SHIMADZU
Excellence in Science

GC-MS GCMS-TQ™8040 NX

Application News

Metabolomic differential analysis of gene-mutated Drosophila using GC/MS

Dr. S. Kashio*1, E. Shimbo*2, Y. Yamada*2, Y. Nakagawa*2
*1: Genetics laboratory, Tokyo University *2: Shimadzu Corporation

User Benefits

- Obtained GC/MS data can be easily visualized by Multi-omics Analysis Package as the software has an intuitive and user-friendly interface.
- Statistical techniques such as principal component analysis, hierarchical clustering, volcano plot and metabolic maps are available to find differences among samples.

■ Introduction

A genetic mutation is a change or defect in the DNA sequence in a cell, which is divided into congenital and acquired mutations. Congenital mutations are mutations found in all the cells that make up the body and are passed down through the generations from parent to child. Acquired mutations, on the other hand, are caused by lifestyle factors such as smoking or environmental factors such as ultraviolet radiation.

Naturally occurring genetic mutations can affect cell division and proliferation, causing cancer or leading to serious diseases caused by metabolic abnormalities. On the other hand, technologies to artificially modify genes have been applied to genetically modified foods, and it is possible to insert a gene from another organism or, conversely, to suppress the function of the original gene.

In addition, because the type and concentration of metabolites that are essential for life support are influenced by an individual's health and constitution, they are attracting attention as biomarkers that show signs of various diseases. Since metabolites are derived from genes in addition to lifestyle, the integrated analysis of metabolites and genomic information is regarded as an important basic research that may lead to the development of drugs to treat diseases and personalized medicine [(1)]. That's where techniques to study the effects of genetic mutations on metabolites become important.

In this application, long-lived gene mutations were generated in yellow Drosophila and metabolites in wild and mutant Drosophila were measured by gas chromatograph mass spectrometer (GC-MS). the Multi-omics Analysis Package was employed to find differences between wild and mutant strains from the measurement results.

Multi-omics Analysis Package is an analysis software with data visualization features such as multivariate analysis (e.g., principal component analysis, class cluster analysis), volcano plots, and metabolic maps. This application news presents examples of analyses using these statistical methods to compare wild and genetically mutated Drosophila and analyzes differences in metabolites.

Fig. 1 GCMS-TQ™8040 NX (Left) and Multi-omics Analysis Package (Right)

■ Experimental

In this application, 50 yellow Drosophila were used for analysis. Of the 50 flies, 20 were wild and the remaining 30 were genetically mutated. One sample was formed with 10 flies and a total of 5 samples were prepared : 2 wild type samples and 3 genetic variant samples.

Drosophila were homogenized by grinding. Metabolites were extracted by shaking the homogenized samples with methanol: water: chloroform (2.5:1:1) extract. After the water phase was separated, it was centrifugally concentrated to remove methanol, and the remaining water phase was dried in a lyophilizer.

Table 1 Instrument Configurations

GC-MS	: GCMS-TQ8040 NX
Auto Injector	: AOC-20i Plus
Auto Sampler	: AOC-20s Plus
Analytical Column	: 5 type (30 m × 0.25 mm I.D., df=1.00 µm)
Glass Insert	: Split liner with wool

Table 2 Analytical Conditions

GC	
Inlet temp.	: 280 ˚C
Injection Mode	: Split
Spit ratio	: 10
Carrier gas	: Helium
Control Mode	: Constant linear velocity (39.0 cm/s)
Column oven temp.	: 100 ˚C (4 min) → (10 ˚C /min) → 320 ˚C (11 min) Total 37.0 mins
Purge flow rate	: 5 mL/min
Sample Inj. volume	: 1 µL
MS	
Ion Source Temp.	: 200 ˚C
Interface temp.	: 280 ˚C
Measurement Mode	: MRM

The dried samples were treated by methoxamine derivatization with methoxamine HCl in pyridine and trimethylsilyl (TMS) derivatization with N-methyl-N-trimethylsilyl trifluoroacetamide (MSTFA). After derivatization, they were transferred to vials and measured by GC-MS.

The pretreatments used in this study are described in the Metabolomics Pretreatment Handbook, and hydrophilic metabolites can be extracted as in this application by pretreating according to the protocol [(2)].

Smart Metabolites Database Ver. 2 was used for GC-MS analysis. The equipment configuration and analysis conditions are shown in Tables 1 and 2, respectively. Using Multiple Reaction Monitoring (MRM) as a data collection mode, it is possible to measure 604 compounds with high sensitivity in 37 minutes of analysis time.

■ Results

Measurements using GC/MS and Smart Metabolites Database detected a total of about 300 metabolites in each sample. Fig. 2 shows a typical chromatogram measured by Scan and MRM. Using MRM, components that could not be detected by Scan were detected with high sensitivity. MRM also separates the peaks of compounds, making it easier to distinguish the peaks of target compounds and reducing analysis time.

The peak area values of the components obtained by measurement were output in CSV format from the data analysis software LabSolutions Insight. Data cleansing was performed on the output results, such as removing missing values and removing saturated compounds. For the 170 compounds obtained after cleansing the data, each peak area value was normalized by subtracting the average area value of the 5 samples and dividing by the standard deviation.

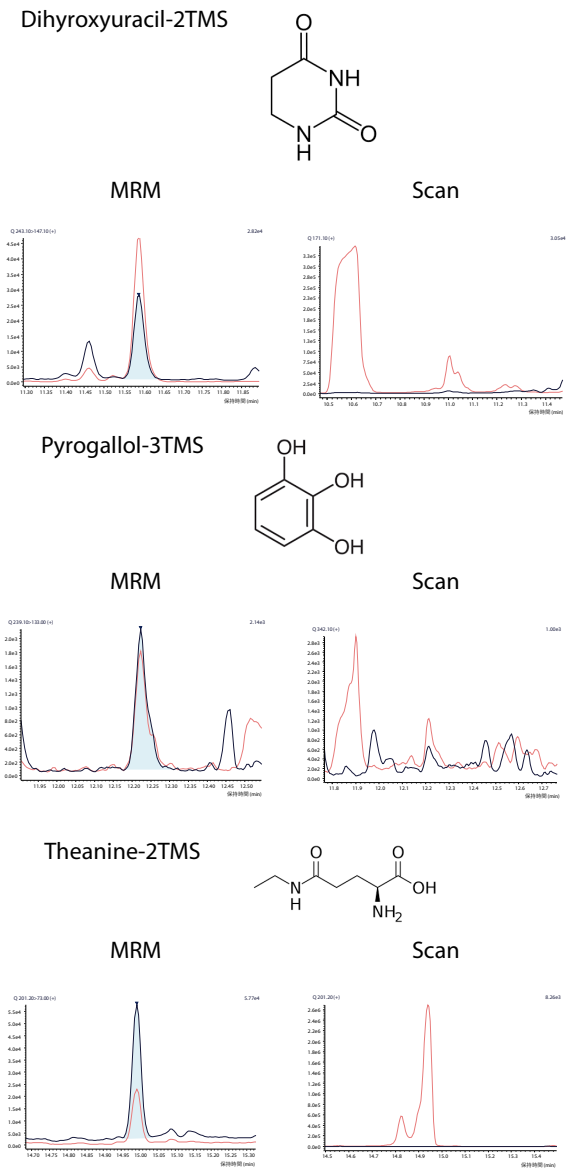

Fig. 2 Chromatogram comparison between MRM and Scan measurement modes

■ Multivariate analysis

The cleansed results were read in Multi-omics Analysis Package for multivariate analysis. This is a statistical method that uses variables (detected compounds) to visualize and infer results. In this application, it was visualized and inferred, by multivariate analysis using 170 component variables, whether there was a dominant difference in metabolites between wild and genetically mutated yellow Drosophila, and which compounds were characteristic in each sample group.

Principal component analysis, a type of multivariate analysis, is a statistical method that summarizes many quantitative explanatory variables into fewer indicators or composite variables. Fig. 3 shows the score plot for principal component analysis. PC1 on the horizontal axis contributed 55.4 %, and when combined with PC2 on the vertical axis, which contributed 18.9 %, the cumulative contribution was 74.3 %. This means that when 170 dimensions of information were compressed into 2 dimensions, only 25.7 % of the information was lost. The wild type and gene variant can be separated only by PC1 (1D) without relying on PC2 (2D). Using PC2, we found that gene-variant sample 3 had slightly different tendencies from other gene-variant samples 1 and 2. Similarly, the two wild type samples were found to have different tendencies. PC2 suggested that metabolites tended to differ from individual to individual within the same sample group.

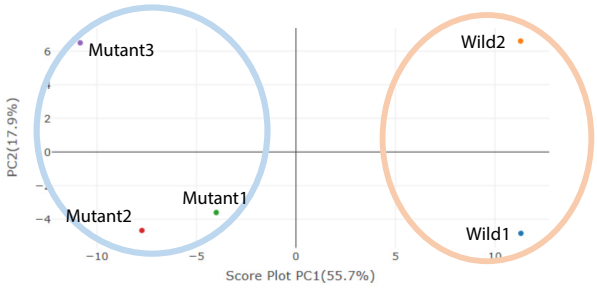

Fig. 3 Score plot by principal component analysis

Figure 4 shows the loading plot of principal component analysis. In conjunction with the score plot, it can be seen which compounds are characteristic of that sample. Being characteristic means that there is a large difference between the wild type and the genetic variant, making them "informative" compounds. Since there are more distinctive compounds in the mutant form than in the wild form, searching for information-rich variables in the wild form may be more efficient for finding differences. Wild-type, informative variables included porphobilinogen, anthranilic acid, urea, glyceraldehyde, uridine, adipic acid, homocysteine, kynurenine, and gluconic acid. On the other hand, genetic variants are thought to have many variables. Cysteine, uracil, dopamine, glutamic acid, and octopamine were considered as variables with high information, but due to the large number of candidates, it was necessary to resort to other multivariate analysis methods.

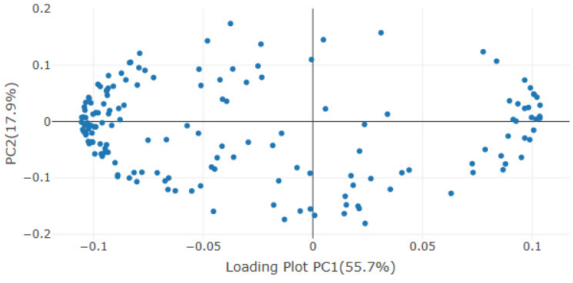

Fig. 4 Loading plot by principal component analysis

A volcano plot to visualize variables with high feature values is shown in Fig. 5. The horizontal axis of the volcano plot, called the fold change, shows how different a particular compound is between the wild type and the genetic variant. For example, a difference of 2 times Log2 equals +/- 2, and a difference of 4 times equals +/- 4. In Fig. 5, positive values represent increases in the gene variant and negative values represent decreases in the gene variant. The vertical axis shows -Log 10 (p-value) between the two groups, with the higher the point, the lower the p-value and the more significant the difference.

Compounds that were more common in the gene variant than in the wild type were cysteine, catechol, uridine, 2-hydroxyglutaric acid, and N-acetyltyrosine. Also, saccharopine, cystathionine, tryptophan, and 2-aminobutyric acid were found to be low in gene variants. In the loading plot of the principal component analysis in Fig. 4, kynurenine was decreased in the gene variant, but tryptophan and kynurenine were in the same metabolic pathway, suggesting a connection between the volcano plot and the results of the principal component analysis.

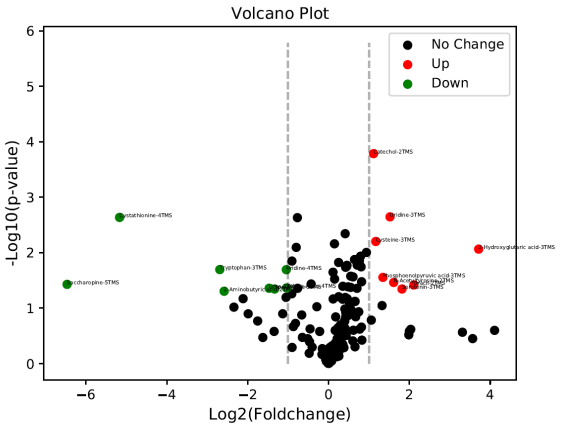

Fig. 5 Volcano plot between the wild and the genetically modified flies

The results of a class cluster analysis to verify the results of principal component analysis are shown in Fig. 6. The wild type and the gene variant were separated by hierarchy. We also found that sample 3 of the gene variant was slightly different from other gene variant samples.

The types of metabolites shown on the horizontal axis are broadly divided into two types (those that increase and those that decrease in the gene variant), and only a few compounds remained unchanged.

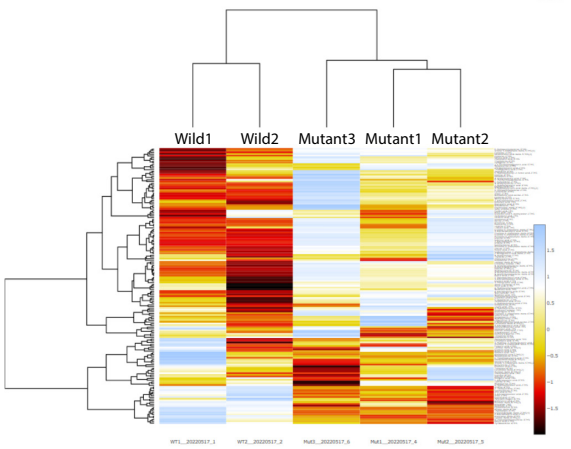

Fig. 6 Hierarchical cluster analysis between the wild and the genetically modified flies

Fig. 7 shows the differences between the samples in a metabolic map.

Metabolic maps visualized compounds that showed increases or decreases in gene variants. Kynurenine and 5-hydroxy-tryptophan, two metabolites of tryptophan, were significantly reduced in the gene variant than in the wild type.

Fig. 8 shows chromatograms of kynurenine and histamine. Like kynurenine, histamine is also a compound that has been confirmed to be reduced in a genetic variant. These compounds, which showed differences on the metabolic map, also showed differences on the chromatogram.

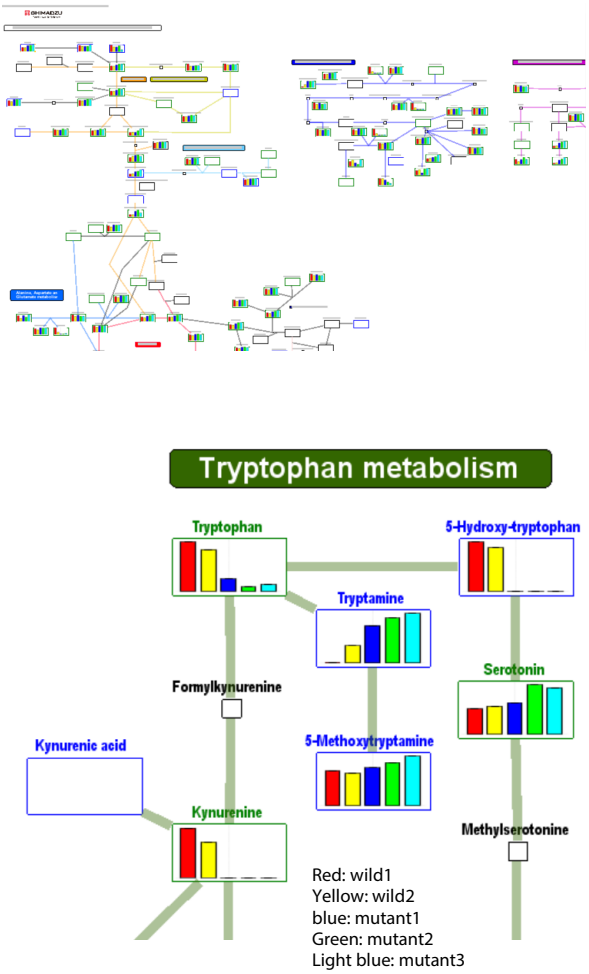

Fig. 7 Metabolic pathway overview (top) and an enlarged image (bottom)

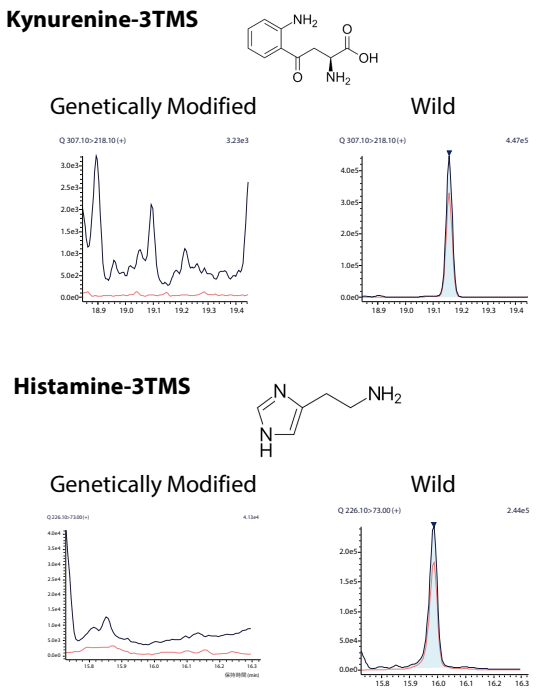

Fig. 8 Example chromatograms of potential indicator metabolites

■ Conclusion

Metabolites from wild and genetically mutated yellow Drosophila were analyzed by gas chromatograph mass spectrometer GCMS-TQ 8040 NX. The results of the comprehensive detection of metabolites were analyzed by principal component analysis, class cluster analysis, volcano plots and metabolic maps using Multi-omics Analysis Package. As a result, the metabolites that differed due to genetic mutations were visualized. Multi-omics Analysis Package is developed based on the tools available on the open research platform "GARUDA," which is mainly operated by the Research Institute of Systems Biology (SBI).

By using Multi-omics Analysis Package, it is possible to objectively judge the results obtained using GC-MS.

<Reference>
(1) Japan Agency for Medical Research and Development, Press Release A part of the effect of different genomes on metabolites, accessed 20th June 2022 https://www.amed.go.jp/news/release_20160818-02.html
(2) Shimadzu Corporation, Metabolomics Pretreatment Handbook, accessed 9th June 2022 https://www.an.shimadzu.co.jp/pdf/c146-2181.pdf

01-00410-EN First Edition: Mar. 2023

SHIMADZU
Shimadzu Corporation
www.shimadzu.com/an/

› Please fill out the survey

Related Products Some products may be updated to newer models.

› GCMS-TQ™8040 NX

Triple Quadrupole Gas Chromatograph Mass Spectrome…

› Multi-omics Analysis Package

LC/MS, GC/MS Data Analysis Software

Related Solutions

› Life Science

› Metabolomics

› Price Inquiry

› Product Inquiry

› Technical Service / Support Inquiry

› Other Inquiry