Journal of Chromatography A, 1374 (2014) 199–206

Contents lists available at ScienceDirect

Journal of Chromatography A

journal homepage: www.elsevier.com/locate/chroma

Comparative evaluation of eight software programs for alignment of gas chromatography–mass spectrometry chromatograms in metabolomics experiments

Weihuan Niu[a], Elisa Knight[b,1], Qingyou Xia[c], Brian D. McGarvey[b,*]

[a] School of Life Science, Chongqing University, Chongqing 400044, China
[b] Southern Crop Protection and Food Research Centre, Agriculture and Agri-Food Canada, 1391 Sandford Street, London, Ontario N5V 4T3, Canada
[c] State Key Laboratory of Silkworm Genome Biology, Southwest University, Chongqing 400715, China

ARTICLE INFO

Article history:
Received 4 January 2014
Received in revised form
26 September 2014
Accepted 4 November 2014
Available online 20 November 2014

Keywords:
Retention time alignment
Software comparison
Gas chromatography–mass spectrometry (GC–MS)
Metabolomics
Principal components analysis (PCA)

ABSTRACT

Since retention times of compounds in GC–MS chromatograms always vary slightly from chromatogram to chromatogram, it is necessary to align chromatograms before comparing them in metabolomics experiments. Several software programs have been developed to automate this process. Here we report a comparative evaluation of the performance of eight programs using prepared samples of mixtures of chemicals, and an extract of tomato vines spiked with three concentrations of a mixture of alkanes. The programs included in the comparison were SpectConnect, MetaboliteDetector 2.01a, MetAlign 041012, MZmine 2.0, TagFinder 04, XCMS Online 1.21.01, MeltDB and GAVIN. Samples were analyzed by GC–MS, chromatograms were aligned using the selected programs, and the resulting data matrices were pre-processed and submitted to principal components analysis. In the first trial, SpectConnect, MetAlign and MetaboliteDetector correctly identified ≥90% of the true positives. In the second trial, MetAlign and MetaboliteDetector correctly identified 87% and 81% of the true positives, respectively. In addition, in both trials >90% of the peaks identified by MetAlign and MetaboliteDetector were true positives.

1. Introduction

Metabolomics is an unbiased approach used to detect and quantify the low molecular weight metabolites in a biological sample [1–3]. Along with genomics, transcriptomics and proteomics, metabolomics is becoming a key technology to measure the ultimate phenotype of an organism [4]. In particular, metabolomics has the advantage of using undirected global screening approaches based on the measurement of signal intensities [5–8].

Metabolomics is complicated by the huge diversity of metabolites present in an organism. The plant kingdom alone has an estimated 100,000–200,000 metabolites and the human metabolome contains approximately 1500 metabolites [9]. It has been suggested that a comprehensive plant metabolic profile should include metabolites from multiple primary and secondary metabolic pathways including carbohydrates, amino acids, organic acids, lipids/fatty acids, vitamins and various other compound classes such as phenylpropanoids, terpenoids, alkaloids, and glucosinolates, depending on the species under study [10].

A number of analytical tools have been used to analyze the highly complex mixtures present in metabolomics experiments [4–8,11]. These include gas chromatography–mass spectrometry (GC–MS) [12], liquid chromatography–mass spectrometry (LC–MS) [13], capillary electrophoresis–mass spectrometry (CE–MS) [14], Fourier transform infrared spectroscopy (FT-IR) and nuclear magnetic resonance spectroscopy (NMR) [8,15]. Advantages of GC–MS include high resolution chromatographic separation and wide applicability through derivatization. It is also a stable, sensitive and rugged platform and offers the ability to identify metabolites through comparing experimental retention times/indices and electron-impact mass spectra with those in commercial or in-house libraries. To date, GC–MS is the most developed analytical platform for plant metabolite profiling [16] and has been called the gold standard for metabolic profiling [17].

Metabolomics is by design non-targeted, and is expected to be responsive to both expected and unexpected differences between

* Corresponding author. Current affiliation: College of Chemistry and Environment Protection Engineering, Southwest University for Nationalities, Chengdu 610041, China. Tel.: +86 13 678148230; fax: +86 28 85524382.
E-mail address: bmcgarve@uwo.ca (B.D. McGarvey).
[1] Current affiliation: Gamma-Dynacare Medical Laboratories, 245 Pall Mall Street, London, Ontario N6A 1P4, Canada.

http://dx.doi.org/10.1016/j.chroma.2014.11.005
0021-9673/

groups of samples. This means that all metabolites are potentially of interest, since it is not possible to know in advance which metabolites will be affected by a treatment or different environmental conditions, etc. When comparing groups of chromatograms, therefore, it is essential that all observed metabolites are included in the comparison. One challenge for GC–MS, as for any chromatographic technique, is that retention times for the same compound always vary slightly from chromatogram to chromatogram. Alignment of chromatograms is therefore an essential step in data processing to ensure that the same compound is being compared in each chromatogram. Since chromatograms in metabolomics experiments may include hundreds or even thousands of peaks, the alignment process must be automated; to confirm that the same compound is being compared for each peak among hundreds of peaks in several or many chromatograms is far too labor-intensive to be done manually.

Several software programs have been released to automate the alignment of GC–MS or LC–MS chromatograms, and some have been compared in earlier studies. An LC–MS study by Lange et al. [18] used two LC–MS proteomics and two LC–MS metabolomics data sets to evaluate six freely available alignment programs, msInspect, MZmine, OpenMS, XCMS, SpecArray and XAlign. They evaluated the performances based on alignment recall and alignment precision, as well as running times. For the proteomics data sets, OpenMS performed best, followed by XAlign, XCMS and MZmine, while for the metabolomics data sets, XCMS performed best, closely followed by MZmine, with OpenMS and XAlign not far behind. In another study, Peters et al. [19] demonstrated a new strategy to objectively determine the quality of an alignment result for GC–MS chromatograms by using several parameters to determine the outcome of the alignment process and select the optimum set of parameters for three software packages, namely MarkerlynxTM, MZmine and MetAlign. Another study by Koh et al. [20] evaluated three freely available programs to analyze GC–MS chromatograms, MetAlign, MZmine, TagFinder, as well as the commercially available Calibration and Statistical Compare features of ChromaTOF. They chose not to compare MET-IDEA, MetaboliteDetector, MetaboAnalyst, MeltDB, ChromA, XCMS and AnalyzerPro because of difficulties in using them to process the GC–MS data. For their comparison they utilized two GC–MS data sets, one prepared using mixtures of standard metabolites and the other consisting of human bladder cancer urine samples. Rather than rely on random variations of retention time between chromatograms, they induced retention time shifts (2 s and 4 s positive and negative) by altering the programmed temperature gradient of the GC. For the first data set, the Calibration feature of ChromaTOF and MZmine gave the best performance, while for the second set, the Calibration and Statistical Compare features of ChromaTOF and MetAlign gave the best performance.

In our own search for tools to automate the analysis of many GC–MS chromatograms in metabolomics experiments, we decided to compare the performance of eight programs, including five not evaluated by Koh et al. The programs we selected for the comparison were SpectConnect [21], MetaboliteDetector 2.01a [22], MetAlign 041012 [23,24], MZmine 2.0 [25], TagFinder 04 [26,27], XCMS Online 1.21.01 [28,29], MeltDB [30–32] and GAVIN [33]. All were downloaded freely for local installation or were web-based applications, and were capable of analyzing GC–MS data in an untargeted manner.

In addition to the problem of retention time variability, GC–MS, even though it is a very high resolution chromatographic technique, sometimes suffers from the problem of co-elution of compounds. One of the advantages of GC–MS is that it is possible to deconvolute overlapping peaks. The program Automated Mass spectral Deconvolution and Identification System (AMDIS) [34], freely available from the National Institute of Standards and Technology (NIST), is often used by chromatographers to deconvolute overlapping peaks.

The alignment programs employ different means of dealing with overlapping peaks. Some, including MetaboliteDetector, MeltDB and MZmine, include an algorithm in the program for deconvolution. Others, including MetAlign and XCMS Online circumvent the need for deconvolution by analyzing individual mass chromatograms. Still others, including SpectConnect and GAVIN require that the chromatograms be deconvoluted using AMDIS or ChromaTOF prior to alignment. TagFinder can handle externally deconvoluted chromatograms, but can also deal with baseline-corrected chromatograms.

Our approach to the evaluation and comparison of the selected programs differed from that of Koh et al. [20]. In order to determine which program would work best in our experiments, we prepared two sets of designed samples. The first was prepared by combining commercially available mixtures of chemicals in varying proportions. The second was prepared by spiking an extract of tomato vines with three different concentrations of a commercial mixture of alkanes. All samples were analyzed by GC–MS. Rather than induce retention time differences, as Koh et al. did, we felt it would be more relevant to simply challenge the alignment software programs with chromatograms whose retention times varied randomly in the normal fashion. Chromatograms were analyzed using each of the eight alignment programs, and the resulting data matrix was subjected to several preprocessing steps prior to analysis by principal components analysis (PCA) to flag compounds that differed between groups of samples. Programs were evaluated based on the average percentage of expected true positives flagged, and the percentage of peaks flagged that were in fact true positives. We present here the results of this comparison of the performance of the eight selected software programs.

2. Materials and methods

2.1. Preparation of sample solutions for Trial 1

Four standard mixtures were purchased, including Programmed Test Mix in methylene chloride (varied concentrations), Supelco® 37 component fatty acid methyl esther (FAME) mix 10 mg/mL in methylene chloride (varied), C4–C24 even carbon saturated fatty acid ethyl esters (FAEES) 1000 μg/mL each component in hexane, and C7–C30 saturated alkanes 1000 μg/mL each component in hexane (Table 1), all from Sigma–Aldrich, Oakville, Ontario, Canada.

For the first trial 16 samples were created by micro-pipetting specific volumes of four standard mixtures into individual vials as summarized in Table 2. Four distinct groups were created, with four separately prepared replicates in each group. Two mixtures were held constant in all samples: C7–C30 alkanes and FAMEs. The fatty acid ethyl esters mixture was left out of two groups, added to a third group, and doubled in concentration in the fourth. The programmed test mix comprising 11 compounds from a variety of classes were varied in concentration up to 12-fold between groups.

2.2. Preparation of sample solutions for Trial 2

For the second trial, tomato vine material including stems, leaves and fruit (33.71 kg) was collected from a commercial greenhouse in Leamington, Ontario on 26 April 2010 and dried at 70 °C for 24 h. Tomato vine tissue was milled to 20 mesh size. For non-polar extraction, 300 mg powder was extracted in 15 mL dichloromethane with ultrasonication for 15 min at room temperature, and centrifuged at 3000 rpm for 15 min, after which the dichloromethane supernatant was decanted into a flask. The