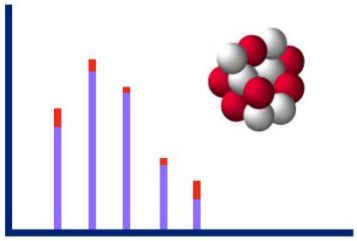


IsoCor

Software Version 1.0

http://metasys.insa-toulouse.fr/software/isocor/

User Manual

Version 1.0

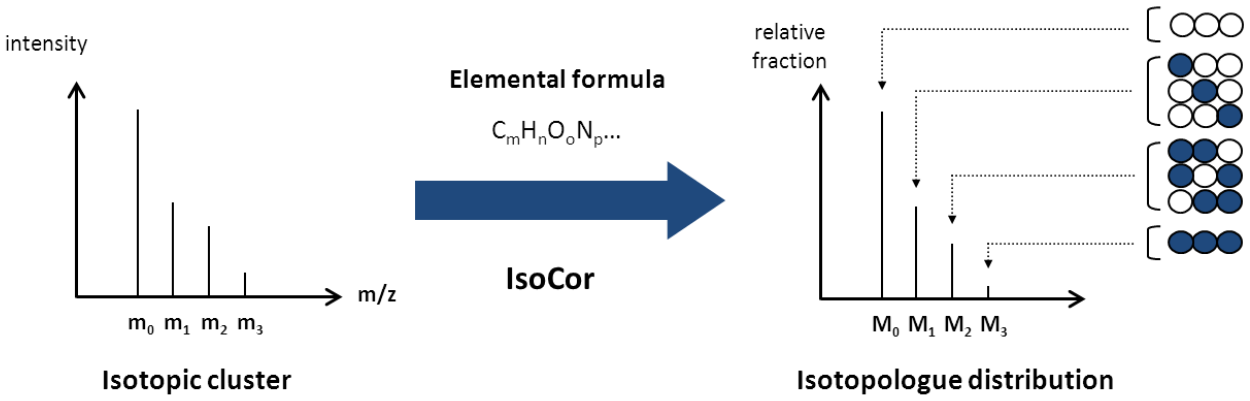
Author: Pierre Millard <pierre.millard@insa-toulouse.fr>

Group MetaSys: Metabolic Systems - UMR5504, UMR792 Ingénierie des Systèmes Biologiques et des Procédés
CNRS, INRA, INSA - 135 Avenue de Rangueil - 31077 TOULOUSE Cedex 4 - FRANCE
Tel/Fax : +33 (0) 561 559 689 - http://metasys.insa-toulouse.fr

Contents

INTRODUCTION

IsoCor is a scientific software designed for the purpose of isotope labeling experiments (ILE). IsoCor correct raw MS data (isotopic clusters) for both all naturally-occurring isotopes and purity of the isotopic tracer. The output of IsoCor is the isotopologue distribution – *i.e*. the relative fractions of molecular entities differing only in the number of isotopic substitutions – of the molecule. It also calculates the mean enrichment – *i.e.* the molecular content in the isotope – in metabolites or their fragments.



For the theory behind these calculations see the following papers:

- Wittmann, C. and Heinzle, E. (1999). Mass spectrometry for metabolic flux analysis, Biotechnol Bioeng, 62, 739-750.
- van Winden, W.A., et al. (2002). Correcting mass isotopomer distributions for naturally occurring isotopes, Biotechnol Bioeng, 80, 477-479.
- Millard, P., et al. (2011). IsoCor: Correcting MS data in isotope labeling experiments. (Submitted)

The main additional value of this program compared to other publicly available software is the generalization of the correction to any stable isotopic tracer from well-known (^{13}C, ^{15}N or ^{18}O) to unusual (^{57}Fe, ^{70}Zn, ^{77}Se…) isotopes.

For more details, see the paper on IsoCor cited before.

LICENSING

The original version of IsoCor software was developed in the MetaSys team in the LISBP, Toulouse, FRANCE.

or implied. See the License for the specific language governing permissions and limitations under the License.

Software and documentation author: Pierre Millard, INRA, France <pierre.millard@insa-toulouse.fr>

REQUIREMENTS AND DEPENDANCIES

IsoCor was developed on Windows but can be used both on Linux (or other UNIX, MacOS included but not yet tested) and Windows platforms.

This software requires Python 2.6+ (not 3.0 or higher) and modules:

- wxPython (v2.8.11.0)
- NumPy (v1.6.0.2)
- SciPy (v0.9.0.1)

These packages are included in pythonxy 2.6.5 (http://www.pythonxy.com), a scientific-oriented Python distribution.

If you are not used to install system wide environments like Python, ask some help from your local computer service. We don’t provide support for installation.

USER'S MANUAL

I. Required information and data files

The information required to perform the correction – *i.e.* the natural abundance of isotopes (Rosman and Taylor, 1998) and a list of metabolites and derivatives which includes their elemental formulas – are provided in flat text files. The structure of these files is provided in section VI. They can be edited and implemented according to the user's needs.

- File 'Isotopes.dat'

This file contains the list of elements for which the isotopic correction can be performed. It contains the list of isotopes to be considered as well as their natural abundance. The element of the isotopic tracer must be declared in this file. All elements not included into this file are supposed to have no isotope - i.e. m_0 = 100%. A file containing the values of the natural abundances of isotopes commonly found in biological compounds (Rosman and Taylor, 1998) is provided with the software.

- File 'Metabolites.dat'

This file contains the list of metabolites, and their elemental formulas, on which the isotopic correction can be applied to. Please refer to the section 'Definition of elemental formula' of the tutorial for more details on the formula. A file containing most common metabolites is provided with the software.

- File 'Derivatives.dat'

This file contains the list of chemical derivatives - with their elemental formulas - that have to be considered for the isotopic correction of metabolites or fragments that have been derivatized prior to MS analysis. CAUTION: ALL elements - including the one of the isotopic tracer - in the elemental formula of the derivative are considered to have their isotope(s) at natural abundance. Please refer to the section 'Definition of elemental formula' of the tutorial for more details on the formula. A file containing some common derivatives is provided with the software.

- Input data files

These files contain the raw MS data (intensities of each peak of the measured isotopic clusters) to be corrected. See examples provided with IsoCor.

II. Graphic user interface

Initially, IsoCor appears with the window shown in figure 1.

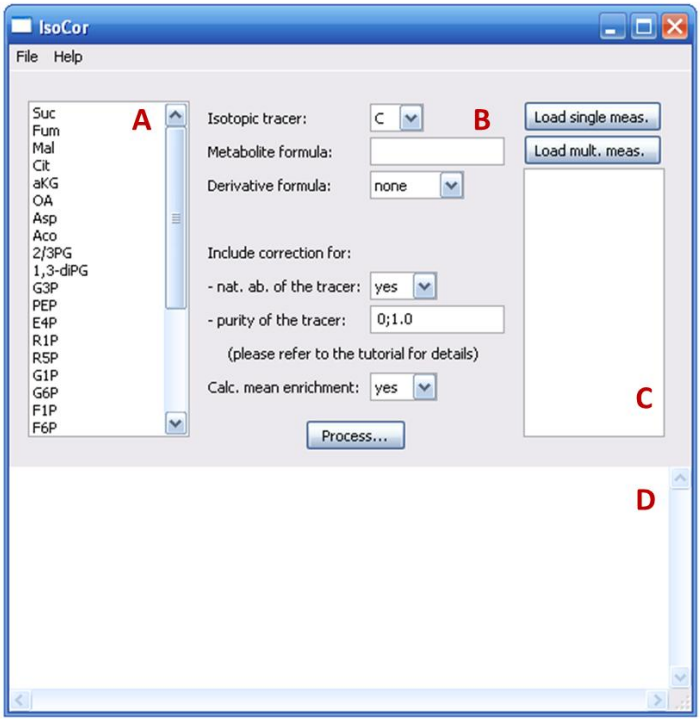


Figure 1. Window of IsoCor. (**A**) List of metabolites. (**B**) Correction parameters panel. (**C**) Selection of experimental data files. (**D**) Calculation results.

III. Performing correction

The correction can be performed for a single metabolite or for multiple metabolites (taken from a unique biological sample or from different samples) at the same time. Input files are plain text files for which the syntax rules are detailed in section VI. For each isotopic cluster to correct, the minimal number of measured intensities – *i.e.* the length of the measured isotopic cluster – must be at least equal to the mass shift of the fully-labelled molecule, without missing value.

<u>For a single metabolite</u>

- Load a data set using the 'Load single meas.' Button.
- Select a metabolite from the list or enter a formula.
- If relevant, select a derivative or enter a new derivative in field 'Derivative formula'.

Note: All elements from the derivative are supposed to be at natural abundance.

- Select options and parameters (detailed in section IV).
- Click on the 'Process' button.
- The output of the calculation is given in the box at the bottom of the IsoCor window (figure 2).

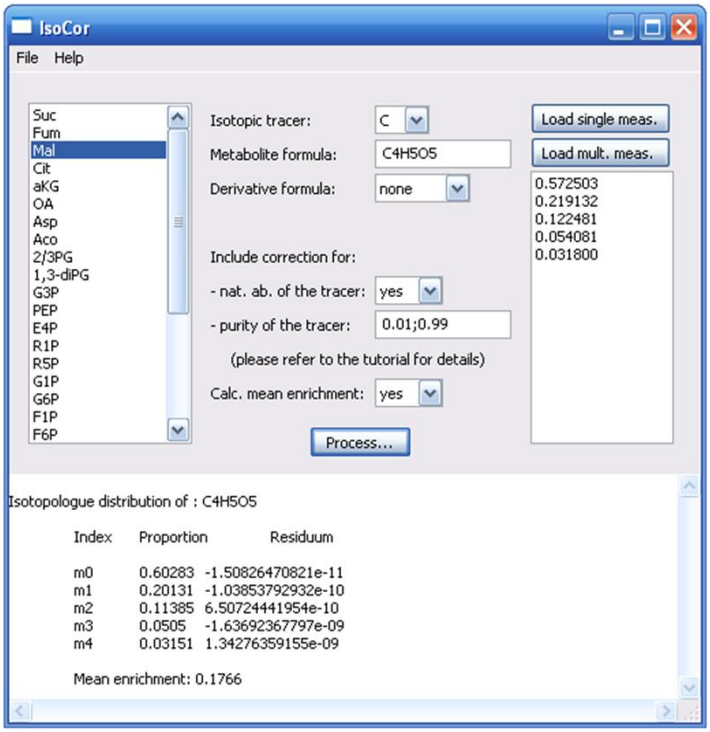


Figure 2. View of IsoCor window after correction of a single metabolite.

For a list of metabolites

- Load a data set 'InputFile.txt' using the 'Load multiple meas.' button.
- Select options and parameters (detailed in section IV).
- Click on the 'Process' button.

IsoCor proceeds automatically to the correction of all the sets of data that are included into the data files, provided the metabolites, derivatives, formulas and MS data entered using appropriate formats (see section VI).

Note: All elements from the derivative are supposed to be at natural abundance.

- The output of the calculations is written in a text file ('InputFile_res.txt'). The structure of tis file is detailed in section V.
- A report file is generated and given in the box at the bottom of the IsoCor window (figure 3). It is also written in a text file ('InputFile_log.txt').

Note: All these files are silently overwritten if already exist. So take care to copy your results elsewhere if you want to protect them from overwriting.

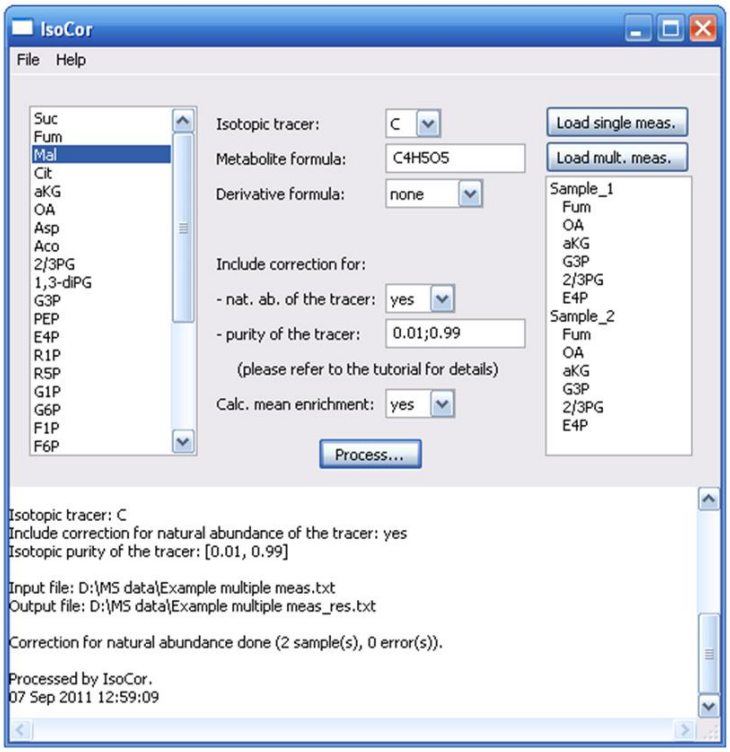


Figure 3. View of IsoCor window after correction of multiple metabolites.

IV. Parameters and options

Here after the available parameters and options are enumerated and detailed.

1. Isotopic tracer

User must declare the element used as isotopic tracer in the ILE.

2. Definition of elemental formulas

The elemental formulas used for correction have to be defined carefully otherwise the correction will be wrong.

- The formula provided in the field 'Metabolite formula' or in the data files must correspond to the formula of the ion that is actually measured by MS, i.e. the ion that gives rise to the measured isotopic cluster. Example: If pyruvic acid ($C_3H_4O_3$) is analyzed by LC-MS using multiple ion monitoring (MIM) in the negative mode, and the detected ion is the $[M\text{-}H]^-$ ion, then the formula to introduce in IsoCor is $C_3H_3O_3$.
- If a derivatization step is required before MS analysis, the elemental composition of the derivative (often common to several metabolites) must be declared, as it is in the MS-detected ion. If the derivative moiety of the MS-detected ion contains atoms corresponding to the isotopic tracer, these atoms are supposed to be at natural abundance for the isotopic tracer. Therefore the derivative moiety, but not always that of the metabolite moiety, must be corrected for natural abundance. To allow the correction of all elements of the derivative moiety

of the MS-detected ion, the formula of the derivative moiety must declared separately from that of the metabolite moiety. Example: alanine ($C_3H_7O_2N$) can be analyzed by GC-MS after t-butyl-dimethyl-silylation (TBDMS derivatization). A fragment that is classically used for isotopomer analysis is the 'M-57' fragment that contains the compound of interest and two TBDMS groups, one of which lost the fragment [C_4H_9]. The elemental composition of the two TBDMS groups excluding that of the fragment - *i.e.* [$Si_2C_8H_{21}$] - must be declared in the 'Derivative formula' field, meanwhile the elemental composition of the alanine moiety of the detected ion [$C_3H_5O_2N$] must be declared in the 'Metabolite formula' field.

IsoCor can be used to correct both low and high resolution MS measurements.

IsoCor includes a pre-established list of metabolites and derivatives - and their elemental formulas - in the 'Metabolites.dat' and 'Derivatives.dat' files, respectively. These files can be edited and implemented according to the user's needs.

3. Correction for the tracer

The options to deal with the tracer are detailed hereafter. These correction parameters have to be defined carefully.

3.1 Correction for natural abundance

Partially-labeled and unlabelled substrates usually contain a fraction of labeled atoms (at natural abundance) into their unlabelled positions. To perform the correction the option 'nat. ab. of the tracer' must be enabled.

CAUTION: This correction is valid only when the isotopes of the tracer element occurs at natural abundance into the unlabeled positions of the input substrate(s), which is in most cases assumed but has to be checked from the manufacturers or determined experimentally.

3.2 Correction for isotopic purity

Labeled substrates usually contain into the labeled positions a fraction of unlabeled atoms for which MS data must be corrected. The fractions of each isotope into the labeled positions must be declared ordered by mass shifts and separated by a semicolon ';'. As example, if the content in ^{13}C atoms into each position of a U-^{13}C-labeled compound is 99%, other 1% being ^{12}C atoms, the purity must be entered as: '0.01;0.99'. If you do not want to correct isotopic clusters for the isotopic purity of the substrate, just let the default value (purity = 100%) in the 'Purity' field.

CAUTION: This correction is valid only when the distribution of the isotopes is binomial into the labeled positions of both the input substrate and the metabolites – *i.e.* there is no isotopic fractionation effects of any kind –, which is in most cases assumed but has to be checked from the manufacturers or determined experimentally. When different labeled substrates are mixed, this correction also requires that all their labeled positions have the same isotopic purity.

4. Calculation of mean enrichment

The mean isotopic enrichment of a metabolite refers to the mean content in isotopic tracer in the metabolite, expressed as the relative fraction of total atoms of its element in the metabolite. This information is particularly useful for the quantification of split ratios between two metabolic pathways resulting in different content of tracer. If the option 'Calc. mean enrichment' is enabled, then IsoCor calculates the mean enrichment using the following formula:

$$ME = \frac{\sum_{i=0}^{n} M_i . i}{n}$$

where *ME* is the mean enrichment of the molecule, M_i is the relative proportion of the isotopologue containing *i* atoms of isotopic tracer and *n* denotes the number of atoms of tracer element in the molecule.

V. Result files

- When multiple metabolites are processed, their corrected isotopic clusters are written in the result file 'InputFile_res.txt', a plain text files with tabulation as separator. The content of each column is detailed in table 1.

Column	*Name*	*Description*
1	Sample	Name of the sample
2	Metabolite	Name of the metabolite
3	Peak index	This value corresponds to the mass shift of the measured peak for the column 'Isotopic cluster' and to the number of isotopic tracer in the molecule for the column 'Isotopologue distribution'
4	Isotopic cluster	Measured isotopic cluster
5	Isotopologue distribution	Corrected isotopic cluster (normalized to 1)
6	Residuum	Difference between measured and simulated isotopic cluster (value relative to the sum of measured intensities)
7	Mean enrichment	Mean content in isotopic tracer in the metabolite (optional)

Table 1. Content of the file 'InputFile_res.txt'.

- The report given in the box at the bottom of the IsoCor window is also written in the file 'InputFile_log.txt'. It contains information such as parameters of correction (isotopic tracer, purity…), paths of input and output files, and errors occurred during the correction process.

VI. Structure of data files

All data files are plain text files with tabulation as separator.

'Isotopes.dat' : contains the elements and their natural abundance, given as relative fraction

Element_1	proportion_m0	proportion_m1	
Element_2	proportion_m0	proportion_m1	proportion_m2
...			

'Metabolites.dat' and 'Derivatives.dat' : contain lists of metabolites and derivatives with their elemental formulas

Suc	C4H5O4
Fum	C4H3O4
G6P	C6H10O5
...	

Input data file : file structure for correction of a single isotopic cluster

Intensity_M0
Intensity_M1
Intensity_M2
Intensity_M3
...

Input data file : file structure for correction of multiple isotopic clusters (different metabolites/samples)

Sample 1	Metabolite 1		Intensity_M0
			Intensity_M1
			Intensity_M2
	Metabolite 2	Derivative 1	Intensity_M0
			Intensity_M1
			Intensity_M2
Sample 2	Metabolite 3	Derivative 1	Intensity_M0
			Intensity_M1
			Intensity_M2
			Intensity_M3
...			

VII. Warning and error messages

Error messages are explicit. The user should examine carefully any warning/error message. After correcting the problem, the user may have to rerun IsoCor (to reload data files 'Isotopes.dat', 'Metabolites.dat' and 'Derivatives.dat' if required) and perform correction.

VIII. Validation of the correction process

For validation purpose, we created several sets of simulated isotopologues with varying numbers of tracer atoms (from 2 to 100) and levels of labeling. Test sets were obtained by adding the expected contribution of other naturally abundant elements (calculated with Molecular Weight Calculator, http://www.alchemistmatt.com/mwtwin.html) to these isotopologues. The correction was performed both on these theoretical test sets and on the same sets with noise added. The implementation appears quite robust and very accurate in both cases.

The tests sets and the results of the correction process are in the file 'Validation\Validation.xlsx'.

LICENSE FOR ISOCOR SOFTWARE

Educational Community License, Version 2.0 [OSI Approved License]

Educational Community License Version 2.0, April 2007

The Educational Community License version 2.0 ("ECL") consists of the Apache 2.0 license, modified to change the scope of the patent grant in section 3 to be specific to the needs of the education communities using this license. The original Apache 2.0 license can be found at: http://www.apache.org/licenses/LICENSE-2.0

TERMS AND CONDITIONS FOR USE, REPRODUCTION, AND DISTRIBUTION

1. Definitions.

"License" shall mean the terms and conditions for use, reproduction, and distribution as defined by Sections 1 through 9 of this document.

"Licensor" shall mean the copyright owner or entity authorized by the copyright owner that is granting the License.

"Legal Entity" shall mean the union of the acting entity and all other entities that control, are controlled by, or are under common control with that entity. For the purposes of this definition, "control" means (i) the power, direct or indirect, to cause the direction or management of such entity, whether by contract or otherwise, or (ii) ownership of fifty percent (50%) or more of the outstanding shares, or (iii) beneficial ownership of such entity

"You" (or "Your") shall mean an individual or Legal Entity exercising permissions granted by this License.

"Source" form shall mean the preferred form for making modifications, including but not limited to software source code, documentation source, and configuration files.

"Object" form shall mean any form resulting from mechanical transformation or translation of a Source form, including but not limited to compiled object code, generated documentation, and conversions to other media types.

"Work" shall mean the work of authorship, whether in Source or Object form, made available under the License, as indicated by a copyright notice that is included in or attached to the work (an example is provided in the Appendix below).

"Derivative Works" shall mean any work, whether in Source or Object form, that is based on (or derived from) the Work and for which the editorial revisions, annotations, elaborations, or other modifications represent, as a whole, an original work of authorship. For the purposes of this License, Derivative Works shall not include works that remain separable from, or merely link (or bind by name) to the interfaces of, the Work and Derivative Works thereof.

"Contribution" shall mean any work of authorship, including the original version of the Work and any modifications or additions to that Work or Derivative Works thereof, that is intentionally submitted to Licensor for inclusion in the Work by the copyright owner or by an individual or Legal Entity authorized to submit on behalf of the copyright owner. For the purposes of this definition, "submitted" means any form of electronic, verbal, or written communication sent to the Licensor or its representatives, including but not limited to communication on electronic mailing lists, source code control systems, and issue tracking systems that are managed by, or on behalf of, the Licensor for the purpose of discussing and improving the Work, but excluding communication that is conspicuously marked or otherwise designated in writing by the copyright owner as "Not a Contribution."

"Contributor" shall mean Licensor and any individual or Legal Entity on behalf of whom a Contribution has been received by Licensor and subsequently incorporated within the Work.

2. Grant of Copyright License.

Subject to the terms and conditions of this License, each Contributor hereby grants to You a perpetual, worldwide, non-exclusive, no-charge, royalty-free, irrevocable copyright license to reproduce, prepare Derivative Works of, publicly display, publicly perform, sublicense, and distribute the Work and such Derivative Works in Source or Object form.

3. Grant of Patent License.

Subject to the terms and conditions of this License, each Contributor hereby grants to You a perpetual, worldwide, non-exclusive, no-charge, royalty-free, irrevocable (except as stated in this section) patent license to make, have made, use, offer to sell, sell, import, and otherwise transfer the Work, where such license applies only to those patent claims licensable by such Contributor that are necessarily infringed by their Contribution(s) alone or by combination of their Contribution(s) with the Work to which such Contribution(s) was submitted. If You institute patent litigation against any entity (including a cross-claim or counterclaim in a lawsuit) alleging that the Work or a Contribution incorporated within the Work constitutes direct or contributory patent infringement, then any patent licenses granted to You under this License for that Work shall terminate as of the date such litigation is filed. Any patent license granted hereby with respect to contributions by an individual employed by an institution or organization is limited to patent claims where the individual that is the author of the Work is also the inventor of the patent claims licensed, and where the organization or institution has the right to grant such license under applicable grant and research funding agreements. No other express or implied licenses are granted.

4. Redistribution.

You may reproduce and distribute copies of the Work or Derivative Works thereof in any medium, with or without modifications, and in Source or Object form, provided that You meet the following conditions:

You must give any other recipients of the Work or Derivative Works a copy of this License; and

You must cause any modified files to carry prominent notices stating that You changed the files; and

You must retain, in the Source form of any Derivative Works that You distribute, all copyright, patent, trademark, and attribution notices from the Source form of the Work, excluding those notices that do not pertain to any part of the Derivative Works; and

If the Work includes a "NOTICE" text file as part of its distribution, then any Derivative Works that You distribute must include a readable copy of the attribution notices contained within such NOTICE file, excluding those notices that do not pertain to any part of the Derivative Works, in at least one of the following places: within a NOTICE text file distributed as part of the Derivative Works; within the Source form or documentation, if provided along with the Derivative Works; or, within a display generated by the Derivative Works, if and wherever such third-party notices normally appear. The contents of the NOTICE file are for informational purposes only and do not modify the License. You may add Your own attribution notices within Derivative Works that You distribute, alongside or as an addendum to the NOTICE text from the Work, provided that such additional attribution notices cannot be construed as modifying the License.

You may add Your own copyright statement to Your modifications and may provide additional or different license terms and conditions for use, reproduction, or distribution of Your modifications, or for any such Derivative Works as a whole, provided Your use, reproduction, and distribution of the Work otherwise complies with the conditions stated in this License.

5. Submission of Contributions. Unless You explicitly state otherwise, any Contribution intentionally submitted for inclusion in the Work by You to the Licensor shall be under the terms and conditions of this License, without any additional terms or conditions. Notwithstanding the above, nothing herein shall supersede or modify the terms of any separate license agreement you may have executed with Licensor regarding such Contributions.

6. Trademarks. This License does not grant permission to use the trade names, trademarks, service marks, or product names of the Licensor, except as required for reasonable and customary use in describing the origin of the Work and reproducing the content of the NOTICE file.

7. Disclaimer of Warranty. Unless required by applicable law or agreed to in writing, Licensor provides the Work (and each Contributor provides its Contributions) on an "AS IS" BASIS, WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied, including, without limitation, any warranties or conditions of TITLE, NON-INFRINGEMENT, MERCHANTABILITY, or FITNESS FOR A PARTICULAR PURPOSE. You are solely responsible for determining the appropriateness of using or redistributing the Work and assume any risks associated with Your exercise of permissions under this License.

8. Limitation of Liability. In no event and under no legal theory, whether in tort (including negligence), contract, or otherwise, unless required by applicable law (such as deliberate and grossly negligent acts) or agreed to in writing, shall any Contributor be liable to You for damages, including any direct, indirect, special, incidental, or consequential damages of any character arising as a result of this License or out of the use or inability to use the Work (including but not limited to damages for loss of goodwill, work stoppage, computer failure or malfunction, or any and all other commercial damages or losses), even if such Contributor has been advised of the possibility of such damages.

9. Accepting Warranty or Additional Liability. While redistributing the Work of Derivative Works thereof, You may choose to offer, and charge a fee for, acceptance of support, warranty, indemnity, or other liability obligations and/or rights consistent with this License. However, in accepting such obligations, You may act only on Your own behalf and on Your sole responsibility, not on behalf of any other Contributor, and only if You agree to indemnify, defend, and hold each Contributor harmless for any liability incurred by, or claims asserted against, such Contributor by reason of your accepting any such warranty or additional liability.

END OF TERMS AND CONDITIONS

APPENDIX: How to apply the Educational Community License to your work

To apply the Educational Community License to your work, attach the following boilerplate notice, with the fields enclosed by brackets "[]" replaced with your own identifying information. (Don't include the brackets!) The text should be enclosed in the appropriate comment syntax for the file format. We also recommend that a file or class name and description of purpose be included on the same "printed page" as the copyright notice for easier identification within third-party archives.

Copyright [yyyy] [name of copyright owner] Licensed under the Educational Community License, Version 2.0 (the "License"); you may not use this file except in compliance with the License. You may obtain a copy of the License at

http://www.osedu.org/licenses/ECL-2.0

Unless required by applicable law or agreed to in writing, software distributed under the License is distributed on an "AS IS" BASIS, WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied. See the License for the specific language governing permissions and limitations under the License.

IsoCor (hereafter "the program") is a scientific software designed to correct MS data for natural abundance.

All its parts are freely distributed according to GPL2 terms http://www.gnu.org/licenses/gpl-2.0.txt (permission to execute, to copy, to modify and to redistribute; obligation to preserve the original copyright notice and to keep your modifications OpenSource only if you redistribute them alone or as part of IsoCor)

WARRANTY

BECAUSE THE PROGRAM IS LICENSED FREE OF CHARGE, THERE IS NO WARRANTY FOR THE PROGRAM, TO THE EXTENT PERMITTED BY APPLICABLE LAW. EXCEPT WHEN OTHERWISE STATED IN WRITING THE COPYRIGHT HOLDERS AND/OR OTHER PARTIES PROVIDE THE PROGRAM "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESSED OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE. THE ENTIRE RISK AS TO THE QUALITY AND PERFORMANCE OF THE PROGRAM IS WITH YOU. SHOULD THE PROGRAM PROVE DEFECTIVE, YOU ASSUME THE COST OF ALL NECESSARY SERVICING, REPAIR OR CORRECTION.

COPYRIGHT 2011, INRA