

understandingData

by Iwan Binanto

Submission date: 12-Aug-2020 03:23PM (UTC+0700)

Submission ID: 1368736104

File name: UnderstandingLCMSData-.docx (1.1M)

Word count: 3615

Character count: 20247

Understanding LCMS data for identification of chemical compounds contained in Rodent Tuber: Timeseries or not

Iwan Binanto^{1,2}, Harco Leslie Hendric Spits Warnars¹, Nesti Fronika Sianipar^{3,4}, Widodo Budiharto¹

¹Computer Science Department, BINUS Graduate Program, Doctor Computer Science Bina Nusantara University, Jakarta, Indonesia

²Informatics Department, Sanata Dharma University, Yogyakarta, Indonesia

³Food Technology Department, Faculty of Engineering, Bina Nusantara University, Jakarta, Indonesia

⁴Research Interest Group Biotechnology, Bina Nusantara University, Jakarta, Indonesia

ABSTRACT

Processing data for specific purposes requires an understanding of the data itself. A special investigation is needed to understand the data. Liquid Chromatography-Mass Spectrometry data are the results of material sample examination, which, in this case, was the sample of Rodent Tuber plants. The data need to be examined and understood whether they are timeseries or not, which is important for further processing. In this paper, we examined the data visually with graphs extracted from the data by human eyes and examination using the Augmented Dickey Fuller Test conducted by python programming with its library. From human eyes visual observations and computation using ADF Test, it can be concluded that Liquid Chromatography-Mass Spectrometry data are stationary timeseries.

Keywords: Understanding data, Chemical compound identification, LCMS data, Timeseries

1. INTRODUCTION

Liquid Chromatography-Mass Spectrometry (LCMS) is a technology for physical separation of thousands of more-comprehensive metabolites [1] - [6] because of its wide scope and high sensitivity to recognize chemical compounds in a sample of plant [7], [8].

Analyzing the plants using LCMS is essential since plant biochemistry is very rich and they have many semipolar compounds, including the main secondary metabolite groups, which can be separated and detected by LCMS [9].

LCMS signals provide information about the mass/charge ratio (m/z) of ionized molecules or fragments of molecules with their retention times. They need to be mapped into metabolites form to understand the produced biochemical processes. The mapping process is still an interpretation and it is complex for it requires an efficient and accurate identification. If it is done manually, it will take a lot of time [8], [10], [11]. This is a barrier in LCMS-based studies [7]. Thus, computing and algorithms are needed in order to simplify and speed up the analysis and interpretation of LC-MS data to extract the existing information [12].

Processing the LCMS data, whether it is done manually or computationally by using certain algorithms, depends on how the researcher understands the data itself; while understanding the data is not an easy process as it requires a lot of time, even for researchers in the domain of their own disciplines [13].

This paper is the product of the researcher's understanding of the LCMS dataset obtained from the studies of Rodent Tuber plants [14] - [21]. The dataset has a Retention Time feature. Thus, based on the definition of timeseries by Shumway et. al, which defines timeseries as a collection of random variables obtained in the exact order as they were in time [22], the dataset can be concluded as timeseries. However, the researchers are not convinced by this conclusion. Therefore, the visual and computational evidences, which show that the dataset is considered as timeseries, are provided since this is important for further processing.

1.1. Basic Principles of Liquid Chromatography-Mass Spectrometer (LC-MS)

Separating components in a mixture in which they are selectively distributed between two phases which are not amalgamated: the mobile phase through the stationary phase is a method called Chromatography [23]. The mobile phase is described as "fluid that seeps through or along a heap of stationary phases in a fixed direction". The fluid can be liquid, gas or supercritical fluid, while the stationary phase can be solid, gel or liquid. If the stationary phase is a liquid, it can be distributed to solids, which may not contribute to the separation process [24]. It is called Liquid Chromatography (LC) since the mobile phase is liquid to transport sample molecules through the stationary phase [12], [23].

[Type here]

The Chromatography process occurs as a result of repetitive absorption or discharge during the process of analyte's movement pass through the stationary phase. The differences in the distribution coefficients of each analyte in the sample causes separation [23].

Liquid Chromatography is a basic separation technique in chemistry and related natural sciences and a universal technique used for separating compounds from mixtures [25]. It separates various organic compounds safely, from tiny molecules of drug metabolites to peptides and proteins [26]. Commonly, recent Liquid Chromatography utilize the High Performance Liquid Chromatography (HPLC) instrument [27]. HPLC facilitates the analysis of chemical compounds with higher polarity and lower volatility in a wider range of mass without derivatization [28].

It is difficult on Liquid Chromatography to make sure specific compounds at a peak, even if there is only one compound in the sample. Mass spectrometry needs to be added as it will provide information at the peak about the mass of all compounds so it can be used to identify the compounds [27].

Mass Spectrometry is based on ions analysis that budge through a vacuum. This produces a mass spectrum, which informs about molecular weight, structure, identity, number, and purity of the sample. This is valuable information that can be utilized to help identify compounds [27], [29].

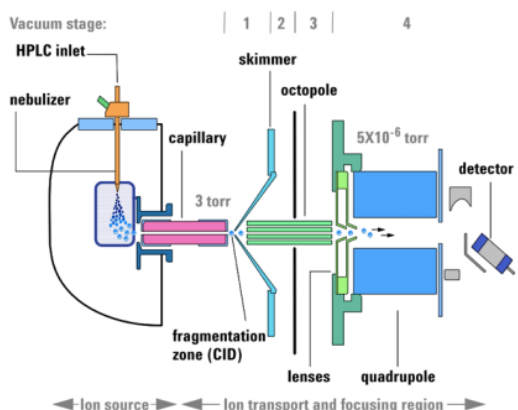


Figure 1. Diagram of LC-MS instrument with quadrupole for mass analyzer [29]

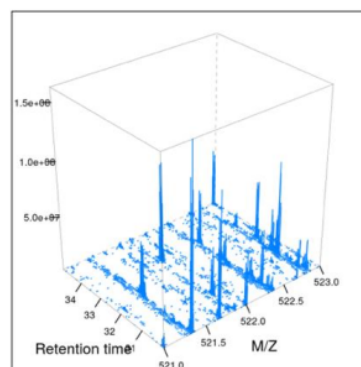


Figure 2. 3D Visualization LC-MS data [31]

Mass Spectrometers usually consist of three main parts: ion source, mass analyzer, and detector. During the process of converting sample molecules into ions, the mass analyzer separates them, in an electromagnetic field and/or in a tube while floating in the air before being measured by the detector. There are available some options for ion sources, which are Electrospray Ionization (ESI), Atmospheric Pressure Chemical Ionization (APCI), Atmospheric Pressure Photoionization (APPI), and Fast Atomic Bombardment (FAB). Mass analyzers can be categorized into quadrupoles, ion traps, Time-Of-Flight (TOF), Orbitrap, and Fourier Transform Ion Cyclotron (FTICR) [30]. The simplest and cheapest mass analyzer tend to Quadrupoles [26]. Mass Spectrometry technology enables development of simultaneous quantification of low and high molecular weight analytes in various concentrations and it is flexible and reliable method [25]. Briefly, Mass Spectrometry is utilized to measure ratio of mass to charge of charged particles namely mass-to-charge (m/z) ratio [27].

Liquid Chromatography combine with Mass Spectrometry (LC-MS) facilitates quantitative determination of compounds and more definite identification [24]. Figure 1 shows LC-MS device's general overview diagram with quadrupole. Output of the LC-MS instrument is a scattered 3D signal with features namely m/z, retention time, and intensity for each feature detected in peak [12] as shown in Figure 2.

1.2. Typhonium Flagelliforme (Rodent Tuber)

Typhonium Flagelliforme, known as Rodent Tuber has been recognized as a potent medicinal herb from the family Araceae (Arum) [32] - [34]. Southeast Asian countries, including India and China, utilize this

2 [Type here]

[Type here]

plant for years as alternative cancer therapy. Typhonium Flagelliforme is a potential health supplement for treatment lung, breast, rectum, liver, prostate, cervical and pancreatic cancer and leukemia [34] - [40].

Characteristics of this plant can be seen on its leaves. Typhonium Flagelliforme leaves vary greatly, from elliptical to ovoid, heart-shaped or arrow-like with 30 cm long petioles. The height of the plant can reach 30 cm with whitish and oval tubers [33], [34] as shown in figure 3.

This plant is very beneficial and has been used as local wisdoms. However, there is not much information related to its existing chemical compounds.



Figure 3. *Typhonium Flagelliforme* plant [17]

2. MATERIALS AND METHOD

2.1. Data Collection

The data of the study were obtained from research [14] - [21]. Those research produce 10 datasets from outputs of LCMS instrument. These datasets are in form of raw data proprietary, that the information is important and crucial, which only can be read by the instrument that produced them. In a computer operating system, this proprietary dataset is read as folders with several files that can be read by a text processing application although the important and crucial data still cannot be read.

These datasets need to be converted into .xlsx files so that they can be read by human and it becomes easier to analyze. There are two stages of doing this, which are (1) conversion of the raw data to open format, and (2) conversion of the open format to .xlsx.

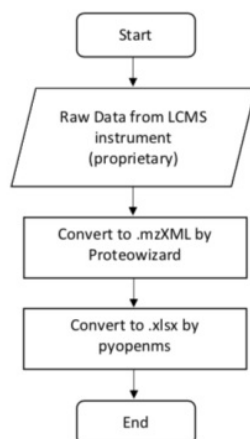


Figure 4. Data Collection Flowchart

The conversion to open format utilized open source namely Proteowizard version 3 for Windows, which is developed by Chambers, et.al. [41]. The ten raw datasets were successfully converted to .mzXML format. Here are names of the files U-BM 8-1_170818_4.mzXML, U-BM 8-2_170818_5.mzXML, U-BM 8-

[Type here]

[Type here]

3_170818_11.mzXML, U-BM 8-5_170818_6.mzXML, U-BM 8-8_170818_7.mzXML, U-KB_170818_8.mzXML, U-KB + KP_170818_9.mzXML, U-PM 8-2_170818_1.mzXML, U-PM 8-3_170818_2.mzXML, U-PM 8-4_170818_3.mzXML. All datasets, which are in the .mzXML open format, then were converted to .xlsx file using the Python language by pyopenms library, which is developed by Rost et.al. [42]. Flowchart of this data collection can be seen in Figure 4.

The conversion this dataset is used to get the data representation in matrix form with dimensions m/z , intensity, and retention time. This 3D matrix representation is often used for LC-MS data processing [28].

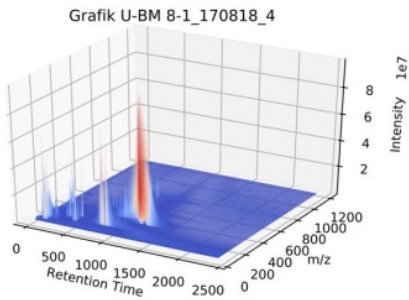


Figure 5. 3D visualization of U-BM 8-1_170818_4 dataset

2.2. Dataset Characteristics dan Visualization

All datasets that have been converted to .xlsx are processed using Python to obtained their features. All datasets are matrices with 3 columns, namely m/z , intensity, and retention time.

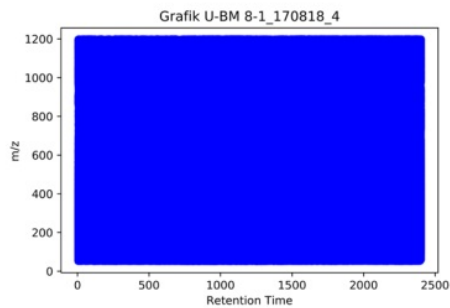


Figure 6a. m/z vs. Retention Time visualization of U-BM 8-1_170818_4 dataset

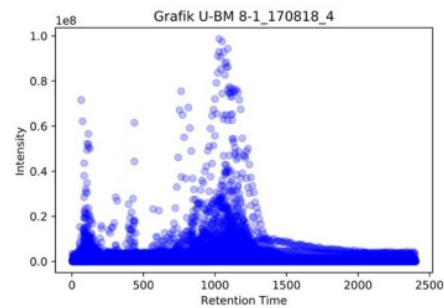


Figure 6b. Intensity vs. Retention Time visualization of U-BM 8-1_170818_4 dataset

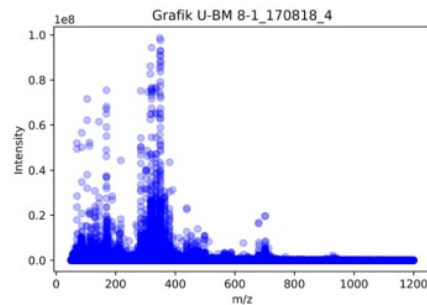


Figure 6c. Intensity vs. m/z visualization of U-BM 8-1_170818_4 dataset

[Type here]

The Retention Time column contains times that have periods with start on 5,022 seconds and a final on 2401 seconds. One time period is 5,022 seconds; so, there are 478 periods. A period has many m/z values with different intensities, so one period has a lot of Retention Time duplication. One Retention Time, for example 5.022 has 1224 duplication. For that reason, one period of Retention Time must be made single. Processing of Retention Time to be single is carried out by utilizing pivoting techniques which are done using Python. The Intensity column contains long integer numbers that describe the existing peaks; while the m/z column is the mass of the molecules present in one sample of the plant being tested.

This paper only presents one dataset as discussion material, namely U-BM dataset 8-1_170818_4.xlsx because all existing datasets have the same features which are Retention Time, m/z, and Intensity. There is only the number of rows that are different. Dataset selection for this discussion is done randomly, which means that there are no specific criteria. U-BM 8-1_170818_4.xlsx has 752,365 lines. To make it easier to understand this dataset, a visualization is needed. The visualization of the data is essential to understand and explore the data since it can explain the structure of the data and the existing patterns, as well as a form of data communication to a wide audience [31], [43]. Figure 5 is the data visualization from the existing features.

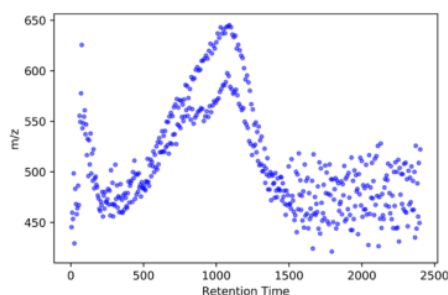


Figure 7a. Pivoted m/z vs. Retention Time visualization of U-BM 8-1_170818_4 dataset

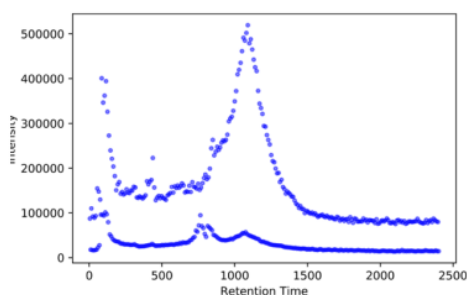


Figure 7b. Pivoted Intensity vs. Retention Time visualization of U-BM 8-1_170818_4 dataset

Besides the data visualization shown in Figure 5., 2D visualization with m/z vs. Retention Time, Intensity vs. Retention Time, and Intensity vs. m/z are done as shown in Figure 6a., Figure 6b., Figure 6c.

Figure 6a. shows that, from the beginning, during a period of Retention Time, many molecular masses were detected. Almost all periods have m/z from small to large so that the graph becomes full and difficult to analyze and understand. In Figure 6b, there is a high intensity at a certain Retention Time. The highest intensity was 98,656,256 at Retention Time in 1029.71 seconds. In figure 6b, the intensity is high at 98,656,256 for m/z 348,2371826171875. These numbers are obtained from the dataset extraction using Python.

3. RESULTS AND ANALYSIS

To facilitate visual analysis, graphics with Retention Time feature will be processed further, since they will be analyzed as timeseries or not. We can see that, from Figure 6a. and 6b., the Retention Time feature has a lot of duplication dipivoted to be single by calculating the mean of the value of other features. The results are shown in Figure 7a. and 7b.

In Figure 7a. and 7b. shows that data in this dataset were obtained over time. So, the dataset is timeseries based on the definitions from Shumway et. al. [22].

```
Hasil perhitungan ADF m/z U-BM 8-1_170818_4
ADF Statistic: -41.66716587473554
p-value: 0.0
Critical Values:
5%: -2.861543841626907
1%: -3.4303586916886233
10%: -2.566772044757177
Reject Ho - Time Series is Stationary
```

Figure 8a. ADF Test result for m/z

```
Hasil perhitungan ADF Intensitas U-BM 8-1_170818_4
ADF Statistic: -60.39892460642201
p-value: 0.0
Critical Values:
5%: -2.86154384218355
1%: -3.4303586929480314
10%: -2.566772045053458
Reject Ho - Time Series is Stationary
```

Figure 8b. ADF Test result for intensity

To make sure that the dataset is timeseries, we utilized the Augmented Dickey Fuller (ADF) Test to see whether the timeseries are stationary or not. The ADF test was done in Python by using statsmodels.tsa.stattools.adfuller library [44]. The tested features were m/z and intensity. The

[Type here]

[Type here]

12

time needed for the ADF test using this library was about 25 minutes with a 2.8 GHz i7 processor and 16GB RAM. The ADF test result shows that both stationary are timeseries and they are shown in Figure 8a. and 8b.

4. CONCLUSION

In this paper, we have describe a set of visualizations and Augmented Dickey Fuller Test results for U-BM 8-1_170818_4 dataset which is dataset from LC-MS process of Rodent Tuber plant. After observing and visually analyzing dataset by human eyes, and also after see the results of ADF test, it is concluded that the dataset examined in this study is a timeseries with stationary. Thus, for further processing, it is better to use algorithms intended for timeseries data.

REFERENCES

- [1] I. Blaženović, T. Kind, J. Ji, and O. Fiehn, "Software tools and approaches for compound identification of LC-MS/MS data in metabolomics," *Metabolites*, vol. 8, no. 2, 2018, doi: 10.3390/metabo8020031.
- [2] M. Katajamaa, J. Miettinen, and M. Orešič, "MZmine: Toolbox for processing and visualization of mass spectrometry based molecular profile data," *Bioinformatics*, vol. 22, no. 5, pp. 634–636, 2006, doi: 10.1093/bioinformatics/btk039.
- [3] S. Y. Lee *et al.*, "Comparison of GC/MS and LC/MS methods for the analysis of propofol and its metabolites in urine," *J. Chromatogr. B Anal. Technol. Biomed. Life Sci.*, vol. 900, pp. 1–10, 2012, doi: 10.1016/j.jchromb.2012.05.011.
- [4] E. R. Perez, J. A. Knapp, C. K. Horn, S. L. Stillman, J. E. Evans, and D. P. Arfsten, "Comparison of LC-MS-MS and GC-MS analysis of benzodiazepine compounds included in the drug demand reduction urinalysis program," *J. Anal. Toxicol.*, vol. 40, no. 3, pp. 201–207, 2016, doi: 10.1093/jat/bkv140.
- [5] B. Rushing, A. Wooten, M. Shawky, and M. I. Selim, "Comparison of LC – MS and GC – MS for the Analysis of Pharmaceuticals and Personal Care Products in Surface Water and Treated Wastewaters," *Curr. Trends Mass Spectrom.*, no. July, pp. 14–17, 2016.
- [6] A. Zhang, H. Sun, P. Wang, Y. Han, and X. Wang, "Modern analytical techniques in metabolomics analysis," *Analyst*, vol. 137, no. 2, pp. 293–300, 2012, doi: 10.1039/c1an15605e.
- [7] B. Zhou, "Computational Analysis of LC-MS/MS Data For Metabolite Identification," Virginia Polytechnic Institute and State University, 2012.
- [8] A. Vaniya and O. Fiehn, "Using fragmentation trees and mass spectral trees for identifying unknown compounds in metabolomics," *TrAC - Trends Anal. Chem.*, vol. 69, pp. 52–61, 2015, doi: 10.1016/j.trac.2015.04.002.
- [9] R. C. H. De Vos, S. Moco, A. Lommen, J. J. B. Keurentjes, R. J. Bino, and R. D. Hall, "Untargeted large-scale plant metabolomics using liquid chromatography coupled to mass spectrometry," *Nat. Protoc.*, vol. 2, no. 4, pp. 778–791, 2007, doi: 10.1038/nprot.2007.95.
- [10] R. J. M. Weber *et al.*, "Computational tools and workflows in metabolomics: An international survey highlights the opportunity for harmonisation through Galaxy," *Metabolomics*, vol. 13, no. 12, pp. 1–5, 2017, doi: 10.1007/s11306-016-1147-x.
- [11] P. J. Ulintz, J. Zhu, Z. S. Qin, and P. C. Andrews, "Improved Classification of Mass Spectrometry Database Search Results Using Newer Machine Learning Approaches," *Mol. Cell. Proteomics*, vol. 5, no. 3, pp. 497–509, 2006, doi: 10.1074/mcp.M500233-MCP200.
- [12] F. Fernández-Albert, "Machine Learning Methods for the Analysis of Liquid Chromatography-Mass Spectrometry datasets in Metabolomics," UNIVERSITAT POLITÈCNICA DE CATALUNYA, 2014.
- [13] L. Koesten, K. Gregory, P. Groth, and E. Simperl, "Talking datasets: Understanding data sensemaking behaviours," *arXiv:1911.09041v3 [cs.HC]*, pp. 1–26, 2020, [Online]. Available: <http://arxiv.org/abs/1911.09041>.
- [14] D. Laurent, N. F. Sianipar, Chelen, Listiari, and A. Wantho, "Analysis of Genetic Diversity of Indonesia Rodent Tuber (*Typhonium flagelliforme* Lodd.) Cultivars Based on RAPD Marker)," in *The 3rd International Conference on Biological Science 2013 (The 3rd ICBS-2013)*, 2015, vol. 2, pp. 139–145.
- [15] N. F. Sianipar, W. Maarisit, and A. Valencia, "Toxic Activities Of Hexane Extract And Column Chromatography Fractions Of Rudent Tuber Plant (*Typhonium Flagelliforme* Lodd.) On *Artemia Salina*," *Indones. J. Agric. Sci.*, vol. 14, no. 1, pp. 1–6, 2013.
- [16] N. F. Sianipar and R. Purnamaningsih, "Enhancement of the Contents of Anticancer Bioactive Compounds in Mutant Clones of Rodent Tuber (*Typhonium flagelliforme* Lodd.) based on GC-MS Analysis," *Pertanika J. Trop. Agric. Sci.*, vol. 41, no. 1, pp. 305–320, 2018.
- [17] R. Purnamaningsih and N. F. Sianipar, "Analysis Of Bioactive Compounds And Morphological Traits In Indonesian Rodent Tuber Mutant Clones Of Pekalongan Accession Using GC-MS," *J. Teknol.*, vol. 80, no. 2, pp. 131–136, 2018.
- [18] N. F. Sianipar, Ariandana, and W. Maarisit, "Detection of Gamma-Irradiated Mutant of Rodent Tuber (*Typhonium flagelliforme* Lodd.) In Vitro Culture by RAPD Molecular Marker," *Procedia Chem.*, vol. 14, pp. 285–294, 2015, doi: 10.1016/j.proche.2015.03.040.
- [19] N. F. Sianipar, D. Laurent, R. Purnamaningsih, and I. Darwati, "Genetic Variation of the First Generation of Rodent Tuber (*Typhonium flagelliforme* Lodd.) Mutants Based on RAPD Molecular Markers," *HAYATI J.*

6 [Type here]

[Type here]

- Biosci.*, vol. 22, no. 2, pp. 98–104, 2015, doi: 10.4308/hjb.22.2.98.
- [20] N. F. Sianipar, R. Purnamaningsih, D. L. Gumanti, Rosaria, and M. Vidiarti, "Analysis Of Gamma Irradiated Fourth Generation Mutant Of Rodent Tuber (Typhonium Flagelliforme Lodd.) Based On Morphology And RAPD Markers," *J. Teknol.*, vol. 78, no. 5–6, pp. 41–49, 2016.
- [21] N. F. Sianipar, R. Purnamaningsih, and Rosaria, "Bioactive compounds of fourth generation gamma-irradiated Typhoniumflagelliforme Lodd . mutants based on gas chromatography-mass spectrometry," 2016, doi: 10.1088/1755-1315/41/1/012025.
- [22] R. H. Shumway and D. S. Stoffer, *Time Series Analysis and Its Applications With R Examples*, 3rd ed. Springer Science+Business Media New, 2011.
- [23] W. M. A. Niessen, *Liquid chromatography-mass spectrometry Third Edition*. 2006.
- [24] R. E. Ardrey, *Liquid Chromatography –Mass Spectrometry: An Introduction*. 2003.
- [25] M. Ludovici, C. Ialongo, and E. Camera, "Principles, current applications, and future perspectives of liquid chromatography-mass spectrometry in clinical chemistry," in *Liquid Chromatography: Applications: Second Edition*, Second Edi., vol. 2, Elsevier Inc., 2017, pp. 727–751.
- [26] Hewlett Packard, *Basics of LC/MS*. 2001.
- [27] P. R. Kumar, S. R. Dinesh, and R. Rini, "Lcms- A Review and a Recent Update," *J. Pharm. Pharm. Sci.*, vol. 5, no. 5, pp. 377–391, 2016, doi: 10.20959/wjpps20165-6656.
- [28] R. Tautenhahn, C. Botcher, and S. Neumann, "Highly sensitive feature detection for high resolution LC/MS," *BMC Bioinformatics*, vol. 9, no. 504, pp. 1–16, 2008, doi: 10.1186/1471-2105-9-504.
- [29] Agilent Technologies Inc., *Agilent 6100 Series Quadrupole LC / MS Systems*. 2010.
- [30] B. Zhou, J. F. Xiao, L. Tuli, and H. W. Ransom, "LC-MS-based metabolomics," *Mol. Biosyst.*, vol. 8, no. 2, pp. 470–481, 2012, doi: 10.1039/C1MB05350G.
- [31] L. Gatto, L. M. Breckels, T. Naake, and S. Gibb, "Visualization of proteomics data using R and Bioconductor," *Proteomics*, vol. 15, no. 8, pp. 1375–1389, 2015, doi: 10.1002/pmic.201400392.
- [32] D. H. Nicolson and M. Sivadasan, "Four frequently confused species of Typhonium Schott (Araceae)," *Blumea J. plant Taxon. plant Geogr.*, vol. 27, no. 2, pp. 483–497, 1981.
- [33] N. I. Rezali, N. Jaafar Sidik, A. Saleh, N. I. Osman, and N. A. Mohd Adam, "The effects of different strength of MS media in solid and liquid media on in vitro growth of Typhonium flagelliforme," *Asian Pac. J. Trop. Biomed.*, vol. 7, no. 2, pp. 151–156, 2017, doi: 10.1016/j.apjtb.2016.11.019.
- [34] M. Singh, D. Kumar, D. Sharma, and G. Singh, "Typhonium Flagelliforme: a Multipurpose Plant," *Int. Res. J. Pharm.*, vol. 4, no. 3, pp. 45–48, 2013, doi: 10.7897/2230-8407.04308.
- [35] C. Y. Choo, K. L. Chan, K. Takeya, and H. Itokawa, "Cytotoxic activity of Typhonium flagelliforme (Araceae)," *Phyther. Res.*, vol. 15, no. 3, pp. 260–262, 2001, doi: 10.1002/ptr.717.
- [36] S. T. Gopukumar, P. Alexander, T. K. S. Fathima, R. Jeniffa, and P. K. Praseetha, "Typhonium Divaricatum : a Promising Endangered Anti-Cancer," *Med. Plants Their Conserv. Manag.*, no. July, pp. 114–119, 2015.
- [37] C. K. Neoh, "Typhonium divaricatum (rodent tuber): A promising local plant in the fight against cancer," *Med. J. Malaysia*, vol. 47, no. 1, pp. 86–88, 1992.
- [38] S. Mohan, A. Bustaman, S. Ibrahim, A. S. Al-Zubairi, and M. Aspollah, "Anticancerous Effect of Typhonium flagelliforme on Human T-Lymphoblastoid Cell Line CEM-ss," *J. Pharmacol. Toxicol.*, vol. 3, no. 6, pp. 449–456, 2008, doi: 10.3923/jpt.2008.449.456.
- [39] A. Bustamam *et al.*, "In vitro ultramorphological assessment of apoptosis on CEMss induced by linoleic acid-rich fraction from Typhonium flagelliforme tuber," *Evidence-based Complement. Altern. Med.*, vol. 2011, 2011, doi: 10.1093/ecam/nej010.
- [40] S. Mohan *et al.*, "Typhonium flagelliforme inhibits the proliferation of murine leukemia WEHI-3 cells in vitro and induces apoptosis in vivo," *Leuk. Res.*, vol. 34, no. 11, pp. 1483–1492, 2010, doi: 10.1016/j.leukres.2010.04.023.
- [41] M. C. Chambers *et al.*, "A cross-platform toolkit for mass spectrometry and proteomics," *Nat. Biotechnol.*, vol. 30, no. 10, pp. 918–920, 2012, doi: 10.1038/nbt.2377.
- [42] H. L. Röst *et al.*, "OpenMS: A flexible open-source software platform for mass spectrometry data analysis," *Nat. Methods*, vol. 13, no. 9, pp. 741–748, 2016, doi: 10.1038/nmeth.3959.
- [43] J. Xia, R. Mandal, I. V. Sinelnikov, D. Broadhurst, and D. S. Wishart, "MetaboAnalyst 2.0-a comprehensive server for metabolomic data analysis," *Nucleic Acids Res.*, vol. 40, no. W1, pp. 127–133, 2012, doi: 10.1093/nar/gks374.
- [44] J. Perktold, S. Seabold, and J. Taylor, "statsmodels.tsa.stattools.adfuller — statsmodels," <https://www.statsmodels.org/>, 2020.
<https://www.statsmodels.org/devel/generated/statsmodels.tsa.stattools.adfuller.html#statsmodels.tsa.stattools.adfuller> (accessed Aug. 01, 2020).

[Type here]

ORIGINALITY REPORT

20%

SIMILARITY INDEX

12%

INTERNET SOURCES

17%

PUBLICATIONS

4%

STUDENT PAPERS

PRIMARY SOURCES

- | | | |
|---|--|-----|
| 1 | Iwan Binanto, Harco Leslie Hendric Spits Warnars, Nesti Fronika Sianipar, Bahtiar Saleh Abbas. "LC-MS Analysis: Mini Review Frequently Used Open Source Softwares", 2019 6th International Conference on Information Technology, Computer and Electrical Engineering (ICITACEE), 2019
Publication | 11% |
| 2 | Submitted to Troy University
Student Paper | 1% |
| 3 | eprints.binus.ac.id
Internet Source | 1% |
| 4 | Submitted to Universiti Teknologi Malaysia
Student Paper | 1% |
| 5 | pdfs.semanticscholar.org
Internet Source | 1% |
| 6 | Pulok K. Mukherjee. "Plant Metabolomics and Quality Evaluation of Herbal Drugs", Elsevier BV, 2019
Publication | 1% |

7	Nur Inani Rezali, Norrizah Jaafar Sidik, Azani Saleh, Nurul Izzati Osman, Nurul Athirah Mohd Adam. "The effects of different strength of MS media in solid and liquid media on in vitro growth of <i>Typhonium flagelliforme</i> ", Asian Pacific Journal of Tropical Biomedicine, 2017 Publication	1 %
8	aaltodoc.aalto.fi Internet Source	<1 %
9	www.docstoc.com Internet Source	<1 %
10	www.nature.com Internet Source	<1 %
11	tel.archives-ouvertes.fr Internet Source	<1 %
12	escholarship.org Internet Source	<1 %
13	Laurent Gatto, Sebastian Gibb, Johannes Rainer. "MSnbase, Efficient and Elegant R-Based Processing and Visualization of Raw Mass Spectrometry Data", Journal of Proteome Research, 2020 Publication	<1 %
14	Castillo, S.. "Algorithms and tools for the preprocessing of LC-MS metabolomics data",	<1 %

Chemometrics and Intelligent Laboratory Systems, 20110815

Publication

15

www.mdpi.com

Internet Source

<1 %

16

tdx.cat

Internet Source

<1 %

17

hdl.handle.net

Internet Source

<1 %

18

J. de Corral, H. Pfister. "Hardware-Accelerated 3D Visualization of Mass Spectrometry Data", VIS 05. IEEE Visualization, 2005., 2005

Publication

<1 %

Exclude quotes On

Exclude matches

< 5 words

Exclude bibliography On