GeoRge: Generalizable statistically-oriented R-workflow for isotopic label-tracking experiments

Jordi Capellades[1,3], Miriam Navarro[2,3], Sara Samino[2], Maria Vinaixa[2,3], Oscar Yanes[2,3]

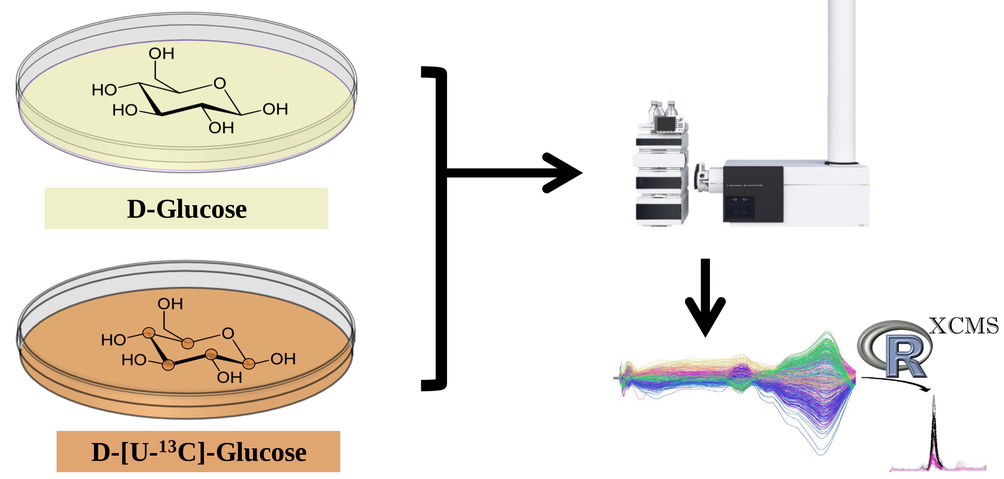
1. Institut d'Investigació Sanitària Pere Virgili (IISPV), Tarragona, Spain; 2. Spanish Biomedical Research Center in Diabetes and Associated Metabolic Disorders (CIBERDEM), Madrid, Spain; 3. Universitat Rovira i Virgili, Tarragona, Spain.

INTRODUCTION

Metabolite concentration is deeply regulated through homeostasis, whose fluctuations are modified by lower level processes, such as gene regulation or protein activity. Stable isotope-labelling metabolomics studies allow an unbiased mapping of fluxes through multiple metabolic pathways. Though, there are no computational solutions to deal with this type of data. We propose an intuitive workflow, written in the open language R, for stable isotope-labelling non-targeted metabolomics experiments, with the aim to calculate the percentage of labelling in identified metabolites.

MATERIALS

ARPE-19 cells were cultured under standard conditions, serum deprived media were prepared with 5.5 mM or 25 mM of either D-Glucose or D-[U-^{13}C]-Glucose, and cultured in normoxic or hypoxic conditions. Each condition was run in triplicate.

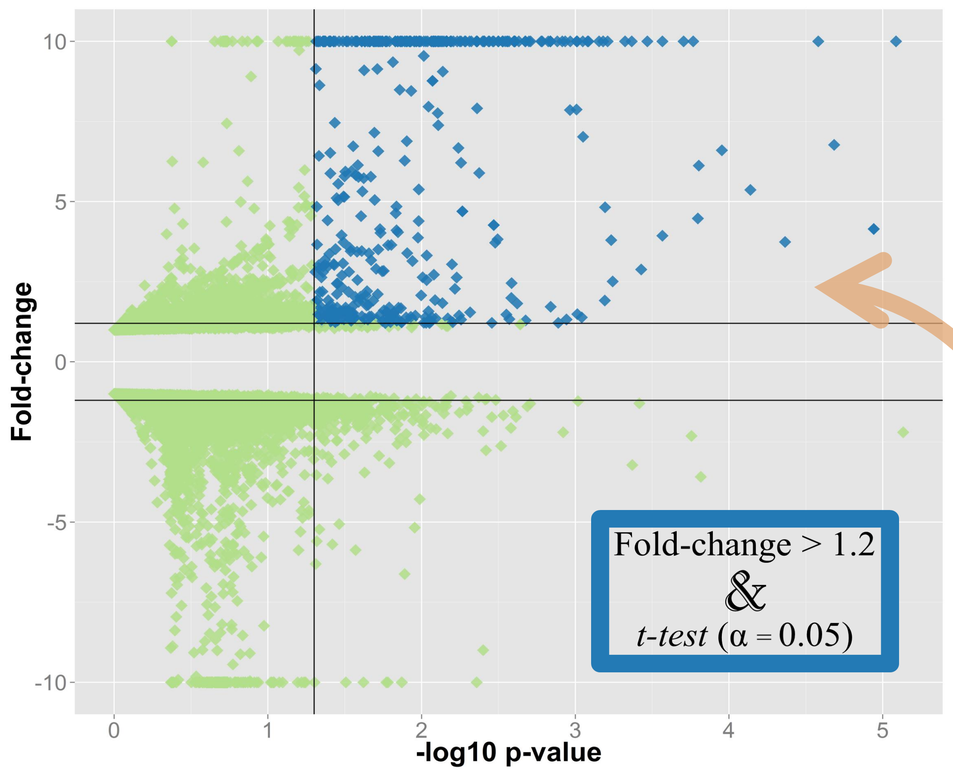


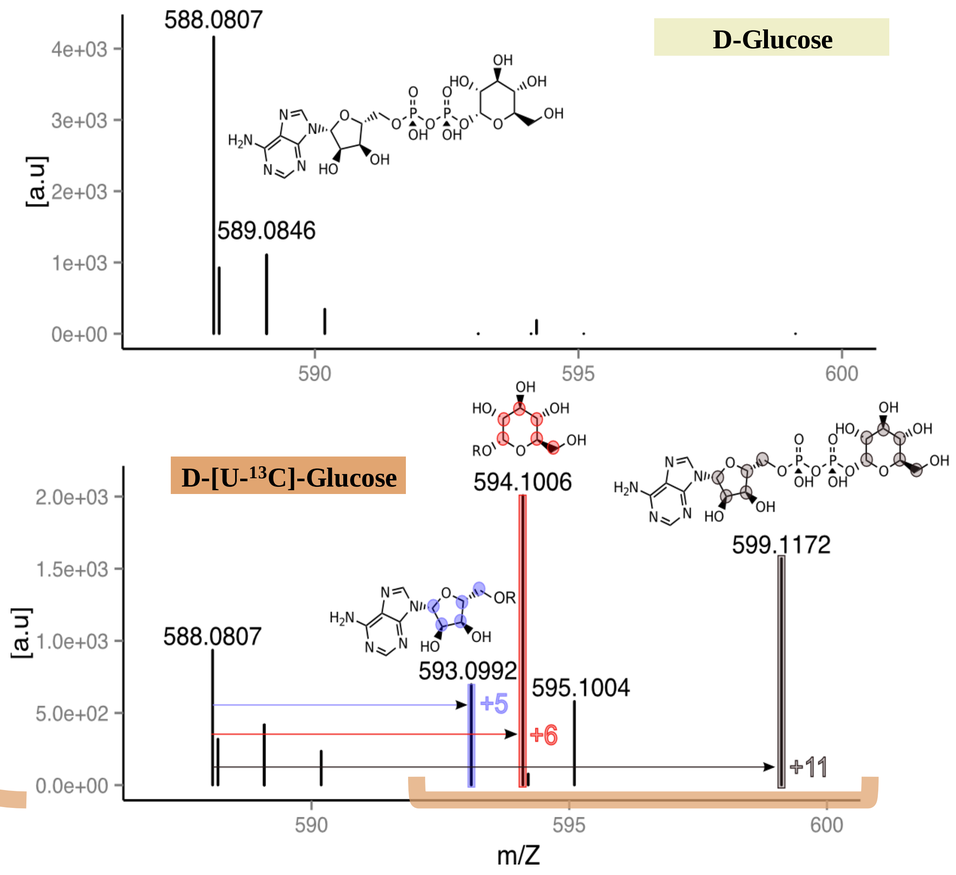
PRINCIPLE

Find those peaks that are not present in unlabeled samples, consistently along the replicates.

- t-test (Compare the distribution of intensities)
- Fold-change (Ratio of the mean peak intensity) labeled / unlabeled samples

These new peaks are masses which are originated by different transformations of the labeled precursor metabolite.



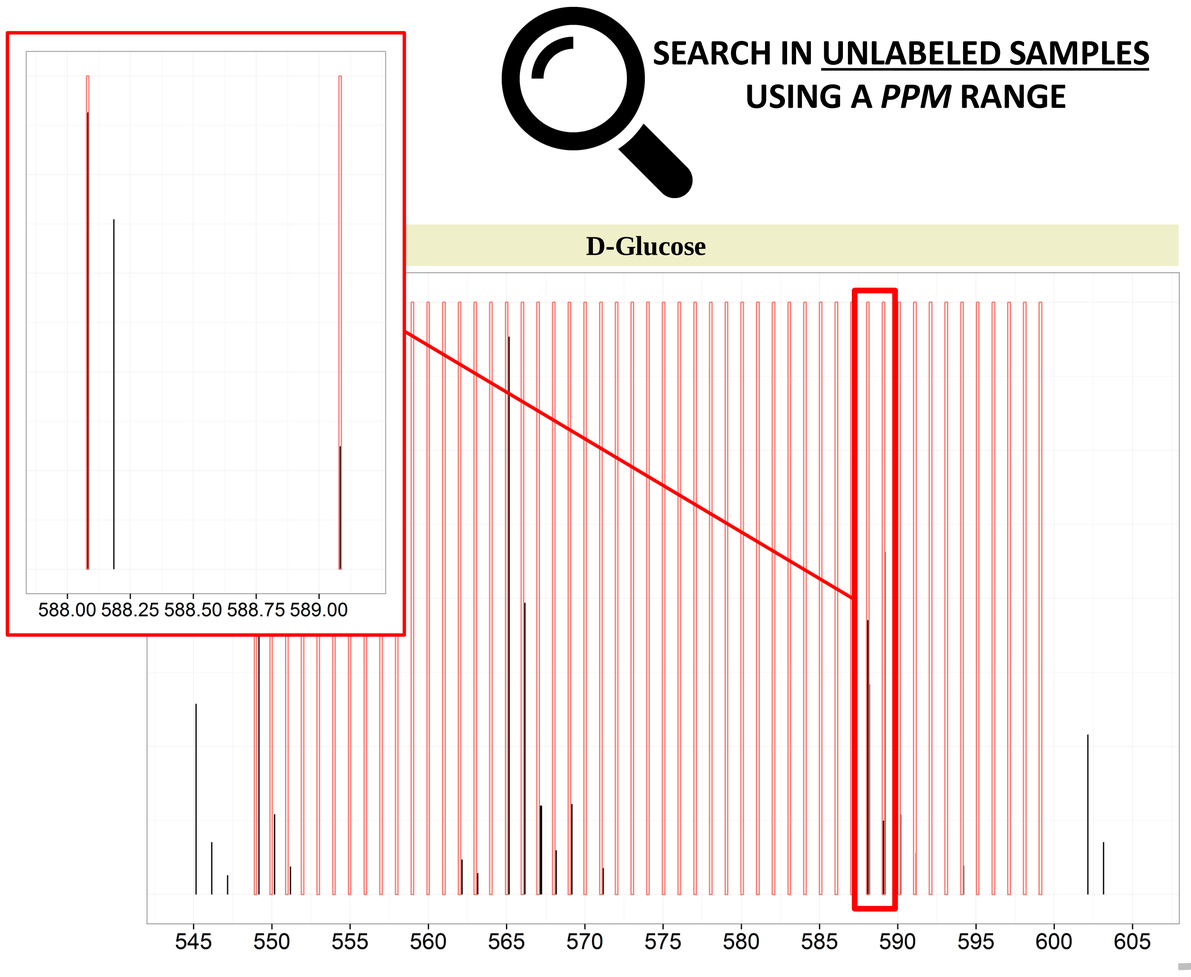


599.1172 / 12.00 ≈ 50

↓

[599.1172, 598.1135,…, 594.1005,…, **588.0804**,…,548.9497]

SEARCH IN <u>UNLABELED SAMPLES</u> USING A *PPM* RANGE



BASE PEAK SEARCH

The masses passing the fold-change and p-value filters are associated with metabolites incorporating a labeled (^{13}C) atom.

$$\forall\, x:$$
$$y = [M + (0, \dots, n)] = \left[x - \left((0, \dots, n) \times k\right)\right]$$
$$where:$$
$$n = \left\lceil \frac{x}{12.00000} \right\rceil$$
$$k = 13.00335 - 12.00000$$

The mass of the new peaks (x) is divided by the mass of the unlabeled atom (in carbon experiments is 12.00) in order to calculate the maximum number of carbon atoms the given mass could have.

This is used to obtain a vector of masses (y), that are searched in the unlabeled samples in order to find the base peak.

If an element of y matches with a peak of the unlabeled samples, the mean intensity of the match needs to have a minimum intensity threshold to be considered a base peak. If more than one match passes the threshold, they are saved as different base peaks of the same y element.

BASE PEAK IDENTIFICATION

The masses found as base peaks are matched against the Human Metabolome Database (HMDB).

PERCENTAGE CALCULATION

The mean intensity of each isotopologue is divided by the sum of all mean intensities.