A Signal Filtering Method for Improved Quantification and Noise Discrimination in Fourier Transform Ion Cyclotron Resonance Mass Spectrometry-Based Metabolomics Data

Tristan G. Payne,[a] Andrew D. Southam,[b] Theodoros N. Arvanitis,[a] and Mark R. Viant[b]

[a] School of Electronic, Electrical and Computer Engineering, University of Birmingham, Birmingham, United Kingdom
[b] School of Biosciences, University of Birmingham, Birmingham, United Kingdom

Direct-infusion electrospray-ionization Fourier transform ion cyclotron resonance mass spectrometry (DI ESI FT-ICR MS) is increasingly being utilized in metabolomics, including the high sensitivity selected ion monitoring (SIM)-stitching approach. Accurate signal quantification and the discrimination of real signals from noise remain major challenges for this approach, with both adversely affected by factors including ion suppression during electrospray, ion–ion interactions in the detector cell, and thermally-induced white noise. This is particularly problematic for complex mixture analysis where hundreds of metabolites are present near the noise level. Here we address relative signal quantification and noise discrimination issues in SIM-stitched DI ESI FT-ICR MS-based metabolomics. Using liver tissue, we first optimized the number of scans (n) acquired per SIM window to address the balance between quantification accuracy versus acquisition time (and thus sample throughput); a minimum of $n = 5$ is recommended. Secondly, we characterized and computationally-corrected an effect whereby an ion's intensity is dependent upon its location within a SIM window, exhibiting a 3-fold higher intensity at the high m/z end. This resulted in significantly improved quantification accuracy. Finally, we thoroughly characterized a three-stage filter to discriminate noise from real signals, which comprised a signal-to-noise-ratio (SNR) hard threshold, then a "replicate" filter (retaining only peaks in r-out-of-3 replicate analyses), and then a "sample" filter (retaining only peaks in $>s$% of biological samples). We document the benefits of three-stage filtering versus one- and two-stage filters, and show the importance of selecting filter parameters that balance the confidence that a signal is real versus the total number of peaks detected. (J Am Soc Mass Spectrom 2009, 20, 1087–1095) © 2009 American Society for Mass Spectrometry

Metabolomics involves the measurement of multiple low molecular weight molecules from a biological sample to generate a unique metabolic fingerprint. The fingerprints from different phenotypes can then be compared to identify metabolic changes that provide biochemical rationale for the phenotypic differences. This approach has been successfully applied to a wide range of subjects, including disease diagnosis, nutrition and food science, environmental science, and toxicology [1–6]. Global metabolomic studies seek to identify and quantify as many metabolites as possible within the metabolome. Consequently, there has been a rapid increase in the number of direct infusion electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry (DI ESI FT-ICR MS) metabolomics studies reported, since this technique benefits significantly from high mass accuracy and resolution, as well as relatively high sensitivity [7–13]. However, the downside of this approach is that the accuracy and precision of metabolite quantification is lower than for other metabolomics approaches, such as nuclear magnetic resonance (NMR) spectroscopy [14], partly due to the noise introduced in the MS measurement. The challenge is how to then reliably filter this noise from the mass spectrum to maximize the number of real metabolites detected by DI ESI FT-ICR MS, while retaining the high-throughput characteristics of this approach compared with chromatographic based MS methods. In this paper, we consider the issues of relative quantification of mass peaks as well as noise discrimination.

Sources of noise in ESI FT-ICR MS include ion–ion interactions inside the ICR detector cell [15], white noise caused by thermal effects [16], ion suppression and

Address reprint requests to Dr. M. R. Viant, School of Biosciences, University of Birmingham, Birmingham, B15 2TT, UK. E-mail: M.Viant@bham.ac.uk

1044-0305/09/$32.00
doi:10.1016/j.jasms.2009.02.001

Published online February 7, 2009
Received December 24, 2008
Revised January 30, 2009
Accepted February 3, 2009

enhancement during ESI [17], incomplete desolvation of the sample [17], and an event whereby the electrospray is temporarily lost due to particulates blocking the spray nozzle (a "drop-out"). When "drop-out" occurs, the total ion current (TIC) of the signal drops significantly (<50% of typical values), and the resulting spectrum shows elevated noise levels and is not representative of the sample composition [18]. Some of this noise can be reduced before the measurement, for example by optimization of the ICR detector cell parameters [15, 19], or after the measurement, during the preprocessing stages [20]. The overall signal-to-noise ratio (SNR) in a mass spectrum can be improved by acquiring more scans [21]: the magnitude of additive white Gaussian noise (AWGN) increases as $n^{1/2}$ (where n is the number of scans acquired) and the signal increases as n, and therefore the overall SNR is in principle proportional to $n^{1/2}$ [22]. In reality, the actual SNR realized is limited by other non-white noise effects, including instrument parameter drift and changes in the sample composition, which can contribute to an increased variability of measured ion intensities over time. Furthermore, increasing the scan time significantly is undesirable for high-throughput methods such as DI ESI FT-ICR MS. Therefore, as part of our optimization of signal quantification, here we measure the actual benefit realized as the number of scans is increased when analyzing a liver tissue extract.

Despite optimizing the signal acquisition stage, ultimately a mass spectrum will contain both signal and noise. Profiling the complete metabolome will necessarily require the detection of low intensity metabolites, particularly as it has been shown that (at least for NMR studies) the number of unique metabolites is inversely proportional to their abundance [18]. The crucial step is in discarding noise while retaining and then measuring real peaks. For a single spectrum, this is typically achieved by setting an arbitrary peak area threshold [23] or a SNR threshold, and retaining only peaks that exceed this specification. SNR thresholds used include 3:1 [24–28] (the limit of detection), 5:1 [29], 10:1 [30] (the limit of quantification), and higher. The calculation of SNR is sometimes unclear and inconsistent; here we use the definition shown in eq 1.

$$\mathrm{SNR} = \frac{\text{height of signal peak in magnitude spectrum}}{\begin{array}{l}\text{standard deviation of the white Gaussian}\\ \text{noise in a signal-free region of the real}\\ \text{and imaginary components of the}\\ \text{complex spectrum (assumed equal)}\end{array}} \quad (1)$$

However, a hard threshold technique such as this either results in many low abundance ions not being detected, or many noise peaks being falsely counted as real. In addition, it can incorrectly remove ions with abundances that are intermittently reduced by adverse effects during ESI. This potential loss of data has been discussed recently by Wilson et al. [31], who develop an algorithm using liquid chromatography mass spectrometry (LC-MS) data and propose first applying a peak intensity threshold across all mass spectra to generate one set of "real" peaks. This matrix of peak intensities will contain multiple missing values (i.e., zero intensities) since peaks present in some spectra will not be detected in others, as they are below the detection threshold. This threshold is then removed, and all the missing values are filled in by reintegrating the areas around every "real" peak. The problem of reliably detecting low intensity ions remains, however, since a hard threshold must initially be used. Other means of filtering noise peaks in mass spectra include by peak variance [32] or resolution [33, 34], and filtering by peak shape is commonly used in LC-MS [35, 30]. While using additional information such as peak resolution, shape, or variance will certainly provide further evidence for the existence of a "real" peak, we have observed that in practice, noise does not appear sufficiently distinct from low intensity signal peaks to differentiate solely on peak parameters or variance.

It is worth noting that ion capture and detection in FT-ICR MS is a discrete process due to the manner in which ions are collected and then "held" by the magnetic field during the detection period in the ICR detector cell. After detection, the ions are released and a newly collected ion package is transferred to the ICR detector cell, and the process repeats. It is therefore appropriate to use a "replicate filter", where a single biological sample is measured multiple times and only those peaks common to a minimum number of the replicates are retained; and/or a "sample filter", where only those peaks common to a minimum number of biological samples within a group (e.g., a treatment group) are kept in the dataset. To date, these filters have been used empirically, and there is no consensus on the optimal number of replicates/samples or the filtering parameters. An example replicate filter applied to DI FT-ICR MS data stipulates that each peak be present in at least two out of three replicate analyses [11], and Quick et al. first applied this filter to DI ESI time-of-flight MS data with the requirement that each peak be present in all 3 replicate analyses [21, 32]. Example sample filters include requiring a peak to occur in 50% [30], 60% to 75% [11], or 80% [23] of the total number of samples in the group. Determination of the optimal filter settings for the analysis of complex mixtures, based upon a theoretical framework, is required.

Previously, we have demonstrated a significant increase in detection sensitivity in DI ESI FT-ICR MS based metabolomics by recording each wide-scan mass spectrum as a series of overlapping selected ion monitoring (SIM) windows [10], an approach that is now being adopted by other laboratories [36]. During further experimentation using this approach, we noticed a systematic intensity error whereby the observed intensity of an ion is dependent upon the location of that ion within the SIM window. In this paper, we first optimize the number of scans acquired for each SIM window, by

DI ESI FT-ICR MS of liver tissue extracts, to address the balance between quantification accuracy and acquisition time (and thus sample throughput). Then we characterize and computationally correct the systematic intensity error in the SIM-stitching approach [10]. Finally, we apply the SIM-stitching DI ESI FT-ICR MS method [10] to both real and simulated data to assess the performance of a three-stage filtering method. This method requires multiple samples to be acquired in triplicate, and applies a SNR hard threshold followed by a replicate filter and then sample filter, to produce a filtered peak list that contains the intensity of each peak in each sample (see Figure 1). We compare the results with one- and two-stage filtering methods, thus showing how an optimal filtering method can be objectively selected.

Experimental Design

Preparation of Liver Extracts

Frozen livers from a wild-caught species of marine flatfish, dab (*Limanda limanda*), were provided by the Centre for Environment, Fisheries, and Aquaculture Science (Cefas, Weymouth, UK). Tissues were kept at −80 °C until extraction. Each liver was homogenized in 8 μL/mg (wet tissue mass) methanol and 2.5 μL/mg water using a Precellys-24 ceramic bead-based homogenizer (Stretton Scientific Ltd., UK) [37]. Next, 8 μL/mg chloroform (pesticide analysis grade, Fisher Scientific, Loughborough, UK) and 4 μL/mg water were added, the biphasic mixture was centrifuged (1500 *g*), and the upper (polar) and lower (nonpolar) phases were isolated and frozen at −80 °C. Before MS analysis, polar extracts were dried using a centrifugal concentrator (Thermo Savant, Holbrook, NY), resuspended in 3 times the volume of original solvent (80:20 methanol/water with 0.25% formic acid), and then centrifuged (15,000 *g*). The nonpolar extracts were not used in this study.

FT-ICR Mass Spectrometry

Samples were analyzed using a hybrid 7-T Fourier transform ion cyclotron resonance mass spectrometer (LTQ FT; Thermo Electron Corp., Bremen, Germany) equipped with a chip-based direct infusion nanoelectrospray ionization source (Triversa; Advion Biosciences, Ithaca, NY). Nanoelectrospray conditions comprised a 200 nL/min flow rate, 0.3 psi backing pressure, and 1.7 kV electrospray voltage, controlled by ChipSoft software (version 8.1.0, Advion Biosciences). A conventional SIM scan mode was employed. Mass resolution was fixed at 100,000 (defined for an ion at m/z 400) throughout. The automatic gain control (AGC, corresponding to the number of charges transferred from the front-stage linear ion trap to the ICR detector cell) target was 1×10^5, and the maximum ion trap fill time was set to 1 s throughout. Data were obtained as transient files (i.e., scans recorded in the time domain) and processed using custom-written MATLAB (version 7; The MathWorks, Natick, MA) scripts as described in [10]. Briefly, SIM scans were first averaged and Fourier-transformed. Data prepared for the Noise Filtering section of this work was additionally internally calibrated using identified metabolites (see Supplementary Content, Table S1, which can be found in the electronic version of this article), and then stitched together to generate one mass spectrum and associated peak list per replicate analysis.

Optimizing Number of Scans

Consider a single scan in which a peak's intensity is measured as a true signal component μ, plus a noise component normally distributed with zero mean and standard deviation σ; the peak intensity will therefore have a normal (μ, σ^2) distribution. By averaging n scans, each with true signal intensity μ and a similarly distributed noise component, the mean peak intensity has a normal $(\mu, \sigma^2/n)$ distribution [38], i.e., the standard deviation (SD) is $\sigma/(n^{1/2})$, and hence the relative (to μ) standard deviation (RSD) is reduced by a factor of $n^{1/2}$. However, as discussed previously, this benefit is limited by other sources of noise. Since a series of replicate spectra are composed of many peaks (each peak with a unique RSD), a median RSD value can be derived for the dataset, which has recently been proposed as a valuable benchmark for spectral variability in a metabolomics experiment [39]. Therefore, to determine the optimal number of scans, n, that should be acquired for each spectrum, a sample was analyzed for at least 95 scans with the instrument configured in SIM scan mode and using three separate m/z range settings: 110–140, 230–260, and 470–500 m/z. For each m/z range setting, the first five scans were discarded, then three sets of n

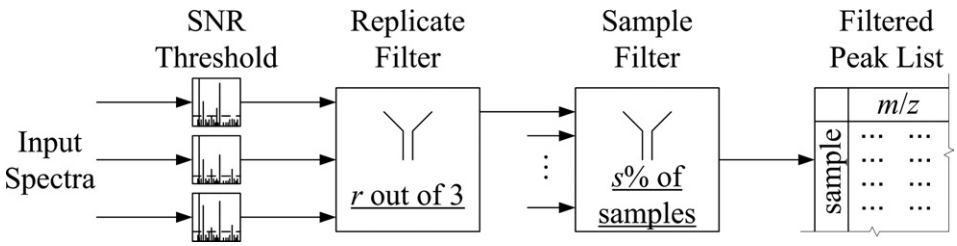


Figure 1. The three stage filtering schema.

(n from 3 to 30) consecutive scans were averaged, e.g., when $n = 4$, scans 1–4, 5–8, and 9–12 were averaged to produce three mean spectra. The location of peaks common to all three spectra were then determined, and the RSD values, calculated as the standard deviation relative to the mean intensity of these peaks, were evaluated. The median RSD value was then taken as a representative measure of spectral reproducibility [39], with a single median RSD value for each unique number of scans, n.

SIM Window Intensity Correction

On-going use of the original SIM-stitching DI ESI FT-ICR MS method for metabolomics [10] has highlighted a systematic error in the intensity measurements, whereby the intensity of a peak is dependent upon its location within the SIM scan range being measured. Since SIM-stitching involves the use of several SIM windows with different m/z ranges, this error could be compounded, resulting in additional errors in quantification. To characterize the error, with the goal to correct this intensity bias, a liver extract was analyzed over a "sliding" SIM window of fixed width 30 m/z and starting location from 50 to 430 m/z in 2 m/z increments, i.e., a total of 191 staggered SIM windows were recorded. High intensity peaks were selected in the range 80–430 m/z that were consistently present and did not neighbor other high intensity peaks, to reduce mutual space-charge effects on the measured peak intensities. As the SIM window is moved, the intensities of the peaks of interest are tracked.

Noise Filtering

Given a spectrum containing signal and noise, the aim of any noise filter is to remove noise while retaining signal. Therefore, to assess the performance of a noise filter it is necessary to know which peaks are due to real signal, and which are due to noise. However, given a spectrum acquired from a complex tissue extract, one cannot easily classify low intensity peaks as signal or noise. To address this problem, we have developed a statistical model as described in the Supplementary Content that allows us to realistically simulate mass spectra, including signal and noise. The simulated FT-ICR mass spectra are generated in the time domain with signal and noise features distributed according to observations from measured data, and provide a means of quantifying the performance of our three-stage noise filtering. The simulated spectra include random noise, both as additive white noise induced by thermal effects, and as random fluctuations in the intensities of "real" signal peaks across replicate spectra caused by variations in the ESI process and ion detection.

Multiple simulated spectra are created, representing replicate spectra of multiple biological samples (e.g., of the same phenotype and treatment group). These are then used as the input to three different configurations of the noise filter:

1. A SNR threshold filter applied to a single replicate spectrum, such that only peaks with a SNR above a threshold are considered real.
2. A SNR threshold filter (as above) followed by a "replicate" filter (the replicate filter can be configured such that a peak must be present in at least two out of three replicates, or a peak must be present in all three replicates to be retained as real).
3. A SNR threshold filter (as above) followed by a "replicate" filter (peak must be present in at least r out of three replicates), and a "sample" filter (peak must be present in at least s% of the number of biological samples, see Figure 1).

By comparing the output (filtered peak list) with the signals used to generate the simulated spectra, the performance of each filter can be quantified.

Results and Discussion

In this section, optimization of quantification in SIM-stitching FT-ICR mass spectra is discussed in terms of the three areas mentioned above: number of scans and correcting for intensity errors across SIM windows (the pre-processing stages), followed by post-acquisition noise filtering. The noise filtering is tested using simulated spectra, and so the model used to create these spectra is also discussed.

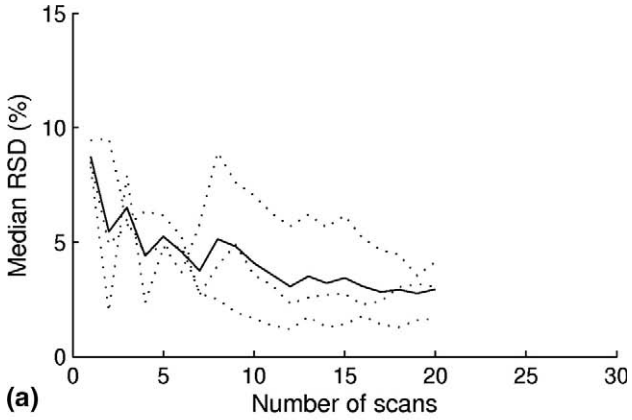
Optimizing Number of Scans

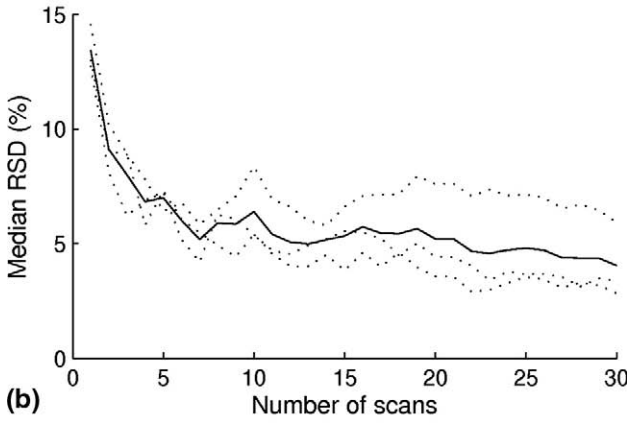
The results in Figure 2 show a rapidly decreasing median RSD with increasing n for all mass ranges. The experiment was repeated three times, which reveals considerable variability between repeats, and this could be caused by slow fluctuations in the electrospray process. The highest decrease in median RSD appears to be up to around $n = 5$ scans, which shows an approximately 2-fold improvement in median RSD of ~5% compared with $n = 1$ scan, and so $n = 5$ is used for the rest of the modeling presented here. Beyond $n = 5$, the median RSD tends to decrease only marginally, potentially due to longer-term variations, which counteract the anticipated gain due to signal averaging. The technical variation associated with the DI ESI FT-ICR measurement of a liver extract (~5%, Figure 2) is comparable with the technical variation observed in NMR studies [39].

SIM Window Intensity Correction

The intensity of several closely-spaced peaks, relative to their overall mean intensity, is shown to vary (Figure 3) as the SIM window location defining the m/z region being scanned is varied. The figure clearly shows that the measured signal intensity increases as the peak

moves across the SIM window. This instrumental measurement error is large, with a measured peak area at the high m/z end of a SIM window ~3 times more intense than if the identical peak is measured at the low m/z end of an adjacent SIM window. To quantify this effect, the gradient of the trend line in Figure 3 was measured for several different peaks over the range 70–500 m/z, and the results are plotted in Figure 4. This





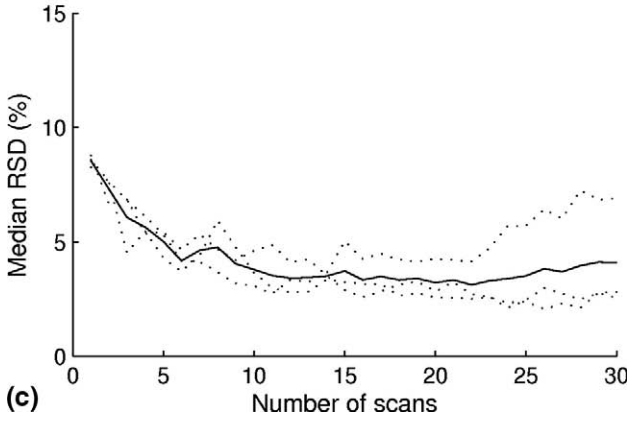


Figure 2. Median RSD of three spectra with mass ranges 110–140 m/z (**a**), 230–260 m/z (**b**), and 470–500 m/z (**c**) as the number of scans averaged to form each spectrum, n, is increased. The experiment was repeated three times each, shown as a dotted line, with the average as a solid line. The maximum number of scans in the range 110–140 m/z is limited by the second experiment repeat, which produced an insufficient number of total scans (66 obtained, 90 required).

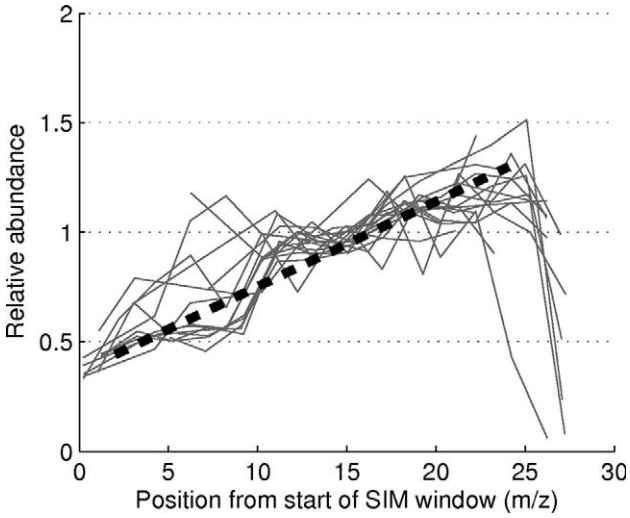


Figure 3. The intensities of several peaks in the range 352–354 m/z relative to their mean values as measured across multiple SIM window m/z ranges, showing considerable variation in peak intensity as the SIM window position moves across the peaks. A trend line is plotted (broken line).

reveals a trend (albeit noisy) in the gradient, which is parameterized as a linear function, and consequently allows the intensity of each peak in the mass spectrum to be corrected based on its location from the start of the SIM window. Figure S4 in the Supplementary Content demonstrates the improvement in peak intensities as a result of this correction. Also evident in Figure 3 is the previously reported "edge effect," which is apparent as a significant decrease in peak intensity (or disappearance of the peak altogether) as peaks near the extremities, particularly at the high m/z end of the SIM window. A more complete experimental characterization of these edge effects, which are easily dealt with

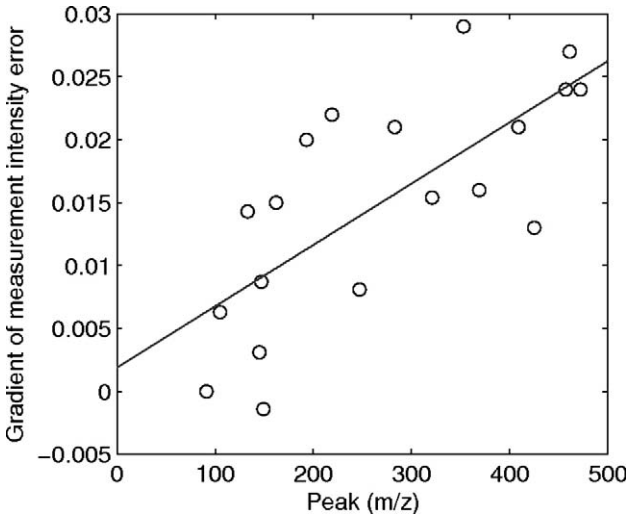


Figure 4. The gradient of the measurement intensity error for a selection of peaks located between 70 and 500 m/z. The line of best-fit is shown.

during spectral processing, was discussed by Southam et al. [10].

Simulated Spectra Model

The spectrum model described in the Supplementary Content generates a mass spectrum with realistic peak intensities and frequencies, and includes signal peaks derived from measured spectra, and random noise. By generating spectra using this model, the performance of the filtering strategy can be assessed. We created a set of 30 random spectra as described above, labeled *1a*, *1b*, *1c*, *2a*, *2b*, *2c*, . . ., *10a*, *10b*, *10c*, where (*a*, *b*, *c*) indicate the replicate spectra and (*1* . . . *10*) the repeat experiments. After applying a certain filtering configuration, the resulting peak list is compared with the "true" signal frequencies, *f*. Those peaks common to *f* to within 2.5 ppm (to allow for error in the frequency estimation caused by noise-induced peak shape distortion while avoiding detection of neighboring peaks) are considered correctly identified, while the remainder are considered noise. As stated previously, one measure of filter performance is the number of real peaks identified. We add to this the empirical probability that a detected peak is real, defined as shown in eq 2.

$$\text{P(detected and real)} = \frac{\text{number of detected peaks which are real}}{\text{number of detected peaks}} \quad (2)$$

SNR Threshold Noise Filtering

This strategy attempts to reduce noise from the spectrum by applying a hard threshold to the SNR of each peak such that only peaks with a SNR above a threshold are retained. For this experiment, we use only the first replicate of each sample (*1a*, . . ., *10a*), a total of 10 spectra. Each spectrum is filtered as described above. The results for typical SNR thresholds of 3.0, 5.0, 8.0 and 10.0 for 10 repeat experiments are shown in Table 1. It is apparent that one can either detect many real peaks (at low SNR threshold) or have high confidence that a detected peak is real (at high SNR threshold), but not both. Since it is necessary to achieve both of these in metabolomics studies, SNR threshold noise filtering alone is not recommended.

Table 1. Metrics for simulated spectra, filtered by SNR threshold only: mean number of real peaks, mean number of noise peaks, and mean probability that a detected peak is real, based on 10 repeats

SNR threshold	N_{real}	N_{noise}	$P_{d,r}$
3.0	3427	10987	0.238
5.0	2159	50	0.977
8.0	1298	14	0.989
10.0	1044	8	0.992

Table 2. Metrics for simulated spectra, filtered by SNR threshold and replicate filtered: mean number of real peaks, mean number of noise peaks, and mean probability that a detected peak is real, based on 10 repeats

SNR threshold	Replicate filter	N_{real}	N_{noise}	$P_{d,r}$
3.0	2 out of 3	3413	807	0.809
3.0	3 out of 3	2332	29	0.988
5.0	2 out of 3	2023	29	0.986
5.0	3 out of 3	1573	10	0.994

SNR Threshold and Replicate Noise Filtering

In this strategy, first SNR threshold filtering is applied as described above, then all remaining peaks are replicate-filtered with $r = 2$ (such that peaks must be present in at least two out of three replicate spectra), and also $r = 3$ (peaks must be present in all three spectra). The experiment is repeated over all 10 samples. The results for typical SNR threshold and replicate filtering with $r = 2$ and also $r = 3$ are shown in Table 2. Since the replicate filter removes a large amount of the noise, it is now feasible to use a low SNR threshold, e.g., with SNR = 3.0, 80.9% of the peaks detected by the two-stage filter are real, compared to only 23.8% using a SNR threshold alone, with both methods detecting just over 3400 real peaks. However, there are still many noise peaks also being detected even when a two out of three replicate filter is applied. This is because the low SNR threshold still allows many noise peaks to pass (Table 1), and consequently even with replicate filtering many of these remain. While applying a three out of three replicate filter ($r = 3$) mitigates these effects, many real peaks are also lost, and therefore this is not a good option. The third filtering method aims to address this shortfall.

Three-Stage Noise Filtering

Here, we present results using the three-stage filter (Figure 1). It is our experience that occasionally peak intensities drop below the noise level, apparently "disappearing", with no indication in the TIC. Therefore, we selected a two out of three replicate filters as optimal, which circumvents this potential loss of real peaks. The two-stage filter is applied first, with a varying SNR threshold and $r = 2$, resulting in 10 peak lists. The final stage acts across these samples to retain peaks present in at least *s*% of them, with *s* set to 10, 20, . . ., 100%. The experiment uses all of the available simulated spectra (*1a*, *1b*, *1c*, . . ., *10a*, *10b*, *10c*) where (*a*, *b*, *c*) indicate the replicate spectra and (*1* . . . *10*) the samples. An important assumption here is that the variation observed between replicate samples also captures the variation between samples. While this is not expected to be true in a real experiment since the concentration of metabolites will vary between biological samples, no methods have been reported to model

the general variation between samples, and so to maintain generality we recognize the approximation that, as is often the case, samples within one biological group will all be similar (e.g., the same phenotype) and accept that variations in metabolite populations will not be extreme (i.e., all metabolites will be present across all samples, albeit in varying concentrations).

The results of the experiment (with selected values of the SNR threshold, r and s, shown in Table 3) reveal that no single solution can definitively be described as "optimal". In general, the compromise available is between high likelihood that a detected peak is real (achievable with high SNR threshold and/or high r and s) and high number of correctly detected peaks but also incorrectly identified noise (achievable with low SNR threshold and/or low r and s).

It should be noted that approximations and assumptions are applied to the model and so emphasis should be placed on the *relative* performance of different filter configurations instead of absolute values such as number of peaks detected and likelihood that a detected peak is real. To this extent, we compared several configurations of the filter, which demonstrate the trade-off between these two metrics (Table 3). It is apparent that selecting the optimal configuration may depend upon the data being filtered. If the samples are likely to be very similar, then a higher s is acceptable since it is anticipated that a high number of peaks will be common across samples. In this case, a lower SNR threshold (2.6–3.0) may be used, yielding ~3800–3900 real peaks with a likelihood of 0.940–0.986 that a detected peak is real. If more variability is observed between samples, reducing s to 10% requires the SNR threshold to be raised to 4.0 to yield a comparable number of real peaks (3825) and probability of a detected peak being real (0.951).

To help select parameters, all results with $r = 2$ are presented as contour plots (Figure 5). The first plot shows that the number of real peaks detected is optimal with low SNR threshold and low sample filter, s (lighter regions). The second plot shows the probability that a detected peak is real, and is optimal with high SNR and high s. Thus, by considering a third metric, the product of these first two metrics, the overall relative performance of the filter can be visualized (third plot). Using this plot, we have identified optimal parameter sets that encompass the lightest region. These are listed in Table 3, and comprise of SNR = 2.6, $r = 2$, s = 40%–60%; SNR = 2.8, $r = 2$, s = 3%–50%; SNR = 3.0, $r = 2$, s = 30%–50%; SNR = 3.5, $r = 2$, s = 20%–30% and SNR = 4.0, $r = 2$, s = 10%. The optimal setting of SNR = 5.0 is

Table 3. Metrics for simulated spectra and three stage filtering: number of real peaks, number of noise peaks, and probability that a detected peak is real, based on a single repeat

SNR threshold	Replicate filter	Sample filter	N_{real}	N_{noise}	$P_{d,r}$
2.6	2 out of 3	40%	3913	150	0.963
2.6	2 out of 3	60%	3646	50	0.986
2.8	2 out of 3	30%	3956	253	0.940
2.8	2 out of 3	50%	3753	62	0.984
3.0	2 out of 3	30%	3929	267	0.959
3.0	2 out of 3	50%	3659	56	0.985
3.5	2 out of 3	20%	3882	156	0.961
3.5	2 out of 3	30%	3705	96	0.975
4.0	2 out of 3	10%	3825	196	0.951
5.0	2 out of 3	10%	2857	84	0.971

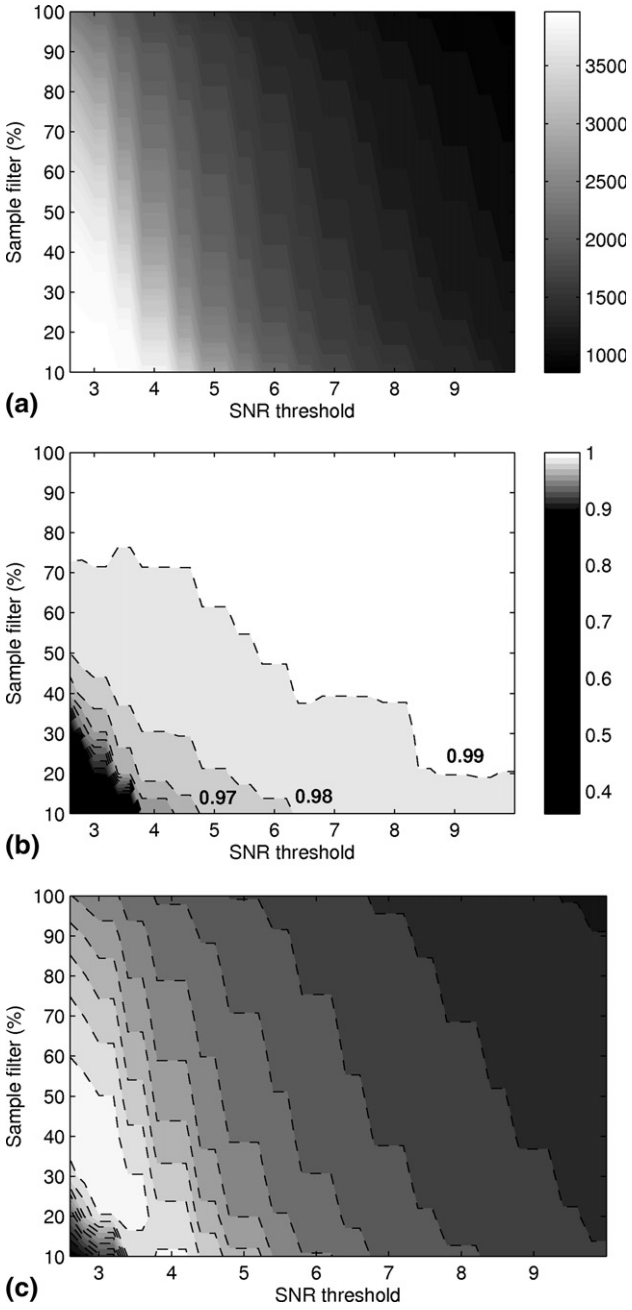


Figure 5. Contour plots showing the quality of the simulated spectrum, after three-stage filtering has been applied with varying SNR threshold and sample filter (s%), and replicate filtering fixed at $r = 2$ (out of 3), in terms of: (**a**) number of real peaks correctly identified, (**b**) probability that a detected peak was real (with reference to the "real" peaks in the model), and (**c**) the product of these two metrics.

also shown in Table 3 for comparison with one- and two-stage filtering (Tables 1 and 2). Our laboratory has used three-stage filtering with settings that favor higher probability of a peak being real: SNR = 3.5, r = 2 and s = 50%, as we recently reported for an FT-ICR MS-based metabolomics study of an aquatic invertebrate [12]. The effect of the filtering stages in this study are reflected by these typical total numbers of positive and negative ion peaks at each stage: 13,979 peaks after the SNR threshold filter, 8532 peaks after replicate filtering, and 5447 peaks after sample filtering. Of these 5447 peaks, more than 1000 metabolites were putatively identified [12]. Overall, this three-stage filter shows significant improvement in either the number of real peaks detected or the probability that a detected peak is real, or both.

Conclusions

This study highlights two important issues that should be addressed when considering DI ESI FT-ICR MS for quantitative metabolomics. The first is that the number of scans acquired for each spectrum greatly affects the reproducibility of the resulting spectrum, and we report at least five scans are required for a median RSD of ~5%. The second consideration concerns potential instrument-induced anomalies, and here we report a significant effect on peak intensity, which is dependent upon the location of the peak in a SIM window. Having addressed these issues, we describe a three-stage noise filtering method and show significant benefit on the quality of the data compared with single- and two-stage filtering methods. We describe how the parameters of the filter should be chosen carefully to maximize the metrics of interest, and report optimal values for these parameters. While the results presented here may vary as a result of the features of the particular empirical data, experiments from a different type of biological sample (spectra of human cell line extracts, with >2-fold increase in peak density) have yielded similar results (data not shown). When devising a data filtering strategy in FT-ICR MS experiments, the parameters should be selected using a scheme such as this as a guide. Such an approach, while applied here to DI ESI FT-ICR MS data, is likely to be relevant to any high-resolution MS data where there are significant numbers of low intensity peaks that would be lost if a simple hard threshold was applied.

Acknowledgments

T.G.P. and T.N.A. thank the EPSRC for a Ph.D. Studentship and DTA support, A.D.S. thanks the BBSRC and Cefas for a Ph.D. CASE Studentship (under Cefas Seedcorn contract DP195), and M.R.V. thanks the NERC for an Advanced Fellowship (NER/J/S/2002/00,618). This work was partly supported by NERC grant NE/D002508/1. The authors also thank Drs. Grant Stentiford, Brett Lyons, and Steve Feist (Cefas) for the fish liver samples, Nadine Taylor (Birmingham) for help preparing samples and collecting data, and Dr. Helen Cooper (Birmingham) for advice and comments on the manuscript.

Appendix A Supplementary Material

Supplementary material associated with this article may be found in the online version at doi:10.1016/j.jasms.2009.02.001.

References

1. Plumb, R. S.; Stumpf, C. L.; Granger, J. H.; Castro-Perez, J.; Haselden, J. N.; Dear, G. J. Use of Liquid Chromatography/Time-of-Flight Mass Spectrometry and Multivariate Statistical Analysis Shows Promise for the Detection of Drug Metabolites in Biological Fluids. *Rapid Commun. Mass Spectrom.* **2003**, *17*, 2632–2638.
2. van Ravenzwaay, B.; Coelho-Palermo Cunha, G.; Leibold, E.; Looser, R.; Mellert, W.; Prokoudine, A.; Walk, T.; Wiemer, J. The Use of Metabolomics for the Discovery of New Biomarkers of Effect. *Toxicol. Lett.* **2007**, *172*, 21–28.
3. Catchpole, G. S.; Beckmann, M.; Enot, D. P.; Mondhe, M.; Zywicki, B.; Taylor, J.; Hardy, N.; Smith, A.; King, R. D.; Kell, D. B.; Fiehn, O.; Draper, J. Hierarchical Metabolomics Demonstrates Substantial Compositional Similarity Between Genetically Modified and Conventional Potato Crops. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 14458–14462.
4. Robertson, D. G. Metabonomics in Toxicology: A Review. *Toxicol. Sci.* **2005**, *2*, 809–822.
5. Gibney, M. J.; Walsh, M.; Brennan, L.; Roche, H. M.; German, B.; van Ommen, B. Metabolomics in Human Nutrition: Opportunities and Challenges. *Am. J. Clin. Nutr.* **2005**, *82*, 497–503.
6. Viant, M. R.; Rosenblum, E. S.; Tjeerdema, R. S. NMR-Based Metabolomics: A Powerful Approach for Characterizing the Effects of Environmental Stressors on Organism Health. *Environ. Sci. Technol.* **2003**, *37*, 4982–4989.
7. Dunn, W.; Ellis, D. Metabolomics: Current Analytical Platforms and Methodologies. *Trends Anal. Chem.* **2005**, *24*, 285–294.
8. Brown, S.; Kruppa, G.; Dasseux, J. Metabolomics Applications of FT-ICR Mass Spectrometry. *Mass Spectrom. Rev.* **2005**, *24*, 223–231.
9. Breitling, R.; Pitt, A.; Barrett, M. Precision Mapping of the Metabolome. *Trends Biotechnol.* **2006**, *24*, 543–548.
10. Southam, A. D.; Payne, T. G.; Cooper, H. J.; Arvanitis, T. N.; Viant, M. R. Dynamic Range and Mass Accuracy of Wide-Scan Direct Infusion Nanoelectrospray Fourier Transform Ion Cyclotron Resonance Mass Spectrometry-Based Metabolomics Increased by the Spectral Stitching Method. *Anal. Chem.* **2007**, *79*, 4595–4602.
11. Han, J.; Danell, R. M.; Patel, J. R.; Gumerov, D. R.; Scarlett, C. O.; Speir, J. P.; Parker, C. E.; Rusyn, I.; Zeisel, S.; Borchers, C. H. Towards High-Throughput Metabolomics Using Ultrahigh-Field Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Metabolomics* **2008**, *4*, 128–140.
12. Taylor, N. S.; Weber, R. J. M.; Southam, A. D.; Payne, T. G.; Hrydziuszko, O.; Arvanitis, T. N.; Viant, M. R. A New Approach to Toxicity Testing in *Daphnia magna*: Application of High Throughput FT-ICR Mass Spectrometry Metabolomics. *Metabolomics* doi: 10.1007/s11306-008-0133-3. In press.
13. Oikawa, A.; Nakamura, Y.; Ogura, T.; Kimura, A.; Suzuki, H.; Sakurai, N.; Shinbo, Y.; Shibata, D.; Kanaya, S.; Ohta, D. Clarification of Pathway-Specific Inhibition by Fourier Transform Ion Cyclotron Resonance/Mass Spectrometry-Based Metabolic Phenotyping Studies. *Plant Physiol.* **2006**, *142*, 398–413.
14. Viant, M. R.; Bearden, D. W.; Bundy, J. G.; Burton, I. W.; Collette, T. W.; Ekman, D. R.; Ezernieks, V.; Karakach, T. K.; Lin, C. Y.; Rochfort, S.; de Ropp, J. S.; Teng, Q.; Tjeerdema, R. S.; Walter, J. A.; Wu, H. International NMR-Based Environmental Metabolomics Intercomparison Exercise. *Environ. Sci. Technol.* **2009**, *43*, 219–225.
15. Gordon, E. F.; Muddiman, D. C. Impact of Ion Cloud Densities on the Measurement of Relative Ion Abundances in Fourier Transform Ion Cyclotron Resonance Mass Spectrometry: Experimental Observations of Coulombically Induced Cyclotron Radius Perturbations and Ion Cloud Dephazing Rates. *J. Mass Spectrom.* **2001**, *36*, 195–203.
16. Kaur; O'Conner. Algorithms for Automatic Interpretation of High Resolution Mass Spectra. *J. Am. Soc. Mass Spectrom.* **2006**, *17*, 459–468.
17. Cech N. B.; Enke, C. G. Practical Implications of Some Recent Studies in Electrospray Ionization Fundamentals. *Mass Spectrom. Rev.* **2001**, *20*, 362–387.
18. Lewis, I. A.; Schommer, S. C.; Hodis, B.; Robb, K. A.; Tonelli, M.; Westler, W. M.; Sussman, M. R.; Markley, J. L. Method for Determining Molar Concentrations of Metabolites in Complex Solutions from Two-Dimensional 1H-13C NMR Spectra. *Anal. Chem.* **2007**, *79*, 9385–9390.
19. Uechi, G. T.; Dunbar, R. C. Space Charge Effects on Relative Peak Heights in Fourier Transform-Ion Cyclotron Resonance Spectra. *J. Am. Soc. Mass Spectrom.* **1992**, *3*, 734–741.

20. Frahm, J. L.; Velez, C. M. C.; Muddiman, D. C. Understanding the Influence of Post-Excite Radius and Axial Confinement on Quantitative Proteomic Measurements Using Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Rapid Commun. Mass Spectrom.* **2007**, *21*, 1196–1204.
21. Dunn, W. B.; Overy, S.; Quick, W. P. Evaluation of Automated Electrospray-TOF Mass Spectrometry for Metabolomic Fingerprinting of the Plant Metabolome. *Metabolomics* **2005**, *1*, 137–148.
22. Marshall, A. G. Theoretical Signal-to-Noise Ratio and Mass Resolution in Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Anal. Chem.* **1979**, *51*, 1710–1714.
23. Bijlsma, S.; Bobeldijk, I.; Verheij, E. R.; Ramaker, R.; Kochhar, S.; Macdonald, I. A.; van Ommen, B.; Smilde, A. K. Large-Scale Human Metabolomics Studies: A Strategy for Data (Pre-) Processing and Validation. *Anal. Chem.* **2006**, *78*, 567–574.
24. Stentiford, G.; Viant, M.; Ward, D.; Johnson, P.; Martin, A.; Wenbin, W.; Cooper, H.; Lyons, B.; Feist, S. Liver Tumors in Wild Flatfish: A Histopathologic, Proteomic, and Metabolomic Study. *OMICS* **2005**, *9*, 281–299.
25. Marshall, A.; Hendrickson, C.; Jackson, G. Fourier Transform Ion Cyclotron Resonance Mass Spectrometry: A Primer. *Mass Spectrom. Rev.* **1998**, *17*, 1–35.
26. Senko, M. W.; Beu, S. C.; McLafferty, F. W. Determination of Monoisotopic Masses and Ion Populations for Large Biomolecules from Resolved Isotopic Distributions. *J. Am. Soc. Mass Spectrom.* **1995**, *6*, 229–233.
27. Gordon, E. F.; Muddiman, D. C. Quantification of Singly Charged Biomolecules by Electrospray Ionization Fourier Transform Ion Cyclotron Resonance Mass Spectrometry Utilizing an Internal Standard. *Rapid Commun. Mass Spectrom.* **1999**, *13*, 164–171.
28. Masselon, C.; Tolmachev, A. V.; Anderson, G. A.; Harkewicz, R.; Smith, R. D. Mass Measurement Errors Caused by "Local" Frequency Perturbations in FTICR Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2002**, *13*, 99–106.
29. Hughey, C. A.; Galasso, S. A.; Zumberge, J. E. Detailed Compositional Comparison of Acidic NSO Compounds in Biodegraded Reservoir and Surface Crude Oils by Negative Ion Electrospray Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Fuel* **2007**, *86*, 758–768.
30. Smith, C.; Want, E.; O'Maille, G.; Abagyan, R.; Siuzdak, G. XCMS: Processing Mass Spectrometry Data for Metabolite Profiling Using Nonlinear Peak Alignment, Matching, and Identification. *Anal. Chem.* **2006**, *78*, 779–787.
31. Sangster, T. P.; Wingate, J. E.; Burton, L.; Teichert, F.; Wilson, I. D. Investigation of Analytical Variation in Metabonomic Analysis Using Liquid Chromatography/Mass Spectrometry. *Rapid Commun. Mass Spectrom.* **2007**, *21*, 2965–2970.
32. Overy, S. A.; Walker, H. J.; Malone, S.; Howard, T. P.; Baxter, C. J.; Sweetlove, L. J.; Hill, S. A.; Quick, W. P. Application of Metabolite Profiling to the Identification of Traits in a Population of Tomato Introgression Lines. *J. Exp. Bot.* **2005**, *56*, 287–296.
33. van der Burgt, Y. E. M.; Taban, I. M.; Konijnenburg, M.; Biskup, M.; Duursma, M. C.; Heeren, R. M. A.; Römpp, A.; van Nieuwpoort, R. V.; Bal, H. E. Parallel Processing of Large Datasets from NanoLC-FTICR-MS Measurements. *J. Am. Soc. Mass Spectrom.* **2007**, *18*, 152–161.
34. Katajamaa, M.; Orešic, M. Processing Methods for Differential Analysis of LC/MS Profile Data. *BMC Bioinf.* **2005**, *6*, 179.
35. Andreev, V. P.; Rejtar, T.; Chen, H.-S.; Moskovets, E. V.; Ivanov, A. R.; Karger, B. L. A Universal Deionizing and Peak Picking Algorithm for LC-MS Based on Matched Filtration in the Chromatographic Time Domain. *Anal. Chem.* **2003**, *75*, 6314–6326.
36. Giavalisco, P.; Hummel, J.; Lisec, J.; Inostroza, A. C.; Catchpole, G.; Willmitzer, L. High-Resolution Direct Infusion-Based Mass Spectrometry in Combination with Whole 13C Metabolome Isotope Labeling Allows Unambiguous Assignment of Chemical Sum Formulas. *Anal. Chem.* **2008**, *80*, 9417–9425.
37. Wu, H.; Southam, A. D.; Hines, A.; Viant, M. R. High-Throughput Tissue Extraction Protocol for NMR and MS-Based Metabolomics. *Anal. Biochem.* **2008**, *372*, 204–212.
38. Pitman, J. *Probability*. Springer-Verlag: New York, 1993; p. 363–364.
39. Parsons, H. M.; Ekman, D. R.; Collette, T. W.; Viant, M. R. Spectral Relative Standard Deviation: A Practical Benchmark in Metabolomics. *The Analyst* **2009**, *134*, 478–485.