Analyst

Accepted Manuscript

This is an *Accepted Manuscript*, which has been through the RSC Publishing peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, which is prior to technical editing, formatting and proof reading. This free service from RSC Publishing allows authors to make their results available to the community, in citable form, before publication of the edited article. This *Accepted Manuscript* will be replaced by the edited and formatted *Advance Article* as soon as this is available.

To cite this manuscript please use its permanent Digital Object Identifier (DOI®), which is identical for all formats of publication.

More information about *Accepted Manuscripts* can be found in the **Information for Authors**.

Please note that technical editing may introduce minor changes to the text and/or graphics contained in the manuscript submitted by the author(s) which may alter content, and that the standard **Terms & Conditions** and the **ethical guidelines** that apply to the journal are still applicable. In no event shall the RSC be held responsible for any errors or omissions in these *Accepted Manuscript* manuscripts or any consequences arising from the use of any information contained in them.

An optimized method for NMR-based plant seed metabolomic analysis with maximized polar metabolite extraction efficiency, signal-to-noise ratio, and chemical shift consistency

Xiangyu Wu[a], Ning Li[a], Hongde Li[a] and Huiru Tang*[a#]

[a] *Key Laboratory of Magnetic Resonance in Biological Systems, State Key Laboratory of Magnetic Resonance and Atomic and Molecular Physics, Wuhan Centre for Magnetic Resonance, Wuhan Institute of Physics and Mathematics, University of Chinese Academy of Sciences, Wuhan, 430071, P. R. China.*

[#] *State Key Laboratory of Genetic Engineering, Biospectroscopy and Metabolomics, School of Life Sciences, Fudan University, Shanghai 200433, P.R. China*

* To whom all correspondence should be addressed. Huiru.tang@wipm.ac.cn; Tel, +86-27-87198430; Fax, +86-27-87199291.

Abstract

Plant metabolomic analysis has become an essential part of functional genomics and systems biology which requires effective extraction of both primary and secondary metabolites from plant cells. To establish an optimized extraction method for the NMR-based analysis, we used the seeds of mungbean (*Vigna radiata* cv. Elü No.1) as model and systematically investigated the dependence of the metabolite composition in plant extracts on various extraction parameters including cell-breaking methods, extraction solvents, number of extraction repeats, tissue-to-solvent ratio, and extract-to-buffer ratio (for final NMR analysis). We also compared two NMR approaches for quantitative metabolomic analysis from completely relaxed spectra directly and from partially relaxed spectra calculated with T_1. By maximizing the extraction efficiency and signal-to-noise ratio but minimizing inter-sample chemical-shift variations and metabolite degradations, we established a parameter-optimized protocol for the NMR-based plant seed metabolomic analysis. We concluded that aqueous methanol was the best extraction solvent with the optimal tissue-to-solvent ratio of about 1:10-1:15 (mg:μL). The combination of tissuelyser homogenization with ultrasonication was the choice of cell-breaking method with three-time repeated extractions being necessary. For NMR analysis, the optimal extract-to-solvent was around 5-8 mg/mL and completely relaxed spectra were ideal for intrinsically quantitative metabolomic analysis although partially relaxed spectra were employable for comparative metabolomics. This optimized method will offer ensured data quality for high-throughput and reliable plant metabolomics studies.

Key words: NMR, plant metabonomics, metabolite extraction, quantitative analysis

Introduction

Metabolomics/metabonomics has now become a well-accepted approach for holistic understanding of the molecular aspects of gene functions, physiology and pathophysiology of mammals and plants. This is because metabolomics (and/or metabonomics) offers information on the global metabolite composition (i.e., metabolome) of a given biological system and its dynamic metabolic responses towards various stimuli[1-3]. Metabolome also serves as a link to genotype and macroscopic phenotype[4, 5] providing the metabolic phenotype (or metabotype) as a new form of molecular phenotype[4]. Therefore, metabolomic analysis has been successfully employed to probe the molecular aspects of pathogenesis[6-9], toxicity[10-13], stress effects on biochemistry[14-16] as well as disease diagnosis and prognosis[17-19]. In general, metabolomic analysis is achieved by measuring the concentration of all detectable metabolites in various biological matrices[11, 20] using mainly NMR and the hyphenated chromatography-mass spectrometry (including LC-MS and GC-MS) techniques although the hyphenated LC-DAD-NMR/MS techniques have become increasingly important for detecting and identifying novel metabolites[2, 21, 22]. Such matrices include biofluids[23, 24], cells[15, 25, 26], tissues[27-29] and extracts[4, 11, 14].

For plants, metabolomic analysis has already become an essential approach for understanding biochemistry, gene functions and stress-adaptations[14, 30-33]. With the presence of tough cell walls, such analysis for plants often has very different requirements from what needs for animal models especially when metabolite extraction is needed. Whilst metabolic profiles and responses towards different stimuli for mammals can be easily sampled *via* both biofluids (e.g., urine and serum) and tissues, metabolic features for plants are normally sampled through specific organs and tissues. Under such circumstances, metabolite extraction is a matter of necessity for detailed metabolomic analysis demanding appropriate cell-breaking methods and extraction solvents even though high-resolution magic-angle spinning NMR (HRMAS-NMR) techniques are useful for non-destructive tissue metabolite analysis[27-29]. Plant cells also contain rich secondary metabolites which are of great importance for plant self-defense and bioactivities[34-36]. Therefore, plant cells and tissues require different metabolite extraction strategies from those used for animal tissues. Furthermore, different tissues from different plant organs (e.g. leaves, seeds and roots) may have fairly different physical properties thus require different metabolite extraction methods during metabolomic analysis. Unlike mammalian samples for which the optimized metabolomic analysis methods have been well established[37-42], the optimized methods remain to be fully established for many plant tissues such as seeds even though many works have already been published on plant metabolomic protocols[43-46] for some

plant tissues including plant roots[4], above-ground tissues[44, 47], stems[22] and whole plants[14].

For metabolomic analysis, plant tissues are often pulverized (into powder) followed by solvent extractions with some cell-breaking methods. For pulverization, mortar and pestle are usable for grinding soft tissues such as leaves and stems[32, 48] in liquid nitrogen whilst an electrical grinder is suitable for pulverizing most frozen and dry plant tissues[33, 49] including above-ground parts[14, 44] and roots[4, 31]. For cell-breaking, an ultrasonic oscillator[31, 44, 50, 51], tissuelyser and their combinations are useable[44, 50]. Several solvent systems have been used for plant metabolite extractions including aqueous methanol[4, 31, 50], aqueous ethanol[52], aqueous acetonitrile[4, 14, 31, 44], water[53], perchloric acid solution[54] and trichloroacetic acid solution[54, 55]. However, it remains unclear which of these is suitable for different plant tissues such as seeds. The optimal tissue-to-solvent ratio (TSR) is another important parameter requiring optimization for plant metabolome extraction with the reported values ranging from 1:3 to 1:50 (mg:μL)[49-51, 56, 57]. How many times the extraction procedures ought to be repeated to achieve a complete metabolite extraction also remains to be optimized for many plant tissues such as seeds. It is understandable that these parameters will affect metabolite extracting efficiency and thus analytical sensitivity. As has been systematically done for plant roots[4, 31] and mammalian samples[11, 40], optimization of these factors is critically important for seeds since they will undoubtedly affect the metabolite composition of the resultant extracts[4, 31, 58].

Moreover, quantification of all detectable metabolites is ideal to achieve quantitative metabolomic analysis especially those metabolites having significant variations. With commonly employed ^{1}H NMR techniques, one can either acquire the completely relaxed or partially relaxed spectra for this purpose. In the former, the molar concentration of a given proton (of a metabolite) is directly proportional to the resonance integral of that proton whereas in the latter such concentration ought to be calculated from the resonance integral and T_1 relaxation time of the given protons[4, 31]. In all cases, nevertheless, an inert internal standard is required with known concentration and its T_1 value. It is worth noting that so far most metabolomic studies have used data from the partially relaxed spectra for the multivariate data analysis. However, it remains unclear whether the same results are obtainable from the completely and partially relaxed spectra especially in the case of comparative plant metabolomic analysis.

In order to establish an optimized method for plant seed metabolomics studies, in this work, we employed the germinated seeds of mungbean (*Vigna radiata* cv. Elü No.1) as a model system and systematically assessed multiple parameters for metabolite extraction including cell-breaking methods, extraction solvents, tissue-to-solvent ratio (TSR), the numbers of repeat extractions, and extract-to-buffer

ratio (EBR). We aimed to maximize the extraction efficiency, signal-to-noise ratio (SNR) and inter-sample chemical shift consistency in NMR detections with assessment of the T_1 relaxation effects on plant metabolomic analysis. Based on these, we established a parameter-optimized method for the NMR-based plant seed metabolomics.

1. Experimental section

1.1 Materials

Mungbean (*Vigna radiata* cv. Elü No.1) seeds were obtained from the Institute of Food Crops, Hubei Academy of Agriculture Science. Methanol, acetonitrile, perchloric acid (pcA), trichloroacetic acid (TCA), potassium carbonate (K_2CO_3), $K_2HPO_4 \cdot 3H_2O$, $NaH_2PO_4 \cdot 2H_2O$ and sodium hypochlorite were all purchased from Sinopharm Chemical Reagent Co. Ltd. (Shanghai) as analytical grade. D_2O (99.9% D) and sodium 3-trimethlysilyl [2,2,3,3-D_4] propionate (TSP) were purchased from Sigma-Aldrich Inc. Double distilled water from an Elix Advantage System (Merck Millipore, Germany) was employed for preparation of all buffers and aqueous solutions used in this study. Aqueous methanol (50%) and acetonitrile (50%) used for metabolite extraction were prepared by mixing organic solvents with water, respectively, and known as aqueous methanol and aqueous acetonitrile in the following text. Phosphate buffer (PB) was prepared by dissolving K_2HPO_4 and NaH_2PO_4 in water with an appropriate molar ratio (about 4:1). Icy and hot buffers were PB buffers (0.1 M, pH 7.4) with a temperature of 0□ and 90□ respectively.

1.2 Treatment of mungbean seeds

Mungbean seeds were sterilized with 0.1% sodium hypochlorite solution for fifteen minutes and then washed with water for seven times. About five hundred such seeds were uniformly placed in ten petri dishes having two layers of sterilized and dampened filter paper. After such treatment at 26□ for sixteen hours, these seeds were collected individually and snap-frozen in liquid nitrogen followed with storage in -80□ freezer until further extraction treatments.

1.3 Optimization of the metabolite extraction

Above seed samples (about 100 mg) were individually ground into powder in liquid nitrogen with a mortar and pestle. The powder so-obtained was employed as raw materials for all optimization. Moisture contents (or dry weight) of seeds were measured from ten parallel samples of ground seed powder according to the standard oven method of the International Seed Testing Association[59].

1.3.1 Effects of cell-breaking methods and number of repeat extractions on metabolite extraction efficiency

The above powder (about 100 mg) was employed to assess the effects of different cell-breaking methods on the metabolite extraction efficiency with 50% aqueous methanol (1 mL) as solvent. In all cases, the mixtures of solvent and seed powder were vortex-mixed for ten seconds and kept in an ice-water bath for 10 minutes followed with the cell-breaking treatments. Four different cell-breaking methods were used here, namely, direct extraction with no treatment, homogenization with a tissuelyser (QIAGEN, Germany),

ultrasonication, and tissulyser homogenization plus ultrasonication. Tissuelyser homogenization was conducted at 20 Hz for 90 s twice with a break (90 s) in between; ultrasonication (40 KHz, 200 W) was done in an ice-bath with fifteen sonication-break (1 min each) cycles. After treatments, the mixtures were subjected to 10 minutes centrifugation (16000 x g, 4℃) to obtain supernatants. In all cases, such extraction procedure was repeated three times to obtain four supernatants for each sample. Following removal of organic solvent with a Savant SpeedVac® concentrator (SC110A, Thermo, Germany), the remaining solution was respectively lyophilized (for about 48 hours) to yield dry extracts.

NMR profiles of these four extracts from above four repeated extractions were analyzed for the optimal number of repeat extractions. The combined extracts from the first two, first three and all four extractions were also subjected to NMR analysis. For each individual extraction step, extract samples for NMR analysis were prepared by re-dissolving the freeze-dried extracts from each extraction, respectively, in 1.2 mL phosphate buffer (0.1 M, pH 7.4) containing 10% D_2O and 0.15 mM TSP. 540 μL of solution from each extraction was employed to assess the metabolite composition of four sequential extractions. The combined extracts from the first two extractions were obtained by mixing 270 μL above solutions of the first and second extracts (from the same sample). The combined extracts from the first three extractions were obtained by mixing 180 μL above solutions of the first, second and third extracts (from the same sample) whereas the combined extracts from the four extractions were prepared by mixing 135 μL above solutions of the first, second, third and forth extracts from the same sample. In any cases, ten biological replicates were used for each assessed group.

Extraction efficiency of extraction procedures was assessed using the extraction ratio (ER) calculated from the weight of extract against the dry weight of raw material.

1.3.2 Optimization of extraction solvents

Six different solvents (1.5 mL) were added to seed powders (about 100 mg) in parallel including icy buffer (IB), hot buffer (HB), aqueous acetonitrile (ACN), aqueous methanol (MeOH), 1 M pcA, and 1 M TCA. Ten biological replicates were used for each solvent group. All the mixtures of solvent and seed powder were vortex-mixed for ten seconds and kept in ice-water for 10 minutes except for samples in HB, which were kept in hot water (90□), instead. The combined tissuelyser homogenization and ultrasonication was used as the cell-breaking method and conducted as described above. Extraction was further repeated three times and four resultant supernatants were pooled together to be assessed. For ACN and MeOH extractions, all extracts were lyophilized for 48 hours and weighted after removal of organic solvents using

a Savant SpeedVac® concentrator. For pcA extraction, the supernatants were neutralized with K_2CO_3 solution (1 M) in an ice-bath followed with centrifugation to remove potassium perchlorate. The solution was then added with 10% (v/v) PB (1.5 M, pH 7.4) followed with lyophilization. Supernatants from TCA, IB and HB extractions were directly lyophilized. For NMR analysis, the dried extracts from MeOH and ACN were reconstituted in 1200 μL PB (0.1 M, pH 7.4) containing 10% D_2O and 0.15 mM TSP. The extracts from pcA, IB and HB were added with 1080 μL water and 120 μL D_2O containing 1.5 mM TSP. The extracts from TCA were added with 920 μL water, 40 μL K_2CO_3 solution (1 M), 120 μL phosphate buffer (1.5 M, pH 7.4) and 120 μL D_2O containing 1.5 mM TSP. For all extracts, 540 μL of the final solution was added to NMR tubes (i.d., 5 mm) respectively for NMR measurements.

1.3.3 Assessments for the optimal tissue-to-solvent ratio (TSR) and extract-to-buffer ratio (EBR)

For TSR optimization, aqueous methanol was used as solvent and the combined tissuelyser homogenization and ultrasonication was employed as cell-breaking method. Six sets of samples were prepared by using 1.2 mL aqueous methanol with seed weights of 240, 160 and 120 mg together with 2 mL solvent with seed weights of 133, 100 and 67 mg, respectively. In this way, six tissue-to-solvent ratios were assessed including 1:5, 1:7.5, 1:10, 1:15, 1:20 and 1:30 (mg:μL), respectively, with ten biological replicates for each set. For EBR optimization, six solutions were prepared with the extract concentration of 2.3, 3.5, 5.2, 8.3, 11.0 and 16.7 milligram freeze-dried extract per milliliter buffer (mg/mL).

1.3.4 Assessment of T_1 relaxation effects on quantitative NMR analysis

Ten seed samples (around 50 mg) were extracted three times using aqueous methanol (500 μL) and the combined tissuelyser homogenization and ultrasonication method described above. Three supernatants from the same sample were pooled together and lyophilized after removal of organic solvent under vacuum. The powder extracts were respectively re-dissolved in 600 μL phosphate buffer (0.1 M, pH 7.4) containing 10% D_2O and 0.15 mM TSP followed with NMR analyses (Figure S1).

1.4 NMR analyses

All NMR analyses for extraction parameter optimization were conducted on a Bruker AVII 500 MHz NMR spectrometer (Bruker BioSpin). ^{1}H NMR spectra were acquired at 298 K using a NOESYGPPR1D pulse sequence with a broad-band inverse probe for 5 mm tubes. Water signal was suppressed with a weak continuous wave (CW) irradiation during recycle delay (2 s) and mixing time (0.1 s). 90° pulse length was adjusted to about 10 μs for each sample and 64 transients were collected with 32 k data-points and the spectral width of 20 ppm (10 kHz).

For quantitative NMR and T_1 relaxation measurements, all ^{1}H NMR spectra were acquired at 298 K on a Bruker AVIII 600 MHz NMR spectrometer (Bruker BioSpin) equipped with an inverse cryogenic probe. The standard NOESYGPPR1D pulse sequence was employed with the recycle delay of 2 s for the partially and 22.5 s for completely relaxed spectra, respectively. In both cases, water suppression was achieved by CW irradiation for 2 s during recycle delay prior to the first 90^o pulse and during mixing time (0.1 s). 90° pulse length was adjusted to about 10 μs for each sample and 64 transients were collected with 64 k data-points and the spectral width of 20 ppm (12 kHz).

T_1 values of metabolites were measured with an inversion-recovery sequence. Water suppression was switched on only during the last two seconds of recycle delay (22.5 s). 90° pulse length was accurately set to 10 μs for each sample individually. Eighteen relaxation delays were employed (i.e., 0.05, 0.1, 0.14, 0.19, 0.27, 0.37, 0.52, 0.72, 1, 1.39, 1.93, 2.68, 5.18, 10, 22, 25, 28 and 32 s). Sixty-four transients were collected with 32 k data-points for each spectrum with the spectral width of 20 ppm (12 kHz). T_1 values for all metabolite protons were calculated from the signal integrals as a function of relaxation delays using TOPSPIN software (V 2.1, Bruker BioSpin). For quantification, peak integrals were obtained from the deconvolution results of ^{1}H NMR spectra using MestReNova (V 8.1, Mestrelab Research, Spain).

For metabolite signal assignment purposes, five two-dimensional NMR (2D NMR) spectra were recorded at 298 K for selected samples and processed with similar parameters described previously[4, 31, 44]. These included ^{1}H *J*-resolved spectroscopy (*J*-RES), ^{1}H-^{1}H correlation spectroscopy (COSY), ^{1}H-^{1}H total correlation spectroscopy (TOCSY), ^{1}H-^{13}C heteronuclear single quantum coherence (HSQC) and heteronuclear multiple bond correlation (HMBC) spectra. These spectra were all acquired on a Bruker AVIII 600 MHz NMR spectrometer equipped with an inverse cryogenic probe.

1.5 NMR data processing and analysis

One dimensional ^{1}H NMR spectra were phase- and baseline-corrected manually using TOPSPIN software with chemical shift referenced to the TSP signal (δ0.00). Signal-to-noise ratios (SNR) in ^{1}H NMR spectra were calculated from the citrate signal (δ 2.50-2.66) and noise region (δ -2.00 to -3.00) using standard method and TOPSPIN software.

For multivariate data analysis, spectral region δ0.80-9.60 was uniformly bucketed into bins with 1.5 Hz in width using AMIX software package (V 3.8.3, Bruker BioSpin). Spectral regions containing residual solvent signals for water (δ 4.50-5.23), methanol (δ 3.35-3.39) and acetonitrile (δ 2.07-2.09) were discarded to eliminate solvent signals. These binned data were normalized to the dry weight of mungbean

seeds (to obtain the absolute concentration information). Principal component analysis (PCA) was conducted with the mean-centered data whereas the orthogonal partial least square-discriminate analysis (OPLS-DA) was conducted from the pareto-scaled data with SIMCA P+ (V 12, Umetrics, Sweeden). For OPLS-DA, 7-fold cross-validation was used with NMR data as *X*-matrix and the group information as *Y*-matrix. Qualities of these OPLS-DA models were evaluated with R^2X values indicating the explained variations and Q^2 values representing the model predictabilities. All these models were further tested for their robustness with CV-ANOVA approach[60] to ensure the significance of intergroup differentiations (with $p<0.05$ as significant level). Loadings plots for OPLS-DA models were generated from the back-transformed data[61] with an in-house developed MATLAB script.

1.6 Inter-sample chemical shift variations for metabolite signals.

Seven typical metabolites were selected to assess the inert-sample chemical shift variations including lactate (containing a carboxyl group), succinate (containing two carboxyl groups), citrate (containing three carboxyl groups), uridine (containing two imide groups), histidine (containing an amino, carboxyl and imidazole group), tryptophan and N1-methyl nicotinate. The latter three metabolites also represented some metabolites containing aromatic rings. Since citrate has pK_a values[38] of 3.13, 4.70 and 6.40 whereas histidine has pK_a values[38] of 1.82, 6.04 and 9.33, these two metabolites are normally regarded as the most difficult metabolites in terms of maintaining their chemical shift consistency[38]. We calculated the standard deviations (SD) of the chemical shifts from ten biological replicates for these metabolites.

2 Results and discussion

An ideal metabolite extraction method ought to have maximized extraction efficiency and analytical sensitivity but minimized the changes to metabolites, interferences to metabolite signals and inter-sample chemical shift variations for metabolite signals. Extraction ratios can be readily used to assess the extraction efficiency whereas the concentration ratios for different metabolites are indicative to the extraction-induced metabolite changes. Furthermore, the effects of solvents on quality of metabolite signals can be easily assessed with the presence of extra signals and variations in signal chemical shift. In optimal conditions, inter-sample chemical shift variations ought to be smaller than half of a bin-width (about 0.75 Hz in this study). To meet these demands, appropriate parameters need considering and optimizing including cell-breaking methods, solvents, tissue-to-solvent ratio (TSR) and extract-to-buffer ratio (EBR) during NMR analysis.

2.1 Solvent effects on the metabolite compositions of extracts from mungbean seeds

Ideal solvents used for metabolite extractions ought to quench enzymatic reactions efficiently without causing the extra chemical reactions for all metabolites but be effective and efficient in extracting most metabolites. Solvents ought to have no signals and cause little changes to the quality of metabolite signals although it is also acceptable if such solvent effects can be easily removed. Therefore, we assessed six commonly used solvents here for their suitability, including icy buffer (IB) and hot buffer (HB), aqueous methanol (MeOH) and acetonitrile (ACN), solutions of perchoric acid (pcA) and trichloroacetic acid (TCA).

^{1}H NMR spectra of mungbean extracts resulting from six different solvents (Figure 1) all showed rich metabolite information. Fifty-four metabolites were readily assignable unambiguously according to the published data[54, 62, 63], which were further confirmed individually with a series of 2D NMR spectra. These mungbean seed metabolites included 16 amino acids, 12 organic acids, 8 nucleotide related metabolites, 12 carbohydrates, 4 choline metabolites, ethanol and some lipids. Inspection of the spectra indicated that extracts from most solvents differed from each other only in terms of metabolite concentration. This indicated that all six solvents were effective in extracting the abundant mungbean metabolites. However, TCA extract contained two extra signals at about δ9.46 (a singlet) and δ2.75 (a doublet) probably resulting from some impurities in TCA. Close inspection of the NMR spectra of mungbean extracts showed that both extracts from icy and hot buffer contained obviously more intense broad lipid signals (δ0.5-2.0) than the other four extracts from two organic solvents and two acid solutions (Figure 1). These lipid signals were

probably from saturated fatty acids in the soluble lipid-oligosaccharide complexes[64]. Broad signals probably from proteins were also evident (δ 6.5-9.0) especially in the icy and hot buffer extracts. The presence of these signals in NMR spectra will inevitably cause some difficulties for analysis of other metabolites in terms of spectral resolution, accuracy of signal integrals and baseline quality. NMR data also showed that the chemical shift variations for histidine were well above 0.75 Hz for extracts from icy and hot buffer (Figure 2a). Compared to the extracts from two organic solvents with good enzyme quenching capability, IB extracts contained substantially more glucose (Figure 2b) probably due to inefficient enzyme-quenching. However, HB extracts contained more sucrose and raffinose family oligosaccharides (RFOS) (Figure 2b) with no outstanding differences in the levels of galactose and glucose. This is consistent with the fact that increased temperature enhances the stress-induced sucrose accumulation[65, 66] and biosynthesis of RFOS[31, 67, 68]. All these indicate that icy and hot buffers are not ideally suitable for plant metabolomics.

TCA and pcA are well reported in literature for their uses of metabolite extractions especially for mammalian tissues[54]. However, our results showed that the inter-sample chemical-shift standard deviations were well above 1 Hz for succinate, citrate, uridine, tryptophan and histidine (about 5 Hz for histidine) in extracts from pcA and TCA (Figure 2a). This probably resulted from the difficulties to maintain uniform pH values for the different extracts with the use of strong acids although effects of high ionic strength could not be ignored. It is well known that chemical shifts of citrate and histidine are sensitive to both pH and ionic strength variations[38, 41] around pH 6-8 since they all have a $p\mathrm{K_a}$ value of about 6.0. We also found that lyophilization failed to remove TCA completely even when the vacuum was 0.02 mbar as reported previously[55] although another previous report suggested otherwise[54]. In both cases, neutralization process cannot be easily done in a high throughput manner and high salt concentration resulting from neutralization will inevitably cause sensitivity losses[38] as well. Such strong acid may also cause hydrolysis to some labile plant metabolites which was evident from the significantly low levels for RFOS and sucrose but high levels for glucose and galactose in both extracts obtained from these two acid solutions (Figure 2b). For this reason, we are not recommending these two solvent systems for plant tissue extraction unless under some special circumstances.

In contrast, the inter-sample chemical-shift variations were well below 0.5 Hz for all these representative metabolites in extracts (Figure 2a) from aqueous methanol and acetonitrile. The levels of RFOS, sucrose, glucose and galactose all showed no fluctuations indicating that both MeOH and ACN did not cause

changes of metabolites. This is consistent with the common knowledge that methanol and acetonitrile can precipitate proteins to quench enzyme activity. OPLS-DA results further showed that aqueous methanol extracts contained higher levels for almost all metabolites than aqueous acetonitrile extracts with exception of lipids (Figure 2c). The extraction ratios from those two solvents were not significantly different even though MeOH yields more extracts (Figure S2). Therefore, it is reasonable to conclude that aqueous methanol is probably the most suitable solvent (amongst these six tested) for plant metabolomic analysis with the solvent costs considered as well.

2.2 Assessment of cell-breaking methods

To assess the suitability of different cell-breaking methods, we employed aqueous methanol as solvent and the mortar-and-pestle ground powder of mungbean seeds as raw materials. ^{1}H NMR spectra (Figure S3) showed that metabolites obtained from all these cell-breaking methods were similar although metabolite ratios differed to some degree especially for signals of N1-methyl nicotinate and aglycone (around δ8.0), 4-hydroxybenzoate (δ7.6) and nucleotides (δ5.5-6.3). The results for extraction ratios (Figure 3a) clearly indicated that all three treatments (namely, tissuelyser homogenization, ultrasonication and the combined tissuelyser homogenization and ultrasonication) gave significantly higher extraction yields than direct extraction (without cell-breaking treatment). Ultrasonication was apparently more effective for metabolite extraction than tissuelyser homogenization when tough cell walls are present probably due to different cell-breaking mechanisms for these two methods. The extract yields were not statistically higher from the combination of tissuelyser homogenization and ultrasonication than from ultrasonication alone (Figure 3a). In any case, inter-sample chemical-shift standard deviations were below 0.75 Hz (Figure 3b) being acceptable for multivariate data analysis. OPLS-DA results (Figure S4) showed that extracts from ultrasonication had more amino acids than these from tissuelyser homogenization alone. This indicates that either ultrasonication or the combined approach is employable for extracting metabolites from plant tissues and the mortar-and-pestle grinding alone is not sufficient. To ensure good extraction rate, therefore, we concluded that the optimal cell-breaking method is the combination of tissuelyser homogenization with ultrasonication. It is worth mentioning that temperature control during ultrasonication step is vitally important and the intermittent interruptions are useful way to avoid temperature-fluctuations during both ultrasonication and tissuelyser homogenization.

2.3 Optimization for tissue-to-solvent ratio (TSR)

For tissue extraction, optimal tissue-to-solvent ratio (TSR) is also critically important especially when

taking into consideration of the saturation effects with high TSR. With TSR ranging from 1:5 to 1:30 (mg:μL), our results showed that extraction ratios (ER) from the first extraction were much higher than the subsequent three extractions (Figure S5). For the first three extractions, the extraction ratios were obviously dependent on TSR and the second extraction still yielded more than 10% of total extracts with the TSR as high as 1:30. This indicated that a single extraction was not enough to achieve sufficient metabolite extraction. Furthermore, no significant differences were observable for the total extraction ratios (from four sequential extractions) with TSR ranging from 1:5 to 1:15 (Figure S5); ER showed no differences for TSR of 1:20 and 1:30 either. Moreover, the extract yields from the forth extraction were much less than 10% of total yields (the sum of all extracts) and become independent on the TSR (Figure S5). In all cases, the extraction ratios from the combined first three extractions were not significantly different from these for the combined four extractions (Figure 4). This further suggested that three sequential extractions were necessary and sufficient to eliminate the TSR effects on metabolite composition. When TSRs were 1:5 and 1:7.5, the extraction ratios from the combined first two extractions were significantly lower than these from the combined first three extractions (Figure 4). In contrast, with TSR ranging from 1:10 to 1:30, the total extraction ratios resulting from the first two extractions were not significantly different from these obtained for the first three extractions (Figure 4). Therefore, the optimal TSR ought to be between 1:10 and 1:15 (mg:μL).

2.4 Optimal extract concentration for NMR analysis.

For NMR-based metabolomic analysis, concentration of extracts (i.e., extract-to-buffer ratio) is also an important parameter affecting the signal quality of metabolites such as signal-to-noise ratio (SNR) and inter-sample chemical shift variations due to inter-metabolite interactions. Our results showed that SNR increased steadily with the rise of extract concentration due to the increase of metabolite concentration (Figure S6). The inter-sample chemical-shift variations (standard deviations) also increased with the rise of extract concentration (Figure 5) for histidine although the other representative metabolites including these having aromatic rings showed little changes. When extract concentration was lower than 8.3 mg/mL, the standard deviation of chemical-shift for histidine signal was smaller than 1 Hz whereas such variations became great than 1.2 Hz when extract concentration was higher than 11.0 mg/mL. Such values were still acceptable for all other metabolites with the extract concentration of 2.3-16.7 mg/mL. However, when the extract concentration was below 5.2 mg/mL, the SNR will suffer an obvious loss. On balance, therefore, we suggest that the extract concentration be about 5-8 mg/mL.

2.5 Quantification of metabolites with NMR spectroscopy

With NMR methods, concentration of metabolites can normally be determined either by using the signal integrals from the fully relaxed spectra or from partially relaxed spectra taking T_1 values into consideration[4, 31, 69]. Our results (Table 1) showed that when the extract concentration was about 6.7 mg/mL, T_1 for all metabolites from mungbean seeds was shorter than 5 s with exception for formate. Sucrose and raffinose were the most abundant metabolites in this extract of mungbean seeds. Since the total repetition time for the fully relaxed spectra (about 25.3 s) was much longer than $5T_1$, our results for metabolite concentration ought to be reliable with the standard deviations estimated to be about 10-20% from ten biological replicates. However, quantity for some metabolites calculated from the partially relaxed spectra appeared to be significantly different from these from the fully relaxed spectra. This probably results from the accumulative errors from T_1 measurements (e.g., citrate) and signal overlapping (e.g., malate). Nevertheless, OPLS-DA results (Figure S7) indicated that the use of partially or completely relaxed data made little differences in the case of comparative metabolomics as long as the same parameters were employed for control and treated groups. For absolute metabolite quantification, the completely relaxed spectra were more suitable with the use of suitable spectral deconvolution methods.

3 Conclusions

Comprehensive assessments conducted in this work enable an optimized method to be established for plant seed metabolomic analysis. Such method starts with sample grinding with a mortar and pestle in liquid nitrogen followed with metabolite extraction and NMR analysis. The combined tissuelyser homogenization with ultrasonication is the ideal cell-breaking method for plant seed metabolomic analysis and aqueous methanol is the ideal extraction solvent for hydrophilic metabolites. The pooled extracts from three sequential extractions are optimal with the tissue-to-solvent ratio of about 1:10 to 1:15 (mg:μL) and procedures based on a single extraction is not recommended. For NMR analysis, the extract concentration of 5-8 mg/mL is suitable unless a large amount of secondary metabolites dominates plant extracts. Completely relaxed spectra are favorable for the NMR-based quantitative plant metabolomic analysis.

Acknowledgement

We acknowledge the financial supports from National Natural Science Foundation of China (21175149, 20825520 and 21221064). Mungbean seeds used in this study were obtained from Institute of Food Crops, Hubei Academy of Agriculture Science as a free gift.

Electronic supplementary information (ESI) available: Figure S1-S7

Figure Legends:

Figure 1. ^{1}H NMR spectra for mungbean seed extracts from aqueous methanol (A), aqueous acetonitrile (B), perchloric acid (C), trichloroacetic acid (D), hot buffer (E) and icy buffer (F). Spectral regions of δ 0.80-3.18, 5.50-8.80 and 8.80-9.60 were vertically expanded for 4, 64 and 32 times respectively. 1: leucine; 2: valine; 3 (isoleucine), 4 (ethanol), 5 (threonine), 6 (lactate), 7 (alanine), 8 (lysine), 9 (citrulline), 10 (acetate), 11 (glutamate), 12 (2-amino-4-oxopentanate), 13 (methionine), 14 (4-aminobutyrate), 15 (malate), 16 (succinate), 17 (2-ketoglutarate), 18 (citrate), 19 (aspartate), 20 (tyrosine), 21 (O-methyl *scyllo*-inositol), 22 (choline), 23 (phosphorylcholine), 24 (glycerophosphorylcholine), 25 (betaine), 26 (*scyllo*-inositol), 27 (inositol), 28 (tartarate), 29 (N1-methyl nicotinate) 30 (α-glucose), 31 (α-arabinose), 32 (α-galactose), 33 (sucrose), 34: raffinose family oligosaccharides (RFOS); 35 (uridine diphosphate-glucose), 36 (uridine diphosphate-mannose), 37 (uracil), 38 (uridine), 39 (uridine monophosphate), 40 (cytidine monophosphate), 41 (cytidine triphosphate), 42 (inosine monophosphate), 43 (inosine), 44 (*cis*-aconitic acid), 45 (fumarate), 46 (4-hydroxybenzoate), 47 (aglycone), 48 (histidine), 49 (tryptophan), 50 (phenylalanine), 51 (formate), 52 (lipid), 53: (β-galactose); 54 (β-glucose), U1-U7: unassigned metabolites.

Figure 2. Metabolite extraction properties of aqueous methanol (A), aqueous acetonitrile (B), perchloric acid (C), trichloroacetic acid (D), hot buffer (E) and icy buffer (F) for mungbean seeds. (a) inter-sample standard deviations (SD) for chemical shifts of seven metabolites in mungbean extracts resulting from six different solvents including lactate (□), succinate (＊), citrate(●), uridine (▲), histidine(□), tryptophan (○) and N1-methyl nicotinate(□); (b) absolute concentrations of some carbohydrates in mungbean seed extracts from six solvents (A-F) where these groups with different letters meant significant difference ($p<0.05$) and RFOS were raffinose family oligosaccharides; (c) OPLS-DA results for mungbean seed extracts from (A) aqueous methanol and (B)aqueous acetonitrile ($R^2X=0.68$, $Q^2=0.84$, $p=4.4e^{-5}$).

Figure 3. Effects of cell-breaking methods on metabolite extraction properties for mungbean seeds. A: no treatment; B: tissuelyser homogenization; C: ultrasonication; D: combined tissuelyser homogenization and ultrasonication. (a) extraction ratios (ER) from four cell-breaking methods; only these groups with different letter had significant differences ($p<0.05$). (b) inter-sample standard deviations (SD) for chemical shifts of lactate (□), succinate (＊), citrate(●), uridine (▲), histidine(□), tryptophan (○) and N1-methyl nicotinate(□), respectively.

Figure 4. Effects of tissue-to-solvent ratios and numbers of repeat extractions on the metabolite extraction ratios (ER) for mungbean seeds. Only these groups with different letters had significant differences

($p<0.05$). E1: the first extract; E12: pooled extracts from the first two extractions; E123: pooled extracts from the first three extractions and E1234: pooled extracts from the first four extractions.

Figure 5. Effects of mungbean seed extract concentrations on the inter-sample standard deviations (SD) for chemical shifts of lactate (□), succinate (*), citrate(●), uridine (▲), histidine(□), tryptophan (○) and N1-methyl nicotinate(□).

Reference

1. J. K. Nicholson, J. C. Lindon and E. Holmes, *Xenobiotica*, 1999, **29**, 1181-1189.
2. H. R. Tang and Y. L. Wang, *Prog. Biochem. Biophys.* , 2006, **33**, 401-417.
3. W. Weckwerth and O. Fiehn, *Curr. Opin. Biotechnol.*, 2002, **13**, 156-160.
4. H. Dai, C. N. Xiao, H. B. Liu, F. H. Hao and H. R. Tang, *J. Proteome Res.*, 2010, **9**, 1565-1578.
5. O. Fiehn, *Plant Mol. Biol.*, 2002, **48**, 155-171.
6. Y. P. An, W. X. Xu, H. H. Li, H. H. Lei, L. M. Zhang, F. H. Hao, Y. X. Duan, X. Yan, Y. Zhao, J. F. Wu, Y. L. Wang and H. R. Tang, *J. Proteome Res.*, 2013, **12**, 3755-3768.
7. M. E. Dumas, R. H. Barton, A. Toye, O. Cloarec, C. Blancher, A. Rothwell, J. Fearnside, R. Tatoud, V. Blanc, J. C. Lindon, S. C. Mitchell, E. Holmes, M. I. McCarthy, J. Scott, D. Gauguier and J. K. Nicholson, *Proc. Natl. Acad. Sci. USA*, 2006, **103**, 12511-12516.
8. F. C. Dong, L. L. Zhang, F. H. Hao, H. R. Tang and Y. L. Wang, *J. Proteome Res.*, 2013, **12**, 2958-2966.
9. W. X. Xu, J. F. Wu, Y. P. An, C. N. Xiao, F. H. Hao, H. B. Liu, Y. L. Wang and H. R. Tang, *J. Proteome Res.*, 2012, **11**, 3423-3435.
10. J. K. Nicholson, J. Connelly, J. C. Lindon and E. Holmes, *Nat. Rev. Drug Discov.*, 2002, **1**, 153-161.
11. L. M. Jiang, J. Huang, Y. L. Wang and H. R. Tang, *J. Proteome Res.*, 2012, **11**, 3848-3859.
12. Y. Semel, N. Schauer, U. Roessner, D. Zamir and A. R. Fernie, *Metabolomics*, 2007, **3**, 289-295.
13. D. G. Robertson, M. D. Reily, R. E. Sigler, D. F. Wells, D. A. Paterson and T. K. Braden, *Toxicol. Sci.*, 2000, **57**, 326-337.
14. J. T. Zhang, Y. Zhang, Y. Y. Du, S. Y. Chen and H. R. Tang, *J. Proteome Res.*, 2011, **10**, 1904-1914.
15. Y. F. Ye, L. M. Zhang, F. H. Hao, J. T. Zhang, Y. L. Wang and H. R. Tang, *J. Proteome Res.*, 2012, **11**, 2559-2566.
16. Y. F. Ye, L. M. Zhang, R. Yang, Q. J. Luo, H. M. Chen, X. J. Yan and H. R. Tang, *J. Agric. Food Chem.*, 2013, **61**, 8356-8363.
17. J. T. Brindle, H. Antti, E. Holmes, G. Tranter, J. K. Nicholson, H. W. L. Bethell, S. Clarke, P. M. Schofield, E. McKilligin, D. E. Mosedale and D. J. Grainger, *Nat. Med.*, 2002, **8**, 1439-1445.
18. J. R. Marchesi, E. Holmes, F. Khan, S. Kochhar, P. Scanlan, F. Shanahan, I. D. Wilson and Y. L. Wang, *J. Proteome Res.*, 2007, **6**, 546-551.
19. Y. H. Wang, L. M. Zhang, W. L. Chen, J. H. Wang, N. Li, J. M. Li, J. Q. Mi, W. N. Zhang, Y. Li, S. F. Wu, J. Jin, Y. G. Wang, H. Huang, Z. Chen, S. J. Chen and H. R. Tang, *J. Proteome Res.*, 2013, **12**, 4393-4401.
20. L. M. Zhang, Y. F. Ye, Y. P. An, Y. Tian, Y. L. Wang and H. R. Tang, *J. Proteome Res.*, 2010, **10**, 614-623.
21. W. B. Dunn and D. I. Ellis, *Trac-Trend Anal. Chem.*, 2005, **24**, 285-294.
22. H. B. Liu, A. M. Zheng, H. L. Liu, H. Y. Yu, X. Y. Wu, C. N. Xiao, H. Dai, F. H. Hao, L. M. Zhang, Y. L. Wang and H. R. Tang, *J. Agric. Food Chem.*, 2012, **60**, 129-135.
23. E. Holmes, R. L. Loo, J. Stamler, M. Bictash, I. K. S. Yap, Q. Chan, T. Ebbels, M. De Iorio, I. J. Brown, K. A. Veselkov, M. L. Daviglus, H. Kesteloot, H. Ueshima, L. C. Zhao, J. K. Nicholson and P. Elliott, *Nature*, 2008, **453**, 396-400.
24. M. S. Sabatine, E. Liu, D. A. Morrow, E. Heller, R. McCarroll, R. Wiegand, G. F. Berriz, F. P. Roth and R. E. Gerszten, *Circulation*, 2005, **112**, 3868-3875.
25. J. L. Griffin, M. Bollard, J. K. Nicholson and K. Bhakoo, *NMR Biomed.*, 2002, **15**, 375-384.
26. J. L. Griffin, J. C. M. Pole, J. K. Nicholson and P. L. Carmichael, *BBA-Gen. Subjects*, 2003, **1619**, 151-158.
27. Y. L. Wang, E. Holmes, E. M. Comelli, G. Fotopoulos, G. Dorta, H. Tang, M. J. Rantalainen, J. C. Lindon, I. E. Corthesy-Theulaz, L. B. Fay, S. Kochhar and J. K. Nicholson, *J. Proteome Res.*, 2007, **6**, 3944-3951.
28. O. Beckonert, H. C. Keun, T. M. D. Ebbels, J. G. Bundy, E. Holmes, J. C. Lindon and J. K. Nicholson, *Nat. Protoc.*, 2007, **2**, 2692-2703.
29. M. E. Bollard, S. Garrod, E. Holmes, J. C. Lindon, E. Humpfer, M. Spraul and J. K. Nicholson, *Magn. Reson. Med.*,

2000, **44**, 201-207.

30. A. Moing, M. Maucourt, C. Renaud, M. Gaudillère, R. Brouquisse, B. Lebouteiller, A. Gousset-Dupont, J. Vidal, D. Granot, B. Denoyes-Rothan, E. Lerceteau-Köhler and D. Rolin, *Funct. Plant Biol.*, 2004, **31**, 889-902.
31. H. Dai, C. N. Xiao, H. B. Liu and H. R. Tang, *J. Proteome Res.*, 2010, **9**, 1460-1475.
32. K. Gallardo, C. Le Signor, J. Vandekerckhove, R. D. Thompson and J. Burstin, *Plant Physiol.*, 2003, **133**, 664-682.
33. A. L. Van Eenennaam, K. Lincoln, T. P. Durrett, H. E. Valentin, C. K. Shewmaker, G. M. Thorne, J. Jiang, S. R. Baszis, C. K. Levering, E. D. Aasen, M. Hao, J. C. Stein, S. R. Norris and R. L. Last, *Plant Cell*, 2003, **15**, 3007-3019.
34. Z. B. Lu, G. J. Nie, P. S. Belton, H. R. Tang and B. L. Zhao, *Neurochem. Int.*, 2006, **48**, 263-274.
35. C. Chen, H. R. Tang, L. H. Sutcliffe and P. S. Belton, *J. Agric. Food Chem.*, 2000, **48**, 5710-5714.
36. H. R. Tang, A. D. Covington and R. A. Hancock, *Biopolymers*, 2003, **70**, 403-413.
37. O. Beckonert, M. Coen, H. C. Keun, Y. L. Wang, T. M. D. Ebbels, E. Holmes, J. C. Lindon and J. K. Nicholson, *Nat. Protoc.*, 2010, **5**, 1019-1032.
38. C. N. Xiao, F. H. Hao, X. R. Qin, Y. L. Wang and H. R. Tang, *Analyst*, 2009, **134**, 916-925.
39. M. Lauridsen, S. H. Hansen, J. W. Jaroszewski and C. Cornett, *Anal. Chem.*, 2007, **79**, 1181-1186.
40. J. F. Wu, Y. P. An, J. W. Yao, Y. L. Wang and H. R. Tang, *Analyst*, 2010, **135**, 1023-1030.
41. L. M. Jiang, J. Huang, Y. L. Wang and H. R. Tang, *Analyst*, 2012, **137**, 4209-4219.
42. A. Beneduci, G. Chidichimo, G. Dardo and G. Pontoni, *Anal. Chim. Acta*, 2011, **685**, 186-195.
43. Y. H. Choi, S. Sertic, H. K. Kim, E. G. Wilson, F. Michopoulos, A. W. M. Lefeber, C. Erkelens, S. D. P. Kricun and R. Verpoorte, *J. Agric. Food Chem.*, 2005, **53**, 1237-1245.
44. C. N. Xiao, H. Dai, H. B. Liu, Y. L. Wang and H. R. Tang, *J. Agric. Food Chem.*, 2008, **56**, 10142-10153.
45. N. J. Kruger, M. A. Troncoso-Ponce and R. G. Ratcliffe, *Nat. Protocols*, 2008, **3**, 1001-1012.
46. H. K. Kim, Y. H. Choi and R. Verpoorte, *Nat. Protoc.*, 2010, **5**, 536-549.
47. C. X. Liu, F. H. Hao, J. Hu, W. L. Zhang, L. L. Wan, L. L. Zhu, H. R. Tang and G. C. He, *J. Proteome Res.*, 2010, **9**, 6774-6785.
48. F. R. G. Terras, K. Eggermont, V. Kovaleva, N. V. Raikhel, R. W. Osborn, A. Kester, S. B. Rees, S. Torrekens, F. Vanleuven, J. Vanderleyden, B. P. A. Cammue and W. F. Broekaert, *Plant Cell*, 1995, **7**, 573-588.
49. O. Fiehn, J. Kopka, R. N. Trethewey and L. Willmitzer, *Anal. Chem.*, 2000, **72**, 3573-3580.
50. X. M. Wang, T. Sakuma, E. Asafu-Adjaye and G. K. Shiu, *Anal. Chem.*, 1999, **71**, 1579-1584.
51. Y. H. Choi, E. C. Tapias, H. K. Kim, A. W. M. Lefeber, C. Erkelens, J. T. J. Verhoeven, J. Brzin, J. Zel and R. Verpoorte, *Plant Physiol.*, 2004, **135**, 2398-2410.
52. J. G. Streeter and C. E. Strimbu, *Anal. Biochem.*, 1998, **259**, 253-257.
53. H. N. Johansen, V. Glitso and K. E. B. Knudsen, *J. Agric. Food Chem.*, 1996, **44**, 1470-1474.
54. T. W. M. Fan, *Prog. Nucl. Magn. Reson. Spectrosc.*, 1996, **28**, 161-219.
55. S. M. Bagnasco, S. Uchida, R. S. Balaban, P. F. Kador and M. B. Burg, *Proc. Natl. Acad. Sci. USA*, 1987, **84**, 1718-1720.
56. D. Cook, S. Fowler, O. Fiehn and M. F. Thomashow, *Proc. Natl. Acad. Sci. USA*, 2004, **101**, 15243-15248.
57. J. H. Liu, J. E. Burdette, H. Y. Xu, C. G. Gu, R. B. van Breemen, K. P. L. Bhat, N. Booth, A. I. Constantinou, J. M. Pezzuto, H. H. S. Fong, N. R. Farnsworth and J. L. Bolton, *J. Agric. Food Chem.*, 2001, **49**, 2472-2479.
58. S. Sagisaka, *Plant Physiol.*, 1976, **57**, 308-309.
59. R. H. Ellis, T. D. Hong and E. H. Roberts, *Ann. Bot.*, 1990, **66**, 341-348.
60. L. Eriksson, J. Trygg and S. Wold, *J. Chemometrics*, 2008, **22**, 594-600.
61. O. Cloarec, M. E. Dumas, J. Trygg, A. Craig, R. H. Barton, J. C. Lindon, J. K. Nicholson and E. Holmes, *Anal. Chem.*, 2005, **77**, 517-526.
62. T. W. M. Fan and A. N. Lane, *Prog. Nucl. Magn. Reson. Spectrosc.*, 2008, **52**, 69-117.

63. J. C. Lindon, J. K. Nicholson and J. R. Everett, *Annu. Rep. NMR Spectrosc.*, 1999, **38**, 1-88.
64. R. J. Staneloni, M. E. Tolmasky, C. Petriella and L. F. Leloir, *Plant Physiol.*, 1981, **68**, 1175-1179.
65. T. Li, Q. H. Liu, R. Ohsugi, T. Yamagishi and H. Sasaki, *Rice Science*, 2006, **13**, 205-210.
66. A. M. Lafta and J. H. Lorenzen, *Plant Physiol.*, 1995, **109**, 637-643.
67. T. Peterbauer, L. B. Lahuta, A. Blochl, J. Mucha, D. A. Jones, C. L. Hedley, R. J. Gorecki and A. Richter, *Plant Physiol.*, 2001, **127**, 1764-1772.
68. A. Nishizawa, Y. Yabuta and S. Shigeoka, *Plant Physiol.*, 2008, **147**, 1251-1263.
69. F. F. Chen, J. T. Zhang, X. S. Song, J. Yang, H. P. Li, H. R. Tang and Y. C. Liao, *J. Proteome Res.*, 2011, **10**, 2273-2285.

Table 1. Concentration and proton T_1 values for some metabolites in mungbean seeds

	δ (moieties)	T_1±SD (s)	concentration (mg/g) [a]	concentration (mg/g) [b]
leucine	0.96 (δ-CH_3)	0.79±0.09	0.16±0.02	0.14±0.02
valine	0.99 (γ-CH_3)	0.96±0.06	0.19±0.01	0.20±0.03
isoleucine	1.01 (γ-CH_3)	0.79±0.10	0.08±0.01	0.09±0.01
ethanol	1.18 (β-CH_3)	2.17±0.12	0.03±0.01	0.02±0.01
lactate	1.33 (β-CH_3)	1.25±0.03	0.54±0.04	0.58±0.07
alanine	1.48 (β-CH_3)	1.30±0.05	0.31±0.03	0.36±0.03*
acetate	1.92 (CH_3)	3.80±0.03	0.12±0.02	0.12±0.01
glutamate	2.08 (β-CH_2)	0.99±0.05	1.36±0.13	1.23±0.16
2-amino-4-oxopentanate	2.14 (CH_3)	1.92±0.02	2.99±0.31	2.99±0.36
methionine	2.17 (S-CH_3)	2.44±0.04	0.56±0.04	0.57±0.05
4-aminobutyrate	2.30 (α-CH_2)	1.15±0.04	0.10±0.01	0.09±0.01
succinate	2.41 (α/β-CH_2)	1.47±0.03	0.10±0.01	0.10±0.02
2-ketoglutarate	2.45 (γ-CH_2)	0.87±0.08	0.23±0.04	0.18±0.03*
citrate	2.55 (α/γ-CH_2)	0.41±0.01	8.81±0.65	9.54±0.77*
aspartate	2.80 (β-CH_2)	1.13±0.07	1.22±0.12	1.31±0.15
tyrosine	2.94 (CH_2)	0.98±0.12	1.06±0.10	1.17±0.10*
choline	3.20 (N-CH_3)	2.18±0.01	1.59±0.14	1.56±0.12
choline phosphate	3.22 (N-CH_3)	1.04±0.02	0.28±0.02	0.26±0.03
glycerophosphoryl choline	3.24 (N-CH_3)	0.90±0.05	0.10±0.02	0.10±0.02
betaine	3.27 (N-CH_3)	1.34±0.05	0.06±0.00	0.06±0.01
scyllo-inositol	3.35 (CH)	1.32±0.03	0.04±0.01	0.06±0.01*
O-methyl *scyllo*-inositol	3.61 (CH_3)	1.28±0.02	5.07±0.50	4.92±0.42
malate	4.31 (α-CH)	3.68±0.02	1.52±0.15	1.89±0.14*
tartarate	4.34 (α/β-CH)	4.43±0.02	0.34±0.03	0.42±0.04*
N1-methyl nicotinate	4.44 (CH_3)	2.29±0.05	1.36±0.13	1.39±0.15
α-glucose	5.23 (1-CH)	2.25±0.35	0.60±0.18	0.78±0.24
α-galactose	5.25 (1-CH)	#	0.10±0.04	#
sucrose	5.42 (Glc-1-CH)	1.02±0.03	17.91±1.48	18.60±1.79
RFOS	5.44 (Glc-1-CH)	0.98±0.05	31.01±2.47	32.04±2.27
uridine	5.91 (6-CH)	#	0.10±0.02	#
aglycone	7.00 (3-CH)	1.13±0.24	0.14±0.09	0.10±0.02
histidine	7.08 (4-CH)	2.45±0.15	0.25±0.03	0.33±0.03*
phenylalanine	7.43 (2/6-CH)	2.42±0.09	0.26±0.03	0.26±0.03
4-hydroxybenzoate	7.64 (2/6-CH)	#	0.04±0.01	#
tryptophan	7.72 (4-CH)	1.81±0.22	0.32±0.03	0.41±0.04*
IMP	8.27 (2-CH)	2.55±0.12	0.27±0.06	0.26±0.05
inosine	8.35 (8-CH)	1.86±0.15	0.15±0.02	0.15±0.03
formate	8.46 (CH)	6.10±0.14	0.02±0.00	0.02±0.00

T_1 values were measured from 6.7 extracts (mg/mL). # not determined; *: $p<0.05$; [a] values from fully relaxed spectra in the forms of mg metabolites per gram dry seeds; [b] values from the partially relaxed spectra by taking into consideration of T_1 in the forms of mg metabolites per gram dry seeds; RFOS: raffinose family oligosaccharides. IMP: inosine monophosphate.

Figure 1

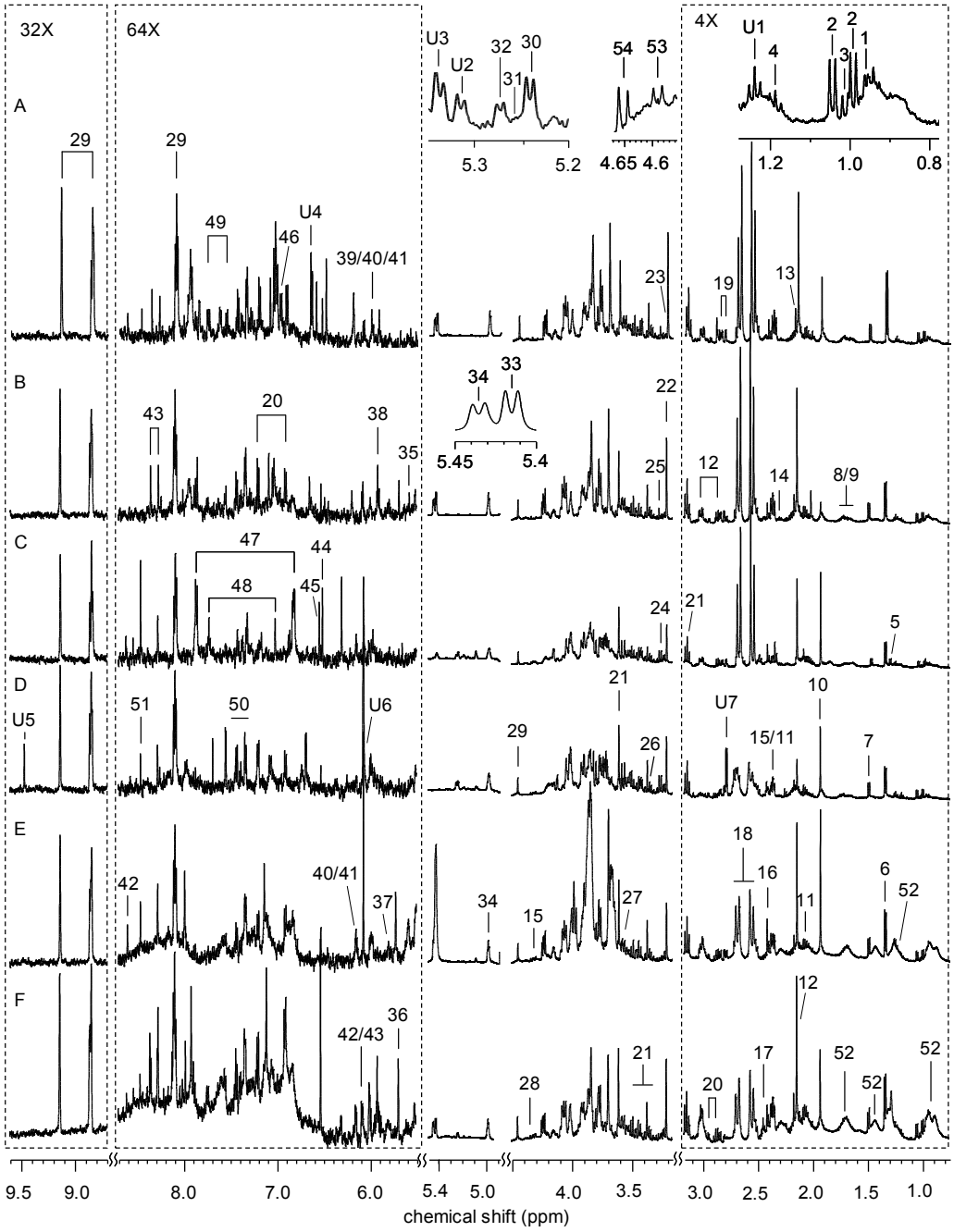


Figure 2

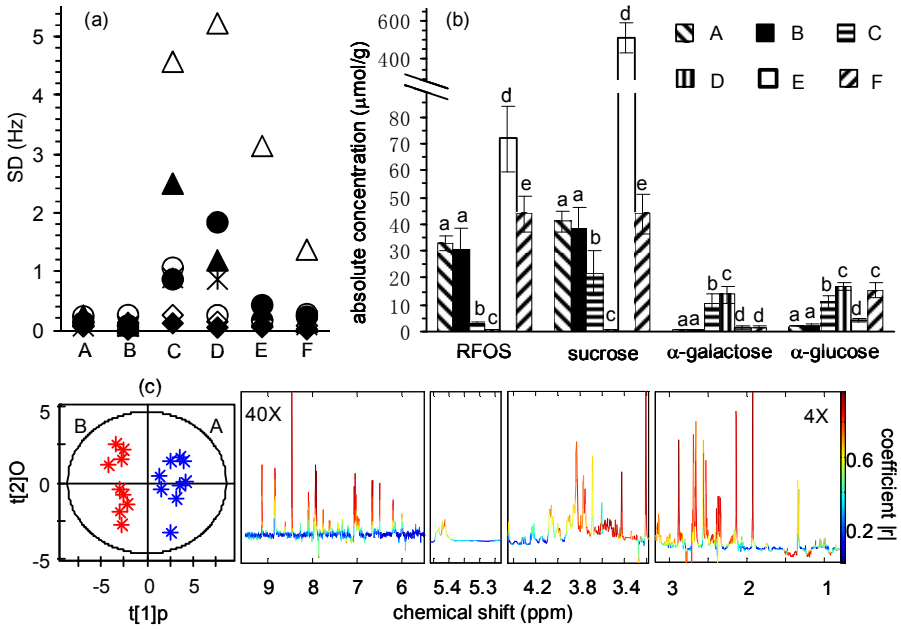


Figure 3

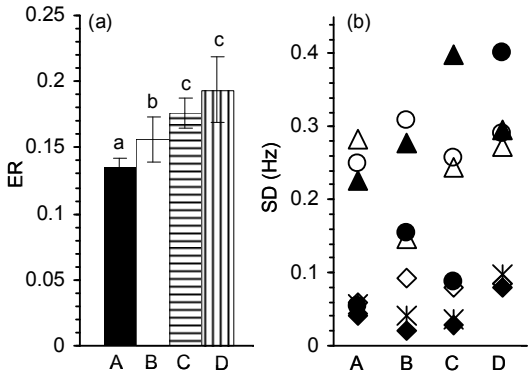


Figure 4

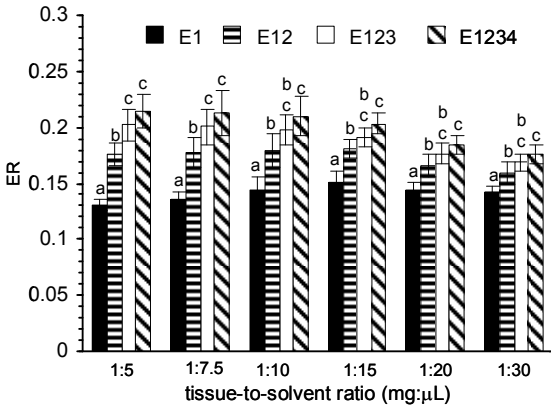
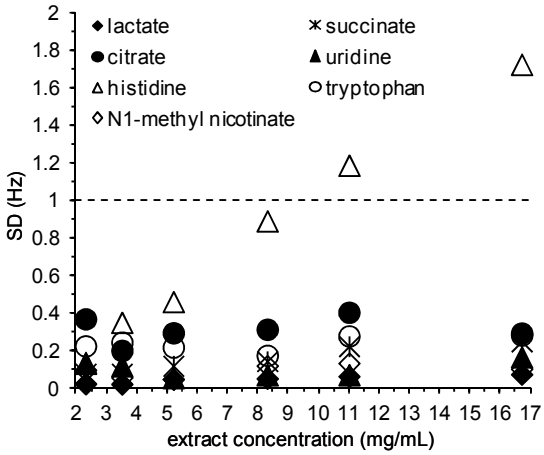


Figure 5



A table of contents entry

An optimized method for NMR-based plant seed metabolomic analysis is established with extraction solvent, cell-breaking method and extract-to-buffer ratio.