Fisher Ratio Method Applied to Third-Order Separation Data To Identify Significant Chemical Components of Metabolite Extracts

ARTICLE *in* ANALYTICAL CHEMISTRY · AUGUST 2006

Impact Factor: 5.64 · DOI: 10.1021/ac0602625 · Source: PubMed

8 AUTHORS, INCLUDING:

Robert E Synovec
University of Washington Seattle
164 PUBLICATIONS **4,658** CITATIONS

Anal. Chem. **2006,** *78,* 5068−5075

Fisher Ratio Method Applied to Third-Order Separation Data To Identify Significant Chemical Components of Metabolite Extracts

Karisa M. Pierce,[†] Jamin C. Hoggard,[†] Janiece L. Hope,[†,‡] Petrie M. Rainey,[§] Andrew N. Hoofnagle,[§] Rhona M. Jack,[∥] Bob W. Wright,[⊥] and Robert E. Synovec*[,†]

Department of Chemistry, Box 351700, and Department of Laboratory Medicine, Box 357110, University of Washington, Seattle, Washington 98195, Children's Hospital & Regional Medical Center, 6901 Sand Point Way NE, P.O. Box 50020, Seattle, Washington 98145-5020, and Pacific Northwest National Laboratory, Battelle Boulevard, P.O. Box 999, Richland, Washington 99352

This report is about applying a Fisher ratio method to entire four dimensional (4D) data sets from third-order instrumentation data. The Fisher ratio method uses a novel indexing scheme to discover the unknown chemical differences among known classes of complex samples. This is the first report of a Fisher ratio analysis procedure applied to entire 4D data sets of third-order separation data, which, in this case, is comprehensive two-dimensional gas chromatography coupled with time-of-flight mass spectrometry analyses of metabolite extracts using all of the collected mass channels. Current analysis methods for third-order separation data use only user-defined subsets of the 4D data set. First, in a validation study, the Fisher ratio method was demonstrated to objectively evaluate and determine the chemical differences between three controlled urine samples that differed by known spiked chemical components. It was determined that, out of more than 600 recognizable chemical components in a single sample, the six spiked components, along with only two other matrix components, differed most significantly in concentration among the control samples. In a second study, the Fisher ratio method was used in a novel application to discover the unknown chemical differences between urine metabolite samples from pregnant women and nonpregnant women. A brief list of the top 11 components that were most significantly different in concentration between the pregnant and nonpregnant samples was generated. Because the Fisher ratio calculation statistically differentiates regions of the chromatogram with large class-to-class variations from regions containing large within-class variations, the Fisher ratio method should generally be robust against biological diversity in a sample population. Indeed, application of principal component analysis in this second study failed due to biological diversity of the samples.

Comprehensive two-dimensional (2D) separation instruments coupled with multichannel detectors provide high peak capacity separations with selective detection that is ideally suited for complex sample analysis. Examples of such instruments include LC × LC,[1,2] LC × CE,[3,4] CE × CE,[5,6] LC × GC,[7,8] and GC × GC[9,10] coupled with multichannel detectors. With each analysis, these instruments produce a trilinear three-way array of data (3D), composed of the two separation dimensions and the detector channel dimension.[11−13] When numerous 3D analyses make up a data set, that data set is now a four-way array (4D) because of the added sample replicate dimension. It is common for analysts to use subsets of the 4D data set to discover significant chemical differences between complex samples. We describe a method that uses all of the 4D data set to objectively discover significant chemical differences between complex samples. The technique described herein is demonstrated using comprehensive two-dimensional gas chromatography coupled with time-of-flight mass spectrometry (GC × GC−TOFMS) separations of organic acid metabolites, but the technique is applicable to data from all 2D separation instruments coupled with multichannel detectors (i.e., third-order separation instruments).[14,15]

* Corresponding author. E-mail: synovec@chem.washington.edu. Phone: 206-685-2328. Fax: 206-685-8665.

[†] Department of Chemistry, University of Washington.

[‡] Current address: Scientific Resources Center, Cargill Research Bldg., 2301 Crosby Rd., Wayzata, MN 55391.

[§] Department of Laboratory Medicine, University of Washington.

[∥] Children's Hospital & Regional Medical Center.

[⊥] Pacific Northwest National Laboratory.

(1) Bushey, M. M.; Jorgenson, J. W. *Anal. Chem.* **1990**, *62*, 161−167.
(2) Holland, L. A.; Jorgenson, J. W. *Anal. Chem.* **1995**, *67*, 3275−3283.
(3) Bushey, M. M.; Jorgenson, J. W. *Anal. Chem.* **1990**, *62*, 978−984.
(4) Zhang, J.; Hu, H.; Gao, M.; Yang, P.; Zhang, X. *Electrophoresis* **2004**, *25*, 2374−2383.
(5) Michels, D. A.; Hu, S.; Dambrowitz, K. A.; Eggertson, M. J.; Lauterbach, K.; Dovichi, N. J. *Electrophoresis* **2004**, *25*, 3098−3105.
(6) Liu, H.; Yang, C.; Yang, Q.; Zhang, W.; Zhang, Y. *J. Chromatogr., B* **2005**, *817*, 119−126.
(7) Quigley, W. C.; Fraga, C. G.; Synovec, R. E. *J. Microcolumn Sep.* **2000**, *12*, 160−166.
(8) Koning, S.; Janssen, H.-G.; Deursen, M. v.; Brinkman, U. A. Th. *J. Sep. Sci.* **2004**, *27*, 397−409.
(9) Gorecki, T.; Harynuk, J.; Panic, O. *J. Sep. Sci.* **2004**, *27*, 359−379.
(10) Xie, L.; Marriott, P. J.; Adams, M. *Anal. Chim. Acta* **2003**, *500*, 211−222.
(11) Prazen, B. J.; Synovec, R. E.; Kowalski, B. R. *Anal. Chem.* **1998**, *70*, 218−225.
(12) Sanchez, E.; Kowalski, B. R. *J. Chemom.* **1990**, *4*, 29−45.
(13) Sinha, A. E.; Hope, J. L.; Prazen, B. J.; Fraga, C. G.; Nilsson, E. J.; Synovec, R. E. *J. Chromatogr., A* **2004**, *1056*, 145−154.

10.1021/ac0602625 CCC: $33.50
Published on Web 06/06/2006

Common approaches to determine the differences between complex samples in a 4D data set (in this case GC × GC−TOFMS separations of metabolites) involve extracting and analyzing target subregions from the original 3D data arrays. However, the user must know a target region beforehand and a large portion of the collected data is not utilized nor explored. Another approach is to compare extracted matrices from the original 3D data arrays (e.g., select a few mass channels or use total ion current chromatograms, TIC).[16−18] In fact, one group of metabolomics researchers discovered biomarkers in GC × GC−TOFMS TIC chromatograms of animal tissue extracts using comparison routines such as a relative weighted peak surface difference between two averaged samples.[19] Utilizing the TIC is actually utilizing a subset of the original 4D data set. We extend this idea to automatically use all of the mass channels (the entire 4D data set) so that information in unexpectedly selective mass channels may be gleaned from the data.

Multiway principal component analysis (PCA) has been applied to extracted matrices of the original 3D data arrays to determine exactly which compounds distinguish classes of samples.[20−22] PCA models the overall variation inherent in the data so variables that are highly loaded correspond with components that distinguish the classes. Since PCA is automated, PCA scripts could potentially be applied to entire 4D data sets as long as the 4D data set is properly indexed into a 2D data set. However, PCA is unsupervised so PCA depends on the scores clustering correctly. Thus, PCA is more apt to fail for data sets where portions of the sample profile (i.e., data for a given sample) have within-class (noise) variations that obscure other portions of the sample profile containing class-to-class (chemical) variations, causing the scores to cluster incorrectly.

On the other hand, the supervised Fisher ratio technique described herein is well suited for sample profiles that may have portions of large within-class variations. Given information about class membership and the assumption that class-to-class variations will be larger than sample-to-sample variations within a class for chromatographic peaks that truly distinguish classes, the Fisher ratio analysis will differentiate the portions of the sample profile that contain class-to-class variations from the portions of the sample profile that contain within-class variations. Since Fisher ratios of the detector signal are a function of only one variable, this analysis is generally used for first-order instrumentation data (2D matrices), not for higher dimensional arrays.[23−25] However, in our work, we provide the means of extending Fisher ratio analysis to third-order separation instrumentation data.

The general problem we address by using Fisher ratio analysis is the validity of our hypothesis that the complex samples contain statistically different component concentrations.[26] This basic idea was published by Fisher in the 1920s.[27] Given the boom in third-order separation instrumentation since the late 1990s, we are presenting a method for reapplying Fisher's basic idea to a third-order data structure using a novel indexing scheme. Our approach to third-order separation instrumentation data is significantly different from multivariate analysis of variance (MANOVA).[28,29] MANOVA has been used for classification of first-order instrumental data that are transformed into higher-order data by augmenting time-resolved information such as treatment/day data for metabolomics studies.[28,30−32] This MANOVA approach allowed statistical comparisons of days, animals, and other issues. The Fisher ratio method is not applicable to this type of pseudo-higher-order data; instead we propose the Fisher ratio method for third-order separation data. MANOVA for first-order instrumental responses requires applying PCA, partial least squares, or some sort of component analysis, forcing the user to determine the number of components to retain in their model.[28,29,33,34] There are also many reports of analysis of variance coupled with PCA for first- or second-order instrumentation.[28,33,35,36]

The report herein is about applying an automated Fisher ratio method to entire 4D GC × GC−TOFMS data sets using a novel indexing scheme to discover the sample components that distinguish complex metabolite samples. In a validation study, the Fisher ratio method is used to objectively "discover" the differences between three control urine samples. The results can then be validated by the known differences in spiked organic acid metabolites. In a second study, the Fisher ratio method is used to objectively discover the unknown differences in organic acid metabolites between urine samples from pregnant women and nonpregnant women.

THEORY

Given the class membership of each sample profile, the Fisher ratio analysis calculates a Fisher ratio at every point in the separation space. A Fisher ratio is the class-to-class variation of the detector signal divided by the sum of the within-class variations

(14) Dalluge, J.; Stee, L. L. P. V.; Xu, X.; Williams, J.; Beens, J.; Vreuls, R. J. J.; Brinkman, U. A. Th. *J. Chromatogr., A* **2002**, *974*, 169−184.
(15) Deursen, M. v.; Beens, J.; Reijenga, J.; Lipman, P.; Cramers, C. *J. High Resolut. Chromatogr.* **2000**, *23*, 507−510.
(16) Welthagen, W.; Shellie, R. A.; Spranger, J.; Ristow, M.; Zimmermann, R.; Fiehn, O. *Metabolomics* **2005**, *1*, 65−73.
(17) Hope, J. L.; Prazen, B. J.; Nilsson, E. J.; Lidstrom, M. E.; Synovec, R. E. *Talanta* **2005**, *65*, 380−388.
(18) Mohler, R. E.; Dombek, K. M.; Hoggard, J. C.; Young, E. T.; Synovec, R. E. *Anal. Chem.* **2006**, *78*, 2700−2709.
(19) Shellie, R. A.; Welthagen, W.; Zrostlikova, J.; Spranger, J.; Ristow, M.; Fiehn, O.; Zimmermann, R. *J. Chromatogr., A* **2005**, *1086*, 83−90.
(20) Giordani, P.; Kiers, H. A. L. *J. Chemom.* **2004**, *18*, 253−264.
(21) Wold, S.; Geladi, P.; Esbensen, K.; Ohman, J. *J. Chemom.* **1987**, *1*, 41−56.
(22) Smilde, A. K.; Westerhuis, J. A.; Jong, S. d. *J. Chemom.* **2003**, *17*, 323−337.
(23) Johnson, K. J.; Synovec, R. E. *Chemom. Intell. Lab. Syst.* **2002**, *60*, 225−237.
(24) Massart, D. L. *Chemometrics: A Textbook*; Elsevier Sciences Ltd.: New York, 1988.
(25) Pierce, K. M.; Hope, J. L.; Johnson, K. J.; Wright, B. W.; Synovec, R. E. *J. Chromatogr., A* **2005**, *1096*, 101−110.
(26) Sharaf, M. A.; Illman, D. L.; Kowalski, B. R. *Chemometrics*; John Wiley & Sons: New York, 1986.
(27) Fisher, R. A. *Statistical Methods for Research Workers*, 14 ed.; A. Constable Ltd.: Edinburgh, 1970.
(28) Smilde, A. K.; Jansen, J. J.; Hoefsloot, H. C. J.; Lamers, R.-J. A. N.; Greef, J. v. d.; Timmerman, M. E. *Bioinformatics* **2005**, *21*, 3043−3048.
(29) Stahle, L.; Wold, S. *Chemom. Intell. Lab. Syst.* **1990**, *9*, 127−141.
(30) Jansen, J. J.; Hoefsloot, H. C. J.; Greef, J. v. d.; Timmerman, M. E.; Smilde, A. K. *Anal. Chim. Acta* **2005**, *530*, 173−183.
(31) Jonsson, P.; Johansson, A. I.; Gullberg, J.; Trygg, J.; Jiye, A.; Grung, B.; Marklund, S.; Sjostrom, M.; Antti, H.; Moritz, T. *Anal. Chem.* **2005**, *77*, 5635−5642.
(32) Jansen, J. J.; Hoefsloot, H. C. J.; Boelens, H. F. M.; Greef, J. v. d.; Smilde, A. K. *Bioinformatics* **2004**, *20*, 2438−2446.
(33) Harrington, P. d. B.; Vieira, N. E.; Espinoza, J.; Nien, J. K.; Romero, R.; Yergey, A. L. *Anal. Chim. Acta* **2005**, *544*, 118−127.
(34) Jackson, J. E. *A User's Guide to Principal Components*; Wiley: New York, 1991.
(35) Greef, J. v. d.; Smilde, A. K. *J. Chemom.* **2005**, *19*, 376−386.
(36) Antti, H.; Bollard, M. E.; Ebbels, T.; Keun, H.; Lindon, J. C.; Nicholson, J. K.; Holmes, E. *J. Chemom.* **2002**, *16*, 461−468.

of the detector signal.[23−25,37] The class-to-class variation is calculated as

$$\sigma_{\mathrm{cl}}{}^2 = \frac{\sum(\bar{x}_\mathrm{i} - \bar{x})^2 n_\mathrm{i}}{(k-1)} \quad (1)$$

where n_i is the number of measurements in the ith class, $\bar{x}_\mathrm{i}$ is the mean of the ith class, $\bar{x}$ is the overall mean, and k is the number of classes. The within-class variation is calculated as

$$\sigma_{\mathrm{err}}{}^2 = \frac{\sum(\sum(\bar{x}_\mathrm{ij} - \bar{x})^2) - (\sum(\bar{x}_\mathrm{i} - \bar{x})^2 n_\mathrm{i})}{(N-k)} \quad (2)$$

where $\bar{x}_\mathrm{ij}$ is the ith measurement of the jth class and N is the total number of sample profiles. A Fisher ratio is then calculated as the ratio between the two variances:

$$\text{Fisher ratio} = \sigma_{\mathrm{cl}}{}^2/\sigma_{\mathrm{err}}{}^2 \quad (3)$$

Since Fisher ratios of the detector signal are a function of only one variable (retention time), Fisher ratio analysis is generally used for 2D data matrices, not for higher dimensional arrays. However, in this work, the 4D data set of GC × GC−TOFMS sample profiles was automatically reduced to 2D subsets using a novel indexing scheme. The indexing scheme ensures that all of the collected data contribute to the final determination of chemical components that statistically differ among sample types.

Figure 1 is a schematic of the indexing scheme for Fisher ratio analysis of third-order instrumental data. Each GC × GC−TOFMS analysis produces 3D arrays of data. When numerous GC × GC−TOFMS analyses make up a data set, that data set is composed of four dimensions. This 4D data set is then divided into subsets, where a subset is composed of an unfolded single m/z chromatogram, extracted from all of the sample replicates, and then augmented in a matrix. Such a subset is automatically made for all of the m/z monitored, and all of the subsets are individually submitted to Fisher ratio calculation. Each submission produces a 2D Fisher ratio plot; thus, if 361 m/z are monitored, then 361 2D Fisher ratio plots are automatically generated, where the 2D refers to the two GC separation dimensions. If the user decides to weight the ratios by chromatographic peak height, each point in the 2D Fisher ratio plot is multiplied by the mean, baseline-corrected chromatographic signal at that point for the respective m/z subset, thus minimizing any regions where chromatographic noise differences may happen to correlate with sample class differences. The 2D Fisher ratio plots of the subsets (with or without weighting) are then mathematically added together along the m/z dimension to create a 2D sum of Fisher ratios plot. The locations with the greatest 2D sum of Fisher ratio values may be projected onto the original GC × GC−TOFMS data arrays to objectively locate, and ultimately discover, the chemical components with concentrations that are statistically different among the samples.

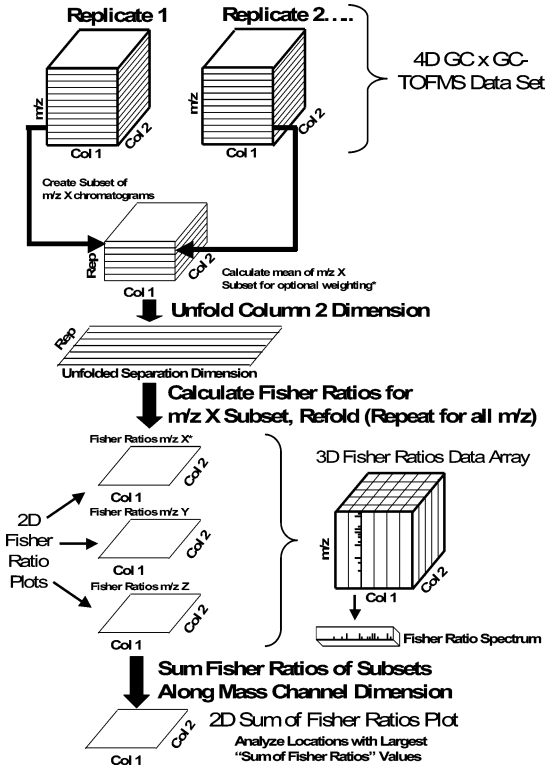

Figure 1. Schematic of novel indexing scheme to reduce 4D data to 2D data for calculation of Fisher ratios. Ultimately, the entire set of data that is collected is automatically (i.e., not manually) submitted to Fisher ratio analysis, and the unknown chemical differences among complex samples are objectively discovered.

EXPERIMENTAL SECTION

For the validation study using controlled samples, urine was collected, extracted, and derivatized at Children's Hospital (Seattle, WA). Further preparation occurred at the University of Washington (Seattle, WA) according to the following Children's Hospital procedure. The urine samples underwent basic biphasic extraction. The aqueous layer was treated with 50 μL of hydroxylamine oxime solution (25 mg/mL). The oximated samples were acidified (pH 1) and underwent a second biphasic extraction. The organic phase was dried and 100 μL of a trimethylsilation reagent, Sylon BFT (Supelco, Sigma Aldrich, St. Louis, MO), was added. The trimethylsilation reaction replaces active hydrogen atoms with trimethylsilyl groups (TMS groups), increasing the volatility and thermal stability of the organic acid metabolites. The extract from multiple patients was combined to increase the volume, producing a homogeneous urinary organic acid mixture (HUOA). Aliquots of HUOA were spiked with one of two organic acid standard solutions. One of the spiked urine samples, referred to as SSP, was an aliquot of HUOA spiked 1:1 (volume HUOA/volume standard solution) with a standard solution of TMS-derivatized succinic acid, salicylic acid, and *p*-hydroxybenzoic acid. The other spiked urine sample, referred to as KOC, was an aliquot of HUOA spiked 1:1 with a standard solution of TMS-derivatized α-ketoglutaric acid, oxalic acid, and citric acid. This resulted in the additional amount of 750 ng/μL of each spiked organic acid in the SSP and KOC validation samples, assuming maximum extraction efficiencies. The spiked urine samples are described in Table 1. Another HUOA aliquot was spiked with a standard solution blank, where the standard solution blank had undergone the same process of sample handling as the organic acid standard solutions, but contained no additional organic acid standard compounds. This resulted in a "benchmark" control sample, referred to as BNC.

(37) Duda, R. O.; Hart, P. E. *Pattern Classifications and Scene Analysis*; Wiley: New York, 1973.

Table 1. Concentration of Each Spiked Analyte Added to the Urine Matrix[a]

sample names	α-ketoglutaric acid (ng/μL)	oxalic acid (ng/μL)	citric acid (ng/μL)	salicylic acid (ng/μL)	succinic acid (ng/μL)	*p*-hydroxy-benzoic acid (ng/μL)	standard solution blank (μL)
KOC	750	750	750				
SSP				750	750	750	
BNC							50

[a] Samples labeled KOC were prepared by spiking 50 μL of the homogeneous urinary organic acid mixture (HUOA) with 50 μL of the α-ketoglutaric acid, oxalic acid, and citric acid standards mixture at the concentrations indicated. Samples labeled SSP were prepared by spiking 50 μL of the HUOA with 50 μL of the salicylic acid, succinic acid, and *p*-hydroxybenzoic acid standards mixture at the concentrations indicated. Samples labeled BNC were prepared by mixing 50 μL of the HUOA with 50 μL of the standard solution blank.

The control samples for the validation study were separated by a Leco Pegasus 4D GC × GC−TOFMS (Leco Corp., St. Joseph, MI). Column 1 was 20 m × 250 μm × 0.5 μm DB-5 (J & W Scientific, Alltech, Deerfield, IL). Column 2 was 2 m × 180 μm × 0.2 μm Rtx-200 (Restek Corp., Bellefonte, PA). The cryogenic modulator had a 2-s modulation period. The carrier gas was helium at 0.8 mL/min in constant flow mode. A volume of 1 μL was injected and split 100:1 with an inlet temperature of 250 °C. Four replicate separations of each sample were collected. The column 1 temperature program was 60 °C for 0.25 min, increased at 8 °C/min to 250 °C and held for 10 min. Column 2 was housed in a separate oven and held at a constant 10 °C higher than the column 1 oven temperature. The solvent delay was 5 min. Mass spectra were collected from *m/z* 40 to 400 at 100 spectra/s.

For the unknown urine samples study, the University of Washington Laboratory Medicine department provided urine samples from six pregnant women and from four women who were not pregnant. The use of patient specimens was approved by the University of Washington Human Subjects Review Committee. Volumes containing 0.7 mg of creatinine were obtained for each urine sample to account for the patient's individual kidney dilution effects. To each sample was added 50 μL of 0.5% (v/v) caproic acid/methanol internal standard. NaCl crystals were added until saturated, and the samples were vortexed. Then, 0.5 mL of NaOH was added to adjust pH to 12, and 1 mL of hydroxylamine hydrochloride was added. The samples were capped, vortexed, and left to sit for 1 h. Then, 3 mL of anhydrous ethyl acetate was added to each sample and the resulting mixture vortexed for 30 s. The samples were centrifuged for 2 min, and the top ethyl ether layer was discarded. A 0.5-mL aliquot of 6M HCl was added to adjust the pH to 1. The organic acids were extracted with 3 mL of ethyl ether, vortexed, and centrifuged, and the top ethyl acetate layer was transferred to a new centrifuge tube. The extraction of the pH 1 solution was repeated by adding 3 mL of ethyl ether, vortexing, centrifuging, and transferring the top ether layer to the tube containing the previous ethyl acetate extract. To the pooled ethyl acetate extracts was added 2 g of sodium sulfate, and the resultant mixture was then vortexed. The supernatant was transferred to a new centrifuge tube. The ethyl ether−acetate extract was evaporated at 60 °C with nitrogen until the volume was reduced, but not completely dried. The remaining extract was transferred to a GC autosampler vial. The extraction tube was rinsed with ethyl acetate, dried down, and added to the vial. The vials were heated at 60 °C until almost dry. To each vial was added 0.1 mL of BSTFA + 1% TMCS. Vials were capped and sat at 60 °C for 1 h.

The 10 urine samples with unknown differences due to pregnant versus nonpregnant status were separated in replicate by the Leco Pegasus 4D GC × GC−TOFMS (five replicates each). Column 1 was 20 m × 250 μm × 0.5 μm RTX-5MS (Restek Corp.). Column 2 was 2 m × 180 μm × 0.2 μm Rtx-200. The cryogenic modulator modulation period was 2 s. The carrier gas was helium at a constant 1.5 mL/min. A volume of 1 μL was injected and split 5:1 with an inlet temperature of 250 °C. The column 1 temperature program was 50 °C for 0.5 min, increased at 8 °C/min to 250 °C, and held for 0.5 min. Column 2 was housed in a separate oven and held at a constant 10 °C higher than the column 1 oven temperature. The solvent delay was 7 min. Mass spectra were collected from *m/z* 41 to 350 at 100 spectra/s.

The Fisher ratio algorithm was written using Matlab 7.0.4 (The Mathworks, Natick, MA). The code contained a subfunction that accessed Leco files, extracted the individual *m/z* chromatograms, and converted them into Matlab variables. We also have a version of the Fisher ratio algorithm that accesses .cdf files to extract the individual *m/z* chromatograms and convert them into Matlab variables. Control samples were normalized by the sum of the total ion current chromatogram. The pregnant and nonpregnant samples were normalized by the caproic acid standard peak at *m/z* 173. Mass spectral similarity searches were performed using NIST MS search 2.0 of the main library and an in-house library (NIST/EPA/NIH Mass Spectral Library, NIST 02).

RESULTS AND DISCUSSION

Each GC × GC−TOFMS analysis produces a 3D array of data as illustrated in Figure 1. An *m/z* 73 GC × GC−TOFMS chromatogram of the BNC validation sample is shown in Figure 2 (contour plots shown for clarity). The chromatographic peaks range over 3 orders of magnitude in this chromatogram, where peaks are defined as local maxima greater than five times the standard deviation of the noise. A peak-finding algorithm was used to determine that there are over 600 recognizable peaks in this single *m/z* chromatogram. A total of 361 *m/z* were monitored; so 361 such chromatograms exist for one GC × GC−TOFMS analysis.

The 12 GC × GC−TOFMS separations of the BNC, KOC, and SSP validation samples were submitted to Fisher ratio analysis with weighting as described in the Theory section. Since 361 *m/z* were monitored, 361 2D Fisher ratio plots were automatically generated and analyzed. For brevity, we begin by considering the 2D Fisher ratio plots for *m/z* 73 and 247. A contour plot of the 2D Fisher ratios for the *m/z* 73 subset is shown in Figure 3A. Peaks in the 2D Fisher ratio plot are defined as local maxima

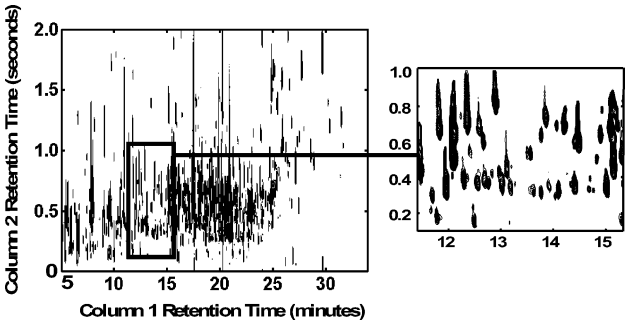

Figure 2. An *m*/*z* 73 GC × GC−TOFMS chromatogram (contour plot) of sample BNC and subregion. There are over 600 peaks in this single *m*/*z* chromatogram with peak signals ranging over 3 orders of magnitude. Peaks are defined as local maxima with signal greater than five times the standard deviation of the noise.

with 2D Fisher ratio values greater than five times the standard deviation of the noise regions. There are many 2D Fisher ratio peaks for m/z 73 because the hydrogen-bonding groups in the organic acid extracts were replaced with TMS groups, which yield m/z 73 fragment ions. The most intense 2D Fisher ratio peaks seen in Figure 3A are located where the spiked organic acids are located in the original chromatograms. Remaining peaks correspond with sample matrix chemical components with abundance variations that moderately correlate with the sample types. A contour plot of the 2D Fisher ratios for the m/z 247 subset is shown in Figure 3B. Some of the most intense 2D Fisher ratio peaks seen in Figure 3B are located where spiked organic acids with m/z 247 fragment ions were located in the original chromatograms. Fisher ratio values are essentially metrics of the signal-to-noise ratio of chemical selectivity, where, in this context, the "signal" corresponds to detected chemical differences. Some m/z contained features only where spiked chemical components eluted, and some m/z contained contributions from matrix components. So, some m/z ions are more selective than others. This suggests that wisely analyzing only selective m/z subsets could be beneficial for targeting families of complex sample components when the families have characteristic fragment ions. For the remainder of the analysis, we choose to analyze all of the m/z collected. Thus, the 361 2D Fisher ratio plots (with weighting) were summed along the m/z dimension to create a 2D sum of Fisher ratio plot, as shown in Figure 3C. Due to the validation study sample design, the spiked components were expected to have the greatest 2D sum of Fisher ratio values. The 2D sum of Fisher ratio peaks in Figure 3C were defined as local maxima with signals greater than five times the standard deviation of the noise. The peaks in Figure 3C with the greatest magnitude of 2D sum of Fisher ratios were located and labeled A−I. Table 2 contains the locations and magnitude of the top nine 2D sum of Fisher ratio peaks (labeled A−I, with and without weighting). These locations were projected onto the original GC × GC−TOFMS chromatograms. The locations of 2D sum of Fisher ratio peaks A−I corresponded with the locations of chromatographic peaks in the original GC × GC−TOFMS chromatograms for the metabolites of interest. Each chromatographic peak was identified using the full mass spectral information in the original GC × GC−TOFMS data. Identities of chromatographic peaks A−I are listed in Table 2, as well as the mass spectral match value at that location. Out of the nine peaks with the greatest 2D sum of Fisher ratios

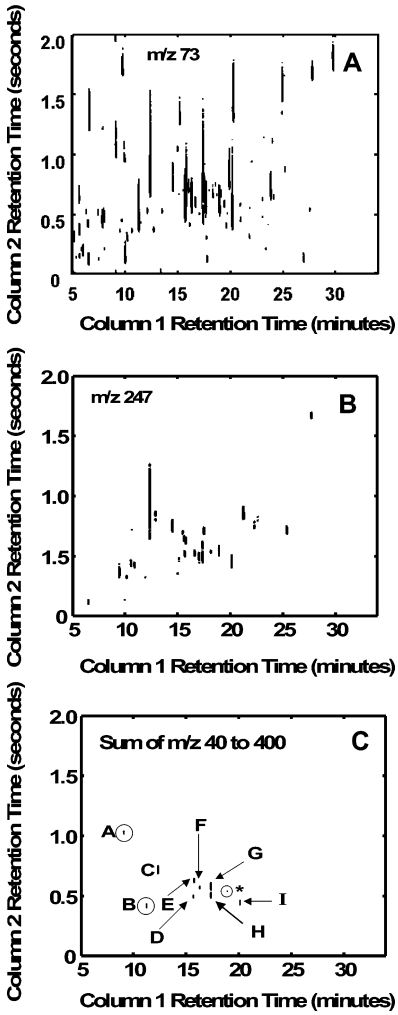

Figure 3. 2D Fisher ratio plots for various *m*/*z*. (A) A contour plot of the 2D Fisher ratios for the *m*/*z* 73 subset (with weighting). The minimum contour plotted is 2×10^{-6}. (B) A contour plot of the 2D Fisher ratios for the *m*/*z* 247 subset (with weighting). The minimum contour plotted was 9×10^{-9}. (C) The 2D sum of Fisher ratios plot (with weighting) for *m*/*z* 40−400. The 361 2D Fisher ratio plots were mathematically added together to create the 2D sum of Fisher ratio plots. The 2D sum of Fisher ratio peaks with the greatest magnitude are labeled A−I. The locations of 2D sum of Fisher ratio peaks A−I were projected onto the original GC × GC−TOFMS data arrays to create a list of chemical components that most distinguished the three samples in Table 2. The minimum contour plotted is 5.1×10^{-3}. *The starred peak matched to a citric acid impurity.

(with and without weighting), seven were the spiked components (α-ketoglutaric acid has two TMS-derivatized conformations). The remaining two labeled 2D sum of Fisher ratio peaks in Figure 3C (3-methylglutaconic acid and tiglylglycine) were urine matrix components. The starred 2D sum of Fisher ratios peak in Figure 3C is a citric acid impurity.

Figure 4 is a bar graph comparison of the TIC chromatographic peak volumes at the nine locations that were found in Figure 3C and listed in Table 2. The Fisher ratio analysis objectively "discovered" that out of the more than 600 recognizable sample components (counted in m/z 73 only), the 6 spiked components, along with only 2 other matrix components, were the most significant for distinguishing these complex samples. These experimental results serve to validate the algorithm and methodology.

Table 2. Control Sample Chemical Components with Greatest 2D Sum of Fisher Ratio Values from the Algorithm Validation Study[a]

feature	$t_{1,R}$ (min)	$t_{2,R}$ (s)	identity	MS match value	2D sum of Fisher ratios	2D sum of weighted Fisher ratios
A	9.1	1.04	oxalic acid	934	17 000	6.2×10^{-3}
B	11.3	0.43	3-methylglutaconic acid	684	29 000	6.9×10^{-3}
C	12.4	0.92	succinic acid	960	230 000	1.6×10^{-1}
D	15.7	0.50	salicylic acid	900	120 000	8.6×10^{-3}
E	15.8	0.63	α-ketoglutaric acid, 2-TMS	692	170 000	3.1×10^{-2}
F	16.3	0.57	tiglylglycine	716	100 000	9.2×10^{-3}
G	17.3	0.60	*p*-hydroxybenzoic acid, 4-TMS	962	37 000	2.0×10^{-1}
H	17.4	0.59	α-ketoglutaric acid 2	793	790 000	2.3×10^{-2}
I	20.1	0.45	citric acid, 3-TMS	920	150 000	4.2×10^{-2}

[a] Mass spectral match values obtained from NIST library search algorithm. Match values range from 0 to 999, with 999 being the best match.

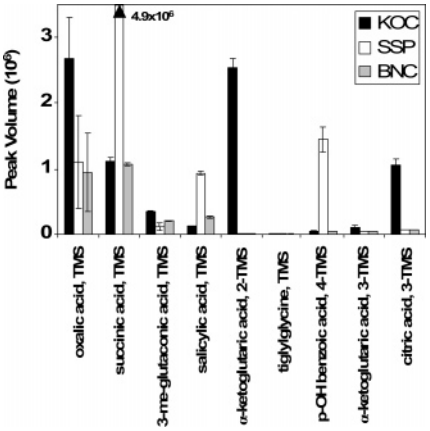

Figure 4. Bar graph of the TIC chromatographic peak volumes at the nine locations of interest that were automatically listed by the Fisher ratio algorithm and described in Table 2.

The Fisher ratio method reduced the original 4D GC × GC−TOFMS data set to a 3D data array of Fisher ratios as a function of column 1 retention time, column 2 retention time, and m/z. We will take a closer look at the information gleaned using all m/z ions in the Fisher ratios 3D data array to further validate the algorithm. It is reasonable to expect that, if the algorithm was properly calculating large Fisher ratios for the spiked components and if a spiked component was separated from all other interferences, then the m/z's with largest Fisher ratio values in the 3D Fisher ratios data array at the spiked component's location should be similar to the m/z's in the observed mass spectrum and matching library mass spectrum at the same location in the original GC × GC−TOFMS chromatograms. Signals in a single mass spectrum have equal selectivity between classes, regardless of m/z value, so an m/z ion with very small intensity may have a large Fisher ratio when the chemical component statistically differs in concentration between classes. Here, for proof of principle, we reduce the comparison to the top 15 m/z in terms of relative intensity. For example, the retention time for 2D sum of Fisher ratios peak B, in Table 2, was $t_{1,R} = 12.4$ min and $t_{2,R} = 0.92$ s. The Fisher ratios as a function of m/z located at 12.4 min and 0.92 s in the chromatographic space are displayed as a "spectrum" in Figure 5A. The top 15 m/z with the greatest Fisher ratio magnitude are labeled in Figure 5A. The observed mass spectrum at this same location in the original GC × GC−TOFMS data array

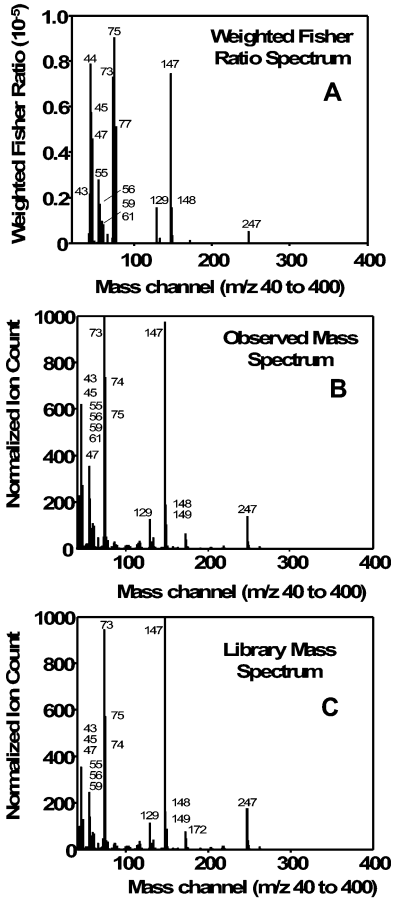

Figure 5. Automatically generated 3D data array of Fisher ratios as a function of peak B column 1 retention time, peak B column 2 retention time, and m/z, where peak B was identified as succinic acid-TMS. (A) Weighted Fisher ratios as a function of m/z located at 12.4 min and 0.92 s. (B) The observed mass spectrum located at 12.4 min and 0.92 s. (C) The library mass spectrum for succinic acid-TMS (that automatically matched to the observed mass spectrum) located at peak B (match value 960). The top 15 most intense m/z's are labeled in each spectrum.

was extracted (shown in Figure 5B) and compared to a library of mass spectra (Figure 5C). The top 15 most intense m/z in the observed mass spectrum are labeled in Figure 5B. The library mass spectrum for succinic acid (shown in Figure 5C) matched the extracted mass spectrum (match value 960). The top 15 most

Table 3. Pregnant and Nonpregnant Chemical Components with Greatest 2D Sum of Fisher Ratio Values from the Discovery Study[a]

$t_{1,R}$ (min)	$t_{2,R}$ (s)	identity	MS match value	2D sum of Fisher ratios	2D sum of weighted Fisher ratios	[Nonpreg]/[Pregnant]
8.3	1.93	glycolic acid[a]	950	1300	4.38×10^{-10}	0.56
9.3	1.89	glyoxylic acid[a]	935	2400	4.44×10^{-10}	0.84
10.6	1.72	2-methyl-3-hydroxybutyric acid, TMS[a,b]	920	9100	5.52×10^{-10}	1.39
12.9	1.9	*n*-valeric acid, TMS[b]	652	5700	1.11×10^{-10}	1.92
14.1	0.14	glutaric acid, TMS[b]	802	7700	1.74×10^{-10}	1.64
14.6	0.43	serine, TMS[b]	678	5300	6.68×10^{-11}	4.60
15.0	1.64	oxalic acid, TMS[b]	610	6000	3.12×10^{-11}	1.82
18.7	1.89	citric acid, 2-TMS[a,b]	855	11000	7.40×10^{-10}	1.82
18.9	0.07	α-*N*-acetylneuraminic acid, TMS[b]	592	5300	4.54×10^{-11}	1.23
19.5	1.97	benzoic acid, 3-TMS[a,b]	860	18000	5.53×10^{-10}	2.94
19.5	0.16	α-D-glucopyranoside, TMS[b]	645	5700	1.67×10^{-10}	1.63

[a] Mass spectral match values obtained from NIST library search algorithm. Match values range from 0 to 999, with 999 being the best match. The ratio of chemical compound chromatographic peak volumes in nonpregnant samples to pregnant samples is listed under the column heading [Nonpreg]/[Pregnant], representing the concentration ratio. [a] Five metabolites with greatest 2D sum of weighted Fisher ratios. [b] Nine metabolites with greatest 2D sum of Fisher ratios.

intense m/z's in the library mass spectrum are labeled in Figure 5C. The m/z's with largest Fisher ratio values at the peak B location (Figure 5A) were similar to the ion fragments in the observed (Figure 5B) and matching library (Figure 5C) mass spectra. The top 15 m/z's in the Fisher ratio spectrum that had the greatest magnitude matched 13 of the top 15 most intense m/z's in the observed mass spectrum. Likewise, the top 15 m/z's in the Fisher ratio spectrum matched 12 of the top 15 most intense m/z's in the library mass spectrum. Essentially, we are demonstrating that the Fisher ratio spectrum is in reasonable agreement with the true mass spectrum.

For chromatographic peaks that are not fully resolved from other peaks, the chromatographic profiles and the mass spectrum are not pure and may benefit from a third-order deconvolution technique like parallel factor analysis (PARAFAC).[38−40] The quantitative results in Figure 4 demonstrated that peak purity was sufficient to warrant use of the TIC for quantification, but using a deconvolution method such as PARAFAC would likely refine the quantitative results and the mass spectral match values.

In an ideal situation, the chromatographic peaks of the control samples would be completely resolved; there would be no systematic noise, no matrix effects, and the Fisher ratio method would return a list of 2D sum of Fisher ratios for which the greatest peaks corresponded with the spiked component peaks. In reality the situation is not ideal, so the analyst should use the weighting option in the Fisher ratio analysis. Weighting each m/z Fisher ratio subset by the mean, baseline-corrected chromatographic signal at that point of the m/z subset decreases the Fisher ratio for chromatographic noise regions that may happen to correlate with sample types. This weighting essentially improves the signal-to-noise-ratio of the chemical selectivity in the 2D Fisher ratio plot. For example, for the original spiked control samples without weighting, all of the spiked peaks were within the top 13 most intense 2D sum of Fisher ratio peaks. However, by weighting the original spiked control samples during Fisher ratio analysis, all of the spiked peaks were within the top nine most intense 2D sum of Fisher ratio peaks (shown in Figure 3C and Table 2).

We now turn our attention to applying the algorithm to true unknowns to demonstrate the utility of discovering differences in chemical composition. A data set of 50 GC × GC−TOFMS separations of urine extracts from four patients who were not pregnant (5 replicates each) and six patients who were pregnant (5 replicates each) was obtained. The data set was submitted to the Fisher ratio algorithm with and without weighting. The peaks in the 2D sum of the Fisher ratio plot were automatically located and sorted in descending order of 2D sum of Fisher ratio values (with and without weighting). The locations of the top five peaks with the greatest 2D sum of Fisher ratio values (with weighting) are listed in Table 3. The locations of the top nine peaks with the greatest 2D sum of Fisher ratio values (without weighting) are listed in Table 3. The peaks in the 2D sum of the Fisher ratio plots (with and without weighting) were located where chromatographic peaks were located. The chromatographic peaks eluting at the 11 listed retention times in the original GC × GC−TOFMS data arrays were identified; the identities and match values are listed in Table 3. Some of the match values were low. This is most likely due to either unresolved peaks or the unknown analyte spectrum not being present in the libraries. Also listed in Table 3 are ratios of the nonpregnant TIC peak volumes to pregnant TIC peak volume. The Fisher ratio method objectively discovered components that significantly differentiated the nonpregnant samples from the pregnant samples.

Because the Fisher ratio calculation differentiates class-to-class variations from within-class variations, it is robust against biological diversity. For instance, one might expect that the list of chromatographic peaks sorted in descending order by the 2D sum of Fisher ratio values (with and without weighting) is going to be the same as the list of chromatographic peaks sorted in descending order by the ratio of the chromatographic peak volumes between classes. According to Table 3, this was not the result because uncontrollable sample matrix differences caused within-class variations that lower the magnitude of some Fisher ratios,

(38) Wu, H.-L.; Shibukawa, M.; Oguma, K. *J. Chemom.* **1998**, *12*, 1−26.
(39) JiJi, R. D.; Andersson, G. G.; Booksh, K. S. *J. Chemom.* **2000**, *14*, 171−185.
(40) Sinha, A. E.; Hope, J. L.; Prazen, B. J.; Nilsson, E. J.; Jack, R. M.; Synovec, R. E. *J. Chromatogr., A* **2004**, *1058*, 209−215.

despite large concentration differentials. The average variation for all 11 peaks in Table 3 was 54% RSD. Biological variations are commonly 10-fold larger than instrumental variations.[41] However, the Fisher ratio analysis was robust against the biological diversity of the sample matrix and yielded a list of specific chemical components that truly differentiated the sample classes. The pregnant and nonpregnant urine samples were used as a proof-of-principle example of discovering unknown differences between complex samples. The clinical significance of the chemical compounds listed in Table 3 is not yet known and is an inspiration for future work.

Compared to PCA, the Fisher ratio method handles biodiversity better by differentiating regions of large within-class variation from regions of large class-to-class variation. To compare the Fisher ratio method with the PCA method, the m/z 73 chromatograms were extracted from all 50 of the GC × GC−TOFMS separations of the pregnant and nonpregnant samples. These chromatograms were submitted to multiway PCA. PCA failed to cluster the scores according to pregnant and nonpregnant sample types (not shown for brevity), so the loadings could not be used to determine the significant sample components. The biological diversity between women within a single class caused PCA to fail, whereas the reported Fisher ratio method yielded a list of components that significantly differed between the classes.

It would be interesting to compare the results in Table 3 to literature results using other analytical methods, but studies regarding changes in organic acid concentrations of urine during normal pregnancies are rare because the technology to discover the differences was not available until recently. The concentrations of the 11 organic acids listed in Table 2 have not been previously reported. However, it has been reported that ~50% of pregnant women have an increased concentration of 3-hydroxyisovaleric acid in their urine over nonpregnant women.[42−44] We located and quantified 3-hydroxyisovaleric acid in our data. The pregnant urine samples did have a higher average concentration of 3-hydroxyisovaleric acid. The ratio of nonpregnant to pregnant 3-hydroxyisovaleric acid peak volumes was 0.39. At the 3-hydroxyisovaleric acid location, the 2D sum of Fisher ratio values was 2900 and the 2D sum of weighted Fisher ratio values was 8.40×10^{-12}. The 3-hydroxyisovaleric acid peak volume for the pregnant samples had a large relative standard deviation (110%).

CONCLUSIONS

The Fisher ratio method provided a means of using all of the data that were collected to determine the chemical components that significantly differed in concentration between sample types. The Fisher ratio method will not perform optimally when retention time variations obscure chemical variations. A comprehensive 2D alignment algorithm was recently published that can process entire sets of 2D separations data, but the algorithm needs further development in order to preserve the spectroscopic dimension of third-order instrumentation data.[45] In the future, the Fisher ratio method will be used as a front end tool for finding subregions of the separation that contain unresolved peaks which can benefit from submission to PARAFAC using only the selective m/z that have high Fisher ratios.[13,40] This method would allow fast PARAFAC processing of the entire 4D data set.

ACKNOWLEDGMENT

This work was supported by the Internal Revenue Service through an Interagency Agreement with the U.S. Department of Energy. The Pacific Northwest National Laboratory is operated by Battelle Memorial Institute for the U.S. Department of Energy under contract DE-AC05-76RLO 1830. The views, opinions, or findings contained in this report are those of the authors and should not be construed as the official Internal Revenue Service position, policy, or decision unless designated by other documentation.

Received for review February 9, 2006. Accepted May 2, 2006.

AC0602625

(41) Fiehn, O.; Kopka, J.; Dormann, P.; Altmann, T.; Trethewey, R. N.; Willmitzer, L. *Nat, Biotechnol.* **2000**, *18*, 1157−1161.
(42) Mock, D. M.; Quirk, J. G.; Mock, N. I. *Am, J. Clin. Nutr.* **2002**, *75*, 295−299.
(43) Mock, D. M.; Jackson, H.; Lankford, G. L.; Mock, N. I.; Weintraub, S. T. *Biomed. Environ. Mass Spectrom.* **1989**, *18*, 652−656.
(44) Mock, D. M.; Stadler, D. D.; Stratton, S. L.; Mock, N. I. *J. Nutr.* **1997**, *127*, 710−716.
(45) Pierce, K. M.; Wood, L. F.; Wright, B. W.; Synovec, R. E. *Anal. Chem.* 2005, *77*, 7735−7743.