African Journal of Microbiology Research Vol. 6(13), pp. 3159-3172, 9 April, 2012
Available online at http://www.academicjournals.org/AJMR
DOI: 10.5897/AJMR11.1269
ISSN 1996-0808 ©2012 Academic Journals

Full Length Research Paper

A comparison of two extraction methods for differentiating and characterising various *Mycobacterium* species and *Pseudomonas aeruginosa* using GC-MS metabolomics

Ilse Olivier and Du Toit Loots*

School for Physical and Chemical sciences, Centre for Human Metabonomics, North-West University, Potchefstroom, 2520, South Africa.

Accepted 23 February, 2012

We investigated the capacity of a metabolomics research approach to characterise and differentiate between various infectious *Mycobacterium* species and *Pseudomonas aeruginosa*, and compared two extraction procedures; 1) extracting the fatty acid metabolome, and 2) extracting the total metabolome, prior to gas chromatography mass spectrometry (GC-MS) and statistical data analyses. Both extraction procedures, as part of a metabolomics study, were able to successfully differentiate between all bacterial groups investigated. The total metabolome extraction method proved the better of the two methods due to its comparative: simplicity; speed (taking less than 4 h), repeatability; extraction capacity (considering the range of compounds extracted and their relative concentrations), and; ability to extract those compounds which allow a better differentiation and characterisation of the investigated sample groups.

Key words: Tuberculosis, *Mycobacterium*, metabolomics, gas chromatography mass spectrometry (GC-MS), biosignature.

INTRODUCTION

Although tuberculosis (TB) was declared a global public health emergency by the World Health Organization (WHO) almost 20 years ago, it is still considered a major health threat today, with 9.4 million newly diagnosed cases and 1.8 million TB related deaths reported worldwide in 2008 alone (WHO, 2010). Although this is a universal epidemic, Africa (30%) and Asia (55%) account for 85% of all global TB cases, and when evaluating recorded adult deaths in low- and middle-income countries, TB is ranked third, after HIV/AIDS and ischemic heart disease (WHO, 2010).

Although *Mycobacterium tuberculosis* is the primary cause of TB, other mycobacterial strains may also result in opportunistic infections in humans and are frequently encountered in clinical specimens (Miguez-Burbano, 2006). These species include: *Mycobacterium avium*, which has the potential to cause pulmonary infection in patients with chronic lung disease; *Mycobacterium kansasii*, which may cause skin and soft tissue infections, skeletal infections, lung infections, surgical site infections and disseminated disease (Davis, 2007) and; *Mycobacterium bovis*, the organism responsible for causing TB in cattle (Todar, 2005). Additionally, other lung pathogens, such as *Pseudomonas aeruginosa*, a genus from the phylum Proteobacteria and a common cause of iatrogenic infections, may also cause symptoms similar to that of pulmonary TB, when intruding the respiratory passages of patients with immuno-compromising illnesses (Todar, 2005).

The relatively new research field of metabolomics encompasses 'the non-biased identification and quantify-cation of all the metabolites in a biological system', using highly sensitive analytical procedures (Dunn et al., 2005).

*Corresponding author. E-mail: dutoit.loots@nwu.ac.za. Tel: +27 (0)18 299 1818. Fax: + 27 (0)18 299 2316.

Lately, metabolomics has successfully been used as a tool to characterise a variety of disease states, in order to better understand the underlying mechanisms of these diseases (Schoeman and Loots, 2011). Only one metabolomics study to date, using electronic nose metabolomics analyses, was capable of differentiating between various infectious *Mycobacterium* species (*M. tuberculosis*, *M. avium*, and *Mycobacterium scrofulaceum)* and *P. aeruginosa* (Fend et al., 2006). However, this approach, using sensory array detection, isn't capable of identifying those compounds best explaining the variation detected, and hence, no marker metabolites characterising these species were determined.

Consequently, we developed and compared two rapid extraction procedures and investigated their respective capacities to extract those compounds which best differentiate the above mentioned infectious *Mycobacterium species* and *P. aeruginosa*, using a GC-MS metabolomics approach. Given that the fatty acid composition of these bacterial species is a well-know characteristic feature, which may potentially differentiate these groups (Lambert et al., 1986; Mosca et al., 2006), the first extraction procedure developed was aimed at selectively isolating these fatty acids. The second extraction procedure, comparatively investigated, was however aimed at extracting a far greater variety of metabolites, considering that the definition of metabolomics is the "unbiased" identification of "all" metabolites present in a biological system (Dunn et al., 2005). Furthermore, using multivariate statistical methods, those compounds contributing most to the variation between the infectious bacterial groups, as detected via each of the extraction methods investigated, prior to GC-MS analyses, were identified. The biological importance of these compounds was then discussed, in order to further validate their use for metabolomics applications. These potential metabolite makers may be of great scientific value, considering their ability to better characterise these organisms and their potential to be used diagnostically.

MATERIALS AND METHODS

Chemicals and reagents

N-methyl-N-(tert.-butyldimethylsilyl)-trifluoroacetamide (MSTFA) with 1% trimethylchlorosilane (TMCS), pyridine, potassium hydroxide (KOH), methyl-nonadecanoic acid (Me-C19), glacial acetic acid, and trimethylpentane were purchased from Merck (Darmstadt, Germany). Methoxyamine hydrochloride and 3-phenyl butyric acid were purchased from Sigma-Aldrich (St. Louis, Mo., USA). All organic solvents used were ultra purity Burdick & Jackson brands (Honeywell International Inc., Muskegon, MI, USA) and were used without any further purification.

Cultures

All the cultured organism samples used in this study were supplied by the Royal Tropical Institute, Amsterdam, Netherlands. All the bacteria (*M. tuberculosis, M. avium, M. bovis, M. kansasii,* and *P. aeruginosa,* 12 repeats of each*)* were cultured in Middlebrook 7H9 medium with oleic acid-albumin-dextrose-catalase enrichment. The bacteria were incubated at 37°C while shaking at 120-150 rpm until an optical density (420 nm) of 0.30 (≈2×10^8 bacteria/ml) was reached. The cells were washed once with PBS, collected via centrifugation, snap frozen and stored at -80°C until extraction and GC-MS metabolomics analysis. For determining the comparative species differentiation capacities of the two respective extraction methods (fatty acid metabolome and total metabolome extractions), samples were re-suspended in ddH_2O at a concentration of 1×10^8 bacteria/ml, prior to extraction, GC-MS analysis and statistical data processing.

Extraction procedure 1 – Fatty acid metabolome extraction

An extraction solvent mixture (1.25 ml), consisting of chloroform, methanol, and ddH_2O, in a ratio of 1:3:1, respectively, was added to 250 µl of the above mentioned sample suspensions in a microcentrifuge tube. The extraction was performed using an MM 400 vibration mill (Retsch GmbH and co. KG, Haan, Germany) at a frequency of 30 Hz for 5 min with a 3 mm tungsten carbide bead (Retsch GmbH and co. KG) for increased extraction efficiency. Chloroform and ddH_2O (250 µl of each) were subsequently added to the extract, followed by centrifugation for 5 min at 550 x g. The organic phase was collected, transported to a kimax tube, and dried under a light stream of nitrogen. Thereafter, 0.5 ml of chloroform, 0.5 ml of methanol, and 1 ml of methanolic KOH (0.2 M) was added to the dried extract. The mixture was then incubated for 30 min at 60°C followed by the addition of 2 ml of hexane, 200 µl of glacial acetic acid (1 M), and 2 ml of ddH_2O. After centrifugation at 550 x g for 5 min, the organic phase was once again collected and transferred to a new kimax tube. The remaining water phase was re-extracted 3 times with hexane and the combined organic phases from each of the previous extractions were dried under nitrogen and re-suspended in 50 µL of external standard (Me-C19) dissolved in trimethylpentane (15.6 ng/ml).

Extraction procedure 2 - Total metabolome extraction

As an internal standard, 50 µl of 3-phenyl butyric acid (26.25 mg/ml) was added to 250 µl of the above mentioned sample suspensions in a microcentrifuge tube, followed by the addition of 1.25 ml of an extraction solvent mixture containing chloroform, methanol, and ddH_2O, in a 1:3:1 ratio. The extraction was again performed using an MM 400 vibration mill at a frequency of 30 Hz, for 5 min, after adding a 3 mm tungsten carbide bead to each tube for increased extraction efficiency. After centrifugation, the supernatant was collected, transferred to a GC-MS sample vial and dried under a light stream of nitrogen. The dried supernatant was subsequently derivatized with 50 µl of methoxyamine hydrochloride in pyridine (15 mg/ml) at 50°C for 90 min. Hereafter, the extract was trimethylsilylated with 50 µl of MSTFA with 1% TMCS at 50°C for 60 min.

GC-MS parameters

The above mentioned prepared extracts were analysed by injecting 1 µl on an Agilent 7890A gas chromatograph (Agilent, Atlanta, GA) coupled to an Agilent 5975 mass selective detector equipped with a 7683B injector, 7683 auto sampler and VF1-MS capillary column (30 m × 250 µm i.d., 0.25 µm film thickness) in the splitless mode. Helium was used as the carrier gas and the pressure was programmed as such that the helium flow was kept constant at 1.2 ml/min. Compounds was detected using the MS in full scan mode

(*m/z* 50-550 for the fatty acid extracts and *m/z* 60-800 for the total metabolite extracts). For the fatty acid metabolome extracts, the most optimal analytical conditions were an injector temperature of 250°C, for the total run time, and an initial GC oven temperature of 50°C for 1 min followed by a primary increase in oven temperature of 20°C/min to 160°C. This was then followed by an increase in oven temperature of 10°C/min to 190°C and then a further increase of 5°C/min to a final temperature of 300°C, at which the temperature was maintained for 5 min. For the total metabolome extracts, the most optimal analytical conditions were an injector temperature of 270°C, for the total run time, and an initial GC oven temperature of 70°C for 2 min, followed by an increase in oven temperature of 4°C/min to a final temperature of 300°C, at which the temperature was held for 2 min.

Data-acquisition

The raw GC-MS data were deconvoluted and analysed using AMDIS software (Automated Mass Spectral Deconvolution and Identification System, V2.65). Alignment of the detected compounds across the samples analysed was achieved by creating a new reference library in AMDIS, which contained the mass spectra of all the compounds detected above a threshold of 0.01% of the total signal, for all the samples analysed. Each analysed sample was subsequently processed using the aforementioned reference library, and the resulting output of each sample was combined into a data matrix containing the relative concentrations (normalised with the internal standard) for all compounds present or absent in each sample. This was done for each of the GC-MS data sets collected from each of the 2 extraction methods investigated.

Statistical data analysis

The statistical packages, "R" (version 2.13.0) and Statistica (version 10) were used for all the statistical data analyses. Data were normalised relative to the respective internal standards in order to compensate for sample loss during the extraction or chromatographic injection. This relative normalisation was used purely as a measure to identify variations in the abundances of the detected compounds between individual samples and sample groups, and not for absolute quantification purposes.

To prevent variables with high concentrations from dominating the multivariate statistical analyses, data pre-treatment using a non-parametric transformation function (Koekemoer and Swanepoel, 2008) was used to scale the data prior to statistical data analyses, after which mean centring was applied.

The repeatability of the 2 extraction methods investigated were compared using the distribution of the calculated coefficient of variation (CV) values of the relative concentrations of all the compounds detected subsequent to GC-MS analysis of the extracted sample repeats.

In order to determine whether or not a natural grouping (differentiation) exists between the investigated sample groups, an unsupervised classification procedure known as principal component analysis (PCA) was applied. PCA reduces the dimension of the input data matrix by using a weighted sum of the compound concentrations, hereby summarising the variation in the data matrix into a model plane, generating scores. A scatter plot of these scores provides an overview of the samples and how they relate to each other.

Partial least squares discriminant analysis (PLS-DA) is a supervised classification procedure, where class membership information is used to build a discrimination model, and may therefore be used to identify those compounds which best characterise the differentiated sample groups, by ranking the metabolites according to the variable influence on the projection (VIP) parameter. VIP is a weighted sum of squares of the PLS-DA weights, indicating the importance of the metabolite to the total model, and consequently, the highest ranked compounds are then identified as metabolite markers for the respective sample groups (Madsen et al., 2010). Using the mass fragmentation patterns generated by the MS, together with their respective GC retention times, the identities or compound names for these metabolite markers were determined by using libraries generated from previously injected standards. Using these libraries, a value, representative of the degree of similarity between the mass spectra of the detected compound to that of a previously injected standard in the library, at a corresponding retention time, was then assigned to each compound. A similarity value higher than 80% was considered a positive identification and the corresponding compound name was assigned. All spectra with a similarity lower than 80% were identified as unknown compounds.

RESULTS AND DISCUSSION

Although the purpose of this investigation wasn't aimed at directly comparing any of the two described extraction methods to those conventionally used for similar purposes, it is of value to mention that the use of a vibration mill as described in this study, significantly improves the extraction efficiency of any extraction procedure, especially when extracting compounds from more robust matrixes (such as cell or tissues samples). This increased extraction efficiency allows for faster extraction times and the analyses of 24 samples simultaneously, using far less solvent volumes, in a 2 ml centrifuge tube.

The fatty acid extraction method, for instance, can be completed within 5 h, including the drying and derivitization steps, using a total solvent volume of only 1.75 ml, which is a vast improvement on the 16 h required for the completion of the original Bligh- Dyer method, which also requires far more solvent.

The absence of an organic solvent washing step for the specific extraction of fatty acids makes the total metabolome extraction method even simpler and faster, requiring even smaller amounts of solvent material (1.25 ml) for completion in less than 4 h.

Extraction capacity

The AMDIS generated reference library (a compilation all the compounds extracted from all 5 sample groups) for the fatty acid metabolome extraction procedure contained 270 compounds, predominantly fatty acids and various hydrocarbons while the AMDIS generated reference library for the total metabolome extraction procedure, contained a total 1194 compounds, of various polarities, belonging to a variety of compound classes, including amongst others: fatty acids, amino acids, alcohols, organic acids, monosaccharides, alkenes, alkanes, purines, pyrimidines, etc. Table 1 indicates the number of compounds extracted via each of the compared extraction methods, for each of the infectious disease sample groups. Considering this, the total metabolome extraction procedure detected, on average, 358 more

Table 1. Number of compounds in the AMDIS generated reference libraries originating from each of the bacterial sample groups (12 sample repeats of each), for the two investigated extraction methods. The standard deviation (SD) values, representing the variation in the number of compounds detected between individual samples, are given in parenthesis.

Bacterial sample group	Number of compounds (SD)	
	Fatty acid metabolome extraction method	Total metabolome extraction method
M. tuberculosis	180 (10.09)	461 (10.93)
M. avium	178 (6.12)	501 (13.24)
M. bovis	174 (13.12)	388 (14.40)
M. kansasii	191 (5.87)	641 (14.94)
P. aeruginosa	161 (11.31)	679 (19.66)
Total number of compounds in library	270	1194

compounds per sample group, which could be expected, as the fatty acid metabolome extraction method is a more targeted approach, and is not aimed at extracting all compounds present in the cell matrix.

It should be noted that this evaluation of the number of detected compounds is merely for descriptive purposes, and may not necessarily be an indication of which of the two described methods performs better for metabolomics applications and the identification of makers characterising the above mentioned infectious diseases. Although, as per definition of metabolomics, the total metabolome extraction procedure should better fulfil the required criteria for identifying "all compounds in an unbiased manner", the increased dimensionality of the resulting extract may potentially overwhelm the dimensionality limits of the GC-MS, leading to co-elution and poor repeatability and consequently larger variation or unwanted "noise" in the collected data, hampering its applications for metabolomics research purposes. Knowing this beforehand, however, GC-MS flow rates and temperature programming was optimised, in order to minimise the chance of this happening.

Repeatability

As previously mentioned, the repeatability of the 2 extraction methods investigated were compared using the distribution of the calculated CV values of the relative concentrations of all the compounds detected subsequent to GC-MS analysis of the extracted sample repeats (Figure 1).

Comparatively, for all the organism groups, an averaged 50 ± 6.7% of all the compounds detected, via the fatty acid extraction procedure, as compared to 62 ± 4.6% for that of the total metabolome extraction, fell below a 50% CV value. This indicates that the total metabolome extraction procedure, despite extracting a larger number and variety of compounds, with potentially larger variations in their comparative concentrations, which could in theory overwhelm the dimensionality of the GC-MS with regards to dynamic range and chromatographic separation, still showed better comparative repeatability.

As the most optimal GC-MS conditions were selected for analysing the extracts derived from both extraction methods, the better repeatability of the total metabolome extraction procedure is most easily explained by its comparative simplicity, resulting in a far lower likelihood for analytical variation than the fatty acid metabolome extraction method, which requires an additional step. An interesting observation was that in both cases, the *M. bovis* sample group showed the lowest repeatability of all the bacterial sample groups analysed. When evaluating these results, one needs to keep in mind that the CV value of a compound is directly correlated to its detected concentration (the lower the concentration, the worse the CV) (Rocke and Lorenza, 1995). Using both extraction methods, all identified compounds were detected in significantly lower concentrations in *M. bovis*, as compared to the other bacterial species investigated, possibly explaining the resulting high CV values. *M. tuberculosis* and *M. avium* were extracted with the best comparative repeatability using the fatty acid and total metabolome extraction procedures, respectively.

Differentiation capacities

Considering the application of these methods to comparative metabolomics, not only is the repeatability and extraction capacity as described above of importance, but also the methods' capacity to extract those compounds most important for differentiating of the groups being compared, in order to answer the biological question at hand. Figure 2 represents examples of the total ion chromatograms obtained after GC analyses of each of the bacterial species investigated, showing a clear visual difference between the species using both of the tested extraction methods.

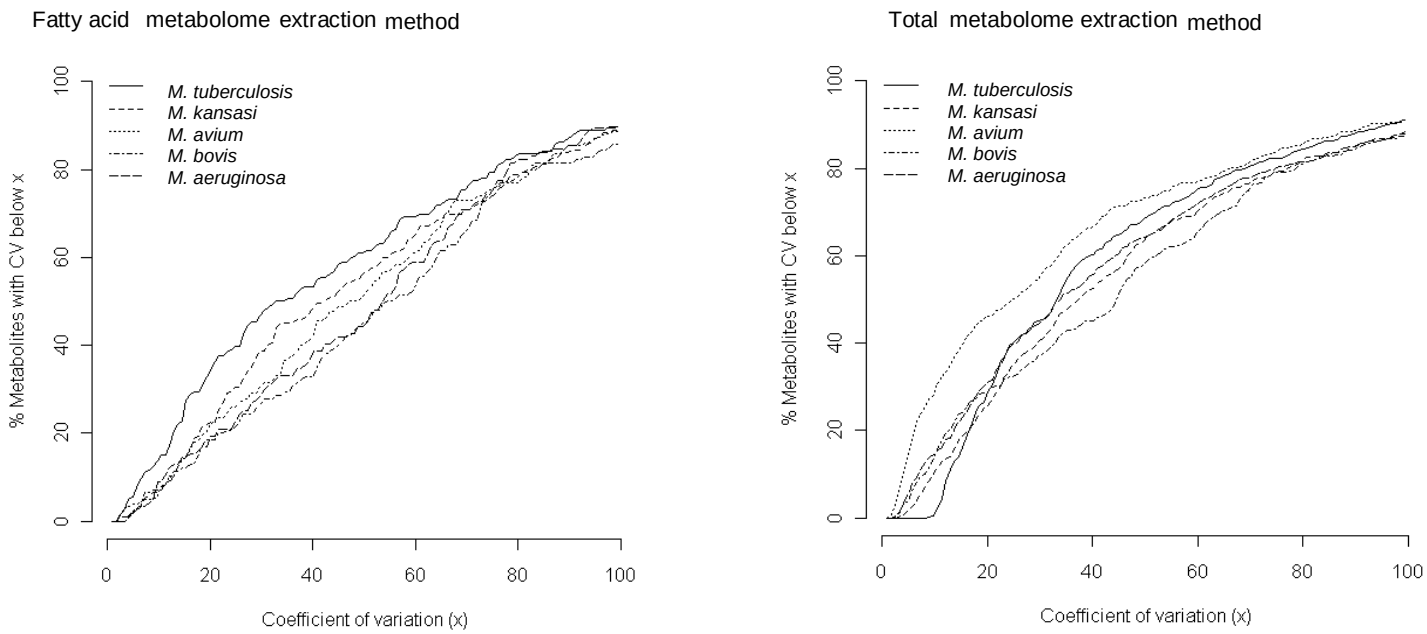


Figure 1. Distribution of the calculated coefficient of variation (CV) values for all the GC-MS detected compounds from each infectious organism group (12 sample repeats of each), extracted using the two extraction methods comparatively.

To determine the ability of the compared extraction procedures to extract those compounds which best explain the variance between the investigated infectious organism groups, two multivariate statistical data analyses (PCA and PLS-DA) were used to summarise and interpret the collected data. Prior to these analyses, a 50% filter was applied to the data collected, in order to exclude those compounds which do not appear in at least 50% of the samples, in one or more of the sample groups, as a data pre-processing step for eliminating unnecessary "noise" in the data.

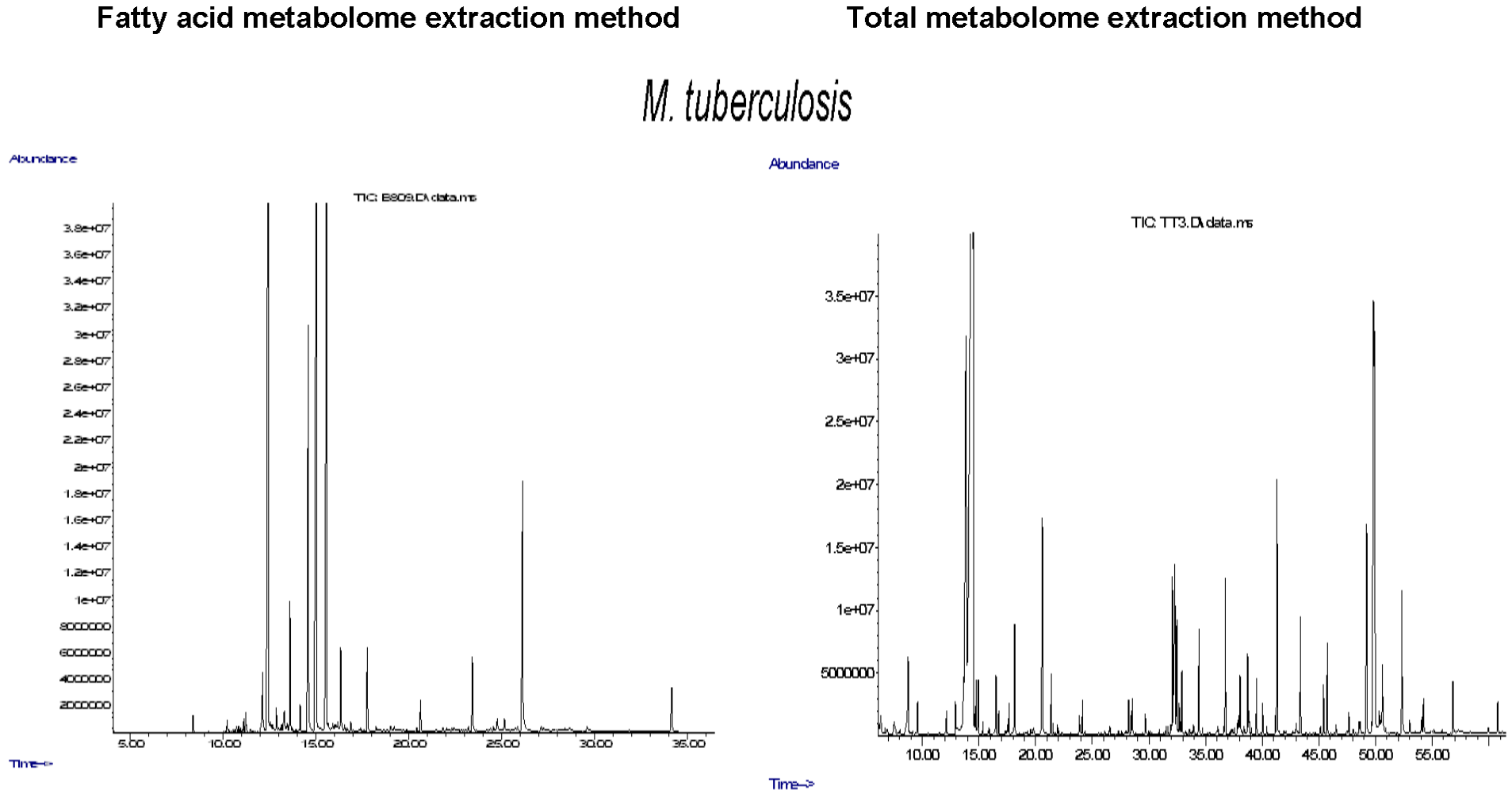
The PCA scores plot using the fatty acid metabolome extraction procedure‘s GC-MS data is shown in Figure 3, and indicates a clear grouping of the individual samples of the respective infectious bacterial groups and a differentiation between these groups. Furthermore, all the related *Mycobacterium* species groups were grouped together, relative to the *P. aeruginosa* group, indicating that organisms with similar genetic characteristics group together due the similarities in their genomes and resultant metabolite profiles.

Figure 4 illustrates the PCA scores plot for the GC-MS data acquired using the total metabolome extraction procedure. Similar to that of the fatty acid metabolome extraction method’s PCA, the related *Mycobacterium* species groups were, once again, grouped together, relative to the *P. aeruginosa* group. This time however, a visibly smaller inter-sample, and hence, within group variation was detected, comparative to that of the fatty acid metabolome extraction procedure, resulting in a better differentiation of the infectious bacterial groups investigated. This smaller within group variation can be explained by the comparatively better extraction repeatability of the total metabolome extraction procedure for those compounds which best described the variation between the infectious organism groups investigated.

Metabolite maker identification

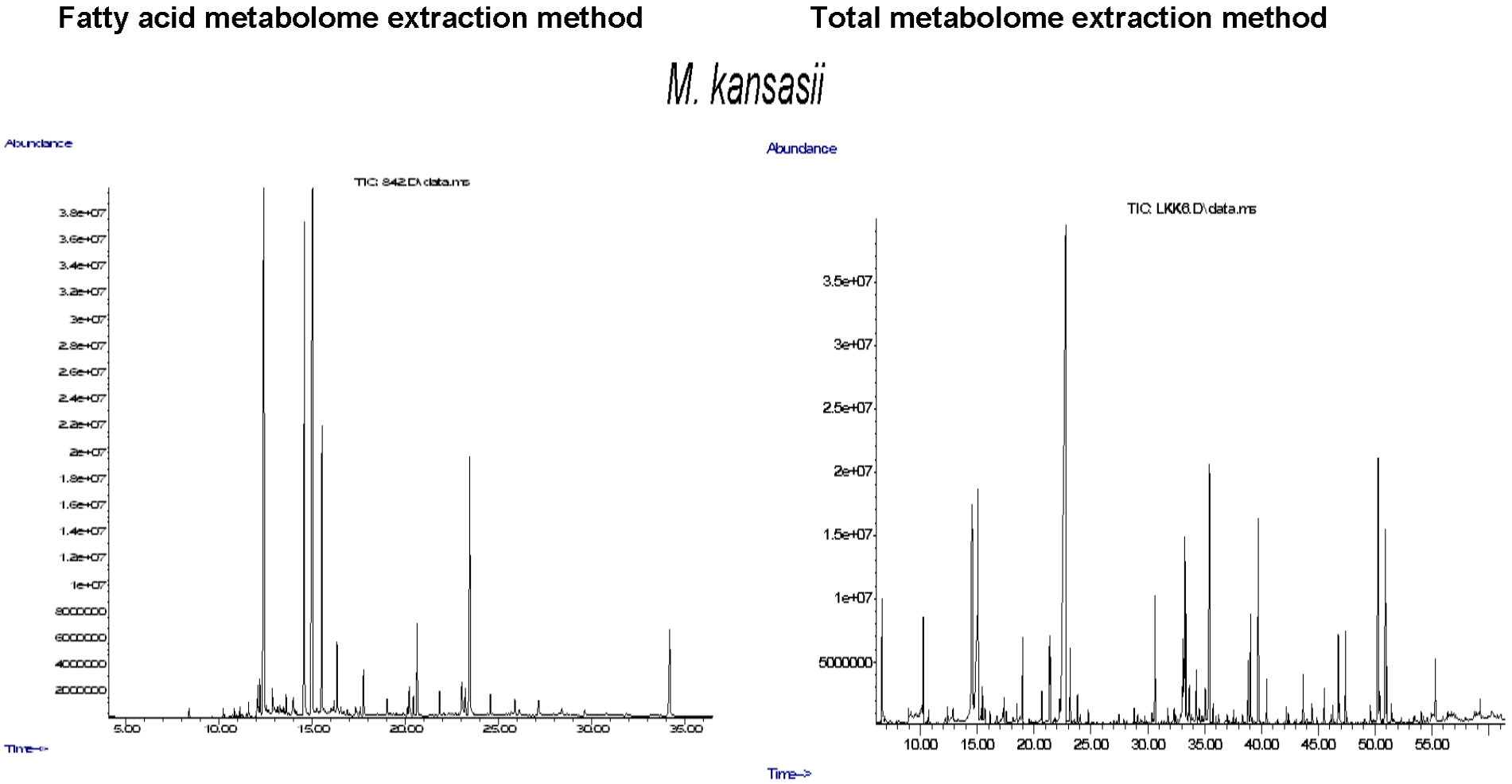
One of the primary functions of a metabolomics research approach is not only determining if a differentiation of various disease states exists on the basis of varying metabolite profiles, but also to identify those metabolite markers which may be used to better characterise the compared sample groups. Considering this, those metabolites with the highest VIP values (identified via PLS-DA), extracted by each of the investigated methods, were identified and evaluated from a functional biology perspective. This was done in order to ensure that the variation detected is in fact due to the extraction of those compounds of biological relevance. In Table 2, the mean relative concentrations of those metabolite markers identified using the fatty acid metabolome extraction approach are presented. The PLS-DA model used for identifying these had a modelling parameter R^2Y (cum) of 93.7%, indicative of the total explained variation of the response Y. Q^2 (cum), the cross-validated variation explained by the response Y, was 89.1%.

From the data in Table 2, it is clear that it is not only the compounds totally unique to a particular group which may be considered as important for differentiating these infectious organisms, but also those compounds which are common to two or more sample groups, showing



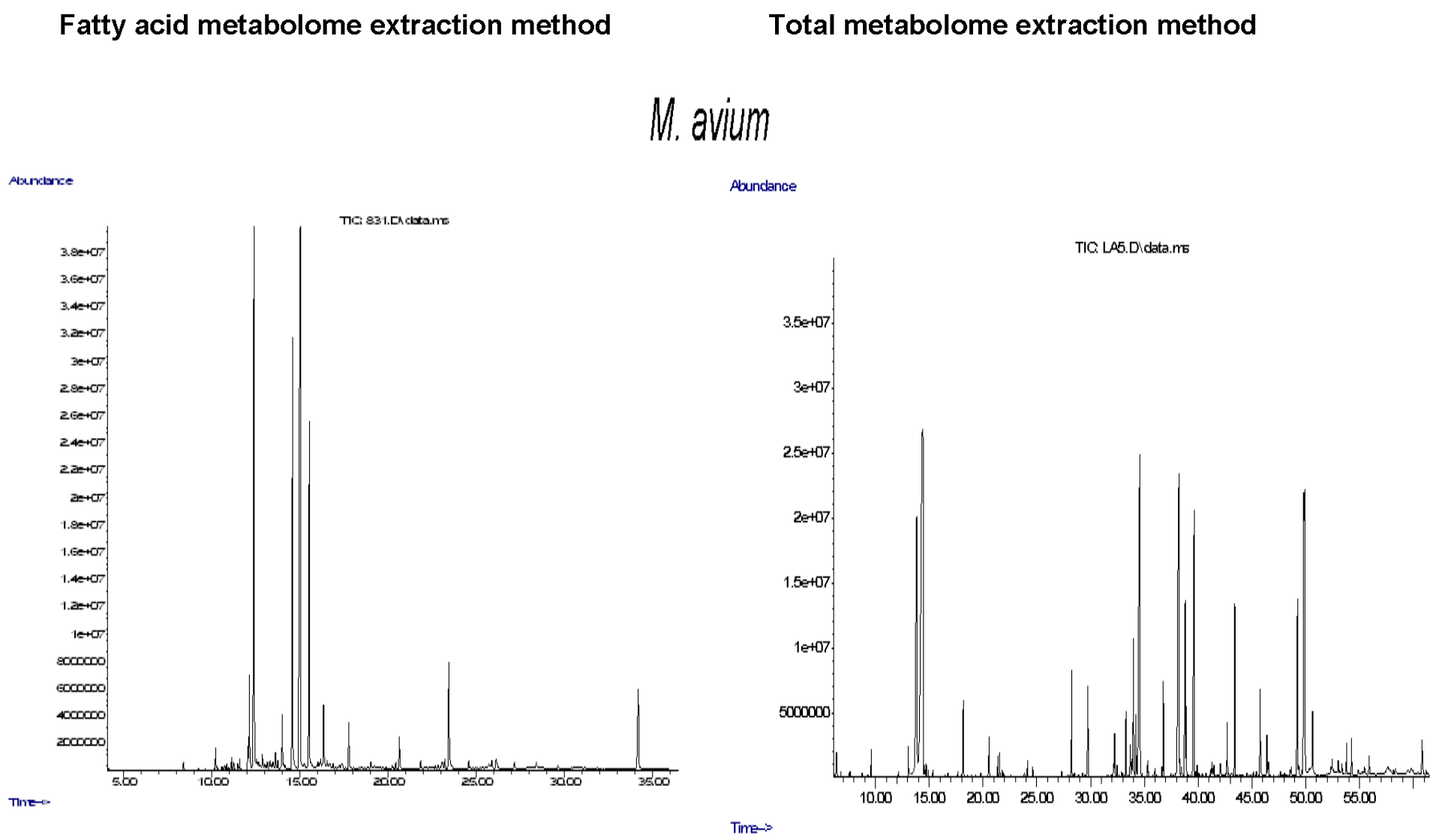
constant differences in their concentrations.

Of the two methods investigated, the fatty acidextraction procedure was selected based on the knowledge that *Mycobacterium* species contain rather unique fatty acids, which are involved in a number of structural and functional processes, including their pathogenecity (Knisley et al*.,* 1985). Consistent with previous reports, we identified the well known characteristic fatty acid, tuberculostearic acid (TBSA), as well as oleic acid (C18:1 ω9c), which is known to be one of the most abundant fatty acids in these bacteria (Lambert et al., 1986; Mosca et al., 2006), as novel metabolite markers for all *Mycobacterium* species groups (Larsson et al*.*, 1987; Stopforth et al*.*, 2004), confirming that this extraction procedure is well suited for characteristic marker metabolite identification, using a metabolomics

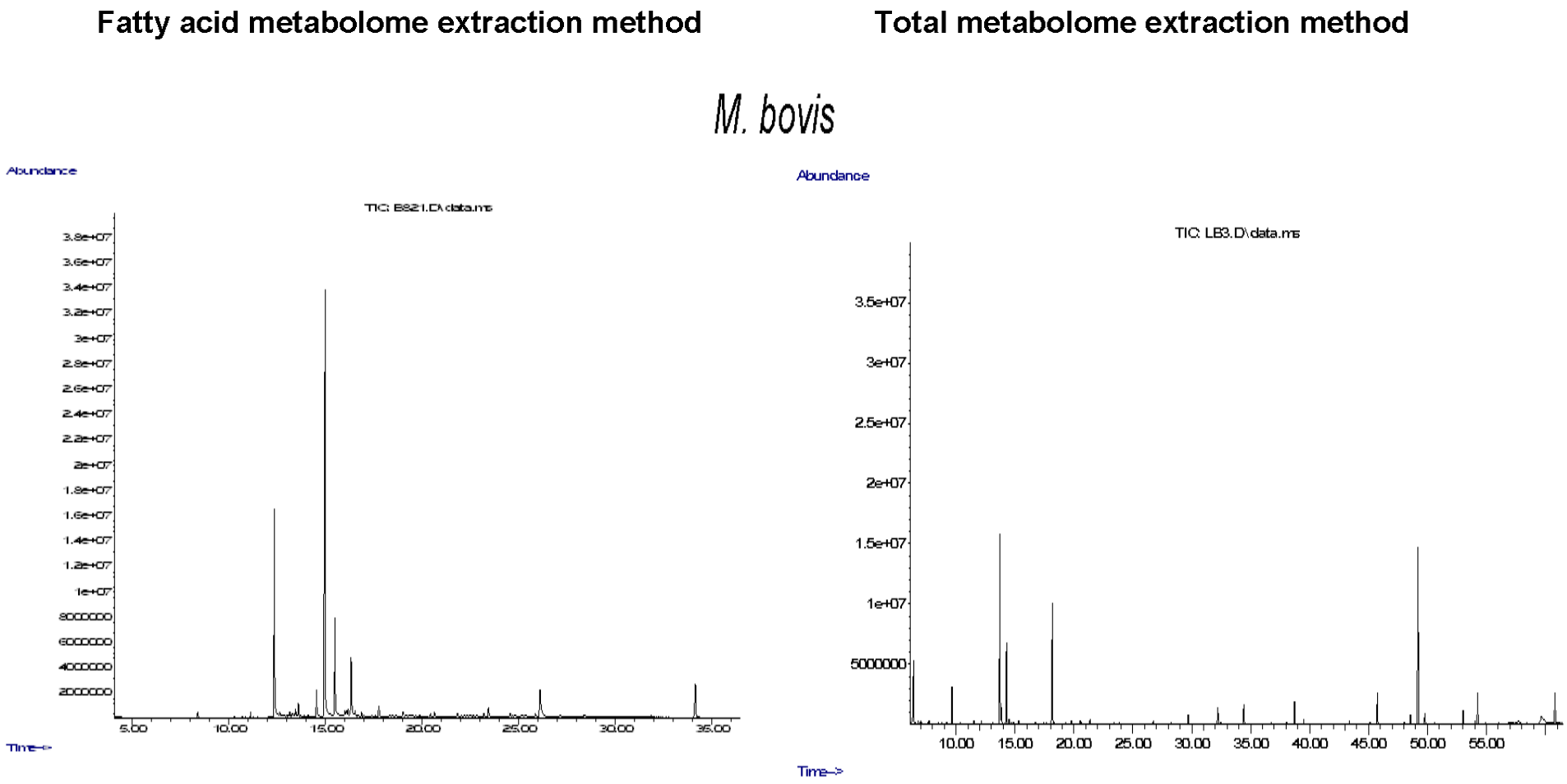


approach. Additionally, the mycolic acid cleavage products (MACPs), eicosanoic acid (C20:0), docosanoic acid (C22:0), which are formed as a result of the chromatographic heat cleavage of the characteristic *Mycobacterium* cell wall components, mycolic acids (Guerrant et al*.,* 1981), were detected in relatively abundant concentrations in all the *Mycobacterium* species, as opposed to only trace amounts in the *P. aeruginosa* sample group.

Another MACP, octacosanoic acid (C28:0), together with gondoic acid (C20:1 ω9c) and two unknown compounds with molecular masses of 466 and 494 respectively, may potentially characterise *Mycobacterium* species, as these were only detected in the *Mycobacterium* species tested, with the exception of *M. bovis*. All identified compounds were detected in significantly



lower concentrations in the *M. bovis* sample group, in comparison to the other bacterial groups tested and, hence, these makers might potentially be present in these samples, however at concentrations below the detection limit. Two unsaturated long-chain fatty acids, erucic acid (C22:1 ω9c) and nervonic acid (C22:1 ω9c), were detected only in the *M. kansasii* and *M. avium* sample groups, whereas pentacosanoic acid



(C25:0) were detected exclusively in the *M. tuberculosis* and *M. bovis* sample groups. Furthermore, as previously identified, 2,4 dimethyltetradecanoic (2,4-DM C14:0) was detected as a novel metabolite for *M. kansasii* (Lambert et al., 1986; Mosca et al., 2006)

We additionally detected another unknown compound, with a molecular mass of 340, specifically characterising *M. tuberculosis* and two compounds, 2-octyl-cyclopropaneoctanoic acid (2-octyl-CPOA) and 2-hexyl-cyclopropaneoctanoic acid (2-hexyl-CPOA) as unique characteristic metabolites of *P. aeruginosa.*

The PLS-DA model used for identifying the characteristic marker metabolites extracted using the total metabolome extraction procedure had a modelling parameter R^2Y (cum) of 97.8%, with a Q^2 (cum) of 98.8%. This extraction method was developed for the purpose of extracting the entire metabolome and consequently the metabolites identified included various fatty acids, amino acids, alcohols, organic acids, monosaccharides, alkenes, alkanes, purines, pyrimidines, etc. These metabolite markers identified are listed in Table 3.

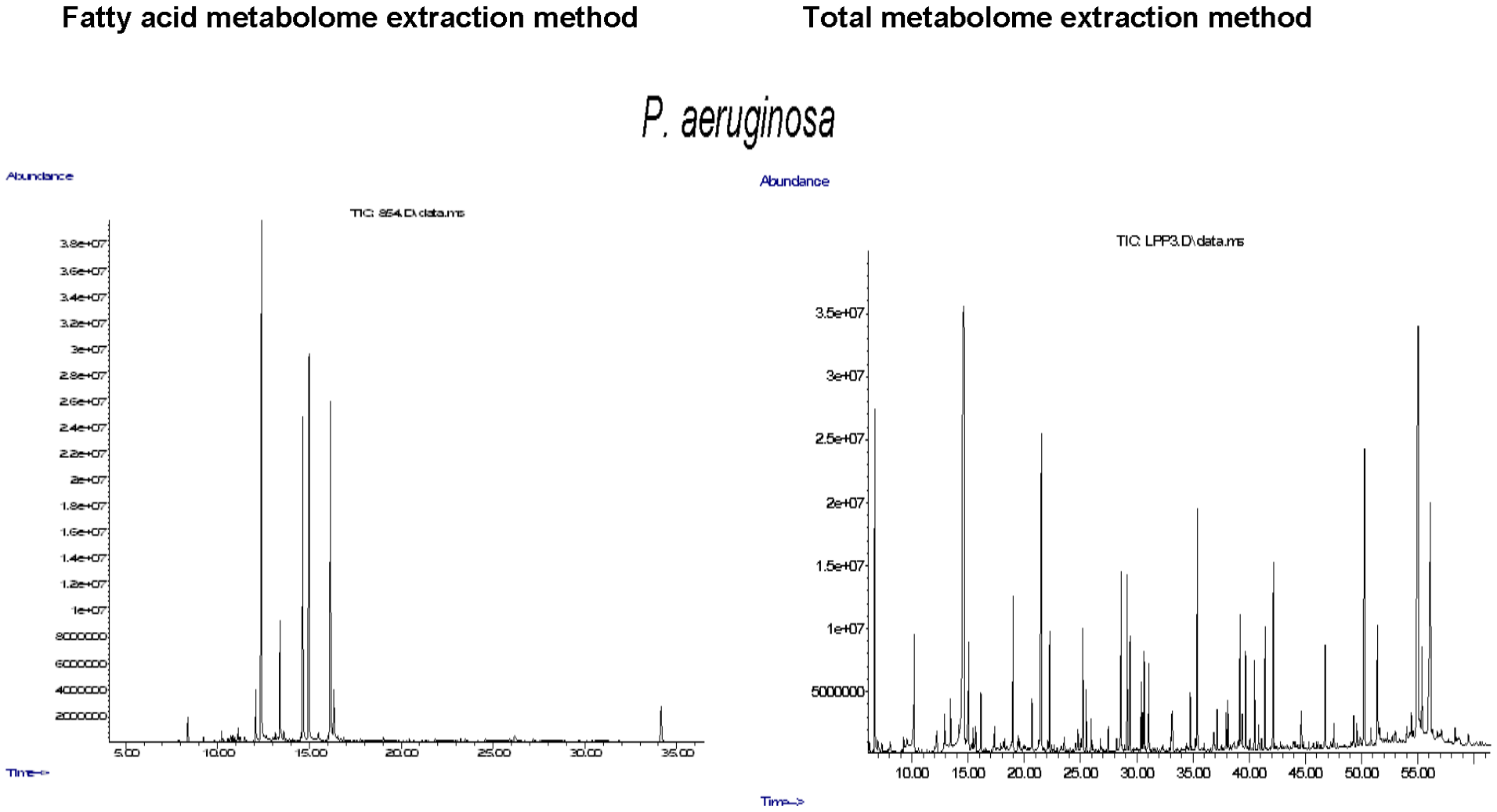


Figure 2. Examples of the total ion chromatograms of the bacterial species investigated, obtained after chromatographic analyses following sample extraction using the two tested extraction methods.

Although both of the investigated methods have the capacity to extract fatty acids, a more complete fatty acid profile was attained using the fatty acid metabolome extraction approach, due to the use of more selective solvents. The total metabolome extraction procedure, on the other

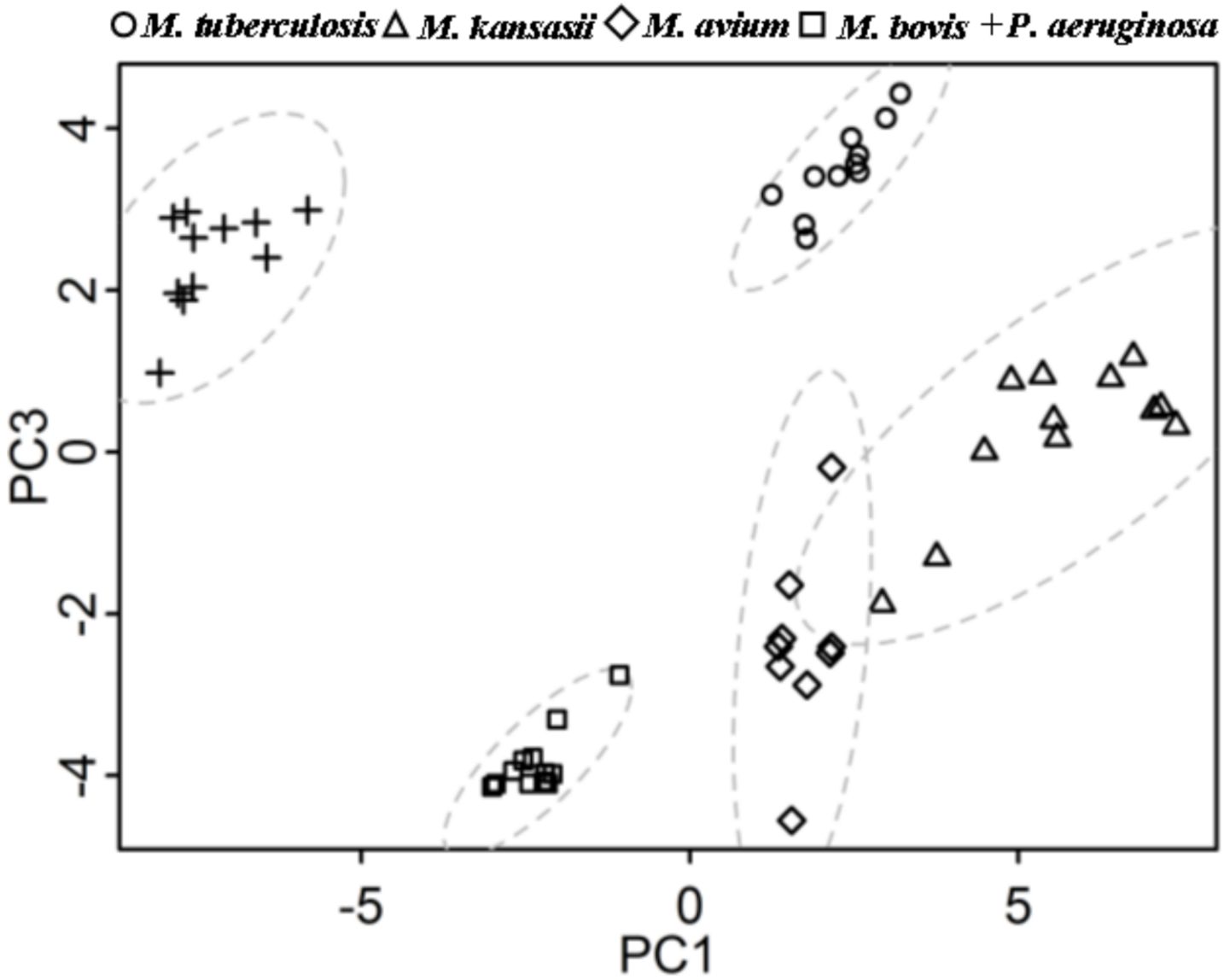


Figure 3. PCA scores plot (PC1 vs. PC3), of the GC-MS data acquired using the fatty acid metabolome extraction method, indicating a grouping of the individual samples into their respective organisms groups on the basis of the variation in the detected fatty acid metabolome. Three PCs were extracted, explaining a total of 64.4% of the variation in the data (R^2X cum), of which PC 1 explained 34.8%, PC 2 explained 15.7%, and PC 3 explained 13.9%. The ellipses represent the 95% confidence interval of the modelled variation.

hand, has the capacity to extract a larger variety of compounds from all compounds classes, including fatty acids, and hence has the advantage of detecting a greater spectrum of characteristic markers. When investigating mycobacteria, however, this may be seen as a minor limitation, due to the fact that some characteristic fatty acids might not be detected in these complex chromatograms. Interestingly, only palmitoleic acid (C16:1ω7c) was identified as a common metabolite marker using both extraction methods, and universally occurred in the highest comparative concentration in the *M. avium* bacterial sample group, and in the lowest concentration in the *M. bovis* sample group. The highest ranked metabolite marker extracted using the total metabolome extraction procedure, and novel to *P. aeruginosa,* was indole-acetic acid. This metabolite has previously been identified in this organism and in other related plant growth-promoting rhizobacteria (Karnwal, 2009). Other novel metabolite markers identified *P. aeruginosa* using this extraction method included: cadaverine, an intermediate in the alternative decarboxylation and catabolism of L-lysine (Fothergill and Guest, 1977; Stewart, 1970); putrescine, an intermediate used in the L-arginine decarboxylase catabolic pathway (Mercenier et al., 1980); purine; valeric acid; and two unknown compounds with molecular masses of 343 and 373. Two

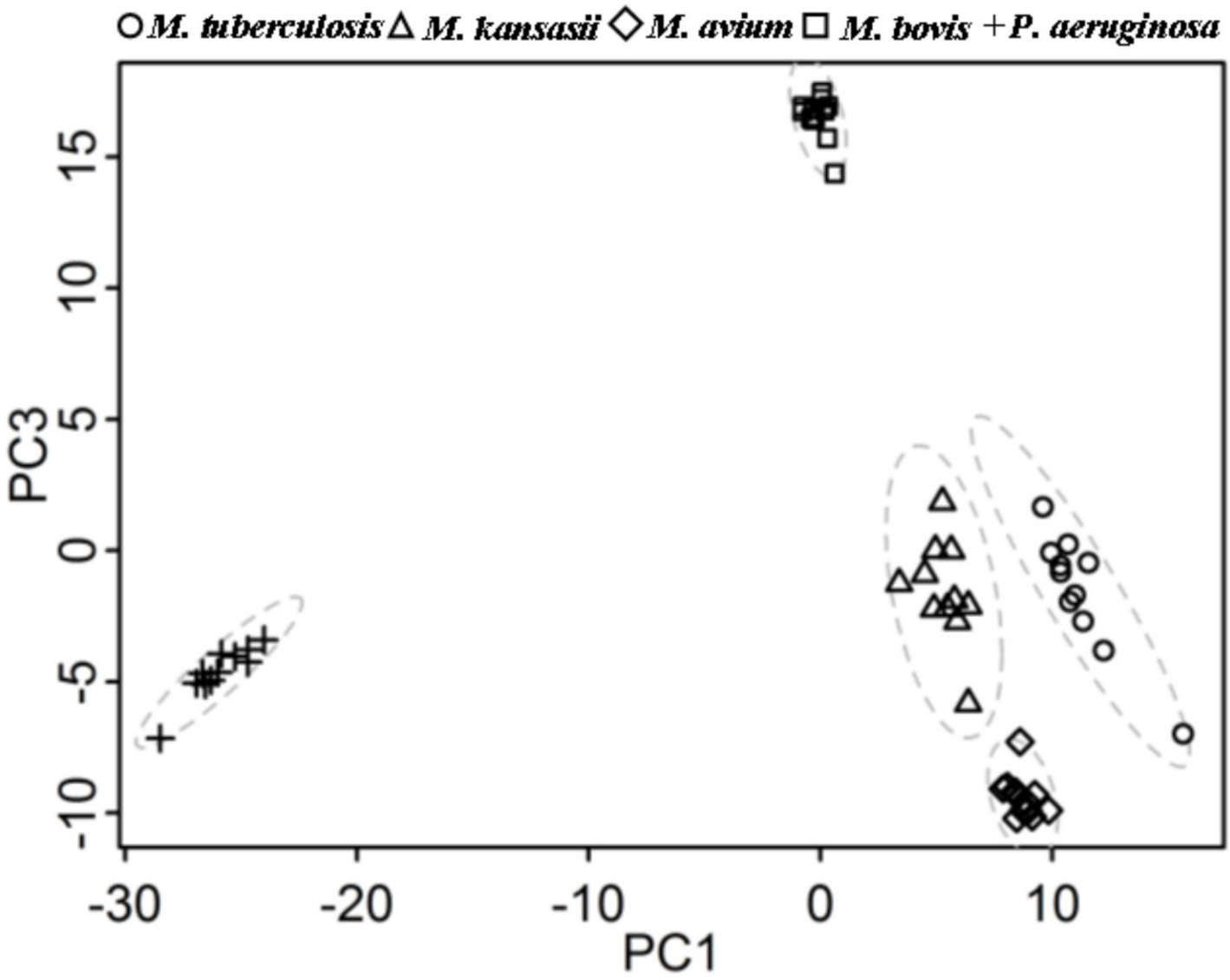


Figure 4. PCA scores plot (PC1 vs. PC3), of the GC-MS data acquired using the total metabolome extraction method, indicating a grouping of the individual samples into their respective organisms groups, with a comparatively smaller within group variation, and hence, bigger inter-group variation, as opposed to the fatty acid extraction procedure. Three PCs were extracted, explaining a total of 48.8% of the variance in the data (R^2X cum), of which PC 1 explained 27.2%, PC 2 explained 14.5%, and PC 3 explained 7.08%. The ellipses represent the 95% confidence interval of the modelled variation.

markers novel to the related *Mycobacterium* species, namely inositol and its stereoisomer, myo-inositol, were also identified using this approach. As summarised by Brown et al. (2007), the myo-inositol supply in mycobacteria is believed to be sustained by *de novo* synthesis initiated by the conversion of glucose-6-phosphate to inositol-1-phosphate followed by dephosphorylation, and is a novel characteristic of these organisms. Succinic acid and an unknown compound with a molecular mass of 268 were identified as novel *M. kansasii* markers, in addition to elevated concentrations of erythritol and citric acid, in comparison to *M. tuberculosis, M. avium* and *M. bovis,* which has never before been documented. It is also important to note that an unknown compound of mass 541 was uniquely detected in *M. tuberculosis,* which could potentially be investigated further for possible diagnostic applications.

Conclusions

Considering these results, the total metabolome extraction procedure has advantages over that of the total fatty acid extraction procedure investigated, considering that: 1) it is simpler; 2) faster; 3) has a comparatively better repeatability, consequently resulting

Table 2. Mean relative concentrations (ug/mg sample) of metabolite markers identified for the various infectious organism groups (12 sample repeats of each) following the fatty acid metabolome extraction method. Standard deviation (SD) values are given in parenthesis.

VIP ranking	Species	*M. tuberculosis*	*M. kansasii*	*M. avium*	*M. bovis*	*P. aeruginosa*
	Compound	Concentration (SD)				
1	TBSA	273.9 (28.2)	171.3 (50.4)	139.8 (27.3)	32.7 (8.7)	0
2	C22:0	14.1 (2.7)	63 (26.1)	17.7 (5.4)	2.4 (0.6)	T
3	2-octyl-CPOA	0	0	0	0	195.9 (25.8)
4	C17:0	43.2 (8.4)	8.7 (3)	6.3 (0.9)	4.5 (1.2)	3.6 (0.6)
5	C22:1 ω9c	0	21.3 (9)	1.5 (0.3)	0	0
6	C24:1 ω9c	0	25.2 (11.7)	2.4 (0.06)	0	0
7	2,4-DM C14:0	0	3.3 (1.8)	0	T	0
8	Unknown 466	0.6 (0.3)	4.8 (2.7)	0.3 (0.6)	0	0
9	C16:1 ω7c	18 (4.5)	20.7 (8.7)	25.5 (8.4)	0.3 (0.3)	19.5 (6.9)
10	C20:0	36.9 (6.9)	29.4 (11.1)	20.7 (3)	5.4 (1.2)	1.2 (0.3)
11	C14:0	4.8 (2.7)	17.4 (10.8)	8.7 (4.2)	0.3 (0.3)	2.1 (1.5)
12	C28:0	3.3 (0.9)	2.7 (1.2)	T	0	0
13	Unknown 494	1.5 (0.3)	1.5 (0.6)	0.9 (0.3)	0	0
14	Unknown 422	5.7 (1.2)	0	0	0.6 (0.3)	0
15	C25:0	4.2 (0.9)	0	0	0.3 (0.03)	0
16	C18:1 ω9c	192.6 (32.1)	309 (86.1)	193.5 (24.9)	10.5 (2.7)	0
17	2-hexyl-CPOA	0	0	0	0	50.1 (8.4)
18	C20:1 ω9c	0.9 (0.3)	5.4 (2.4)	0.9 (0.6)	0	0
19	Unknown 340	1.8 (0.3)	0	0	0	0

TBSA, tuberculostearic acid; CPOA, cyclopropaneoctanoic acid; DM, dimethyl; T, trace amounts.

Table 3. Mean relative concentrations (ug/mg sample) of metabolite markers identified for the various infectious organism groups (12 sample repeats of each) following the total metabolome extraction procedure. Standard deviation (SD) values are given in parenthesis.

VIP ranking	Species	*M. tuberculosis*	*M. kansasii*	*M. avium*	*M. bovis*	*P. aeruginosa*
	Compound	Concentration (SD)				
1	Indole-acetic acid	0	0	0	0	123.6 (0.6)
2	Unknown 343	0	0	0	0	216.6 (36.6)
3	Citric acid	0	213.3 (55.2)	0	4.5 (0.3)	0
4	Cadaverine	0	0	0	0	84.3 (15.6)
5	Purine	0	0	0	0	399 (67.5)
6	Inositol	262.8 (138.6)	6.9 (6.0)	168.6 (20.1)	0.3 (0.6)	0
7	5'-Adenylic acid	0	0	55.2 (7.8)	0	735 (111.3)
8	C16:1 ω7t	0.3 (0.6)	0.9 (1.2)	62.1 (5.7)	0	0.3 (1.2)
9	Myo-Inositol	227.1 (149.1)	27.9 (18.0)	348 (50.4)	1.5 (1.8)	0
10	Succinic acid	0	119.4 (19.8)	0	0	0
11	Putrescine	0	0	0	0	191.1 (42.9)
12	C10:0	0	T	19.8 (1.5)	0	53.1 (13.5)
13	Unknown 268	0	22.5 (0)	0	0	0
14	C16:1 ω7c	6.6 (2.2)	7.2 (4.5)	303 (32.1)	0	53.1 (9.9)
15	Erythritol	2.4 (1.2)	19.2 (5.4)	12 (0.9)	0	0
16	Unknown 367	0	0.6 (0.9)	15.9 (1.8)	0	39 (7.2)
17	Unknown 373	0	0	0	0	60.6 (14.7)
18	Unknown 541	84 (42)	0	0	0	0
19	Valeric acid	0	0	0	0	114.9 (15)

in less inter-sample variation, and; 4) isolates the total metabolome, increasing the chance for identifying unique metabolite markers from a variety of compound classes. Despite this however, both these methods proved to extract characterising compounds which best differentiate the various TB causing mycobacteria from one another, and from other unrelated lung pathogens, which are known to induced a similar disease state and symptoms. Hence, both methods could potentially play a valuable role in identifying the metabolite markers which could be used to successfully detect, diagnose, and better characterise various infectious diseases.

REFERENCES

Brown AK, Meng G, Ghadbane H, Scott DJ, Dover, LG, Nigou, J, Besra GS, Fütterer K (2007). Dimerization of inositol monophosphatase *Mycobacterium tuberculosis* SuhB is not constitutive, but induced by binding of the activator Mg2+. BMC Struct. Biol., 7: 55.

Davis R (2007). Tuberculosis and the intensivist. Curr. Anaesthesia Crit. Care, 18: 76-85.

Dunn WB, Bailey NJC, Johnson HE (2005). Measuring the metabolome: current analytical technologies. Analyst., 30: 606-625.

Fend R, Kolk AH, Bessant C, Buijtels P, Klatser PR, Woodman, AC (2006). Prospects for clinical application of electronic-nose technology to early detection of *Mycobacterium tuberculosis* in culture and sputum. J. Clin. Microbiol., 44: 2039-2045.

Fothergill JC, Guest JR (1977). Catabolism of L-Lysine by *Pseudomonas aeruginosa*. J. Genl. Microbiol., 99: 139-155.

Guerrant GO, Lambert MA, Moss CW (1981). Gas-chromatographic analysis of mycolic acid cleavage products in mycobacteria. J. Clin. Microbiol., 13: 899-907.

Karnwal A (2009). Production of indole acetic acid by fluorescent *Pseudomonas* in the presence of L-Tryptophan and rice root exudates. J. Plant Pathol., 91(1): 61-63.

Knisley CV, Damato JJ, McClatchy K, Brennan PJ (1985). Rapid and sensitive identification of *Mycobacterium tuberculosis.* J. Clin. Microbiol., 22(5): 761-767.

Koekemoer G, Swanepoel JWH (2008). A semi-parametric method for transforming data to normality. Stat. Comput., 18: 241-257.

Lambert MA, Moss CW, Silcox VA, Good RC (1986). Analysis of mycolic acid cleavage products and cellular fatty acids of *Mycobacterium* species by capillary gas chromatography. J. Clin. Microbiol., 23: 731-736.

Larsson L, Odham G, Westerdahl G, Olsson B (1987). Diagnosis of pulmonary tuberculosis by selected-ion monitoring: improved analysis of tuberculostearate in sputum using negative-ion mass spectrometry. J. Clin. Microbiol., 893-896.

Li X, Xu Z, Lu X, Yang X, Yin P, Kong H, Yu Y, Xu G (2009). Comprehensive two-dimensional gas chromatography/time-of-flight mass spectrometry for metabonomics: Biomarker discovery for diabetes mellitus. Anal. Chim. Acta., 633: 257–262.

Madsen R, Lundstedt T, Trygg J (2010). Chemometrics in metabolomics—A review in human disease diagnosis. Anal. Chim. Acta., 659: 23–33.

Mercenier A, Simon J, Haas D, Stalon V (1980). Catabolism of L-Arginine by *Pseudomonas aeruginosa*. J. Genl. Microbiol., 116: 381-389.

Mosca A, Russo F, Miragliotta L, Iodice MA, Miragliotta G (2006). Utility of gas chromatography for rapid identification of mycobacterial species frequently encountered in clinical laboratory. J. Microbiol. Methods., 86: 392-395.

Miguez-Burbano MJ, Flores M, Ashkin D, Rodriguez A, Granada AM, Quintero N, Pitchenik A (2006). Non-tuberculous mycobacteria disease as a cause of hospitalization in HIV-infected subjects. Int. J. Infect. Dis. 10: 47-55.

Rocke DM, Lorenza S (1995). Two-component model for measurement error in analytical chemistry. Technometrics., 37(2): 176-184.

Schoeman JC, Loots DT (2011). Improved disease characterisation and diagnostics using metabolomics: a review. J. Cell Tissue Res., 11: 2673-2683.

Stewart DJ (1970). The decarboxylation of L-lysine by *Pseudomonas aeruginosa*. J. Appl. Bact., 33: 501-504.

Stopforth A, Tredoux A, Crouch A, van Helden P, Sandra P (2004). A rapid method of diagnosing pulmonary tuberculosis using stir bar sorptive extraction-thermal desorption-gas chromatography-mass spectrometry. J. Chromatogr. A., 1071: 135-139.

Todar K (2005). Todar's Online Textbook of Bacteriology. University of Wisconsin-Madison Department of Bacteriology. http://www.textbookofbacteriology.net/tuberculosis.html

WHO (2010). The global plan to stop TB 2011-2015: transforming the fight towards elimination of tuberculosis. WHO Press. Geneva, Switzerland.