Journal of Chromatography A, 1361 (2014) 265–276

Contents lists available at ScienceDirect

Journal of Chromatography A

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

Urinary metabolic fingerprinting of mice with diet-induced metabolic derangements by parallel dual secondary column-dual detection two-dimensional comprehensive gas chromatography☆

Davide Bressanello[a], Erica Liberto[a], Massimo Collino[a], Stephen E. Reichenbach[b], Elisa Benetti[a], Fausto Chiazza[a], Carlo Bicchi[a], Chiara Cordero[a,*]

[a] Dipartimento di Scienza e Tecnologia del Farmaco, Università degli Studi di Torino, Via Pietro Giuria 9, I-10125 Torino, Italy
[b] Computer Science and Engineering Department, University of Nebraska, 1400 R Street, Lincoln, NE 68588-0115, USA

ARTICLE INFO

Article history:
Received 27 June 2014
Received in revised form 31 July 2014
Accepted 4 August 2014
Available online 14 August 2014

Keywords:
Two-dimensional comprehensive gas chromatography–mass spectrometry
Parallel dual secondary column-dual detection
Diet-induced metabolic derangements
Quantitative metabolomics
Urine samples profiling
Fingerprinting

ABSTRACT

This study investigates the potential of a parallel dual secondary column-dual detection two-dimensional comprehensive GC platform (GC × 2GC–MS/FID) for metabolic profiling and fingerprinting of mouse urine. Samples were obtained from a murine model that mimics a typical unhealthy Western diet featuring both high fat and sugar (HFHS) intake, which induces obesity, dyslipidemia, and insulin resistance. Urines collected at different steps of the study were used to obtain pivotal and comparative data on the presence and relative distributions of early markers of metabolic disease. The data elaboration and interpretation work-flow includes an advanced untargeted fingerprinting approach, with peak-region features to locate relevant features to be quantified by external standard calibration. The reliability of untargeted fingerprinting is confirmed by quantitative results on selected relevant features that showed percentages of variations consistent with those observed by comparing raw data quantitative descriptors (2D peak-region volumes and percent of response). Analytes that were up-regulated with % of variation ranging from 30 to 1000, included pyruvic acid, glycerol, fructose, galactose, glucose, lactic acid, mannitol and valine. Down-regulation was evidenced for malonic acid, succinic acid, alanine, glycine, and creatinine. Advanced fingerprinting also is demonstrated for effectively evaluating individual variations during experiments, thus representing a promising tool for personalized intervention studies. In this context, it is interesting to observe that informative features that were not discriminant for the entire population may be relevant for individuals.

1. Introduction

Early changes in metabolite profiles of biofluids (e.g., plasma and urine samples) are considered reliable biomarkers of early metabolic dysfunction and often are used to characterize clinical manifestations of metabolic disorders, mainly type 1 and type 2 diabetes. A key pathogenic mechanism is the disruption of glucose homeostasis, which leads to the development of insulin resistance and impaired insulin production. Disturbances in both the secretion and action of insulin impact the global regulation of metabolism, affecting the composition of blood, urine, and other body fluids. Traditionally, to get a vision of the physiopathologic responses related to metabolic glucose deregulation, single metabolites or classes of small molecules are measured using targeted analytical assays. In that approach, the relationships among diverse metabolites and multiple pathways are ignored, hindering a useful integrated vision in the assessment of complex diseases. More recently, the identification of potential disease biomarkers has been greatly facilitated by the upsurge in new technologies for comprehensive metabolic profiling, which are often collectively termed metabolomics [1–3]. For instance, recent epidemiological studies used metabolomics to predict incident diabetes and revealed branched-chain and aromatic amino acids, including, isoleucine, leucine, valine, tyrosine, and phenylalanine, as highly significant predictors of future diabetes [3,4].

In this context, two-dimensional comprehensive gas chromatography with mass spectrometry (GC × GC–MS) represents one of the most advanced and informative hyphenated GC platforms currently available for medium-to-low molecular weight metabolite profiling. Thanks to its superior separation power, sensitivity,

☆ Presented at 38th International Symposium on Capillary Chromatography and 11th GCxGC Symposium, 18–23 May 2014, Riva del Garda, Italy.

* Corresponding author. Tel.: +39 011 6707662; fax: +39 011 2367662.
E-mail address: chiara.cordero@unito.it (C. Cordero).

http://dx.doi.org/10.1016/j.chroma.2014.08.015
0021-9673/

and informative bi-dimensional (2D) separation patterns, detailed profiles and fingerprints of complex biological samples can be comprehensively evaluated. However, to reveal the so-called *metabolic fine print* [4], analytical efforts must be directed to low molecular weight organic compounds (<1500 Da) with a great diversity of chemical properties and wide concentration ranges. As a consequence, robust, reproducible, accurate, and informative methods are needed to enable reliable samples comparisons.

From this perspective, when a GC × GC–MS platform is adopted, the system configuration represents a critical but challenging aspect requiring careful tuning of the columns' stationary phase chemistry and dimensions (especially length and inner diameter) to maximize the system separation power and simultaneously avoid second-dimension (^{2}D) column overloading, thereby improving quantitation accuracy and response linearity over a wider range of concentrations [5]. Quantitative metabolomics, which includes not only the detailed profiling of metabolites but also their true quantitation, is required to realize the potential of biomarker investigations.

To date, most studies of metabolic profiling by GC × GC–MS have used a conventional column setup consisting of a non-polar primary (^{1}D) column of 30 m × 0.25 mm d_c × 0.25 μm d_f and a single mid-polarity secondary (^{2}D) column of 1–2 m × 0.1 mm d_c × 0.1 μm d_f [6]. However, to overcome some limits of conventional column configurations in these earlier studies, Koek et al. demonstrated that wider bore ^{2}D columns (i.e., 0.25 mm d_c) with higher mass loadability provided more precise and accurate quantitative results, although the overall system peak capacity was lower [6]. More recently, Rocha et al. [7] investigated the composition of human urine volatilome, adopting an apolar (DB-5) ^{1}D column of 30 m × 0.32 d_c × 0.25 μm d_f coupled to a polar (DB-FFAP) ^{2}D column of 0.79 m × 0.25 d_c × 0.25 μm d_f. That column setup provided appropriate orthogonality and suitable mass loadability for the analytes under study.

Generally, GC × GC detection requires fast detectors, for example flame ionization detector (FID) or electron impact (EI) fast acquisition time-of-flight mass spectrometers (TOFMS). However, reliable and consistent results both in terms of analyte identification and quantitation also can be obtained with modern fast quadrupoles, operating at high frequencies [8,9]. These MS detectors are experiencing a growing popularity in GC × GC applications, confirmed by the increasing number of publications appearing in the literature [10]. Last but not least, high-resolution TOFMS (HR-TOFMS) detectors are emerging as valuable tools in hyphenated multidimensional analytical platforms for metabolomics because of their informative potential in analyte identifications based on accurate mass detection [11].

This study investigates the potential of a parallel dual secondary column-dual detection two-dimensional comprehensive GC platform (GC × 2GC–MS/FID) for metabolic profiling and fingerprinting of urine samples obtained from a murine model of diet-induced metabolic derangements. The animal model mimics a typical unhealthy Western diet featuring both high fat and sugar (HFHS) intake, which induces obesity, dyslipidemia, and insulin resistance [13,14]. Urine samples collected at different times during the study were used to obtain pivotal and comparative data on the presence and relative distribution of early markers of metabolic disease.

The instrumental platform operates at close-to-optimal ^{2}D linear velocities in both chromatographic dimensions and doubles secondary column loading capacity with positive effects on overall system orthogonality, efficiency, and selectivity [12]. In addition, the improved information potential due to the dual detection (by MS and FID), poses challenges in terms of data elaboration but offers the opportunity to cross-validate results of both targeted-quantitative and untargeted profiling.

The data elaboration and interpretation work-flow includes an advanced untargeted fingerprinting approach with peak-region features [11,15,16] to locate relevant features to be quantified by external standard calibration. Accuracies of both targeted and untargeted elaboration are assessed by comparing MS and FID results. The advantages of dual detector/dual pattern information cross-matching in terms of exploiting the overall system potential for comparative analysis and quantitative metabolomics are apparent in the results.

2. Experimental

2.1. Chemicals

All chemicals were from Sigma–Aldrich (Milan, Italy), in particular:

(a) pure standards of *n*-alkanes (from *n*-C9 to *n*-C25) for system evaluation, flow/pressure optimization, and linear retention index (I^T_S) determination;
(b) pure standards for quantitative determinations of pyruvic acid, lactic acid, malonic acid, succinic acid, malic acid, 2-ketoglutaric acid, hippuric acid, L-alanine, L-valine, glycine, L-threonine, L-tyrosine, creatinine, phenylalanine, xylitol, ribitol, glycerol, fructose, galactose, glucose, mannitol, and myo-inositol, and the internal standard (ISTD) gallic acid;
(c) derivatization reagents O-methylhydroxylamine hydrochloride (MOX) and N-methyl-N-(trimethylsilyl)trifluoroacetamide (MSTFA);
(d) HPLC-grade solvents: methanol, pyridine, *n*-hexane, and dichloromethane.

2.2. Samples

Four-week-old male C57BL/6J mice (n = 16) (Harlan-Italy; Udine, Italy) were housed in a controlled environment at 25 ± 2 °C. All the animals were fed with a normal pellet diet for 1 week prior to the experimentation. The animals then were allocated to one of two dietary regimens, either normal (control, n = 8) or a high-fat high-sugar diet (HFHS, n = 8), for 12 weeks. The HFHS diet contained 45% fat, 20% protein, and 35% carbohydrate. Animal care was in compliance with the "Principles of laboratory animal care" (NIH publication 85–23, 1985) and the experimental protocol has been approved by the Turin University Ethics Committee. Urine samples were collected at 1 week (basal) and after 6, 9, and 12 weeks (W6, W9, and W12) then immediately quenched on liquid nitrogen and stored at −80 °C until derivatization/analysis. For urine collection, conscious mice were individually placed in metabolic cages with free access to water for 16 h.

Urine samples were submitted to a standard derivatization protocol [17] consisting of the following steps: 200 μL of urine and a suitable volume of ISTD (gallic acid solution at 10 g/L) were diluted with methanol up to 1000 μL and carefully mixed (Whirlimixer vortex, Fisher Scientific, Loughborough, Leicestershire, UK). Then, 30 μL of MOX were added to 20 μL of that solution and the resulting solution was incubated for 2 h at 60 °C. Next, 30 μL of MSTFA were added and the mixture was incubated at 100 °C for 60 min. The resulting sample solution diluted in *n*-hexane was immediately analyzed in duplicate or stored at −80 °C until analysis.

2.3. GC × 2GC–MS/FID instrument set-up

GC × GC analyses were run with the following system configuration: an HT280T multipurpose sampler (HTA, Brescia, Italy) was integrated with an Agilent 6890 GC unit coupled to an Agilent

5975C MS detector (Agilent, Little Falls, DE, USA) operating in EI mode at 70 eV. The GC transfer line was set at 300 °C. An *Auto Tune* option was used and the scan range was set to m/z 50–350 with a scanning rate of 12,500 amu/s to obtain a spectra generation frequency of 25 Hz. The flame ionization detector (FID) was operated as follows: base temperature 300 °C, H_2 flow 40 mL/min, air flow 240 mL/min, make-up (N_2) 450 mL/min, and sampling frequency 150 Hz.

The column set consisted of primary column of 30 m × 0.25 mm d_c × 0.25 μm d_f SE52 (95% polydimethylsiloxane, 5% phenyl) connected to two secondary columns of equivalent length of 1.4 m × 0.1 mm d_c × 0.10 μm d_f OV1701 (86% polydimethylsiloxane, 7% phenyl, 7% cyanopropyl). Connections between the primary and the two secondary columns were by a SilFlow™ GC 3 Port Splitter (SGE Ringwood, Victoria, Australia). The secondary column toward the MS detector was connected to a Quick Swap unit (G3185, Agilent, Little Falls, DE, USA) and to an auxiliary electronic pressure controller (EPC) consisting of a one channel Pneumatics Control Module (G2317A, Agilent, Little Falls, DE, USA). The restrictor capillary in the GC–MS transfer line was of 0.17 m × 0.1 mm d_c. A schematic picture of the system configuration is provided as a supplementary file (supplementary Fig. 1 (SF1)). All columns and capillaries were from Mega (Legnano, Milan, Italy). The carrier gas was helium delivered at constant flow with initial head pressure p_i 296.0 KPa and the auxiliary gas for MS outlet pressure correction (He) was delivered at 39.9 KPa (relative). The split ratio (MS/FID) was 50:50.

Supplementary material related to this article can be found, in the online version, at http://dx.doi.org/10.1016/j.chroma.2014.08.015.

Injections for the analysis of both urine samples and *n*-alkanes for linear retention indices determination I^T_S, was by a HT280T sampler (HTA, Brescia, Italy) under the following conditions: split/splitless injector, split mode, split ratio 1/10, injector temperature 280 °C, and injection volume 2 μL. The oven temperature programme was 50 °C (1 min) to 300 °C (10 min) at 4.0 °C/min.

The system was equipped with a two-stage KT 2004 loop thermal modulator (Zoex Corporation, Houston, TX) cooled with liquid nitrogen controlled by Optimode™ V.2 (SRA Instruments, Cernusco sul Naviglio, MI, Italy). The hot jet pulse time was set at 350 ms, modulation time was 5 s, and the cold-jet total flow was progressively reduced with a linear function from 30% of mass flow controller (MFC) at initial conditions to 5% at the end of the run. Loop dimensions were chosen on the basis of the expected carrier linear velocities to ensure that two-stage band focusing and release were performed for each modulation stage. The first 0.6 m of the ^{2}Ds was wrapped in the metal slit of the loop-type modulator.

2.4. Data acquisition and pattern elaboration

Samples of urine from mice exposed to either normal (control, $n = 8$) or a high-fat high-sugar diet (HFHS, $n = 8$) for 12 weeks collected at 1 week (W1-basal) and after 6, 9, and 12 weeks (W6, W9, and W12) were processed with the proposed method. With eight subjects (or experimental replicates) for each of two subject conditions (control and HFHS diet), four sample collection periods (at 1, 6, 9, and 12 weeks), and two chromatographic replicates per sample, a total of 128 chromatograms were acquired and analyzed.

Data were acquired by Agilent MSD ChemStation ver D.02.00.275 and processed with GC Image GC × GC Software version 2.4b2 (GC Image, LLC, Lincoln, NE, USA). Statistical analysis was performed with SPSS 14.0 (SPSS Inc., Chicago, IL, USA) and heat map visualization by GENE-E v 3.0.77 (Broad Institute, Inc., Cambridge, MA, USA).

2.5. Parallel dual column-dual detection system setup and performance verification

The system adopted consists of two ^{2}D columns of 1.4 m. A minimal outlet pressure compensation toward MS was required because of the low resistance of the two parallel ^{2}Ds instead of a single one. The mid-point pressure (between the ^{1}D and the two ^{2}D columns) was therefore 182.6 kPa, with an average linear velocity in the first dimension of ($^1\bar{u}$) 34.2 cm/s and in the ^{2}D columns of 179.9 cm/s (hold-up: 0.8 s) [18,19].

The outlet pressure correction of 40 kPa was verified for correctness and consistency by isothermal analysis (at 150 °C) of linear hydrocarbons from C13 to C15 at 296 kPa head-pressure. Supplementary Fig. 2 (SF2) reports overlaid raw chromatograms (FID and MS) and the FID 2D plot. System hold-up times were 1.910 min and 0.88 s in the ^{1}D and ^{2}D, respectively.

Supplementary material related to this article can be found, in the online version, at http://dx.doi.org/10.1016/j.chroma.2014.08.015.

A further verification was done on the reference mixture of target analytes for quantitation. The consistency of the alignment is evident from Fig. 1 (upper part), which shows the overlaid traces of MS total ion current and FID of a 1 mg/L calibration solution.

2.6. Quantitative metabolomics validation parameters

Method validation was run on a three-week protocol, over two months, and the following parameters were characterized: precision, linearity, accuracy, and limit of quantitation (LOQ). Precision data (intra and inter-week precision on retention times and 2D peak volumes on analytes *Ti*) were evaluated by replicating analyses during two months. Linearity was assessed by linear regression analyses within the working range, over at least six different concentration levels, and for each detector (i.e., MS and FID). Experimental results on linearity assessment are reported in Table 1 (including calibration ranges, regression lines, and determination coefficients R^2). Calibration solutions for quantitative determination of relevant analytes identified by peak-region feature fingerprinting (pyruvic acid, lactic acid, malonic acid, succinic acid, malic acid, 2-ketoglutaric acid, hippuric acid, L-alanine, L-valine, glycine, L-threonine, L-tyrosine, creatinine, phenylalanine, xylitol, ribitol, glycerol, fructose, galactose, glucose, mannitol, and myo-inositol) were prepared by mixing single-component standard mother solutions at 10 g/L in suitable solvents and adjusting the final volume up to the required concentration. Each solution then was submitted to derivatization steps detailed in Section 2.3 and directly analyzed. Calibration levels investigated were: 40 mg/L, 30 mg/L, 15 mg/L, 10 mg/L, 5 mg/L, 1 mg/L, 0.5 mg/L, and 0.1 mg/L. Gallic acid, i.e., the internal standard for data normalization and quality control, was at 10 mg/L.

Precision, also reported in Table 1, is expressed as RSD% on normalized 2D volumes from the calibration solution level at 10 mg/L (arbitrarily selected among the others) collected over two months of validations (for a total of nine replicates). The accuracy was assessed by cross-comparison of quantitative results obtained by MS and FID detection (correlation function $R^2 = 0.998$) and by the absolute error (≤5%).

The limit of quantification (LOQ) was determined experimentally by analyzing decreasing concentrations of standard calibrating solutions. Each sample was analyzed in triplicate, and the LOQ was the lowest concentration for which instrumental response (2D peak volume on *Ti*) reported an RSD% below 30%, across replicate analyses. LOQ values also are reported in Table 1.

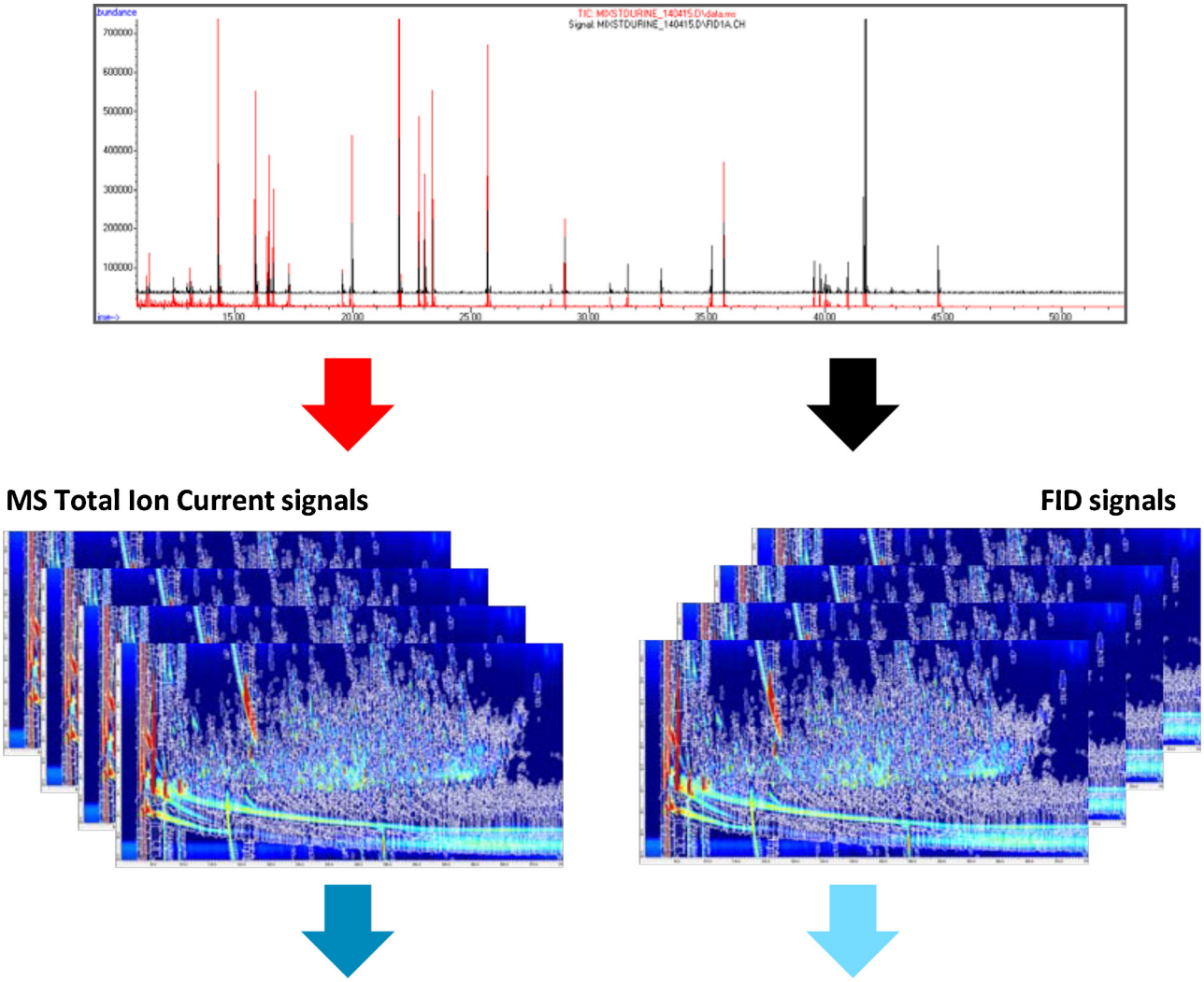

Fig. 1. Schematic diagram of the work-flow followed in the present study.

Table 1
Summary of validation data (FID and MS signals) from the three-week protocol adopted. Target analytes are reported together with target ions and qualifiers (m/z), retention times in the two dimensions (^{1}D Rt (min) and ^{2}D Rt (s)), ^{1}D linear retention index (I^T), calibration range (mg/L), regression equations formulae and coefficient of correlation (R^2), limit of quantitation LOQ (mg/L), and precision on normalized 2D volumes expressed as relative standard deviation (RSD%) on 10 mg/L calibration solution analyzed over the entire validation period.

Analyte	MS signal data									FID signal data				
	Target ion and qualifiers (m/z)	^{1}D Rt (min)	^{2}D Rt (s)	I^T	Range (mg/L)	Regression equation	R^2	LOQ (mg/L)	Precision RSD% (10 mg/L) ($n=9$)	Regression equation	R^2	LOQ (mg/L)	Precision RSD% (10 mg/L) ($n=9$)	
Pyruvic acid	73;174;45	4.08	0.72	787	40–5	$y=0.019x+0.014$	0.978	0.5	4.30	$y=0.026x+0.019$	0.985	0.5	2.43	
Lactic acid	147;117;73	14.25	1.74	1056	40–0.1	$y=0.066x+0.032$	0.979	0.5	2.75	$y=0.091x+0.013$	0.988	0.5	1.53	
Alanine	116;73;147	15.84	1.74	1100	40–5	$y=0.127x$	0.969	1	4.45	$y=0.179x+0.073$	0.977	1	3.65	
Malonic acid	147;73;233	19.51	2.12	1200	40–0.5	$y=0.259x-0.03$	0.989	0.05	3.87	$y=0.589+0.01$	0.998	0.05	3.97	
Valine	144;73;218	19.92	1.74	1211	40–5	$y=0.127x$	0.978	1	3.14	$y=0.028x+0.018$	0.979	1	2.13	
Glycerol	147;73;205	21.91	1.61	1269	40–0.1	$y=0.078x$	0.984	0.05	3.55	$y=0.608x+0.008$	0.988	0.05	4.55	
Glycine	174;73;147	23.11	1.87	1300	40–0.1	$y=0.258x$	0.996	0.05	1.13	$y=0.694x+0.033$	0.997	0.05	3.13	
Succinic acid	147;73;247	23.34	2.08	1310	40–1	$y=0.237x$	0.988	0.05	5.67	$y=0.699+0.021$	0.989	0.05	2.34	
Threonine	73;218;117	25.67	1.78	1380	40–5	$y=0.084x$	0.906	1	3.35	$y=0.116x+0.078$	0.906	1	5.49	
Malic acid	73;147;233	28.92	2.00	1481	40–5	$y=0.205x$	0.995	0.05	6.88	$y=0.572x+0.018$	0.997	0.1	7.01	
Creatinine	115;73;143	30.84	2.12	1545	40–5	$y=0.028x-0.013$	0.978	1	5.61	$y=0.028x+0.017$	0.978	1	4.43	
2-Ketoglutaric acid	73;147;198	31.59	2.22	1571	40–5	$y=0.175x-0.3$	0.982	1	4.65	$y=0.138x+0.012$	0.989	1.5	3.97	
Phenylalanine	73;218;192	33.01	2.17	1620	40–5	$y=0.079x-0.064$	0.966	1	3.21	$y=0.109x+0.04$	0.965	1	1.23	
Xylitol	73;217;147	35.17	1.66	1697	40–0.5	$y=0.229x-0.1$	0.997	0.05	4.88	$y=0.122x+0.033$	0.993	0.05	5.07	
Ribitol	73;217;147	35.69	1.68	1715	40–0.1	$y=0.254x-0.01$	0.996	0.05	1.65	$y=0.465x+0.017$	0.997	0.05	3.52	
Hippuric acid	105;73;206	39.00	3.48	1841	40–5	$y=0.021x-0.138$	0.984	1	1.27	$y=0.027x+0.013$	0.984	1	0.75	
Fructose[a]	73;103;217	39.18	1.75	1860	40–0.1	$y=0.217x-0.307$	0.999	0.05	5.26	$y=0.299x+0.007$	0.998	0.05	3.66	
Galactose[a]	73;205;319	40.17	1.78	1887	40–0.5	$y=0.156x-0.13$	0.999	0.05	2.55	$y=0.349x+0.008$	0.996	0.1	3.45	
Glucose[a]	73;147;205	40.93	1.75	1917	40–0.5	$y=0.657x-0.12$	0.996	0.05	4.76	$y=0.475x+0.043$	0.997	0.1	2.86	
Mannitol	73;319;205	40.9	1.74	1924	40–0.1	$y=0.346x-0.136$	0.998	0.05	3.46	$y=0.450x+0.013$	0.998	0.1	5.06	
Tyrosine	218;73;280	41.25	2.72	1931	40–5	$y=0.176x-0.2$	0.999	0.05	6.54	$y=0.162x+0.018$	0.999	0.1	4.97	
Myo-inositol	73;305;217	44.76	0.64	2081	40–0.1	$y=0.293x-0.1$	0.997	0.05	3.89	$y=0.284x+0.019$	0.994	0.05	2.79	

[a] Aldose.

3. Results and discussion

The goal of the study was to evaluate the potential of a parallel dual secondary column-dual detection GC × GC configuration for metabolic profiling and fingerprinting of urine samples obtained from a murine model of diet-induced metabolic derangements. Specifically, the investigation focuses on: (a) the reliability of comparative untargeted fingerprinting results obtained by a well established pattern recognition data elaboration approach, i.e., template matching based on *peak-region* features [11,15,16]; (b) the consistency of quantitative results of a targeted fingerprinting, performed on the two series of data deriving from parallel ^{2}D separation/detection, and finally (c) the possibility to reliably monitor changes in individuals' fingerprints through the kinetics of experimentation.

The following sections describe the investigation strategy adopted and critically discuss experimental results.

3.1. Untargeted fingerprinting based on peak-region features

Complex patterns, deriving from GC × GC separations, need suitable data elaboration to exploit fully their information potential. For this purpose, different approaches have been proposed and some of them integrated in commercial software for data treatment [6,11,16,20–23].

One method, inspired by pattern recognition procedures, adopts a template of non-targeted (or targeted) *peak features* for data alignment across samples. This approach, known as *template matching fingerprinting*, was developed in 2009 by Reichenbach et al. [24] and successively improved, including MS signature information for peak matching. The template records a prototypical pattern of peaks with associated metadata (MS signature, diagnostic ions, chemical identities, compound-group membership, etc.) extracted from a reference sample(s). The template then is matched to the detected peaks in subsequent chromatograms and metadata are copied from the template to identify the corresponding peaks. The matching algorithm can compensate for retention time shifts and pattern distortions by determining the geometric transformation in the retention-times plane that best fits the expected peak pattern in the template to the target chromatograms.

For multiple chromatograms, automated template matching can be performed by dedicated software and a *composite template* (collecting all peak features reliably matching across all chromatograms of the set) can be built up for comparative analysis, including fingerprinting [15,16]. Although straightforward and intuitive, this approach encounters some limits when applied to very complex samples due to the difficulty of treating peak detection errors and/or the inherent ambiguity of matching. For example, trace peaks may be detected in some samples but not in others, coeluting analytes may be resolved in some chromatograms but not in others, etc.

A *peak-region features* approach to overcome these challenges has been developed and validated over different applications including breast cancer metabolomics [11] and bio-oils characterization [25]. This approach attempts to define one region (i.e., a small 2D retention-times window) per peak over the chromatographic plane to achieve the one-feature-to-one-analyte selectivity of peak features methods but with the implicit matching of region features.

Briefly, this approach: (a) detects and records the peak patterns in individual chromatograms, (b) fixes a few peaks (named registration peaks) that can be reliably matched across samples, (c) aligns and combines the sample chromatograms to create a composite chromatogram, and (d) defines a pattern of region features from the peaks detected (using the Watershed algorithm [25]) in the composite chromatogram. Then, when a target chromatogram is analyzed, (e) the *registration peaks* are matched to target chromatogram pattern, the feature regions are aligned relative to those peaks, and the characteristics of those features are computed to create a feature vector for the target chromatogram, and finally (*f*) the feature vector is used for cross-sample analysis (e.g., classification, discriminant analysis, clustering, etc.).

Fig. 1 illustrates the data elaboration work-flow and Fig. 2 shows results from this sequence of operations on a *control* mouse urine sample. Fig. 2A shows *registration peaks* (indicated by circles) that were determined to be reliably matched across a set of 16 chromatograms (MS and FID signals were separately elaborated) of urine samples collected at different experiment times (4 basal, 4 W6, 4 W9, 4 W12). The 97 reliable peaks (*registration peaks*) which were found to be reliably matched across these chromatograms were used to align the chromatograms and to form a *composite chromatogram* for each channel (MS and FID) [11,26]. A portion of MS total ion current (TIC) of the composite chromatogram of control samples is shown in Fig. 2A–C. Fig. 2B shows the peak-regions (delineated with gray and blue outlines and shaded) formed from the peaks of the *composite chromatogram*, with one region for each detected peak. The template of reliable peaks and composite peak-regions areas (*peak-region reliable template*) obtained by processing the complete set of 128 chromatograms was thus used to align peak-region features and extract comparable information for classification purposes.

The two detection modes generate two distinct data matrices for untargeted fingerprinting. The two fingerprints for each sample, one for each detector mode, record information (selected among different options) about its peak-region features, including retention times in both dimensions, 2D volumes, percent response, and MS fragmentation patterns. The peak-region features of the two detector fingerprints were aligned by cross-matching the *peak-region reliable templates*, thus establishing correspondences between the retention times (location) of the chromatographic areas (regions) and their unique area identifiers (ID#). The complete data matrix is visualized as heat-map in supplementary Fig. 3 (SF3).

Supplementary material related to this article can be found, in the online version, at http://dx.doi.org/10.1016/j.chroma.2014.08.015.

Table 2 reports some of the reliable *peak-regions* that showed the largest *F* values obtained by analysis of variance (ANOVA) on normalized 2D peak-region volumes. The criterion adopted for features selection excluded those with a $F_{calc} < F_{crit}$; with $F_{crit}(1, 6) = 5.99$ for $\alpha = 0.5$. Fig. 3 shows a graphical summary of the discriminant potential of the selected features for MS (3A) and FID (3C) channels, bubble dimensions corresponds to *F* values.

Table 2 lists the 20 peak-regions with the largest F_{calc} values for both down-regulated analytes and up-regulated analytes for both detectors, where down-regulated analytes have a negative difference between the diet and control samples and up-regulated analytes have a positive difference between the diet and control samples. In the present study, because of the limited number of analyzed samples for each class, ANOVA was considered sufficiently reliable to select informative variables. For larger numbers of samples, as well as larger numbers of classes/groups to be compared, supervised multivariate analysis (MVA) methods, such as Soft Independent Modeling of Class Analogies (SIMCA) or partial least squares-discriminant analysis (PLS-DA), might be preferable [27].

The accuracy of fingerprinting based on data from the two detection channels was confirmed by comparing discriminant features rankings. Although these detectors show different analyte response factors and are characterized by different dynamic ranges, the results were coherent. In addition, fingerprinting on FID signals was expected to be affected by higher-rates of false-positive matches due to the lower specificity of the matching algorithm that does

Table 2
List of the 20 *peak-regions* with the largest *F* values on normalized 2D peak-region volumes ($F_{calc} > F_{crit}$ (1, 6) > 5.99 for $\alpha = 0.5$) for both down-regulated and up-regulated analytes. Features are reported together with retention times in the two dimensions (^{1}D Rt (min) and ^{2}D Rt (s)), ranking, 2D absolute volume mean difference (diet *vs.* control), identified metabolite, and HMDB database identifier (IDc). Note that the correspondences between MS and FID features are indicated by the ranking columns; features on the same line do not necessarily correspond.

ID#	MS signal data						HMDB ID[c]	FID signal data				
	^{1}D Rt (min)	^{2}D Rt (s)	Ranking		Mean difference (diet *vs.* control)	Metabolite (I^T)		Ranking		^{1}D Rt (min)	^{2}D Rt (s)	Mean difference (diet *vs.* control)
			MS	FID				FID	MS			
Down-regulated features												
1	41.68	2.09	1	1	−17,442,646	–		1	1	41.67	1.97	−902
2	21.97	2.39	2	–	−10,344,515	–		2	–	10.72	1.72	−568
3	33.42	3.11	3	3	−6,149,324	*n*-hexanoylglycine (1635)	HMDB00701	3	3	33.42	3.07	−228
4	32.39	3.56	4	4	−5,719,193	–		4	4	32.36	3.50	−227
5	38.23	2.16	5	6	−4,690,077	Isocitric acid (1812)	HMDB00193	5	6	39.98	3.45	−169
6	39.98	3.55	6	5	−2,653,283	N-Phenylacetylglycine (1881)	HMDB00821	6	5	38.23	2.06	−153
7	29.99	2.82	7	9	−2,364,652	Pyroglutamic acid (1517)	HMDB00267	7	–	50.09	2.05	−147
8	30.35	1.99	8	14	−2,005,449	–		8	9	52.00	2.09	−134
9	52.02	2.28	9	8	−1,755,994	–		9	7	29.98	2.82	−88
10	31.07	1.82	10	13	−1,706,343	Threonic acid (1578)	HMDB00943	10	11	29.76	3.43	−87
11	29.74	3.42	11	10	−1,697,173	3-Methylcrotonylglycine (1508)	HMDB00943	11	12	27.67	3.29	−82
12	27.66	3.26	12	11	−1,610,925	N-Butyrylglycine (1442)	HMDB00808	12	16	34.18	2.85	−75
13	36.19	2.63	13	–	−1,345,204	–		13	10	31.07	1.79	−70
14	14.15	1.77	14	16	−1,316,931	Lactic acid (1056)	HMDB00190	14	8	30.35	1.98	−63
15	33.02	2.16	15	–	−1,146,529	Phenylalanine (1620)	HMDB00159	15	–	36.68	2.28	−57
16	34.18	2.90	16	12	−1,144,392	–		16	14	14.13	1.99	−55
17	35.71	1.78	17	–	−1,113,926	Ribitol (1715)	HMDB00508	17	–	51.45	3.49	−51
18	23.11	1.90	18	–	−1,040,659	Glycine (1300)	HMDB00123	18	–	48.51	2.39	−50
19	41.45	1.98	19	19	−1,037,333	–		19	19	41.45	2.02	−49
20	41.76	1.86	20	–	−1,017,252	–		20	–	44.40	2.62	−48
Up-regulated features												
21	39.49	1.75	1	2	15,097,922	Fructose[a] (1858)	HMDB00660	1	10	21.98	2.49	457
22	33.3	2.00	2	3	6,572,281	Tartaric acid (1632)	HMDB00956	2	1	39.45	1.63	417
23	39.18	1.75	3	–	6,010,050	Fructose[c] (1845)	HMDB00660	3	2	33.30	1.96	169
24	31.59	2.22	4	–	4,437,145	2-Ketoglutaric acid (1603)	HMDB00208	4	5	38.89	2.23	137
25	38.88	2.28	5	4	4,153,796	–		5	17	39.14	1.67	127
26	39.85	1.76	6	6	3,416,896	Glucose[b] (1880)	HMDB00122	6	6	39.78	1.65	123
27	40.95	1.72	7	8	2,362,473	Glucose[a] (1917)	HMDB00122	7	16	31.56	2.15	109
28	17.03	2.08	8	–	1,658,476	–		8	7	40.93	1.62	74
29	34.25	2.12	9	13	755,287	Adipic acid (1664)	HMDB00448	9	–	52.80	1.77	33
30	22.16	2.25	10	1	741,237	–		10	–	55.29	1.83	31
31	41.54	2.24	11	–	676,510	–		11	20	12.46	2.04	31
32	39.74	1.81	12	17	668,625	Galactose[b] (1870)	HMDB00143	12	15	41.09	1.63	27
33	22.65	1.85	13	14	521,987	–		13	9	34.01	2.18	27
34	23.35	2.14	14	–	411,050	Succinic acid (1310)	HMDB00254	14	13	22.49	2.22	25
35	31.54	2.07	15	7	392,132	α-Hydroxyglutaric acid (1568)	HMDB00694	15	–	59.84	1.91	17
36	41.11	1.75	16	12	392,132	Mannitol (1924)	HMDB00765	16	–	57.60	1.86	15
37	39.50	1.74	17	5	375,012	Tyrosine (1931)	HMDB00158	17	12	40.11	1.67	15
38	28.93	2.04	18	–	288,346	Malic acid (1481)	HMDB00744	18	–	31.53	2.04	13
39	18.38	1.98	19	–	255,055	Isobutyric acid (1175)	HMDB01873	19	–	65.82	2.51	12
40	12.18	1.73	20	11	238,816			20	–	40.51	2.17	6

[a] Aldose.
[b] Pyranose.
[c] Furanose (S. Bouatra, F. Aziat, R. Mandal, A.C. Guo, M.R. Wilson, et al., The human urine metabolome. PLoS ONE 8 (9) (2013) e73076, doi:10.1371/journal.pone.0073076).

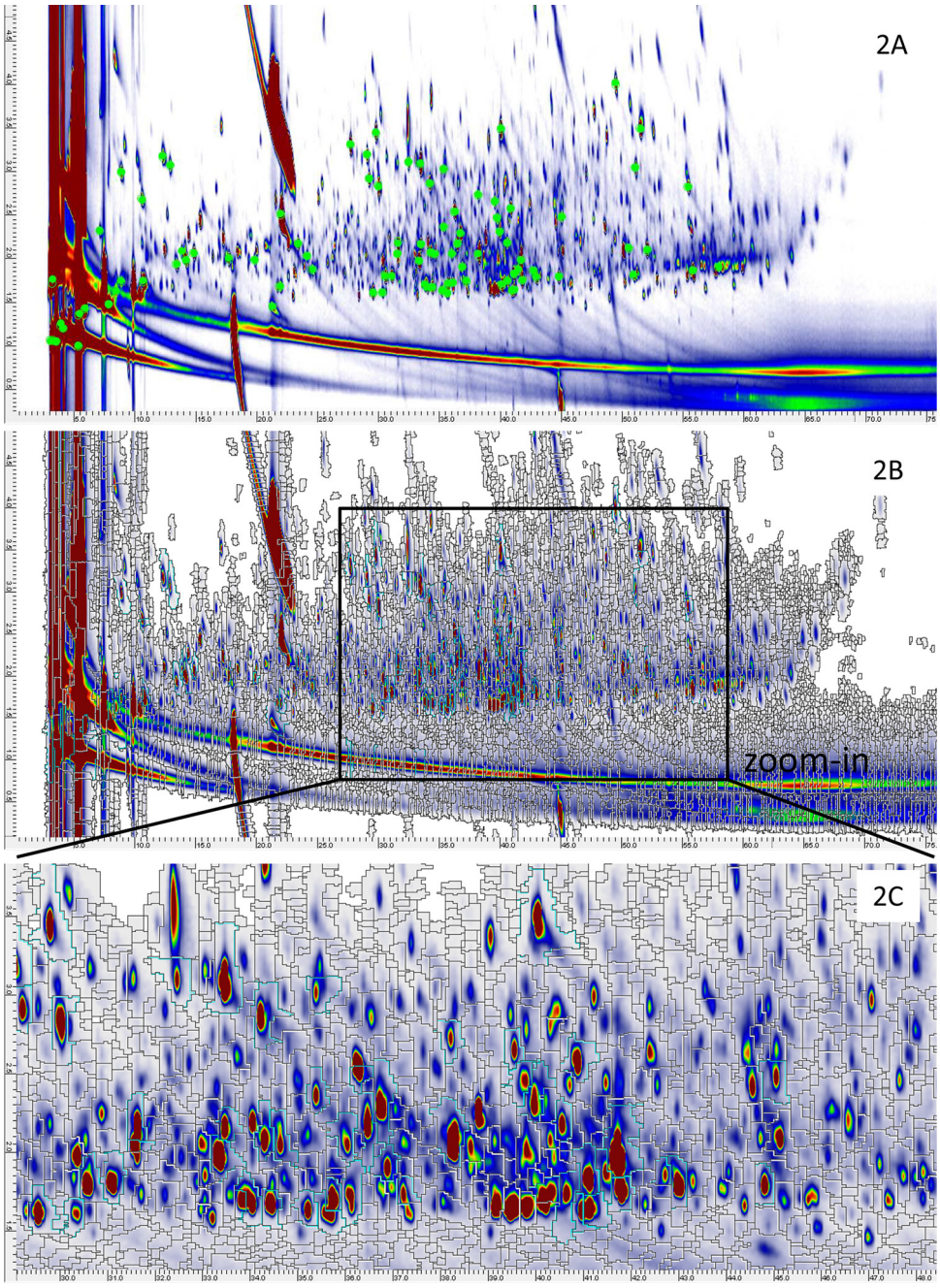

Fig. 2. A pseudocolor image of the MS total ion current (TIC) of the cumulative chromatogram of urine samples collected from four mice at different experiment times (4 basal, 4 W6, 4 W9, 4 W12). (A) *Registration peaks* that reliably matched across chromatograms are indicated by circles. (B) Peak-region areas are delineated by gray outlines (or blue outlines for peak regions of registration peaks) and shaded. (C) A portion of the chromatogram with peak-regions. The feature template consists of the registration peaks and the peak-regions. The registration peaks are used to align the template to the chromatograms and the peak-regions are used characterize features across chromatograms. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this article.)

not include the third dimension of information (i.e., the MS spectrum). The optimization of the bi-dimensional separation in terms of resolution, separation space used, and of 2D column loadability obtained by adopting the parallel dual ^{2}D columns configuration was here the key-factor that minimized matching errors.

The most informative MS *peak-regions* reported in Table 2 found reliable correspondences within those selected by applying the same criteria to FID results. In the "ranking" columns of Table 2 features are reported according with descending order of *absolute mean difference* for the specific detector and the corresponding position in the ranking on the other detector. For example, the peak region #2 at 21.97 min and 2.39 s, ranked as 2nd for the MS trace did not found any correspondence within the first 20 most informative peak-regions of the FID channel. Conversely, feature #4 at 32.39 min and 3.56 s had the same ranking position on the two detection channels. It is noted that small variations of retention times between the two detectors are reasonably due to the different acquisition frequencies of the two channels and on the related resolution. In particular, cross-matching between detectors covers up to 70% within this set of selected features.

Some discrepancies in features ranking between MS and FID were expected because of the different operative principles of the two detectors and, as a consequence, of their analyte(s) response function(s). This aspect positively affects the robustness of the

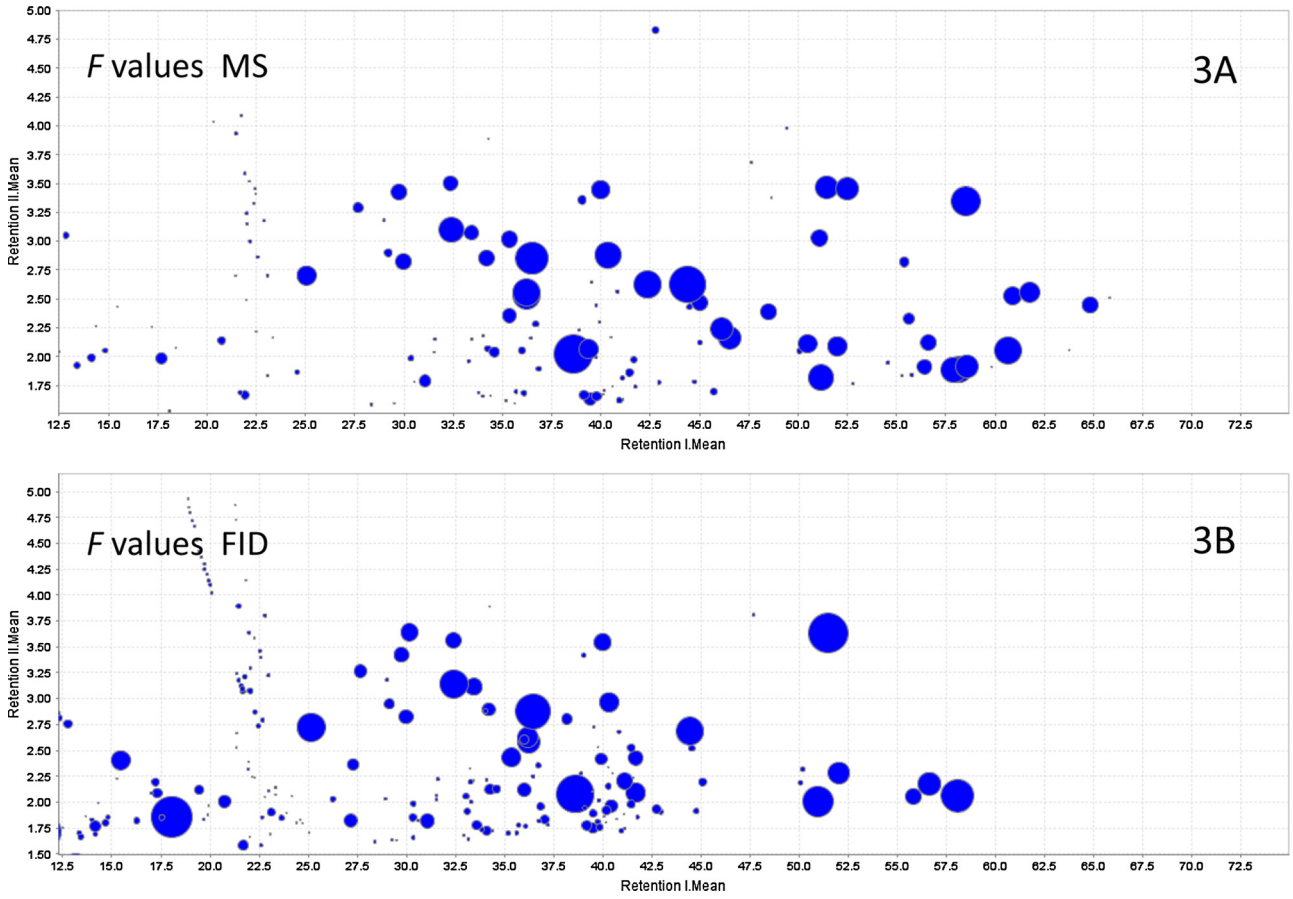

Fig. 3. Graphical summary of discriminant potential of features for MS (A) and FID (B) channels based on *F* values.

results in terms of feature selection because it avoids *a priori* exclusion of potentially informative features whose detector-specific response function is flatter, and consequently smaller 2D Volume variance. In addition, it is noteworthy that the consistency would be expected to increase with more samples, which would provide more reliable composite chromatograms.

From these results, relevant peak features were identified on the basis of the MS fragmentation pattern and linear retention index (I^T) in the first dimension and/or reference standard confirmation, and a subset of them was submitted to quantitative determination.

The following section presents targeted analysis results.

3.2. Targeted quantitative fingerprinting

Quantitation was based on an external standard calibration *versus* internal standard (i.e., gallic acid) normalization. Analytes submitted to quantitative metabolomics were selected from untargeted fingerprinting results and from a careful survey of the scientific literature within those metabolic markers with relevant information potential. It was expected that untargeted profiling would produce some false-positives due to the small number of individuals considered, the analytes-specific response factors, and the detector(s) dynamic range. Inconsistencies might be addressed by true quantitation of target metabolites.

Results on method performance parameters for both detection channels are summarized in Table 1. Linearity and precision were comparable between the two detectors with slightly better performances for FID. On the other hand, LOQ values were lower for MS, thanks to the possibility of isolating the response of diagnostic ions with relatively high *m*/*z* values. This parameter, although not crucial for these application, is here important because the actual system sensitivity is halved compared to a single detection channel configuration, due to the effluent splitting in the two second-dimension columns.

Quantitation was performed on the entire set of samples and the results for the week 12 samples are summarized in Table 3 for the control and HFHS diet groups. The accuracy of quantitative data was verified by regression analysis by computing quantitative results for MS as the independent variable (x) and for FID as dependent variable (y). The correlation coefficient (R^2) was 0.998 and indicated good consistency

Quantitative results, expressed as mg/L, are reported as the median of sixteen measurements per class (eight biological and two analytical replicates) for control mice and for HFHS diet mice at the end of the experimentation (i.e., week 12 – W12) where diet induced damage were expected to be relevant. The % of variation refers to the mean analyte's relative increment/decrement *versus* the control level. Arrows facilitate numerical data interpretation.

Some analytes, as expected and consistent with fingerprinting results, markedly increased their concentration in the HFHS diet group, such as some sugars and polyalcohols, including fructose (+522%), glucose (+79%), and mannitol (+79%). Glycerol was the metabolite that reported the highest variation. These changes reflect the high sugar and fat content in the diet, and their behavior finds confirmation in available data collected in reference databases (Human Metabolome Database Version 3.6).

On the contrary, levels of several amino acids were drastically decreased, including alanine (−36%), threonine (−23%), and glycine (−97%). This decrement in amino acids indicates decreased excretion, possibly, reflecting their increased clearance in hepatic

Table 3
Results, expressed as mg/L of urine, of targeted quantitative fingerprinting for *control* and *HFHS diet* groups at 12 weeks of experimentation. Data is reported as the median of eight measurements (four biological and two analytical replicates). The % of variation refers to single analyte relative increment/decrement *versus* the control level.

Analyte	Controls (mg/L) week 12				HFHS diet week 12				% of variation	Trend
	Median	Min	Max	n[a]	Median	Min	Max	n[a]		
Pyruvic acid	1.89	1.46	4.00	8	8.35	7.1	18.18	8	342	↑↑↑
Lactic acid	2.63	1.86	7.07	8	3.68	2.21	6.39	8	40	↑
Alanine	0.31	0.04	0.58	8	0.2	0.01	0.6	8	−36	↓
Malonic acid	4.41	0.60	8.20	8	1.02	0.93	1.05	8	−77	↓
Valine	0.07	0.01	0.12	8	0.17	0.01	0.03	8	155	↑↑
Glycerol	1.30	1.10	1.40	8	15.87	13.09	33.73	8	1117	↑↑↑↑
Glycine	0.98	0.00	1.96	8	0.03	0.02	0.04	8	−97	↓
Succinic acid	2.52	0.46	2.87	8	0.52	0.52	1.39	8	−79	↓
Threonine	0.29	0.27	1.09	8	0.22	0.12	1.95	8	−23	↓
Malic acid	0.44	0.18	0.79	8	0.43	0.23	0.91	8	−1	↔
Creatinine	3.47	3.06	3.80	8	5.26	5.12	5.95	8	−63	↑
2-Ketoglutaric acid	4.81	3.44	8.09	8	6.71	4.4	7.17	8	39	↑
Phenylalanine	5.22	2.42	8.90	8	5.58	2.98	7.58	8	7	↔
Xylitol	1.35	0.98	2.26	8	1.30	1.28	6.94	8	−3	↔
Ribitol	2.80	1.07	5.04	8	3.91	3.04	6.65	8	40	↑
Hippuric acid	8.04	7.90	8.54	8	–	–	–	8	–	-
Fructose	0.78	0.74	1.34	8	4.85	4.81	14.25	8	522	↑↑↑
Galactose	1.61	1.25	2.79	8	2.13	2.1	2.8	8	32	↑
Glucose	1.22	1.03	2.38	8	2.18	2.07	2.75	8	79	↑
Tyrosine	2.24	2.19	2.74	8	2.46	2.44	4.94	8	10	↔
Mannitol	1.22	1.06	3.68	8	2.18	2.07	5.33	8	79	↑
Myo-inositol	0.62	0.56	1.46	8	0.73	0.68	1.9	8	18	↑

[a] Four mice and two analytical replicates.

gluconeogenesis. However, levels of other amino acids, such as phenylalanine and tyrosine remained invariant, suggesting that the perturbations in amino acid metabolism were more complex than simply an increase in gluconeogenesis. Some acids were up-regulated after dietary impairment, including pyruvic acid (+342%), lactic acid (+40%), and 2-ketoglutaric acid (+39%). Although the majority of these biomarkers are implicated in known diabetic processes, others have not yet been reported as possible biomarkers for diabetes. Thus, biological validation studies are necessary to determine the biological reproducibility and soundness of these initial findings from a small non-targeted metabolomic study.

3.3. Individual kinetics of urine fingerprint

An additional interesting aspect investigated in this study was the possibility of monitoring individual variations of feature patterns during experimentation. A specific application of such metabolomics is personalized medicine [5]. The aim is to monitor subtle changes in the metabolic fine print of individual patients to subsequently adapt pharmacological protocols or dietary interventions as a function of the specific responses/reactions.

Untargeted fingerprinting and targeted quantitative fingerprinting (Table 3) results showed large ranges of variations, within homogeneous groups (controls and HFSC diet), in the features'

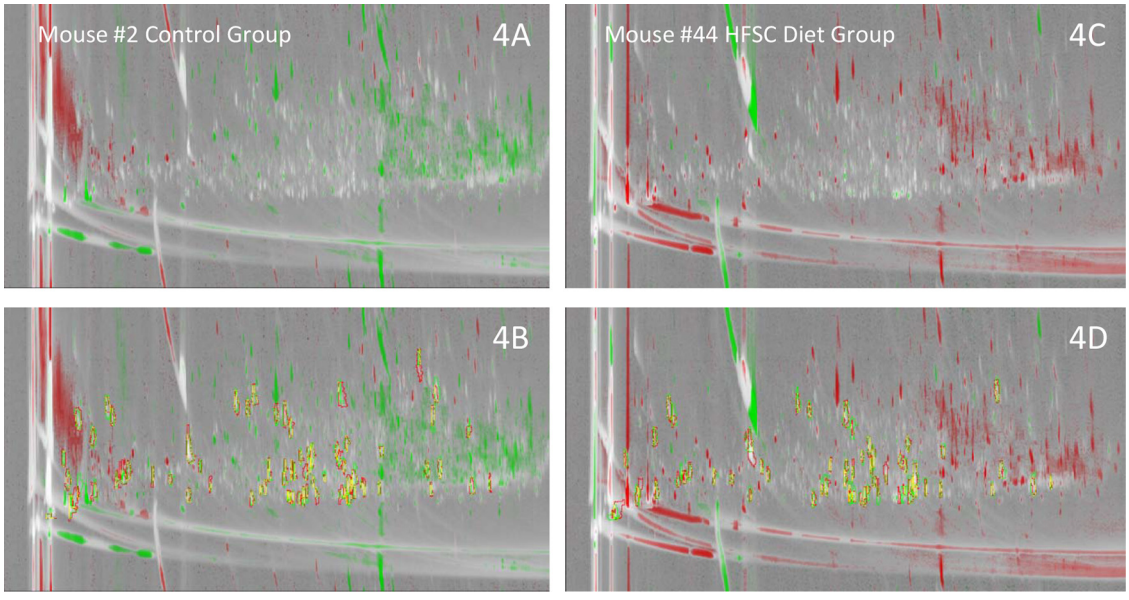

Fig. 4. Comparative visualization of W1-basal (*reference*) and W9 (*analyzed*) urine samples of mouse #2 (A and B) belonging to the control group and mouse #44 (C and D) belonging to the HSHF diet group. Specific template peaks-regions (tailored template peak regions) are yellow highlighted in B and D. Features that reported the highest variation between individuals as a function of the diet also are reported together with tentative identification (based on EI-MS spectrum similarity and LRI), ^{1}D and ^{2}D retention times, 2D normalized response at W1-basal and W9 time course and % of variation. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this article.)

quantitative descriptors (2D peak volumes, 2D peak-region volumes, etc.). This variability for small metabolites among different subjects, whichever biological fluid is under study, is expected because of well-known inter-individual differences due to behavioral, genotypic, and phenotypic factors, and necessitates data from many metabolic snapshots. The intra-individual variability that can be detected in a single subject during the progression of the disease then could be used for pathological staging of the disease and possibly to evaluate therapeutic efficacy.

In this context, the possibility of using advanced (reliable) fingerprinting of urine metabolic patterns for individuals to promptly monitor the occurrence and/or extent of impairment is exciting. The investigation strategy proposed here enables diagnostic fingerprints that might be used to classify individual patterns and/or establish the degree of a dysfunction. Thus, urine samples from two individuals, a control (mouse #2) and a HFSC diet group (mouse #44) mouse, were collected at 1 week (basal) and after 6, 9, and 12 weeks (W6, W9, and W12) over a 16-h period, placing the animals in metabolic cages. Urine samples then were spun in microcentrifuge tubes at 13,200 rpm for 1 min to pellet any contaminating particles. The urine was then pipetted off the particle pellet and stored at −80 C. Urine samples then were submitted to advanced fingerprinting by adopting a "tailored" peak-regions template obtained by extracting informative features from the complete data matrix including the most up- and down-regulated features shown in Table 2. The tailored template was applied to 2D chromatograms of urine samples collected at different times and the corresponding peak-regions were aligned across patterns for comparative purposes.

Subsequently, comparative image visualizations between basal (*reference*) urine and W9 (*analyzed*) samples of mice #2 and #44 were constructed and tailored template peaks distribution observed. The comparative visualization consisted of arithmetic subtraction of a *sample* (or *analyzed*) 2D-chromatogram from a *reference* to reveal differences in the chemical pattern. For a reliable visual comparison, corresponding *peaks* from 2D chromatogram pairs were aligned and normalized in terms of peak-region response [28]. Results are shown in Fig. 4A and B for the control mouse (#2) and C and D for the HFSC diet mouse (#44), tailored template peaks are outlined in Fig. 4B and D.

In the visual comparison of Fig. 4, the *colorized fuzzy difference* [28] visualization uses the Hue-Intensity-Saturation (HIS) color space to color each pixel in the retention-times plane. The method first computes the difference at each data point. The pixel hue is set to green if the difference is positive and red if it is negative (i.e., green if the value in the W9 chromatogram is smaller than the value in the W1 chromatogram and red if the value is greater). The pixel intensity is set to the larger of the two values (scaled by the largest intensity value in the chromatogram) and the pixel saturation is set to the magnitude of the difference between the data points (scaled by the magnitude of the largest difference in the chromatogram). Peaks are visible because large-valued data points yield bright pixels and small-valued data points yield dark pixels. If the difference is relatively large, the color is saturated with red or green (depending on the largest data point); if the difference is relatively small, the color saturation is low, producing a grey level from black to white depending on intensity. Peaks with large differences therefore appear red or green and peaks with small differences appear white or grey. The fuzzy difference is computed as the difference between a data point and a small region of data points in the other chromatogram (scaled by the largest fuzzy difference in the chromatogram). Thus, the colors are saturated with red or green only when the relative difference is large. Fig. 4B and D includes outlines of the tailored template peaks. Table 4 details most informative tailored template peaks and their % of Variation, calculated as the difference between 2D peak-region volume at the beginning of the

Table 4
Features that reported the highest variation between individuals as a function of the diet together with tentative identification (based on EI-MS spectrum similarity and LRI), ^{1}D and ^{2}D retention times, 2D Normalized response at basal and W9 time course and % of Variation.

Analyte	Rat #2 control group							Rat #44 HFSC diet group						
	^{1}D Rt (min)	^{2}D Rt (s)	W9	Basal	Absolute difference	% normalized difference	Trend	^{1}D Rt (min)	^{2}D Rt (s)	W9	Basal	Absolute difference	% normalized difference	Trend
2-Methyl butanal	3.92	1.23	857	3730	−2873	−77	↓	3.92	1.24	3769	503	3266	+649	↑↑↑
Dimethylethanolamine	8.83	1.74	106	33	73	+223	↑↑	8.83	1.73	30	96	−67	−69	↓
Glycerol	21.92	1.67	47	44	3	+7	↑	21.92	1.67	51	38	13	+36	↑
N-Isovaleroylglycine	29.00	3.19	40	21	19	+89	↑	29.00	3.19	80	33	47	+144	↑↑
Erythritol	29.50	1.61	84	107	−23	−21	↓	29.50	1.61	286	140	146	+104	↑↑
Threonic acid	30.68	1.79	60	77	−17	−22	↓	30.50	1.79	243	101	141	+140	↑↑
Tartaric acid	33.25	1.97	357	111	247	+223	↑↑	33.33	1.97	1617	302	1315	+436	↑↑↑
p-Hydroxyphenylacetic acid	33.42	2.17	68	48	21	+44	↑	33.42	2.16	144	45	99	+217	↑↑
Xylose	33.75	1.70	32	44	−12	−28	↓	33.75	1.69	154	43	111	+255	↑↑
Fructose	39.18	1.75	63	71	−8	−11	↓	39.15	1.76	146	75	71	+94	↑
Glucose	40.93	1.77	104	178	−74	−41	↓	40.88	1.75	551	178	373	+210	↑↑

experiments W1 (i.e., basal) and the corresponding value at W9 and then normalized *versus* the basal level. As can be seen from experimental data reported in Fig. 4, most of the selected features are up-regulated after dietary manipulation, in line with quantitative fingerprinting results (e.g., fructose, glucose and glycerol), but several others that fall outside the ranking of Table 2, are informative of the biological phenomenon under study because of their up-regulation in mouse #44. Among others, erythritol, threonic acid, *p*-hydroxy phenylacetic acid, and xylose were not evidenced as discriminant markers of the population under study, but showed remarkable differences in certain individuals.

Percent of variation values, estimated for raw data, were between – 40 and 400%, with the exception of 2-methyl butanal (+650% in mouse #44 after 12 weeks of HSHF diet), which was reported here as secondary product of Maillard reaction whose extent is clearly much more relevant in individuals with a higher level of blood sugars [7,29]. These metabolites deserve further investigation to assess their role in homeostasis.

4. Conclusions

The advantages of a parallel dual secondary column-dual detection GC × GC system in an integrated platform for urine metabolites profiling have been discussed together with some practical aspects concerning data elaboration strategies that enabled a cross-validation of untargeted fingerprinting results for relevant biomarker discovery.

In addition, the experimental conditions produced consistent separation patterns from both detectors in both dimensions that, working in close-to-optimal ^{2}D linear velocities and a doubled secondary column loading capacity, showed positive effects on overall system orthogonality, resolution, and fingerprinting accuracy.

The reliability of the untargeted fingerprinting results has been confirmed by quantitative results on selected relevant features that showed % of variations consistent with those observed by comparing raw data quantitative descriptors (2D peak-region volumes and percent of response).

Last but not least, advanced fingerprinting also demonstrated its effectiveness in the evaluation of individual variations during experiments, thus representing a potentially valuable tool for personalized intervention studies. In this context, it is interesting to observe that informative features that were not discriminant for the entire population can become relevant for single individuals.

Acknowledgement

This study was supported by Ricerca Finanziata da Università – Fondo per la Ricerca Locale (Ex 60%) Anno 2013.

References

[1] W.B. Dunn, N.J. Bailey, H.E. Johnson, Measuring the metabolome: current analytical technologies, Analyst 130 (2005) 606–625.
[2] N. Friedrich, Metabolomics in diabetes research, J. Endocrinol. 215 (2012) 29–42.
[3] A. McKillop, Emerging applications of metabolomic and genomic profiling in diabetic clinical medicine, Diabetes Care 34 (2011) 2624–2630.
[4] P. Hunter, Reading the metabolic fine print. The application of metabolomics to diagnostics, drug research and nutrition might be integral to improved health and personalized medicine, EMBO Rep. 10 (2009) 20–23.
[5] M.F. Almstetter, P.J. Oefner, K. Dettmer, Comprehensive two-dimensional gas chromatography in metabolomics, Anal. Bioanal. Chem. 402 (2012) 1993–2013.
[6] M.M. Koek, B. Muilwijk, L.L. van Stee, T. Hankemeier, Higher mass loadability in comprehensive two-dimensional gas chromatography–mass spectrometry for improved analytical performance in metabolomics analysis, J. Chromatogr. A 1186 (2008) 420–429.
[7] S.M. Rocha, M. Caldeira, J. Carrola, M. Santos, N. Cruz, I.F. Duarte, Exploring the human urine metabolomic potentialities by comprehensive two-dimensional gas chromatography coupled to time of flight mass spectrometry, J. Chromatogr. A 1252 (2012) 155–163.
[8] G. Purcaro, P. Q- Tranchida, C. Ragonese, P. Dugo, L. Conte, P. Dugo, G. Dugo, L. Mondello, Evaluation of a rapid-scanning quadrupole mass spectrometer in an apolar × ionic-liquid comprehensive two-dimensional gas chromatography system, Anal. Chem. 82 (2010) 8583–8590.
[9] B.J. Gonçalves Silva, P.Q. Tranchida, G. Purcaro, M.E. Costa Queiroz, L. Mondello, F.M. Lanças, Evaluation of comprehensive two-dimensional gas chromatography coupled to rapid scanning quadrupole mass spectrometry for quantitative analysis, J. Chromatogr. A 1255 (2012) 177–183.
[10] P.Q. Tranchida, P. Donato, F. Cacciola, M. Beccaria, P. Dugo, L. Mondello, Potential of comprehensive chromatography in food analysis, TrAC 52 (2013) 186–205.
[11] S. Reichenbach, X. Tian, Q. Tao, E. Ledford, Z. Wu, O. Fiehn, Informatics for cross-sample analysis with comprehensive two-dimensional gas chromatography and high-resolution mass spectrometry (GC × GC–HRMS), Talanta 83 (4) (2011) 1279–1288.
[12] M. Collino, M. Aragno, S. Castiglia, G. Miglio, C. Tomasinelli, G. Boccuzzi, C. Thiemermann, R. Fantozzi, Pioglitazone improves lipid and insulin levels in overweight rats on a high cholesterol and fructose diet by decreasing hepatic inflammation, Br. J. Pharmacol. 160 (2010) 1892–1902.
[13] M. Collino, E. Benetti, M. Rogazzo, R. Mastrocola, M.M. Yaqoob, M. Aragno, C. Thiemermann, R. Fantozzi, Reversal of the deleterious effects of chronic dietary HFCS-55 intake by PPAR-δ agonism correlates with impaired NLRP3 inflammasome activation, Biochem. Pharmacol. 85 (2013) 257–264.
[14] L. Nicolotti, C. Cordero, D. Bressanello, C. Cagliero, E. Liberto, F. Magagna, P. Rubiolo, B. Sgorbini, C. Bicchi, Parallel dual secondary column-dual detection: a further way of enhancing the informative potential of two-dimensional comprehensive gas chromatography, J. Chromatogr. A (2014), http://dx.doi.org/10.1016/j.chroma.2014.07.081 (in press).
[15] C. Cordero, E. Liberto, C. Bicchi, P. Rubiolo, S.E. Reichenbach, X. Tian, Q. Tao, Targeted and non-targeted approaches for complex natural sample profiling by GC × GC–qMS, J. Chromatogr. Sci. 48 (2010) 251–261.
[16] S.E. Reichenbach, X. Tian, C. Cordero, Q. Tao, Features for non-targeted cross-sample analysis with comprehensive two-dimensional chromatography, J. Chromatogr. A 1226 (2012) 140–148.
[17] Q. Zhang, G. Wang, Y. Du, L. Zhu, A. Jiye, GC/MS analysis of the rat urine for metabonomic research, J. Chromatogr. B 854 (2007) 20–25.
[18] R. Shellie, P. Marriott, P. Morrison, L. Mondello, Effects of pressure drop on absolute retention matching in comprehensive two-dimensional gas chromatography, J. Sep. Sci. 27 (2004) 504–512.
[19] J. Beens, H.G. Janssen, M. Adahchour, U.A.Th. Brinkman, Flow regime at ambient outlet pressure and its influence in comprehensive two-dimensional gas chromatography, J. Chromatogr. A 1086 (2005) 141–150.
[20] K.M. Pierce, J.C. Hoggard, R.E. Mohler, R.E. Synovec, Recent advancements in comprehensive two-dimensional separations with chemometrics, J. Chromatogr. A 1184 (2008) 341–352.
[21] K.M. Pierce, B. Kehimkar, L.C. Marney, J.C. Hoggard, R.E. Synovec, Review of chemometric analysis techniques for comprehensive two dimensional separations data, J. Chromatogr. A 1255 (2012) 3–11.
[22] S. Castillo, I. Mattila, J. Miettinen, M. Orešič, T. Hyötyläinen, Data analysis tool for comprehensive two-dimensional gas chromatography/time-of-flight mass spectrometry, Anal. Chem. 83 (8) (2011) 3058–3067.
[23] T. Hyötyläinen, Novel methodologies in metabolic profiling with a focus on molecular diagnostic applications, Expert Rev. Mol. Diagn. 12 (2012) 527–538.
[24] S.E. Reichenbach, P.W. Carr, D.R. Stoll, Q. Tao, Smart templates for peak. Two-dimensional liquid chromatography, J. Chromatogr. A 1216 (2009) 3458–3466.
[25] S.E. Reichenbach, M. Ni, V. Kottapalli, A. Visvanathan, Information technologies for comprehensive two-dimensional gas chromatography, J. Chromatogr. A 71 (2004) 107–120.
[26] S.E. Reichenbach, X. Tian, A.A. Boateng, C.A. Mullen, C. Cordero, Q. Tao, Reliable peak selection for multisample analysis with comprehensive two-dimensional chromatography, Anal. Chem. 85 (2013) 4974–4981.
[27] R.G. Brereton, Applied Chemometrics for Scientists, Wiley West Sussex, UK, 2007.
[28] B.V. Hollingsworth, S.E. Reichenbach, Q.P. Tao, A. Visvanathan, Comparative visualization for comprehensive two-dimensional gas chromatography, J. Chromatogr. A 1105 (2006) 51.
[29] A. Lapolla, D. Fedele, R. Seraglla, P. Traldi, The role of mass spectrometry in the study of non-enzymatic protein glycation in diabetes: an update, Mass Spec. Rev. 25 (5) (2006) 775–797.