

Discovering Barley Intake Biomarkers in Urine by UPLC-MS Based Untargeted Metabolomics

Tu Hu *

November 2018

Abstract

Barley is one of the most produced cereal grains. Food industry and customers recently showed increasing interest to use barley as healthy food ingredients. However, its health beneficial effects has not been clarified yet due to lack of subjective measurement of food exposure. An UPLC-MS based untargeted metabolomics study was performed to discover biomarkers of barley intake in urine from a randomized cross-over study in 14 healthy volunteers having consumed whole grain barley bread and whole grain wheat bread. 6 candidate markers were proposed and putatively identified: two of them were reported with molecular weight and possible formulas. Other two were putatively identified as glucuronate conjugates. One was putatively identified as polyphenol or phenol metabolite. The rest one was identified as phytosterol metabolite. Taken together, promising candidate biomarkers of barley intake appear to be phytochemicals and their metabolites. For cereal products, unique phytochemical metabolome pattern could indicate their intakes.

*supervisor: Gözde Gürdeniz

Contents

1 Overview	6
2 Background	6
2.1 Research Questions and Study Design	6
2.2 CVD Risk Factors and Other Health Status Indicators	7
2.3 Conclusions	7
3 Introduction	7
3.1 Barley	7
3.1.1 Barley in the farm	7
3.1.2 Barley on the table	7
3.1.3 Health Benefits of Whole Grain Barley	8
3.1.4 Renewd Interest of Barley in Food Industry	8
3.2 Whole Grain Cereal Intake Biomarkers	9
3.2.1 Biomarkers of Food Intake (BFIs)	9
3.2.2 Alkylresorcinols	9
3.3 Novel Dietary Biomarkers Identification by Untargetted Metabolomics	10
3.3.1 Metabolome Profiling Techniques	10
3.3.2 Metabolomics data analysis	11
3.3.3 Compound Identification and Validation	11
4 Materials and Methods	13
4.1 Chemicals	13
4.2 Apparatus	13
4.3 Whole Grain Barley	13
4.4 LC-MS Analysis of Urine samples and Whole Grain Barley . . .	13
4.5 Data Preprocessing	14
4.6 Data Analysis and Variable Selections	14
4.7 Tandem Mass Spectrometry (MS/MS) Experiment and Identifi- cation	15
5 Results	15
5.1 LC-MS Analysis of Whole Grain Barley	15
5.2 Data-preprocess	15
5.3 PCA Modeling	16
5.4 PLSDA Modeling and Variable Selections	16
5.5 Tandem Mass Spectrometry (MS/MS) and Identification of Po- tential Biomarkers	18
5.5.1 m/z 231.0870	18
5.5.2 m/z 775.3401	18
5.5.3 m/z 216.9995 - phenol or polyphenol metabolite	18
5.5.4 m/z 517.3030 and 501.3080 - glucuronide conjugates . .	20
5.5.5 m/z 291.2683 - sterol derivative	21

6	Conclusions and Discussions	24
7	Perspectives	25
8	Acknowledgements	26
9	Appendix	27
9.1	Chromatogram of Barley Flour	27
9.2	Tandem Mass Spectrometry (MS/MS) spectra of phytosterols .	28
	References	29

Acronyms

AB After Barley. 14, 15, 17

AW After Wheat. 14, 15

BB Before Barley. 14, 17

BFIs Biomarkers of Food Intake. 2, 8, 10, 11

BW Before Wheat. 14, 17

CRP C-Reactive Protein. 6

CVD cardiovascular diseases. 2, 5–8

GC-MS Gas Chromatography-Mass Spectrometry. 9

GNPS Global Nature Products Social Molecular Networking. 11

HDL-C High Density Lipoprotein Cholesterol. 6

HMDB Human Metabolome Database. 10

LC-MS Liquid Chromatography-Mass Spectrometry. 2, 9, 12–14, 23

LDL-C Low Density Lipoprotein Cholesterol. 6

MS/MS Tandem Mass Spectrometry. 2–4, 11, 14, 17–20, 23, 27

NMR Nuclear Magnetic Resonance Spectroscopy. 9

PCA Principle Component Analysis. 2, 13, 15, 23

PLSDA Partial Least Squares Discriminant Analysis. 2, 13, 15, 23

PQN Probabilistic Quotient Normalization. 13

TC Total Cholesterol. 6

TG Triglycerides. 6

TIMS-PASEF Trapped Ion Mobility Spectrometry with Parallel Accumulation
Serial Fragmentation. 10

VIP variable importance in projection. 14, 16

WBB Whole Barley Bread. 5

WWB Whole Wheat Bread. 5

List of Figures

1	Schema of Study Design	6
2	PCA score plot (AW= After Wheat, AB= After Barley, POOL= pooled samples)	16
3	Variables Selected by PLSDA modeling but not classified as potential barley intake biomarkers (AW= After Wheat, AB= After Barley, POOL= pooled samples)	17
4	Intensity differences of potential barley intake biomarkers in different intervention time-points (AW= After Wheat, AB= After Barley, POOL= pooled samples)	17
5	Possible formulas of ion m/z 231.0870	18
6	Possible structures of ion m/z 216.9995	19
7	MS/MS spectra of ion(m/z =501 and 517) in urine sample, collision energy=42 eV	20
8	Secondary fragmentation of the glucuronyl moiety in negative-ion MS/MS spectra of glucuronide conjugates.	20
9	Chromatograms of ion(m/z =291.2683) in bran powder and endosperm powder. Top (bran); bottom (endosperm).	22
10	MS/MS spectra of ion (m/z =291.2683) in urine and barley sample with different collision energies	22
11	MS/MS spectra of ion(m/z =291.2683) in barley sample (Collision energy = 24 eV)	23
12	Major sterols in barley)	23
13	Putatively identified structure	24
14	Identification summary	24
15	Chromatogram of Barley Flour (from top to the bottom: bran flour and endosperm flour)	27
16	MS/MS spectra of common phytosterols	28

1 Overview

This report summarized the study results of 'Investigation of Barley Intake Biomarker' (Barley) project. 'Barley' project was conducted by Tu Hu and supervised by Gözde Gürdeniz.

'Barley' project started from September, 2018 and was registered as 'Project outside Course Scope MSc' with 15 ECTS equivalent to 412 hours workload. The exam should be a written report submitted on 6th, Nov followed by an oral defense on 9th, Nov. The examiners will be Henrik Roager and Gözde Gürdeniz.

2 Background

This project was based on a previous master thesis 'The Effect of Whole Barley Bread on the Risk of Cardiovascular Disease'[1]. Urine samples were collected in the previous research. The relevant contents of previous study[1] were summarized as the Background in this report.

2.1 Research Questions and Study Design

A double-blind randomized cross-over intervention was conducted. The research aimed at investigating the correlation between whole grain barley bread intake and cardiovascular diseases (CVD) risk factors.

Figure 1 (adapted from [1]) shows the schema of study design. Each intervention period lasted three weeks and was separated by a 2-week wash-out period. During intervention period, the subjects consumed 2 custom-made rolls of Whole Barley Bread (WBB) or Whole Wheat Bread (WWB) per day while maintaining a habitual diet. 14 healthy volunteer (6 men and 8 women) participated.

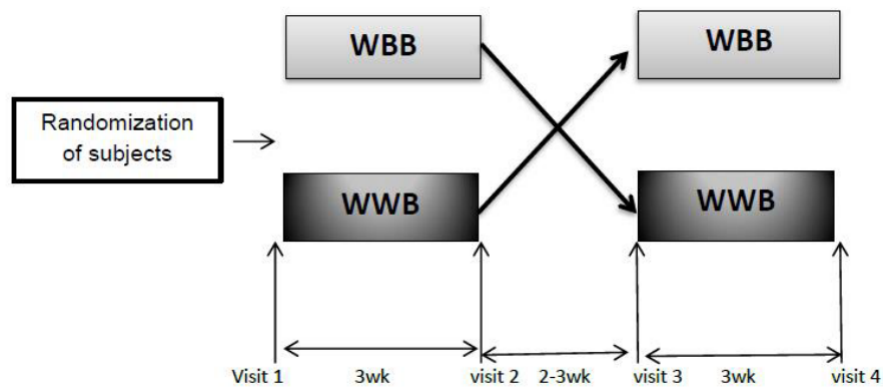


Figure 1: Schema of Study Design

2.2 CVD Risk Factors and Other Health Status Indicators

Total Cholesterol (TC), Low Density Lipoprotein Cholesterol (LDL-C), High Density Lipoprotein Cholesterol (HDL-C), Triglycerides (TG), C-Reactive Protein (CRP), fasting glucose and fasting insulin were measured both before and after intervention.

Besides CVD risk factors, body weight and waist circumference were also measured. Urine and blood samples for this research were also collected during each visit.

2.3 Conclusions

No significant effect of treatment was detected for TC ($p = 0.31$), Triglycerides ($p = 0.11$), CRP ($p = 0.62$), fasting insulin and glucose ($p = 0.78$ and 0.51 , respectively), HDL-C and LDL-C ($p = 0.32$ and 0.25 , respectively). This could be attributed to small sample size and short intervention period. In addition, the storage and food processing of bread could also affect the results[1].

3 Introduction

3.1 Barley

3.1.1 Barley in the farm

Historically, barley (*Hordeum vulgare*) has been an important cereal grain. It was among the first domesticated plants during the hundreds or thousands of years. It played an important role for human's transition from a hunting and gathering to agrarian lifestyle[2, 3].

Nowadays, barley is one of the most produced grains in the farm. In 2016, 143 million tonnes barley were produced, ranked as 4th most produced grains, behind maize, rice and wheat. European Union countries produced 63% barley of the world. Denmark produced 3.9 million tonnes barley in 2016 which was ranked as 11th most produced country of barley[4].

From the perspective of plant science, barley is arguably the most widely adapted cereal grain species with good drought, cold, and salt tolerance[5]. It is cultivated both in highly productive agricultural systems and also in marginal and subsistence environments[6]. Barley's ability to adapt to multiple biotic and abiotic stresses[6] will be crucial to its future exploitation to deal with global food security issue.

3.1.2 Barley on the table

The status of barley on the table is incompatible with its large production in the farm. Barley for direct food use only remains important in few areas, e.g. Asia and northern Africa. Barley has been majorly used as a feed, malting and

brewing grain[2, 6]. About two-thirds were used for feed, one-third for malting and only about 2% of total production were used for food directly[2].

Wheat and rice surpassed barley on the table, partially because wheat and rice products give a better mouth-feel and quality[2]. Other reasons include relatively few efforts or attempts have been devoted to systematically breed and develop barley varieties for food uses, as well as food processing technique and product development[2]. By far, the quality standard of barley for food use has not been well established, making it difficult for food manufacturers to select raw materials suitable for use in specific food products[2].

3.1.3 Health Benefits of Whole Grain Barley

Currently, research results of whole grain barley's health beneficial effects are still controversial.

Some positive research results attributed barley's health benefit effects to its beta-glucan contents or rich phytochemicals. Barley contains 4.6% beta-glucan by dry weight, while wheat, rye and oats contains 0.8%, 1.5% and 5.0% respectively[7]. Cereal beta-glucan was claimed to be capable of reducing blood serum cholesterol and regulating blood glucose levels[7]. Some epidemiological studies also showed whole-grain cereals can protect against obesity, diabetes, CVD and cancers[8]. Besides fibers of cereal, new hypothesis was also proposed that polyphenols, sulfur compounds, lignin and phytic acid, their antioxidant activities also contribute to barley's health benefit effects[9].

However, research results from University of Copenhagen showed that although beta-glucan of barley can induce satiety and reduce energy intake[10], it did not affect cholesterol metabolism[11]. Barley's health beneficial effects and associated metabolic pathways still need further research results to achieve a clear conclusion.

A major problem underlying estimating barley's health benefit effects is lacking an objective measurement of its exposure. It will be further discussed in later text.

3.1.4 Renewd Interest of Barley in Food Industry

In the past decades, barley has attracted a lot of interest of food producers and consumers to be used as a healthy food ingredient[2], especially in baking industry[12].

Renewed interest of barley for food uses largely centres around the effects of beta-glucans on lowering blood cholesterol levels and glycemic index. In addition, whole grain barley foods were considered to be capable of increasing satiety, reducing energy intake and weight loss[2].

There is a great potential for food industry to develop barley-based healthy food as a substitute for popularly used cereal grains such as wheat, oat, rice, and maize[2].

3.2 Whole Grain Cereal Intake Biomarkers

3.2.1 Biomarkers of Food Intake (BFIs)

The research interests and tasks of modern nutrition science have shifted gradually these years [13]. Especially in developed or newly-developed countries, e.g. Denmark and China, nutrition associated diseases trends from essential nutrients deficiency disease to chronic diseases, e.g. obesity, diabetes, CVD and cancer. The research objects also changed from 7 types of essential nutrients to a large group of non-essential nutrients. It attracts increasing interest to research non-essential nutrients' health benefit effects, chronic disease prevention effects, their metabolic and signaling pathways etc.

This trend in nutrition science research also raised a question, 'how to accurately, objectively and economically measure the intake of these non-essential nutrients?' By far, the application of Biomarkers of Food Intake (BFIs) has been proven that it can provide a more objective method to measure dietary exposures with a high precision and detail[13].

Dietary exposure has traditionally been measured by self-reported methods, e.g. 24-hour dietary recalls or food-frequency questionnaires[14]. However, these methods consist a lot of subjective factors, such as recall bias and difficulty in assessing portion sizes[13]. Additionally, such as cigarettes and alcohol consumption could also be biased by self-reporting methods because of cultural and societal attitudes towards these consumptions. Hence, these self-reporting methods intrinsically consist bias and errors.

The wrongly-estimated dietary exposure will further affect the study results of correlations between dietary exposures and disease frequencies, drawing controversial or even wrong conclusions[13].

3.2.2 Alkylresorcinols

Alkylresorcinols and their metabolites can be used as whole grain cereal intake biomarkers. It was reported and validated in around a dozen of researches both in intervention study and population study, shown in the table. Alkylresorcinols and their metabolites can be detected both in urinary and plasma samples.

However, vast majority of these studies mentioned above were focused on whole grain wheat or rye, or taking into account of whole grain cereals group. By far, there exists no biomarker specifically to indicate barley intakes.

No	Authors	Experimental methods	Food types	Compounds	Subjects	Matrix	Reference
1	Wierzbicka, R etc.	three-day weighed food record	Whole grain cereals	alkylresorcinol metabolites	69 Swedish	urine	[15]
2	Zhu, YD etc.	Diet Intervention	Whole grain wheat	alkylresorcinol metabolites, benzoxazinoid derivatives, phenolic acid derivatives	12 healthy participants	urine	[16]
3	Garcia-Aloy, M etc.	Self-reported food frequency questionnaires	whole grain bread	phytochemicals (benzoxazinoids, alkylresorcinol metabolites)	155 subjects	urine	[17]
4	Magnusdottir, OK etc.	controlled diet	whole grain rye	alkylresorcinol C17:0/C21:0 ratio	93 metabolic syndrome patients in Nordic countries	plasma	[18]
5	Lappi, J etc.	Diet Intervention	whole grain and fibre riched rye bread	alkylresorcinol		plasma	[19]
6	Ma, JT etc.	Self-reported food frequency questionnaires	whole grain cereals	alkylresorcinol	407 elders	plasma	[20]
7	Ross, AB etc.	Diet Intervention	whole grain food (including wheat, oats, brown basmati rice, corn, rice, barley)	alkylresorcinol	316 overweight and obese participants	plasma	[21]
8	Andersson, A etc.	Food records	whole grain wheat and rye	alkylresorcinol	72 Swedish adults	nonfasting and fasting plasma	[22]
9	Landberg, R etc.	semi-quantitative food frequency questionnaires	rye bread	alkylresorcinol	360 post-menopausal women	plasma	[23]
10	Montonen, J. etc.	Self-reported food frequency questionnaires	Whole grain food	alkylresorcinol	100 healthy adults	plasma	[24]
11	Guyman, LA etc.	three-day food record and food frequency questionnaires	Whole grain food	3-(3,5-dihydroxyphenyl)-1-propanoic acid		urine	[25]
12	Landberg, R etc.	Diet Intervention	whole grain wheat and rye	alkylresorcinol	22 women and 8 men	plasma	[26]

3.3 Novel Dietary Biomarkers Identification by Untargetted Metabolomics

3.3.1 Metabolome Profiling Techniques

With the development of high-throughout analytical technologies, 3 techniques has been majorly used in untargetted metabolomics study to identify novel dietary biomarkers. They are Liquid Chromatography-Mass Spectrometry (LC-MS), Gas Chromatography-Mass Spectrometry (GC-MS) and Nuclear Magnetic Resonance Spectroscopy (NMR) [13].

The detailed comparisons of these techniques are out of this project scope. But, within these 3 techniques, LC-MS has an outstanding role and is briefly described:

It has been proven with the advantages of:

(1) high sensitivities: the low detection limit allows tiny amount of metabolites detection.

(2) wide coverages of analytes: abundant commercially available columns, different mobile phase combination, gradient elution and temperature etc. can

be optimized to adapt to different analytical tasks.

However, compared its omic-cousins, metabolomics is still not a global technique. Not all metabolites can be profiled in a single analysis.

3.3.2 Metabolomics data analysis

Metabolomics data analysis generally consists data preprocessing, data alignment, data normalization and signal correction followed by the analysis through various statistical methods [13]

3.3.3 Compound Identification and Validation

Compound identification has been recognized as the most difficult aspect hindering the discovery of novel biomarkers. Not only limited to BFLs, compound identification troubles almost all Metabolomics researches, including disease diagnose biomarkers etc.

In the 1st Copenhagen Clinical Metabolomics Conference held in Hellerup, Denmark on 25-26 Oct 2018, experts from diverse Metabolomics research areas proposed several insightful suggestions that currently or in the future could impact the methods for compound identifications. These orally presented proposals were summarized as following:

- **Improvement of analytical instruments for metabolome profiling:** as proposed by Karolina¹, the signals of biomarkers could be very low and the detection can easily be interfered by noise or other compounds in metabolome profiling.

The development of novel analytical instruments could increase analytes coverage or increased resolving power which may have an impact on compound identification in the future.

For example, Trapped Ion Mobility Spectrometry with Parallel Accumulation Serial Fragmentation (TIMS-PASEF) can profile metabolome with high resolution generating a 4-dimensional dataset (ion mobility, retention time, mass to charge ratio and intensity). The unique 4-D fingerprint can make compounds identification easier.

- **Bioinformatics, database and data sharing:** as proposed by Cristina², Gabi³, David⁴, Niles⁵ and Pieter⁶, compound identification could be revolutionized by using bioinformatics such as computational chemistry to predict MS/MS spectra of metabolites. Then these data can be shared via database such as Human Metabolome Database (HMDB) [27] and be accessible by the whole community.

¹Karolina Sulek, Novo Nordisk Foundation Center for Protein Research, University of Copenhagen, Denmark

²Cristina Legido-Quigley, Steno Diabetes Center, Denmark

³Gabi Kastenmüller, Helmholtz Zentrum München, Denmark

⁴David Wishart, University of Alberta, Canada

⁵Nils Færgeman, University of Southern Denmark, Denmark

Pieter⁶ also pointed out a lot of spectra were generated but not open to the community. He proposed that metabolomics researchers should share MS/MS spectra in Global Nature Products Social Molecular Networking (GNPS) [28]. These spectra can be accessed by the whole community. In addition, a computer algorithm namely 'molecular network' can calculate the similarities between unknown compounds and annotated compounds in the database. Hence, researchers can get annotation suggestions of unknown compounds.

- **Deepening the understanding of biological aspect of metabolism:** Marta⁷ proposed that, the technical aspect of metabolomics is relatively mature, such as analytical instruments and bioinformatic tools. On the contrary, the real bottleneck limiting deeper understanding of metabolomics is its biological aspect. If we can deepen the understanding of its biological aspects (such as signaling, metabolic pathway etc.), we can better understand metabolomics itself in return. Therefore, the compound identification could be easier.

- **'Correctly' identify the identifiable compounds:** Mesut⁸ is an expert in shotgun mass spectrometry based lipidomics. He proposed a unique perspective for compound identification but could be inspiring.

He pointed out that, unlike its omic cousins, identifying all metabolites is currently impossible in metabolomics research. Therefore, it is important to 'correctly' identify those identifiable compounds.

A systematic protocol for mammalian lipidome analysis has been built up by him and his colleagues[29]. The key ideas were briefly summarized as following:

An extensive in-house database was built covering all 'biologically-feasible' lipids based on their chain length and unsaturation degree. In addition, he emphasized a strict analytical conditions. For example, the lipid extraction should always be conducted under 4 degree to avoid plastic dissolving into the organic solvents. By the way, based on a lot of preliminary experiments, he chose specific brands of solvents, plastic containers, experimental temperature etc to avoid contaminants to the maximum extent. Further, he stucked strictly to these conditions. Therefore, it can reach a correct identification as well as quantification.

Besides these proposals from experts, compound identification, especially for Biomarkers of Food Intake discovery, identification also relies on the knowledge or progress from other disciplines, such as food chemistry, plant science. A systematic literature research is suggested to 'dig' information for compound identification[30].

⁶Pieter Dorrestein, University of California San Diego, USA

⁷Marta Cascante, University of Barcelona, Spain

⁸Mesut Bilgin, Danish Cancer Society, Denmark

4 Materials and Methods

4.1 Chemicals

80% ethanol (CAS 64-17-5, 96%, diluted by MiliQ water), solvent A (miliQ water + 0.1% formic acid), solvent B (0.1% formic acid in 70% acetonitrile and 30% methanol)

4.2 Apparatus

Rotational vacuum concentrator (RVC 2-35, Martin Christ, Germany), kitchen high speed blender, ultrasonic bath (Branson 3800), vortex mixer, centrifuge.

Two sets of Liquid Chromatography-Mass Spectrometry (LC-MS) systems were used:

UPLC-QToF: An ultraperformance liquid chromatography (UPLC) system coupled to a quadrupole time-of-flight mass spectrometer (Premier QTOF, Waters Corporation, Manchester, UK) was used for urine sample analysis.

UPLC-Vion QoF: An ultraperformance liquid chromatography (UPLC) system coupled to Vion IMS QToF Ion Mobility Quadrupole Time-of-flight Mass Spectrometry, Waters, USA)

More details regarding LC-MS analytical condition can be found elsewhere[31, 32]

4.3 Whole Grain Barley

Whole grain barley cereal (1.5 kg, Aurion, Denmark) was purchased from Helsehuset, Gammel Kongevej 92, Frederiksberg. Whole grain barley was first grained by a high-speed kitchen blender and then filtered by a sieve (2 mm * 2 mm). Two fractions were achieved after sieve filtration.

The fraction that can pass through the sieve had smaller particle size (smaller than 2 mm). This fraction demonstrated light brown colour and majorly came from barley outer layer, i.e. bran part.

The other fraction had bigger particle size. This fraction demonstrated white colour, but containing sparsely-distributed light brown particles. This fraction majorly came from inner barley, i.e. endosperm part. They were further smashed by a mortar and filtered by the sieve (2 mm * 2mm).

4.4 LC-MS Analysis of Urine samples and Whole Grain Barley

Urine samples were collected 24 hours after intake of test bread. Samples were diluted 10 times by MiliQ-water then stored in -80 °C until analysis. The metabolome was profiled by UPLC-QToF, together with internal standard, metabolomics standard and pooled samples.

Whole grain barley powders were extracted by 80% ethanol twice. The ratio (v/m) between ethanol and barley was 0.015, i.e. 15 mL per gram. Each

extraction session lasted 30 min with the assistance of ultrasonic in room temperature. The supernatants were collected after 1st extraction session. Then, fresh ethanol solution were added to start 2nd extraction session. In the end of the extraction, supernatants from both sessions were combined in brown bottles. 2 mL supernatants from each bottle were further transferred to eppendorf tubes. Solvents in eppendorf tubes were evaporated by a rotational vacuum concentrator overnight.

Outer layer and inner powders were treated separately. Both of them were extracted in triplicates. The extraction was conducted in a yellow light room. All glasswares and eppendorf tubes were brown or covered by aluminium foils to protect light sensitive extracts.

200 μ L solvent A was used to reconstitute the extracts. Further they were diluted 10-fold to avoid ion suppression. The samples were centrifuged and the supernatants were transferred into vials for injection.

4.5 Data Preprocessing

Raw data was first converted into '.cdf' format by DataBridge, an LC-MS data file conversion program built-in MassLynx developed by Waters company.

Then, data was preprocessed by MZmine (v2.31) following the steps: peak detection, deisotoping, alignment and gap filling. Positive mode and negative were separately processed because of different noise level and in-source reaction.

In the end, the detected features, including information of mass to charge ratio (m/z), retention time (rt) and intensities were output as '.csv' files for further investigation.

4.6 Data Analysis and Variable Selections

Data analysis was performed in MATLAB R2018a (v9.4.0.813654). Data was analyzed by Principle Component Analysis (PCA), Partial Least Squares Discriminant Analysis (PLSDA) and cross validation in PLS_Toolbox (v8.6.2, Eigenvec-tor Research Inc).

Further, potential biomarkers were selected for identification, which will be described in later texts.

PCA analysis is a unsupervised multivariable statistical method. Multivariate analysis is most commonly used for exploratory analysis of metabolome profiling data [33]. It was often considered as starting point to analyze metabolome profiling data [13]. Because PCA can provide an objective assessment of the principal patterns in the data set (e.g. intake or non-intake)[13]. PCA can also detect outliers to control the data quality[34].

For both modes, PCA modeling used autoscale and Probabilistic Quotient Normalization (PQN) as preprocessing methods.

PLSDA is the most commonly used supervised multivariable analysis[13]. PLSDA was used to differentiate the barley and wheat intake.

Variables were selected by repeatedly removing variables with selectivity ratio and variable importance in projection (VIP) values lower than 1 until no further increase in the cross-validation classification errors could be observed. Final models with selected variables were evaluated using test set misclassification. The variables that were selected in at least 75% of the models were recorded for further investigation.

Potential barley intake biomarkers were selected from these variables based on the criteria:

- They should have low intensities in group Before Barley (BB), Before Wheat (BW), After Wheat (AW)
- They should have high intensities in intervention group After Barley (AB)

4.7 Tandem Mass Spectrometry (MS/MS) Experiment and Identification

Tandem Mass Spectrometry (MS/MS) can select specific ions of interest and fragment them. MS/MS is an important method in structure elucidation. MS/MS was performed by Vion. Collision energy of 14 eV, 28 eV, 42 eV were used to achieve different extent of fragmentation. Database HMDB, m/z cloud and Chemspider were used. MS/MS spectrum was also input into SIRIUS (version 4.0.1, build 3) to predict the structure.

5 Results

5.1 LC-MS Analysis of Whole Grain Barley

The elution patterns of barley flour chromatograms (shown in Appendix) were mostly similar. Elution time of several peaks slightly deviated, but all within 0.02 min.

However, one major difference of the chromatograms was, bran powder had higher intensities than endosperm powders for some peaks. For example, in negative mode, bran powder had double intensities for peaks with retention time (min) 3.84, 3.97, 4.43 compared with endosperm powder. The large difference could be explained by bran powder containing more 'whole grain' part. The compound in this part was more soluble in ethanol. But errors originated from the extraction was also possible.

5.2 Data-preprocess

After data-preprocessing, 1719 features were detected in positive mode; in negative mode, 3304 features were detected. Based on previous experience, it was very common to observe that more features were detected in negative mode than positive mode.

It could be explained by, either the instrument performs better in negative mode having a higher resolving power, or, some metabolites could be ionized easier in negative mode than positive mode, e.g. glucuronate conjugates.

5.3 PCA Modeling

PCA modeling results were shown in Figure 2. In score plot, After Wheat (AW) and After Barley (AB) can not be well separated.

PCA can also detect outliers and be used for data quality control[13]. In both modes, pooled samples located tightly near the center, which indicated that the data had a high quality. Two subjects were detected as outliers (outside 95% confidence level limit) in both modes. We investigated raw data and observed that these two subjects almost showed high intensities in all variables. The reason could be their urine samples were too concentrated.

This is why metabolomics research could be very complex. Although subjects were treated with the same intervention and biofluids were collected, processed and analyzed in the same way. But the metabolome could vary a lot between individuals. Unlike laboratory animals, humans are free-living individuals. Metabolome could be affected by a lot of external factors, such as genes, living environments and health statues.

In this case, concentrated urine could be simply attributed to that these individuals did not drink enough water. These two subjects were not removed in further study.

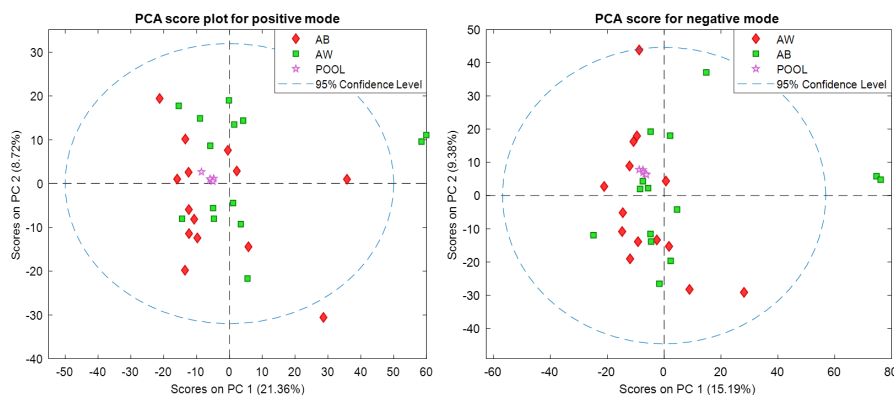


Figure 2: PCA score plot (AW= After Wheat, AB= After Barley, POOL= pooled samples)

5.4 PLSDA Modeling and Variable Selections

PLSDA modeling selected 72 variables out of 1719 features in positive mode; 86 variables out of 3304 features in negative mode. These selected variables

had high VIP value and selectivity ratio.

Although they can differentiate barley and wheat intake, not all of them were classified as potential biomarkers for barley intake.

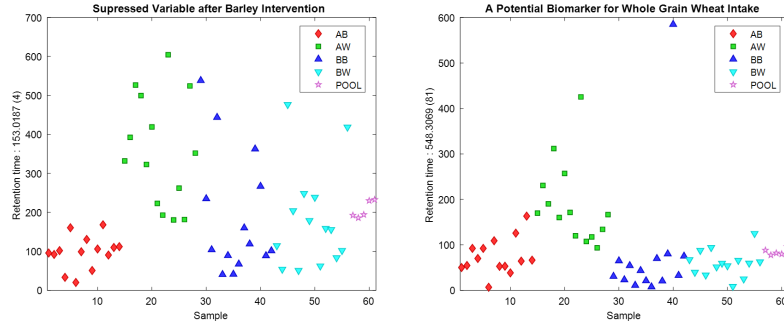


Figure 3: Variables Selected by PLSDA modeling but not classified as potential barley intake biomarkers (AW= After Wheat, AB= After Barley, POOL= pooled samples)

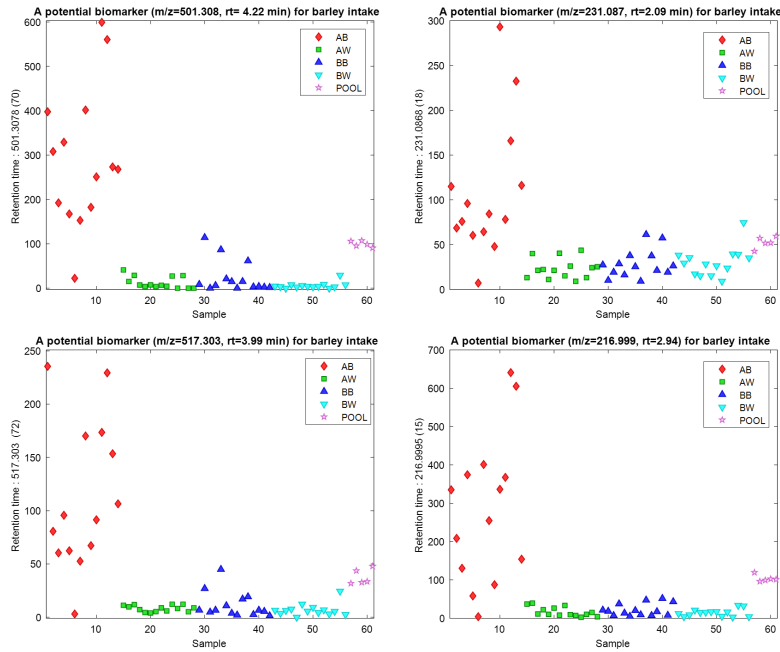


Figure 4: Intensity differences of potential barley intake biomarkers in different intervention time-points (AW= After Wheat, AB= After Barley, POOL= pooled samples)

For example, as shown in Figure 3 (left), the intensity of this variable got suppressed after barley intake. This could be an endogenous metabolite suppressed by barley intake. In Figure 3 (right), this variable could be a potential biomarker for whole grain wheat intake. Because it had lower value in AB, BB, BW but increased intensities after whole grain wheat intake.

Potential biomarkers for barley intake were further selected based on the criterials as described before to perform MS/MS analysis for identification. Within these potential biomarkers, 5 were from negative mode and 1 was from positive mode. Figure 4 showed intensities of biomarkers (part) in different time-points before and after intervention.

5.5 Tandem Mass Spectrometry (MS/MS) and Identification of Potential Biomarkers

5.5.1 m/z 231.0870

This ion was detected in negative mode. Intensity of this ion was too low to be detected in MS/MS. Moreover, there was no feasible hints in the database. Possible formulas were summarized in Figure 5.

Monoisotopic mass (Da)	Formula	Deviation (ppm)
232.0947	C ₁₀ H ₁₆ O ₆	2
232.0888	C ₁₇ H ₁₂ O	24
232.10	C ₁₆ H ₁₂ N ₂	25
232.0848	C ₁₂ H ₁₂ N ₂ O ₃	41
232.1059	C ₉ H ₁₆ N ₂ O ₅	50
232.1099	C ₁₄ H ₁₆ O ₃	68

Notes: Monoisotopic mass and formula were calculated on neutral molecular. Deviation was calculated based on deprotonated form (ESI⁻).

Figure 5: Possible formulas of ion m/z 231.0870

5.5.2 m/z 775.3401

This ion was detected in negative mode. It was annotated as the dimer of [C₁₈H₂₈O₉-H]⁻ with neutral mass 388.1735. However, the intensities of both monomer and dimer were too low to be detected in MS/MS. No feasible hints existed in the database.

5.5.3 m/z 216.9995 - phenol or polyphenol metabolite

This ion was detected in negative mode. Although this potential biomarker can

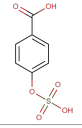
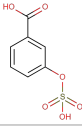
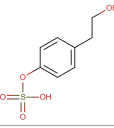
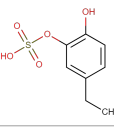
Monoisotopic mass(Da)	Formula	Name	Structure	Food Source
217.9885	$C_7H_6O_6S$	4-hydroxybenzoic acid-4-O-sulphate		Tea (metabolite of benzoic acid)
		3-hydroxybenzoic acid-3-O-sulphate		Tea (metabolite of benzoic acid)
218.0248	$C_8H_{10}O_5S$	Tyrosol 4-sulfate		Virgin olive oil (metabolite of polyphenols or phenols)
		(5-ethyl-2-hydroxyphenyl)oxidanesulfonic acid		Predicted (metabolite of polyphenols or phenols)

Figure 6: Possible structures of ion m/z 216.9995

distinguish barley and wheat intake, the intensity of this ion was too low to be detected in MS/MS. In addition, this ion was not detected in barley samples. The reason could be either it was not ionized or it was bio-transferred.

From HMDB database[27], several feasible compounds matched its mass. Figure 6 summarized the feasible structures. These compounds were all either detected or predicted metabolites of polyphenols, phenols and benzoic acids.

4-hydroxybenzoic acid-4-O-sulphate and 3-hydroxybenzoic acid-3-O-sulphate were isomers and had the same formula, $C_7H_6O_6S$. They were both detected in urine after tea intake and identified as benzoic acid metabolites[35].

Tyrosol 4-sulfate was detected in plasma after virgin olive oil intake and identified as metabolite of polyphenols or phenols[36]. (5-ethyl-2-hydroxyphenyl) oxidanesulfonic acid was a predicted metabolite of polyphenols or phenols[37].

Because whole grain barley is rich in phenols, polyphenols and benzoic acids [38]. Meanwhile, sulfidation is a common phase II metabolic reaction[39]. These compounds could also present in the urine after barley intake.

Although whole grain wheat also contains polyphenols, different concentration, structure or metabolic pathways could be the reason why barley and wheat intake can be distinguished by these compounds. This proposal needs to be further validated.

5.5.4 m/z 517.3030 and 501.3080 - glucuronide conjugates

These two potential biomarkers were detected in negative mode. Their neutral molecular formulas were differentiated by an oxygen atom, $C_{30}H_{46}O_7$ and $C_{30}H_{46}O_6$. They were putatively identified as glucuronide conjugates of barley metabolites.

MS/MS spectra (Figure 7) showed a highly-similar pattern between m/z 50-200. A characteristic lost (-176) of glucuronide conjugates[39] was detected for both ions. The un-conjugated ions (341.26754, 325.27394) were detected as well. However, un-conjugated ions were not detected in barley in both modes. Either they were not ionized, or they were bio-transformed.

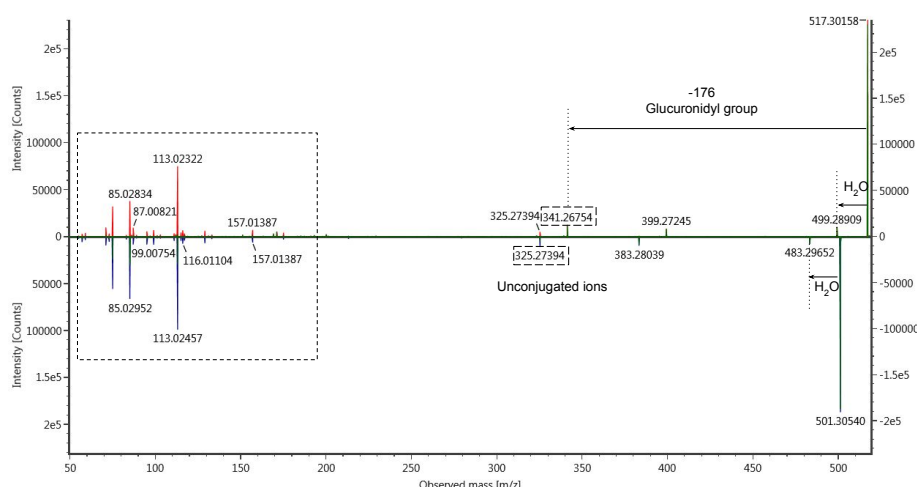


Figure 7: MS/MS spectra of ion(m/z=501 and 517) in urine sample, collision energy=42 eV

In low mass region (m/z 50-200), it was difficult to differentiate the origins of fragments were from glucuronic acid or from un-conjugated molecular. Secondary fragments of glycuronyl moiety (Figure 8, adapted from [39]) interfered the identification.

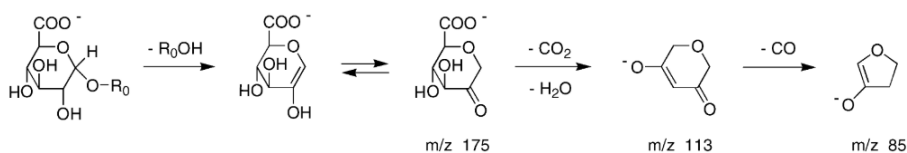


Figure 8: Secondary fragmentation of the glycuronyl moiety in negative-ion MS/MS spectra of glucuronide conjugates.

Conjugation reactions with glucuronic acid occur in animals with glucurono-

transferase after activation of the acid to uridine-5'-diphosphate glucuronic acid (UDP-5'- glucuronic acid). The glucuronic acid is bound to the corresponding molecule by a glycosidic linkage[39]. Besides glucuronide conjugates, glucosides, malonylglucosides, sulfates, acetates, methyl and glycine conjugates were all reported as common phase II metabolites of drug and food in urine[39].

Although MS/MS analysis can distinguish different types of conjugate, the exact conjugation site needs to be identified by additional 1-NMR[39], which is based on the spin state of proton nuclei.

5.5.5 m/z 291.2683 - sterol derivative

This ion ($m/z=291.2683$) was detected in positive mode both in urine and barley sample. Its deprotonated form was not detected in negative mode. The reason could be this ion was not ionized in negative mode or interfered by other ions.

This compound could originate from barley's 'whole grain' (bran) part. In barley samples, this ion showed higher intensity in bran powder (around 9000 ion counts) than in endosperm powder (around 4500 ion counts). Figure 9 showed the chromatogram.

This molecular was not bio-transformed during metabolism. Retention time (min) in urine was 6.71 and in barley sample was 6.88. Considering they were from different matrices, this small deviation was within normal range. Moreover, the MS/MS spectra (Figure 10) from both barley and urine matched with each other.

This ion should have alkyl chain or ring structure. The MS/MS spectra (Figure 11) demonstrated consecutive '-14' ($-\text{CH}_2$) lost between major fragments. Around the major fragments in the spectra, '-2' or '+2' ions exist, which could be attributed to the double bond rearrangement..

Calculating from its molecular weight, the neutral molecular had the formula ' $\text{C}_{20}\text{H}_{34}\text{O}$ '. The degree of unsaturation was 4 double-bond-equivalent. Based on knowledge of phytochemicals presented in barley[38] and MS/MS spectra of other sterols (Figure 16, in Appendix), we inferred that it could be a compound derived from sterol.

Normally, sterols had mass over 400 Da. This compound was either a unreported new compound or alkyl chain dissociated due to reasons such as storage, heat or processing.

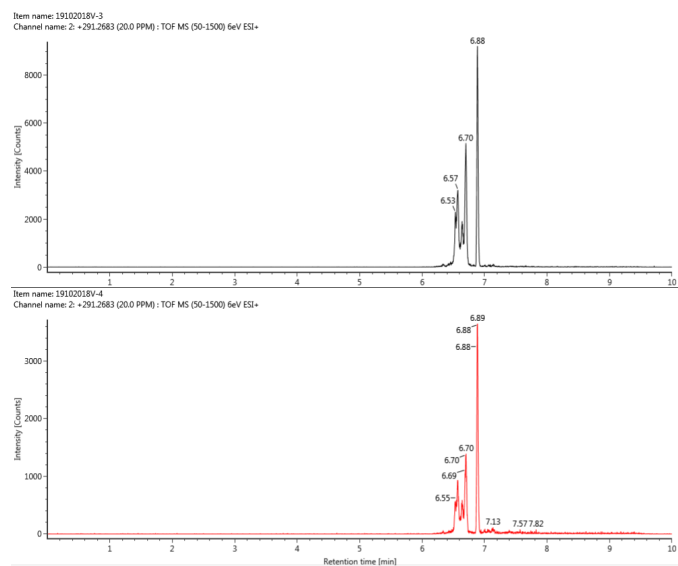


Figure 9: Chromatograms of ion($m/z=291.2683$) in bran powder and endosperm powder. Top (bran); bottom (endosperm).

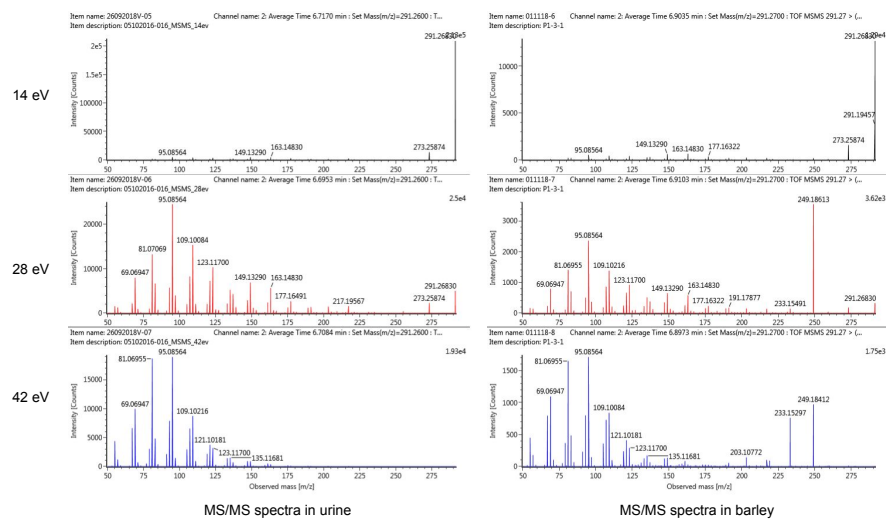


Figure 10: MS/MS spectra of ion ($m/z=291.2683$) in urine and barley sample with different collision energies

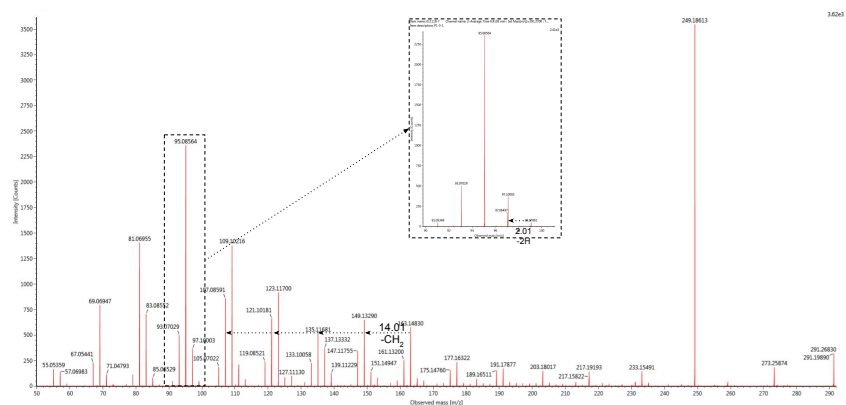


Figure 11: MS/MS spectra of ion($m/z=291.2683$) in barley sample (Collision energy = 24 eV)

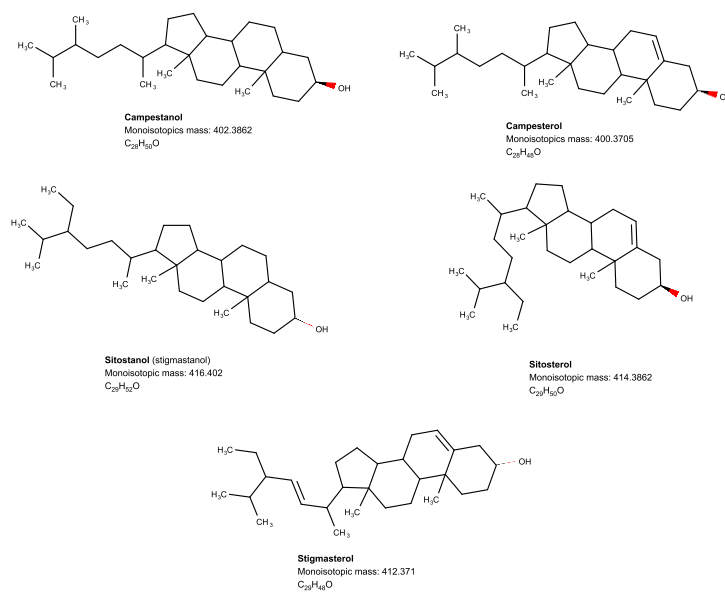
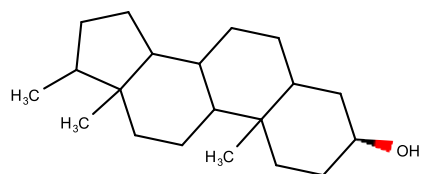


Figure 12: Major sterols in barley)

Figure 12 showed 5 major existed phytosterols in barley and were quantified[40]. These 5 major types makes up around 90% of total phytosterols[40]. Still some minor sterols has not been identified yet. We putatively identified the structure of this compound13. This proposal needs to be further validated.



Putatively identified structure
 Monoisotopics mass: 290.2609
 $C_{20}H_{34}O$

Figure 13: Putatively identified structure

6 Conclusions and Discussions

An LC-MS metabolomics study was performed to discover biomarkers for barley intake. LC-MS was used to analyze whole grain barley (bran part and endosperm part separately). Further, LC-MS was used to profile metabolome of urine from cross-over intervention. Multivariable data analysis (PCA and PLSDA) was used to control data quality and select significant variables that can distinguish barley and wheat intake.

summary of identification									
no	m/z	rt (vion qtof)	rt (qtof)	ms/ms fragments	neutral formula	annotation	mode	detected in	origins
1	517.3030	6.48 6.52	4.22	116.0096 224.0600	$C_{30}H_{46}O_7$	[Unknown1+O+glucuronate-H] ⁻	ESI-	urine	phase II metabolites (glucuronate conjugate)
	341.2675	N/A			$C_{20}H_{38}O_4$	[Unknown1+O] ⁻	ESI-	N/A	
	343.2815	N/A			$C_{20}H_{38}O_4$	[Unknown1+O+H] ⁺	ESI+	N/A	
2	501.3080	6.70	3.99	116.0110 169.1222 171.1372 325.2739	$C_{30}H_{46}O_6$	[Unknown1+glucuronate-H] ⁻	ESI-	urine	phase II metabolites (glucuronate conjugate)
	325.2739	N/A			$C_{20}H_{38}O_3$	[Unknown1-H] ⁻	ESI-	N/A	
	327.2879	N/A			$C_{20}H_{38}O_3$	[Unknown1+H] ⁻	ESI+	N/A	
3	775.3401	N/A	3.78	N/A	$C_{36}H_{56}O_{18}$	[2Unknown3 - H] ⁻	ESI-	urine	
	387.1666	N/A	3.78	N/A	$C_{18}H_{28}O_9$	[Unknown3] ⁻	ESI-	urine	
4	216.9995	5.31	2.94	N/A	$C_7H_6O_8$	[M - H] ⁺	ESI-	urine	metabolites of polyphenol/phe nol
5	231.0870	N/A	2.10	N/A	Figure 5	[M - H] ⁺	ESI-	urine	
6	291.268	6.71	4.21	69.06947 81.07079 95.08564 109.10084 123.117	$C_{20}H_{34}O$	[M + H] ⁺	ESI+	urine & barley	metabolites of sterol

Figure 14: Identification summary

Further selected potential biomarkers were identified by MS/MS and data-

base searching. In the end, we proposed 6 potential biomarkers and putatively identified them for barley intake from urine shown in Figure 14. Within these 6 markers:

- Two potential markers (m/z 231.0870, 775.3401), we can only report their molecular weight and possible molecular formulas.
- One marker was putatively identified as polyphenol or phenol metabolite.
- Two markers were putatively identified as glucuronate conjugates. Their unconjugated moleculars were not clear yet.
- One marker was putatively identified as sterol metabolite.

These markers should be further investigated and validated to confirm the structure.

These candidate biomarkers of barley intake appeared to be phytochemicals and their metabolites. Phytochemicals are chemical compounds produced by plants to help them thrive or thwart competitors, predators, or pathogens. As a term, phytochemicals is generally used to describe plant compounds that are under research with unestablished effects on health and are not scientifically defined as essential nutrients.

In whole grain cereals, alkylresorcinols, benzoxazinoids, flavonoids, lignans and phytosterols were reported[41]. From the perspective of systems biology, these phytochemicals are metabolites of plants controlled by gene expression. Therefore, different cereals, or more generally different plant-source food, should have their unique phytochemical profile.

While human consumes the plant-source food, human's metabolome from biofluids should also demonstrate unique pattern that can be tracked back to these phytochemicals. Therefore, unique metabolome pattern of phytochemical could differentiate and indicate plant-source food intakes.

However, food is very complex biological and chemical mixtures. Phytochemicals only make up its small proportion. Identification, quantification methods and databases have not been established yet. Discovery of biomarkers for these plant-based phytochemicals could be progressed with more insights into these phytochemicals, but vice versa.

7 Perspectives

Plasma samples were also collected. They can be further studied to provide hints for compound identifications.

In this study, we detected some potential wheat intake biomarkers as well. These markers can also be identified and compared with barley markers. This could get some insights about specificity of cereal intake.

8 Acknowledgements

I sincerely acknowledge Gözde Gürdeniz's helpful, patient and detailed supervision for me. She was always available for my questions.

In addition, I acknowledge Lars' coordination job and sharing his knowledge of phytosterols and all fellows from Metabolomics group and all technicians. Without their kind help, I can not successfully finalize my project. I also thank the opponent (censor) Henrik to give his critical comments to the project.

I also thank Simon's Danish knowledge and his generous help in searching for the whole grain cereals. I acknowledge Inge sharing her knowledge and experience on whole grain cereals and Nikolin sharing her knowledge on Chemistry.

In the end, I acknowledge my girlfriend Jingyao, my grandfather Mr. Wang and all other dear family members who live in the time zone 6 or 7 hours ahead of me, but cares every details of my life.

This 15 ECTS project accounts only for 12.5% of my master study. But it did open the door of a '*brave new world*', Metabolomics, to me.

'O wonder!

How many goodly unknowns are there here!

How beauteous metabolome is! O brave new world,

That has barley in't!

9 Appendix

9.1 Chromatogram of Barley Flour

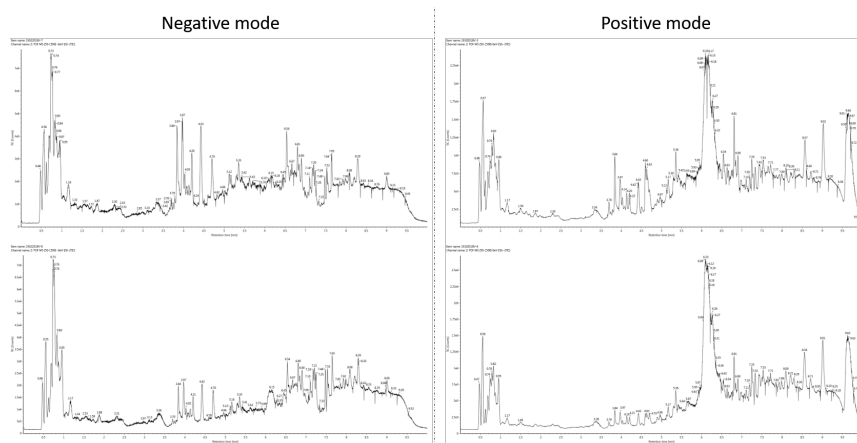


Figure 15: Chromatogram of Barley Flour (from top to the bottom: bran flour and endosperm flour)

9.2 Tandem Mass Spectrometry (MS/MS) spectra of phytosterols

MS/MS spectra of common phytosterols are shown in Figure 16, adapted from [42]

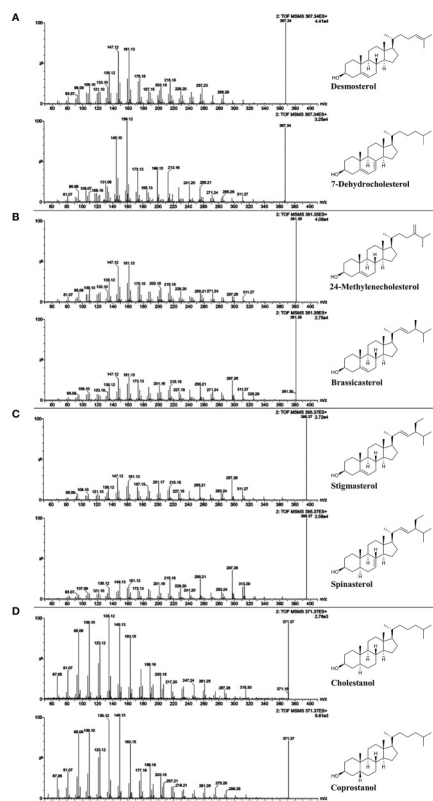


Figure 16: MS/MS spectra of common phytosterols

References

1. Popovici, V. *The Effect of Whole Barley Bread on the Risk of Cardiovascular Disease* Thesis (University of Copenhagen, 2016), 1–65.
2. Baik, B.-k. & Ullrich, S. E. Barley for food : Characteristics , improvement , and renewed interest. *Journal of Cereal Science* **48**, 233–242 (2008).
3. Ullrich, S. E. *Barley: Production, improvement, and uses* (ed Ullrich, S. E.) ISBN: 0470958626 (John Wiley & Sons, 2010).
4. FAO. *FAO Statistics* ISBN: 9789251088029 (2016).
5. Ullrich, S. E. *Barley: Production, improvement, and uses* (ed Ullrich, S. E.) ISBN: 0470958626 (John Wiley & Sons, 2010).
6. Newton, A. C. *et al.* Crops that feed the world 4. Barley: a resilient crop? Strengths and weaknesses in the context of food security. *FOOD SECURITY* **3**, 141–178. ISSN: 1876-4517 (June 2011).
7. Frølich, W., Åman, P. & Tetens, I. Whole grain foods and health—a Scandinavian perspective. *Food & nutrition research* **57**, 18503 (2013).
8. Fardet, A. New hypotheses for the health-protective mechanisms of whole-grain cereals: what is beyond fibre? *Nutrition Research Reviews* **23**, 65–134. ISSN: 0954-4224. http://www.journals.cambridge.org/abstract_S0954422410000041 (June 2010).
9. Fardet, A. New hypotheses for the health-protective mechanisms of whole-grain cereals: what is beyond fibre? *Nutrition Research Reviews* **23**, 65–134. ISSN: 0954-4224. http://www.journals.cambridge.org/abstract%7B%5C_%7DS0954422410000041 (June 2010).
10. Kristensen, M. *et al.* Oat and barley β -glucans induce satiety and reduce energy intake - a study on acute and longer-term effects. *Obesity Reviews* **12**, 61. ISSN: 1467-7881 (2011).
11. Ibrugger, S. *et al.* Extracted oat and barley beta-glucans do not affect cholesterol metabolism in young healthy adults. *eng. The Journal of nutrition* **143**, 1579–1585. ISSN: 1541-6100 (Electronic) (Oct. 2013).
12. Sullivan, P., Arendt, E. & Gallagher, E. The increasing use of barley and barley by-products in the production of healthier baked goods. *Trends in Food Science & Technology* **29**, 124–134. ISSN: 0924-2244. <http://dx.doi.org/10.1016/j.tifs.2012.10.005> (2013).
13. Scalbert, A. *et al.* The food metabolome: a window over dietary exposure. *The American Journal of Clinical Nutrition* **99**, 1286–1308. ISSN: 0002-9165. <https://academic.oup.com/ajcn/article/99/6/1286/4577352> (June 2014).
14. Rutishauser, I. H. E. Dietary intake measurements. *eng. Public health nutrition* **8**, 1100–1107. ISSN: 1368-9800 (Print) (Oct. 2005).

15. Wierzbicka, R., Zamaratskaia, G., Kamal-Eldin, A. & Landberg, R. Novel urinary alkylresorcinol metabolites as biomarkers of whole grain intake in free-living Swedish adults. *MOLECULAR NUTRITION & FOOD RESEARCH* **61**. ISSN: 1613-4125 (July 2017).
16. Zhu, Y., Wang, P., Sha, W. & Sang, S. Urinary Biomarkers of Whole Grain Wheat Intake Identified by Non-targeted and Targeted Metabolomics Approaches. *SCIENTIFIC REPORTS* **6**. ISSN: 2045-2322 (Nov. 2016).
17. Lind, M. V. *et al.* Plasma Alkylresorcinols Reflect Gluten Intake and Distinguish between Gluten-Rich and Gluten-Poor Diets in a Population at Risk of Metabolic Syndrome. *JOURNAL OF NUTRITION* **146**, 1991–1998. ISSN: 0022-3166 (Oct. 2016).
18. Magnusdottir, O. K. *et al.* Whole Grain Rye Intake, Reflected by a Biomarker, Is Associated with Favorable Blood Lipid Outcomes in Subjects with the Metabolic Syndrome - A Randomized Study. *PLOS ONE* **9**. ISSN: 1932-6203 (Oct. 2014).
19. Sawicki, C. M. *et al.* Phytochemical pharmacokinetics and bioactivity of oat and barley flour: A randomized crossover trial. *Nutrients* **8**, 1–15. ISSN: 20726643 (2016).
20. Ma, J. *et al.* Plasma Alkylresorcinols, Biomarkers of Whole-Grain Intake, Are Related to Lower BMI in Older Adults. *JOURNAL OF NUTRITION* **142**, 1859–1864. ISSN: 0022-3166 (Oct. 2012).
21. Ross, A. B. *et al.* Plasma alkylresorcinols as a biomarker of whole-grain food consumption in a large population: results from the WHOLEheart Intervention Study. *AMERICAN JOURNAL OF CLINICAL NUTRITION* **95**, 204–211. ISSN: 0002-9165 (Jan. 2012).
22. Andersson, A., Marklund, M., Diana, M. & Landberg, R. Plasma Alkylresorcinol Concentrations Correlate with Whole Grain Wheat and Rye Intake and Show Moderate Reproducibility over a 2-to 3-Month Period in Free-Living Swedish Adults. *JOURNAL OF NUTRITION* **141**, 1712–1718. ISSN: 0022-3166 (Sept. 2011).
23. Landberg, R. *Alkylresorcinols as Biomarkers of Whole Grain Wheat and Rye Intake* ISBN: 978-91-86195-58-8 (2009).
24. Montonen, J. *et al.* Reliability of fasting plasma alkylresorcinol concentrations measured 4 months apart. *EUROPEAN JOURNAL OF CLINICAL NUTRITION* **64**, 698–703. ISSN: 0954-3007 (July 2010).
25. Guyman, L. A. *et al.* Urinary 3-(3,5-dihydroxyphenyl)-1-propanoic acid, an alkylresorcinol metabolite, is a potential biomarker of whole-grain intake in a US population. *JOURNAL OF NUTRITION* **138**, 1957–1962. ISSN: 0022-3166 (Oct. 2008).
26. Landberg, R., Kamal-Eldin, A., Andersson, A., Vessby, B. & Aman, P. Alkylresorcinols as biomarkers of whole-grain wheat and rye intake: plasma concentration and intake estimated from dietary records. *AMERICAN JOURNAL OF CLINICAL NUTRITION* **87**, 832–838. ISSN: 0002-9165 (Apr. 2008).

27. Wishart, D. S. *et al.* HMDB 4.0: the human metabolome database for 2018. eng. *Nucleic acids research* **46**, D608–D617. ISSN: 1362-4962 (Electronic) (Jan. 2018).
28. Wang, M. *et al.* Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking. eng. *Nature biotechnology* **34**, 828–837. ISSN: 1546-1696 (Electronic) (Aug. 2016).
29. Nielsen, I. Ø., Maeda, K. & Bilgin, M. in *Lipidomics: Methods and Protocols* (ed Bhattacharya, S. K.) 123–139 (Springer New York, New York, NY, 2017). ISBN: 978-1-4939-6996-8. https://doi.org/10.1007/978-1-4939-6996-8_12.
30. Praticò, G. *et al.* Guidelines for Biomarker of Food Intake Reviews (BFIRev): How to conduct an extensive literature search for biomarker of food intake discovery. *Genes and Nutrition* **13**. ISSN: 18653499 (2018).
31. Jensen, M. G., Meier, S., Bech, L., Lund, E. & Dragsted, L. O. Detecting Beer Intake by Unique Metabolite Patterns (2016).
32. Barri, T., Holmer-Jensen, J., Hermansen, K. & Dragsted, L. O. Metabolic fingerprinting of high-fat plasma samples processed by centrifugation- and filtration-based protein precipitation delineates significant differences in metabolite information coverage. eng. *Analytica chimica acta* **718**, 47–57. ISSN: 1873-4324 (Electronic) (Mar. 2012).
33. Worley, B., Halouska, S. & Powers, R. Utilities for quantifying separation in PCA/PLS-DA scores plots. eng. *Analytical biochemistry* **433**, 102–104. ISSN: 1096-0309 (Electronic) (Feb. 2013).
34. Gürdeniz, G. *Data Handling Strategies in Nutritional Metabolomics: Illustrated Using Metabolic States and Trans-fat Exposures* PhD thesis (2013). ISBN: 9788776115616.
35. Van Der Hooft, J. J. J. *et al.* Structural elucidation and quantification of phenolic conjugates present in human urine after tea intake. *Analytical Chemistry* **84**, 7263–7271. ISSN: 00032700 (2012).
36. Suarez, M. *et al.* Improved method for identifying and quantifying olive oil phenolic compounds and their metabolites in human plasma by microelution solid-phase extraction plate and liquid chromatography-tandem mass spectrometry. eng. *Journal of chromatography. B, Analytical technologies in the biomedical and life sciences* **877**, 4097–4106. ISSN: 1873-376X (Electronic) (Dec. 2009).
37. DJOUMBOU FEUNANG, Y. *Cheminformatics Tools for Enabling Metabolomics* PhD thesis (University of Alberta, 2017).
38. Idehen, E., Tang, Y. & Sang, S. Bioactive phytochemicals in barley. *Journal of Food and Drug Analysis* **25**, 148–161. ISSN: 10219498. <http://dx.doi.org/10.1016/j.jfda.2016.08.002> (2017).

39. Levsen, K. *et al.* Structure elucidation of phase II metabolites by tandem mass spectrometry: an overview. *Journal of Chromatography A* **1067**, 55–72. ISSN: 00219673. <http://linkinghub.elsevier.com/retrieve/pii/S0021967304015936> (Mar. 2005).
40. Lampi, A.-M., Moreau, R. A., Piironen, V. & Hicks, K. B. Pearling barley and rye to produce phytosterol-rich fractions. *eng. Lipids* **39**, 783–787. ISSN: 0024-4201 (Print) (Aug. 2004).
41. Koistinen, V. M. & Hanhineva, K. Mass spectrometry-based analysis of whole-grain phytochemicals. *Critical Reviews in Food Science and Nutrition* **57**, 1688–1709. ISSN: 15497852. <http://dx.doi.org/10.1080/10408398.2015.1016477> (2017).
42. Muenger, L. H., Boulos, S. & Nystroem, L. UPLC-MS/MS Based Identification of Dietary Steryl Glucosides by Investigation of Corresponding Free Sterols. *FRONTIERS IN CHEMISTRY* **6**. ISSN: 2296-2646 (Aug. 2018).