	−17	−15						
PYR	15.7	87	43	−33	−10	−13	−7						
SUC	22.8	117	73	−50	−3	−15	−2	30	−4,500	550	90	90	9
MAL	24.4	133	73	−50	−9	−23	−11						
αKG	25.5	145	101	−43	−5	−11	−9.6						
FUM	26.4	115	71	−37	−4	−12	−11						
3PGA	26.5	185	79	−42	−5	−46	−11						
FBP	27	339	97	−66	−8	−26	−11						
CIT/ICT	27.4	191	75	−60	−10	−25	−10						
RUBP	27.6	309	97	−68	−6	−28	−12						
PEP	27.8	167	79	−55	−5	−21	−7						

Labeling of selected metabolites measured by monitoring transitions from selected precursor ions (MS1) to daughter ions (MS2) in multiple reaction monitoring (MRM) mode. Compound specific MS parameters are abbreviated as follows: declustering potential (DP), entrance potential (EP), collision energy (CE), collision exit potential (CXP), curtain gas (CUR), ion spray velocity (IS), temperature (TEM), ion source gas 1 (GS1), ion source gas 2 (GS2), and collision (CAD). List of analyte abbreviations can be found in Subheading 4 (*see* **Note 15**)

1. Transfer sample to LC injection vial.
2. Run sample on linear ion-trap triple quadrupole LC-MS/MS with the following settings:
   (a) Injection volume: 10 μL.
   (b) Column: Phenomenex Synergi Hydro-RP (or similar), 150 mm × 2.1 mm ID, 4 μm particle size.
   (c) Column flow: 0.3 mL/min.
   (d) Column temp: 25 °C.
   (e) Eluent A: Solution of 10 mM tributylamine + 15 mM acetic acid (*see* **Note 10**).
   (f) Eluent B: HPLC-grade methanol.
   (g) Gradient profile: 0 % B (0 min), 8 % B (8 min), 22 % B (18 min), 40 % B (28 min), 60 % B (32 min), 90 % B (34 min), 90 % B (37 min), 0 % B (39 min), 0 % B (49 min).
   (h) MS settings: Negative-mode electrospray ionization (ESI), MRM mode (*see* Table 2 and **Note 11**).

3.8 MS Data Processing

Analysis of MS data requires (1) identification of chromatographic peaks and fragment ions associated with target analytes of interest, (2) integration of ion chromatograms over time to quantify relative abundance of specific isotope peaks, and (3) assessment of measurement standard errors.

1. Identify the chromatographic peaks associated with the analytes of interest based on both the retention time (RT) and the MS fingerprint of the peak.
2. Identify ions to be used for mass isotopomer analysis and determine their molecular composition. The best GC-EI-MS fragment ions are highly abundant ions with masses greater than 150 Da, since these are less likely to be contaminated by interfering fragment ions of similar mass.
3. Integrate the mass isotopomer peaks using either custom or commercial software.
4. Correct MIDs for natural isotope abundance (optional). The method of Fernandez et al. [32] can be applied to perform the correction.
5. Calculate the mean and standard error of MIDs for each metabolite at each time point.

3.9 Isotopically Nonstationary Metabolic Flux Analysis (INST-MFA)

A flow chart of a typical INST-MFA process is shown in Fig. 4. INST-MFA is concerned with solving an "inverse problem" where fluxes and pool sizes are estimated from measured labeling patterns and extracellular rates through the means of an iterative least-squares fitting procedure. At each iteration, a "forward problem" is solved where an isotopomer model is used to simulate labeling measurements for a given metabolic network and a given set of parameter estimates. The discrepancy between the simulated and measured labeling patterns is then assessed, and the parameter estimates are updated to achieve an improving fit. Once convergence to the best-fit solution is obtained, the procedure terminates and the optimal flux and pool size estimates are returned.

1. *Build an isotopomer model for INST-MFA.* In order to perform INST-MFA, it is necessary to reconstruct a metabolic network from biochemical literature and the annotated genome of the organism of interest. This network must prescribe both (1) the stoichiometry of all enzymatic reactions under consideration and (2) atom transitions for each reaction (*see* **Note 12**). Reactions must also be classified as either reversible or irreversible.
2. *Construct the stoichiometric matrix S.* Fluxes are required to satisfy the stoichiometric constraint

$$S \cdot v = 0 \tag{1}$$

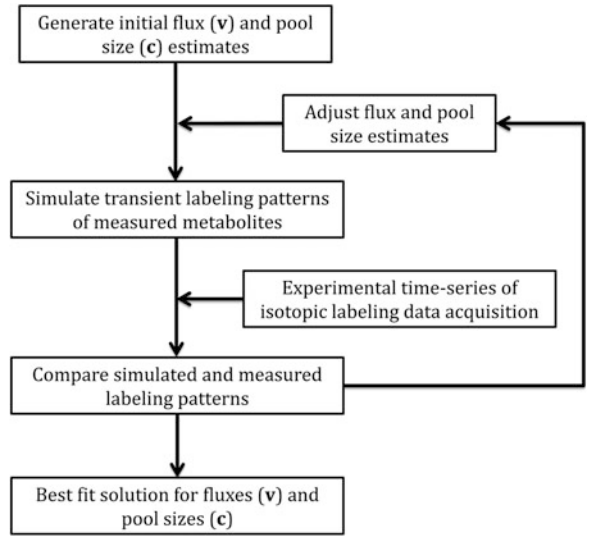

Fig. 4 Flow chart showing the overall schematic of ^{13}C INST-MFA. Following the labeling experiment and MS analysis of the measured metabolites, computational analysis of the dynamic changes in isotope labeling patterns can be used to estimate metabolic pathway fluxes and pool sizes. This involves solving an inverse problem whereby the vectors of flux (**v**) and pool size (**c**) parameters are iteratively adjusted until the mismatch between simulated and experimentally measured data sets is minimized

where S is the stoichiometric matrix and v is the flux vector. In Eq. 1, the stoichiometric matrix S is a $k \times j$ matrix, with k metabolites and j fluxes. Reversible reactions should be modeled as separate forward and backward fluxes so that all fluxes are nonnegative. Appendix shows a simple network model example, which illustrates the process of setting up the stoichiometric matrix as well as subsequent steps discussed below.

3. *Identify the free fluxes of the network and the null space matrix.* From a computational standpoint, it is more convenient to work with "free" fluxes rather than the "true" network fluxes [33, 34]. Free fluxes can be obtained from the general solution to Eq. 1:

$$v = N \cdot u \tag{2}$$

where N is the null space matrix of S and u is the vector of free fluxes. There are many methods to calculate a valid null space matrix, and generally there is not a unique null space matrix for any given stoichiometric matrix [35]. The size of the null space matrix and the number of independent flux variables are determined by the rank of the stoichiometric matrix. With $r = \text{rank}(S) \leq k$, the null space matrix is a $j \times j - r$ matrix and the number of free fluxes is $j - r$.

4. *Identify the minimal set of EMUs required to simulate the MIDs of the measured metabolites.* In INST-MFA, the isotopomer balances are described by a system of ordinary differential equations, which is significantly more difficult to solve than the algebraic systems that describe steady-state labeling. Due to this additional difficulty, algorithms for solving the forward problem of INST-MFA need to be carefully designed so that computational expense does not become prohibitive. The most efficient approach involves first decomposing the isotopomer network into elementary metabolite units (EMUs) [7, 8]. By only solving for the isotopomer distributions of EMUs that contribute to the available measurements, this approach minimizes the number of ODEs that need to be integrated and thereby enables the forward problem to be solved thousands of times faster than previous methods. This, in turn, increases the efficiency of solving the inverse problem of INST-MFA because each iteration of the parameter estimation procedure can be completed in minimal time.

   An EMU is defined as a distinct subset of a metabolite's atoms and can exist in a variety of mass states depending on its isotopic composition. In its lowest mass state, an EMU is referred to as M0, while an EMU that contains one additional atomic mass unit (e.g., as a result of a ^{13}C atom in place of ^{12}C atom) is referred to as M1, with higher mass states described accordingly. An MID is a vector that contains the fractional abundance of each mass state of an EMU. To solve the forward problem of simulating metabolite labeling in INST-MFA, the isotopomer network is first systematically searched to enumerate all EMUs that contribute to measurable MS fragment ions [8]. The main advantage of the EMU decomposition is that metabolites are never broken into smaller pieces than is strictly required to describe the labeling state of the selected metabolites (*see* Appendix).

5. *Set up EMU balances and simulate the labeling distribution.* The EMU reactions identified from network decomposition form the new basis for generating system equations. In INST-MFA, these EMUs are grouped into mutually dependent blocks using a Dulmage–Mendelsohn decomposition [36, 37] (*see* **Note 13**). Therefore, by definition, all EMUs within a particular block have the same number of atoms and must be solved simultaneously and not sequentially. The decoupled blocks can be arranged into a cascaded system of ODEs with the following form:

$$\mathbf{C}_n \cdot \frac{\mathrm{d}\mathbf{X}_n}{\mathrm{d}t} = \mathbf{A}_n \cdot \mathbf{X}_n + \mathbf{B}_n \cdot \mathbf{Y}_n \tag{3}$$

   Level n of the cascade represents the network of EMUs within the nth block. The rows of the state matrix $\mathbf{X}_n$ correspond to MIDs of EMUs within the nth block. The input matrix $\mathbf{Y}_n$ is

analogous but with rows that are MIDs of EMUs that are previously calculated inputs to the nth block (or MIDs of source EMUs that are unbalanced). The concentration matrix $\mathbf{C}_n$ is a diagonal matrix whose elements are pool sizes corresponding to EMUs represented in $\mathbf{X}_n$. The system matrices $\mathbf{A}_n$ and $\mathbf{B}_n$ describe the network as follows:

$$\mathbf{A}_n(i,j) = \begin{cases} -\text{sum of fluxes consuming } i\text{th EMU in } \mathbf{X}_n & i = j \\ \text{flux to } i\text{th EMU in } \mathbf{X}_\text{n} \text{ from } j\text{th EMU in } \mathbf{X}_n & i \neq j \end{cases} \quad (4)$$

$$\mathbf{B}_n(i,j) = \{\text{flux to } i\text{th EMU in } \mathbf{X}_n \text{ from } j\text{th EMU in } \mathbf{Y}_n \quad (5)$$

The system matrices $\mathbf{A}_n$ and $\mathbf{B}_n$ can be evaluated directly once the "true" flux vector (v) has been determined from the free fluxes (u) and the null space matrix (N) (*see* Appendix).

6. *Optimize parameters.* Fluxes and pool sizes are estimated by minimizing the difference between measured and simulated data according to the following equation [5, 8]:

$$\begin{aligned} &\min_{u,c} \phi = [m(u,c,t) - \widehat{m}(t)]^T \cdot \sum\nolimits_{\mathbf{m}}^{-1} \cdot [m(u,c,t) - \widehat{m}(t)] \\ &\text{s.t. } N \cdot u \geq 0, c \geq 0 \end{aligned} \quad (6)$$

where ϕ is the objective function to be minimized, u is a vector of free fluxes, c is a vector of metabolite concentrations, t is time, $m(u,c,t)$ is a vector of simulated measurements, $\widehat{m}(t)$ is a vector of observed measurements, $\sum_{\mathbf{m}}$ is the measurement covariance matrix, and N is the null space of the stoichiometric matrix. A reduced gradient method can be implemented to handle the linear constraints of this problem within a Levenberg–Marquardt nonlinear least-squares solver [38, 39]. Alternatively, gradient-free optimization approaches have been applied by Noh et al. [5]. **Steps 7** and **8** do not need to be solved when using a gradient-free method.

7. *Differentiate the isotopomer balances with respect to each free flux to generate the sensitivity equations. Solve the sensitivity equations to calculate the first derivatives of the measured MIDs with respect to the free parameters.* Estimation of both the unknown fluxes and pool sizes using INST-MFA is accomplished by finding a best-fit solution to the inverse problem. Efficient solution of this problem typically relies on optimization algorithms that choose the search direction based on the gradient of the least-squares objective function (*see* Eq. 6) with respect to all adjustable parameters. The most accurate and least expensive way to obtain the required gradient information is to integrate a system of sensitivity equations whose solution describes how the calculated MIDs vary in response to changes in the model parameters. Implicit differentiation of Eq. 3 yields the following sensitivity equation:

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathbf{X}_n}{\partial p} = \mathbf{C}_n^{-1}\cdot\mathbf{A}_n\cdot\frac{\partial \mathbf{X}_n}{\partial p} + \frac{\partial\left(\mathbf{C}_n^{-1}\cdot\mathbf{A}_n\right)}{\partial p}\cdot\mathbf{X}_n + \mathbf{C}_n^{-1}\cdot\mathbf{B}_n\cdot\frac{\partial \mathbf{Y}_n}{\partial p} + \frac{\partial\left(\mathbf{C}_n^{-1}\cdot\mathbf{B}_n\right)}{\partial p}\cdot\mathbf{Y}_n \tag{7}$$

where p is the vector of adjustable flux and pool size parameters. This system of equations can be solved in tandem with those of Eq. 3, and the time-dependent sensitivities can be used to evaluate the objective function gradient during each iteration of the INST-MFA inverse problem.

8. *Calculate the Hessian and gradient of the objective function.* Differentiation of the objective function with respect to the flux and pool size parameters will yield equations for the Hessian matrix **H** and the gradient vector **J**:

$$\mathbf{J} = \left(\frac{\partial m}{\partial p}\right)^{\mathrm{T}}\cdot\sum\nolimits_m^{-1}\cdot(m-\widehat{m}) \tag{8}$$

$$\mathbf{H} = \left(\frac{\partial m}{\partial p}\right)^{\mathrm{T}}\cdot\sum\nolimits_m^{-1}\cdot\left(\frac{\partial m}{\partial p}\right) \tag{9}$$

where $(\partial m/\partial p)$ is the matrix of sensitivities of the simulated measurements with respect to the parameters. Both the Hessian and gradient will be evaluated at each iteration if a gradient-based optimization algorithm is employed.

9. *Evaluate the goodness of fit.* Flux and pool size estimation is initiated with random values for all fluxes and pool sizes. The estimation algorithm will continue until no further improvements are made to the sum of squared residuals (SSRES) of the objective function. However, the fact that INST-MFA yields a set of fluxes that minimize the difference between the observed and simulated measurements does not mean that the flux model is adequate. Testing the goodness of fit will determine whether the optimal solution is statistically acceptable based on the minimized SSRES. At convergence, the minimized variance-weighted SSRES is a stochastic variable drawn from a chi-square distribution with n–p degrees of freedom (DOF), where n is the number of independent measurements and p is the number of estimated parameters. The SSRES that is calculated should therefore be in the interval $\left[\chi^2_{\alpha/2}, \chi^2_{1-\alpha/2}\right]$, where α is a chosen threshold value corresponding to the desired confidence level (e.g., 0.05 for 95 % confidence or 0.01 for 99 % confidence). The model fit is accepted when the SSRES falls within the limits of the expected chi-square range [40]. Additionally, the distribution of residuals should be assessed for normality. The standard deviation-weighted residuals should be normally distributed with a mean of zero and standard deviation of one.

One approach that can be used to evaluate the hypothesis that the residuals are normally distributed is the Lilliefors test [41]. Various plots can also be constructed to assess normality of the residuals.

10. *Calculate the 95 % confidence intervals using either continuation methods or Monte Carlo analysis.* Once an optimal solution has been obtained, nonlinear confidence intervals on the fitted parameters should be computed using robust, global methods instead of relying solely upon local standard errors. The local standard errors can be easily obtained from the parameter covariance matrix at the optimal solution; however, they do not accurately reflect changing sensitivities at points removed from the optimal solution. Parameter continuation can be performed to calculate accurate upper and lower bounds on the 95 % confidence interval for each flux or pool size parameter [40]. This determines the sensitivity of the minimized SSE to varying a single parameter away from its optimal value while allowing the remaining parameters to adjust in order to minimize $\Delta\phi$. Large confidence intervals indicate that the flux cannot be estimated precisely. On the other hand, small confidence intervals indicate that the flux is well determined. Monte Carlo simulation can also be used to calculate the 95 % confidence intervals. This method is typically more expensive than the parameter continuation approach but is expected to yield similar results.
11. *Report the flux values and flux uncertainties.* Once an acceptable fit to the experimental measurements has been achieved and confidence intervals have been computed for all parameters, the results are best summarized visually in the form of a flux map. Figure 5 shows an example plant flux map under photoautotrophic growth conditions using INST-MFA, involving multiple subcellular compartments. Several software tools have been recently developed, which aid in the construction of these maps (*see* **Note 14**).

4 Notes

1. For labeling experiments involving *Arabidopsis thaliana*, it is suggested that 4-week-old plants be used. This is because 4-week-old plants at metabolic steady state provide the maximum biomass, RuBisCO content, and chlorophyll content without the production of reproductive organs (i.e., flowers and seeds). Thus, the 4-week-old plants are active and vegetative. Additionally, metabolic steady state can be determined by measuring the sucrose and starch production of leaves; if the production of sucrose and starch remains consistent, then metabolic steady state can be established.

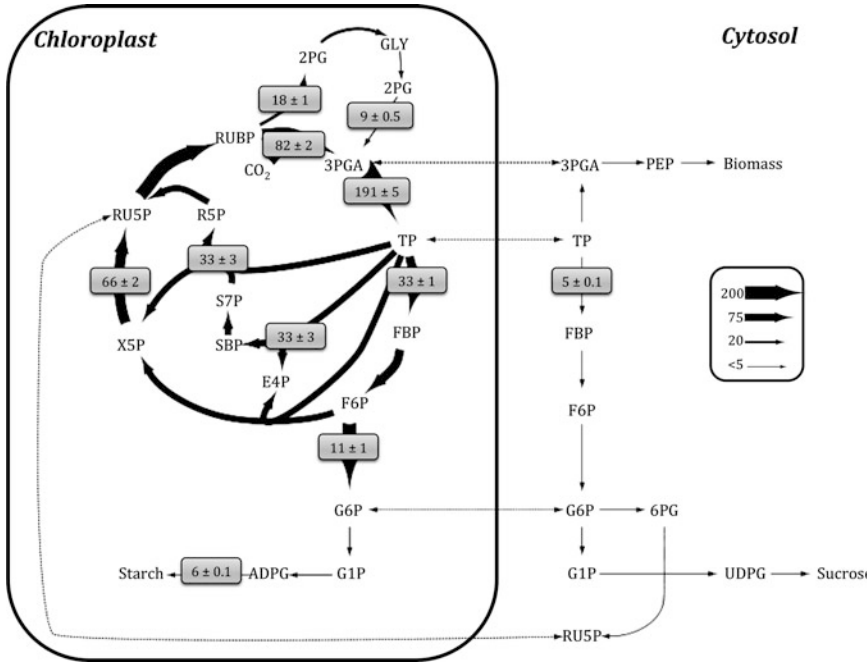

Fig. 5 Example flux map for a plant INST-MFA study under photoautotrophic growth conditions. This flux map shows a hypothetical flux map associated with a plant INST-MFA study involving multiple subcellular compartments. Values are represented as $M \pm$ SE, where M is the median of the flux confidence interval and SE is the estimated standard error of M. *Arrow* thickness is scaled proportional to net flux. *Dotted arrows* indicate transport reactions between compartments

2. Derivatization agents such as methoxyamine (MOX), trimethylsiloxane (TMS), or *tert*-butyl dimethylsiloxane (TBDMS) are typical for GC-MS analysis. The MOX reaction protects ketone and aldehyde functional groups and thereby prevents the formation of multiple TMS or TBDMS derivatives. This step is unnecessary if no ketone or aldehyde functional groups are present in the analytes of interest. TMS and TBDMS derivatives produce several characteristic fragment ions that facilitate identification [42]. Huege et al. [10] provide a list of some metabolites derivatized with TMS used for isotopomer analysis via GC-MS analysis. Ahn and Antoniewicz [43] provide a similar list for TBDMS-derivatized metabolites for GC-MS analysis.
3. Available GC-MS freeware include AMDIS (http://chemdata.nist.gov/mass-spc/amdis/) and Wsearch32 (http://www.wsearch.com.au/wsearch32/wsearch32.htm). Two popular freeware programs for LC-MS/MS data analysis are MZmine and XCMS, the latter of which runs in the R statistical programming environment. Both programs require the user to convert raw data files into a nonproprietary format such as mzXML, NetCDF, or mzData. Conversion to mzXML format

can be accomplished using one of several instrument-specific software tools developed and maintained by the Seattle Proteome Center (http://tools.proteomecenter.org/software.php).

4. Samples should be collected more frequently near the beginning of the tracer experiment, as the isotopic labeling will be changing most rapidly during this initial time period. For example, Wiechert et al. [5] have recommended using an approach where the length of each time interval increases exponentially (e.g., 1, 2, 4, 8, 16) following an initial period where uniformly spaced samples are collected at the maximum rate.
5. The results from the fractional labeling measurements are not immediately employable in their acquired form but must be deconvoluted to determine the actual MIDs for the metabolite pools in each compartment. To do this we first describe the contribution of a metabolite to each fraction. This is a function of the metabolite pool size and the presence of the compartment in the fraction, as described by Eq. 10:

$$F_{i,j} = A_i \cdot E_{i,j} \tag{10}$$

where $F_{i,j}$ is the contribution of metabolite A of compartment i to fraction j, A_i is the pool size of the metabolite in compartment i, and $E_{i,j}$ is the relative compartmental activity of compartment i in fraction j. The fractional contribution can then be normalized such that

$$B_{i,j} = \frac{A_i \cdot E_{i,j}}{\sum_i A_i \cdot E_{i,j}} \tag{11}$$

Using this relationship we can write calculated fractional MIDs as a function of $B_{i,j}$ and the actual compartmental MIDs:

$$\mathrm{MID}_{\mathrm{calc},j} = \sum_i B_{i,j} \cdot \mathrm{MID}_i \tag{12}$$

This problem can then be solved for all MID_i and A_i simultaneously using a nonlinear least-squares fitting to minimize the difference between the calculated fractional MIDs for all fractions and those observed experimentally, and confidence intervals can be obtained accordingly.

6. For *A. thaliana*, the NAF gradient range to achieve maximum separation of biomass is 1.43–1.62 g/cm^3 [22]. For soybeans, this range has been found to be 1.28–1.59 g/cm^3 [44].
7. The split ratio (in split mode) or purge activation time (in splitless mode) can be varied to achieve total ion counts in the desired range. Split mode is appropriate for concentrated samples while splitless mode is most appropriate for dilute samples. Typical split ratios vary from 1:5 to 1:100, with 1:10 representing a good initial value. Typical purge times range from 0.5 to 2 min, with 1 min representing a good initial value. Note that GC inlet liners are specific for either split or splitless operation and are not

interchangeable. Also note that when using splitless mode, the initial column temperature should be near or below the boiling point of the solvent in which the sample is dissolved [42].

8. The optimal linear velocity is in the range 20–40 cm/s for helium. Agilent's FlowCalc tool or the instrument control software can help determine what the linear velocity will be for a particular combination of flow rate, column diameter, temperature, and pressure.
9. This step may be necessary for dilute samples in order to ensure complete conversion. However, it may be skipped for concentrated samples.
10. In order to dissolve tributylamine completely in water, the tributylamine and acetic acid are first mixed together in a dry flask before the requisite amount of ultrapure water is added. The solution is then filtered through a 0.45 μm membrane prior to use. The final pH should be 4.5–5.
11. To increase sensitivity, the method should be divided into multiple time segments with different MRM transitions scanned in each interval. Dwell time of each transition should be optimized such that the total cycle time in each time segment does not exceed 2 s. This will provide at least 10–15 scans of each chromatographic peak as it elutes from the column.
12. Networks used for MFA typically include glycolysis, pentose phosphate pathway, amino acid metabolism, TCA cycle, and various amphibolic pathways that interact with the TCA cycle. This backbone of central metabolic pathways may be further augmented by additional reactions of interest to a particular investigation. Some helpful online public libraries include Kyoto Encyclopedia of Genes and Genomes (KEGG; http://www.genome.jp/kegg/), BioCyc (http://biocyc.org/), metaTIGER (http://www.bioinformatics.leeds.ac.uk/metatiger/), ENZYME (http://enzyme.expasy.org/), and BRENDA (http://www.brenda-enzymes.info/).
13. Blocks are defined by sets of EMUs whose MIDs are mutually dependent within the context of the EMU reaction network. The EMUs are arranged into blocks where the EMU reaction network is regarded as a directed graph, where the nodes represent EMUs and edges represent EMU reactions. An $N \times N$ adjacency matrix is constructed for the directed graph, where N is the total number of EMUs. A nonzero entry $a(i, j)$ of the adjacency matrix indicates the dependence of the ith EMU's MID on the jth EMU's MID. A Dulmage–Mendelsohn decomposition is performed on the adjacency matrix, returning an upper block triangular matrix from which the diagonal blocks are extracted. Blocks can be arranged

so that each is a self-contained subproblem that depends on the outputs of previously solved blocks, creating a cascaded system.

14. Several tools have been recently developed for flux visualization in the context of metabolic networks, such as FluxMap [45], FluxViz [46], fa-BINA [47], Omix [48], BioCyc Omics Viewer [49], Reactome Skypainter [50], Pathway Projector [51], MetaFluxNet [52], and OptFlux [53].
15. Abbreviations: *3PGA*, 3-phosphoglycerate; *αKG*, alpha-ketoglutarate; *CIT*, citrate; *DHAP*, dihydroxyacetone phosphate; *F6P*, fructose-6-phosphate; *FBP*, fructose 1,6-bisphosphate; *FUM*, fumarate; *G6P*, glucose 6-phosphate; *GAP*, glyceraldehyde 3-phosphate; *ICT*, isocitrate; *MAL*, malate; *PEP*, phosphoenol-pyruvate; *PYR*, pyruvate; *R5P*, ribose 5-phosphate; *RU5P*, ribulose 5-phosphate; *RUBP*, ribulose 1,5-bisphosphate; *S7P*, sedoheptulose 7-phosphate; *SUC*, succinate; *X5P*, xylulose 5-phosphate.

Acknowledgements

This work was supported by NSF EF-1105249 and DOE DE-SC0008118. LJJ was supported by DOE SCGF DE-AC05-06OR23100 and GAANN P200A090323.

Appendix: Simple Network Example for INST-MFA Calculations

A simple metabolic network appears in Fig. 6 as an example of how to construct the stoichiometric matrix *S*, identify the set of EMUs required to simulate MIDs of measured metabolites, and set up

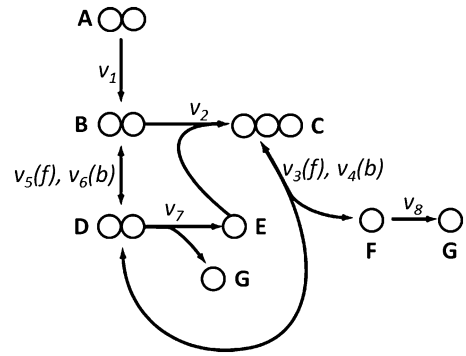

Fig. 6 Simple metabolic network used to illustrate the decomposition into EMUs. Atom transitions for the reactions in this model are given in Table 3. The network fluxes are assumed to be constant since the system is at metabolic steady state. Extracellular metabolite A is assumed to be at a fixed state of isotopic labeling to which intracellular metabolites B, C, D, E, F, and G adapt over time

Table 3
Stoichiometry and atom transitions for the reactions in the example metabolic network

Reaction number	Reaction stoichiometry	Atom transitions
1	A → B	ab → ab
2	B + E → C	ab + c → abc
3 and 4	C ↔ D + F	abc ↔ cb + a
5 and 6	B ↔ D	ab ↔ ab
7	D → E + G	ab → b + a
8	F → G	a → a

dynamic isotopomer balances on these EMUs. Table 3 delineates the atom transitions for the network. In this network example, metabolite A is the sole substrate and metabolite G is the only final product. The intermediary metabolites B, C, D, E, and F are assumed to be at metabolic steady state but isotopically nonstationary.

The stoichiometric matrix S is shown below, which has $k = 5$ intermediary metabolites and $j = 8$ fluxes, resulting in a 5×8 matrix:

$$S = \begin{bmatrix} 1 & -1 & 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & 1 & -1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 1 & -1 & -1 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & 0 & -1 \end{bmatrix}$$

Therefore, $S \cdot v = 0$ is expressed in vector form as

$$\begin{bmatrix} v_1 - v_2 - v_5 + v_6 \\ v_2 - v_3 + v_4 \\ v_3 - v_4 + v_5 - v_6 - v_7 \\ -v_2 + v_7 \\ v_3 - v_4 - v_8 \end{bmatrix} = 0$$

A systematic method of EMU network decoupling in which metabolite units are grouped into mutually dependent blocks is described through this simple network example. For this example, we will set up the simplest possible model to simulate the MID of metabolite C, i.e., EMU C_{123}. First, we need to identify all the possible EMUs that contribute to the formation of C_{123}—in this reaction model, C_{123} is formed from the condensation of B_{12} + E_1 and D_{12} + F_1 in reactions 2 and 4, respectively. This is recorded and the process is then repeated for all new EMUs, starting with the largest EMU in size; in this case, all EMUs of size 3 have already been identified. Next, the process is repeated to determine all the

Table 4
Complete list of EMU reactions generated for metabolite C

Reaction number	EMU reaction	EMU reaction size balance	Block
2	$B_{12} + E_1 \rightarrow C_{123}$	$2 + 1 = 3$	6
4	$D_{12} + F_1 \rightarrow C_{123}$	$2 + 1 = 3$	6
2	$B_2 + E_1 \rightarrow C_{23}$	$1 + 1 = 2$	5
4	$D_{12} \rightarrow C_{23}$	$2 = 2$	5
3	$C_{23} \rightarrow D_{12}$	$2 = 2$	5
5	$B_{12} \rightarrow D_{12}$	$2 = 2$	5
6	$D_{12} \rightarrow B_{12}$	$2 = 2$	5
1	$A_{12} \rightarrow B_{12}$	$2 = 2$	5
2	$B_1 \rightarrow C_1$	$1 = 1$	4
4	$F_1 \rightarrow C_1$	$1 = 1$	4
3	$C_1 \rightarrow F_1$	$1 = 1$	4
2	$E_1 \rightarrow C_3$	$1 = 1$	3
4	$D_1 \rightarrow C_3$	$1 = 1$	3
3	$C_3 \rightarrow D_1$	$1 = 1$	3
5	$B_1 \rightarrow D_1$	$1 = 1$	3
6	$D_1 \rightarrow B_1$	$1 = 1$	3
1	$A_1 \rightarrow B_1$	$1 = 1$	3
7	$D_2 \rightarrow E_1$	$1 = 1$	2
2	$B_2 \rightarrow C_2$	$1 = 1$	1
4	$D_2 \rightarrow C_2$	$1 = 1$	1
1	$A_2 \rightarrow B_2$	$1 = 1$	1
6	$D_2 \rightarrow B_2$	$1 = 1$	1
5	$B_2 \rightarrow D_2$	$1 = 1$	1
3	$C_2 \rightarrow D_2$	$1 = 1$	1

Subscripts denote atoms that are part of their respective EMUs. The EMU reactions are also divided into their respective blocks after Dulmage–Mendelsohn decomposition has been applied

EMUs of size 2 that were previously identified, starting with D_{12}. D_{12} is formed from two different reactions—from B_{12} in reaction 5 and from C_{23} in reaction 3. Following this, we determine which reactions form C_{23}; C_{23} is formed from D_{12} in reaction 4 and $B_2 + E_1$ in reaction 2. Finally, we need to determine which reactions form B_{12}. B_{12} is formed from A_{12} and D_{12} in reactions 1 and 6, respectively. A_{12} is a network substrate and is not produced by any other reactions and D_{12} has already been considered in the previous step. Therefore, all EMU reactions of size 2 have been identified. The process is repeated once again for EMUs of size 1, until all the EMUs have been traced back to network substrates or previously identified EMUs. Table 4 shows the complete EMU decomposition of this system, which involves 24 EMU reactions connecting 16 EMUs.

After EMU decomposition, the reaction network can be further decoupled into blocks, which group together minimal sets of

mutually dependent metabolite units that must be solved simultaneously. Figure 7 shows the EMU network decomposition for the simple network example after block decoupling. The blocks are arranged so that each one is a self-contained subproblem, which will depend on the outputs of the previously solved blocks. Therefore, EMUs in block 1 should first be solved, then block 2, etc.

The EMU reactions obtained from network decomposition and block decoupling form the new basis for generating system equations. The decoupled blocks can be arranged into a cascaded system of ODEs with the following form, as described in Subheading 3.9:

$$\mathbf{C}_n \cdot \frac{\mathrm{d}\mathbf{X}_n}{\mathrm{d}t} = \mathbf{A}_n \cdot \mathbf{X}_n + \mathbf{B}_n \cdot \mathbf{Y}_n$$

The concentration matrix $\mathbf{C}_n$ is a diagonal matrix whose elements are pool sizes corresponding to EMUs represented in $\mathbf{X}_n$. $\mathbf{X}_n$ comprises row vectors that represent the MIDs of each EMU and $\mathrm{d}\mathbf{X}_n/\mathrm{d}t$ is the time derivative of $\mathbf{X}_n$. Analogously, the input matrix $\mathbf{Y}_n$ also comprises row vectors that represent MIDs of EMUs that have been previously calculated. The system matrices $\mathbf{A}_n$ and $\mathbf{B}_n$ come from calculating the "true" flux vectors (v) based on the chosen free fluxes (u) and null space matrix (N). Furthermore, in the decoupled blocks, the full MID of products formed from condensation reactions can be obtained from the convolution (or Cauchy product, denoted by "×") of MIDs of preceding EMUs. In the case of C_{123}, these MIDs are B_{12} and E_1 or D_{12} and F_1, i.e., $C_{123} = B_{12} \times E_1$ or $C_{123} = D_{12} \times F_1$. The following equations represent the system of ODEs for the simple network example:

$$\begin{bmatrix} C_\mathrm{C} & 0 & 0 \\ 0 & C_\mathrm{B} & 0 \\ 0 & 0 & C_\mathrm{D} \end{bmatrix} \begin{bmatrix} \frac{\mathrm{dC_2}}{\mathrm{d}t} \\ \frac{\mathrm{dB_2}}{\mathrm{d}t} \\ \frac{\mathrm{dD_2}}{\mathrm{d}t} \end{bmatrix} = \begin{bmatrix} -v_2 - v_4 & v_2 & v_4 \\ 0 & -v_1 - v_6 & v_6 \\ v_3 & v_5 & -v_3 - v_5 \end{bmatrix} \begin{bmatrix} \mathrm{C_2} \\ \mathrm{B_2} \\ \mathrm{D_2} \end{bmatrix} + \begin{bmatrix} 0 \\ v_1 \\ 0 \end{bmatrix} [\mathrm{A_2}]$$

$$[C_\mathrm{E}] \left[\frac{\mathrm{dE_1}}{\mathrm{d}t} \right] = [-v_7][\mathrm{E_1}] - [v_7][\mathrm{D_2}]$$

$$\begin{bmatrix} C_\mathrm{C} & 0 & 0 \\ 0 & C_\mathrm{D} & 0 \\ 0 & 0 & C_\mathrm{B} \end{bmatrix} \begin{bmatrix} \frac{\mathrm{dC_3}}{\mathrm{d}t} \\ \frac{\mathrm{dD_1}}{\mathrm{d}t} \\ \frac{\mathrm{dB_1}}{\mathrm{d}t} \end{bmatrix} = \begin{bmatrix} v_2 + v_4 & -v_4 & 0 \\ -v_3 & v_3 + v_5 & -v_5 \\ 0 & -v_6 & v_1 + v_6 \end{bmatrix} \begin{bmatrix} \mathrm{C_3} \\ \mathrm{D_1} \\ \mathrm{B_1} \end{bmatrix} + \begin{bmatrix} v_2 & 0 \\ 0 & 0 \\ 0 & v_1 \end{bmatrix} \begin{bmatrix} \mathrm{E_1} \\ \mathrm{A_1} \end{bmatrix}$$

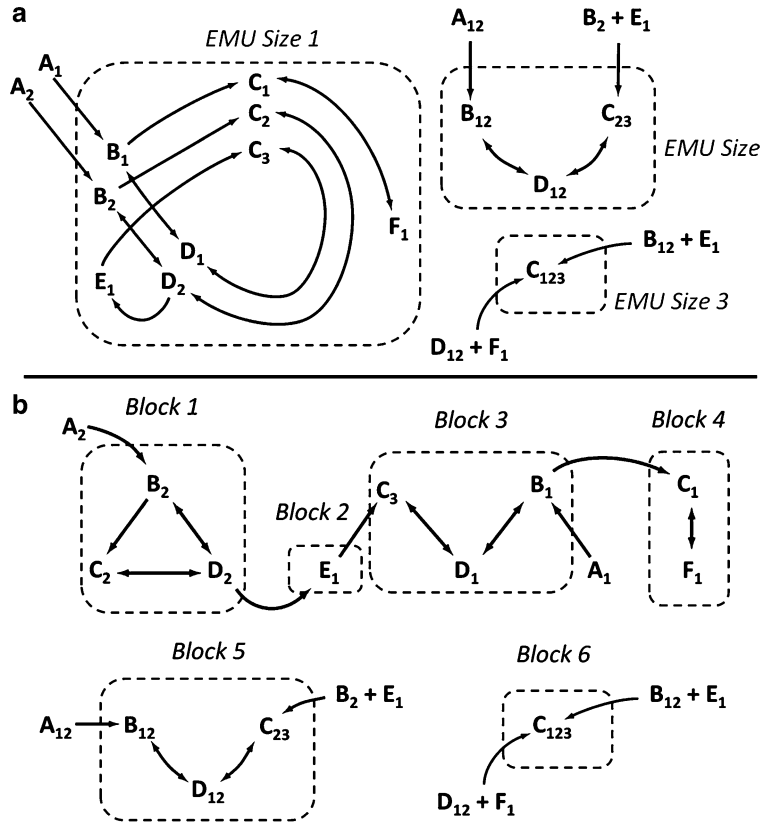

Fig. 7 (**a**) EMU network decomposition for simple example network (Fig. 6) generated to simulate the labeling of metabolite C. The EMU network was decoupled based on EMU size and network connectivity. (**b**) EMU network decomposition for the same network using block decoupling. *Subscripts* refer to the atoms that are contained within the EMU

$$\begin{bmatrix} C_{\mathrm{C}} & 0 \\ 0 & C_{\mathrm{F}} \end{bmatrix} \begin{bmatrix} \frac{\mathrm{dC_1}}{\mathrm{d}t} \\ \frac{\mathrm{dF_1}}{\mathrm{d}t} \end{bmatrix} = \begin{bmatrix} -v_2 - v_4 & v_4 \\ v_3 & -v_3 \end{bmatrix} \begin{bmatrix} \mathrm{C_1} \\ \mathrm{F_1} \end{bmatrix} + \begin{bmatrix} v_2 \\ 0 \end{bmatrix} [\mathrm{B_1}]$$

$$\begin{bmatrix} C_{\mathrm{C}} & 0 & 0 \\ 0 & C_{\mathrm{D}} & 0 \\ 0 & 0 & C_{\mathrm{B}} \end{bmatrix} \begin{bmatrix} \frac{\mathrm{dC_{23}}}{\mathrm{d}t} \\ \frac{\mathrm{dD_{12}}}{\mathrm{d}t} \\ \frac{\mathrm{dB_{12}}}{\mathrm{d}t} \end{bmatrix} = \begin{bmatrix} -v_2 - v_4 & v_4 & 0 \\ v_3 & -v_3 - v_5 & v_5 \\ 0 & v_6 & -v_1 - v_6 \end{bmatrix} \begin{bmatrix} \mathrm{C_{23}} \\ \mathrm{D_{12}} \\ \mathrm{B_{12}} \end{bmatrix} + \begin{bmatrix} v_2 & 0 \\ 0 & 0 \\ 0 & v_1 \end{bmatrix} \begin{bmatrix} \mathrm{B_2 \times E_1} \\ \mathrm{A_{12}} \end{bmatrix}$$

$$[C_C]\left[\frac{dC_{123}}{dt}\right] = [-v_2 - v_4][C_{123}] + [\,v_2 \quad v_4\,]\begin{bmatrix} B_{12} \times E_1 \\ D_{12} \times F_1 \end{bmatrix}$$

Solving this system of ODEs will simulate the EMU labeling trajectories needed to calculate the time-dependent MID of metabolite C. The flux and pool size parameters can then be adjusted iteratively using an optimization search algorithm to converge on parameter values that minimize the lack of fit with experimental mass isotopomer data.

References

1. Wiechert W (2001) 13C Metabolic flux analysis. Metab Eng 3:195–206
2. Sauer U (2006) Metabolic networks in motion: 13C-based flux analysis. Mol Syst Biol 2:62
3. Nöh K, Grönke K, Luo B et al (2007) Metabolic flux analysis at ultra short time scale: isotopically non-stationary 13C labeling experiments. J Biotechnol 129:249–267
4. Shastri AA, Morgan JA (2007) A transient isotopic labeling methodology for 13C metabolic flux analysis of photoautotrophic microorganisms. Phytochemistry 68:2302–2312
5. Nöh K, Wahl A, Wiechert W (2006) Computational tools for isotopically instationary 13C labeling experiments under metabolic steady state conditions. Metab Eng 8:554–577
6. Wiechert W, Nöh K (2005) From stationary to instationary metabolic flux analysis. Adv Biochem Eng Biotechnol 92:145–172
7. Young JD, Walther JL, Antoniewicz MR et al (2008) An elementary metabolite unit (EMU) based method of isotopically nonstationary flux analysis. Biotechnology 99:686–699
8. Antoniewicz MR, Kelleher JK, Stephanopoulos G (2007) Elementray metabolite units (EMU): a novel framework for modeling isotopic distributions. Metab Eng 9:68–86
9. Arrivault S, Guenther M, Ivakov A et al (2009) Use of reverse-phase liquid chromatography, linked to tandem mass spectrometry, to profile the Calvin cycle and other metabolic intermediates in Arabidopsis rosettes at different carbon dioxide concentrations. Plant J 59: 826–839
10. Huege J, Sulpice R, Gibon Y et al (2007) GC-EI-TOF-MS analysis of in vivo carbon-partitioning into soluble metabolite pools of higher plants by monitoring isotope dilution after 13CO2 labelling. Phytochemistry 68:2258–2272
11. Römisch-Margl W, Schramek N, Radykewicz T et al (2007) 13CO2 as a universal metabolic tracer in isotopologue perturbation experiments. Phytochemistry 68:2273–2289
12. Sekiyama Y, Kikuchi J (2007) Towards dynamic metabolic network measurements by multi-dimensional NMR-based fluxomics. Phytochemistry 68:2320–2329
13. Geigenberger P, Tiessen A, Meurer J (2011) Use of non-aqueous fractionation and metabolomics to study chloroplast function in Arabidopsis. Methods Mol Biol 775:135–160
14. Ausloos P, Clifton CL, Lias SG et al (1999) The critical evaluation of a comprehensive mass spectral library. J Am Soc Mass Spectrom 10:287–299
15. Kopka J, Schauer N, Krueger S et al (2005) GMD@CSB.DB: the golm metabolome database. Bioinformatics 21:1635–1638
16. Kind T, Wohlgemuth G, Lee DY et al (2009) FiehnLib: mass spectral and retention index libraries for metabolomics based on quadrupole and time-of-flight gas chromatography/mass spectrometry. Anal Chem 81: 10038–10048
17. Smith CA, O'Maille G, Want EJ et al (2005) METLIN: a metabolite mass spectral database. Ther Drug Monit 27:747–751
18. Wishart DS, Tzur D, Knox C et al (2007) HMDB: the human metabolome database. Nucleic Acids Res 35:D521–D526
19. Horai H, Arita M, Kanaya S et al (2010) MassBank: a public repository for sharing mass spectral data for life sciences. J Mass Spectrom 45:703–714
20. Gerhardt R, Heldt HW (1984) Measurement of subcellular metabolite levels in leaves by fractionation of freeze-stopped material in nonaqueous media. Plant Physiol 14:542–547

21. Bergmeyer H (1983) Methods of enzymatic analysis. Verlag Chemie, Deerfield Beach, FL
22. Krueger S, Giavalisco P, Krall L et al (2011) A topological map of the compartmentalized Arabidopsis thaliana leaf metabolome. PLoS One 6:e17806
23. Sigma-Aldrich Enzymatic assay of phospho (enol)pyruvate carboxylase. Presented at the (1994).
24. Maeda H, Yoo H (2011) Prephenate aminotransferase directs plant phenylalanine biosynthesis via arogenate. Nat Chem Biol 7:19–21
25. Li Y-T, Primate DR (1967) Studies on the glycosidases in jack bean meal. J Biol Chem 242:5474–5480
26. Antoniewicz MR, Kelleher JK, Stephanopoulos G (2007) Accurate assessment of amino acid mass isotopomer distributions for metabolic flux analysis. Anal Chem 79:7554–7559
27. Wittmann C, Hans M, Heinzle E (2002) In vivo analysis of intracellular amino acid labelings by GC/MS. Anal Biochem 307:379–382
28. Roessner U, Wagner C, Kopka J et al (2000) Technical advance: simultaneous analysis of metabolites in potato tuber by gas chromatography-mass spectrometry. Plant J 23:131–142
29. Yang L, Kasumov T, Yu L et al (2006) Metabolomic assays of the concentration and mass isotopomer distribution of gluconeogenic and citric acid cycle intermediates. Metabolomics 2:85–94
30. Luo B, Groenke K, Takors R et al (2007) Simultaneous determination of multiple intracellular metabolites in glycolysis, pentose phosphate pathway and tricarboxylic acid cycle by liquid chromatography-mass spectrometry. J Chromatogr A 1147:153–164
31. Young JD, Shastri AA, Stephanopoulos G et al (2011) Mapping photoautotrophic metabolism with isotopically nonstationary (13)C flux analysis. Metab Eng 13:656–665
32. Fernandez CA, Des Rosiers C, Previs SF et al (1996) Correction of 13C mass isotopomer distributions for natural stable isotope abundance. J Mass Spectrom 31:255–262
33. Schmidt K, Carlsen M, Nielsen J et al (1997) Modeling isotopomer distributions in biochemical networks using isotopomer mapping matrices. Biotechnol Bioeng 55:831–840
34. Wiechert W, Siefke C, de Graaf AA et al (1997) Bidirectional reaction steps in metabolic networks: II. Flux estimation and statistical analysis. Biotechnol Bioeng 55:118–135
35. Foster LV (1986) Rank and null space calculations using matrix decomposition without column interchanges. Linear Algebra Appl 74:47–71
36. Dulmage AL, Mendelsohn NS (1958) Coverings of bipartite graphs. Can J Math 10: 517–534
37. Pothen A, Fan CJ (1990) Computing the block triangular form of a sparse matrix. ACM Trans Math Softw 16:303–324
38. Gill PE, Murray W, Wright MH (1981) Practical optimization. Academic, London
39. Madsen K, Nielsen HB, Tigleff O (2004) Methods for non-linear least squares problems. http://www2.imm.dtu.dk/pubdb/views/publication_details.php?id=3215
40. Antoniewicz MR, Kelleher JK, Stephanopoulos G (2006) Determination of confidence intervals of metabolic fluxes estimated from stable isotope measurements. Metab Eng 8:324–337
41. Conover WJ (1999) Practical nonparameteric statistics. Wiley, New York
42. Kitson FG, Larsen BS, McEwen CN (1996) Gas chromatography and mass spectrometry: a practical guide. Academic, San Diego
43. Ahn WS, Antoniewicz MR (2011) Metabolic flux analysis of CHO cells at growth and non-growth phases using isotopic tracers and mass spectrometry. Metab Eng 13:598–609
44. Benkeblia N, Shinano T, Osaki M (2007) Metabolite profiling and assessment of metabolome compartmentation of soybean leaves using non-aqueous fractionation and GC-MS analysis. Metabolomics 3:297–305
45. Rohn H, Hartmann A, Junker A et al (2012) FluxMap: a VANTED Add-on for the visual exploration of flux distributions in biological networks. BMC Syst Biol 6:33
46. König M, Holzhütter H-G (2010) Fluxviz—cytoscape plug-in for visualization of flux distributions in networks. Genome informatics. International Conference on Genome Informatics. 24:96–103
47. Hoppe A, Hoffmann S, Gerasch A et al (2011) FASIMU: flexible software for flux-balance computation series in large metabolic networks. BMC Bioinformatics 12:28
48. Droste P, Miebach S, Niedenführ S et al (2011) Visualizing multi-omics data in metabolic networks with the software Omix: a case study. Biosystems 105:154–161
49. Paley SM, Karp PD (2006) The pathway tools cellular overview diagram and Omics Viewer. Nucleic Acids Res 34:3771–3778
50. Matthews L, Gopinath G, Gillespie M et al (2009) Reactome knowledgebase of human biological pathways and processes. Nucleic Acids Res 37:D619–D622

51. Kono N, Arakawa K, Ogawa R et al (2009) Pathway projector: web-based zoomable pathway browser using KEGG atlas and Google Maps API. PLoS One 4:e7710
52. Lee SY, Lee D-Y, Hong SH et al (2003) MetaFluxNet, a program package for metabolic pathway construction and analysis, and its use in large-scale metabolic flux analysis of Escherichia coli. Genome informatics. International Conference on Genome Informatics. 14:23–33
53. Rocha I, Maia P, Evangelista P et al (2010) OptFlux: an open-source software platform for in silico metabolic engineering. BMC Syst Biol 4:45