An optimised sample preparation method for NMR-based faecal metabonomic analysis†

Junfang Wu,[ab] Yanpeng An,[ab] Jianwu Yao,[ab] Yulan Wang[a] and Huiru Tang*[a]

Received 30th December 2009, Accepted 3rd February 2010
First published as an Advance Article on the web 17th February 2010
DOI: 10.1039/b927543f

Faecal metabonomic NMR analysis plays an essential role in investigating the interactions between mammalian metabolism and symbiotic gut microbiota. However, the faecal metabolite extraction method remains to be optimised and standardised to take into consideration signal-to-noise ratios, pH and chemical shift consistency. In the current investigation, we compared extraction consistency of three homogenisation methods including manual ultrasonication, automatic homogenization with tissuelyser and their combination, and systematically optimised faecal metabolite extraction parameters, including the faeces-to-buffer ratio ($W_f : V_b$), extraction repetition times and duration. We found that automatic homogenisation with tissuelyser was the choice of extraction method owning to its good metabolite extraction consistency and high throughput. We also recommend $W_f : V_b$ of 1 : 10 (mg μl^{-1}) and use of the combined first two extracts as the resultant samples to represent faecal metabolite composition. Such recommendation is based on considerations of maximisation of the spectral signal-to-noise ratio, pH and chemical shift consistency, completeness of metabolite extraction and sample preparation throughput so that the method is suitable for analysing a large number of samples especially in human population studies.

Introduction

Mammals harbour trillions of symbiotic gut microbes with essential functions and hence have been regarded as "superorganism".[1] The symbiotic gut microbiota (microbiome) interacts with mammalian metabolism to contribute to the hosts' biology[2–4] and play important roles in mammalian pathology[5] and efficacy of xenobiotic (*e.g.* drugs) interventions.[6] For example, these interactions are implicated in the effects of phytomedicines,[7] hepatotoxins[8,9] and parasitic infections[10–13] as well as in the development of glucose intolerance and insulin resistance.[14] The host-microbe co-metabolism appeared to be one of the major aspects of such interactions in modulating and regulating the host physiology.[15] The well-known enterohepatic recirculation is by far the most efficient route of host-microbiome co-metabolism of bile acids and drugs.

The microbiome and mammalian metabolism interactions can be investigated with microecological techniques,[16] which directly analyses the composition of microbiotal species. The interactions can also be investigated indirectly using the metabolic profiling of the host urine samples since a number of urinary metabolites are from host-microbe co-metabolisms.[15] For example, phenylacetylglycine in urine is formed in the mammalian liver through glycine conjugation of phenylacetic acid which is formed from intestinal bacterial conversion of phenylalanine.[17] Another way of investigating such interactions is to analyse the metabolite compositions of faeces as documented in a number of published reports.[18–22] For example, the faecal metabolite composition of patients with Crohn's disease having dysfunctions of gut microbiota showed significant changes in methylamine, trimethylamine and short chain fatty acids (SCFAs),[19] which were produced by the gut microbiota *via* the fermentation of complex carbohydrates.

The efficient extraction method is a prerequisite for effective and robust faecal metabonomic analysis. Thus far, such methods for extracting faecal metabolites have not been optimised although several studies have been reported on the metabolic profiling of faecal extracts.[18,22] The ideal extraction methods ought to be of high consistency for good reproducibility, high extraction consistency to represent the metabolite composition of faecal materials and high throughput to enable multiple sample analysis, such as in the case of human population studies. When the NMR-based metabonomics approaches are employed for metabolite composition analysis, the demands for extraction method will also include optimal concentrations and consistent pH values for the resultant extracts so as to maximise signal-to-noise ratio (SNR) and minimise the inter-sample chemical shift variations.

In the previous work, several extraction methods have been employed including direct sample filtration,[19] ultrasonic homogenisation,[22] freeze-thaw and combinations of different extraction solvents.[18] However, the faeces weight-to-solvent ratio (mg μl^{-1}) was not optimised; hence the inter-sample pH uniformity, chemical-shift consistency and optimal SNR in the resultant NMR spectra were not comprehensively taken into consideration. The previous works employed the faeces

[a]*State Key Laboratory of Magnetic Resonance and Atomic and Molecular Physics, Wuhan Center for Magnetic Resonance, Wuhan Institute of Physics and Mathematics, Chinese Academy of Sciences, Wuhan, 430071, People's Republic of China. E-mail: Huiru.tang@wipm.ac.cn; Fax: +86-27-87199291; Tel: +86-27-87198430*
[b]*Graduate School of the Chinese Academy of Sciences, Beijing, 100049, People's Republic of China*
† Electronic supplementary information (ESI) available: Fig. S1–4. See DOI: 10.1039/b927543f

weight-to-solvent ratios varying from 1 : 20[18] to 1 : 2[20,21] without offering reasons and such ratio is even higher when a direct filtration method was used.[23] Additionally, buffers are normally employed as extraction solvent for the purpose of maintaining pH consistency of the resultant extracts. It is conceivable in such cases that if the faeces weight-to-solvent ratio is too small, the inter-sample pH and chemical shift consistency can be easily maintained but with the sacrifice of the SNR. On the other hand, when the faeces weight-to-solvent ratio is too high, saturation effects can possibly cause incomplete extraction of certain metabolites, large inter-sample pH variations and thus chemical shift inconsistency. Furthermore, several extraction solvents were employed in previous work, including phosphate buffer,[22,24] methanol, alkalized (with NaOH) and acidified (with formic acid) solutions.[18] Strong basic and acidic solutions may cause the degradation of some metabolites (*e.g.* through hydrolysis) and difficulties in maintaining the final pH, especially when the high throughput metabonomic measurements are considered. The problem associated with methanol extraction is that some of the important volatile metabolites, such as endogenous methanol and SCFAs including acetic acid, propionic acid and butyric acid could be lost during removal of methanol. Extracts from a single extraction were often employed in those studies without testing whether such extraction was efficient or not.

Moreover, bacterial metabolites are important parts of faecal metabolite composition, hence consistent and complete bacterial cell lysis is essential. A well-known efficient cell lysis method widely used in molecular biology studies, ball-milling lysis with a tissue-lyser, has not been tested for faecal metabolite extraction. Such a lysis method ought to be well suited for high throughput analysis since such tissuelysers are readily available with minimal costs and combined cell lysis mechanisms including agitation, grinding and homogenisation.[25,26] Lyophilisation has also been frequently employed to remove successive volumes of solvent in previous faecal metabonomic analysis. However, such procedures ought to be employed with extra care since lyophilisation will inevitably cause concentration changes for volatile but important faecal metabolites such as SCFAs.[22]

In this work, we compared the extraction consistency of three different homogenisation methods for metabonomic studies using mouse faeces samples and systematically optimised the extraction parameters including the faeces weight to buffer volume ratios ($W_f : V_b$), repetition times and homogenisation durations with consideration of inter-sample pH and chemical shift consistency, spectral SNR and extraction throughputs. The aim of this work is to establish an effective and optimised sample preparation method for the NMR-based faecal metabonomics studies.

Experimental

Chemicals

Double distilled water was used for preparation of all solutions. Phosphate buffer solution (0.1 M, K_2HPO_4/NaH_2PO_4, pH 7.37) was prepared as faecal extraction solvent which contained 10% D_2O (99.9% D) to provide a field lock for the NMR spectrometer, 0.01% NaN_3 (w/v) as preservatives and 0.5 mM sodium 3-trimethylsilyl (2,2,3,3-^{2}H4) propionate (TSP) as chemical shift reference.

Faecal sample collection and preparation

Six female healthy BALB/C mice (7 weeks old) were purchased, as part of another investigation (Permission No. SYXK (E) 2008-0013), from the Animal Experimental Centre of Wuhan University, China, and allowed to have free access of normal chow and water with regulated light/dark cycle (12 h each). Following 4 weeks assimilation, faeces samples were collected from individual animals between 09 : 00 and 11 : 00 on the same day. All stool samples were taken immediately after being voided and quenched in liquid nitrogen followed by storage at −80 °C until required for further extraction.

pH measurements

pH measurements for faecal extracts were performed at room temperature (25 ± 1 °C) using a Mettler Toledo pH meter (Delta 320) equipped with a Mettler Toledo combination glass electrode, which was calibrated using commercial standard buffers (pH 4.01 and 7.00).

Optimisation of faecal extraction methods

To optimise the faeces-to-buffer ratio (FBR, $W_f : V_b$, mg μl^{-1}), faecal extraction was performed on ice, individually, by adding 700 μl phosphate buffer (0.1 M, pH 7.37) to 21 mg, 35 mg, 70 mg, and 100 mg thawed stool samples as illustrated in Fig. 1a. The resultant samples therefore have four different FBRs, namely, 1 : 33, 1 : 20, 1 : 10 and 1 : 7 (mg μl^{-1}). After vortex mixing for about 30 s, the mixed slurry was subjected to freeze-thaw treatments (3 times) and followed with ultrasonication cycles for 5 or 10 times in order to optimise the extracting durations. Ultrasonication was conducted in an ice bath in the form of ultrasonication (20 s)–vortex (10 s)–waiting (30 s). After 10 mins centrifugation (16 000 × g at 4 °C), 600 μl of the supernatant was withdrawn followed with pH and ^{1}H NMR measurements. The remaining residues of each sample were further subjected to the above mentioned procedure twice. For the convenience of discussion, the resultant first, second and third extracts were designated as F_1, F_2 and F_3 respectively. The pH values and NMR spectra were also measured for the combined extracts from the first two (F_1 and F_2) and first three extracts (F_1, F_2 and F_3) with equal volumes, which were designated as F_{12} and F_{123} respectively.

We further compared the effects of different homogenisation methods on the pH and NMR profiles of the extracts. In such cases, a $W_f : V_b$ ratio of 1 : 20 (mg μl^{-1}) was used with the procedures described in Fig. 1b. In brief, following the addition of phosphate buffer to the faeces sample and 30 s of vortexing, the mixtures were subjected to freeze-thaw treatments for 3 times. Then the samples were homogenised with 3 different methods, including manual ultrasonication (Fig. 1a), 90 s homogenisation with a tissuelyser (QIAGEN, Hilden, Germany) at 20 Hz, and the combination of both. In the latter case, samples were treated with 90 s homogenisation with a tissuelyser at 20 Hz followed with 20 s ultrasonication. Supernatants were designated as F^U_1, F^T_1 and F^{UT}_1, respectively and pH values and NMR profiles were recorded. The remaining residuals were further extracted once in the exactly same manner to obtain samples F^U_2, F^T_2 and F^{UT}_2. The equal volume of the first and second extracts

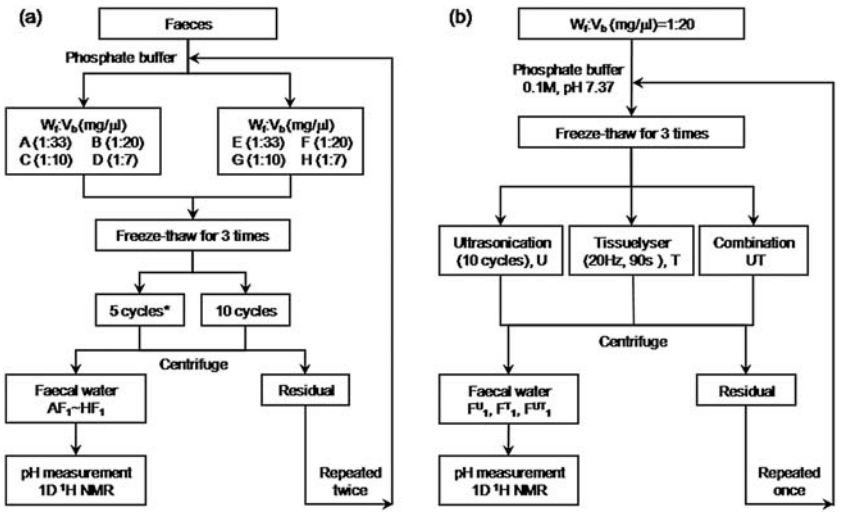

Fig. 1 Procedures for faecal metabolite extractions. (a) The flow chart shows procedures for optimising the faeces-to-solvent ratio (W_f : V_b, mg μl^{-1}) and ultrasonication periods. (b) The flow chart shows procedures for optimising homogenisation methods (manual ultrasonication, automatic tissuelyser and the combination of both). *1 cycle denotes ultrasonication for 20 s, vortex for 10 s and break for 30 s.

from the same samples were combined (*i.e.*, $F^U_1 + F^U_2$, $F^T_1 + F^T_2$ and $F^{UT}_1 + F^{UT}_2$) to give F^U_{12}, F^T_{12} and F^{UT}_{12} respectively, which were also subjected to pH and NMR measurements.

^{1}H NMR spectroscopy

All the extracts (600 μl) were individually transferred into 5 mm NMR tubes for NMR spectroscopic analysis. ^{1}H NMR spectra were recorded at 298 K on a Bruker AVII 500 MHz NMR spectrometer (Bruker Biospin, Germany), operating at 500.13 MHz for proton frequency, using a broad band inverse (BBI) probe with z-gradient. One-dimensional NMR spectra were acquired using the first increment of NOESY pulse sequence (recycle delay-G_1-90°-t_1-90°-G_2-t_m-90°-acquisition) with water presaturation during both the recycle delay (2s) and mixing time (t_m, 100 ms). For each sample, the 90° pulse length was adjusted to about 10 μs and 32 scans were collected into 32k data points with a spectral width of 20 ppm.

For resonance assignment purposes, standard 2D NMR spectra were also acquired on a Bruker AVIII 600 MHz NMR spectrometer (Bruker Biospin, Germany) equipped with an inverse cryogenic probe for selected samples, including ^{1}H-^{1}H correlation spectroscopy (COSY) and total correlation (TOCSY), ^{1}H-^{13}C heteronuclear single quantum coherence spectroscopy (HSQC) and ^{1}H-^{13}C heteronuclear multiple bond correlation (HMBC). For COSY and TOCSY experiments, 80 transients for each of 128 increments were acquired into 2k data points with spectral width of 10.5 ppm for both dimensions. For TOCSY experiments, MLEV17 was employed as the spin-lock scheme with the mixing time of 100 ms and TPPI as the phase increment scheme. For HSQC, typically 2k data points with 240 scans per increment and 120 increments were acquired with the spectral widths of 10.0 ppm and 150 ppm for ^{1}H and ^{13}C respectively. The HMBC experiment was conducted in the phase insensitive mode with the gradient selected pulse sequence and the long-range coupling constant of 6 Hz. A total of 400 transients were acquired into 4k data points for each of 80 increments with the spectral widths of 10.5 ppm and 220 ppm for ^{1}H and ^{13}C respectively. These data were zero-filled into a 4 × 2k data matrix and applied with appropriate apodization functions prior to Fourier transformation (FT) with complex forward linear predictions.

Metabolite concentration ratios

The metabolite concentration ratios (MCRs) were calculated from the signal integrals of glucose (δ 5.265–5.225), acetate (δ 1.949–1.897), butyrate (δ 0.869–0.919) and succinate (δ 2.399–2.415). In such cases, the number of protons giving such signals was taken into consideration with an assumption that T_1 values of these protons were not drastically different in different samples.

Data reduction and multivariate data analysis

All free induction decays (FIDs) from 1D ^{1}H NMR were multiplied with an exponential function with a line-broadening factor of 1 Hz prior to FT. ^{1}H NMR spectra obtained were corrected manually for phase and baseline distortion and calibrated to TSP at δ 0.0 using TOPSPIN (V2.0, Bruker Biospin). The spectral region δ 0.5–9 was segmented into buckets with the equal width of 0.002 ppm (1 Hz) using the AMIX package (V3.8.3, Bruker Biospin, Germany). The region δ 5.15–4.4 was discarded to eliminate the effects of imperfect water suppression. Each bucket was normalised to the total integral of the spectrum prior to multivariate data analysis using the SIMCA-P+ package (V.11, Umetrics AB, Sweden). Principal component analysis (PCA) was performed using mean-centered data to detect the general trends and outliers. Projection to latent structure discriminant analysis (PLS-DA) was applied with two components calculated and 5-fold cross-validation. The quality of models was evaluated with the R^2X and Q^2 values, reflecting the explained variables and the model predictability, and further assessed with rigorous permutation tests (n = 200).[27,28]

Results and discussion

^{1}H NMR spectroscopy of faecal extracts

Fig. 2 shows the ^{1}H NMR spectra of faecal extracts (F_1, F_2 and F_3) obtained from a female BALB/C mouse. Resonances were assigned to specific metabolites based on literature data[3,22,29] and public databases[30,31] with further confirmation by 2D NMR spectra. The mouse faecal extracts are dominated by about 40 metabolites (Fig. 2) including 4 SCFAs (formate, acetate, propionate and butyrate), 4 monosaccharides (glucose, arabinose, galactose and xylose), 18 amino acids such as glycine, alanine and histidine, tricarboxylic acid (TCA) cycle intermediates such as fumarate and succinate, bile acids, pyrimidine and purine metabolites such as uracil, adenine and hypoxanthine. Among them, Most of the metabolites have been reported in previous work.[3,22] Monosaccharides (except for glucose) were reported for the first time with NMR in faecal extracts although they were detected with HPLC previously.[32] Monosaccharides are listed in Table 1 and likely originated from non-digestive fibers. SCFAs are products of the gut microbiota fermentation of fibrous residues whereas amino acids, glucose, TCA cycle intermediates, pyrimidine and purine metabolites are likely from the endogenous metabolism of gut microbiota. Approximately 90% of bile acids are re-absorbed *via* enterohepatic circulation and the presence of a small amount of bile acids in faecal extracts are from the remaining non-absorbed residues. Therefore, the metabolite composition of faecal extracts contains metabolic information from host, diet, symbiotic gut microbiota and their symbiotic interactions.

Visual inspection of the NMR spectra showed that the overall spectral signal intensities of the third extracts (F_3) were substantially lower than the first one (F_1). However, the second extract still contained intense signals from a number of metabolites especially from SCFAs, amino acids, pyrimidine and purine metabolites. This observation indicates that a single extraction may not completely obtain all metabolites from a faecal sample and the combined first and second extracts is probably a better choice of sample to represent the metabolite composition of faeces for the purpose of metabolic profiling. In this way, faecal extracts contain all metabolites presented in a faecal sample but with limited dilution, hence maintaining reasonable SNR in NMR measurements.

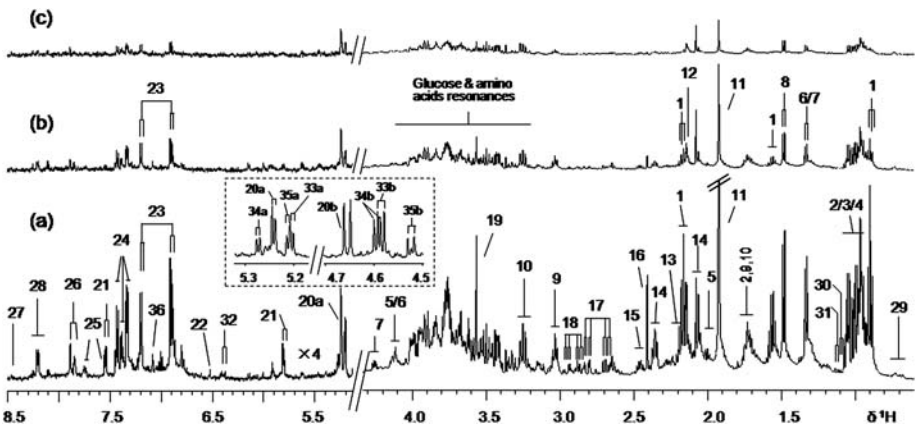

Fig. 2 500 MHz ^{1}H NMR spectra of faecal extracts obtained from a BALB/C mouse using the ratio of 1 : 10 (W_f : V_b, mg μl^{-1}) and 10 ultrasonic cycles. The spectra in the region (δ 5.1–8.5) were vertically expanded 4 times compared with the region (δ 0.6–4.4). (a) the first extracts. (b) the second extracts. (c) the third extracts. Key: 1, n-butyrate; 2, leucine; 3, isoleucine; 4, valine; 5, proline; 6, lactate; 7, threonine; 8, alanine; 9, lysine; 10, arginine; 11, acetate; 12, methionine; 13, propionate; 14, glutamine; 15, glutamate; 16, succinate; 17, aspartate; 18, asparagine; 19, glycine; 20a, α-glucose; 20b, β-glucose; 21, uracil; 22, fumarate; 23, tyrosine; 24, phenylalanine; 25, tryptophan; 26, hypoxanthine; 27, formate; 28, adenine; 29, bile acids; 30, α-keto-β-methyl-*N*-valerate; 31, α-keto-isovalerate; 32, urocanate; 33a, α-xylose; 33b, β-xylose; 34a, α-galactose; 34b, β-galactose; 35a, α-arabinose; 35b, β-arabinose; 36, histidine.

Table 1 NMR data for monosaccharides in the faecal extracts

Metabolites (Key)	Proton group	δ ^{1}H (multiplicity)(ppm)[a]	δ ^{13}C/ppm
α-glucose (20a)	1-CH; 2-CH; 6-CH; 6-CH′	5.24(d); 3.54(dd); 3.87(m); 3.75(dd)	93.1 (5.24)[b]; 75.2(3.54)
β-glucose (20b)	1-CH; 2-CH; 6-CH; 6-CH′	4.65(d); 3.24(dd); 3.75(dd); 3.90(dd)	97.1 (4.65); 77.9 (3.24)
α-xylose (33a)	1-CH; 2-CH; 3-CH	5.20(d); 3.53(dd); 3.68(m)	93.12 (5.20)
β-xylose (33b)	1-CH; 2-CH; 3-CH	4.58(d); 3.24(dd); 3.47(t)	97.75 (4.58)
α-galactose (34a)	1-CH; 2-CH; 3-CH	5.27(d); 3.81(dd); 3.97(m)	93.28 (5.27); 72.1 (3.81)
β-galactose (34a)	1-CH; 2-CH; 3-CH	4.59(d); 3.49(dd); 3.67(m)	97.75 (4.59)
α-arabinose (35a)	1-CH; 2-CH	5.21(d); 3.87(dd)	92.01 (5.21);63.8 (3.87)
β-arabinose (35a)	1-CH; 2-CH; 3-CH	4.52(d); 3.52(dd); 3.69(m)	97.7 (4.52)

[a] d, doublet; t, triplet; m, multiplets; dd, doublet of doublet. [b] the numbers in parentheses indicate chemical shifts of the directly attached protons obtained from HSQC spectra.

pH consistency of faecal extracts

Inter-sample pH consistency is crucial to minimise metabolite chemical shift variations for effective data analysis in the NMR-based metabonomics studies. Such pH consistency is normally achievable by using buffer systems during sample preparation.[33,34] Previous results have already shown that when pH variation is smaller than 0.1 unit around neutral pH (*i.e.*, pH 7.35 $\pm$ 0.1), the chemical shift consistency for most metabolites can be effectively maintained.[35] In this study, we comprehensively evaluated the effects of four FBRs ($W_f : V_b$) and two different (duration of) ultrasonic treatments on the pH consistency of the resultant faecal extracts from the six mice. Fig. 3 shows the distributions of pH values for the faecal extracts as a function of $W_f : V_b$ and from two extraction procedures. Ultrasonication durations (5 or 10 cycles) appear to have little effects on the sample pH values.

However, the inter-sample pH variations increase clearly with the rise of $W_f : V_b$ for the first extracts (F_1) and the combined extracts F_{12} (*i.e.*, $F_1 + F_2$) with the largest pH variation (about 0.3 unit) observed for F_1 with $W_f : V_b$ of 1 : 7 (mg μl^{-1}). The inter-sample variations for pH values of F_1 were also greater than 0.1 unit required for pH consistency when $W_f : V_b$ was greater than 1 : 10. Although some F_1 extracts ($W_f : V_b$ below 1 : 20) and all F_2 and F_3 extracts had good pH consistency (variation smaller than 0.1), they cannot be the best choice due to incomplete extraction for F_1 and sample dilutions (in the case of F_2 and F_3) which may cause severe loss of SNRs. The combination extracts (F_{12}) had reasonably small pH variations (less than 0.1 unit) for all samples with $W_f : V_b$ up to 1 : 10 (mg μl^{-1}).

The intra-group averaged SNR of six parallel samples was calculated from the same region of alanine (1.51 to 1.483 ppm) against the same baseline region (−2 to −3 ppm). Alanine was used here as an example compound because its resonance was well separated from other signals in ^{1}H NMR spectra of faecal extracts. The results showed obvious SNR reduction with the decreasing $W_f : V_b$ ratio. For example, SNR was reduced more than 50% when $W_f : V_b$ was decreased from 1 : 10 to 1 : 33–1 : 20 (mg μl^{-1}) (Fig. 4) for F_1, F_{12} and F_{123} extracts. When $W_f : V_b$ was 1 : 10, such SNR reduction was more than 50% from F_1 to F_{123} whereas only less than a 25% SNR reduction was evident from F_1 to F_{12}. Therefore, taking into consideration of pH consistency, maximum SNR and completeness of metabolite extraction, the optimum $W_f : V_b$ ought to be about 1 : 10 (mg μl^{-1}) with the combination of F_1 and F_2.

Chemical-shift consistency of faecal extracts

We further assessed the metabolite chemical-shift consistency although such consistency was often ensured with good pH consistency.[35,36] For this purpose, we calculated the standard deviations (δ, Hz)[35,36] for metabolite chemical shifts from six faecal samples extracted with different $W_f : V_b$ ratios and ultrasonication periods. Since the line-widths of ^{1}H NMR signals are

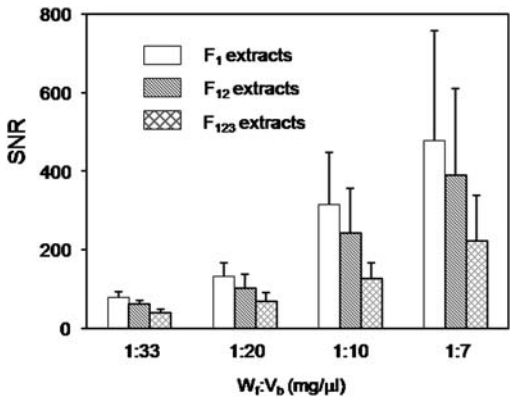

Fig. 4 Averaged NMR signal-to-noise ratios of alanine methyl signal from the first faecal extracts (F_1) and the combined extracts (F_{12} and F_{123}) using 10 ultrasonication cycles.

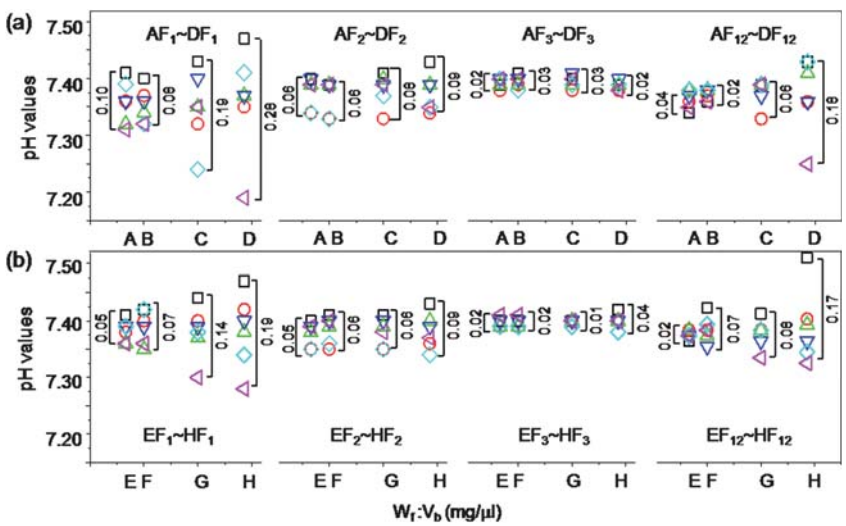

Fig. 3 The pH values obtained from faecal extracts as function of faeces-to-buffer ratios with (a) 5 ultrasonication cycles and (b) 10 ultrasonication cycles. Keys: A to D respectively represent four $W_f : V_b$ ratios (1 : 33; 1 : 20; 1 : 10; 1 : 7 mg μl^{-1}) with 5 ultrasonication cycles; E to H represent $W_f : V_b$ ratios of 1 : 33, 1 : 20, 1 : 10 and 1 : 7 (mg μl^{-1}), respectively, with 10 ultrasonication cycles; extraction fractions are denoted as F_1, F_2 and F_3 for the first, second, third extracts, respectively, and F_{12} is the combination of F_1 and F_2 with equal volume. The coloured symbols represented pH values of samples.

normally about 2 Hz taking line-broadening factors into consideration, it is thus required to limit the standard deviation of the metabolite chemical shifts to 2 Hz so as to maintain data resolution in the subsequent statistical analysis. Here five typical faecal metabolites were subjected to such assessment as representatives for monocarboxylic acids (acetate and lactate), dicarboxylic acids (succinate), amino acids with two ionisable groups (glycine) and multiple ionisable groups (histidine), respectively. Fig. 5 shows that all metabolites in faecal extracts (both F_1 and F_{12}) have increased chemical-shift variations with the increase of $W_f : V_b$ ratios although such changes are much more obvious for histidine. For all metabolites concerned, the standard deviation of chemical shifts was clearly smaller than 2 Hz for the first extracts (F_1) when $W_f : V_b$ ratios were smaller than 1 : 20 whereas it was about or greater than 2 Hz for histidine signals when $W_f : V_b$ ratios were greater than 1 : 10.

In contrast, such standard deviation of chemical shifts was safely kept within 2 Hz for all faecal metabolites in the combined extract (F_{12}) when $W_f : V_b$ ratios was 1 : 10 or smaller, being acceptable for high resolution multivariate data analysis. Moreover, the results from 10 min ultrasonication (Fig. 5b) appeared to have slight improvements for the metabolite chemical-shift consistency than the 5 min treatment (Fig. 5a). Therefore, the combined first two extracts with $W_f : V_b$ ratio of 1 : 10 and 10 min ultrasonication appeared to be optimum for chemical-shift consistency as well. Since two imidazole protons of histidine had pK_a (~6.04) close to our targeted pH (~7.4), their signals tended to have relatively large chemical-shift inconsistency which agreed well with the findings in the case of urine samples.[35] Careful attention and occasional corrections might be necessary for these two peaks prior to multivariate data analysis

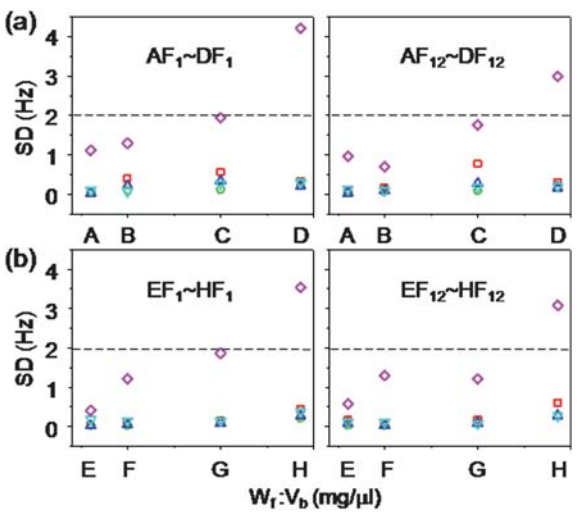

Fig. 5 Standard deviation of chemical shifts for five representative metabolites in extracts obtained from (a) 5 ultrasonication cycles and (b) 10 ultrasonication cycles. Key: A–D: A to D respectively represent four $W_f : V_b$ ratios (1 : 33; 1 : 20; 1 : 10; 1 : 7 mg μl^{-1}) with 5 ultrasonication cycles; E to H represent $W_f : V_b$ ratios of 1 : 33, 1 : 20, 1 : 10 and 1 : 7 (mg μl^{-1}), respectively, with 10 ultrasonication cycles; extraction fractions are denoted as F_1, for the first extracts, and F_{12} as the combination of F_1 and F_2 with equal volume. Key for symbols: histidine (◇), glycine (□), succinate (○), acetate (△), lactate (▽).

although our present results showed acceptable chemical-shift consistency.

The metabolite concentration ratios (MCRs) were calculated to further evaluate the above considerations in terms of composition differences of the faecal extracts. If MCRs were the same in F_1, F_2 and F_3, then F_1 would be a good representation of faecal metabolite compositions. Four metabolites, namely, glucose, acetate, butyrate and succinate, were chosen for this purpose to represent monosaccharides, SCFAs and TCA intermediates since these four metabolites had well resolved signals. The MCRs for Glu/But, Glu/Ace and Glu/Suc (Table 2) in the first extracts (F_1) were significantly different from the second extracts (F_2) except for the Glu/But ratio at 5 sonication cycles. This implies that the metabolite composition of F_1 cannot represent that of faecal samples and the combined extracts are probably a better choice. Furthermore, no significant differences for MCRs were found between F_{12} and F_{123}. This indicates that F_{12} is a good choice of sample representing faecal metabolite composition with consideration of SNR losses in F_{123}. We further analysed the composition-wide differences between F_{12} and F_{123} using multivariate data analysis approach. Although PLS-DA models showed visual separations between F_{12} and F_{123}, such models failed rigorous permutation tests (Fig. S1, ESI†). Therefore, the assumption that F_{12} and F_{123} were two separate groups was not valid; thus no significant compositional differences were present between F_{12} and F_{123}. Based on the above discussions, the combined first two extracts, F_{12}, stood out as the choice of samples with good representation for faecal extract composition, acceptable pH consistency (thus chemical shift consistency) and minimised signal-to-noise ratios (Fig. 4).

Comparison of three different extraction methods

The faecal metabonomic analysis requires complete and efficient extraction methods with reasonable simplicity to meet the requirement of high throughout.

Based on the parameters optimised above, we conducted a comparative metabonomic analysis on faecal extracts obtained from three different homogenisation methods, including ultrasonication, automatic homogenisation with tissuelyser and their combination. ^{1}H NMR spectral analysis showed that the numbers of detected metabolites were the same for faecal extracts obtained

Table 2 Averaged concentration ratios for some metabolites in faecal extracts[a,c]

	Glu/But	Glu/Ace	Glu/Suc
Extracts obtained with 5 sonication cycles			
F_1	1.65 ± 0.92	0.60 ± 0.20	13.80 ± 4.23
F_2	2.59 ± 1.26	1.27 ± 0.38[b]	29.03 ± 7.14[b]
F_{12}	2.34 ± 1.31	0.80 ± 0.29	17.93 ± 4.39
F_{123}	2.62 ± 1.32	0.92 ± 0.30	20.57 ± 5.09
Extracts obtained with 10 sonication cycles			
F_1	1.48 ± 0.32	0.53 ± 0.10	12.25 ± 4.63
F_2	2.46 ± 0.54[b]	1.11 ± 0.23[b]	24.45 ± 3.30[b]
F_{12}	1.99 ± 0.51	0.67 ± 0.14	15.07 ± 4.86
F_{123}	1.96 ± 0.28	0.72 ± 0.15	17.05 ± 5.08

[a] The faeces-to-buffer ratio was 1 : 10 ($W_f : V_b$, mg μl^{-1}). [b] Significant differences were found between F_1 and F_2 ($P < 0.05$). [c] No significant differences were found between F_{12} and F_{123}.

Table 3 PCA and PLS-DA results for faecal extracts obtained from different extraction methods

		PCA		PLS-DA		
		R^2X	Q^2	R^2X	Q^2	Permutation Y or N[a]
First extracts	F^U_1 *vs.* F^T_1[b]	0.414	0.062	0.328	0.097	N
	F^U_1 *vs.* F^{UT}_1	0.412	0.097	0.163	−0.506	N
	F^T_1 *vs.* F^{UT}_1	0.435	0.141	0.191	−0.526	N
Combined extracts	F^U_{12} *vs.* F^T_{12}[c]	0.41	0.048	0.319	0.418	N
	F^U_{12} *vs.* F^{UT}_{12}	0.396	0.065	0.141	−0.521	N
	F^T_{12} *vs.* F^{UT}_{12}	0.432	0.102	0.374	−0.37	N

[a] Y: the model is valid. N: the model is invalid. [b] F^U_1, F^T_1 and F^{UT}_1 denote the first extracts obtained from ultrasonication, tissuelyser and the combination of both treatments, respectively. [c] F_{12}: Combination of the first two extracts.

from these three methods (Fig. S2, ESI†). Further multivariate data analysis was conducted to assess whether there were metabolite composition differences resulting from extraction methods. Although PLS-DA scores plots (Fig. S3, S4, ESI†) showed some visual separations between faecal extracts obtained from three extraction methods, validity assessment results from Q^2 (Table 3) and permutation tests revealed that these PLS-DA models were invalid when considering the faecal extracts from three different extraction methods as different classes of samples (Fig. S3, S4, ESI†). This implies that there were no significant metabolite-composition differences between extracts obtained from these three homogenisation methods. Therefore, we recommended the tissuelyser extraction as the method of choice since it is efficient, simple to use and capable of high throughput studies.

Conclusions

NMR-based faecal metabonomic analysis offers a great opportunity for understanding interactions between mammalian hosts and gut microbiota. Based on an extensive optimisation of extraction methods and parameters, we recommend 1 : 10 (mg μl^{-1}) as optimum faeces-to-buffer ratio, phosphate buffer (0.1M, pH 7.4) as solvent and the combination of the first two extracts as targeted samples. We also recommend the use of semi-automatic tissuelyser as the choice of homogenisation method. Such protocol enables the maximum extractability, high throughput and reasonable SNRs together with good inter-sample pH and chemical shift consistency. In the case of unavailability for tissuelyser, manual ultrasonication methods can still fulfil the same tasks with limitation to throughput; in such case, the recommended optimal ultrasonication duration is ten cycles in the manner of (20 s) ultrasonication–(10 s) vortex–30 s) break.

Acknowledgements

We acknowledge the financial supports from the Ministry of Science and Technology (2007CB914701 for HT, 2009CB118804 for YW), National Natural Science Foundation of China (20775086, 20825520 for HT and 20775087 for YW) and the Knowledge Innovation Program of the Chinese Academy of Sciences (KSCX1-YW-02 for HT, KJCX2-YW-W11 for YW). We are also grateful for some useful comments from anonymous reviewers which enabled us to improve our manuscript.

References

1 J. Lederberg, *Science*, 2000, **288**, 287–293.
2 F. P. Martin, M. E. Dumas, Y. Wang, C. Legido-Quigley, I. K. Yap, H. Tang, S. Zirah, G. M. Murphy, O. Cloarec, J. C. Lindon, N. Sprenger, L. B. Fay, S. Kochhar, P. van Bladeren, E. Holmes and J. K. Nicholson, *Mol. Syst. Biol.*, 2007, **3**, 112.
3 F. P. Martin, N. Sprenger, I. K. Yap, Y. Wang, R. Bibiloni, F. Rochat, S. Rezzi, C. Cherbut, S. Kochhar, J. C. Lindon, E. Holmes and J. K. Nicholson, *J. Proteome Res.*, 2009, **8**, 2090–2105.
4 F. P. Martin, Y. Wang, N. Sprenger, I. K. Yap, T. Lundstedt, P. Lek, S. Rezzi, Z. Ramadan, P. van Bladeren, L. B. Fay, S. Kochhar, J. C. Lindon, E. Holmes and J. K. Nicholson, *Mol Syst Biol*, 2008, **4**, 157.
5 D. A. Peterson, D. N. Frank, N. R. Pace and J. I. Gordon, *Cell Host Microbe*, 2008, **3**, 417–427.
6 I. D. Wilson and J. K. Nicholson, *Curr. Pharm. Des.*, 2009, **15**, 1519–1523.
7 Y. Wang, H. Tang, J. K. Nicholson, P. J. Hylands, J. Sampson and E. Holmes, *J. Agric. Food Chem.*, 2005, **53**, 191–196.
8 L. Ding, F. Hao, Z. Shi, Y. Wang, H. Zhang, H. Tang and J. Dai, *J. Proteome Res.*, 2009, **8**, 2882–2891.
9 I. K. Yap, T. A. Clayton, H. Tang, J. R. Everett, G. Hanton, J. P. Provost, J. L. Le Net, C. Charuel, J. C. Lindon and J. K. Nicholson, *J. Proteome Res.*, 2006, **5**, 2675–2684.
10 Y. Wang, E. Holmes, J. K. Nicholson, O. Cloarec, J. Chollet, M. Tanner, B. H. Singer and J. Utzinger, *Proc. Natl. Acad. Sci. U. S. A.*, 2004, **101**, 12676–12681.
11 Y. Wang, J. Utzinger, J. Saric, J. V. Li, J. Burckhardt, S. Dirnhofer, J. K. Nicholson, B. H. Singer, R. Brun and E. Holmes, *Proc. Natl. Acad. Sci. U. S. A.*, 2008, **105**, 6127–6132.
12 Y. Wang, S. H. Xiao, J. Xue, B. H. Singer, J. Utzinger and E. Holmes, *J. Proteome Res.*, 2009, **8**, 5442–5450.
13 J. F. Wu, E. Holmes, J. Xue, S. H. Xiao, B. H. Singer, H. R. Tang, J. Utzinger and Y. L. Wang, *Int. J. Parasitol.*, 2010, DOI: 10.1016/j.ijpara.2009.11.003.
14 X. Zhang, Y. Wang, F. Hao, X. Zhou, X. Han, H. Tang and L. Ji, *J. Proteome Res.*, 2009, **8**, 5188–5195.
15 M. Li, B. Wang, M. Zhang, M. Rantalainen, S. Wang, H. Zhou, Y. Zhang, J. Shen, X. Pang, H. Wei, Y. Chen, H. Lu, J. Zuo, M. Su, Y. Qiu, W. Jia, C. Xiao, L. M. Smith, S. Yang, E. Holmes, H. Tang, G. Zhao, J. K. Nicholson, L. Li and L. Zhao, *Proc. Natl. Acad. Sci. U. S. A.*, 2008, **105**, 2117–2122.
16 R. E. Ley, M. Hamady, C. Lozupone, P. J. Turnbaugh, R. R. Ramey, J. S. Bircher, M. L. Schlegel, T. A. Tucker, M. D. Schrenzel, R. Knight and J. I. Gordon, *Science*, 2008, **320**, 1647–1651.
17 J. Delaney, W. A. Neville, A. Swain, A. Miles, M. S. Leonard and C. J. Waterfield, *Biomarkers*, 2004, **9**, 271–290.
18 D. M. Jacobs, N. Deltimple, E. van Velzen, F. A. van Dorsten, M. Bingham, E. E. Vaughan and J. van Duynhoven, *NMR Biomed.*, 2008, **21**, 615–626.
19 J. R. Marchesi, E. Holmes, F. Khan, S. Kochhar, P. Scanlan, F. Shanahan, I. D. Wilson and Y. Wang, *J. Proteome Res.*, 2007, **6**, 546–551.
20 D. Monleón, J. M. Morales, A. Barrasa, J. A. Lopez, C. Vazquez and B. Celda, *NMR Biomed.*, 2009, **22**, 342–348.
21 M. Ndagijimana, L. Laghi, B. Vitali, G. Placucci, P. Brigidi and M. E. Guerzoni, *Int. J. Food Microbiol.*, 2009, **134**, 147–153.
22 J. Saric, Y. Wang, J. Li, M. Coen, J. Utzinger, J. R. Marchesi, J. Keiser, K. Veselkov, J. C. Lindon, J. K. Nicholson and E. Holmes, *J. Proteome Res.*, 2008, **7**, 352–360.
23 J. Pettersson, P. C. Karlsson, Y. H. Choi, R. Verpoorte, J. J. Rafter and L. Bohlin, *Biol. Pharm. Bull.*, 2008, **31**, 1192–1198.
24 J. Saric, J. V. Li, Y. Wang, J. Keiser, J. G. Bundy, E. Holmes and J. Utzinger, *PLoS Neglected Trop. Dis.*, 2008, **2**, e254.
25 R. Nakaune and M. Nakano, *J. Virol. Methods*, 2006, **134**, 244–249.
26 E. Veronesi, P. P. Mertens, A. E. Shaw, J. Brownlie, P. S. Mellor and S. T. Carpenter, *J. Med. Entomol.*, 2008, **45**, 129–132.
27 C. M. Slupsky, K. N. Rankin, J. Wagner, H. Fu, D. Chang, A. M. Weljie, E. J. Saude, B. Lix, D. J. Adamko, S. Shah, R. Greiner, B. D. Sykes and T. J. Marrie, *Anal. Chem.*, 2007, **79**, 6995–7004.
28 J. A. Westerhuis, J. C. Hoefsloot, S. Smit, D. J. Vis, A. K. Smilde, E. J. van Velzen, J. P. M. van Duijnhoven and F. A. van Dorsten, *Metabolomics*, 2008, **4**, 81–89.

29 W. M. T. Fan, *Prog. Nucl. Magn. Reson. Spectrosc*, 1996, **28**, 161–219.
30 Q. Cui, I. A. Lewis, A. D. Hegeman, M. E. Anderson, J. Li, C. F. Schulte, W. M. Westler, H. R. Eghbalnia, M. R. Sussman and J. L. Markley, *Nat. Biotechnol.*, 2008, **26**, 162–164.
31 D. S. Wishart, C. Knox, A. C. Guo, R. Eisner, N. Young, B. Gautam, D. D. Hau, N. Psychogios, E. Dong, S. Bouatra, R. Mandal, I. Sinelnikov, J. Xia, L. Jia, J. A. Cruz, E. Lim, C. A. Sobsey, S. Shrivastava, P. Huang, P. Liu, L. Fang, J. Peng, R. Fradette, D. Cheng, D. Tzur, M. Clements, A. Lewis, A. De Souza, A. Zuniga, M. Dawe, Y. Xiong, D. Clive, R. Greiner, A. Nazyrova, R. Shaykhutdinov, L. Li, H. J. Vogel and I. Forsythe, *Nucleic Acids Res.*, 2009, **37**, D603–610.
32 J. L. Slavin, J. A. Marlett and M. J. Neilson, *J Nutr*, 1983, **113**, 2353–2359.
33 M. A. Constantinou, E. Papakonstantinou, M. Spraul, S. Sevastiadou, C. Costalos, M. A. Koupparis, K. Shulpis, A. Tsantili-Kakoulidou and E. Mikros, *Anal. Chim. Acta*, 2005, **542**, 169–177.
34 A. W. Nicholls, E. Holmes, J. C. Lindon, J. P. Shockcor, R. D. Farrant, J. N. Haselden, S. J. P. Damment, C. J. Waterfield and J. K. Nicholson, *Chem. Res. Toxicol.*, 2001, **14**, 975–987.
35 C. Xiao, F. Hao, X. Qin, Y. Wang and H. Tang, *Analyst*, 2009, **134**, 916–925.
36 M. Lauridsen, S. H. Hansen, J. W. Jaroszewski and C. Cornett, *Anal. Chem.*, 2007, **79**, 1181–1186.

