Methods for improving the visualization and deconvolution of isotopic signals

GUILLAUME TCHERKEZ1,2, JALEH GHASHGHAIE2 & HOWARD GRIFFITHS3

1Plateforme Métabolisme-Métabolome, IFR 87, Bât. 630, 2Laboratoire d’écophysiologie végétale, CNRS UMR 8079, Bât. 362, Université Paris Sud-XI, 91405 Orsay cedex, France and 3Department of Plant Sciences, University of Cambridge, CB2 3EA, UK

ABSTRACT

Stable isotopes and their associated mechanistic frameworks have provided the means to model biological transformations from inorganic sources to organic sinks, and their impact on long-term terrestrial carbon reservoirs. However, stable isotopes have also the potential to diagnose the mechanisms by which biological systems operate, when using ‘multidimensional’ analyses of the isotopic outputs. For this purpose, we suggest that isotopic signals may be treated as mathematical vectors to reveal their interrelationships. These visualizations may be achieved, thanks to multidimensional representations (the ‘isotopology’ method), or by appreciating the colinearity of isotopic vectors to develop a clustering analysis (the ‘isotopomic’ method) similar to that used in molecular biology. Both methods converge to the same mathematical form, i.e. both lead to a covariation approach. Using simple practical examples, we argue that such procedures allow the deconvolution of plant biological systems by revealing the hierarchy of contributory physiological processes.

Key-words: stables isotopes; fractionation; vectors; array; clustering analysis.

INTRODUCTION

Earth and plant systems biology requires an understanding of both interannual biospheric processes and its contributory molecular mechanisms, and stable isotopes can be used to integrate such geochemical and biological systems (Yakir 2002). Systematic shifts in natural isotopic ratios (such as 13C/12C or 18O/16O) are based on fractionation processes which lead to discriminations between isotopes in biological cornerstone reactions like photosynthesis [in which ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco) discriminates against 13CO2] (Farquhar, Ehleringer & Hubick 1989; Tcherkez, Farquhar & Andrews 2006), leaf and canopy transpiration (discrimination against H218O and exchange as either C18O16O or water vapour) (Farquhar, Barbour & Henry 1998; Helliker & Griffiths 2007) or plant nitrate assimilation (discrimination against 15NO3–) (Ledgard, Woo & Bergersen 1985; Werner & Schmidt 2002; Tcherkez and Farquhar 2006).

There have been a number of recent reviews which have evaluated biological and ecological stable isotope transformations, their impact on carbon reservoirs in the atmosphere, plant and soil, and coupling with the hydrological cycle (Griffiths 1998; Dawson et al. 2002; Flanagan, Pataki & Ehleringer 2005; Griffiths & Jarvis 2005). Most recently, it has been suggested that transformations and isotopic distributions can be viewed as ‘isoscapes’ which trace interannual variations or geographical linkages (West et al. 2006).

In this paper, we propose that isotope data should be visualized more clearly to reveal the connectivity between isotopic signals, and diagnose underlying mechanisms within plant systems. This may be achieved with ‘isotopologies’, which represent multidimensional maps of observed isotopic distributions, or ‘isotopomics’, in which systematic isotopic enrichments and depletions (isotopic shifts) within biological systems are correlated as vectors using multivariate, differential displays and cluster analysis.

A MULTIDIMENSIONAL ‘ISOTOPOLOGY’

The clearest representation of a global isotopic output is the annual dance of photosynthetic carbon uptake and sequestration in summer, followed by respiratory CO2 release in winter, which drives seasonal variations in CO2 concentration. Particularly in the Northern Hemisphere, the heterogeneity and temporary disequilibrium in isotopic mass balance of atmospheric CO2 lead to a striking intraannual variation in 13C/12C isotope composition (denoted as δ13C). This interaction is classically represented as a ‘rug’ or ‘flying carpet’ diagram (whose data are available at the NOAA CMDL Carbon Cycle Cooperative Global Air Sampling Network website: http://wwwsr.cmdl.noaa.gov/ccgg/iadv), and encapsulates our vision of an isotopic landscape or an ‘isotopology’. We suggest that this kind of approach can be used for other isotopic data, which are normally only shown as one-dimensional isotopic diagrams.

Multiparametric representation: visualizing the multidependence of a given isotopic signal

To illustrate this point practically at the leaf level, we start from a multidimensional representation of exported
assimilates that are subsequently distributed in several pools and could be used for respiration or growth (Fig. 1). It shows how the time-course of the $\delta^{13}C$ signal in phloem sucrose depends on both circadian metabolic changes and internal parameters (such as the rate of starch synthesis) that are in turn modulated by environmental conditions. Briefly, the fructose-producing aldolase reaction of the chloroplast favours $^{13}C$ (enriching starch in this isotope). The remaining triose phosphates are used in the cytosol to produce ($^{13}C$-depleted) sucrose, so that sucrose found in the phloem during the day is $^{13}C$-depleted, contrary to sucrose produced during the night, simply because the latter comes from depolymerization of starch. This example typifies how the compartmentation of metabolic reactions (chloroplastic starch synthesis versus cytoplasmic sucrose synthesis), when associated with an isotope effect, scales up to the plant level. Multidimensional representations involving isotopes are useful to deconvolute the complexity of physiological processes, and more importantly, challenge the common interpretation of delta values; as can be seen from Fig. 1, a given $\delta^{13}C$ of sucrose may correspond to several possible drivers (starch synthesis, relative time), as well as environmental effects such as the impact of temperature on metabolic processes.

Such diel variations will be, among others, a factor contributing to the cellulose isotope signatures of tree rings, which in turn integrate seasonal climatic conditions (Schleser et al. 1999; MacCarroll & Loader 2004). Thus, as a further multidimensional representation, we provide a simple conceptual framework depicting the physiological complexity of ring deposition by the cambium (Fig. 2). A physical analogy involving transfer functions has also been proposed to improve the analyses of isotopic signals in rings (Schleser et al. 1999). Cellulose deposition is influenced by the (leaf) carbon source (as determined by $c_i/c_e$, the ratio of intracellular/external $CO_2$), and also by the contribution that reserves make as substrate for cellulose deposition (whether sucrose from starch in the dark, or sucrose during the day) and respiratory losses (Fig. 2a).

When modelled as a three-dimensional representation (Fig. 2b), the deviation of the $\delta^{13}C$ value of ring cellulose from the $\delta^{13}C$ value of fixed $CO_2$ can be seen as a function of the proportion of day sucrose (versus night sucrose) as a carbon source ($L$) and the relative respiratory rate ($r$), assuming a fixed respiratory discrimination of $-6\%$. It is clear that both parameters have an effect on the carbon isotope signature of cellulose, and a quite large deviation is obtained under certain conditions (Fig. 2b). For example, the larger the respiratory losses ($^{13}C$-enriched), the more $^{13}C$ depleted the remaining deposited material (so that $\delta^{13}C$–$\delta^{13}C$ increases). Further, the larger the contribution of ($^{13}C$-depleted) day sucrose to cellulose deposition, the lower the isotope composition of cellulose (so that $\delta^{13}C$–$\delta^{13}C$ increases as well). This representation integrates the key determinants of internal carbon signature, which may in turn be changed by environmental conditions (e.g. the increase of respiration with temperature, etc.).

The multidimensional representations of both Figs 1 and 2 are classical representations, in which the isotopic signal is a function of physiological drivers. Still, they are useful to deconvolute the complexity of physiological processes, and visualize how a given isotopic signal value may correspond to several possible values of the drivers. That said, multidimensional representations plotting several isotopic signals are perhaps more useful for deconvoluting interactions between bio(geo)chemical pathways, as we now show.

**Multiple isotopic signals: partitioning biochemical pathways from isotopic compositions**

When partitioning occurs during metabolism, graphical representations can be used to visualize the network of correlations. This can also be allied with additional techniques such as correlative clustering analysis, developed as follows. It will become apparent that both analyses – graphical multidimensional and clustering – are similar in the way that
the final mathematical formalism converges, but differ in terms of their visual realization.

As an example, we reanalyse data generated from the isotope composition of different metabolites extracted from a leaf during a series of labelling experiments (Nogués et al. 2004). A conventional multidimensional representation would show each compound on a separate axis. However, to display the interrelationships between compounds in such a graphical presentation, such data can be interpreted using a Bayesian analysis; for a given isotopic signal in compound number \( i \), one may look at (the) other compound(s) in the pathway which covary in association with that signal (Fig. 3). Thus, in mathematical terms, two isotopic signals \( i \) and \( j \) are best correlated if they follow each other in the graphical presentation, independently of others (denoted here as \( n \)). This means that signals \( n \) are in a perpendicular space, with the scalar product of \( i \) and \( j \) maximal, and the scalar products of \( i \) and any \( n \), or \( j \) and any \( n \), zero. In its general form, the scalar product would be as follows:

\[
\langle x_i, x_j \rangle = \sum_{k=1}^{K} x_{ik} x_{jk}
\]

where isotopic signal \( x_{ik} \) is that of compound \( i \) obtained during the measurement (or experiment) number \( k \). This is very similar to what the similarity-based clustering analysis discussed as follows gives.

The associated visualization may be, however, complicated when more than three isotopic signals are considered. While a classical bidimensional graph may be used for all the isotopic signals, a polar graph is, perhaps, a more convenient way to associate isotopic shifts in the different compounds, and determine whether they follow the same pathway. Figure 3 gives such an example using the data of Nogués et al. (2004), collected during labelling experiments allowing Phaseolus leaves to fix varying quantities of isotopically defined CO₂. We now plot the offset in isotopic signals between respired CO₂ (after illumination) and compound-specific signals in sucrose, starch, organic acids,

![Diagram showing the carbon isotope composition of cellulose in tree rings and the associated visualization.](image-url)
Figure 3. Representation of the isotopic shifts in the compound-specific δ¹³C values after labelling leaves with industrial CO₂ (data from Nogués et al. 2004), plotted as a polar graph in which each increment corresponds to a different experiment (increasing fixed C amount) and the angles (denoted as θ in the main text) represent the isotopic shifts normalized to 100 angular degrees. Suc, sucrose; OA, organic acids; Lip, lipids; Prot, heat-precipitated proteins.

The isotopic shifts are normalized to 100° (angular degrees) and plotted as angles (denoted as θ just below), and each labelling condition (experiment) corresponds to a radial increment. This allows one to see whether the isotopic pathways have the same shape, or rather, diverge. From a quantitative point of view, we note that two pathways A and B are exactly similar if the angles (θₐ and θ₈) are similar for each radial increment, i.e. if the corresponding cosine [cos(θₐ – θ₈)] value is 1, then the scalar product of the associated vectors is maximal. It can be seen in Fig. 3 that the pathway followed by CO₂ is closer to that of carbohydrates and correlates less well with other compounds. In other words, this indicates that the origin of CO₂ respired by leaves after illumination is carbohydrates. Less apparent is, however, the molecule from which CO₂ is derived (starch or sucrose). Another representation, coupled to a more quantitative correlative technique, is thus described as follows.

**ISOTOPOMICS**

The isotopomics concept is illustrated by Fig. 4. Here, we show how statistical methods derived from genomics can be used to deconvolute isotope signals. Traditional cluster analysis of gene transcription using microarrays (Eisen et al. 1998) indicates the abundance of many transcripts by colours. The cluster analysis of genes, where correlation coefficients or related mathematical parameters use similarities in transcription patterns to detect the coregulation of gene expression, can also be used to infer metabolic interactions and associations (Gachon et al. 2005). Similarly, arrays can be used to show isotopic modifications for different compounds of a biological system (in rows), induced by different environmental conditions (in columns); the clusters show the isotopic proximity of compounds or partners within a network of fluxes comprising the particular biological system of interest. This provides an effective mean of seeing the correlative network between compounds in the given biological system. In other words, we use here isotopes not only as classical tracers, but additionally as ‘correlators’. From a mathematical point of view, the principle is similar to the genetic clustering analysis, i.e. uses correlations of isotopic shift vectors as calculated, for example, by the normalized scalar product. If we denote x as the vectors of isotopic shifts (array conditions or ‘experiments’ are numbered from 1 to K), the similarity (or colinearity) index (denoted as S) between two vectors i and j has the following general form:

\[ S_{ij} = \frac{\sum_{k=1}^{K} x_{ik} x_{jk}}{\sqrt{\sum_{k=1}^{K} x_{ik}^2 \sum_{k=1}^{K} x_{jk}^2}} \]

The xᵢₖ values may be relative isotopic shifts in delta values or ratios or % in δ¹³C (in case of strong labelling), etc. All the Sᵢᵢ values generate a matrix which may be in turn treated by the clustering programme to generate clusters.

As a simple example, such an approach can be used at the leaf scale with compound-specific isotope composition following isotopic labelling (Fig. 4). Again, data from Nogués et al. (2004) (in which leaves were labelled with ¹³C-depleted CO₂) are used to generate a matrix of relative isotopic shifts from the initial value. Array conditions (columns) correspond to different labelling level (denoted as Labᵢ, where i is the amount of carbon fixed during the labelling period in Fig. 4). The colour brightness indicates the amplitude of the relative isotopic shift. The cluster analysis can be used to generate a phylogenetic-style tree indicated on the left-hand side of Fig. 4. In this particular example, metabolic associations can be easily concluded...
with this technique: CO₂ evolved in darkness correlates more closely with starch and sucrose; and lipids, on the other hand, show little correlation with other compounds because of their slow turnover. Thus, starch appears to be the most likely substrate feeding leaf respiration following illumination.

**PERSPECTIVES**

At the outset, we emphasized the current use of isotopes to understand the architecture of plant metabolism and communities in ecosystems. Isotopes indeed provide a key to these approaches, linking global carbon mass balance to basic enzymatic discrimination (e.g. Rubisco and respiratory processes) via elementary physiological events, which ultimately define the CO₂ signature in the atmosphere. However, more correlative approaches using isotopes, which have been undertaken in genomic and metabolomic approaches, have not been realistically conducted to date. If integrative methods such as isotopomics and isotopologies were to be adopted, this would help to understanding connectivity between metabolite pools or drivers of ecosystem dynamics. This would help to elucidate interactions between biogeochemical factors controlling carbon sequestration and respiratory release.

In the present paper, we have used leaf-scale data as an example to explain the principle of the approaches. In a similar manner, we propose that isotopomic approaches could be extended to larger-scale studies, to investigate the architecture of the carbon trafficking networks in soil or ecosystem scales, showing, for example, transfer of carbon from leaves to soil-respired CO₂ (Bowling et al. 2002; Knoll et al. 2005). We nevertheless recognize that this would require a large body of isotopic data associated with several components of ecosystems. However, there is no doubt that this may be overcome in the very near future, thanks to the use of automated continuous-flow spectrometric measurements or tunable diode laser techniques (as CO₂ isotope composition is concerned) (Bowling et al. 2002). Ultimately, we suggest that by mapping the distribution of isotopic signals in biological systems, these approaches will improve our understanding of ecological processes by weighting the contributory physiological processes, and provide a more effective means to visualize their interrelationships.

**REFERENCES**


Received 19 April 2007; received in revised form 27 April 2007; accepted for publication 29 April 2007