

This document is the Accepted Manuscript version of a Published Work that appeared in final form in Analytical Chemistry, copyright © American Chemical Society after peer review and technical editing by the publisher. To access the final edited and published work see <http://pubs.acs.org/doi/abs/10.1021/acs.analchem.5b03684> .

Capturing plant metabolome with direct-immersion *in vivo* solid phase microextraction of plant tissues

Sanja Risticvic¹, Erica A. Souza-Silva^{1#}, Jennifer R. DeEll², Jack Cochran^{3§}, Janusz Pawliszyn^{1*}

¹ Department of Chemistry, University of Waterloo, 200 University Avenue West, Waterloo, ON, N2L 3G1, Canada; ² Ontario Ministry of Agriculture, Food and Rural Affairs, 1283 Blueline Rd. at Hwy #3, Box 587, Simcoe, ON, N3Y 4N5, Canada; ³ Restek, 110 Benner Circle, Bellefonte, PA 16823, USA; [§] The Pennsylvania State University, Forensic Science, University Park, PA 16801, USA.

* Phone: +1 519-888-4641, Fax: +1 519-888-4348, E-mail: janusz@uwaterloo.ca.

ABSTRACT: For the first time, an *in vivo* sampling mode of direct immersion - solid phase microextraction (DI-SPME) was employed to capture the metabolome of living plant specimens, using apple (*Malus x domestica* Borkh.) as a model system. Metabolites were extracted from apple tissues and introduced by thermal desorption into a comprehensive two-dimensional gas chromatography - time-of-flight mass spectrometry (GCxGC-ToFMS) instrument. The feasibility of this sampling approach, based on exploitation of microextraction principles, including negligible depletion of free analyte concentrations, solventless sampling and sample preparation and on-site compatibility was determined in global metabolite analysis. Rather than adopting an approach of traditional sample preparation, requiring metabolism quenching and laborious sample preparation, the objective of the study was to capture metabolome *in vivo*, evaluate the feasibility of approach to provide unbiased extraction coverage and compare analytical precision when different SPME sampling modes are employed. The potential of *in vivo* DI-SPME in quantitative plant metabolomics was assessed by evaluating changes in metabolic fingerprints in response to fruit maturation. The *in vivo* SPME sampling approach has been demonstrated as capable of sampling living systems with high reproducibility, considering that nearly 50% of hundreds of evaluated compounds included in the determination of analytical performance met the 15% RSD FDA criterion. Esters were extracted with high repeatability (% RSD for hexyl butanoate and butyl butanoate 16.5 and 5.9, respectively, 9 determinations in 3 apples) and found to be up-regulated in response to apple fruit maturation.

Introduction

In recent years, metabolomics has been increasingly adopted as a tool for comprehensive characterization of food metabolome and assurance of food quality and safety.¹ Development of suitable extraction methodology remains one of the main challenges in designing an appropriate metabolomics platform.² The choice of sample preparation has enormous impact on the quality of metabolomics data and the array of metabolite classes covered.³ To ensure broad metabolome coverage and a minimum degree of discrimination against metabolite classes present in the investigated

biological system, sample preparation and extraction methodology should be unbiased and non-selective.

SPME was introduced in 1989 to promote rapid sample preparation and provide an efficient way toward integration of sample preparation with separation and detection systems.⁴ In SPME, analytes are extracted by exposing a small volume of polymeric extracting phase directly to the sample matrix or to the headspace above it for a predetermined period of time.⁵⁻⁶ After analyte concentration reaches the distribution equilibrium between the sample matrix and fibre coating, or after the desired sensitivity is obtained in non-equilibrium conditions, the extracting phase is introduced into a gas

chromatography (GC) inlet for thermal desorption or an appropriate desorption solvent for coupling with liquid chromatography (LC).⁵⁻⁶

Solvent-free sample preparation, miniaturized format, and non-exhaustive analyte recovery are features that make SPME an ideal sampling and sample preparation method for *in vivo* analysis of biological systems.⁷ Additionally, according to SPME theory, the amount of analyte extracted by SPME becomes independent of sample volume under the condition of negligible depletion, which is fulfilled when analyzing large sample volumes or compounds with low fibre coating/sample matrix distribution constants (K_{fs}).⁶ Minimized perturbation to the living system is ensured by the small dimensions of the SPME probe relative to the size of many naturally occurring systems and negligible non-exhaustive extraction of free analyte concentration. In addition to studies on determination of environmental pollutants, pharmaceuticals, and personal-care products in tissues and biological fluids of living, freely moving animals, *in vivo* SPME has been employed for the analysis of biologically active compounds in plant, insect, and animal emissions.⁸⁻¹¹

In response to the increasing demand in development of analytical technologies with high specificity and powerful quantitative and qualitative capabilities in metabolomics and considering that in plants, gene expression is directly related to endogenous, rather than emitted volatiles, the objective of the current study was to test the feasibility of *in vivo* DI-SPME for high-resolution metabolite profiling of plant tissues by using apple as a model system. Considering the heterogeneity of many naturally existing food and plant matrices, high sensitivity and selectivity for volatile analytes and enhanced fibre coating lifetime, the use of headspace-SPME (HS-SPME) has provided rewarding method performance characteristics both in terms of sensitivity and characterization of volatile metabolome complement.¹²⁻¹⁴ However, despite assuring protection of extracting phase from non-volatile matrix components, HS-SPME sampling is biased against less volatile and polar metabolites.¹⁵ On the other hand, the proposed *in vivo* approach was based on the utilization of microextraction principles of SPME, including miniaturized format and on-site compatibility of sampling device. Minimized perturbation against studied specimens is ensured by negligible depletion of free analyte concentrations. In order to assure a high degree of sensitivity and specificity when capturing the metabolome and reliability in metabolite annotation, *in vivo* SPME was hyphenated to GCxGC-ToFMS. Analytical performance data was assessed by providing a global evaluation of analytical precision for both *ex vivo* and *in vivo* approaches of sample preparation and analysis. The potential of *in vivo* DI-SPME in quantitative plant metabolomics was

evaluated by statistical evaluation of the collected data in order to detect changes in metabolic fingerprints in response to fruit maturation. Finally, in response to the need for acquisition of representative metabolic datasets and better understanding of matrix effects encountered during DI-SPME, a careful examination of secondary chromatography effects manifested by isovolatility curves and tailing peak profiles for analytes resulting from Maillard reactions was conducted for different SPME extraction phase chemistries.

Experimental Section

***In vivo* DI-SPME procedure and sampling set up**

Considering the previous studies conducted on systematic comparison of extraction efficiency and metabolite coverage for commercially available SPME fibre coatings, 50/30 μm divinylbenzene/carboxen/polydimethylsiloxane (DVB/CAR/PDMS) fibre assemblies (stableflex, 23-gauge needle size, automated format) (Supelco, Oakville, ON, Canada) were used for all experiments.¹⁵⁻¹⁶ DVB/CAR/PDMS coatings overcoated with an external layer of PDMS (as per procedure outlined elsewhere) were employed to monitor metabolome coverage and chromatographic profiles acquired following more effective clean-up of coating surface prior to desorption.¹⁷ SPME fibre coatings for *in vivo* sampling were conditioned as per the supplier recommendations. Prior to *in vivo* sampling, each fibre was additionally conditioned for 5 min at the employed desorption temperature, and needles of SPME fibre assemblies were sealed with Teflon caps. The sampling design consisted of exposing fibre coatings into fruit tissue from directions that were perpendicular with respect to the fruit stem. The inserted coatings were either kept in such a way as to sample the metabolome from three distinct sides of the fruit (sampling design 1, Supplementary Figure 1), or to ensure that a close distribution of sampling positions (1.5 cm distance) was preserved (sampling design 2). Fibre coatings were exposed to the tissue for 60 min at a sampling depth of 3 cm. Triplicate *in vivo* determinations were performed per each apple. Coatings were wiped with Kim Wipes following extraction, washed in water for 10 s and wiped with Kim Wipes again before withdrawing into their respective needles. During transportation, SPME fibre assemblies were stored in dry ice at $-70\text{ }^{\circ}\text{C}$, followed by desorption immediately after arrival to the laboratory or storage in freezer at $-30\text{ }^{\circ}\text{C}$ before analysis. Details on the conditions applied during *ex vivo* extraction can be found in Supplementary Information.

Conditions for acquisition of metabolomics data on GCxGC-ToFMS

A LECO Pegasus 4D GCxGC-ToFMS system equipped with the Agilent 6890N GC and a high speed ToF mass spectrometer (LECO, St. Joseph, MI, USA) was employed in the study. Modulation was performed with a dual-stage quad-jet cryogenic modulator (licensed from Zoex, Houston, TX, USA), and a MultiPurpose Sampler (MPS 2) autosampler was used for automation of the SPME procedure (Gerstel GmbH, Mulheim an der Ruhr, Germany). The column ensemble consisted of a 5% phenyl 95% dimethylpolysiloxane Rxi-5SilMS (30 m x 0.25 mm ID x 0.25 μ m) capillary column from Restek (Bellefonte, PA, USA) in the first dimension coupled to polar polyethylene glycol columns, including BP 20 (SGE Incorporated, Austin, TX, USA) and Stabilwax (Restek, Bellefonte, PA, USA) in the second dimension. The performance of secondary columns from different manufacturers was tested during the course of the study in order to obtain a suitable degree of robustness in terms of second dimension retention time stability when SPME extracts obtained from samples high in water content were introduced into GCxGC. Limited usability, sequentially decreasing retention times, loss of resolution in the second dimension, and distorted peak profiles were found with time for thinner stationary phase polyethylene glycol columns even when long sequences of samples prepared by HS-SPME were concerned. In order to ensure effective metabolite alignment and increased robustness, apple samples intended for statistical comparison were analyzed on a Stabilwax (1 m x 0.25 mm ID x 0.25 μ m) second dimension column; this configuration ensured minimum retention shifts in the second dimension. Thus, for the compounds most susceptible to retention time shifts, including 2-alcohols, 1-alcohols, monoterpene alcohols, and sesquiterpene alcohols, the RSDs of second dimension retention times obtained in HS-SPME analysis of control samples throughout the duration of the metabolomics study were 0.5, 0.8, 1.0, and 1.1%, respectively. Additional details on GCxGC conditions and data processing procedure can be found in Supplementary Information.

Results and Discussion

Global evaluation of analytical precision of *in vivo* DI-SPME and *ex vivo* HS-SPME metabolomics platforms

An evaluation and interpretation of the analytical precision of the *in vivo* DI-SPME – GCxGC-ToFMS metabolomics platform were conducted by considering both the analytical and

biological sources of variation. For sampling design 1, after the implementation of careful data processing and reduction strategy, 357 true metabolites (retention time coordinates shown in Supplementary Figure 2) were considered in the global evaluation of the analytical precision for intra-fruit repeatability. The median % relative standard deviation (% RSD) for this comprehensive evaluation was 22.0% with minimum and maximum % RSD values of 0.3 and 102.3%, respectively. Furthermore, for 41.5% of the peaks, the 15% RSD criterion established by the Food and Drug Administration (FDA) for analytical variability in targeted bioanalysis was satisfied.³ In order to understand the trends in analytical variability of the data set generated when different sides of fruit were sampled, 53 chemically and functionally distinct metabolites (Supplementary Figure 3) were tentatively identified with the aid of mass spectral similarity, retention index, injection of authentic standards and/or GCxGC structurally ordered separations. Homologous groups of structurally related compounds, including aldehydes, 1-alcohols, 2-alkenals, 2-ketones, delta-lactones, gamma-lactones, and terpenoids, were included as well in order to examine the association between molecular weight and analytical precision. The results of the intra-fruit repeatability study are illustrated in Supplementary Table 1. Figure 1 illustrates the correlation between molecular weight and analytical variability for intra-fruit determination for homologous groups of aldehydes, 1-alcohols, and terpenes.

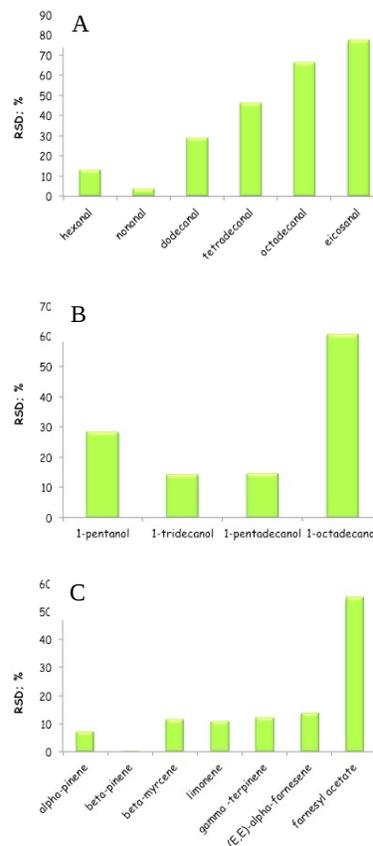


Figure 1. Correlations between molecular weight and analytical precision of *in vivo* DI-SPME – GCxGC-ToFMS metabolomics platform for selected homologous groups of metabolites including A – aldehydes, B – 1-alcohols, and C – terpenoids. With the exception of terpene group containing several metabolites with same molecular weights (alpha-pinene, beta-pinene, beta-myrcene, limonene and gamma-terpinene), metabolites were listed in the order of increasing molecular weight on x-axis.

The obtained results clearly demonstrate significant correlations between metabolite molecular weights and analytical precision. In particular, for all groups of structurally related metabolites, a trend showing an increase in % RSD values with respect to increasing hydrophobicity was observed.

In a subsequent adopted sampling design approach (sampling design 2), all sampling positions were set apart 1.5 cm from each other. Intra- and inter-fruit repeatability were evaluated for 111 one-dimensional peak entries in total (Supplementary Figure 4). The median % RSD for intra-fruit repeatability involving three fibre coatings and a selected apple of earlier harvest maturity (HC-O) was 37.1%, with a minimum of 0.7% (for unidentified compound with retention time coordinates of 1106 and 1.71 s) and a maximum of 125.6% (for metabolite with retention time coordinates of 3108 and 0.910 s). A comparison of analytical precision for two *in vivo* sampling designs (sampling design 1 and sampling design 2) was made by considering homologous series of esters (Supplementary Figure 5). RSDs (%) obtained for hexyl butanoate in apples of earlier harvest maturity with apple codes HC-O-2, HC-O-3, and HC-O-5 were 12.0, 6.5, and 13.5%, respectively (Figure 2A). Similarly, for butyl butanoate (Figure 2B), intra-fruit repeatability is represented by 5.9, 4.5 and 9.0% RSD for HC-O-2, HC-O-3, and HC-O-5 apples, respectively. Considering that % RSDs for sampling design 1 obtained on the basis of triplicate intra-fruit determinations were 37.3 for hexyl butanoate and 18.6 for butyl butanoate, the implementation of the sampling design in which fibres were inserted in the apple at a distance of 1.5 cm apart from each other resolved limitations associated with high analytical variability observed for higher molecular weight compounds. Figure 2 illustrates that in addition to excellent intra-fruit repeatability, analytical precision for inter-fruit determinations was also satisfactory (16.5 and 5.9 % RSD for hexyl butanoate and butyl butanoate, respectively).

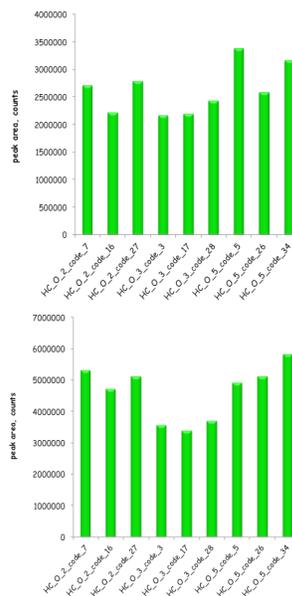


Figure 2. Extraction efficiencies of DVB/CAR/PDMS coatings in *in vivo* sampling of hexyl butanoate (A) and butyl butanoate (B) (three apples of earlier harvest maturity – HC-O, $n=3$ extractions per apple).

When the three sampling positions were adjusted as to sample the metabolome from different sides across the fruit cortex (sampling design 1), hydrophobic metabolites were not sampled with high repeatability. However, inter-fibre repeatability was satisfactory based on the extraction efficiencies obtained for a wide spectrum of analyte properties in HS-SPME analysis of water samples spiked with representative metabolites (system precision mixture). Based on the fundamentals of SPME extraction, the amount of analyte extracted is governed by the distribution constant if equilibrium is reached.⁵⁻⁶ In accordance with theoretical principles, as well as experimentally verified observations from a previous study on matrix effects and analyte extraction rates during SPME extraction of complex systems, slow equilibration was characteristic of compounds with high hydrophobicity and high K_{fs} .¹⁶ Even though hydrophobic compounds partition very strongly into the SPME coating, their diffusion process through a semi-solid sample matrix is very slow. When such analytes were extracted from apple homogenate, equilibrium was not reached within 60 min, and the rate of mass transfer governed the extraction efficiency and extraction speed.¹⁶ Even though precision is compromised for short exposure times in the steeper regions of the extraction time profile curve, the negligible extraction time variations during *in vivo* sampling are not expected to result in such significant relative errors in amounts extracted.⁵⁻⁶ Similarly, mass transfer rates between the investigated tissue and the fibre coatings should not be affected by sampling

position, provided that the same compartments are sampled. Based on the performance of individual coatings *in vivo* and the observation that the enhancement in extraction recovery was not specific to the particular coating and sampling position, potential adsorption of macromolecules on the coating surface and subsequent fouling of the extraction phase should not lead to such dramatic differences in extraction recoveries for the three determinations (Supplementary Figure 6).

Each plant organ, tissue, and compartment is composed and characterized by a specific set of metabolites that are present in specific distributions and are very often differentially affected by external stimuli. For example, other authors reported non-uniform glucosinolate distribution throughout the leaf tissue of *Arabidopsis thaliana*, differing carbohydrate profiles obtained from petiole recesses and leaf disks taken from the same leaf of *Cucurbita maxima*, and dependence of spatial localizations of volatile metabolites on the distribution of their primary and secondary metabolite precursors.¹⁸⁻²⁰

However, in addition to metabolite spatial localizations between different plant compartments, the intra-compartmental variations in metabolite content have been extensively reported in particular in response to environmental stimuli.²¹ For example, Rudell et al. determined that scald in 'Granny Smith' apples was eliminated on the side of the fruit directly exposed to UV-vis irradiation treatment after harvest.²² Increased light treatment duration affected scald susceptibility and resulted in metabolome alterations, including decreased levels of α -farnesene.²² In addition to artificial light, sunlight-exposed apple portions were less susceptible to scald and had heightened phenolic levels, while on the contrary, bagging the fruit to limit sunlight exposure during fruit ripening increased scald incidence.²² In strawberry fruit, the effect of environmental and growing conditions, including harvest date and shading effect, was investigated with respect to levels of volatile metabolites by Watson et al.²³ Considering that 47% shading treatment caused a significant reduction in the concentration of volatile metabolites as compared to control fruit of the same harvest, it was suggested that brief light integral had a significant effect on strawberry flavour quality.²³

Differences in extracted amounts for high molecular weight metabolites observed between the three sampling positions may hence be attributed to differences in spatial inter- and intra-compartmental distribution of analytes. In the sampling design implemented, the perpendicular direction of inserted fibre coatings with respect to apple stem and consistent exposure depth were implemented to ensure *in vivo* sampling of same compartments within apple. However, as

mentioned previously, intra-compartmental metabolome variations are also possible in response of the plant to environmental stimuli.

Comparison of *in vivo* DI-SPME and *ex vivo* HS-SPME metabolomics platforms in terms of analytical precision

For comparative purposes, the identified analytes listed in Supplementary Table 1 and several additionally included metabolites (total of 40 compounds) were also considered for determination of performance characteristics of *ex vivo* HS-SPME metabolomics. Determination of long term stability of samples for a complex matrix of highly dynamic nature after they were thawed, placed in extraction vials and stored on autosampler tray before extraction was conducted. Stability of metabolomics samples previously subjected to proper metabolism quenching procedure has also impact on degree of throughput of a metabolomics platform because it defines the number of samples that can be prepared and subsequently submitted to automated extraction procedure. The median % RSD corresponding to intra-fruit determination and the long term stability of samples subjected to HS-SPME ($n = 10$) was 13.6%, with minimum and maximum RSDs of 2.5 and 88.6% for 2-hexenal and limonene, respectively. The profiles of extracted responses for selected compounds were unusual as the storage time on the autosampler tray was increasing (0 to 12 hours) (Supplementary Figure 7). For example, for 6-methyl-5-hepten-2-one, *beta*-myrcene, (2E)-2-octenal, *trans-beta*-damascenone, and (Z,Z)-farnesol, % RSDs were 35.2, 27.3, 20.9, 41.2, and 37.5%, respectively (Supplementary Figure 7). Conversely, intra-fruit reproducibility and long-term stability for the same compounds were excellent for HS-SPME analysis of samples for which storage on autosampler tray was avoided and extraction was conducted immediately after sample thawing. The % RSDs ($n=10$ fresh sample preparations) for 6-methyl-5-hepten-2-one, *beta*-myrcene, (2E)-2-octenal, *trans-beta*-damascenone, and (Z,Z)-farnesol determined in freshly analyzed samples were 7.0, 6.7, 7.3, 6.6, and 19.5%, respectively.

However, although fresh sample preparation resolved some stability-related issues, this approach did not reduce analytical variability for hydrophobic and higher molecular weight compounds extracted by HS-SPME. While % RSDs for 1-tridecanol, 1-pentadecanol, and 2-heptadecanone obtained for HS-SPME extractions of apple homogenate ($n=10$) conducted immediately after sample preparation were 41.4, 58.7, and 60.6 %, respectively, % RSDs corresponding to *in vivo* extraction of intact apple ($n=3$) and subsequent GCxGC-ToFMS analysis

were 14.4, 14.7, and 43.0, respectively (Supplementary Table 2).

The high analytical variability in HS-SPME extraction of hydrophobic compounds was at first observed in one of the earlier studies.¹⁶ Poor solubility in aqueous medium and subsequent competitive adsorption onto the vial walls also attributed to a narrow linear dynamic range in HS-SPME analysis of aqueous samples containing highly hydrophobic analytes.¹⁶ However, in addition to glass walls, the presence of competing phases in naturally occurring heterogeneous systems, including the apple homogenate, further reduces the free concentration of high K_{fs} metabolites. According to the fundamentals of SPME, when different phases compete for analytes of interest, the amount extracted is dependent on the volume of competing phase and the analyte distribution constant between the competing phase and fibre coating.⁶ The samples included in determination of analytical precision for *ex vivo* HS-SPME were processed in same way with respect to length of extraction, thawing time and minimum storage on autosampler tray. Each individual determination was based on thawing of homogenate corresponding to same apple fruit stored in different vials during freeze storage. The high analytical variability is likely attributed to competition of different phases for analytes of interest, as well as differing amounts of competing phase/skin tissue in stored vials, and in different aliquots from the same vial corresponding to the same apple fruit. Under these circumstances, among different determinations of same apple homogenate, losses of hydrophobic metabolites are expected to occur to a different extent (Supplementary Figure 8). Even though centrifugation is expected to improve analytical precision, the experimental findings (Supplementary Table 2) demonstrate that the use of *in vivo* SPME should resolve challenges in sampling of highly hydrophobic analytes.

A summary of analytical precision results obtained when different *ex vivo* and *in vivo* SPME sample preparation approaches were employed in apple metabolomics is given in Figure 3. Analytical precision obtained when sampling *in vivo* was comparable to, or superior in performance when compared to traditional HS-SPME sampling. Furthermore, *in vivo* sampling addressed issues encountered during high-throughput HS-SPME analysis of labile and hydrophobic metabolites.

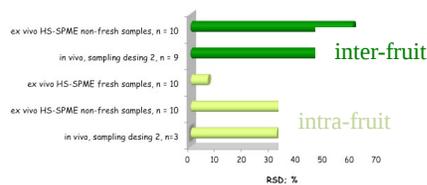
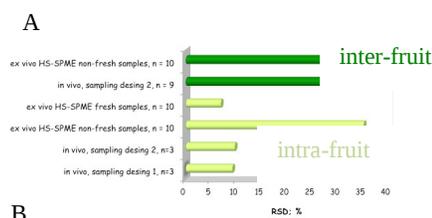


Figure 3. Intra-fruit (light green color) and inter-fruit (dark green color) variability for 6-methyl-5-hepten-2-one (plot A) and *trans-beta*-damascenone (plot B) obtained for *ex vivo* HS-SPME (stored and freshly analyzed samples) and *in vivo* DI-SPME (sampling designs 1 and 2).

For *ex vivo* SPME, issues with metabolite stability when samples are stored on the autosampler tray are likely related to heterogeneity and the dynamic nature of the investigated matrix, composed of thousands of chemically diverse analytes. Based on the literature reports, even freeze storage is not exempt from issues related to alterations of fruit volatile metabolite composition and loss of sample integrity.^{20,24} A review on the chemistry of beer ageing published by Vanderhaegen and co-workers highlights the nature and extent of changes occurring in the chemical composition of this food commodity during storage.²⁵ These authors reported a number of reactions and mechanisms responsible for the formation and/or decomposition of selected groups of metabolites, including carbonyl compounds evaluated in this study. The summary of aging reactions producing carbonyl compounds, including oxidation of higher alcohols, aldol condensation, oxidation of unsaturated fatty acids (leading to production of aldehydes such as (2E)-2-octenal), Strecker degradation of amino acids, and formation of *trans-beta*-damascenone was thoroughly presented. The formation of (2E)-nonenal and other linear C₄-C₁₀ alkenals and alkanals in beer during storage correlates to increasing levels of (2E)-2-octenal during storage of apple samples on the autosampler tray.²⁵

Statistical interpretation of data and potential of *in vivo* DI-SPME in quantitative metabolomics

Evaluation of analytical precision for fruit-to-fruit *in vivo* determinations (9 fibre coatings, sampling design 2, apples of earlier harvest maturity considered, including HC-O-2, HC-O-3 and HC-O-5) resulted in a median RSD of 55.6% (minimum 10.6%, maximum 214.8%) for 111 compounds. This is to be expected considering that in global metabolomics, biological variability significantly exceeds analytical variability. As concluded by Watson et al., significant strawberry fruit-to-fruit variation in concentrations of volatile metabolites was observed and attributed to different maturity levels of fruit.²³ The effect of

uneven maturity on high fruit-to-fruit variability in metabolite content is even more exacerbated here for 'Honeycrisp' apples, as for this cultivar maturity issues were recognized in previous studies.²⁶ In addition, the apples considered in this global evaluation of inter-fruit repeatability were grown in different locations with respect to the tree, light supplement and nutrient availability and all these factors induce significant alterations in metabolome profile. The findings of a study reported by Tikunov et al. on application of HS-SPME multivariate analysis platform for analysis of tomato samples are relevant for highlighting repeatability obtained during determination of volatile and semivolatile metabolome composition.²⁷ The authors reported the range of fruit-to-fruit variation within genotype, starting from 8 (2E-heptenal) to 35 (2-methylbutanol) % RSD. On the other hand, biological variation between genotypes ranged between 28% and 198%.²⁷ Fruit-to-fruit variability was therefore significant despite the implementation of traditional and well-known headspace SPME approach of preconcentration and isolation. Considering the use of multiple fibre coatings, small temperature and sunlight exposure variations during large-scale on-site sampling, differences in intra-compartmental localizations of metabolites and potential differential adhesion of matrix macrocomponents on surface of different coatings, higher analytical variability in semi-quantitative determination of volatile and semivolatile metabolome fraction by *in vivo* SPME is expected.

Since gene expression and associated enzyme and metabolite complements change as fruit tissue develops, the maturity of the fruit at the time of harvest is a major determinant of the specific volatile metabolome complement.²⁸ In order to determine whether substantial intra-fruit and inter-fruit variability in extraction efficiencies obtained by *in vivo* DI-SPME have the potential to adversely affect the quality of metabolomics interpretation, statistical treatment of data (12 apple samples per maturity group, in total two harvest maturity groups) was performed. Initially, one-way analysis of variance (ANOVA) was conducted on extracted SPME responses represented in terms of global means for each apple. The metabolites considered included butyl propanoate, butyl butanoate, ethyl hexanoate, butyl 2-methylbutanoate and estragole (Supplementary Table 3). The selection of metabolites was influenced by the study performed by Schaffer and coworkers who generated a transgenic line of 'Royal Gala' apple silenced for production of ethylene-induced ripening.²⁹ Specific pathways of aroma biosynthesis are ethylene-dependent, including the activity of alcohol *o*-acyltransferase (AAT) which catalyzes esterification of alcohols and acyl-CoAs derived from fatty acid and amino acid metabolism to form volatile esters.³⁰⁻³¹ Hence in

response to application of external ethylene, these fruits underwent a normal climacteric burst leading to production of increased levels of ester, polypropanoid, and terpene volatile metabolites.²⁹ Data presented in Supplementary Table 3 suggest that at 95% confidence level, the contents of target metabolites were significantly over-expressed in the group of samples having higher maturity index. The results are clearly correlated to the study by Schaffer et al., where the contents of butyl propanoate, butyl butanoate, ethyl hexanoate, butyl 2-methylbutanoate and estragole increased over an 8-day period following ethylene application.

Furthermore, the levels of 225 volatile and semivolatile metabolites comprising a wide array of analyte properties (processing and data reduction strategies are outlined in Supplementary Information) and extracted by *in vivo* DI-SPME were used to characterize samples according to maturity stage by means of a principal components analysis (PCA) model. Several principal components were extracted, among which PC1, PC2, and PC5 accounted for 36.3, 34.5, and 3.9% of total variability in the data set, respectively. Therefore, despite substantial fruit-to-fruit variability, the scores plot presented in Figure 4 indicates quantitative capabilities of *in vivo* SPME since two maturity stages could be clearly defined by the levels of hundreds of chemically distinct metabolites which were extracted by SPME *in vivo*.

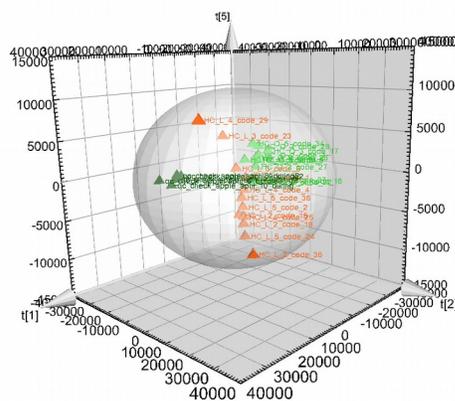


Figure 4. Scores plot corresponding to Principal Components Analysis (PCA) of *in vivo* DI-SPME data for HC-L apples (higher maturity index, represented by orange triangles) and HC-O apples (lower maturity index, represented by light green triangles). The distribution of QC samples is outlined by dark green triangles.

Metabolome coverage and matrix effects

The impressive metabolome coverage obtained with *in vivo* DI-SPME is reflected by the extraction of several classes of chemically diverse compounds. The metabolome is comprised of a

homologous series of alkanes, acetate esters, en-1-yl acetate esters, propanoate esters, 2-methylpropanoate esters, butanoate esters, 2-methylbutanoate esters, aliphatic aldehydes, (2E)-2-alkenals, (3Z)-3-alkenals, (2E,4E)-dienals, saturated aliphatic ketones, cyclic ketones, enones, cyclic enones, diketones, saturated primary alcohols, saturated tertiary alcohols, unsaturated aliphatic alcohols, diols, aromatics, di-aromatics, alkenyl benzenes, benzyl ethers and alkenyl benzyl ethers, benzyl diethers, saturated carboxylic acids, aromatic aldehydes, aromatic ketones, benzyl alcohols, benzyl acetates, and glycol ethers. Supplementary Table 4 illustrates a list of 198 true metabolites from a reduced data matrix (mass spectral similarity greater than 750, elimination of metabolites for which separation efficiency and modulator effectiveness were not optimum, elimination of metabolites having streaking one-dimensional peak profiles characteristic of analyte decomposition) that were tentatively identified on the basis of retention time and mass spectral comparison with reference standards, retention index comparison, and GCxGC molecular structure-retention relationships. Supplementary Figure 9 illustrates peak apex plot for metabolite grouping into respective classes of structurally related compounds.

The chromatographic profile corresponding to the *in vivo* DI-SPME metabolic fingerprint (Figure 5) was highly complex, with several components exhibiting overloaded and distorted peak profiles. Under the conditions of non-linear chromatography, overloading in GCxGC may occur on either of the two columns and is controlled by temperature and stationary phase and solute chemistries.³² However, examination of *in vivo* chromatograms did not indicate such relationships for the overloaded components. Furthermore, considering that non-linearity is dependent on the injected amount and, in the case of *in vivo* SPME, distorted peak profiles were not observed for metabolites having high extraction sensitivity, these secondary chromatography effects were not related to overloading due to increased extraction efficiency. Peak broadening due to slow desorption from the inner carboxen layer of the DVB/CAR/PDMS coating is ruled out in light of our previous investigations which confirmed the occurrence of distorted peak profiles when CAR/PDMS coating was employed.¹⁵ Furthermore, the absence of the relationship between solute chemical property and peak distortion excluded the possibility of analyte adsorption on active sites in the inlet. Isovolatility and streaking profiles in the GCxGC experiment have been associated with decomposition and dissociation of chemical compounds into products that are GC amenable.³² In the case of *in vivo* SPME, the formation of solutes occurred during thermal desorption, followed by their continuous

introduction into GCxGC. The tentative annotation of analyte identities revealed that major compounds exhibiting such behaviour were intermediates and products of the Maillard reaction, including 2-Furancarboxaldehyde (Furfural), 2,5-Furandicarboxaldehyde, and 5-(Hydroxymethyl)-2-furfural.³³⁻³⁶ Maillard reaction occurs at high thermal treatments between the free amino group of lysine and/or other amino acids and the carbonyl groups of reducing sugars such as glucose and maltose.³³⁻³⁶

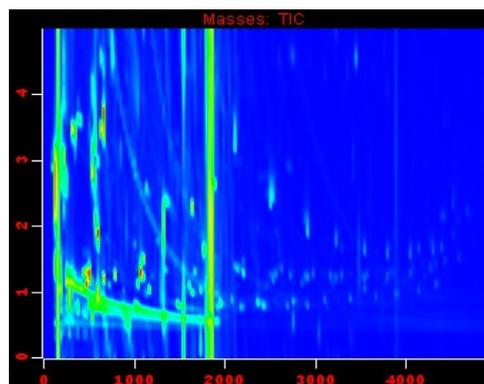


Figure 5. Contour plot of GCxGC-ToFMS TIC chromatogram corresponding to *in vivo* DI-SPME sampling

The contour plot of the GCxGC chromatographic profile corresponding to the elution window of one of the major decomposition products, 5-(Hydroxymethyl)-2-furfural, is shown in Supplementary Figure 10. While chromatographic profiles corresponding to *ex vivo* DI-SPME also contained decomposition products of the Maillard reaction (comparison between *ex vivo* and *in vivo* sampling approaches is given in Supplementary Table 5), these compounds were generated to a lesser extent, possibly due to high sample dilution with water during metabolism quenching or potential residual enzymatic activity leading to hydrolysis of sugars in samples having high aqueous content.³

The ineffective clean-up of fibre coatings following direct immersion extraction of complex systems, the incompatibility of extraction phase chemistry with the sample matrix and the potential attachment of non-volatile matrix constituents on the coating surface are factors that may have adverse implications in global metabolomics of food systems. In addition to potentially changing the chemical properties of the extraction phase, enhancement/suppression of the amount of analyte extracted may occur depending on the nature of the matrix components attached on the surface. In addition to reduced coating reusability and method precision, the production of artifacts and loss of resolution for trace metabolites overlapping with overloaded components degrade the quality of the metabolomics data set.

In the current study, these secondary decomposition products represented 3% of the total metabolome complement captured by *in vivo* SPME and were recognized by the ability of the GCxGC experiment to provide relationships between molecular structure and retentions in the two-dimensional separation space and identify various non-linear chromatography effects.^{32,37} Due to complexity of the chromatograms, manual data processing was conducted and, whenever possible, multi-dimensional features of GCxGC-ToFMS were exploited for quantification of components that suffered overloaded peak overlap. In addition, for PDMS-overcoated DVB/CAR/PDMS fibre coatings (Supplementary Figure 11), the production of decomposition products of the Maillard reaction was reduced due to the more effective clean-up of the outer PDMS coating layer.

Conclusions

In this study, *in vivo* sampling mode of DI-SPME was employed in metabolomics of living plant specimens, using apple as a model system. Solvent-free sample preparation, miniaturized format, non-exhaustive analyte recovery and independence of the amount extracted on sample volume under the conditions of negligible depletion are features that make SPME an ideal sample preparation tool for *in vivo* analysis of biological systems. The proposed *in vivo* approach should ensure minimized perturbation to the living system due to the small dimensions of the SPME probe relative to the size of many naturally occurring systems and negligible depletion of free analyte concentrations. Hence, in comparison to traditional methods of sample preparation and *ex vivo* SPME which require laborious sample preparation and metabolism quenching steps to assure sample integrity is preserved, on-site implementation of *in vivo* SPME should lead to acquisition of more accurate metabolome and capture of unique metabolic fingerprints. This will be the focus of future investigations highlighting the potential of *in vivo* SPME in plant metabolomics. The future studies should also investigate the effect of *in vivo* DI-SPME extraction on potential metabolome perturbations and activation of plant defense mechanisms.

Based on the analytical precision data obtained for hundreds of components comprising a wide spectrum of physicochemical properties, novel *in vivo* DI-SPME sampling approach is capable of sampling living systems with high reproducibility. Despite the variability in amounts extracted by *in vivo* SPME that was observed for some compounds and attributed to differences in spatial metabolome localizations between different sampling positions, nearly 50% of evaluated compounds in extensive metabolomics

data set met 15% RSD FDA criterion for analytical variability in targeted bioanalysis. Intra-fruit and fruit-to-fruit repeatability for esters assured detection of metabolic changes resulted by fruit ripening and these compounds that represent crucial biomarkers of contribution to aroma, consumer acceptance and varietal apple differentiation were sampled with high reproducibility. Hence, analytical precision obtained by *in vivo* DI-SPME was comparable to, or superior in performance when compared to traditional HS-SPME sampling. In addition, in comparison to *ex vivo* HS-SPME, the implementation of *in vivo* approach resulted in improvement of intra-fruit repeatability for labile and hydrophobic metabolites (Supplementary Table 5). The statistical treatment of data obtained for hundreds of chemically distinct metabolites that were extracted by *in vivo* DI-SPME allowed detection of changes in metabolic profiles resulted by different fruit maturation levels and discrimination of apple samples on the basis of one of the most important attributes of food quality.

Future DI-SPME *in vivo* platforms should exploit spatial resolution advantages of SPME. The definite elucidation of spatial metabolome localizations has so far been hindered by the lack of *in situ* sampling methods and fruit-to-fruit variability with respect to maturity and developmental level. The implementation of a miniaturized SPME device for *in vivo* sampling allows repeated and multiple samplings of individual fruit, thus eliminating the manifestation of biological inter-species variations in data interpretation. Considering the impressive potential of spatial metabolomics to reveal biochemical interactions that can be implemented for future flavour design efforts in fruits, the capability of *in vivo* SPME in this field should be further explored.

Future studies will also emphasize more comprehensive evaluation of matrix effects encountered during complex sample analysis by DI-SPME for a variety of matrices. These investigations are currently underway and may include, for example, studying the nature of matrix components that are likely to attach to coating surface and cause fibre fouling, the effects of different co-extracted matrix components on extraction recoveries for target analytes, and the design of SPME coatings and clean-up strategies allowing improved compatibility with extraction of complex samples.

AUTHOR INFORMATION

Corresponding Author

* Phone: +1 519-888-4641. Fax: +1 519-888-4348.
E-mail: janusz@uwaterloo.ca.

Present Addresses

#: Instituto de Química, Universidade Federal do Rio Grande do Sul (UFRGS), Av. Bento Gonçalves, 9500, 91501-970 Porto Alegre-RS, Brazil

ACKNOWLEDGMENT

The authors thank the Natural Sciences and Engineering Research Council of Canada (NSERC), Ontario Apple Growers (OAG, Vineland Station, ON, Canada), as well as LECO (St. Joseph, MI, USA), and Gerstel (Linthicum, MD, USA) for collaborative support.

ASSOCIATED CONTENT

Supporting Information

Additional information pertaining to Experimental Section and section on Results and Discussion as noted in text. This material is available free of charge via the Internet at <http://pubs.acs.org>.

REFERENCES

- (1) Castro-Puyana, M.; Herrero, M. *Trends Anal. Chem.* **2013**, *52*, 74–87.
- (2) Hall, R. D. *New Phytol.* **2006**, *169*, 453–468.
- (3) t'Kindt, R.; Morreel, K.; Deforce, D.; Boerjan, W.; Bocxlaer, J. V. *J Chromatogr B* **2009**, *877*, 3572–3580.
- (4) Arthur, C. L.; Pawliszyn, J. *Anal. Chem.* **1990**, *62*, 2145–2148.
- (5) Pawliszyn, J. In *Solid phase microextraction: Theory and practice*; Pawliszyn, J., Ed.; Wiley-VCH: New York, 1997; pp 43–96.
- (6) Pawliszyn, J. In *Handbook of solid phase microextraction*; Pawliszyn, J., Ed.; Chemical Industry Press: Beijing, 2009; pp 13–54.
- (7) Musteata, F. M.; Vuckovic, D. In *Handbook of solid phase microextraction*; Pawliszyn, J., Ed.; Chemical Industry Press: Beijing, 2009; pp 345–377.
- (8) Zhang, X.; Oakes, K. D.; Wang, S.; Servos, M. R. *Trends Anal. Chem.* **2012**, *32*, 31–39.
- (9) Ouyang, G.; Vuckovic, D.; Pawliszyn, J. *Chem. Rev.* **2011**, *111*, 2784–2814.
- (10) Vuckovic, D.; Risticvic, S.; Pawliszyn, J. *Angew. Chem. Int. Ed.* **2011**, *50*, 5618–5628.
- (11) Augusto, F.; Valente, A. L. P. *Trends Anal. Chem.* **2002**, *21*, 428–438.
- (12) Risticvic, S.; Lord, H.; Górecki, T.; Arthur, C. L.; Pawliszyn, J. *Nat. Protoc.* **2010**, *5*, 122–139.
- (13) Risticvic, S.; Chen, Y.; Kudlejova, L.; Vatinno, R.; Baltensperger, B.; Stuff, J. R.; Hein, D.; Pawliszyn, J. *Nat. Protoc.* **2010**, *5*, 162–176.
- (14) Risticvic, S.; Niri, V. H.; Vuckovic, D.; Pawliszyn, J. *Anal. Bioanal. Chem.* **2009**, *393*, 781–795.
- (15) Risticvic, S.; DeEll, J. R.; Pawliszyn, J. *J. Chromatogr. A* **2012**, *1251*, 208–218.
- (16) Risticvic, S.; Pawliszyn, J. *Anal. Chem.* **2013**, *85*, 8987–8995.
- (17) Silva, E. A. S.; Pawliszyn, J. *Anal. Chem.* **2012**, *84*, 6933–6938.
- (18) Kueger, S.; Steinhauser, D.; Willmitzer, L.; Giavalisco, P. *Plant J.* **2012**, *70*, 39–50.
- (19) Moing, A.; Aharoni, A.; Biais, B.; Rogachev, I.; Meir, S.; Brodsky, L.; Allwood, J. W.; Erban, A.; Dunn, W. B.; Kay, L.; de Koning, S.; de Vos, R. C. H.; Jonker, H.; Mumm, R.; Deborde, C.; Maucourt, M.; Bernillon, S.; Gibon, Y.; Hansen, T. H.; Husted, S.; Goodacre, R.; Kopka, J.; Schjoerring, J. K.; Rolin, D.; Hall, R. D. *New Phytol.* **2011**, *190*, 683–696.
- (20) Ryan, D.; Robards, K. *Sep. Purif. Rev.* **2006**, *35*, 319–356.
- (21) Kim, H. K.; Verpoorte, R. *Phytochem. Anal.* **2010**, *21*, 4–13.
- (22) Rudell, D. R.; Mattheis, J. P. *Postharvest Biol. Technol.* **2009**, *51*, 174–182.
- (23) Watson, R.; Wright, C. J.; McBurney, T.; Taylor, A. J.; Linforth, R. S. T. *J. Exp. Bot.* **2002**, *53*, 2121–2129.
- (24) Allwood, J. W.; De Vos, R. C. H.; Moing, A.; Deborde, C.; Erban, A.; Kopka, J.; Goodacre, R.; Hall, R. D. *Methods Enzymol.* **2011**, *500*, 299–336.
- (25) Vanderhaegen, B.; Neven, H.; Verachtert, H.; Derdelincks, G. *Food Chem.* **2006**, *95*, 357–381.
- (26) Cline, J.; Gardner, J. OMAFRA Horticulture News **2009**.
- (27) Tikunov, Y.; Lommen, A.; de Vos, C. H. R.; Verhoeven, H. A.; Bino, R. J.; Hall, R. D.; Bovy, A. J. *Plant Physiol.* **2005**, *139*, 1125–1137.
- (28) Lara, I.; Echeverría, G.; Graell, J.; López, M. L. *J. Agric. Food Chem.* **2007**, *55*, 6087–6095.
- (29) Schaffer, R. J.; Friel, E. N.; Souleyre, E. J. F.; Bolitho, K.; Thodey, K.; Ledger, S.; Bowen, J. H.; Ma, J.-H.; Nain, B.; Cohen, D.; Gleave, A. P.; Crowhurst, R. N.; Janssen, B. J.; Yao, J.-L.; Newcomb, R. D. *Plant Physiol.* **2007**, *144*, 1899–1912.
- (30) Villatoro, C.; Altisent, R.; Echeverría, G.; Graell, J.; López, M. L.; Lara, I. *Postharvest Biol. Technol.* **2008**, *47*, 286–295.
- (31) Ortiz, A.; Graell, J.; Lara, I. *Sci. Hortic.* **2011**, *131*, 6–14.
- (32) Marriott, P. J.; Massil, T.; Hügel, H. *J. Sep. Sci.* **2004**, *27*, 1273–1284.
- (33) Delgado-Andrade, C.; Seiquer, I.; Haro, A.; Castellano, R.; Pilar Navarro, M. *Food Chem.* **2010**, *122*, 145–153.
- (34) Rooijen, C.; Bosch, G.; Wierenga, P. A.; Hendriks, W. H.; van der Poel, A. F. B. *Anim. Feed Sci. Tech.* **2014**, *198*, 238–247.
- (35) Rufián-Henares, J. A.; Delgado-Andrade, C.; Morales, F. J. *Food Chem.* **2009**, *114*, 93–99.
- (36) Chávez-Servín, J. L.; Carbot, K. T.; García-Gasca, T.; Castellote, A. I.; López-Sabater, M. C. *Food Chem.* **2015**, *166*, 486–491.
- (37) Gardner, J. Y.; Brillhart, D. E.; Benjamin, M. M.; Dixon, L. G.; Mitchell, L. M.; Dimandja, J.-M. D. *J. Sep. Sci.* **2011**, *34*, 176–185.

for TOC only

