Original Research Article

**Differentiation of bread made with whole grain and refined wheat (*T. aestivum*)
flour using LC/MS-based chromatographic fingerprinting and chemometric
approaches**

Ping Geng, James M. Harnly, and Pei Chen*

Food Composition and Methods Development Laboratory, Beltsville Human Nutrition
Research Center, Agricultural Research Service, U.S. Department of Agriculture,
Beltsville, Maryland, 20705, United States

* *Correspondence to:* Pei Chen, Food Composition and Methods Development Laboratory,
Beltsville Human Nutrition Research Center, Agricultural Research Service, USDA,
Building 161, BARC-East, 10300 Baltimore Avenue, Beltsville, MD 20705, USA.
E-mail: pei.chen@ars.usda.gov; Phone: 1(301)-504-8144; Fax: 1(301)-504-8314

Abstract

A fuzzy chromatography mass spectrometric (FCMS) fingerprinting method combined with chemometric analysis was established to differentiate between breads made from whole wheat (WW) flour and refined wheat (RW) flour. The chemical compositions of the bread samples were profiled using ultrahigh-performance liquid chromatography with high-resolution accurate-mass multistage mass spectrometry (UHPLC-HRAM-MSⁿ) with atmospheric pressure chemical ionisation (APCI) in positive ionisation mode. Principal component analysis (PCA) and soft independent modelling of class analogy (SIMCA) of the FCMS fingerprints revealed the components responsible for the chemical differences between WW and RW flour/bread samples. Alk(en)ylresorcinols (ARs) were demonstrated to be the most important markers for differentiation between WW and RW flour/breads. Diglycerides (DGs), and phosphatidylethanolamine (PE) also contributed significantly to the classification. Bread crumb and crust showed no difference in chemical profiles in this study. SIMCA, using WW modelling, could be a potent and robust tool for authentication of WW breads.

KEYWORDS: Alk(en)ylresorcinols; phosphatidylethanolamine; diglycerides; UHPLC-HRMS; APCI; FCMS; PCA; SIMCA; Food composition; Food analysis

Comment [1]: Please confirm spelling: alk (en) ylresorcinol or alk(en)ylresorcinol

Comment [2]: Please validate addition of these 2 keywords by JFCA editor.

Chemical compounds studied in this article: 5-*n*-heneicosylresorcinol (PubChem CID: 155461); 5-*n*-nonadecylresorcinol (PubChem CID: 161858); 5-*n*-tricosylresorcinol (PubChem CID: 155462); 5-*n*-pentadecylresorcinol (PubChem CID: 155463); PE(16:0/18:2) (PubChem CID: 9546747); linoleic acid (PubChem CID: 5280450)

Abbreviations used

FCMS fuzzy chromatography mass spectrometric

WW whole wheat

RW refined wheat

AR alk(en)ylresorcinol

DG diglyceride

TG triglyceride

PE phosphatidylethanolamine

PC principal component

1. Introduction

Cereal grain kernels consist of three main parts: endosperm, bran, and germ. The vast majority of cereal products are traditionally prepared from refined flour after removal of bran and germ, the two parts of grain kernels containing most of the dietary fibre and other bioactive components. Compared to refined grains, most whole grains provide more protein, fibre, and other traditional nutrients, including minerals (calcium, magnesium, and potassium) and many phytochemicals. Consumption of whole grain foods has been linked to reduced risk of obesity (Flint et al., 2013; Vitaglione et al., 2015), type 2 diabetes (Parker et al., 2013; Ye et al., 2012), cardiovascular diseases (Katcher et al., 2008; Ye et al., 2012), and certain cancers (Kyro et al., 2013; Kyrø et al., 2013). As public awareness of eating healthy foods grows, whole grain flour in substitution for refined flour is increasingly used in preparation of bread, cookies, crackers, breakfast cereals, and snacks.

Differentiation of a whole grain product from a refined grain product can be very challenging. Food labels can often be misleading. The Whole Grains Council has created an official packaging symbol, the Whole Grain Stamp, to help consumers. The 100% Whole Grain Stamp assures you that a food contains a full serving or more of whole grain in each labelled serving and that all the grain used is whole grain, whereas the basic Whole Grain Stamp indicates the products contain at least 8 g of whole grain per labelled serving.

Bread is one of the most commonly eaten grain sources. For several thousand years, bread has been one of the major constituents of the human diet, and wheat is by far the

most important cereal in bread making. The impact of processing and baking steps on the levels and the composition of flour constituents has been well documented (Ktenioudaki et al., 2015). However, the objective of the present work is to develop a method that can differentiate breads baked from whole wheat (WW) and refined wheat (RW). This is more challenging than differentiating the raw flours as the added baking ingredients (water, yeast, sugar, salt, fats, etc.) and baking process make bread more chemically complex than raw flour.

To date, an analytical method to differentiate whole grain products from refined ones is not available. The chemical differences between germ, bran, and endosperm were investigated in our previous work (Geng et al., 2015). A fuzzy chromatography-mass spectrometric (FCMS) fingerprinting and chemometric method was successfully used to differentiate WW flour from RW flour. The method was expanded in the current study to differentiate breads made from them.

Two chromatographic methods were used in this study. The full UHPLC separation profiled the subtle differences between WW and RW samples. Fuzzy chromatography was applied to achieve high throughput data acquisition by using the shortest elution time possible on a short column. For FCMS, the goal is not optimised separation between individual components, but to alleviate the potential of ion suppression effects presented in complex matrices when a flow injection mass spectrometric (FIMS) method (Chen et al., 2014; Huang et al., 2015; Sun and Chen, 2011) is used.

Principal component analysis (PCA) is usually used to reduce a set of original variables (m/z values detected in flour and bread samples in this study) and to extract a small number of latent factors (principal components, PCs) for the analysis of relationships between groups of samples. Soft independent modelling of class analogy (SIMCA) is a typical class-modeling method based on the evaluation of the PCs of each category. There are two steps involved in SIMCA: 1) modelling, building one PCA model for each category, and 2) classifying new samples, where unknown samples are compared to the class models. Samples may belong to one of the categories, more than one category, or none of the categories. One of the advantages of SIMCA is that all pertinent results can be displayed graphically with an exceptional insight regarding the specific data structures behind the modelled patterns.

2. Experimental

2.1 Materials and Chemicals

The wheat flour samples from 12 different brands were purchased from local grocery stores or online. Unbleached all-purpose flour (RW) and whole wheat flour (WW) from each brand were selected into RW and WW group, respectively, to achieve two groups of flour samples. Each group contained 12 flour samples, one from each brand. RED STAR[®] active dry yeast, Domino[®] sugar, and MORTON[®] salt were bought from local stores.

HPLC grade methanol, formic acid, and *n*-butanol were purchased from VWR International, Inc. (Clarksburg, MD). HPLC water was prepared from distilled water using a Milli-Q system (Millipore Laboratories, Bedford, MA).

2.2 Bread Making and Sample Preparation

Breads were prepared on a 100-g scale based on the AACC Intl. Approved Method 10-10.03, “Optimized Straight-Dough Bread-Making Method” but without using shortening. Flour (100.0 g on a 14.0% moisture base), water, sugar (6.0% on flour basis), active dry yeast (5.3% on flour basis) and salt (1.5% on flour basis) were mixed with optimum mixing time and water absorption. (Baig and Hosney, 1977; Finney, 1984) The bread dough was fermented for 90 min, proofed for 30 min at 30 °C, and then baked at 215 °C for 15 min.

The finished breads were lyophilised, milled and sieved (60 mesh). Crumb and crust were processed separately. One hundred milligrams of sample (raw flour, bread crumb, or bread crust) were extracted with 4.0 mL of water saturated *n*-butanol (Bahrami et al., 2014; Fisher et al., 1964; MacMurray and Morrison, 1970) using a sonicator (Advanced Sonic Processing Systems, Oxford, CT) at 16 kHz with 300 W power for 40 min. The slurry mixture was centrifuged at 13000 *g* for 15 min in a Contifuge 28RS centrifuge (Heraeus, South Plainfield, NJ) at 4 °C. The supernatant was filtered through a 17-mm (0.20 µm) PVDF syringe filter (VWR Scientific, Seattle, WA) and stored at –80 °C until analysis.

2.3 UHPLC-HRMSⁿ Conditions

The UHPLC-HRMSⁿ system consists of an LTQ Orbitrap XL MS with an Accela 1250 binary pump, a PAL HTC Accela TMO autosampler, an Accela PDA detector (Thermo Fisher Scientific, San Jose, CA), and an Agilent G1316A column compartment (Agilent, Santa Clara, CA). APCI positive mode was operated at 450 °C, the heated capillary temperature at 275 °C, the corona discharge current at 5.0 µA, and sheath gas (N₂) and auxiliary gas (N₂) were at 60 and 5 arbitrary units, respectively. The scan event cycle used a FT full scan mass spectrum at a resolving power of 30,000 and three data-dependent MS², MS³, MS⁴ events acquired by LTQ triggered by the most intensive ion from each previous scan event. The automated gain control (AGC) target value was set to 200,000 for FT full scan, and 10,000 for IT MSⁿ scan, respectively. MSⁿ activation parameters used an isolation width of 1.0 amu, max ion injection time of 200 ms, normalisation collision energy at 35%, and activation time of 30 ms.

The full LC separation was carried out on a Poroshell 120EC-C₁₈ column (2.1 × 150 mm, 2.7 µm; Agilent, Santa Clara, CA) with a flow rate of 0.4 mL/min. The mobile phase consisted of a combination of **A** (0.1% formic acid in water) and **B** (0.1% formic acid in methanol). A gradient elution program was employed as follows: 85% **B** at 0–3.0 min, 90% **B** at 5.0–7.0 min, 95–100% **B** over 15.0–20.0 min, and held at 100% **B** until 120.0 min, 100.0–120.0 min convert to waste. The column temperature was 40 °C.

The fuzzy chromatographic separation was carried out on a Hypersil Gold C₁₈ Javelin HTS column (10 mm × 2.1 mm, 1.9 µm particle size) (Thermo Fisher Scientific, San Jose, CA) with a flow rate of 0.4 mL/min. The same mobile phase with full LC method was used with the following gradient: 85% **B** at 0–0.5 min, 100% **B** at 2–7 min, and

maintained at 100% B until 15 min for column washing, 10–15 min divert to waste. The column temperature was 40 °C.

2.4 Data processing

The FCMS fingerprint (averaged spectra 0.5–3.0 min) of each sample consisted of a two-dimensional matrix (ion counts *versus* mass for m/z 100–1000) and was exported to Excel (Microsoft, Inc., Bellingham, WA) as nominal masses for data preprocessing. The preprocessing in Excel involved combining FCMS fingerprints of all the samples into one matrix, sorting the data by sample names, and aligning the masses (each spectrum might have a different number of masses because the m/z below the detection threshold would not be exported into the mass list from the spectrum). A zero was inserted for each missing m/z , using an Excel macro written in-house. The resulting two-dimensional 72×901 matrix (samples $\times$ m/z variables, 24 samples for flour/bread crumb/bread crust) was then imported to SIMCA-P for PCA and SIMCA.

3. Results and discussions

3.1 LC-MS chromatographic fingerprint of the bread and flour samples

The full UHPLC-MS chromatograms of flour and bread samples are shown in [Fig. 1](#). The chromatograms of RW samples (RW flour, Fig. 1A; RW bread crumb, Fig. 1B; and RW bread crust, Fig. 1C) look very similar as do the chromatograms of WW samples (WW flour, Fig. 1D; WW bread crumb, Fig. 1E; and WW bread crust, Fig. 1F). It appears that the processing and baking procedure did not significantly affect the sample composition

(for compounds detectable under experimental conditions). The major differences between RW and WW samples were observed between retention times 4.5–25 min. More peaks were observed in WW samples, which indicated that those peaks could be potential discriminatory compounds for RW and WW samples.

Twenty-six main peaks between 4.5–25 min were putatively identified based on accurate mass measurements, MS^2 – MS^4 spectra, ring-plus-double-bond equivalent (RDB), chromatographic behaviour, and comparison with data in the literature (Carretero et al., 2008; Fang and Barcelona, 1998; Holčapek et al., 2003; Jakab et al., 2002; Knödler et al., 2007; Knödler et al., 2008; Murphy and Harrison, 1994; Neff et al., 2001; Pulfer and Murphy, 2003; Rozenberg et al., 2003; Taguchi et al., 2005; Zhu et al., 2011) or METLIN database (Tautenhahn et al., 2012). The peaks were putatively identified as free fatty acids, alk(en)ylresorcinols (ARs), phytosterols, diglycerides (DGs), and phosphatidylethanolamine (PE), as listed in [Table 1](#). The characteristic MS^n spectra of each compound category were shown in [Supplementary spectra Figs. S1–S5](#). The triglycerides (TGs) were characterised in our previous work (Geng et al., 2015) and will not be discussed here. Although it is possible to distinguish RW and WW samples by visual observation of the full LC-MS chromatograms, running a long LC-MS method is very time consuming. Therefore, a FCMS/chemometric approach was established for high throughput discrimination.

3.2 PCA of RW and WW samples using FCMS fingerprint datasets

The FCMS fingerprinting method roughly separated fatty acids, alk(en)ylresorcinols (ARs), phytosterols, phospholipids, DGs, and TGs into different groups (Fig. S6). The

averaged spectra between 0.5–3.0 min were selected as fingerprints for PCA and SIMCA, due to the more discriminating chromatographic fingerprints, omitting TGs. The first three principal components accounted for 93.6% of the total variance, with values for PC1, PC2 and PC3 of 83.7%, 6.5% and 3.4%, respectively. The score plot of PC1 vs. PC2 (Fig. 2A) shows two distinct clusters as the RW group and WW group are clearly separated. The flour and bread samples were separated in the PC1 vs. PC3 score plot (Fig. 2B). RW flour, RW bread, WW flour, and WW bread formed four clusters, the crust and crumb samples could not be separated.

The variables responsible for the clustering are shown in the PCA loading plot (Fig. 3). Eight ions (m/z 405, 377, 433, 461, 375, 403, 349, and 447) contributing strongly to the separation were previously identified to be ARs (Table 1). ARs are phenolic lipids present in high amounts in the bran fraction of wheat and rye and generally absent in refined cereal products (Landberg et al., 2008). The variable line plot of m/z 405, 377, 433 shown in Fig. 4 demonstrated their distribution across samples. The normalised intensities of ions at m/z 405, 377, 433 in WW samples were higher than those in RW samples. What's more, the intensity curve of flour samples was slightly above their bread counterparts, and it was speculated that this difference of intensity led to the separation of flour and bread samples in the PCA score plot (Fig. 2B). The result was consistent with Winata and Lorenz's report that AR content in wheat sourdough bread was reduced during fermentation and/or the baking process (Winata and Lorenz, 1997). It was not surprising that the ARs could be detected in the RW samples. Even under optimum conditions, the milling process cannot completely separate the bran from the endosperm,

so a small amount of fine bran particles escapes separation and passes through to the flour streams and “contaminates” them (Mousia et al., 2004) .

Other significant contributors for separation were ions at m/z 618/619, 600/601, 576/577, 552/553, 616/617, 602/603, 578/579, 598/599, 279, 397, and 383. These ions could be observed in both loading plots (Fig.3A, PC1 vs. PC2 and Fig.3B PC1 vs. PC3). The variable with nominal mass 618/619 corresponded to the ion at m/z 617.5132 and its M+1 isotope ion observed in HRAM spectra; likewise, 600/601 corresponds to 599.5033 and its isotope ion. Those ions were identified as $[M+H]^+$ and $[M+H-H_2O]^+$ ions of DG (18:2/18:2), respectively. Part of intensity of 600/601 was contributed by $[M+H-141]^+$ ion of PE (18:2/18:2), which was the characteristic ion for PE produced by losing the polar head group phosphoryl ethanolamine (Pulfer and Murphy, 2003; Taguchi et al., 2005). The absolute abundance of DG peaks was high enough to give rise to the significance of their isotope ions in the loading plot. 576/577 were deduced to be $[M+H-H_2O]^+$ ion of DG (16:0/18:2) and/or $[M+H-141]^+$ ion of PE (16:0/18:2); 552/553 corresponded to $[M+H-H_2O]^+$ ion of DG (16:0/16:0); 616/617 to $[M+H]^+$ ion of DG (18:3/18:2); 602/603 came from $[M+H-H_2O]^+$ ion of DG (18:1/18:2) and/or $[M+H-141]^+$ ion of PE (18:1/18:2); 578/579 from $[M+H-H_2O]^+$ ion of DG (16:0/18:1); and 598/599 from $[M+H-H_2O]^+$ ion of DG (18:3/18:2). The ion at m/z 279.2316 was identified to be α -linolenic acid. Ions at m/z 383.3670 and 397.3829 were characteristic $[M+H-H_2O]^+$ ions of phytosterols campesterol and β -sitosterol, respectively.(Rozenberg et al., 2003) Detailed information of these discriminative compounds was shown in Table 1 and Supplementary spectra Figs. S1–S5.

3.3 Modelling and classification of flour and bread samples

RW and WW models were independently created based on flour and bread samples and boundaries at the 95% confidence level were constructed. The optimum number of PCs of each model was determined by cross-validation. With the default setting in SIMCA-P, 1/7 of the observations were excluded during each cross-validation round. The Si vs. Hi plot (Distance vs. Leverage)(Esbensen et al., 2002), also called a membership plot, can be used to visualise the classification of samples (Fig. 5). “Si” (the residual deviation) is the measurement of the distance of a sample to the model plane, corresponding to “DModX” in software SIMCA-P (version 13.0; Umetrics, Sweden). “Hi” (leverage) displays the distance to the model centre for each selected sample, corresponding to “Hotelling T²” in SIMCA-P. With the Si vs. Hi plot, how far the new object is positioned from the model plane (DModX), and how a difference in the object can be considered from the others (Hotelling's T²) are clearly demonstrated. Two limits, one from each statistical measure, were used in the classification. Objects that fall inside these limits are very likely to belong to the model at the chosen significance level. The thresholds at significance level of 0.05 are shown as crossed dotted lines: horizontal for DModX and vertical for Hotelling T² limit.

In Fig. 5A, RW flour and bread samples were used to build an RW model (training set), all the RW and WW samples (test set) were classified by cross-validation using RW model. Based on DModX, RW samples are statistically different from WW at the 95% confidence limit. Two of 36 RW samples (5.6%) are above the 95% limit. This result is

close to the statistical expectation. All the WW samples fall outside the 95% confidence limit and are rejected by the RW model. In Fig. 5B, WW flour and bread samples were used to build WW model (training set); all the RW and WW samples (test set) were classified by cross-validation using WW model. Based on DModX, WW samples are statistically different from RW ones at the 95% confidence limit. One of 36 WW samples is above the 95% limit, accounting for about 2.8%, within the statistical expectation. All the RW samples fall outside the 95% confidence limit and are rejected by the WW model.

The performance of a class model can be evaluated by means of a number of parameters, the most common of which are sensitivity and specificity. The samples belonging to the class being modelled are called true positive (TP) if they are correctly found inside or false negative (FN) if they fall outside. By analogy, samples extraneous to that class are referred to as false positive (FP) if they are found within the boundaries or true negative (TN) if they are correctly outside.

$$\text{Sensitivity} = \frac{TP}{TP+FN}$$

$$\text{Specificity} = \frac{TN}{TN+FP}$$

In this study, for the RW model, the sensitivity (RW samples identified as RW) and specificity (WW samples rejected by RW model) was 94.4% and 100%, respectively; for the WW model, the sensitivity (WW samples identified as WW) and specificity (RW samples rejected by WW model) was 97.2% and 100%, respectively.

Based on DModX, the observations are assigned to the correct class with high sensitivity or rejected from the wrong class with 100% specificity. However, based on Hotelling's T^2 ,

the observations could be classified correctly (Fig 5B) or randomly assigned to a class (Fig. 5A). Therefore, for this data set, the DModX dominates the Hotelling's T^2 . DModX only is efficient enough to discriminate between RW and WW classes, and, originally, Wold used only the Q statistic (DModX) to determine the classification (Harnly et al., 2013; Wold, 1976). From the value of Hotelling T^2 of observations, the distance of each WW sample to the RW model centre is much closer than the distance of each RW sample to WW model centre. For these reasons, WW modelling is a better choice for the verification of WW-authenticity claims than RW modelling.

4. Conclusions

The rapid FCMS fingerprinting method in combination with chemometric approaches was developed to discriminate bread made from whole grain and refined wheat flours efficiently, irrespective of their brands. ARs were confirmed to be the markers of whole grain bread (Geng et al., 2015) and, in addition, some DGs, PEs, and free fatty acids were found to be responsible for the chemical difference between WW and RW bread. SIMCA, based on WW model, was demonstrated to be a potent and robust approach to authenticate WW bread.

Acknowledgment

The authors would like to thank Ms. Shirley Cabiness for her excellent technical assistance in bread baking. This research is supported by the Agricultural Research Service of the U.S. Department of Agriculture and an Interagency Agreement with the Office of Dietary Supplements of the National Institutes of Health.

Supplementary material

The characteristic spectra of free fatty acid (Fig. S1), alkylresorcinol (Fig. S2), phytosterol (Fig. S3), diglyceride (Fig. S4), and phosphatidylethanolamine (Fig. S5); and the typical FCMS total ion chromatogram of bread samples (Fig. S6) are available online.

Notes

The authors declare no competing financial interest.

References

- Bahrami, N., Yonekura, L., Linforth, R., Carvalho da Silva, M., Hill, S., Penson, S., Chope, G., Fisk, I.D., (2014). Comparison of ambient solvent extraction methods for the analysis of fatty acids in non-starch lipids of flour and starch. *Journal of the Science of Food and Agriculture* 94(3), 415-423.
- Baig, M., Hoseney, R., (1977). Effects of mixer speed, dough temperature, and water absorption on flour-water mixograms. *Cereal Chemistry* 54(3), 605-615.
- Carretero, A.S., Carrasco-Pancorbo, A., Cortacero, S., Gori, A., Cerretani, L., Fernández-Gutiérrez, A., (2008). A simplified method for HPLC-MS analysis of sterols in vegetable oil. *European Journal of Lipid Science and Technology* 110(12), 1142-1149.
- Chen, P., Sun, J., Ford, P., (2014). Differentiation of the Four Major Species of Cinnamons (*C. burmannii*, *C. verum*, *C. cassia*, and *C. loureiroi*) Using a Flow Injection Mass Spectrometric (FIMS) Fingerprinting Method. *Journal of Agricultural and Food Chemistry* 62(12), 2516-2521.
- Esbensen, K.H., Guyot, D., Westad, F., Houmoller, L.P., (2002). *Multivariate data analysis: in practice: an introduction to multivariate data analysis and experimental design*. Multivariate Data Analysis.
- Fang, J., Barcelona, M.J., (1998). Structural determination and quantitative analysis of bacterial phospholipids using liquid chromatography/electrospray ionization/mass spectrometry. *Journal of Microbiological Methods* 33(1), 23-35.

Finney, K.F., (1984). An Optimized, Straight-Dough, Bread-Making Method after 44 Years. *Cereal Chemistry* 61(1), 20-27.

Fisher, N., Broughton, M., Peel, D., Bennett, R., (1964). The lipids of wheat. II.—lipids of flours from single wheat varieties of widely varying baking quality. *Journal of the Science of Food and Agriculture* 15(5), 325-341.

Flint, A.J., Field, A., Rimm, E., (2013). Childhood whole-grain intake and young adult obesity. *The FASEB Journal* 27(1_MeetingAbstracts), 847.830.

Geng, P., Harnly, J.M., Chen, P., (2015). Differentiation of Whole Grain from Refined Wheat (*T. aestivum*) Flour Using Lipid Profile of Wheat Bran, Germ, and Endosperm with UHPLC-HRAM Mass Spectrometry. *Journal of Agricultural and Food Chemistry* 63(27), 6189-6211.

Harnly, J., Chen, P., Harrington, P.d.B., (2013). Probability of identification: adulteration of American ginseng with Asian ginseng. *Journal of AOAC International* 96(6), 1258-1265.

Holčapek, M., Jandera, P., Zderadička, P., Hrubá, L., (2003). Characterization of triacylglycerol and diacylglycerol composition of plant oils using high-performance liquid chromatography–atmospheric pressure chemical ionization mass spectrometry. *Journal of Chromatography A* 1010(2), 195-215.

Huang, H., Sun, J., McCoy, J.-A., Zhong, H., Fletcher, E.J., Harnly, J., Chen, P., (2015). Use of flow injection mass spectrometric fingerprinting and chemometrics for differentiation of three black cohosh species. *Spectrochimica Acta Part B: Atomic Spectroscopy* 105, 121-129.

Jakab, A., Héberger, K., Forgács, E., (2002). Comparative analysis of different plant oils by high-performance liquid chromatography–atmospheric pressure chemical ionization mass spectrometry. *Journal of Chromatography A* 976(1), 255-263.

Katcher, H.I., Legro, R.S., Kunselman, A.R., Gillies, P.J., Demers, L.M., Bagshaw, D.M., Kris-Etherton, P.M., (2008). The effects of a whole grain–enriched hypocaloric diet on cardiovascular disease risk factors in men and women with metabolic syndrome. *The American journal of clinical nutrition* 87(1), 79-90.

Knödler, M., Berardini, N., Kammerer, D.R., Carle, R., Schieber, A., (2007). Characterization of major and minor alk (en) ylresorcinols from mango (*Mangifera indica* L.) peels by high-performance liquid chromatography/atmospheric pressure chemical ionization mass spectrometry. *Rapid Communications in Mass Spectrometry* 21(6), 945-951.

Knödler, M., Kaiser, A., Carle, R., Schieber, A., (2008). Profiling of Alk (en) ylresorcinols in cereals by HPLC-DAD-APCI-MS n. *Analytical and Bioanalytical Chemistry* 391(1), 221-228.

Ktenioudaki, A., Alvarez-Jubete, L., Gallagher, E., (2015). A review of the process-induced changes in the phytochemical content of cereal grains: The breadmaking process. *Critical Reviews in Food Science and Nutrition* 55(5), 611-619.

Kyro, C., Olsen, A., Landberg, R., Skeie, G., Loft, S., Aman, P., Leenders, M., Dik, V., Siersema, P., Pischon, T., (2013). Alkyresorcinols (biomarkers of whole-grain intake) and risk of colorectal cancer in the European prospective investigation into cancer and nutrition. *Annals of Nutrition and Metabolism* 63(Supplement 1), 1207-1208.

Kyrø, C., Olsen, A., Landberg, R., Skeie, G., Loft, S., Åman, P., Leenders, M., Dik, V.K., Siersema, P.D., Pischon, T., (2013). Plasma alkylresorcinols, biomarkers of whole-grain wheat and rye intake, and incidence of colorectal cancer. *Journal of the National Cancer Institute* 106(1), djt352.

Landberg, R., Kamal-Eldin, A., Salmenkallio-Marttila, M., Rouau, X., Aman, P., (2008). Localization of alkylresorcinols in wheat, rye and barley kernels. *Journal of Cereal Science* 48(2), 401-406.

MacMurray, T., Morrison, W., (1970). Composition of wheat-flour lipids. *Journal of the Science of Food and Agriculture* 21(10), 520-528.

Mousia, Z., Edherly, S., Pandiella, S.S., Webb, C., (2004). Effect of wheat pearling on flour quality. *Food Research International* 37(5), 449-459.

Murphy, R.C., Harrison, K.A., (1994). Fast atom bombardment mass spectrometry of phospholipids. *Mass Spectrometry Reviews* 13(1), 57-75.

Neff, W.E., Byrdwell, W.C., List, G.R., (2001). Triacylglycerol structures of food fats high in saturated acids by HPLC and mass spectrometry. *Journal of Liquid Chromatography & Related Technologies* 24(6), 837-854.

Parker, E.D., Liu, S., Van Horn, L., Tinker, L.F., Shikany, J.M., Eaton, C.B., Margolis, K.L., (2013). The association of whole grain consumption with incident type 2 diabetes: the Women's Health Initiative Observational Study. *Annals of Epidemiology* 23(6), 321-327.

Pulfer, M., Murphy, R.C., (2003). Electrospray mass spectrometry of phospholipids. *Mass Spectrometry Reviews* 22(5), 332-364.

Rozenberg, R., Ruibal-Mendieta, N.L., Petitjean, G., Cani, P., Delacroix, D.L., Delzenne, N.M., Meurens, M., Quetin-Leclercq, J., Habib-Jiwan, J.-L., (2003). Phytosterol analysis and characterization in spelt (*Triticum aestivum* ssp. *spelta* L.) and wheat (*T. aestivum* L.) lipids by LC/APCI-MS. *Journal of Cereal Science* 38(2), 189-197.

Sun, J., Chen, P., (2011). A flow-injection mass spectrometry fingerprinting method for authentication and quality assessment of *Scutellaria lateriflora*-based dietary supplements. *Analytical and Bioanalytical Chemistry* 401(5), 1577-1584.

Taguchi, R., Houjou, T., Nakanishi, H., Yamazaki, T., Ishida, M., Imagawa, M., Shimizu, T., (2005). Focused lipidomics by tandem mass spectrometry. *Journal of Chromatography B* 823(1), 26-36.

Tautenhahn, R., Cho, K., Uritboonthai, W., Zhu, Z., Patti, G.J., Siuzdak, G., (2012). An accelerated workflow for untargeted metabolomics using the METLIN database. *Nature Biotechnology* 30(9), 826-828.

Vitaglione, P., Mennella, I., Ferracane, R., Rivellese, A.A., Giacco, R., Ercolini, D., Gibbons, S.M., La Storia, A., Gilbert, J.A., Jonnalagadda, S., (2015). Whole-grain wheat consumption reduces inflammation in a randomized controlled trial on overweight and obese subjects with unhealthy dietary and lifestyle behaviors: role of polyphenols bound to cereal dietary fiber. *The American Journal of Clinical Nutrition*, ajcn. 088120.

Winata, A., Lorenz, K., (1997). Effects of fermentation and baking of whole wheat and whole rye sourdough breads on cereal alkylresorcinols. *Cereal Chemistry* 74(3), 284-287.

Wold, S., (1976). Pattern recognition by means of disjoint principal components models. *Pattern recognition* 8(3), 127-139.

Ye, E.Q., Chacko, S.A., Chou, E.L., Kugizaki, M., Liu, S., (2012). Greater whole-grain intake is associated with lower risk of type 2 diabetes, cardiovascular disease, and weight gain. *The Journal of nutrition* 142(7), 1304-1313.

Zhu, Y., Conklin, D.R., Chen, H., Wang, L., Sang, S., (2011). 5-Alk(en)ylresorcinols as the major active components in wheat bran inhibit human colon cancer cell growth. *Bioorganic & Medicinal Chemistry* 19(13), 3973-3982.

Figure captions

Fig. 1. Typical LC-MS chromatograms of flour and bread samples: A) RW flour, B) RW bread crumb, C) RW bread crust, D) WW flour, E) WW bread crumb, F) WW bread crust, G) zoom-in spectrum of range 4.5-24.6 min

Fig. 2. Score plot of PC1 vs. PC2 (A) and PC1 vs. PC3 (B). FL = flour; CS = crust; CM = crumb

Fig. 3. Loading plot of PC1 vs. PC2 (A) and PC1 vs. PC3 (B)

Fig. 4. Line plots of variables at m/z 405, 377, and 433. 0 = RW, 1 = WW, CS = crust, CM = crumb, FL = flour.

Fig. 5. Si vs. Hi plot (DModX vs. Hotelling T^2 plot) for SIMCA using RW and WW models. A: all refined samples, including RW flour and RW bread crumb and crust, were used to build RW model; B: all WW samples, including WW flour and WW bread crumb and crust, were used to build WW model. Sample icons: ●RW, and ▲WW.

Tables

Table 1 Putatively identified potential discriminatory compounds

Peak No.	Ions <i>m/z</i>	RT	Adduct	Formula	Error (ppm)	MS ² –MS ⁴ data	Tentative identification
1	279.2316	5.04	[M+H] ⁺	C ₁₈ H ₃₀ O ₂	0.919	MS2[279.2]: 261.2(100) ^a , 243.2(58), 223.1(27), 209.2(30), 195.1(21), 137.2(31), 123.2(42), 109.2(35), 95.2(38); MS3[279.2→261.2]: 243.3(100) , 173.2(13), 135.2(13); MS4[279.2→261.2→243.3]: 215.2(12), 201.2(44), 187.3(61), 173.2(100), 161.2(28), 159.2(45), 147.3(22), 145.2(50), 133.2(20), 131.3(26)	α-linolenic acid (C18:3)
2	281.2472	6.22	[M+H] ⁺	C ₁₈ H ₃₂ O ₂	1.091	MS2[281.2]: 263.2(100) , 245.2(19) ; MS3[281.2→263.2]: 245.2(100) ; MS4[281.2→263.2→245.2]: 217.1(26), 203.2(43), 189.2(93), 175.2(96), 163.3(39), 161.2(100), 149.2(37), 147.1(71), 135.1(29), 133.2(55), 121.2(28), 119.1(26)	linoleic acid (C18:2)
3	283.2628	7.50	[M+H] ⁺	C ₁₈ H ₃₄ O ₂	1.260	MS2[283.3]: 265.3(100) , 247.3(27); MS3[283.3→265.3]: 247.3(100) ; MS4[283.3→265.3→247.3]: 205.2(26), 191.2(65), 177.3(97), 165.2(19), 163.2(100), 149.1(98), 135.1(73), 121.2(48), 109.2(30), 95.2(34)	oleic acid (C18:1)
4	349.3098	8.77	[M+H] ⁺	C ₂₃ H ₄₀ O ₂	0.879	MS2[349.3]: 334.4(2), 123.2(4), 111.0(100) ; MS3[349.3→111.0]: 93.2(100)	5- <i>n</i> -heptadecylresorcinol (AR-C17:0)
5	375.3255	9.18	[M+H] ⁺	C ₂₅ H ₄₂ O ₂	0.685	MS2[375.3]: 123.1(100) ; MS3[375.3→123.1]: 95.1(100), 77.0(5), 67.1(17)	(14' <i>Z</i>)-5-(nonadeca-14'-enyl)resoreinolresorcinol (AR-C19:1)
6	338.3416	9.46	[M+H] ⁺	C ₂₂ H ₄₃ ON	0.418	MS2[338.3]: 321.3(100) , 303.3(77) ; MS3[338.3→321.2]: 303.4(100) ^b	13-Doosenamidedocosenamide

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman, Italic

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman, Italic

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Peak No.	Ions m/z	RT	Adduct	Formula	Error (ppm)	MS ² →MS ⁴ data	Tentative identification
7	377.3413	10.94	[M+H] ⁺	C ₂₅ H ₄₄ O ₂	0.284	MS2[377.3]: 127.0(2), 123.1(2) , 111.2(100) ; MS3[377.3→→111.2]: 93.2(100)	5-<i>n</i>-nonadecylresorcinol (AR-C19:0)
8	403.3566	11.42	[M+H] ⁺	C ₂₇ H ₄₆ O ₂	1.134	MS2[403.4]: 123.1(100) ; MS3[403.4→→123.1]: 95.2(100), 57.1(14)	(12'Z)-5-(heneicos-12'-enyl) resorcinol —(AR-C21:1)
9	447.3828	11.76	[M+H] ⁺	C ₂₉ H ₅₀ O ₃	1.055	MS2[447.4]: 429.4(100) , 177.3(12), 163.1(39) , 149.1(30); MS3[447.4→→429.4]: 347.3(10), 333.4(9), 177.2(14), 163.2(85), 149.1(100) , 123.1(34)	5-(2'-oxotricosyl) resorcinol (AR-C23:oxo)
10	405.3727	13.46	[M+H] ⁺	C ₂₇ H ₄₈ O ₂	0.018	MS2[405.4]: 387.4(40), 251.2(28), 169.1(27), 140.8(79), 127.2(75), 124.8(67), 123.3(100)	5-<i>n</i>-heneicosylresorcinol (AR-C21:0)
11	617.5132	15.09	[M+H] ⁺	C ₃₉ H ₆₈ O ₅	1.217	MS2[617.5]: 599.5(100) ; MS3[617.5→→599.5]: 581.6(16), 543.5(100), 525.7(55), 507.7(14), 337.4(23) , 319.2(31) , 263.3(33) , 245.3(14)	DG(18:2/18:2)
	599.5033	15.09	[M+H- H ₂ O] ⁺	C ₃₉ H ₆₇ O ₄ ⁺	0.145	MS2[599.5]: 581.6(16), 543.5(100), 525.7(55), 507.7(14), 337.4(23) , 319.2(31) , 263.3(33) , 245.3(14) MS2[383.4]: 301.3(45), 299.3(31), 287.4(31), 273.3(39), 257.3(65), 247.3(46), 243.2(100), 229.3(33), 215.2(47), 203.3(58), 189.2(55), 175.3(45), 161.2(79), 149.3(34), 147.3(56), 135.3(57); MS3[383.4→→243.2]: 215.3(31), 201.2(34), 187.3(84), 173.2(34), 161.2(80), 149.3(47), 147.2(100), 135.2(75), 133.3(50)	campesterol
12	383.3670	15.77	[M+H- H ₂ O] ⁺	C ₂₈ H ₄₈ O	0.595	MS2[433.4]: 415.7(72), 405.5(70), 279.4(50), 265.1(41), 251.4(49), 237.5(41), 197.2(61), 182.7(41), 169.0(52), 167.4(56), 155.2(100), 138.1(41), 123.2(59)	5-<i>n</i>-tricosylresorcinol (AR-C23:0)
13	433.4040	16.08	[M+H] ⁺	C ₂₉ H ₅₂ O ₂	0.017		

Peak No.	Ions m/z	RT	Adduct	Formula	Error (ppm)	MS ² →MS ⁴ data	Tentative identification
14	740.5220	16.15	[M+H] ⁺	C ₄₁ H ₇₄ O ₈ NP	0.650	MS2[740.5]: 599.5(100) ; MS3[740.5→599.5]: 582.0(14), 525.6(40), 507.7(31), 319.3(100), 263.4(88), 245.3(28)	PE(18:2/18:2)
	599.5027	16.15	[M+H-141] ⁺	C ₃₉ H ₆₇ O ₄ ⁺	1.146	MS2[599.5]: 543.4(17), 525.4(23), 507.6(12), 337.3(10), 319.4(97), 263.2(100), 245.3(32) , 189.2(15), 175.3(18) MS2[397.4]: 315.3(50), 299.3(35), 287.3(36), 261.3(37), 257.2(78), 243.3(100), 229.2(43), 215.2(48), 203.2(62), 189.2(59), 175.2(50), 161.3(79), 149.3(38), 147.2(62), 135.2(51); MS3[397.4→243.3]: 215.3(24), 187.1(70), 173.2(29), 161.3(58), 149.3(39), 147.2(100), 135.2(49), 133.3(34)	
15	397.3829	16.85	[M+H- H ₂ O] ⁺	C ₂₉ H ₅₀ O	0.055	MS2[397.4→243.3]: 215.3(24), 187.1(70), 173.2(29), 161.3(58), 149.3(39), 147.2(100), 135.2(49), 133.3(34) MS2[716.5]: 575.5(100) ; MS3[716.5→575.5]: 501.3(24), 319.4(38), 263.3(100), 245.2(29), 239.2(36)	β-sitosterol
16	716.5215	17.17	[M+H] ⁺	C ₃₉ H ₇₄ O ₈ NP	1.369	MS2[716.5]: 575.5(100) ; MS3[716.5→575.5]: 501.3(24), 319.4(38), 263.3(100), 245.2(29), 239.2(36)	PE(16:0/18:2)
	575.5026	17.17	[M+H-141] ⁺	C ₃₇ H ₆₇ O ₄ ⁺	1.368	MS2[575.5]: 501.6(34), 319.3(67), 313.4(26), 263.4(100), 245.2(32), 239.2(40)	
17	742.5371	17.69	[M+H] ⁺	C ₄₁ H ₇₆ O ₈ NP	1.389	MS2[742.5]: 724.5(10), 710.6(6), 601.6(100)	PE(18:1/18:2)
	601.5187	17.69	[M+H-141] ⁺	C ₃₉ H ₆₉ O ₄ ⁺	0.560	MS2[601.5]: 545.4(13), 527.6(30), 339.2(11), 337.4(3), 321.2(16), 319.5(49) , 301.4(12), 265.3(100), 263.2(80), 247.3(29), 245.3(28) MS2[615.5]: 597.6(100) , 517.5(8), 503.5(6), 491.6(5), 337.4(6), 317.4(4) ; MS3[615.5→597.6]: 579.6(47), 561.5(20), 541.5(81), 523.4(100), 505.6(35), 317.4(27)	
18	615.4972	17.87	[M+H] ⁺	C ₃₉ H ₆₆ O ₅	1.790	MS2[615.5]: 597.6(100) , 517.5(8), 503.5(6), 491.6(5), 337.4(6), 317.4(4) ; MS3[615.5→597.6]: 579.6(47), 561.5(20), 541.5(81), 523.4(100), 505.6(35), 317.4(27)	DG(18:3/18:2)
	597.4876	17.87	[M+H- H ₂ O] ⁺	C ₃₉ H ₆₅ O ₄ ⁺	0.229	MS2[597.5]: 541.4(100), 523.4(63), 505.5(25), 337.4(18), 319.3(19), 317.3(34), 263.2(19), 261.3(22), 245.3(11), 243.2(16)	

Formatted: Font: Times New Roman

Peak No.	Ions m/z	RT	Adduct	Formula	Error (ppm)	MS ² →MS ⁴ data	Tentative identification	
19	461.4350	18.75	[M+H] ⁺	C ₃₁ H ₅₆ O ₂	0.666	MS2[461.4]: 111.2(100)	5- <i>n</i> -pentadecylresorcinol (AR-C25:0)	Formatted: Font: Italic
20	617.5129	19.37	[M+H] ⁺	C ₃₉ H ₆₈ O ₅	1.703	MS2[617.5]: 599.5(100) ; MS3[617.5→599.5]: 581.6(48), 563.6(27), 543.5(100), 525.5(96), 515.5(11), 507.6(52), 501.5(13), 443.5(13), 359.4(12), 345.4(14), 339.4(14) , 331.5(13), 317.4(23) , 261.4(12) , 243.4(12)	mix of DG(18:2/18:2) and DG(18:3/18:1)	
	599.5032	19.37	[M+H-H ₂ O] ⁺	C ₃₉ H ₆₇ O ₄ ⁺	0.312	MS2[599.5]: 581.6(12), 543.6(100), 525.5(55), 507.6(21), 501.4(7), 487.5(6), 473.3(6), 431.5(6), 417.4(6), 339.4(11) , 337.4(25) , 319.4(32) , 317.3(14) , 263.3(29) , 261.3(6) , 245.3(15) , 243.2(6) , MS2[593.5]: 575.7(79), 574.9(26), 501.4(5), 353.4(4), 337.4(89) , 319.3(5) , 313.3(100) , 263.3(17) , 257.2(7), 245.3(14) , 239.2(6)		
21	593.5137	20.60	[M+H] ⁺	C ₃₇ H ₆₈ O ₅	0.424	MS2[593.5]: 575.7(79), 574.9(26), 501.4(5), 353.4(4), 337.4(89) , 319.3(5) , 313.3(100) , 263.3(17) , 257.2(7), 245.3(14) , 239.2(6)	DG(16:0/18:2)	
	575.5032	20.60	[M+H-H ₂ O] ⁺	C ₃₇ H ₆₇ O ₄ ⁺	0.325	MS2[575.5]: 501.5(11), 319.4(60) , 313.3(100) , 263.3(51) , 245.4(31) , 239.1(13) , 221.3(5) , 175.3(11)		
22	619.5292	21.17	[M+H] ⁺	C ₃₉ H ₇₀ O ₅	0.648	MS2[619.5]: 601.5(100) , 545.5(22), 527.5(37), 339.4(4) , 337.3(5) ; MS3[619.5→601.5]: 583.6(19), 545.4(47), 527.5(100), 317.4(4)	DG(18:1/18:2)	Formatted: Font: Times New Roman
	601.5189	21.17	[M+H-H ₂ O] ⁺	C ₃₉ H ₆₉ O ₄ ⁺	0.228	MS2[601.5]: 583.6(14), 545.4(100), 527.5(71), 509.5(31), 339.5(28) , 319.4(29) , 265.3(26) , 263.4(25) , 247.3(13) , 245.4(19)		
23	551.5030	21.90	[M+H-H ₂ O] ⁺	C ₃₅ H ₆₇ O ₄ ⁺	0.702	MS2[551.5]: 313.2(4) , 297.3(3), 295.4(5) , 257.4(3), 239.1(100) , 221.1(12)	DG(16:0/16:0)	
24	595.5290	22.49	[M+H] ⁺	C ₃₇ H ₇₀ O ₅	1.010	MS2[595.5]: 577.6(100) , 339.4(100) , 313.4(100) , 283.3(12), 265.4(25) , 257.4(17), 247.3(11) , 239.2(14)	DG(16:0/18:1)	Formatted: Font: Times New Roman
	577.5183	22.49	[M+H-H ₂ O] ⁺	C ₃₇ H ₆₉ O ₄ ⁺	1.276	MS2[577.5]: 503.5(8), 321.4(8) , 313.3(28) , 303.4(4), 265.4(100) , 247.4(39) , 239.2(27) , 221.2(4)		

Peak No.	Ions m/z	RT	Adduct	Formula	Error (ppm)	MS ² →MS ⁴ data	Tentative identification
25	621.5449	23.10	[M+H] ⁺	C ₃₉ H ₇₂ O ₅	0.566	MS2[621.5]: 603.7(100) , 547.5(12), 529.5(39), 511.6(8), 339.3(74) , 321.2(4) , 265.4(22) , 247.3(10)	DG(18:1/18:1)
	603.5344	23.10	[M+H-H ₂ O] ⁺	C ₃₉ H ₇₁ O ₄ ⁺	0.476	MS2[603.5]: 529.6(19), 511.6(6), 339.3(24) , 321.4(12) , 265.3(100) , 247.3(31) ; MS3[603.5→ 265.3]: 247.4(100)	
26	621.5449	23.47	[M+H] ⁺	C ₃₉ H ₇₂ O ₅	0.566	NA	DG(18:0/18:2)
	603.5344	23.47	[M+H-H ₂ O] ⁺	C ₃₉ H ₇₁ O ₄ ⁺	0.476	MS2[603.5]: 529.6(15), 341.4(100) , 319.3(59) , 263.4(43) , 245.2(25)	

^a Characteristic ions for identification are in bold.

^b compared with spectra in METLIN database (METLIN ID 85041 and 64926)

Figures

Fig. 1

Fig. 2

Fig. 3

Fig. 4

Fig. 5

Original research article

Differentiation of bread made with whole grain and refined wheat (*T. aestivum*) flour using LC/MS-based chromatographic fingerprinting and chemometric approaches

Ping Geng, James M. Harnly, and Pei Chen*

Food Composition and Methods Development Laboratory, Beltsville Human Nutrition Research Center, Agricultural Research Service, U.S. Department of Agriculture, Beltsville, Maryland, 20705, United States

*Correspondence to: Pei Chen, Food Composition and Methods Development Laboratory, Beltsville Human Nutrition Research Center, Agricultural Research Service, USDA, Building 161, BARC-East, 10300 Baltimore Avenue, Beltsville, MD 20705, USA.

E-mail: pei.chen@ars.usda.gov; Phone: 1(301)-504-8144; Fax: 1(301)-504-8314

Supplementary Figures (pp. 2–7)

- Fig. S1** MSⁿ spectra of free FA Linolenic acid (C18:3) in APCI positive mode
- Fig. S2** MS² and MS³ spectra of 5-n-nonadecylresorcinol (AR-C19:0) in APCI positive mode
- Fig. S3** MS² spectra for campesterol (A), and β -sitosterol (B) in APCI positive mode
- Fig. S4** MS² and MS⁴ spectra of diglyceride DG(18:2/18:2) in APCI positive mode
- Fig. S5** MS² and MS³ spectra of phosphatidylethanolamine in APCI positive mode
- Fig. S6** Typical FCMS total ion chromatogram of bread samples

Fig. S1. MSⁿ spectra of free FA Linolenic acid (C18:3) in APCI positive mode

Fig. S2. MS² and MS³ spectra of 5-n-nonadecylresorcinol (AR-C19:0) in APCI positive mode

Fig. S3. MS² spectra for campesterol (A), and β -sitosterol (B) in APCI positive mode

Fig. S4. MS² and MS⁴ spectra of diglyceride DG(18:2/18:2) in APCI positive mode

6

Fig. S6. Typical FCMS total ion chromatogram of bread samples.

Fatty acids (FAs), alk(en)ylresorcinols (ARs), phospholipids, phytosterols, and diglycerides (DGs) were eluted before 3.0 min, triglycerides (TGs) were eluted after 3.0 min.

Research Highlights

- Fuzzy chromatography mass spectrometric (FCMS) fingerprinting method was developed
- PCA and SIMCA based on FCMS fingerprints differentiate whole/refined wheat bread
- Alk(en)ylresorcinols discriminate whole grain and refined bread