OXFORD

## Systems biology

# DExSI: a new tool for the rapid quantitation of $^{13}$C-labelled metabolites detected by GC-MS

## Michael J. Dagley* and Malcolm J. McConville*

Department of Biochemistry and Molecular Biology, Bio21 Institute of Molecular Science and Biotechnology, University of Melbourne, Parkville, VIC 3010, Australia

*To whom correspondence should be addressed.
Associate Editor: Jonathan Wren

## Abstract

**Summary:** Stable isotope directed metabolomics is increasingly being used to measure metabolic fluxes in microbial, plant and animal cells. Incorporation of $^{13}$C/$^{15}$N isotopes into a wide range of metabolites is typically determined using gas chromatography-mass spectrometry (GC/MS) or other hyphenated mass spectrometry approaches. The DExSI (**D**ata **Ex**traction for **S**table **I**sotope-labelled metabolites) pipeline is an interactive graphical software package which can be used to rapidly quantitate isotopologues for a wide variety of metabolites detected by GC/MS. DExSI performs automated metabolite annotation, mass and positional isotopomer abundance determination and natural isotope abundance correction. It provides a range of output options and is suitable for high throughput analyses.

**Availability and implementation:** DExSI is available for non-commercial use from: https://github.com/DExSI/DExSI/. For Microsoft Windows 7 or higher (64-bit).

**Contact:** malcolmm@unimelb.edu.au or michael.dagley@unimelb.edu.au

**Supplementary information:** Supplementary data are available at *Bioinformatics* online.

## 1 Introduction

Analysis of metabolic fluxes is important for understanding the metabolic and physiological state of microbes, plants and animals, as well as for understanding how complex metabolic networks are regulated and respond to different genetic, environmental and pharmacological perturbations. One of the most powerful approaches for measuring metabolic fluxes in intact cells involves the use of stable isotope (i.e. $^{13}$C-, $^{15}$N- or $^{2}$H)-labelled tracers and the measurement of label into different metabolite pools using sensitive hyphenated mass spectrometry platforms, such as gas chromatography-mass spectrometry (GC-MS) (Eylert *et al.*, 2008; Hartel *et al.*, 2012; Kloehn *et al.*, 2015; MacRae *et al.*, 2012; Obata *et al.*, 2013). Quantitation of stable isotope enrichment at different time points or at isotopic steady state can be used to generate local or genome scale metabolite flux maps using metabolic flux analysis (MFA) or instationary MFA.

GC-MS can be used to quantitate the steady-state levels of many intermediates in central carbon metabolism as well as stable isotope enrichment in these metabolites, and is amenable to high throughput analyses through automation of sample derivatization and analytical runs (Buescher *et al.*, 2015). However, there is currently no software that

allows the automated detection of multiple metabolites and quantitation of corresponding mass isotopologues in GC-MS chromatograms with downstream correction for stable isotope abundance (Supplementary Table S1; Perez de Souza *et al.*, 2017). In particular, the available vendor-supplied and open-source software pipelines require substantial manual processing during peak identification (as the mass spectra of unlabelled standards differ from those of heavily labelled metabolites) and/or quantitation of multiple isotopologues of the same metabolite. Corrections for the abundance of naturally occurring isotopes and calculation and visualization of isotopologue abundances are also typically done in separate packages (Buescher *et al.*, 2015).

## 2 Materials and methods

DExSI (**D**ata **Ex**traction for **S**table **I**sotope-labelled metabolites) is a graphical and user-friendly software package for the automated identification annotation and quantitation of stable-isotope labelled metabolites detected by GC-MS. Data can be collected in either SIM or SCAN mode. In brief, DExSI utilizes a user-specified library comprising a list of ions (monoisotopic and related mass isotopologues) to identify a targeted list of metabolites. Using a scoring algorithm which accounts

for retention time, the isotopic series of the metabolite, peak height and optionally, user-specified qualifying ions, the best matching peak for each metabolite is annotated independently of the extent of labelling (Supplementary Table S3). Following automated peak annotation and integration, the graphical user interface (Supplementary Fig. S1) allows for user curation and modification of integration boundaries, or the manual reassignment of the best peak, if required. Prior to exporting data, DExSI applies natural isotope abundance correction calculations to accurately determine the extent of labelling and the distribution of isotopologues using the approach developed by Nanchen *et al.* (2007) and van Winden *et al.* (2002). This workflow is represented graphically in Supplementary Figure S2 using data obtained from the unicellular eukaryotic pathogen, *Leishmania mexicana*, cultivated in the presence of D-glucose-$^{13}C_6$. Here DExSI is used to identify 58 metabolites in GC-MS chromatograms.

## 3 Results

Using DExSI, fractional labelling plots can be generated (Supplementary Fig. S3A) and absolute metabolite abundance can be determined through the calculation of a response factor for the known amount of a metabolite of interest in a reference sample, relative to the integrated area of an internal standard (Supplementary Fig. S3B). Data from the software can be exported in a range of ready-to-present formats: raw data tables or graphs of fractional labelling, isotopologue distribution or absolute abundance can be generated. For both absolute abundance and fractional labelling data (Supplementary Fig. S2B), heat maps can be prepared, and pre-formatted Excel tables can be generated for direct import into the VANTED metabolite pathway mapping and visualization software (Supplementary Fig. S2C) (Junker *et al.*, 2006).

Data processing is performed using established methods for background correction and smoothing. First, GC-MS data is imported from vendor-neutral CDF-formatted files, with the entire data set subjected to both background correction, by applying top-hat background filtering, and noise reduction, using Savitzky–Golay smoothing, using the functions as implemented in the open-source SciPy toolkit (Savitzky and Golay, 1964; van der Walt *et al.*, 2014). This approach to background correction has been previously applied in the PyMS and OpenMS software packages (Lange *et al.*, 2005; O'Callaghan *et al.*, 2012). Peaks are identified as local maxima in each extracted ion chromatogram, and metabolite searches are restricted to peaks which contain a minimum number of co-eluting ions (Biller and Biemann, 1974). Following the above processing of the raw data, metabolite annotation is performed, as described earlier, and the best-matching metabolite peak is integrated.

The background correction, peak smoothing and integration functions of DExSI were compared to Agilent MSD ChemStation using a timecourse series of D-glucose-$^{13}C_6$-labelled samples for 1035 peak areas ranging from $5 \times 10^2$ to $10^8$ units, with least-squares linear regression yielding an $R^2$ value of 0.9999817 (Supplementary Fig. S4). Peak detection accuracy was compared between three biological data sets using a metabolite library of ∼60 compounds, yielding >90% correct identifications (Supplementary Table S2).

## 4 Conclusion

DExSI provides the first integrated, graphical software package for the automated processing and visualization of GC-MS data from stable-isotope labelling experiments (features summarized in Supplementary Table S1). It can be used to identify all isotopologues of multiple metabolites in complex biological mixtures irrespective of variation in the extent of labelling, greatly reducing processing time compared with current tools. It implements a range of calculations commonly used for flux analysis (natural isotope abundance correction and fractional labelling determination) and metabolite quantitation, as well as multiple graphing/heat map options and the ability to export data in a number of formats. DExSI can be used to create a highly automated workflow which will greatly improve the utility and throughput of stable isotope labelling experiments and the implementation of $^{13}C$-MFA.

## References

Biller,J.E. and Biemann,K. (1974) Reconstructed mass spectra, a novel approach for the utilization of gas chromatograph—mass spectrometer data. *Anal. Lett.*, **7**, 515–528.

Buescher,J.M. *et al.* (2015) A roadmap for interpreting (13)C metabolite labeling patterns from cells. *Curr. Opin. Biotechnol.*, **34**, 189–201.

Eylert,E. *et al.* (2008) Carbon metabolism of Listeria monocytogenes growing inside macrophages. *Mol. Microbiol.*, **69**, 1008–1017.

Hartel,T. *et al.* (2012) Characterization of central carbon metabolism of Streptococcus pneumoniae by isotopologue profiling. *J. Biol. Chem.*, **287**, 4260–4274.

Junker,B.H. *et al.* (2006) VANTED: a system for advanced data analysis and visualization in the context of biological networks. *BMC Bioinformatics*, **7**, 109.

Kloehn,J. *et al.* (2015) Characterization of metabolically quiescent Leishmania parasites in murine lesions using heavy water labeling. *PLoS Pathog.*, **11**, e1004683.

Lange,E. *et al.* (2005) OPENMS; a generic open source framework for chromatography/MS-based proteomics. *Mol. Cell Proteomics*, **4**, S25–S25.

MacRae,J.I. *et al.* (2012) Mitochondrial metabolism of glucose and glutamine is required for intracellular growth of Toxoplasma gondii. *Cell Host Microbe*, **12**, 682–692.

Nanchen,A. *et al.* (2007) Determination of metabolic flux ratios from 13C-experiments and gas chromatography-mass spectrometry data: protocol and principles. *Methods Mol. Biol.*, **358**, 177–197.

O'Callaghan,S. *et al.* (2012) PyMS: a Python toolkit for processing of gas chromatography-mass spectrometry (GC-MS) data. Application and comparative study of selected tools. *BMC Bioinformatics*, **13**, 115.

Obata,T. *et al.* (2013) Gas-chromatography mass-spectrometry (GC-MS) based metabolite profiling reveals mannitol as a major storage carbohydrate in the coccolithophorid alga Emiliania huxleyi. *Metabolites*, **3**, 168–184.

Perez de Souza,L. *et al.* (2017) From chromatogram to analyte to metabolite. How to pick horses for courses from the massive web resources for mass spectral plant metabolomics. *Gigascience*, **6**, 1–20.

Saunders,E.C. *et al.* (2015) Use of (13)C stable isotope labelling for pathway and metabolic flux analysis in Leishmania parasites. *Methods Mol. Biol.*, **1201**, 281–296.

Savitzky,A. and Golay,M.J.E. (1964) Smoothing and differentiation of data by simplified least squares procedures. *Anal. Chem.*, **36**, 1627–1639.

van der Walt,S. *et al.* (2014) scikit-image: image processing in Python. *PeerJ*, **2**, e453.

van Winden,W.A. *et al.* (2002) Correcting mass isotopomer distributions for naturally occurring isotopes. *Biotechnol. Bioeng.*, **80**, 477–479.