Doctoral Dissertation

Multidisciplinary metabolomics approach reveals meglutol as an important bioactive metabolite in tempe, an Indonesian traditional fermented food

Marvin Nathanael Iman

June 2023

Division of Advanced Science and Biotechnology

Graduate School of Engineering,

Osaka University

博士学位論文

インドネシアの伝統的発酵食品「テンペ」に含まれるメグルトールが重要な生理活性代謝物であることを明らかにするための学際的メタボロミクスアプローチ

Marvin Nathanael Iman

2023 年 6 月

大阪大学大学院工学研究科

生物工学専攻

Table of Contents

Table of Contents 3

List of Abbreviations 6

Chapter 1: **General Introduction** 8

1.1 Functional food 8

1.2 Tempe 9

1.3 Bioactive components of tempe 10

1.4 Food metabolomics 11

1.5 Objective and strategy 12

1.6 Dissertation outline 12

Chapter 2: **GC-MS and LC-MS-based identification of bioactive metabolites in tempe, an Indonesian traditional fermented food** 15

2.1 Introduction 15

2.2 Materials and methods 17

2.2.1 Sample preparation 17

2.2.2 Metabolite extraction for gas chromatography-mass spectrometry (GC-MS) analysis 17

2.2.3 Derivatization for GC-MS analysis 18

2.2.4 GC-MS analysis 19

2.2.5 High performance liquid chromatography–photo diode array–mass spectrometry (HPLC-PDA-MS) analysis 19

2.2.6 Untargeted LC-MS analysis 20

2.2.7 Data processing and statistical analysis 20

2.3.1 Metabolites identified in soybean, germinated soybean, and soybean tempe 21
2.3.2 Bioactive metabolite identification 23
2.3.3 Metabolome profiles of soybean, germinated soybean, and soybean tempe 24
2.3.4 Soybean metabolites significantly modulated by germination and tempe fermentation 26
2.3.5 Soybean bioactive metabolite levels following tempe fermentation and germination 30
2.4 Conclusion 32
Chapter 3: **Population-based examination of plasma 3-hydroxy-3-methylglutaric acid (meglutol) and cardiometabolic risk** 33
3.1 Introduction 33
3.2 Materials and methods 35
3.2.1 Study participants 35
3.2.2 Blood sample collection and plasma metabolomics profiling 35
3.2.3 Outcome assessment 36
3.2.4 Covariate assessment 37
3.2.5 Statistical analysis 38
3.3 Results and discussion 40
3.4 Conclusion 46
Chapter 4: **Optimization of meglutol concentrations in tempe using raw material substitutions** 47
4.1 Introduction 47

4.2 Materials and methods 49
4.2.1 Legume tempe samples 49
4.2.2 Metabolite extraction for gas chromatography-mass spectrometry (GC-MS) analysis 49
4.2.3 Derivatization for GC-MS analysis 50
4.2.4 GC-MS analysis 51
4.2.5 Data processing and statistical analysis of tempe metabolite data 51
4.3 Results and discussion 52
4.4 Conclusion 60
Chapter 5: **Conclusion and future perspectives** 62
5.1 Conclusion 62
5.2 Future perspectives 63
References 64
Supplementary Materials 74

List of Abbreviations

BMI	Body mass index
BPHRS	The Boston Puerto Rican Health Study
CAGR	Compound annual growth rate
ChEBI	Chemical Entities of Biological Interest
CRP	C-reactive protein
CV%	Coefficients of variation
CVD	Cardiovascular disease
EI	Electron ionization
FFQ	Food frequency questionnaire
GABA	Gamma-aminobutyric acid
GC-MS	Gas chromatography-mass spectrometry
HDL-C	High-density lipoprotein cholesterol
HMGCR	3-hydroxy-3-methylglutaryl-CoA reductase ()
HOMA-IR	Homeostatic model assessment of insulin resistance ()
HPLC-PDA-MS	High-performance liquid chromatography-photodiode array-mass spectrometry
LC-MS	Liquid chromatography-mass spectrometry
LDL-C	Low-density lipoprotein cholesterol
METs	Metabolic equivalent score
MRM	Multiple reaction monitoring
MS	Mass spectrometry
MSTFA	N-Methyl-N-trimethylsilyltrifluoroacetamide

NMR	Nuclear magnetic resonance
PC	Principal component
PCA	Principal component analysis
PDA	Photodiode array
PTFE	Polytetrafluoroethylene
QC	Quality control
RTI	Rumah Tempe Indonesia
SD	Standard deviations
SE	Standard errors
SOALS	The San Juan Overweight Adult Longitudinal Study
T2D	Type-2 diabetes
TCA	Tricarboxylic acid
TG	Triglyceride
WC	Waist circumference

Chapter 1

General Introduction

1.1 Functional food

Functional foods are foods that provide health benefits beyond basic nutrition through the presence of physiologically active components commonly referred as bioactive compounds(Iwatani and Yamamoto 2019, Tadesse and Emire 2020). Some examples of functional foods are green tea and berries. Green tea contains catechins (e.g., epigallocatechin gallate), theanine, and caffeine, which contribute to antioxidant, anti-inflammatory, and potentially cancer-preventive effects, as well as improved brain function, metabolism, and weight management(Kochman, Jakubczyk et al. 2020). Berries, rich in anthocyanins, flavonols, and ellagic acid, exhibit antioxidant, anti-inflammatory, anticancer, and neuroprotective effects(Golovinskaia and Wang 2021).

The functional food market has witnessed remarkable growth in recent years, establishing itself as one of the most rapidly expanding segments within the food industry. The global market of functional food is expected to reach $176.7 billion by 2023, with a compound annual growth rate (CAGR) of 7.4%(Misra, Pandey et al. 2021). This surge in popularity can be attributed to the increasing consumer awareness and demand for healthier and more personalized food choices(Alongi and Anese 2021).

One key advantage of functional foods is the absence of recognized adverse side-effects commonly associated with pharmaceutical interventions. Many individuals are wary of the potential risks and complications that can arise from the long-term use of medications. The natural components present in functional foods are generally regarded as safe and well-

tolerated by the human body, making them an attractive option for those seeking alternatives to conventional drugs. This perception of safety further supports the preference for functional foods as a more natural and gentler approach to managing disease risk(Megson, Whitfield et al. 2016).

Furthermore, the use of functional foods as a means to manage disease risk aligns with the broader concept of personalized medicine. Each individual's genetic makeup and physiological characteristics contribute to their unique health needs and response to treatments. Functional foods offer the potential for personalized nutrition, where specific food components can be tailored to an individual's requirements, addressing their specific health concerns. This personalized approach allows for a targeted and potentially more effective strategy in disease prevention and management(Portanguen, Tournayre et al. 2019).

1.2 Tempe

One example of functional food is tempe. Tempe is an Indonesian traditional fermented food made from soybean, well-known for its high nutritional and functional contents. Tempe is made through a controlled fermentation of soybeans using the fungus *Rhizopus oligosporus*. The initial step involves soaking the soybeans, followed by dehulling and cooking to ensure optimal fermentation conditions. Afterward, the cooked soybeans are inoculated with the fungal spores, which initiate the fermentation process. The fermentation takes place at a specific temperature and humidity, allowing the fungus to proliferate and form a cohesive cake-like structure. This distinctive fermentation process gives tempe its characteristic texture, aroma, and flavor profile(Nout and Kiers 2005).

Tempe stands out not only for its unique taste but also for its remarkable nutritional profile. It is a rich source of high-quality protein, containing all the essential amino acids required by the human body(Rahmawati, Astawan et al. 2021). Furthermore, tempe is a notable source of dietary fiber, vitamins (such as B vitamins), and minerals (such as iron, calcium, and magnesium). Its low carbohydrate and cholesterol content also make it an excellent choice for individuals aiming to maintain a healthy diet(Ahnan‐Winarno, Cordeiro et al. 2021).

1.3 Bioactive components of tempe

In addition to its excellent nutritional composition, tempe offers numerous health benefits. Regular consumption of tempe has been associated with improved cardiovascular health, as it may help regulate cholesterol levels and reduce the risk of heart disease(Astawan, Mardhiyyah et al. 2018). The presence of beneficial bioactive compounds in tempe contributes to its potential as a functional food with antioxidant and anti-inflammatory properties. These compounds have been linked to various health benefits, including enhanced immune function(Hosaka, Itoh et al. 2021), cognitive function(Handajani, Turana et al. 2021), liver health, and reduced risk of chronic diseases(Afifah, Nabilah et al. 2020).

To date, most of the bioactive properties of tempe are attributed to tempe's isoflavone contents. Tempe is an excellent source of soy isoflavones such as daidzein and genistein(Harun, Susanto et al. 2017, Astawan, Mardhiyyah et al. 2018, Huang, Wu et al. 2018, Ahnan‐Winarno, Cordeiro et al. 2021). Isoflavones in tempe have also been found to be more bioavailable compared to isoflavones in unfermented soy products. Numerous

research studies have been conducted focusing on the isoflavones found in tempe. Notable examples include the profiling of tempe isoflavone composition and its dynamics during fermentation(Kuligowski, Pawłowska et al. 2016), as well as the investigation of the effects of different processing methods on tempe isoflavones(Mo, Kariluoto et al. 2013, Kuligowski, Sobkowiak et al. 2022).

Recent studies have also investigated the bioactive peptides found in tempe. Tamam *et al.* employed proteomics to identify bioactive peptides in tempe, identifying Val-His and Ala-Leu-Glu-Pro as significant peptides commonly found in tempe from various producers(Tamam, Syah et al. 2019). Sitanggang *et al.* conducted a comprehensive mapping of bioactive peptides in optimally prepared tempe using membrane-based techniques(Sitanggang, Lesmana et al. 2020).

While there has been extensive research conducted on the functionality of tempe, previous efforts to understand its effects have predominantly focused on the overall bioactivity of tempe, lacking specific identification of the bioactive metabolites responsible for these effects. Alternatively, certain studies have solely concentrated on a limited set of well-studied bioactive components, such as soy isoflavones or bioactive peptides, without comprehensively exploring the entire tempe metabolome. Although these studies are valuable and insightful, a comprehensive investigation into the metabolome of tempe would offer significant advantages by ensuring that potential bioactive compounds are not overlooked(Caesar, Kellogg et al. 2019).

1.4 Food metabolomics

Metabolomics is a field of study that focuses on the comprehensive analysis of small molecules, known as metabolites, present in biological systems. Metabolomics involves

the identification, quantification, and analysis of metabolites using advanced analytical techniques, such as mass spectrometry (MS) and nuclear magnetic resonance (NMR)(Putri, Nakayama et al. 2013).

Metabolomics techniques can be used to detect and identify bioactive metabolites found in foods as it allows for a comprehensive analysis on the metabolome of food. In recent years, the use of metabolome profiling to profile for bioactive compounds in foods has been gaining popularity(Yuliana, Khatib et al. 2011, Caesar, Kellogg et al. 2019). Several notable examples include the identification of bioactive prenol lipids and soyasaponins from soy roots and the identification bioactive organic acids in strawberry leaves using LC-MS(Bragagnolo, Funari et al. 2021).

1.5 Objective and strategy

This research aims to profile the bioactive components of tempe using a comprehensive metabolomics-based techniques. Furthermore, we aim to characterize the bioactive components of tempe using a nutritional epidemiological analysis and investigate the improvement of tempe’s bioactive components using a raw material substitution strategy.

1.6 Dissertation outline

This dissertation is divided into five chapters, each addressing different aspects of the bioactive components of tempe as a functional food. Chapter 1 serves as a general introduction, covering functional food, tempe, bioactive components, and the concept of food metabolomics. It emphasizes the existing research gap in the field of metabolomics profiling of tempe's bioactive components, which forms the main objective of this research.

In Chapter 2, the focus shifts to the profiling of bioactive components in mature (yellow) and immature (green) soybean (*Glycine max*), germinated mature (yellow) and immature (green) soybean, and mature (yellow) and immature (green) soybean tempe. To enable this metabolomics analysis, three analytical techniques, namely a widely-targeted GC-MS, untargeted LC-MS, and targeted HPLC-PDA-MS/MS, were employed. This chapter reports the identification of sixteen bioactive metabolites found in tempe.

Building upon the findings of Chapter 2, Chapter 3 delves into the characterization of a specific tempe bioactive metabolite called 3-hydroxy-3-methylglutaric acid (meglutol), as previously reported in Chapter 2. Epidemiological cohorts BPHRS and SOALS were utilized to examine the associations between plasma meglutol levels and LDL-C levels. The results indicate an overall association between higher plasma meglutol levels and lower LDL-C levels. Notably, metabolically compromised patients exhibited stronger associations between plasma meglutol and lower LDL-C.

Chapter 4 focuses on the improvements of meglutol levels in tempe. This chapter quantifies and reports the concentrations of meglutol in various foods, highlighting tempe as the food with the highest meglutol content. Additionally, raw material substitution was implemented to further improve meglutol concentration in tempe. As a result, green pea tempe emerged as the variety with the highest level of meglutol.

Finally, Chapter 5 offers a comprehensive conclusion to the dissertation, summarizing the key findings and their implications. Additionally, it provides a future outlook on the topics explored throughout the dissertation, highlighting potential avenues for further research and exploration in this field.

Chapter 2

GC-MS and LC- MS-based identification of bioactive metabolites in tempe, an Indonesian traditional fermented food

2.1 Introduction

Tempe, a traditional fermented food originating from Indonesia, has gained considerable attention due to its remarkable nutritional and functional qualities. This soybean-based fermented product stands out for its high protein content and low levels of carbohydrates and lipids, making it an invaluable dietary resource. Moreover, extensive research has revealed that tempe is rich in a wide range of bioactive compounds, including antioxidants and anti-inflammatory agents. Notably, the consumption of tempe has been associated with numerous health benefits, such as reducing the risk of cardiovascular diseases(Astuti, Meliala et al. 2000).

Tempe has garnered significant recognition and popularity in recent years, attributed to its exceptional nutritional and functional qualities. Despite the increasing popularity and the potential health benefits associated with tempe consumption, comprehensive research exploring its functional quality remains limited. Previous studies have primarily focused on a small number of well-known bioactive metabolites, such as soy isoflavones(Kuligowski, Pawłowska et al. 2016) or GABA(Koh, Jamaluddin et al. 2012), without conducting a comprehensive investigation of the metabolome of this fermented food. While the characterization of these metabolites in tempe is insightful and valuable, a comprehensive investigation into the metabolome of tempe would offer significant

advantages by ensuring that potential bioactive compounds are not overlooked(Caesar, Kellogg et al. 2019).

Recent research has highlighted the significance of incorporating metabolome information through the utilization of metabolomics, a highly valuable technique for the identification of bioactive metabolites(Stewart, McDougall et al. 2007, Yuliana, Khatib et al. 2011, Caesar, Kellogg et al. 2019). Particularly noteworthy is the successful application of metabolomics in profiling the bioactive constituents of underutilized soy parts, such as soy roots, leaves, and branches(Bragagnolo, Funari et al. 2021). However, the comprehensive characterization of bioactive components in tempe using metabolomics analysis remains an unexplored area of research.

This study aims to fill the current research gap by using metabolomics techniques to comprehensively investigate the metabolome of tempe and profile its bioactive metabolites. To maximize metabolite coverage and capture a wide range of bioactive metabolites with differing chemical and physical characteristics, metabolome profiling using a widely targeted GC-MS analysis, untargeted LC-MS analysis, and a targeted HPLC-PDA-MS/MS analysis are used complementarily. The levels of detected bioactive metabolites following soybean tempe fermentation are analyzed and compared to the changes observed during soybean germination as controls. The findings from this research will provide valuable insights into the functional characteristics of tempe, enriching our understanding of its bioactive properties.

2.2 Materials and methods

2.2.1 Sample preparation

Table S1 presents the sample information used in this study. The preparations of germinated legume and tempe samples were conducted at Osaka University's Laboratory of Bioresource Engineering. The method employed for germinated legumes was adapted from Puteri et al. (2018) with minor modifications(Puteri, Astawan et al. 2018). In brief, legume samples were first primed in 200 mL of water for a duration of 7 hours. Subsequently, the primed legumes were germinated in wet tissue paper for 28 hours, with regular rinsing intervals of every 4 hours. The tempe production followed the Rumah Tempe Indonesia (RTI) method previously reported(Rahmawati, Astawan et al. 2021). Specifically, 50 g of legume samples were soaked and cooked in water, after which they were dehulled and dried at room temperature. The dehulled legumes were then inoculated and incubated at 30°C for 48 hours using the Raprima starting culture.

2.2.2 Metabolite extraction for gas chromatography-mass spectrometry (GC-MS) analysis

To extract metabolites from the biological samples, a protocol involving a combination of organic solvents was employed. The samples were freeze-dried and homogenized using a multi-beads shocker (Yasui Kikai, Japan). Each sample consisted of 10 mg and was collected into 2 mL tubes. The entire process was conducted in triplicate to ensure precision and consistency of the results. A mixed solvent comprising methanol, water, and chloroform in a ratio of 5:2:2, supplemented with an internal standard of 100 µg/mL ribitol, was added to each tube. Subsequently, the tubes were incubated at 37°C for 30 minutes with agitation at 1200 rpm (Eppendorf Ltd., Germany).

After sample collection, centrifugation was performed at 4°C for 3 minutes at a speed of 10,000 rpm. The resulting supernatant, measuring 400 μL, was carefully transferred into fresh 1.5 mL tubes. Following this, 300 μL of water was added, and the tubes were centrifuged again at 4°C for 3 minutes at 10,000 rpm. Subsequently, 200 μL of the aqueous phase was transferred to fresh 1.5 mL tubes for subsequent processing. To ensure the accuracy and reliability of the obtained results, a set of quality control (QC) samples was prepared by combining 200 μL of the aqueous phase from all previously collected samples.

Before analysis, both the samples and the quality control (QC) samples underwent centrifugation using centrifugal concentrator (Taitec Co., Japan) to remove excess liquid. The centrifugation process was carried out for 1 hour at ambient temperature, followed by overnight freeze-drying.

2.2.3 Derivatization for GC-MS analysis

Freeze-dried samples were initially treated with 100 μL of methoxyamine hydrochloride. The treated samples were then incubated at 30°C for 90 minutes with continuous agitation at 1200 rpm. After the incubation period, 50 μL of N-Methyl-N-trimethylsilyltrifluoroacetamide (MSTFA) was added to each sample, followed by an additional 30 minutes of incubation at 37°C with agitation set at 1200 rpm (Eppendorf Ltd., Germany). Subsequently, the prepared samples were carefully transferred into GC vials for subsequent analysis using GC-MS.

2.2.4 GC-MS analysis

The GC-MS analysis procedure was performed using a GC-MS-QP2010 Ultra instrument (Shimadzu, Japan) equipped with an Inert-Cap MS/NS column (GL Sciences). The sample was injected into the GC using an AOC-20i/s autosampler (Shimadzu, Japan) in a 25:1 (v/v) split mode with an injection temperature of 230°C. Helium was employed as the carrier gas at a flow rate of 1.12 mL/min, corresponding to a linear velocity of 39 cm/s. The GC analysis conditions were as follows: an initial temperature of 80°C was maintained for 2 minutes, followed by a programmed increase to 330°C at a rate of 15°C/minute, with a 6-minute hold at the final temperature. Electron ionization (EI) was employed with an ion source temperature of 200°C and an *m/z* range of 85-500, scanned at a rate of 6.67 scans/s. The retention index (RI) was determined using a standard alkene mixture. The GC-MS analysis procedure was conducted at the Laboratory of Bioresource Engineering, Osaka University.

2.2.5 High performance liquid chromatography–photo diode array–mass spectrometry (HPLC-PDA-MS) analysis

To analyze soy isoflavones, a targeted metabolomics analysis using HPLC-PDA-MS instrument (Shimadzu, Japan) was employed. Metabolites were extracted from prepared samples using 70 percent ethanol. The HPLC column used in this analysis was an InertSustainSwift C18 (1.9 μm, 2.1 x 150 mm) column from GL Sciences, Japan. Electrospray ionization (ESI) was employed in positive ionization mode. The injection volume was set to 1 μL, and the column oven temperature was maintained at 45°C. The mobile phase A consisted of a 150:850:5 acetonitrile:water:acetic acid (v:v:v) solution, while mobile phase B was composed of a 800:200:5 acetonitrile:water:acetic acid (v:v:v)

solution. The total flow rate was set to 0.2 mL/minute. Initially, 0% of mobile phase B was retained for 10 minutes to initiate the analysis, followed by a gradual increase to 100% of mobile phase B over a period of 50 minutes. Subsequently, the concentration of mobile phase B was reduced to 0% in one minute and maintained for twenty minutes. For the analysis, multiple reaction monitoring (MRM) mode was employed. PDA detection was performed at a wavelength of 254 nm.

2.2.6 Untargeted LC-MS analysis

The untargeted analysis was conducted using the UHPLC Vanquish Tandem Q Exactive Plus Orbitrap HRMS (ThermoScientific) system, equipped with an Accucore C18, 100 x 2.1 mm, 1.5 µm (ThermoScientific) column. Sample extraction was performed using 70% methanol, followed by sonication and filtration using a PTFE 0.2 µm membrane. The flow rate was set at 0.2 mL/min, and the column oven was maintained at 30°C. For the mobile phases, water containing 0.1% formic acid was used as mobile phase A, while acetonitrile containing 0.1% formic acid was used as mobile phase B. Gradient elution was employed, starting with a 5% concentration of mobile phase B for the initial minute of elution. Subsequently, the concentration of mobile phase B was gradually increased to 95% over a period of 24 minutes and held for 3 minutes, after which it was decreased to 5% within 5 minutes to conclude the analysis. The mass range was set from 100 to 1500 *m/z*, and both positive and negative ionization modes were utilized to enhance the coverage of metabolites.

2.2.7 Data processing and statistical analysis

For the widely-targeted GC-MS analysis, baseline correction, peak identification, and alignment procedures from the resulting GC-MS spectrum data were carried out using

GCMSsolution (Shimadzu, Japan) and MS-DIAL 4.9 (RIKEN, Japan). Metabolite annotations were carried out by comparing RI and MS values to an in-house GC-MS-5MP Library (RIKEN, Japan). LabSolutions (Shimadzu, Japan) was used to analyze HPLC-PDA-MS/MS data, and isoflavone annotation was accomplished via co-injection with standard isoflavone compounds. Compound Discoverer (ThermoFischer) equipped with mzCloud and ChemSpider mass spectral databases was used to process the untargeted LC/MS data and putatively annotate the metabolites. MetaboAnalyst 5.0 and Origin 2022 (OriginLab Corporation, US) were utilized to conduct statistical analyses. Principal component analysis (PCA) was performed with SIMCA-P 13.0 (Umetrics, Sweden).

To predict potential health benefits associated with the compounds, pharmacological data for each annotated compound were collected from reputable sources such as PubChem (https://pubchem.ncbi.nlm.nih.gov/), DrugBank (https://go.drugbank.com/), and Chemical Entities of Biological Interest (ChEBI) (https://www.ebi.ac.uk/chebi/). These databases provide comprehensive information on the properties and activities of various compounds.

2.3 Results and discussion

2.3.1 Metabolites identified in soybean, germinated soybean, and soybean tempe

To comprehensively evaluate the metabolites found in soybean, germinated soybean, and soybean tempe, a widely-targeted GC-MS, untargeted LC-MS, and targeted HPLC-PDA-MS metabolomics analyses were carried out in this study. **Figure 2.1** and **Figure S1** provides a summary of the annotated metabolites obtained from the three different metabolomics approaches.

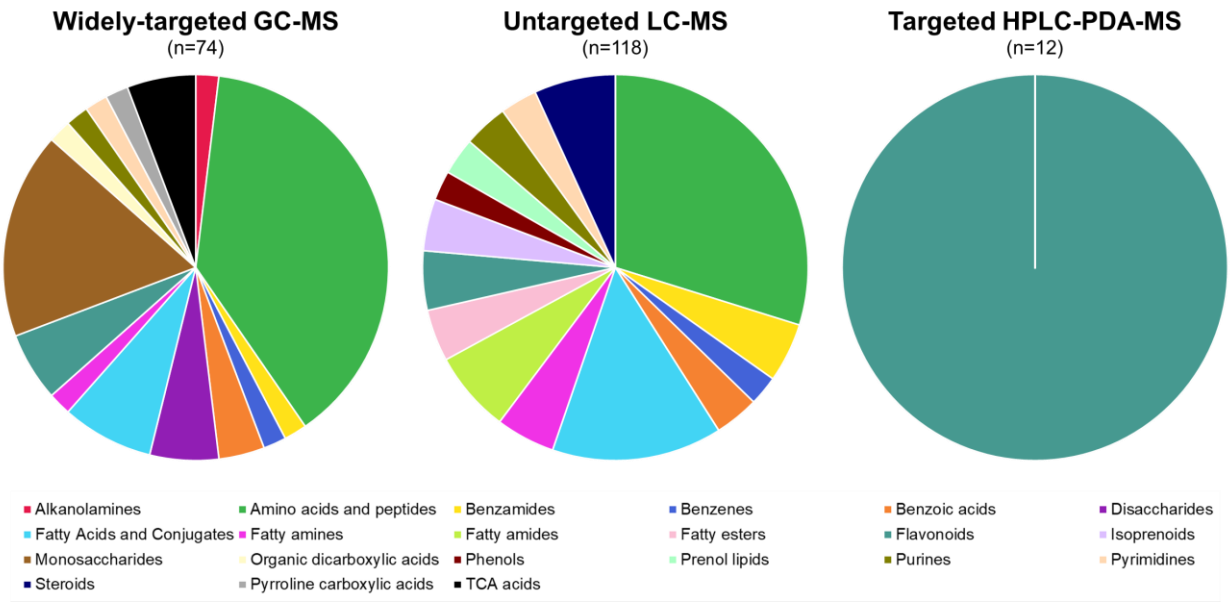


Figure 2.1. Analysis of metabolites found in soybean, germinated soybean, and soybean tempe using: widely-targeted GC-MS, untargeted LC-MS, and targeted HPLC-PDA-MS analysis.

The employed widely-targeted GC-MS analysis enables the profiling of low molecular weight primary metabolites. The reliability of metabolite annotation in this analysis is attributed to the utilization of an in-house, column-specific library. In total, 74 metabolites were annotated from the GC-MS analysis (**Table S2**), with the majority belonging to the amino acids and peptides group, sugar groups, and TCA acids group. Furthermore, the untargeted LC-MS analysis annotated 118 metabolites, predominantly from the amino acids and peptides group, as well as the fatty acids group. This wide coverage of metabolites is made possible by the untargeted nature of the analysis and the annotation method employed. The putative annotation of metabolites in the untargeted LC-MS analysis benefits from publicly available mass spectral databases. Lastly, the targeted HPLC-PDA-MS method was employed to profile the isoflavones present in the samples. Isoflavone peak assignments (**Table S3**), representative chromatogram (**Figure S2**), and

the PDA spectra (**Figure S3**) are shown in the supplementary materials This analysis successfully detected the presence of 12 soy isoflavones, including the aglycones, glucosides, acetyl glucosides, and malonyl glucosides.

2.3.2 Bioactive metabolite identification

To identify bioactive metabolites among the annotated metabolites, the pharmacological function of each annotated metabolites was predicted using publicly available pharmacological databases PubChem, DrugBank, and Chemical Entities of Biological Interest (ChEBI).

A total of 16 metabolites was predicted to have bioactive properties based on previous reports (**Table 2.1**). This list of bioactive metabolites consists of 12 well-characterized soy isoflavones daidzin, glycitin, genistin, malonyl daidzin, malonyl genistin, malonyl glycitin, acetyl daidzin, acetyl genistin, acetyl glycitin, daidzein, glycitein, genistein, as well as 3,4-dihydroxybenzoic acid, 3-hydroxyanthranillic acid, 3-hydroxy-3-methylglutaric acid (meglutol), and 4-aminobutyric acid (GABA).

Table 2.1. Identified bioactive metabolites in soybean, germinated soybean, and soybean tempe with their reported pharmacological functions.

Bioactive Metabolites	Pharmacological Function	Reference
3,4-Dihydroxybenzoic acid	Anti-inflammatory, antioxidant, anticancer	PMID: 26647619
3-Hydroxyanthranilic acid	Antioxidant	PMID: 8955104
3-hydroxy-3-methylglutaric acid (meglutol)	Hypolipidemic	PMID: 6083597
GABA	Neuronal excitability reduction	PMID: 16971751
Daidzein	Anti-inflammatory, antioxidant	PMID: 22484990
Genistein		

Glycitein		
Daidzin		
Genistin		
Glycitin		
Malonyl daidzin		
Malonyl genistin		
Malonyl glycitin		
Acetyl daidzin		
Acetyl genisin		
Acetyl glycitin		

Although the presence of the 12 soy isoflavones, the neurotransmitter 4-aminobutyric acid (GABA), and 3-hydroxyanthranillic acid in tempe have been reported and highlighted as bioactive metabolites in previous studies, to the best of our knowledge, this is the first study to highlight 3,4-dihydroxybenzoic acid and 3-hydroxy-3-methylglutaric acid (meglutol) as bioactive metabolites in tempe.

2.3.3 Metabolome profiles of soybean, germinated soybean, and soybean tempe

The modulability of the 16 identified bioactive metabolites and the overall metabolome profile modification of soybean as a result of germination and tempe fermentation were then investigated. The metabolome profiles of soybean, germinated soybean, and soybean tempe were constructed from metabolites annotated with high identification confidence levels (annotated from GC-MS and HPLC-PDA-MS/MS analyses using an in-house library or via standard compound co-injection). These metabolites were then subjected to unsupervised principal components analysis (PCA) to determine the effect of germination and tempe fermentation on the soybean metabolome, as shown in **Figure 2.2**.

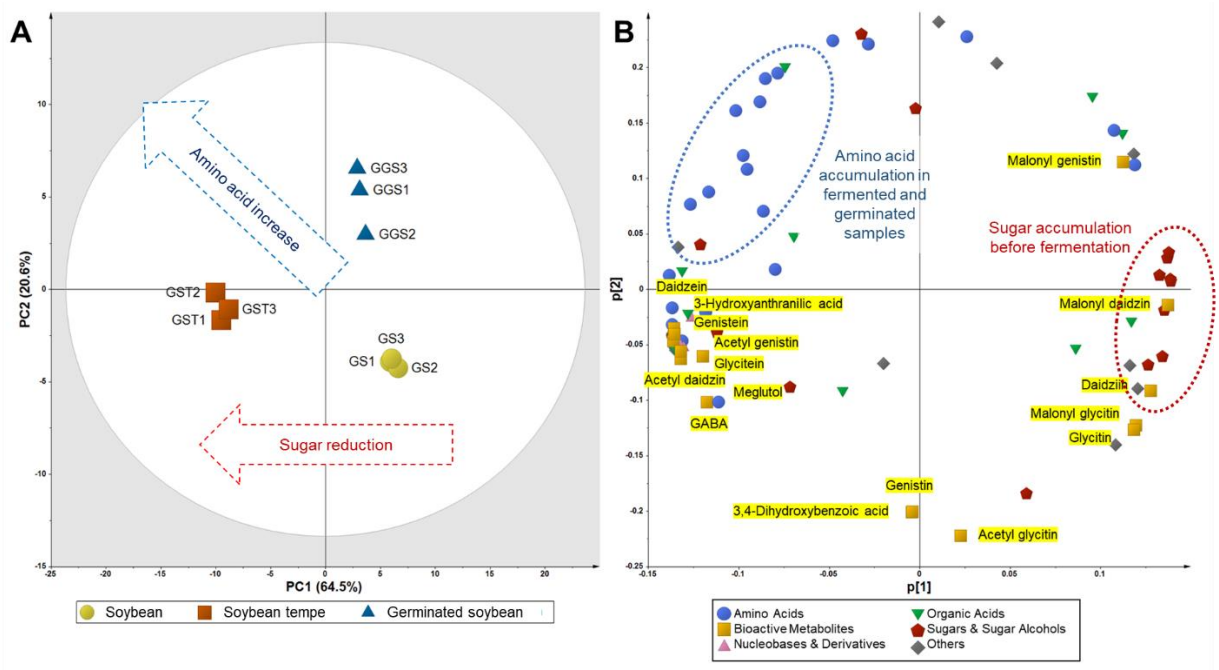


Figure 2.2. Soybean, germinated soybean, and soybean tempe metabolome profile comparison. **A)** PCA score plot. Soybean samples are indicated by circle symbols, soybean tempe samples are indicated by square symbols, germinated samples are indicated by triangle symbols. **B)** PCA loading plot. Amino acids are denoted by circle points, bioactive metabolites by square points, nucleobases and derivatives by triangle points, organic acids by inverted triangle points, sugars and sugar alcohols by pentagon points, and other metabolites by diamond points.

In the PCA score diagram presented in **Figure 2.2A**, three distinct clusters were observed: soybean cluster, germinated soybean cluster, and soybean tempe cluster. The separation of clusters based on the PC2 axis indicates a distinction between soybean samples before and after germination, while the PC1 axis separates clusters of soybean samples before and after tempe fermentation. **Figure 2.2B**, depicting the PCA loading plot, reveals that eight out of the sixteen identified bioactive metabolites cluster on the extreme negative side of PC1, suggesting an accumulation of these metabolites in soybean tempe

compared to unfermented soybean. Additionally, genistin, acetyl glycitin, and 3,4-dihydroxybenzoic acid were concentrated on the extreme negative side of the PC2, indicating an accumulation of these metabolites in non-germinated immature (green) soybean relative to germinated soybean.

The PCA plot reveals a clear trend of amino acid clustering on the positive side of PC2, as depicted in the loading plot. Moreover, the majority of these amino acids exhibit a tendency towards the negative side of PC1, positioned in the top-left region of the PCA loading plot (**Figure 2.2B**). This observed grouping suggests that both the germination and tempe fermentation processes have the potential to enhance the amino acid contents in soybean.

2.3.4 Soybean metabolites significantly modulated by germination and tempe fermentation

To investigate further the effects of tempe fermentation on soybean, metabolites significantly modulated by this process were studied and contrasted against metabolites significantly modulated by germination (**Figure 2.3**).

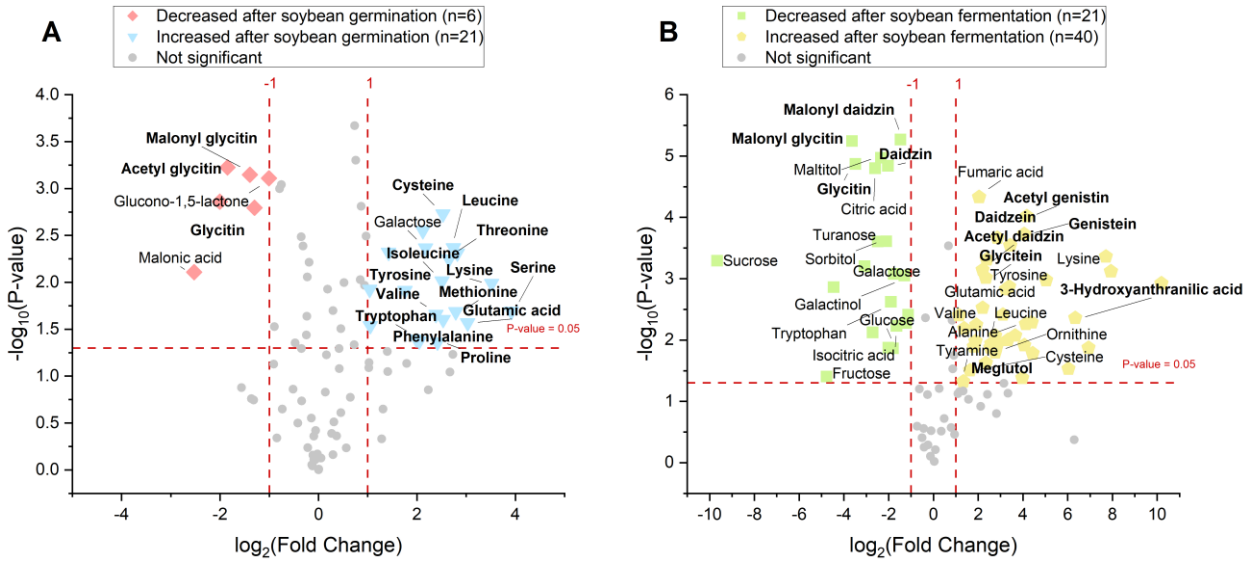


Figure 2.3. Impact of soybean processing on metabolome profiles: **A)** germination and **B)** tempe fermentation induced substantial modulation of metabolites. Metabolites exhibiting a significant increase ($p<0.05$) with a minimum 2-fold change are denoted by triangular and pentagonal points. Conversely, metabolites demonstrating a significant reduction ($p<0.05$) with a minimum 2-fold change are represented by diamond and square points. Metabolites that displayed no significant change ($p>0.05$) or less than a twofold alteration are indicated by circular points.

Twenty-one metabolites increased by at least twofold ($p<0.05$) following germination of soybeans (**Figure 2.3A**). Included in this group are twelve amino acids whose concentrations increased significantly (**Figure 2.4E, J-M, O-T, V**). The observed elevation in amino acid levels can be attributed to the heightened metabolic rate and increased protein synthesis during seed germination(Sattar and Akhtar 1990). While no significant modulation was observed in the identified bioactive metabolites, it is worth noting the detection of decreased levels of isoflavone glucosides alongside increased levels of the isoflavone aglycones daidzein and genistein, albeit the modulation being less than twofold. This change indicates a reduction in complex, conjugated forms of isoflavones, accompanied by an increase in simpler isoflavone aglycones, which is likely due to the activity of β-glucosidases. These enzymes facilitate the hydrolysis of glycosidic bonds in conjugated isoflavones, thereby transforming them into aglycones(Ribeiro, Mandarino et al. 2006).

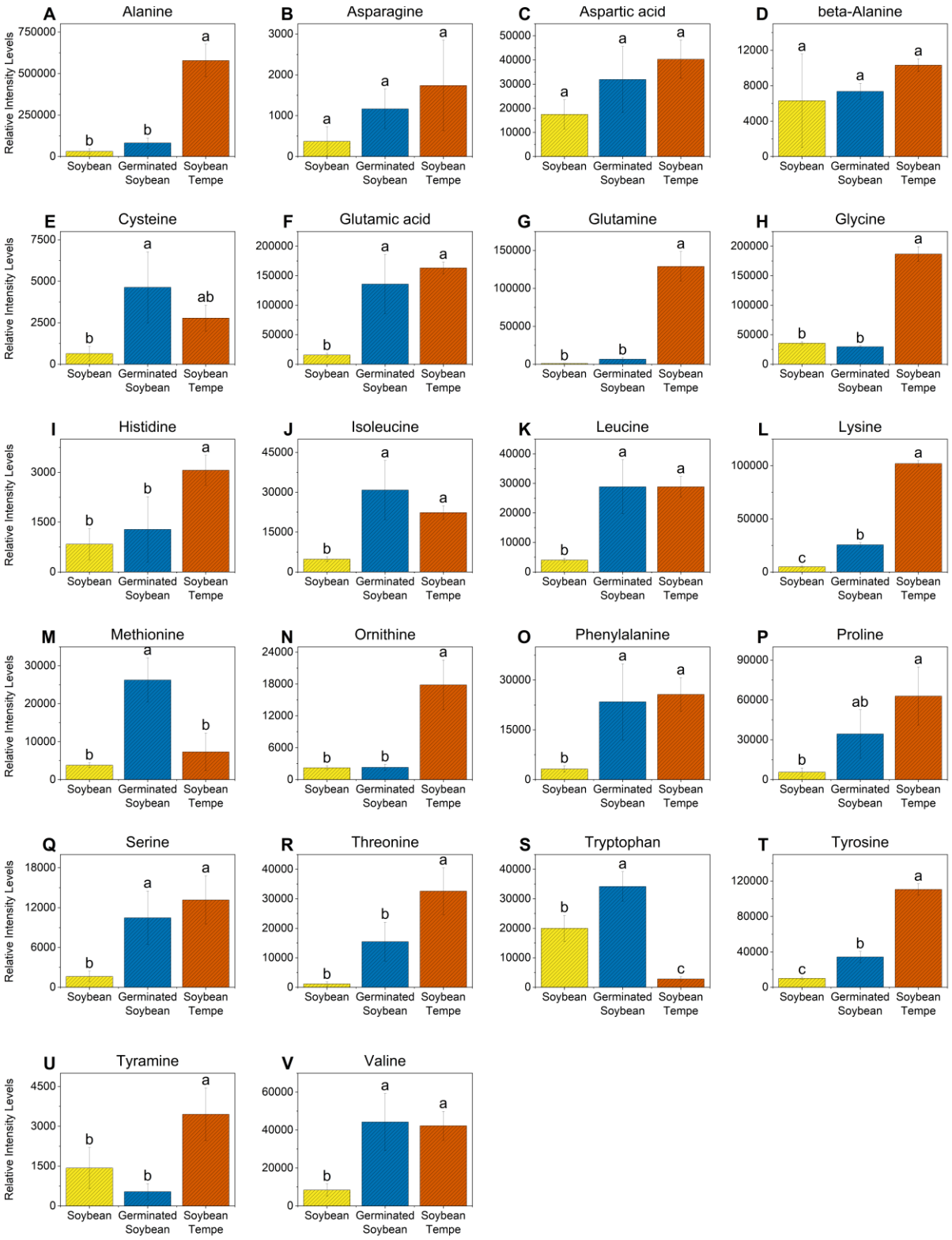


Figure 2.4. Levels of amino acid detected in soybean, germinated soybean, and soybean tempe. **A)** alanine; **B)** asparagine; **C)** aspartic acid; **D)** beta-alanine; **E)** cysteine; **F)** glutamic acid; **G)** glutamine; **H)** glycine; **I)** histidine; **J)** isoleucine; **K)** leucine; **L)** lysine;

M) methionine; **N)** ornithine; **O)** phenylalanine; **P)** proline; **Q)** serine; **R)** threonine; **S)** tryptophan; **T)** tyrosine; **U)** tyramine; **V)** valine. Vertical axis shows relative intensity levels. Error bars signify ±1SD from the mean. Groups labeled with different letters are significantly different ($p < 0.05$).

Tempe fermentation process led to a notable increase of at least two-fold in 40 metabolites, while 21 metabolites experienced a significant decrease (**Figure 2.3B**). These results indicate that tempe fermentation has a more substantial impact on the metabolome of soybeans compared to germination. Notably, seven bioactive metabolites were found to be significantly upregulated after tempe fermentation, including the three soy isoflavone aglycones (daidzein, glycitein, and genistein), as well as acetyl genistin, acetyl daidzin, 3-hydroxyanthranillic acid, and meglutol. The increased levels of these bioactive metabolites can be attributed to the activity of *Rhizopus* spp. fungi, which are known to break down isoflavone glucosides into aglycones using β-glucosidases(Mo, Kariluoto et al. 2013). Furthermore, previous studies have reported that the metabolic activity of *Rhizopus* spp. can stimulate the synthesis of bioactive compounds, which may also explain the observed elevation in bioactive metabolites in this study(Esaki, Onozaki et al. 1996).

Similar to post-germination, a significant elevation in the levels of 14 amino acids was also seen following tempe fermentation (**Figure 4A, E, G-L, N, Q, R, T-V**). The observed increases in amino acid concentrations are also attributable to the fungus *Rhizopus* spp. *Rhizopus* spp. breaks down proteins into more digestible amino acids during tempe fermentation, thereby contributing to the overall increase in amino acid concentrations(Rahmawati, Astawan et al. 2021). Tryptophan level, however, was observed

to be significantly lower after fermentation (Figure 2.4S). A significant drop in a number of sugars and sugar alcohols during soybean tempe fermentation was also observed.

2.3.5 Soybean bioactive metabolite levels following tempe fermentation and germination

The bioactive metabolite levels after soybean tempe fermentation were evaluated individually. These levels were then compared between mature (yellow) and immature (green) soybean varieties, as well as against germinated soybean as controls (**Figure 2.5**).

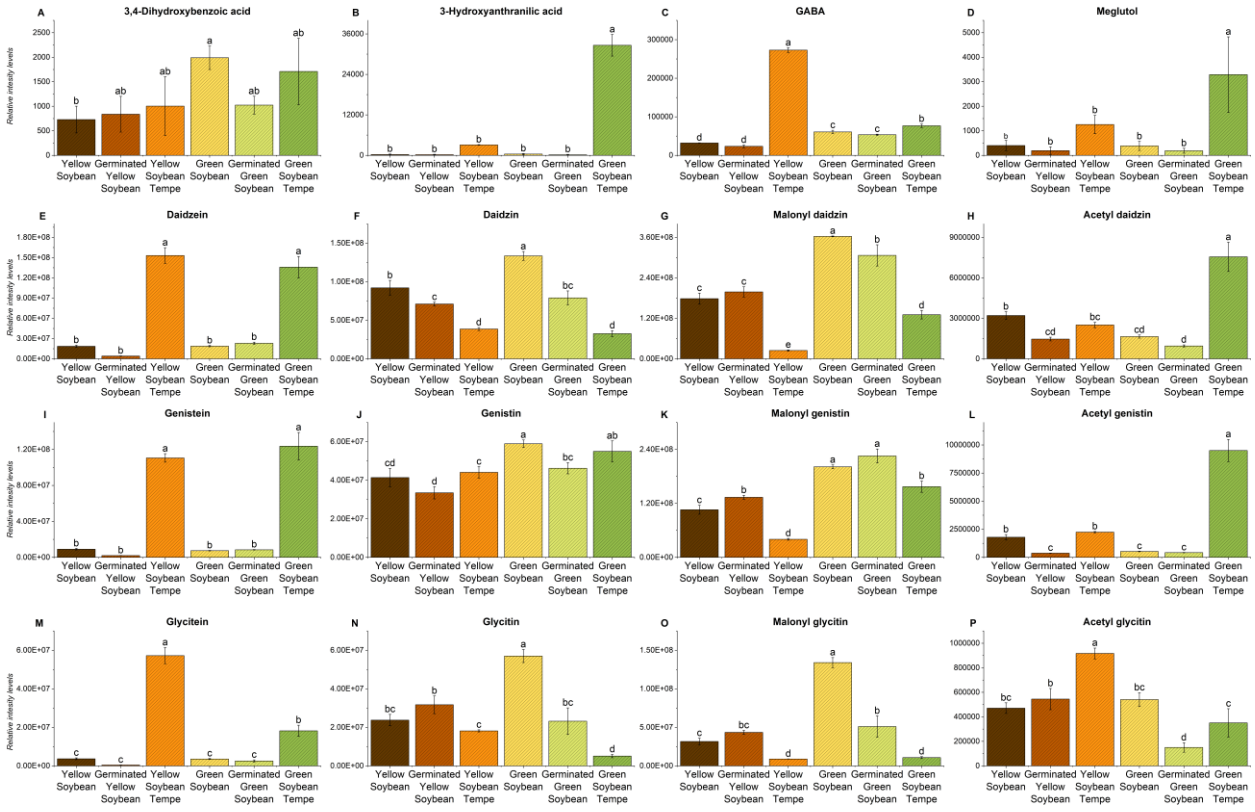


Figure 2.5. Levels of bioactive metabolite in mature (yellow) and immature (green) soybean after germination and fermentation processes. **A)** 3,4-hydroxybenzoic acid; **B)** 3-hydroxyanthranillic acid; **C)** GABA; **D)** meglutol; **E)** daidzein; **F)** daidzin; **G)** malonyl daidzin; **H)** acetyl daidzin; **I)** genistein; **J)** genistin; **K)** malonyl genistin; **L)** acetyl genistin; **M)** glycitein; **N)** glycitin; **O)** malonyl glycitin; **P)** acetyl glycitin. Vertical axis shows

relative intensity levels. Error bars signify ±1SD from the mean. Groups labeled with different letters are significantly different ($p < 0.05$).

Immature (green) soybean tempe showed the highest levels of acetyl daidzin, acetyl genistin, meglutol, and 3-hydroxyanthranilic acid compared to all other groups tested. In addition, all three isoflavone aglycones were substantially higher in green and mature (green) soybean tempe than in other non-fermented soybean sample groups. In contrast, mature (green) soybean tempe contained the highest levels of GABA, acetyl glycitin, and glycitein. Immature (green) soybean contained the highest levels of isoflavone glucosides, their malonyl derivatives, and 3-hydroxybenzoic acid. The overall higher levels of bioactive metabolites contained in immature (green) soybean and immature (green) soybean tempe compared to their mature (yellow) counterparts (**Figure S4**) is likely due to the difference in maturity level. Immature (green) soybeans and mature (green) soybeans are harvested at different stages of maturation. While mature (green) soybeans are harvested when they are completely mature, immature (green) soybeans are picked when they are only approximately 80% mature. Previous studies have indicated that isoflavone levels in soybeans decrease as they mature, suggesting that these bioactive metabolites may be consumed during the maturation process(Simonne, Smith et al. 2000).

Finally, among the 16 bioactive metabolites identified in soybean, germinated soybean, and soybean tempe, 3-hydroxy-3-methylglutaric acid (meglutol) was found to be strongly retained in human blood plasma samples. The detection of meglutol in blood plasma implies that this metabolite undergoes minimal metabolic breakdown or conversion into less biologically active forms during digestion, thereby reinforcing its potential as a bioactive compound. Additionally, meglutol was one of the few bioactive metabolites

found to be significantly elevated after tempe fermentation. Despite this, studies on meglutol's health benefits and its relation with food are currently very limited. For these reasons, in the next few chapters further characterization of meglutol's health benefits, as well as meglutol in food will be explored.

2.4 Conclusion

In this chapter, a comprehensive investigation was conducted to explore the bioactive metabolites in soybean tempe using metabolomics-based approaches. The study employed a combination of widely-targeted GC-MS, untargeted LC-MS, and targeted HPLC-PDA-MS analyses to perform comprehensive metabolome profiling. This approach successfully identified 16 bioactive metabolites in samples of soybean, germinated soybean, and soybean tempe. The identified bioactive metabolites consist of 12 well-characterized soy isoflavones daidzin, glycitin, genistin, malonyl daidzin, malonyl genistin, malonyl glycitin, acetyl daidzin, acetyl genistin, acetyl glycitin, daidzein, glycitein, genistein, as well as 3,4-dihydroxybenzoic acid, 3-hydroxyanthranillic acid, 3-hydroxy-3-methylglutaric acid (meglutol), and 4-aminobutyric acid (GABA). To the best of our knowledge, this is the first study to highlight 3,4-dihydroxybenzoic acid and 3-hydroxy-3-methylglutaric acid (meglutol) as bioactive metabolites in tempe. Following this, the effect of tempe fermentation and germination to the metabolome of soybean was investigated with special attention to the identified bioactive metabolites. Overall, the results presented in this chapter further contribute to the current understanding of tempe functionality and lay a foundation for further enhancing the presence of bioactive metabolites in tempe.

Chapter 3

Population-based examination of plasma 3-hydroxy-3-methylglutaric acid (meglutol) and cardiometabolic risk

3.1 Introduction

Cardiovascular disease (CVD) is a major public health problem that affects millions of people around the world and accounts for a large proportion of global mortality and morbidity. An estimated 17.9 million people died from CVDs in 2019, representing 32% of all global deaths(Bachheti, Worku et al. 2022). CVD encompasses a range of disorders that affect the heart and blood vessels, such as coronary artery disease, stroke, heart failure, and peripheral arterial disease(McDermott, Mandapat et al. 2003). Several factors can increase the risk of developing CVD, some of which are non-modifiable (e.g., age, sex, genetics) and some of which are modifiable (e.g., smoking, hypertension, diabetes, obesity)(Mohammadnezhad, Mangum et al. 2016). Among the modifiable risk factors of CVD, elevated low-density lipoprotein (LDL) cholesterol is considered to be one of the most significant. LDL cholesterol, a lipoprotein responsible for transporting cholesterol from the liver to peripheral tissues, plays a crucial role in the pathogenesis of CVD(Penson, Pirro et al. 2020). Excessive levels of LDL cholesterol can lead to the accumulation of cholesterol in arterial walls, resulting in the formation of plaques that constrict the arterial lumen and impede blood flow. This process, referred to as atherosclerosis, poses a substantial risk of ischemia and tissue damage in various organs, particularly the heart and brain. As a result, the reduction of LDL cholesterol levels emerges as a pivotal strategy in both the prevention and treatment of CVD(Ference, Ginsberg et al. 2017).

Meglutol (3-hydroxy-3-methylglutaric acid) is a hypolipidemic agent naturally found in some food sources, such as oranges, seeds, and beans. It inhibits cholesterol synthesis by acting as a competitive inhibitor to 3-hydroxy-3-methylglutaryl-CoA reductase (HMGCR), an enzyme in the cholesterol biosynthesis pathway. In a previous report, meglutol administration was observed to successfully lower plasma LDL cholesterol concentrations in patients with familial hypercholesterolemia (type IIa or hyper-beta-lipoproteinemia) during an eight-week treatment period(Lupien, Moorjani et al. 1979).

Despite this, limited data is available on the prospective association between plasma meglutol concentrations and changes in cardiometabolic risk factors in larger sample sizes and in larger populations. In recent years, advances in the field of epidemiology and metabolomics have allowed for the collection of valuable metabolome data from participants involved in cohort studies. Traditionally, cohort studies have relied on self-reported information, medical records, and occasionally, genetic data. However, the incorporation of metabolome data has revolutionized the scope and depth of information that can be gathered. With the integration of metabolomics-based profiling to human samples such as blood plasma or urine, the molecular basis of diseases and their relationship with participants' metabolic states can be inferred(Ala-Korpela, Kangas et al. 2012).

In this chapter, the relationship between plasma meglutol and clinical cardiometabolic risk factors is investigated in population-based studies.

3.2 Materials and methods

3.2.1 Study participants

The association between plasma meglutol concentration and cardiometabolic risk factors was examined in two well-characterized, population-based longitudinal cohort studies of Puerto Rican adults. The Boston Puerto Rican Health Study (BPRHS) included 1,500 adults aged 45-75 living in or near Boston, MA, and the San Juan Overweight Adult Longitudinal Study (SOALS) included 1,300 overweight or obese [BMI $\geq$ 25 kg/m^2] adults from the San Juan, PR municipality and surrounding areas. Participants in BPRHS completed 2-year and 5-year follow-up visits, and participants in SOALS completed a 3-year follow-up visit. Design and data collection details for these cohorts have been published elsewhere(Tucker, Mattei et al. 2010, Andriankaja, Jiménez et al. 2015, Pérez, Muñoz et al. 2017, Joshipura, Muñoz-Torres et al. 2018). The cross-sectional analysis included participants with baseline blood samples, plasma metabolomics profiling, and cardiometabolic risk factors. Participants were excluded if they lacked cardiometabolic risk factors at follow-up or if their 2-year standardized change in each outcome was above or below four standard deviations (SD) from the mean.

3.2.2 Blood sample collection and plasma metabolomics profiling

Baseline measurements were established for each participant in the BPRHS and SOALS studies after they had fasted for 10 or 12 hours, respectively. In the BPRHS, blood samples were collected at the participants' residences. These samples were then separated using a portable centrifuge, transported to the laboratory in freezers filled with dry ice, processed, and stored at -70°C. Similarly, in the SOALS, samples were collected at the study center, centrifuged, and frozen at -80°C. Following the previously described

methods(Haslam, Liang et al. 2021), metabolomic profiling of blinded samples was conducted by Metabolon (Durham, North Carolina)(Evans, DeHaven et al. 2009). Metabolomic profiling was performed on blinded samples from 714 BPRHS (371 with T2D, 339 without T2D, and 4 missing T2D status) and 1,011 SOALS (77 with T2D, 938 without T2D) participants in 2017 and 2019, respectively. An inverse normal transformation was used to account for the skewed distribution of plasma meglutol, and undetectable values were imputed at half the minimum.

3.2.3 Outcome assessment

Cardiometabolic outcomes were assessed at multiple time points throughout the study. For participants in the BPRHS group, evaluations took place at baseline, 2-3 years, and 5-7 years, while for SOALS participants, assessments were conducted at baseline and 3 years. Plasma cholesterol, triglyceride (TG), and HDL cholesterol (HDL-C) concentrations were measured using an enzymatic endpoint reaction on an Olympus AU400 for the BPRHS group, following standard operating procedures. For the SOALS group, commercially available assays were employed. The LDL cholesterol (LDL-C) concentration was calculated using the Friedewald formula, unless the TG concentration exceeded 400 mg/dL(Friedewald, Levy et al. 1972). Standard assays were utilized to determine plasma glucose levels, with the BPRHS group employing the Olympus America Inc. OSCR6121 assay from Melville, NY, and the SOALS group using the Vitros System 250 instrument. Insulin levels were measured using the Diagnostic Products Immulite 1000 Insulin Kit (LKIN1) for BPRHS and the TOSOH analyzer for SOALS. HbA1c levels were assessed using the Roche Unimate HbA1c kit from Roche Diagnostics in Indianapolis, Indiana, for BPRHS, and the Siemens Kit for DCA 2000 for SOALS. The intra- and inter-assay

coefficients of variation (CV%) for all measurements were below 10%(Tucker, Mattei et al. 2010, Joshipura, Muñoz-Torres et al. 2018). The waist circumference (WC) was recorded at the umbilical level, with measurements rounded to the nearest 0.1 cm. Height and weight were measured at baseline and follow-up, and body mass index (BMI) was calculated by dividing weight (kg) by height (m^2). The homeostatic model assessment of insulin resistance (HOMA-IR) was computed by multiplying fasting insulin (mU/L) by fasting glucose (mmol/L) and dividing the result by 22.5.

3.2.4 Covariate assessment

During all study visits, comprehensive data were collected using validated questionnaires to gather information on various factors, including age, education, household income, medication use, family history, and health behaviors such as smoking and alcohol consumption. Participant tension levels were assessed using a Spanish version of the Perceived Tension Scale(Cohen, Kamarck et al. 1983). In the BPRHS, dietary habits and nutrient intake were evaluated through a validated food frequency questionnaire (FFQ) specifically adapted for the Puerto Rican population(Tucker, Bianchi et al. 1998). Additionally, diet quality was determined by calculating compliance with the American Heart Association Diet Score(Bhupathiraju, Lichtenstein et al. 2011). White blood cell counts were measured using standard procedures in both cohorts. In the SOALS study, participants' physical activity levels were assessed by determining a metabolic equivalent (MET) score, which was derived from a questionnaire capturing the duration and frequency of physical activities performed on a weekly basis. In the BPRHS, physical activity was measured using an adjusted version of the Paffenbarger questionnaire(Paffenbarger Jr, Hyde et al. 1993, Paffenbarger Jr, Wing et al. 1995), while acculturation levels were

determined using the Acculturation Scale for Hispanics. The scale ranges from 0 to 100, with a score of 100 indicating complete assimilation into US culture(Marin and Gamba 1996).

3.2.5 Statistical analysis

Meglutol concentrations and cardiometabolic risk factors were assessed at baseline and follow-up examinations for 636 people in the BPRHS and 1,011 participants in the SOALS using multivariable linear regression models. At baseline, standardized plasma meglutol concentration (per 1-SD) was evaluated in relation to cardiometabolic risk factors using these models. The primary outcome of interest was LDL-C, and secondary outcomes included HDL-C, TG, WC, BMI, glucose, insulin, HbA1c, HOMA-IR, and CRP. The models were adjusted for several covariates. Initially, coefficients and standard errors (SE) were extracted from the regression models. Model 1 adjusted for age and sex. Model 2 further adjusted for smoking status (never, current, former), education level (≤8th grade, 9th-12th grade or GED, college/some graduate school), physical activity (BPRHS: continuous score; SOALS: METs), alcohol intake (non-consumer, moderate consumer [females: 1 drink/day; males: 1-2 drinks/day], heavy consumer [females: >1 drink/day; males: >2 drinks/day]), T2D status, lipid-lowering medication status, hypertension medication status, income (<$20,000/year, ≥ $20,000/year), acculturation (BPRHS only: %), diet quality (BPRHS only: AHA-DS continuous score), perceived-stress score, family history of T2D (SOALS only), and white blood cell count (CRP outcome only: mm3). Finally, Model 3 added BMI (kg/m^2) and WC (cm) as additional covariates to the previous models. This comprehensive approach aimed to investigate the associations

between meglutol and cardiometabolic risk factors while accounting for relevant confounding factors.

Secondly, we investigated associations between baseline plasma meglutol concentration and longitudinal changes in cardiometabolic risk factors over the course of the follow-up period. To assess these changes, we calculated the standardized two-year differences in cardiometabolic risk factors by subtracting the baseline value from the previous examination, dividing by the number of years of follow-up, and multiplying by two. To analyze the data, we employed multivariable linear regression models with the standardized two-year changes in cardiometabolic risk factors as the outcome. We extracted the coefficients and standard errors (SE) from these models. In the longitudinal analyses, model 1 accounted for baseline values of each associated cardiometabolic risk factor. Additionally, model 3 replaced the adjustment for baseline waist circumference with standardized changes in waist circumference (WC) over a two-year period. To address the issue of multiple observations across participants during the study, we utilized linear mixed-effects regression models in the BPRHS. These models accounted for the variations in measurements obtained at baseline, 2-3-year, and 5-7-year visits.

Thirdly, in our analysis, we explored the presence of effect modification by T2D status (specifically in the BPRHS dataset, as the number of participants with T2D in the SOALS dataset was limited), lipid-lowering medication use, and BMI. We divided participants into two groups based on BMI: obese individuals (BMI $\geq$ 30 kg/m2) and those with normal/overweight BMI (BMI < 30 kg/m2). These analyses were conducted for both cross-sectional and longitudinal assessments of the primary outcome, which focused on LDL-C levels. In the fully adjusted models (Model 3), we obtained β coefficients and

standard errors (SE) from stratified analyses and assessed p-values for the multiplicative interaction terms. To combine estimates from each cohort, we performed meta-analyses using inverse-variance weighted random effects. All statistical analyses were conducted using R (version 3.6.0) statistical software(Team 2013).

3.3 Results and discussion

The demographic characteristics of participants in the Boston Puerto Rican Health Study (BPRHS) and the San Juan Overweight Adult Longitudinal Study (SOALS) are presented in **Table S4**. The average age at baseline for BPRHS participants was 57.1 years, whereas for SOALS participants, it was 50.7 years. A meta-analysis of data from individuals in both studies revealed a significant association between lower baseline meglutol concentration (per 1-SD) and decreased LDL-C concentration (β [SE]: -5.5 [1.6], P = 0.0005) (**Figure 3.1**).

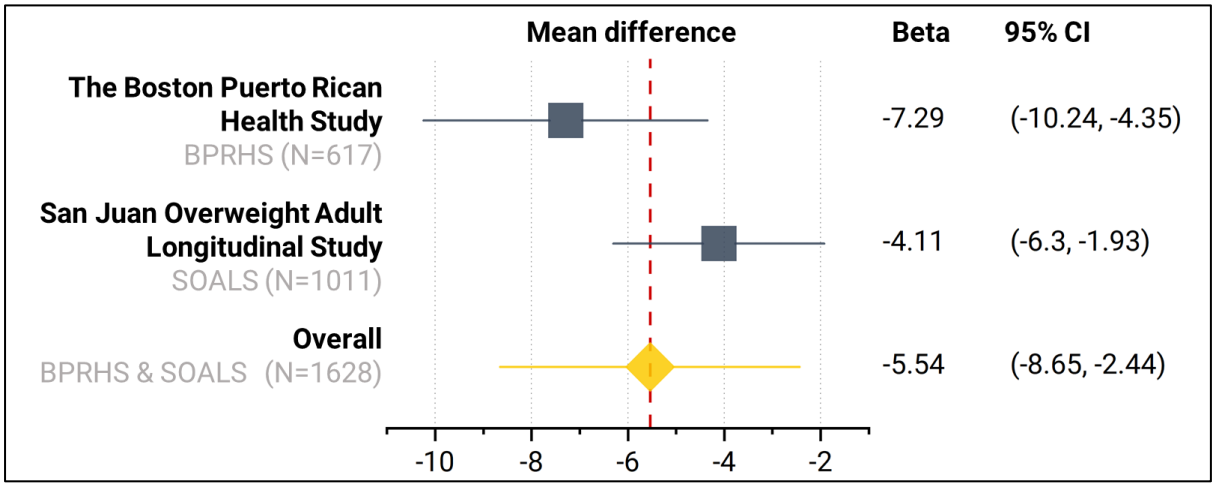


Figure 3.1. Cross-sectional associations between meglutol concentration and LDL-C concentration in Puerto Rican adults. Results from a random effects meta-analysis of the BPRHS and SOALS participants demonstrate that higher baseline meglutol concentration (per 1-SD) is significantly associated with lower LDL-C concentration.

However, when examining longitudinal changes in cardiometabolic risk factors within the combined BPRHS and SOALS cohorts, no significant associations were found between meglutol and these factors. In secondary analyses, after adjusting for relevant variables, we observed that higher baseline meglutol concentration (per 1-SD) was significantly associated with increased BMI (β [SE]: 0.3 [0.2], P = 0.04).

The association of meglutol and LDL-C concentration exhibited significant effect size modifications based on the use of lipid-lowering medication and the presence of T2D. Among individuals taking LDL-lowering medication, the association between baseline meglutol and LDL-C was notably stronger (β [SE]: -7.6 [1.7], P < 0.0001) compared to those not taking such medication (β [SE]: - 3.7 [1.7], P < 0.0001) (Pint = 0.002) (**Figure 3.2**).

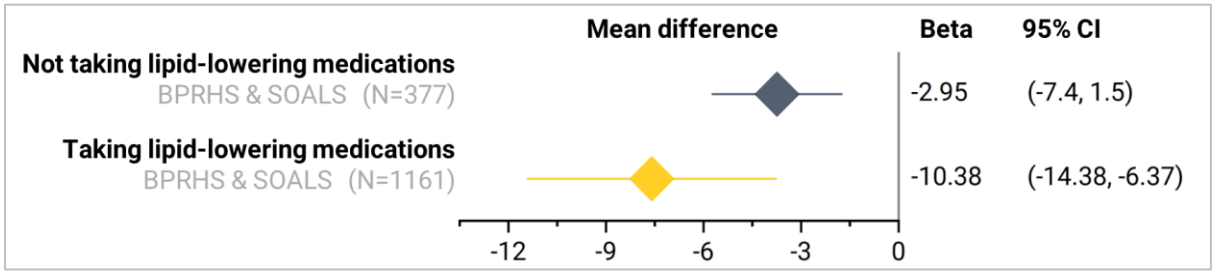


Figure 3.2. Cross-sectional associations between meglutol concentration and LDL-C concentration in Puerto Rican adults were observed to be stronger among individuals taking LDL-lowering medication (β [SE]: -7.6 (1.7), P < 0.0001) compared to those not taking LDL-C lowering medication (β [SE]: -3.7(1.7), P < 0.0001) (Pint = 0.002)

In the BPRHS, baseline meglutol was significantly associated with LDL-C concentration in individuals with T2D (β [SE]: -10.4 [2.0], P < 0.0001), but not in those without T2D (β [SE]: -3.0 [2.3], P = 0.19) (Pint < 0.0001) (**Figure 3.3**).

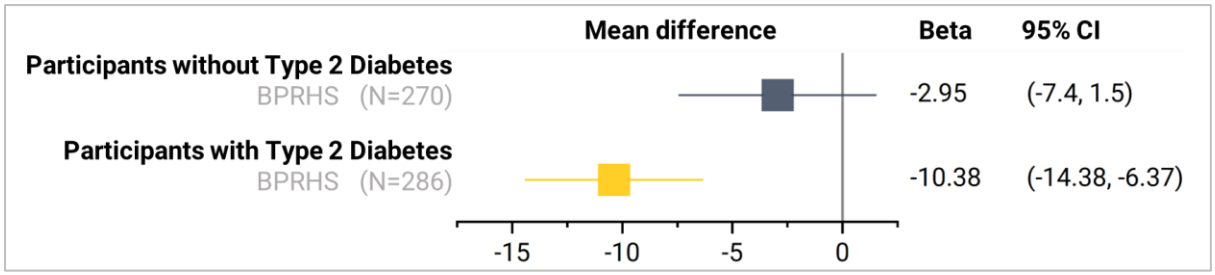


Figure 3.3. Cross-sectional associations between meglutol concentration and LDL-C concentration in BPRHS participants, considering the interaction with T2D. Baseline meglutol concentration was found to be significantly associated with LDL-C among individuals with T2D (β [SE]: -10.4 (2.0), P < 0.0001). However, this association was not observed among individuals without T2D (β [SE]: -3.0 (2.3), P = 0.19) (Pint < 0.0001).

Furthermore, the association between baseline meglutol and LDL-C was found to be statistically significant only in individuals with obesity (BMI ≥ 30 kg/m2; β [SE]: -6.2 [1.1], P < 0.0001), while no significance was observed in those without obesity (BMI < 30 kg/m2; β [SE]: -3.6 [2.5], P = 0.15). However, the test for interaction did not yield statistical significance (Pint = 0.07) (**Figure 3.4**). Importantly, in longitudinal models (**Figure S6**), no statistically significant effect modifications were observed in relation to lipid-lowering medication use, T2D status, or BMI.

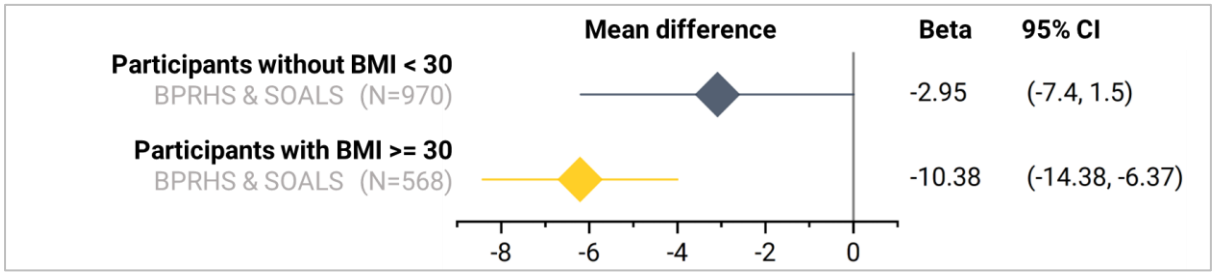


Figure 3.4. Cross-sectional association between baseline meglutol concentration and LDL-C concentration in Puerto Rican adults, considering the interaction with BMI. The

interaction effect of BMI was nearly significant (obese [BMI ≥ 30 kg/m2] vs. overweight/normal weight [BMI < 30 kg/m2]; p-value for interaction = 0.07).

Our study presents a novel finding regarding the association between plasma meglutol concentration and reduced LDL-C, based on two population-based cohort studies. To the best of our knowledge, no previous epidemiological studies have reported this specific association. This finding aligns with earlier meglutol supplementation trials conducted on hypercholesterolemia patients(Lupien, Moorjani et al. 1979, Afifah, Nabilah et al. 2020), which demonstrated reductions in LDL and total cholesterol levels. However, it is important to note that we did not observe any associations between meglutol and longitudinal variations in LDL-C. This could be attributed to the limitations of our study, including the lack of long-term follow-up and the relatively small sample sizes in the BPRHS and SOALS cohorts. Furthermore, in a previous study that investigated 1,300 metabolites in Qatari adults, the association between meglutol and LDL-C did not remain significant after multiple testing correction, although it still showed a significant association ($\beta = -0.10$; $P = 0.01$)(Yousri, Suhre et al. 2022). Interestingly, the same study revealed a positive association between meglutol and prevalent T2D, which was observed in two independent population studies. Our previous studies involving participants from the BPRHS and SOALS also found positive associations between meglutol and prevalent T2D(Haslam, Liang et al. 2021), although incident T2D was not observed(Rivas-Tumanyan, Pacheco et al. 2022). Additionally, another study found a positive correlation between plasma meglutol levels and coronary artery calcium score among 236 European-Americans(Chevli, Freedman et al. 2021). These contradictory associations between meglutol and cardiometabolic risk factors underscore the need for further research to

elucidate the determinants of plasma meglutol and its potential role in influencing cardiometabolic disease risk.

Furthermore, we found that the association between plasma meglutol and LDL-C at baseline to be stronger among individuals taking lipid-lowering medication, those with T2D, and those with obesity. Collectively, these participants make up a less metabolically healthy group, which may suggest that meglutol's LDL-lowering effects are more pronounced in individuals with lower metabolic health than among those with better metabolic health. In agreement with this finding, we also observed a minor positive association between meglutol and BMI in the BPRHS and SOALS cohorts. While our epidemiological approach does not allow us to establish the causal factors underlying these associations, we can speculate on potential drivers. One plausible explanation is that upregulated inflammatory pathways may trigger the endogenous increase in plasma meglutol production as a negative feedback mechanism to counteract the constant upregulation of cholesterol synthesis caused by inflammation(Lupien, Moorjani et al. 1979). Additionally, due to competitive binding of meglutol and drug substrates to enzyme active sites, it is possible that plasma meglutol concentration is upregulated in individuals using LDL-C-lowering medications. Lastly, it is possible that meglutol may have other unknown detrimental effects on metabolic health that outweigh its reported benefits in reducing LDL-C levels. Definitive insights into the overall metabolism and regulation of plasma meglutol necessitate controlled metabolic experiments, such as in vitro, animal, and human studies, which can provide causal understanding.

The epidemiological design presented in this study offers several key advantages. Firstly, it includes two complementary population-based epidemiological studies

conducted among ethnic minority populations. These studies provide a comprehensive understanding of the research topic by capturing diverse perspectives. Secondly, the availability of detailed medical history data, plasma cardiometabolic risk factors, and lifestyle factors allows for their incorporation into multivariable regression models. This comprehensive approach enhances the accuracy and robustness of the analyses. Furthermore, the inclusion of plasma meglutol concentration, a recently added metabolite in high-throughput plasma metabolomic profiling platforms, adds an additional dimension to the investigation.

However, it is important to acknowledge that the observational nature of these studies imposes limitations on establishing causality. The potential for residual confounding must be considered when interpreting the findings. Additionally, the modest number of participants and limited follow-up in the longitudinal analyses necessitate a primary focus on the cross-sectional findings from the baseline examinations. This limitation further reinforces the need to address potential residual confounding. Nevertheless, the meticulous data collection procedures and adjustment for possible confounders significantly reduce the risk of residual confounding. To gain a more comprehensive understanding of the role of plasma meglutol concentrations in long-term health and potential confounders in this relationship, larger studies with extended longitudinal follow-up are required. These studies should examine the complex relationship between dietary intakes, plasma meglutol, and cardiovascular disease events. Such endeavors will contribute to a more nuanced understanding of this topic. It is important to note that the findings of this study primarily pertain to individuals of Puerto Rican descent. Therefore, the generalizability of the results to other racial/ethnic groups may be limited.

Additional research focusing on diverse racial/ethnic groups is necessary to validate and expand upon these findings.

3.4 Conclusion

In summary, the associations of meglutol with cardiovascular risks are investigated in this chapter. Two well-characterized Puerto-Rican cohorts in the US, BPRHS and SOALS, were used to asses these associations. In a meta-analysis, we identified a significant association between lower baseline meglutol concentration (per 1-SD) and lower LDL-C concentration (β [SE]: -5.5 [1.6, P = 0.0005). Furthermore, following stratifications of the constructed models, we found that the effect sizes are stronger in metabolically compromised groups, including: patients with T2D, high BMI levels, and patients taking lipid-lowering medication. The results presented in this chapter are the first epidemiological evidence of the LDL-lowering effect of meglutol which further highlight the potential and importance of meglutol as an effective bioactive compound.

Chapter 4

Optimization of meglutol concentrations in tempe using raw material substitutions

4.1 Introduction

Meglutol (3-hydroxy-3-methylglutaric acid) is a hypolipidemic agent naturally found in some food sources. It inhibits cholesterol synthesis by acting as a competitive inhibitor to 3-hydroxy-3-methylglutaryl-CoA reductase (HMGCR), an enzyme in the cholesterol biosynthesis pathway. In a previous report, meglutol administration was observed to successfully lower plasma LDL cholesterol concentrations in patients with familial hypercholesterolemia (type IIa or hyper-beta-lipoproteinemia) during an eight-week treatment period(Lupien, Moorjani et al. 1979). While it has been identified as a dietary supplement, few foods containing high concentrations of meglutol have been thoroughly characterized.

In previous chapters of this dissertation, our research has revealed the presence of meglutol in soybean samples, with its levels significantly increasing during the process of tempe fermentation. Tempe, an Indonesian traditional fermented food prepared from soybean, is widely recognized for its exceptional nutritional and functional attributes(Astuti, Meliala et al. 2000). However, it is important to note that while our findings demonstrate the existence of meglutol in tempe, a direct comparison of meglutol levels between tempe and other food sources has not yet been conducted. This direct comparison is crucial because it allows us to gain insights into the relative abundance of meglutol in tempe compared to other commonly consumed foods. Understanding the comparative levels of

meglutol in various food sources such as oranges, seeds, and beans can provide valuable information regarding the dietary significance and potential health implications associated with meglutol intake. By examining the concentrations of meglutol in different food items, we can better evaluate the contribution of tempe to overall meglutol exposure in the diet and potentially elucidate any unique characteristics or advantages of consuming tempe in terms of meglutol content.

Furthermore, neither the modulation nor the extent to which meglutol concentrations in food can be modulated have been studied. Investigating the modulation of meglutol concentrations can have profound implications for human health. As explored in previous chapters of this dissertation, meglutol has antilipidemic properties. By elucidating the factors and mechanisms underlying the modulation of meglutol levels, we can potentially develop dietary interventions to increase meglutol content in foods, providing individuals with targeted approaches to optimize their intake of this bioactive compound for health promotion or management of specific health conditions.

The substitution of raw materials for making tempe has been widely investigated as a means to enhance its nutritional and functional properties(Njoku, Ofuya et al. 1991, Matsuo 2006, Bujang and Taib 2014, Pertiwi, Marsono et al. 2020). Previous studies have reported notable improvements when utilizing alternative ingredients. For instance, the production of tempe using chickpeas has shown enhanced antioxidant activity and increased protein levels(Sánchez-Magana, Cuevas-Rodríguez et al. 2014). Likewise, the utilization of grass pea as a raw material has demonstrated improved protein bioavailability while simultaneously reducing antinutrient content(Stodolak, Starzyńska-Janiszewska et al. 2020). Recently, previous studies have also explored the potential of other legumes such as

red kidney beans, white beans, and green peas to further enhance the nutritional and functional qualities of tempe(Rahmawati, Astawan et al. 2021).

This study aims to address the existing research gap by employing metabolomics techniques to assess and examine the meglutol levels in foods that have been previously identified as containing meglutol, such as tempe. Additionally, this research delves into the potential modulation of meglutol levels through the process of tempe fermentation and by substituting raw materials in tempe production.

4.2 Materials and methods

4.2.1 Legume tempe samples

It has been reported that alternative tempes alter the nutritional qualities and properties of tempe. In this investigation, we prepared red kidney bean, green pea, white bean, and edamame tempes, all of which were reported to contain meglutol(Rahmawati, Astawan et al. 2021). Sample information is provided in **Table S5**. All tempe samples were produced in December 2021 at Osaka University's Laboratory of Bioresource Engineering. Raprima was used as the initial culture to produce samples of tempe according to the Rumah Tempe Indonesia (RTI) method(Rahmawati, Astawan et al. 2021). Briefly, 50 g of each legume was soaked, simmered, and dehulled prior to the inoculation of the starter culture. The samples of tempe were then packaged and fermented at 30°C for 48 hours in an incubator (EYELA, Tokyo, Japan).

4.2.2 Metabolite extraction for gas chromatography-mass spectrometry (GC-MS) analysis

To extract metabolites from the biological samples, a protocol comprising a mixture of organic solvents was utilized. Samples were subjected to freeze-drying and homogenization using a multi-beads shocker (Yasui Kikai, Japan). Samples were homogenized and 10 mg of each sample were collected into 2 mL tubes. The process was repeated in triplicate to ensure accuracy and consistency of the results. A mixed solvent comprising methanol, water, and chloroform in the ratio of 5:2:2 containing an internal standard of 100 µg/mL ribitol was added to each tube. The tubes were then incubated at 37°C for 30 minutes with agitation at 1200 rpm (Eppendorf Ltd., Germany). Following the collection of samples, a centrifugation process was carried out at a temperature of 4°C for a duration of 3 minutes at a speed of 10,000 rpm. The resulting supernatant, measuring 400 µL, was then carefully transferred into fresh 1.5 mL tubes. Subsequently, 300 µL water was added, followed by a centrifugation step at 4°C for 3 minutes at 10,000 rpm. Next, a volume of 200 µL of the aqueous phase was transferred to fresh 1.5 mL tubes to be utilized for subsequent processing. In order to ensure the accuracy and reliability of the obtained results, a set of quality control (QC) samples were prepared. This was achieved by combining 200 µL of the aqueous phase obtained from all samples that were previously collected. Prior to analysis, the samples and quality control (QC) were subjected to centrifugation using a Taitec Co. centrifugal concentrator in order to remove excess liquid. The centrifugation process was carried out for a duration of 1 hour at ambient temperature, followed by overnight freeze-drying.

4.2.3 Derivatization for GC-MS analysis

First, freeze-dried samples were treated with 100 µL of methoxyamine hydrochloride and then incubated at 30°C for 90 minutes with 1200 rpm agitation. Following that, 50 µL

of N-Methyl-N-trimethylsilyltrifluoroacetamide (MSTFA) was added to the samples, then subjected to 30 minutes of incubation at 37°C with 1200 rpm agitation (Eppendorf Ltd., Germany). The collected samples were then placed in GC vials for GC-MS analysis.

4.2.4 GC-MS analysis

GC-MS analysis procedure was conducted using a GC-MS-QP2010 Ultra instrument (Shimadzu, Japan) equipped with an Inert-Cap MS/NS column (GL Sciences). The sample was injected into the GC using an AOC-20i/s autosampler (Shimadzu, Japan) in a 25:1 (v/v) split mode and 230°C injection temperature. Helium was utilized as the carrier gas at a flow rate of 1.12 mL/min and a linear velocity of 39 cm/s. The conditions for the GC analysis were as follows: an initial temperature of 80°C for 2 minutes was maintained, followed by an increase to 330°C at a rate of 15°C/minute and a 6-minute hold. Electron ionization (EI) was used with an ion source temperature of 200°C and a mass-to-charge ratio (m/z) range of 85-500 at a scan rate of 6.67 scans/s. The retention index (RI) was determined using a standard alkene mixture. The GC-MS analysis procedure was conducted at Osaka University's Laboratory of Bioresource Engineering.

4.2.5 Data processing and statistical analysis of tempe metabolite data

Baseline correction, denoising, peak detection, alignment, and annotation procedures were carried out using MS-DIAL 4.0 (RIKEN, Japan)(Lai, Tsugawa et al. 2018). Annotation was done using the RI and MS data from the GC/MS-5MP Library (RIKEN, Japan). Meglutol annotation was specifically verified and afterwards quantified by co-injection analysis utilizing a genuine standard (Wako Pure Chemical Industries, Japan).

The peak intensities of the identified metabolites were standardized against the peak intensity of ribitol, which served as the internal standard. Filtering was done to retain only the metabolites with QC standard deviation scores of 30% or lower. SIMCA P+ ver. 13.0 (Umetrics, Sweden) was used to perform principal component analysis (PCA). MetaboAnalyst 5.0(Pang, Chong et al. 2021) was used to conduct heatmap analysis and partial least squares-discriminant analysis (PLS-DA).

4.3 Results and discussion

Several foods previously reported to contain meglutol based on mass spectrometry (MS) or nuclear magnetic resonance (NMR) spectroscopy data are listed in **Table 4.1**. Some foods identified to have meglutol based on our internal food metabolome database screening are also included in the list. A substantial proportion of foods found to contain meglutol were legumes.

Table 4.1. Foods containing meglutol. Legumes and legume-based products were prominent contributors to the list, with notable inclusion of tempe, a traditional Indonesian food derived from fermented soybeans.

Food product	Detection method	Reference
Orange	GC-MS	(Klimek-Szczykutowicz, Szopa et al. 2020)
Flax seed	LC-MS	(Thiombiano, Gontier et al. 2020)
Oat	GC-MS	(Xu, Chen et al. 2021)
Lentils	GC-MS	(Farag, Khattab et al. 2018)
Lupinus seeds	GC-MS	(Farag, Khattab et al. 2018)
Cacao bean	GC-MS	Internal database
Soybean	GC-MS	Internal database
Red kidney bean	GC-MS	Internal database
White bean	GC-MS	Internal database
Green pea	GC-MS	Internal database
Edamame	GC-MS	Internal database

We measured the determined concentration of meglutol in foods listed in **Table 4.1**. Absolute quantification to measure meglutol concentrations was performed using GC-MS with authentic meglutol standards (Figure 4.2a).

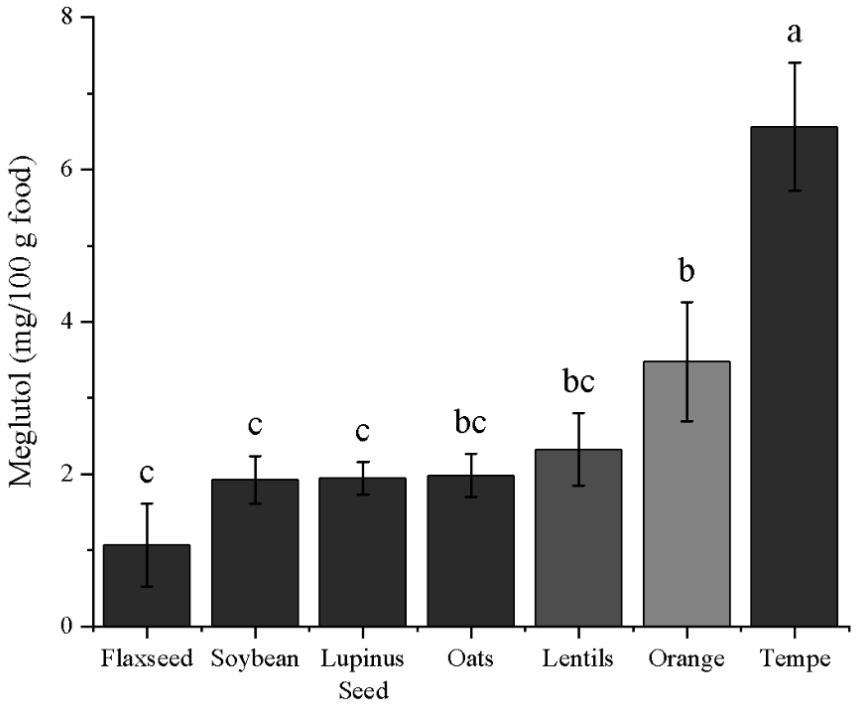


Figure 4.2. Meglutol concentrations in foods previously reported to contain meglutol. Meglutol quantification was conducted using GC-MS-based absolute quantification, validated with co-injection of a meglutol standard compound. Statistical differences between groups are denoted by different letters, adjusted with Tukey's method ($P < 0.05$). Error bars represent ±1SD from the mean measurements. Compared to other meglutol-containing foods, tempe exhibited a significantly higher meglutol concentration.

Our results indicated that tempe has significantly higher meglutol concentrations when compared to other previously reported foods. Tempe contained more than three times as

much meglutol as previously reported meglutol-containing foods on average (p = 6.763e-8) (**Table S6**). Tempe is an Indonesian traditional food made from soybean fermented with *Rhizopus* fungi(Nout and Kiers 2005). Tempe is a nutritious alternative food source that is abundant in bioactive compounds that promote health, such as isoflavones.

In Chapter 2, we demonstrated the ability of tempe fermentation to modulate the metabolome profile of soybeans, resulting in improved concentrations of bioactive compounds. Building upon this foundation, here, we further explore the impact of tempe fermentation on metabolome modulation in alternative legumes, with the aim of optimizing meglutol concentration in tempe. To achieve this, first, we produced tempes using alternative legumes (**Figure 4.3**) white bean, red kidney bean, green pea, and edamame, with *Rhizopus oligosporus* as the starter culture.

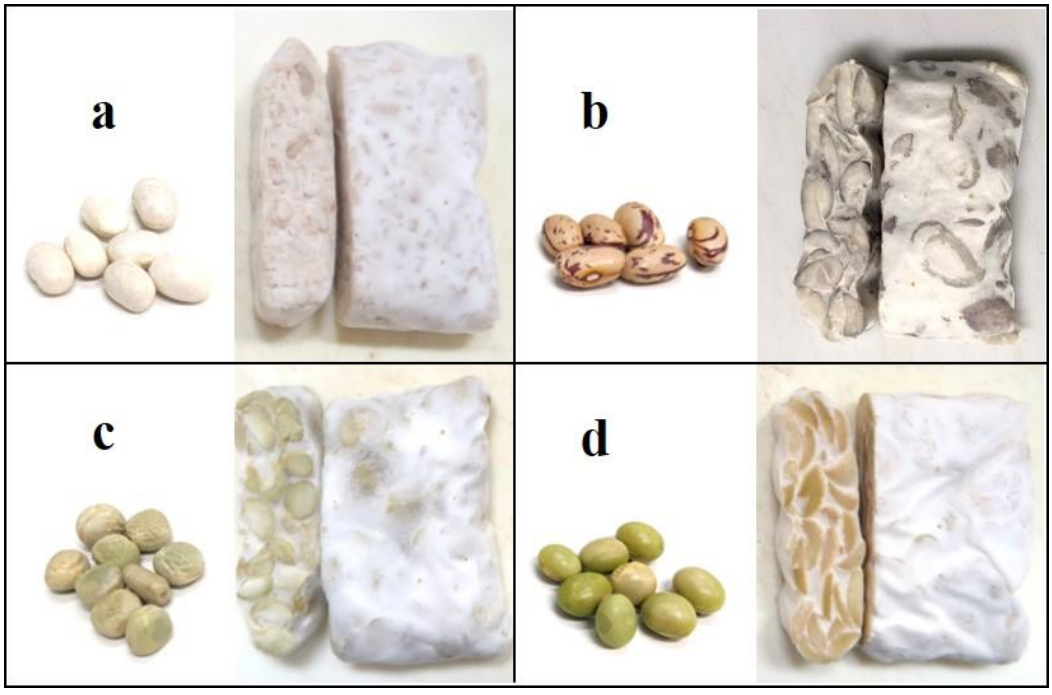
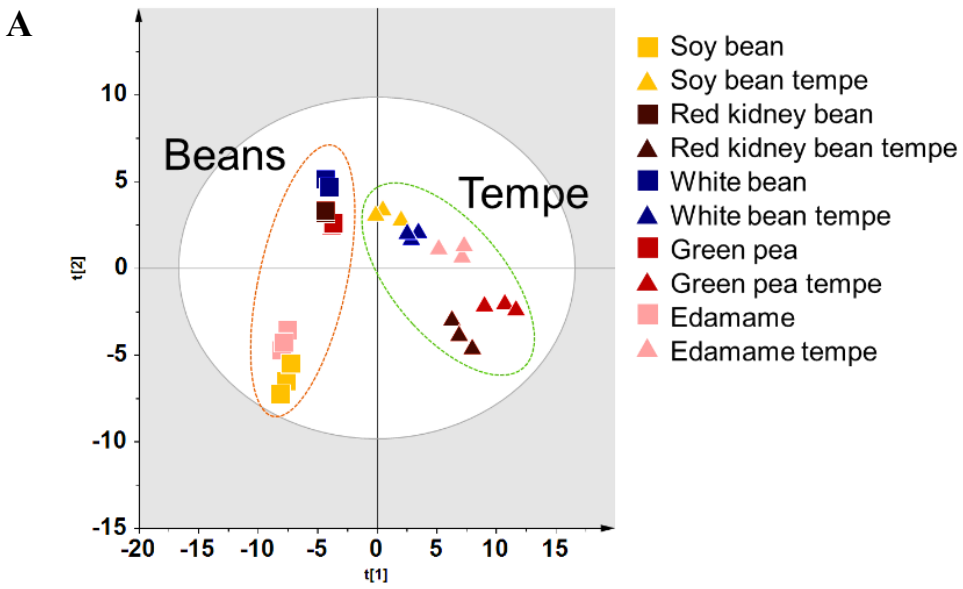


Figure 4.3. Legumes used as substitute raw materials to make tempes. **A)** White bean and white bean tempe; **B)** red kidney bean and red kidney bean tempe; **C)** green pea and green pea tempe; **D)** edamame and edamame tempe.

Next, we conducted a comparative analysis of the metabolome profiles of these different legumes before and after the 48-hour tempe fermentation process. Through a comparison of retention index (RI) values and mass spectra with our in-house library, we successfully identified and annotated a total of 68 primary metabolites from these legumes and their resulting tempes (**Table S7**). To ensure accuracy, the presence of meglutol was specifically confirmed using a standard compound co-injection. Subsequently, we employed unsupervised principal components analysis (PCA) to examine the impact of tempe fermentation on the metabolome of these legumes, as illustrated in **Figures 4.3**.



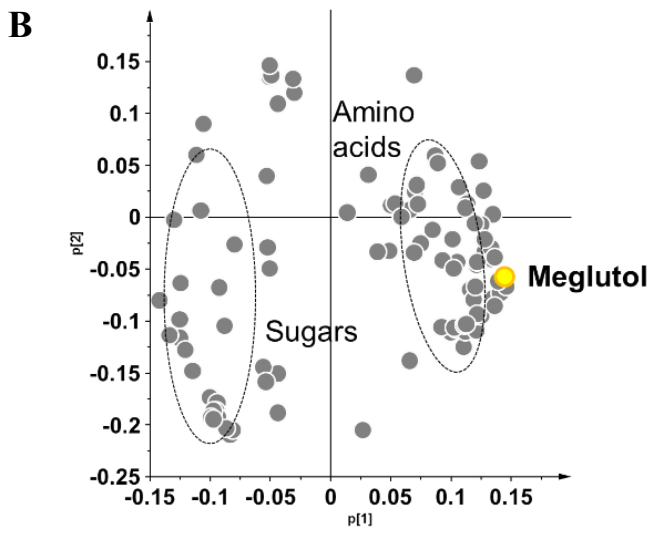


Figure 4.3. Metabolome profiles of legumes pre- and post-tempe fermentation. **A)** Principal Component Analysis (PCA) score plot illustrating the distribution of annotated metabolites obtained from GC-MS analysis. Square points denote legume samples before fermentation, while triangle points represent tempe samples after fermentation. **B)** PCA loading plot showcasing the annotated metabolites obtained from GC-MS analysis. The light-yellow circle corresponds to meglutol.

The PCA analysis revealed two distinct clusters along the PC1 axis of the score plot (**Figure 4.3A**). All legume samples prior to fermentation exhibited a cluster on the negative side of the PC1 axis, while the resulting tempes from the fermentation process formed a unified cluster on the positive side of the PC1 axis. These findings strongly suggest that tempe fermentation induces a consistent alteration in the metabolome of various legumes. By examining the loading plot (**Figure 4.3B**), it became evident that sugars were concentrated on the negative side of the PC1 axis, while amino acids accumulated predominantly on the positive side of PC1. This observation suggests that fermentation leads to a reduction in sugar content and an increase in amino acid content within the legumes. The heatmap analysis (**Figure 4.4**) provides a visual representation that highlights

the modulation of individual metabolites. Intriguingly, meglutol appeared to be clustered on the extreme positive side of PC1 in the loading plot, providing further confirmation of our earlier observation that tempe fermentation enhances meglutol concentration.

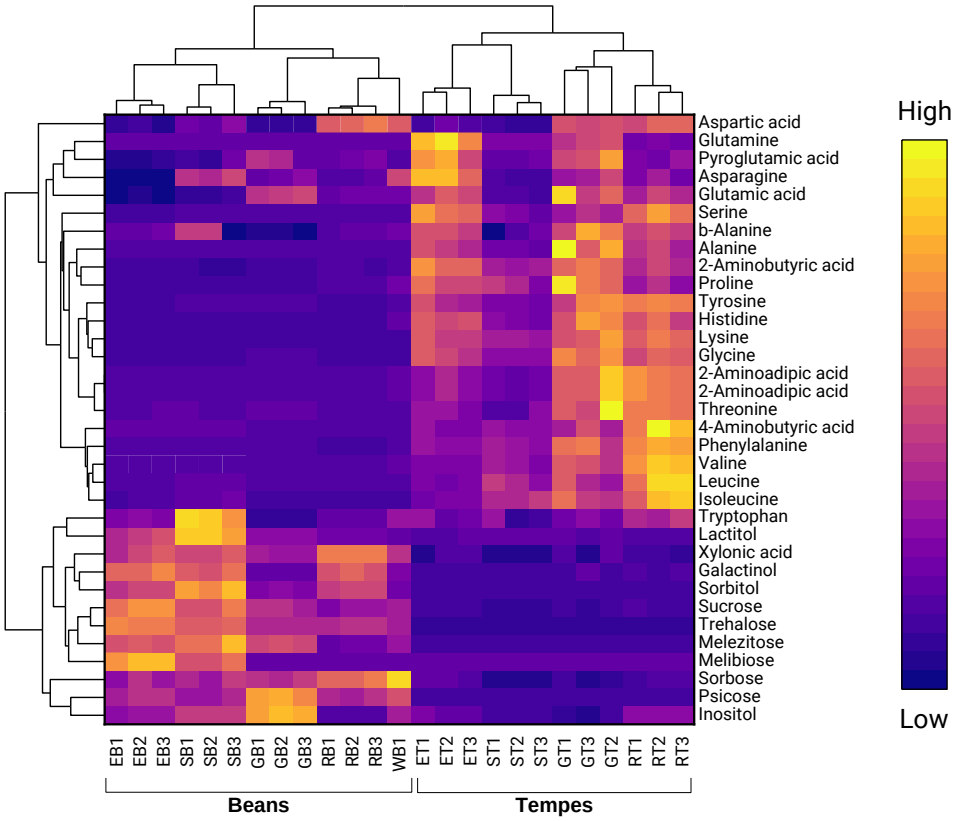


Figure 4.4. heatmap analysis highlighting the modulations of sugars and amino acid levels during tempe fermentation.

In addition, the variance of metabolites belonging to legume cluster prior to tempe fermentation were relatively higher compared to those of the tempe cluster, which can be observed along the PC2 axis of the loading plot (**Figure 4.3B**). This observation suggests that the tempe fermentation agent, Rhizopus oligosporus, similarly affects the metabolome of these diverse legumes, despite their initial metabolome differences.

We conducted GC-MS-based metabolite quantification to determine the levels of meglutol in various legumes and tempe individually. The results are presented in **Figure 4.5**, which illustrates the amount of meglutol present in 100 grams of each legume and tempe sample.

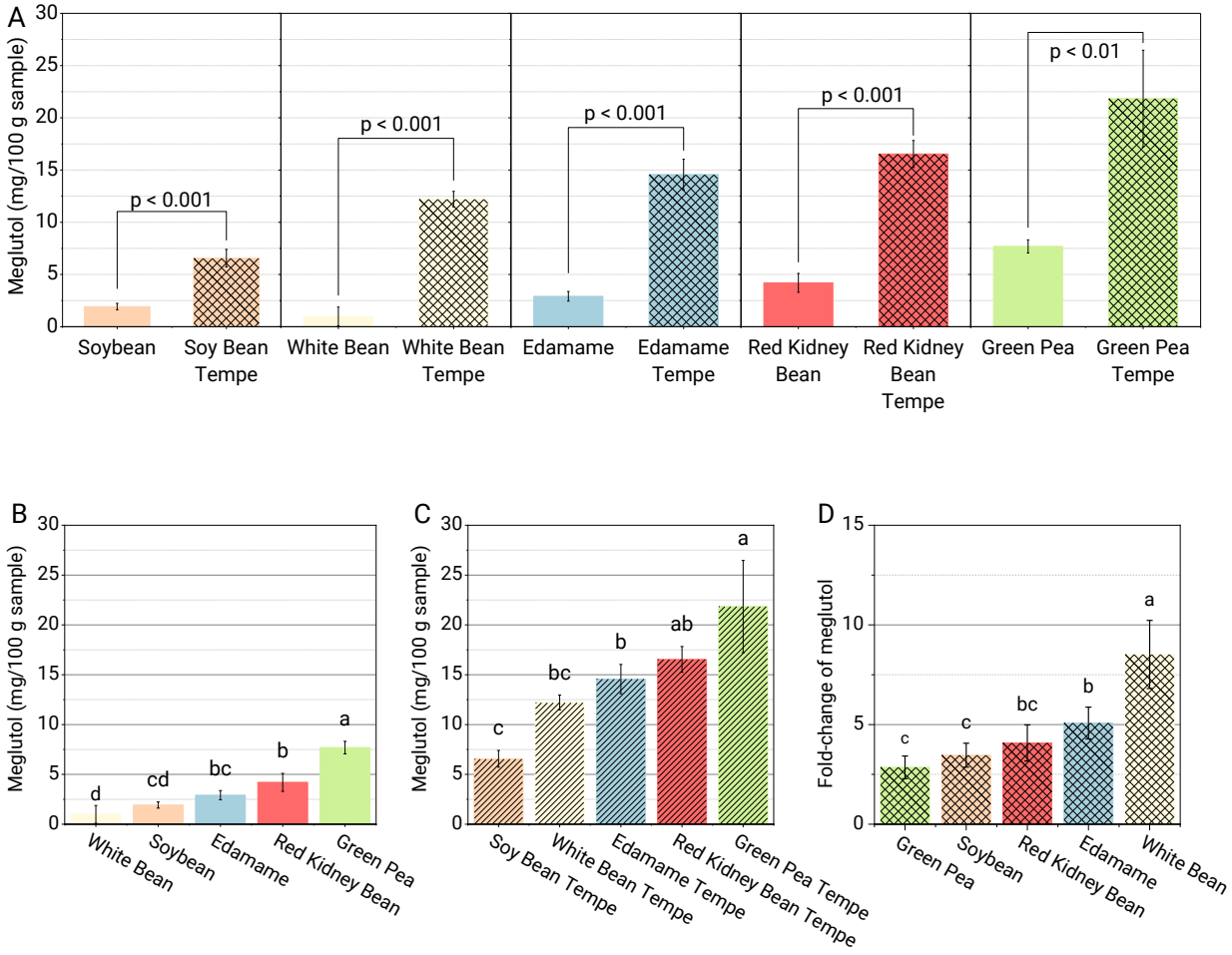


Figure 4.5. Modulation of meglutol levels following tempe fermentation in different legumes. Meglutol concentration was quantified using GC-MS with a meglutol standard compound co-injection. Error bars represent ±1SD from the mean. Groups labeled with different letters are significantly different (Tukey adjustment, $P < 0.05$). **A)** meglutol concentrations before and after fermentation across all legumes. **B)** comparison of meglutol concentrations in legumes prior to tempe fermentation. **C)** comparison of meglutol

concentrations among different types of tempes after fermentation. **D)** fold-change in meglutol concentrations before and after fermentation in the different legumes.

Notably, the meglutol content of all tested legumes displayed a significant increase following tempe fermentation, as depicted in **Figures 4.5a-e**. Interestingly, regardless of their initial meglutol levels prior to fermentation, all alternative legume-based tempes exhibited higher meglutol content compared to traditional soybean tempe. Among the legumes tested, green pea contained the highest meglutol content, measuring 7.69 ± 0.63 mg meglutol/100 g bean, surpassing even the meglutol content of soybean tempe. Furthermore, green pea tempe demonstrated the highest meglutol content among the tested samples, with a concentration of 21.8 ± 4.6 mg meglutol/100 g tempe (**Table S8**). It is worth noting that while green pea tempe displayed the highest meglutol content, white bean exhibited the greatest increase in meglutol concentration following tempe fermentation.

All types of tempes fermented from alternative legumes tested resulted in higher meglutol concentrations compared to traditional soybean tempe. Among these alternatives, green pea tempe stood out with the highest concentration of meglutol, indicating its potential as an excellent dietary source of this compound. Interestingly, despite its high meglutol concentration, green pea tempe showed the lowest change in meglutol levels relative to its legume counterpart. In contrast, white bean tempe, which displayed the lowest meglutol concentration among the alternative tempe samples, exhibited the highest change in meglutol levels relative to its legume. These findings suggest that green peas may have lower bioavailability of meglutol precursors, thereby limiting the *Rhizopus* fungus' access to sufficient resources for producing high meglutol concentrations during fermentation.

Conversely, white beans may contain abundant and bio accessible meglutol precursors, enabling the *Rhizopus* fungus to effectively generate a substantial amount of meglutol during the fermentation process.

The consumption of tempe has been widely recognized for its positive impact on cardiovascular health, as it has been shown to effectively lower LDL-C concentration and total cholesterol levels(Afifah, Nabilah et al. 2020). Previous studies have primarily attributed these health benefits to the isoflavone and phenolic components present in tempe(Harun, Susanto et al. 2017, Astawan, Mardhiyyah et al. 2018, Huang, Wu et al. 2018, Ahnan‐Winarno, Cordeiro et al. 2021). However, our research revealed a novel finding – the high concentration of meglutol in tempe. This finding opens up new possibilities for generating hypotheses regarding the underlying mechanisms through which tempe exerts its positive effects on cardiovascular health.

4.4 Conclusion

In summary, this chapter provides a comprehensive exploration of meglutol in relation to food. Our investigation involved a thorough survey of the meglutol content in various foods previously documented to contain meglutol. Our findings reveal that tempe contains significantly higher meglutol levels compared to other known meglutol-containing foods. Building upon this discovery, we delved into the modulation of meglutol through tempe fermentation and implemented raw material substitution strategies to optimize meglutol concentrations in tempe. Our research demonstrates a significant increase in meglutol levels in soybean, edamame, white bean, red kidney bean, and green pea following tempe fermentation. Notably, green pea tempe exhibited the highest concentration of meglutol at 21.8±4.6 mg/100g, which is over ten times higher than that

found in regular soybean. This study highlights the potential of tempe, particularly green pea tempe, as a functional food rich in meglutol which may have significant implications for controlling LDL-C levels and promoting cardiovascular health.

Chapter 5

Conclusion and future perspectives

5.1 Conclusion

This research dissertation employs metabolomics-based techniques to investigate and analyze bioactive metabolites in tempe. The primary objectives include profiling the bioactive metabolites, characterizing their bioactive properties, and optimizing them using principles of food engineering. The initial step involves a comprehensive investigation of tempe's metabolome to identify bioactive metabolites. This analysis employs widely-targeted GC-MS, untargeted LC-MS, and targeted HPLC-PDA-MS techniques, resulting in the identification of 16 bioactive metabolites. Notably, this study highlights the presence of a hypolipidemic bioactive metabolite, 3-hydroxy-3-methylglutaric acid (meglutol), in tempe for the first time.

To provide a thorough characterization of meglutol's bioactive potential, its association with cardiovascular risk is examined through epidemiological analyses involving two well-characterized cohorts. A meta-analysis reveals a substantial correlation between lower baseline meglutol concentration (per 1-SD) and decreased LDL-C concentration (β [SE]: -5.5 [1.6, P = 0.0005), with a more pronounced effect observed in metabolically-compromised patient groups.

Furthermore, the concentrations of meglutol in various foods are investigated, and tempe stands out as having significantly higher levels compared to other known meglutol-containing foods. Subsequently, the study explores the modulation of meglutol during tempe fermentation and implements strategies for optimizing meglutol concentrations by

substituting raw materials. The findings show that green pea tempe exhibits the highest concentration of meglutol at 21.8±4.6 mg/100g, surpassing regular soybean tempe by over tenfold.

5.2 Future perspectives

The findings of this research shed light on the promising potential of meglutol as a food-based alternative for controlling LDL-C levels. The identification and characterization of meglutol as a hypolipidemic bioactive metabolite in tempe open new avenues for exploring its applications in cardiovascular health management. Further investigations can delve deeper into the mechanisms of action of meglutol, elucidating its interactions with lipid metabolism pathways and its effects on cholesterol homeostasis. These studies can provide a more comprehensive understanding of meglutol's therapeutic potential and contribute to the development of targeted interventions for dyslipidemia.

To further establish the health benefits of consuming tempe as a meglutol-rich functional food, well-designed human intervention studies are warranted. These studies can provide valuable evidence regarding the effects of meglutol on lipid profiles, cardiovascular health parameters, and overall metabolic status in human subjects. By conducting controlled trials with appropriate sample sizes and rigorous study designs, we can evaluate the efficacy and safety of meglutol-rich tempe consumption in diverse populations, including individuals with dyslipidemia, metabolic syndrome, or other cardiovascular risk factors.

References

Afifah, D. N., N. Nabilah, G. T. Supraba, S. N. Pratiwi, N. Nuryanto and M. Sulchan (2020). "The effects of tempeh gembus, an Indonesian fermented food, on lipid profiles in women with hyperlipidemia."

Ahnan-Winarno, A. D., L. Cordeiro, F. G. Winarno, J. Gibbons and H. Xiao (2021). "Tempeh: A semicentennial review on its health benefits, fermentation, safety, processing, sustainability, and affordability." Comprehensive Reviews in Food Science and Food Safety **20**(2): 1717-1767.

Ala-Korpela, M., A. J. Kangas and P. Soininen (2012). "Quantitative high-throughput metabolomics: a new era in epidemiology and genetics." Genome medicine **4**: 1-5.

Alongi, M. and M. Anese (2021). "Re-thinking functional food development through a holistic approach." Journal of Functional Foods **81**: 104466.

Andriankaja, O. M., J. J. Jiménez, F. J. Muñoz-Torres, C. M. Pérez, J. L. Vergara and K. J. Joshipura (2015). "Lipid-lowering agents use and systemic and oral inflammation in overweight or obese adult Puerto Ricans: the San Juan overweight adults longitudinal study (SOALS)." Journal of clinical periodontology **42**(12): 1090-1096.

Astawan, M., Y. S. Mardhiyyah and C. H. Wijaya (2018). "Potential of bioactive components in tempe for the treatment of obesity." Jurnal Gizi dan Pangan **13**(2): 79-86.

Astuti, M., A. Meliala, F. S. Dalais and M. L. Wahlqvist (2000). "Tempe, a nutritious and healthy food from Indonesia." Asia Pacific Journal of Clinical Nutrition **9**(4): 322-325.

Bachheti, R. K., L. A. Worku, Y. H. Gonfa, M. Zebeaman, D. Pandey and A. Bachheti (2022). "Prevention and treatment of cardiovascular diseases with plant

phytochemicals: A review." Evidence-Based Complementary and Alternative Medicine **2022**.

Bhupathiraju, S. N., A. H. Lichtenstein, B. Dawson-Hughes and K. L. Tucker (2011). "Adherence index based on the AHA 2006 diet and lifestyle recommendations is associated with select cardiovascular disease risk factors in older Puerto Ricans." The Journal of nutrition **141**(3): 460-469.

Bragagnolo, F. S., C. S. Funari, E. Ibáñez and A. Cifuentes (2021). "Metabolomics as a tool to study underused soy parts: In search of bioactive compounds." Foods **10**(6): 1308.

Bujang, A. and N. A. Taib (2014). "Changes on amino acids content in soybean, garbanzo bean and groundnut during pre-treatments and tempe making." Sains Malaysiana **43**(4): 551-557.

Caesar, L. K., J. J. Kellogg, O. M. Kvalheim and N. B. Cech (2019). "Opportunities and limitations for untargeted mass spectrometry metabolomics to identify biologically active constituents in complex natural product mixtures." Journal of natural products **82**(3): 469-484.

Chevli, P. A., B. I. Freedman, F.-C. Hsu, J. Xu, M. E. Rudock, L. Ma, J. S. Parks, N. D. Palmer and M. D. Shapiro (2021). "Plasma metabolomic profiling in subclinical atherosclerosis: the Diabetes Heart Study." Cardiovascular diabetology **20**(1): 1-12.

Cohen, S., T. Kamarck and R. Mermelstein (1983). "A global measure of perceived stress." Journal of health and social behavior: 385-396.

Esaki, H., H. Onozaki, S. Kawakishi and T. Osawa (1996). "New antioxidant isolated from tempeh." Journal of Agricultural and Food Chemistry **44**(3): 696-700.

Evans, A. M., C. D. DeHaven, T. Barrett, M. Mitchell and E. Milgram (2009). "Integrated, nontargeted ultrahigh performance liquid chromatography/electrospray ionization

tandem mass spectrometry platform for the identification and relative quantification of the small-molecule complement of biological systems." Analytical chemistry **81**(16): 6656-6667.

Farag, M. A., A. R. Khattab, A. Ehrlich, M. Kropf, A. G. Heiss and L. A. Wessjohann (2018). "Gas chromatography/mass spectrometry-based metabolite profiling of nutrients and antinutrients in eight lens and lupinus seeds (Fabaceae)." Journal of agricultural and food chemistry **66**(16): 4267-4280.

Ference, B. A., H. N. Ginsberg, I. Graham, K. K. Ray, C. J. Packard, E. Bruckert, R. A. Hegele, R. M. Krauss, F. J. Raal and H. Schunkert (2017). "Low-density lipoproteins cause atherosclerotic cardiovascular disease. 1. Evidence from genetic, epidemiologic, and clinical studies. A consensus statement from the European Atherosclerosis Society Consensus Panel." European heart journal **38**(32): 2459-2472.

Friedewald, W. T., R. I. Levy and D. S. Fredrickson (1972). "Estimation of the concentration of low-density lipoprotein cholesterol in plasma, without use of the preparative ultracentrifuge." Clinical chemistry **18**(6): 499-502.

Golovinskaia, O. and C.-K. Wang (2021). "Review of functional and pharmacological activities of berries." Molecules **26**(13): 3904.

Handajani, Y. S., Y. Turana, Y. Yogiara, N. T. Widjaja, T. P. Sani, G. A. M. Christianto and A. Suwanto (2021). "Tempeh consumption and cognitive improvement in mild cognitive impairment." Dementia and geriatric cognitive disorders **49**(5): 497-502.

Harun, I., H. Susanto and A. Rosidi (2017). "Pemberian tempe menurunkan kadar malondialdehyde (MDA) dan meningkatkan aktivitas enzim superoxide dismutase

(SOD) pada tikus dengan aktivitas fisik tinggi." Jurnal Gizi dan Pangan **12**(3): 211-216.

Haslam, D. E., L. Liang, D. D. Wang, R. S. Kelly, C. Wittenbecher, C. M. Pérez, M. Martínez, C.-H. Lee, C. B. Clish and D. T. Wong (2021). "Associations of network-derived metabolite clusters with prevalent type 2 diabetes among adults of Puerto Rican descent." BMJ Open Diabetes Research and Care **9**(1): e002298.

Hosaka, Y., K. Itoh, S. Matsutani, S. Kawate, A. Miura, Y. Mizoura, S. Yamada, H. Konno, E. Grave and K. Nagata (2021). "Fermented food Tempeh induces interleukin 12 and enhances macrophage phagocytosis." Journal of Food Biochemistry **45**(11): e13958.

Huang, Y.-C., B.-H. Wu, Y.-L. Chu, W.-C. Chang and M.-C. Wu (2018). "Effects of Tempeh Fermentation with Lactobacillus plantarum and Rhizopus oligosporus on Streptozotocin-Induced Type II Diabetes Mellitus in Rats." Nutrients **10**(9): 1143.

Iwatani, S. and N. Yamamoto (2019). "Functional food products in Japan: A review." Food Science and Human Wellness **8**(2): 96-101.

Joshipura, K. J., F. J. Muñoz-Torres, B. A. Dye, B. G. Leroux, M. Ramírez-Vick and C. M. Pérez (2018). "Longitudinal association between periodontitis and development of diabetes." Diabetes research and clinical practice **141**: 284-293.

Klimek-Szczykutowicz, M., A. Szopa and H. Ekiert (2020). "Citrus limon (Lemon) phenomenon—a review of the chemistry, pharmacological properties, applications in the modern pharmaceutical, food, and cosmetics industries, and biotechnological studies." Plants **9**(1): 119.

Kochman, J., K. Jakubczyk, J. Antoniewicz, H. Mruk and K. Janda (2020). "Health benefits and chemical composition of matcha green tea: A review." Molecules **26**(1): 85.

Koh, S., A. Jamaluddin, N. Alitheen, N. Mohd Ali, H. Mohd Yusof and K. Long (2012). "Nutritional values of tempe inoculated with different strains of Rhizopus: its γ–aminobutyric acid content and antioxidant property." J Trop Agric Food Sc **40**(2): 181-192.

Kuligowski, M., K. Pawłowska, I. Jasińska-Kuligowska and J. Nowak (2016). "Isoflavone composition, polyphenols content and antioxidative activity of soybean seeds during tempeh fermentation." CyTA - Journal of Food: 1-7.

Kuligowski, M., D. Sobkowiak, K. Polanowska and I. Jasińska-Kuligowska (2022). "Effect of different processing methods on isoflavone content in soybeans and soy products." Journal of Food Composition and Analysis **110**: 104535.

Lai, Z., H. Tsugawa, G. Wohlgemuth, S. Mehta, M. Mueller, Y. Zheng, A. Ogiwara, J. Meissen, M. Showalter, K. Takeuchi, T. Kind, P. Beal, M. Arita and O. Fiehn (2018). "Identifying metabolites by integrating metabolome databases with mass spectrometry cheminformatics." Nature Methods **15**(1): 53-56.

Lupien, P.-J., S. Moorjani, D. Brun and P. Bielmann (1979). "Effects of 3-Hydroxy-3-methylglutaric Acid on Plasma and Low-Density Lipoprotein Cholesterol Levels in Familial Hypercholesterolemia." The Journal of Clinical Pharmacology **19**(2-3): 120-126.

Marin, G. and R. J. Gamba (1996). "A new measurement of acculturation for Hispanics: The Bidimensional Acculturation Scale for Hispanics (BAS)." Hispanic Journal of Behavioral Sciences **18**(3): 297-316.

Matsuo, M. (2006). "Preparation and preferences of peanut-tempeh, peanuts fermented with Rhizopus oligosporus." Food science and technology research **12**(4): 270-274.

McDermott, M. M., A. L. Mandapat, A. Moates, M. Albay, E. Chiou, L. Celic and P. Greenland (2003). "Knowledge and attitudes regarding cardiovascular disease risk and prevention in patients with coronary or peripheral arterial disease." Archives of internal medicine **163**(18): 2157-2162.

Megson, I. L., P. D. Whitfield and I. Zabetakis (2016). "Lipids and cardiovascular disease: where does dietary intervention sit alongside statin therapy?" Food & Function **7**(6): 2603-2614.

Misra, S., P. Pandey and H. N. Mishra (2021). "Novel approaches for co-encapsulation of probiotic bacteria with bioactive compounds, their health benefits and functional food product development: A review." Trends in Food Science & Technology **109**: 340-351.

Mo, H., S. Kariluoto, V. Piironen, Y. Zhu, M. G. Sanders, J.-P. Vincken, J. Wolkers-Rooijackers and M. R. Nout (2013). "Effect of soybean processing on content and bioaccessibility of folate, vitamin B12 and isoflavones in tofu and tempe." Food chemistry **141**(3): 2418-2425.

Mohammadnezhad, M., T. Mangum, W. May, J. J. Lucas and S. Ailson (2016). "Common modifiable and non-modifiable risk factors of cardiovascular disease (CVD) among pacific countries." World Journal of Cardiovascular Surgery **6**(11): 153.

Njoku, H., C. Ofuya and J. Ogbulie (1991). "Production oftempeh from the African yam bean (Sphenostylis stenocarpa Hams)." Food Microbiology **8**(3): 209-214.

Nout, M. and J. Kiers (2005). "Tempe fermentation, innovation and functionality: update into the third millenium." Journal of applied microbiology **98**(4): 789-805.

Paffenbarger Jr, R. S., R. T. Hyde, A. L. Wing, I.-M. Lee, D. L. Jung and J. B. Kampert (1993). "The association of changes in physical-activity level and other lifestyle

characteristics with mortality among men." New England journal of medicine **328**(8): 538-545.

Paffenbarger Jr, R. S., A. L. Wing and R. T. Hyde (1995). "Physical activity as an index of heart attack risk in college alumni." American Journal of Epidemiology **142**(9): 889-903.

Pang, Z., J. Chong, G. Zhou, A. David, L. Chang, M. Barrette, C. Gauthier, P.-É. Jacques, S. Li and J. Xia (2021). "MetaboAnalyst 5.0: narrowing the gap between raw spectra and functional insights." Nucleic Acids Research **49**(W1): W388-W396.

Penson, P. E., M. Pirro and M. Banach (2020). "LDL-C: lower is better for longer—even at low risk." BMC medicine **18**: 1-6.

Pérez, C. M., F. Muñoz, O. M. Andriankaja, C. S. Ritchie, S. Martínez, J. Vergara, J. Vivaldi, L. López, M. Campos and K. J. Joshipura (2017). "Cross-sectional associations of impaired glucose metabolism measures with bleeding on probing and periodontitis." Journal of clinical periodontology **44**(2): 142-149.

Pertiwi, M. G. P., Y. Marsono and R. Indrati (2020). "In vitro gastrointestinal simulation of tempe prepared from koro kratok (Phaseolus lunatus L.) as an angiotensin-converting enzyme inhibitor." Journal of food science and technology **57**(5): 1847-1855.

Portanguen, S., P. Tournayre, J. Sicard, T. Astruc and P.-S. Mirade (2019). "Toward the design of functional foods and biobased products by 3D printing: A review." Trends in Food Science & Technology **86**: 188-198.

Puteri, N. E., M. Astawan, N. S. Palupi, T. Wresdiyati and Y. Takagi (2018). "Characterization of biochemical and functional properties of water-soluble tempe flour." Food Science and Technology **38**: 147-153.

Putri, S. P., Y. Nakayama, F. Matsuda, T. Uchikata, S. Kobayashi, A. Matsubara and E. Fukusaki (2013). "Current metabolomics: practical applications." Journal of bioscience and bioengineering **115**(6): 579-589.

Rahmawati, D., M. Astawan, S. P. Putri and E. Fukusaki (2021). "Gas chromatography-mass spectrometry-based metabolite profiling and sensory profile of Indonesian fermented food (tempe) from various legumes." Journal of Bioscience and Bioengineering **132**(5): 487-495.

Ribeiro, M., J. Mandarino, M. CarrÃO-Panizzi, M. Oliveira, C. Campo, A. Nepomuceno and E. Ida (2006). "β-glucosidase activity and isoflavone content in germinated soybean radicles and cotyledons." Journal of Food Biochemistry **30**(4): 453-465.

Rivas-Tumanyan, S., L. S. Pacheco, D. E. Haslam, L. Liang, K. L. Tucker, K. J. Joshipura and S. N. Bhupathiraju (2022). "Novel Plasma Metabolomic Markers Associated with Diabetes Progression in Older Puerto Ricans." Metabolites **12**(6): 513.

Sánchez-Magana, L. M., E. O. Cuevas-Rodríguez, R. Gutiérrez-Dorado, A. E. Ayala-Rodríguez, A. Valdez-Ortiz, J. Milán-Carrillo and C. Reyes-Moreno (2014). "Solid-state bioconversion of chickpea (Cicer arietinum L.) by Rhizopus oligosporus to improve total phenolic content, antioxidant activity and hypoglycemic functionality." International journal of food sciences and nutrition **65**(5): 558-564.

Sattar, A. and M. Akhtar (1990). "Irradiation and germination effects on phytate, protein and amino acids of soybean." Plant Foods for Human Nutrition **40**: 185-194.

Simonne, A., M. Smith, D. Weaver, T. Vail, S. Barnes and C. Wei (2000). "Retention and changes of soy isoflavones and carotenoids in immature soybean seeds (Edamame) during processing." Journal of Agricultural and Food Chemistry **48**(12): 6061-6069.

Sitanggang, A. B., M. Lesmana and S. Budijanto (2020). "Membrane-based preparative methods and bioactivities mapping of tempe-based peptides." Food Chemistry **329**: 127193.

Stewart, D., G. J. McDougall, J. Sungurtas, S. Verrall, J. Graham and I. Martinussen (2007). "Metabolomic approach to identifying bioactive compounds in berries: advances toward fruit nutritional enhancement." Molecular Nutrition & Food Research **51**(6): 645-651.

Stodolak, B., A. Starzyńska-Janiszewska, M. Mika and A. Wikiera (2020). "Rhizopus oligosporus and Lactobacillus plantarum co-fermentation as a tool for increasing the antioxidant potential of grass pea and flaxseed oil-cake tempe." Molecules **25**(20): 4759.

Tadesse, S. A. and S. A. Emire (2020). "Production and processing of antioxidant bioactive peptides: A driving force for the functional food market." Heliyon **6**(8): e04765.

Tamam, B., D. Syah, M. T. Suhartono, W. A. Kusuma, S. Tachibana and H. N. Lioe (2019). "Proteomic study of bioactive peptides from tempe." Journal of bioscience and bioengineering **128**(2): 241-248.

Team, R. C. (2013). "R: A language and environment for statistical computing."

Thiombiano, B., E. Gontier, R. Molinie, P. Marcelo, F. Mesnard and R. Dauwe (2020). "An untargeted liquid chromatography–mass spectrometry-based workflow for the structural characterization of plant polyesters." The Plant Journal **102**(6): 1323-1339.

Tucker, K. L., L. A. Bianchi, J. Maras and O. I. Bermudez (1998). "Adaptation of a food frequency questionnaire to assess diets of Puerto Rican and non-Hispanic adults." American journal of epidemiology **148**(5): 507-518.

Tucker, K. L., J. Mattei, S. E. Noel, B. M. Collado, J. Mendez, J. Nelson, J. Griffith, J. M. Ordovas and L. M. Falcon (2010). "The Boston Puerto Rican Health Study, a longitudinal cohort study on health disparities in Puerto Rican adults: challenges and opportunities." BMC Public health **10**(1): 1-12.

Xu, Z., X. Chen, X. Lu, B. Zhao, Y. Yang and J. Liu (2021). "Integrative analysis of transcriptome and metabolome reveal mechanism of tolerance to salt stress in oat (Avena sativa L.)." Plant Physiology and Biochemistry **160**: 315-328.

Yousri, N. A., K. Suhre, E. Yassin, A. Al-Shakaki, A. Robay, M. Elshafei, O. Chidiac, S. C. Hunt, R. G. Crystal and K. A. Fakhro (2022). "Metabolic and Metabo-Clinical Signatures of Type 2 Diabetes, Obesity, Retinopathy, and Dyslipidemia." Diabetes **71**(2): 184-205.

Yuliana, N. D., A. Khatib, Y. H. Choi and R. Verpoorte (2011). "Metabolomics for bioactivity assessment of natural products." Phytotherapy Research **25**(2): 157-169.

Supplementary Materials

Table S1. Sample details

Species	**Region of origin**	**Harvest year**	**Producer**
Glycine max	Tohoku, Yamagata, Japan	2022	Suzuya Kokumotsu

Table S2. List of the annotated metabolites from the GC-MS analysis

2-Aminoethanol	Glutamic acid	Pipecolic acid
3,4-Dihydroxybenzoic acid	Glutamine	Proline
3-Hydroxy-3-Methylglutaric acid	Glyceric acid	Putrescine
3-Hydroxyanthranilic acid	Glycerol	Pyroglutamic acid
4-Aminobutyric acid	Glycine	Raffinose
Alanine	Glycitein	Saccharic acid
Allantoic acid	Histidine	Serine
Allantoin	Hypoxanthine	Sorbitol
Asparagine	Inositol	Succinic acid
Aspartic acid	Isocitric acid	Sucrose
beta-Alanine	Isoleucine	Threonine
Cadaverine	Leucine	Thymine
Citric acid	Lysine	Trehalose
Cysteine	Malic acid	Tryptophan

Daidzein	Maltitol	Turanose
Fructose	Maltose	Tyramine
Fumaric acid	Mannitol	Tyrosine
Galactinol	Meso erythritol	Valine
Galactose	Methionine	Xanthine
Genistein	N-acetyl-alpha-D-glucosamine 1-phosphate	Xylitol
Gluconic acid	Ornithine	Xylonic acid
Glucono-1,5-lactone	Paeoniflorin	
Glucosamine	p-Coumaric acid	
Glucose	Phenylalanine	
Glucose 6-phosphate	Phosphate	

Table S3. Isoflavone peak assignments in the HPLC-PDA-MS analysis

Isoflavone	**RT (min)**	**$[M+H]^+$ *(m/z)***	**MS fragment ion *(m/z)***
Daidzin	7.166	417	255
Glycitin	8.297	447	285
Genistin	12.253	433	271
Malonyl daidzin	13.031	503	255
Malonyl genistin	16.28	519	271

Malonyl glycitin	13.599	533	285
Acetyl daidzin	15.789	459	255
Acetyl genistin	19.533	475	271
Acetyl glycitin	16.428	489	285
Daidzein	17.994	255	-
Glycitein	19.206	285	-
Genistein	22.762	271	-

Table S4. Characteristics of study participants in the BPRHS and SOALS at baseline*

	BPRHS	**SOALS**
n	636	1,011
Age, years	57.1 (6.9)	50.7 (6.8)
Sex, % Female	73	73
Type 2 Diabetes, %	52	8
Income**	18 (18)	46
Perceived stress score	23.5 (9.2)	12.6 (6.8)
Acculturation, %	18.1 (6.4)	-
Diet Quality***	43.5 (9.4)	-
Physical Activity Level****		
Inactive	46	-
Low Activity	50	45
Medium Activity	3	55
High Activity	1	-
Alcohol consumption		
None in the past year	60	56
Moderate	34	42
Heavy	5	2
Education level, % greater than high school	14	45
Smoking status, %		
Never	46	64
Former	30	18
Current	23	18
Hypertension medication, %	55	11
Lipid-lowering medication, %	43	13
LDL-C, mg/dl	108 (34)	123 (33)
HDL-C, mg/dl	45 (12)	48 (13)
TG, mg/dl, geometric mean (IQR)	141 (96)	132 (81)
WC, cm	101.9 (14.9)	106.3 (14.1)
BMI, kg/m^2	32.1 (6.7)	33.3 (6.2)
Insulin, µIU/mL, geometric mean (IQR)	13.6 (10.7)	9.2 (7.0)
Glucose, mg/dl	120 (50)	96 (20)
Glycosylated hemoglobin, %	7.0 (1.7)	5.8 (0.6)
HOMA-IR	5.9 (9.7)	2.6 (1.8)
CRP, mg/L	6.3 (8.7)	5.9 (6.3)

Abbreviations: BPRHS, Boston Puerto Rican Health Study; BMI, body mass index; CRP, C-reactive protein; HCHS/SOL, Hispanic Community Health Study/Study Of Latinos; HDL-C, high-density lipoprotein cholesterol; HOMA-IR, homeostatic model assessment of insulin resistance; LDL-C, low-density lipoprotein cholesterol; SOALS, San Juan Overweight Adult Longitudinal Study; TG, triglyceride; WC, waist circumference.

Table S5. Legumes used to make alternative tempes

Legume	Species	Region of origin
Soybean	*Glycine max*	Indonesia
Red kidney bean (うずら豆)	*Phaseolus vulgaris*	Ontario, Canada
White bean (白いんげん豆)	*Phaseolus vulgaris*	Hokkaido, Japan
Green pea (青豌豆)	*Pisum sativum*	United Kingdom
Edamame (秘伝豆)	*Glycine max*	Yamagata, Japan

Table S6. Meglutol concentrations of foods. Quantification was performed using GC-MS with meglutol annotation confirmed using standard compound co-injection.

Food	**Meglutol (mg/100g sample)**		
Flaxseed	1.07	±	0.54
Soybean	**1.92**	**±**	**0.31**
Lupinus Seed	1.94	±	0.22
Oats	1.98	±	0.28
Lentils	2.32	±	0.48
Orange	3.48	±	0.78
Tempe	**6.56**	**±**	**0.84**

Table S7. Metabolites annotated from soybean, white bean, red kidney bean, green pea, edamame, and their respective resulting tempes. Metabolome profiling was performed using GC-MS with meglutol annotation confirmed using standard compound co-injection.

2,6-Pyridinedicarboxylic acid	Glutamine	Oleic acid	Tyrosine
2-Aminoadipic acid	Glyceric acid	Oxalate	Uracil
2-Aminobutyric acid	Glycerol	Pantothenate	Urea
2-Aminoethanol	Glycine	Phenylalanine	Uric acid
3,4-Dihydroxybenzoate	Glycolic acid	Plamitic acid	Valine
3-Hydroxy-3-methylglutarate	Histidine	Proline	Xylitol
3-Hydroxyanthranilic acid	Inositol	Propyleneglycol	Xylonic acid
4-Aminobutyric acid	Isocitric acid	Putrescine	
4-Hydroxyphenethyl alcohol	Isoleucine	Pyroglutamic acid	
Alanine	Lactic acid	Ribitol	
Asparagine	Lauric acid	Ribose	
Aspartic acid	Leucine	Serine	
b-Alanine	Lysine	Sorbitol	
Cadaverine	Malic acid	Sorbose	
Citric acid	Malonic acid	Stearic acid	
Cysteine	Maltotriose	Threonine	
Fumaric acid	Melezitose	Thymine	
Gentiobiose	Melibiose	Trehalose	
Glucarate	Methionine	Tryptophan	
Glutamic acid	Myristic acid	Turanose	

Table S8. Meglutol concentrations of various legumes and tempes. Quantification was performed using GC-MS with meglutol annotation confirmed using standard compound co-injection

	Food	Meglutol (mg/100g sample)
Legume	Soy Bean	1.92 ± 0.31
	Red Kidney Bean	4.20 ± 0.90
	White Bean	0.99 ± 0.89
	Green Pea	7.69 ± 0.63
	Edamame	2.91 ± 0.46
Tempe	Soy Bean Tempe	6.56 ± 0.84
	Red Kidney Bean Tempe	16.54 ± 1.30
	White Bean Tempe	12.21 ± 0.75
	Green Pea Tempe	21.84 ± 4.63
	Edamame Tempe	14.56 ± 1.47

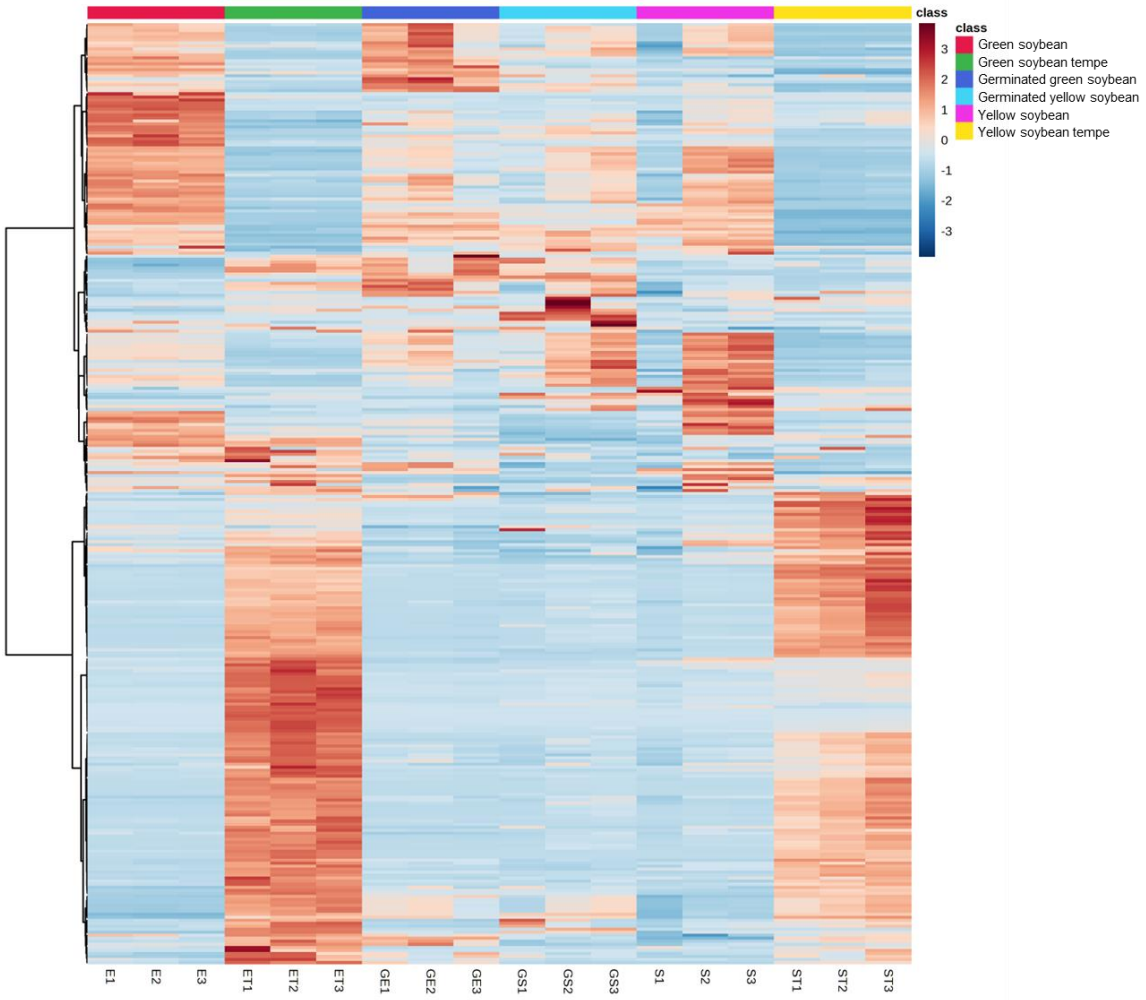


Figure S1. Heatmap highlighting the metabolites identified from the metabolomics analyses separating analyzed samples based on their metabolome

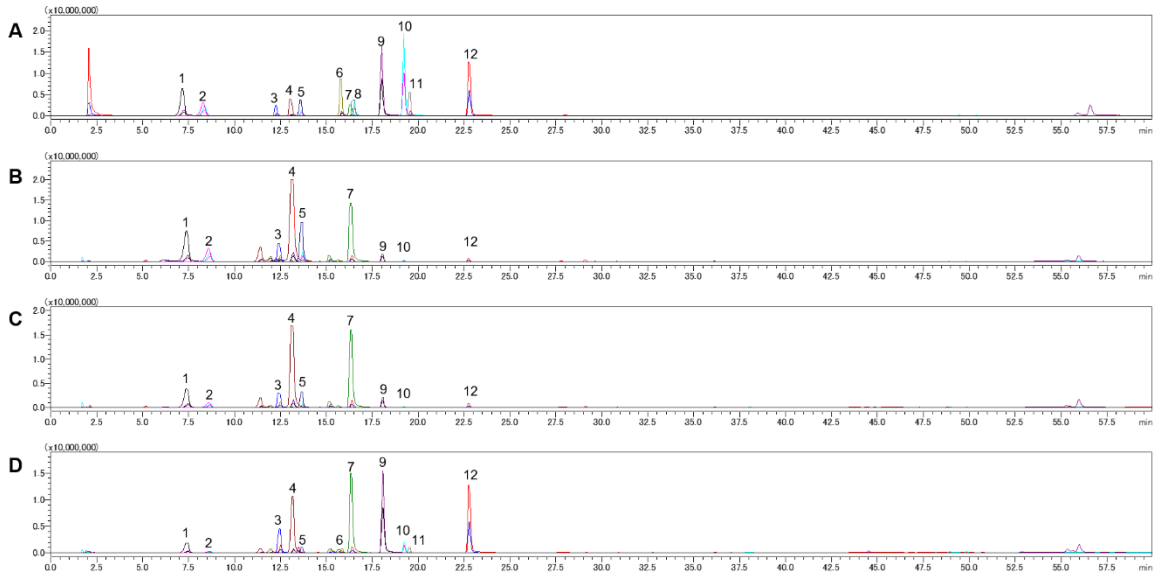


Figure S2. Representative chromatograms from the HPLC-PDA-MS analysis. **A)** isoflavone mixed standard; **B)** immature (green) soybean; **C)** germinated immature (green) soybean; **D)** immature (green) soybean tempe. Soy isoflavones annotated include: 1) daidzin; 2) glycitin; 3) genistin; 4) malonyl daidzin; 5) malonyl glycitin; 6) acetyl daidzin; 7) malonyl genistin; 8) acetyl glycitin; 9) daidzein; 10) glycitein; 11) acetyl genistin; 12) genistein.

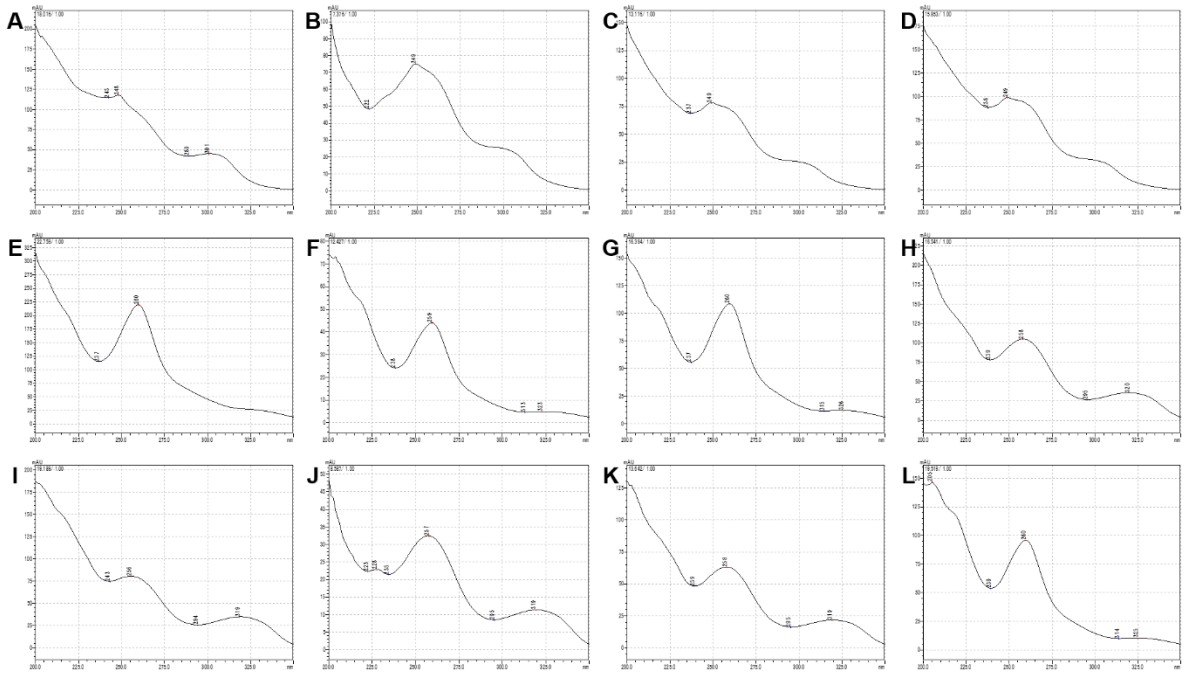


Figure S3. PDA spectral data of the 12 detected soy isoflavones from the HPLC-PDA-MS analysis. **A)** daidzein; **B)** daidzin; **C)** malonyl daidzin; **D)** acetyl daidzin; **E)** genistein; **F)** genistin; **G)** malonyl genistin; **H)** acetyl genistin; **I)** glycitein; **J)** glycitin; **K)** malonyl glycitin; **L)** acetyl glycitin.

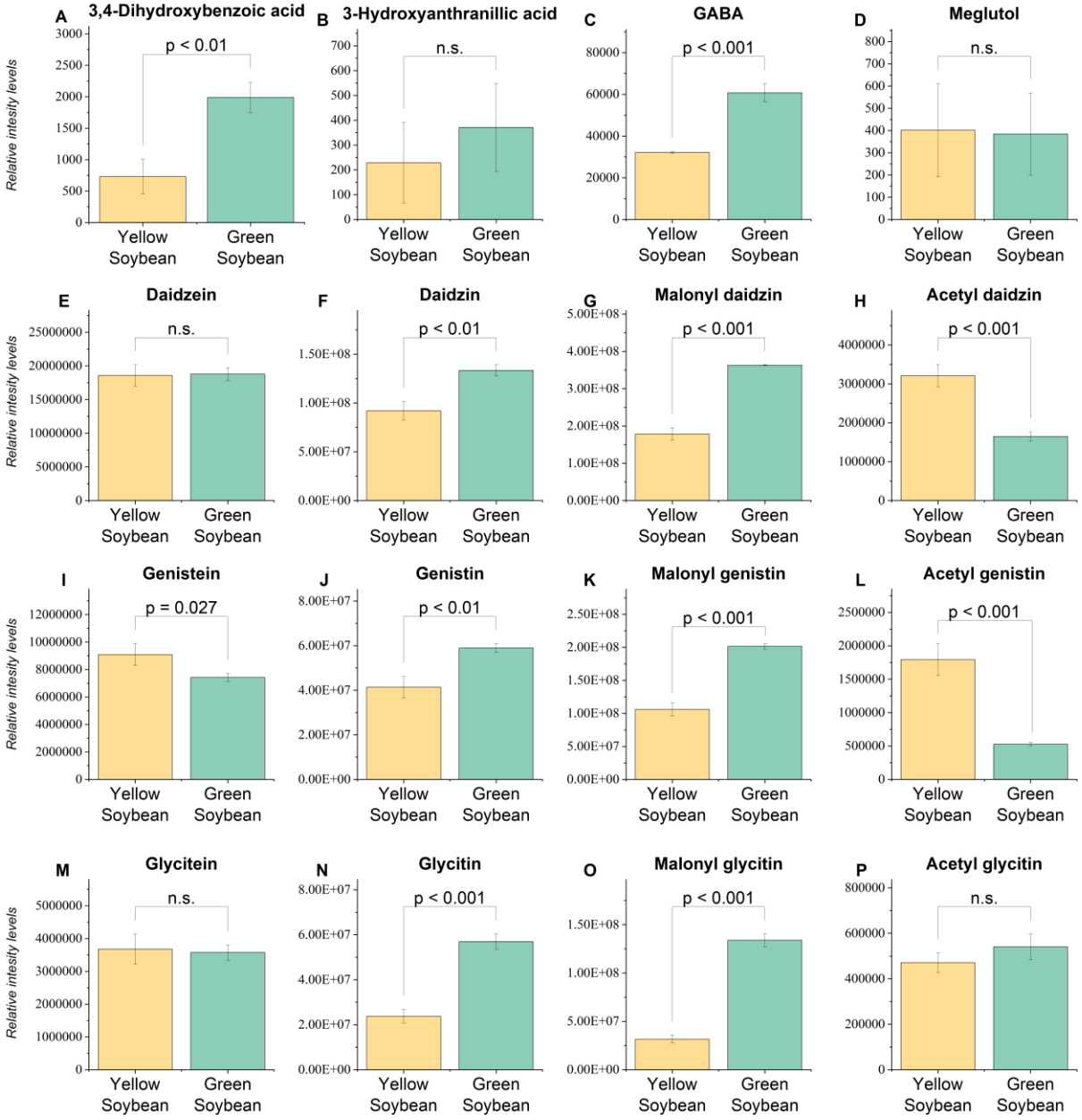


Figure S4. Bioactive metabolite levels in mature (yellow) and immature (green) soybean samples, including: **A)** 3,4-hydroxybenzoic acid; **B)** 3-hydroxyanthranillic acid; **C)** GABA; **D)** meglutol; **E)** daidzein; **F)** daidzin; **G)** malonyl daidzin; **H)** acetyl daidzin; **I)** genistein;

J) genistin; **K)** malonyl genistin; **L)** acetyl genistin; **M)** glycitein; **N)** glycitin; **O)** malonyl glycitin; **P)** acetyl glycitin. Vertical axis represents relative intensity levels. Error bars indicate ±1SD from the measured mean. Significance levels are measured using Tukey's test.