Manual for

FlavonoidSearch FsTool

16 May 2017

Contents

Introduction

FlavonoidSearch[1] is a system for annotating flavonoid aglycones using mass spectra obtained from multiple-stage mass spectrometry (MS^n) analyses. FlavonoidSearch consists of two parts: FsDatabase, a database of predicted mass fragments for 6867 known flavonoids; and FsTool, a Java-based tool to search FsDatabase. FsDatabase was manually constructed based on new structure- and fragmentation-related rules and heuristics as well as know fragmentation rules in the literature. The system showed high identification accuracy for the flavonoid aglycone when compared to existing tools and revealed high accuracy for discrimination between the flavonoid aglycone and other compounds.

This manual is for FsTool which can be used both as an easy-to-use graphical user interface (GUI) tool and a command line tool for high-throughput calculations on servers.

The tool is available at the FlavonoidSearch website:
http://www.kazusa.or.jp/komics/software/FlavonoidSearch

License

This software is available free of charge for academic purposes. The core search program (flavonoidsearch.jar) is an open source software licensed under the GNU Lesser General Public License, Version 2.1 (LGPL 2.1). The source codes of flavonoidsearch.jar are included in the jar file.

Computer requirements

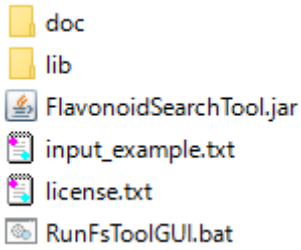
A PC with Java Runtime Environment (version 1.6 or later) is required to run FsTool. See the Oracle's website (https://www.java.com/ja/download/help/download_options.xml) for installation of Java. The tool is tested on the following operating systems:

Windows10, Mac OSX 10.9.5 and CentOS 7.2 Linux

Use as a GUI tool

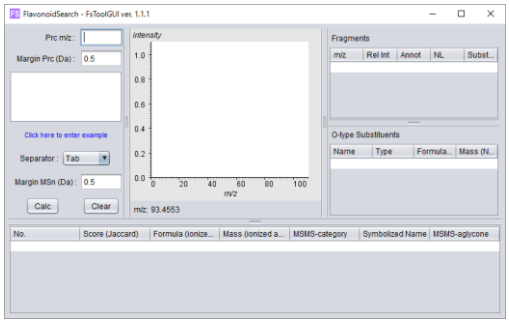
Run / Exit

Decompress the zip file of the program using a decompression software such as 7zip. Following files and folders will be generated.



Run on Windows

Double click the file 'RunFsToolGUI.bat'. The main window FsTool will be displayed.



* A black console window will be displayed too. Do not close this window while FsTool is running, or FsTool closes too.

Run on Mac OSX / Linux

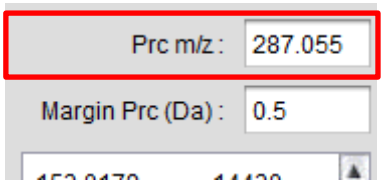
Run a terminal software and go to the directory where the decompressed files exist. Execute the following command:

```
java -jar FlavonoidSearchTool.jar
```

To exit the tool, click 'x' button at top-right corner of the window.

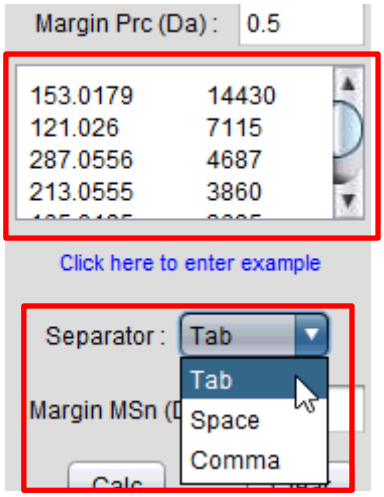
Searching

Enter *m/z* value in the 'Prc m/z' field.

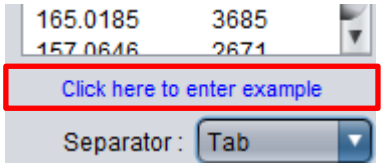


* Only spectra obtained by positive mode can be searched in FlavonoidSearch.
* Enter *m/z* values that are measured using mass spectrometers without any consideration with adducts. The mass values of the precursor ions have been calculated in consideration with the types of adduct ($[M+H]^+$ or $[M]^+$) in the predicted fragment database (FsDatabase).

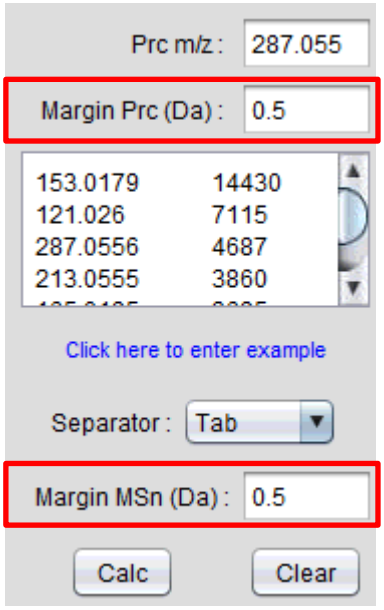
Enter a query spectrum data in the text area. A pair of *m/z* value and intensity of the MS^n fragment ion separated by tab, space or comma should be written in each row. Select the type of the separator from the 'Separator' pull down list.



* Example values will be entered by clicking on the text 'Click here to enter example'.

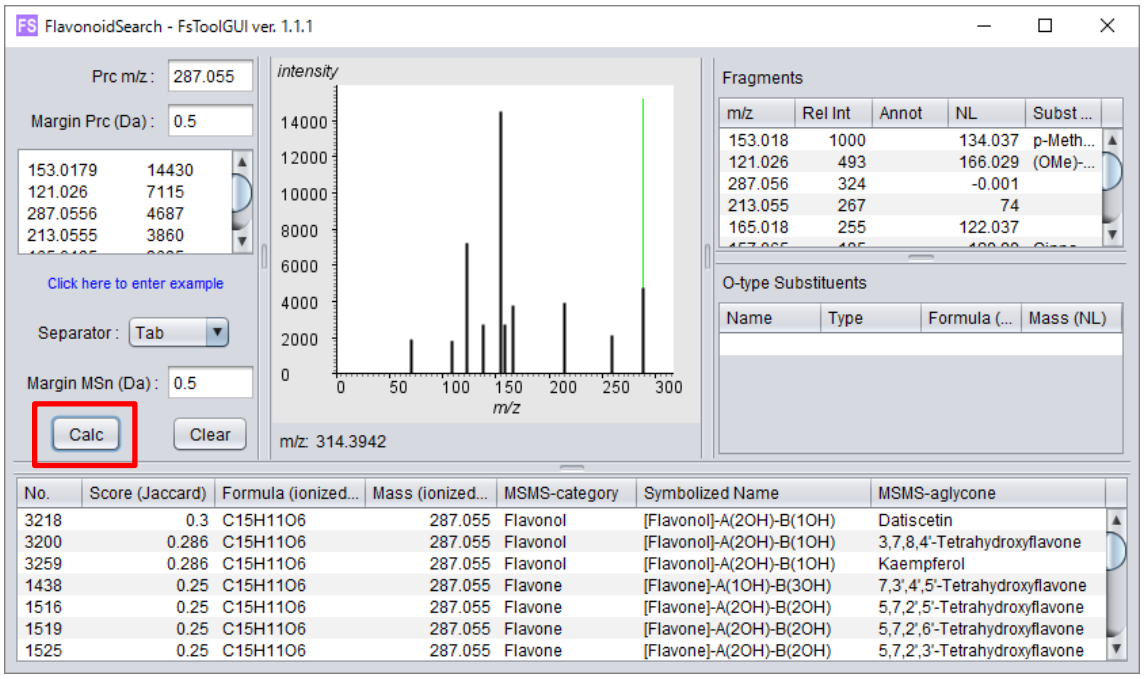


Enter the mass tolerances (given in daltons, Da) for the precursor and the MS^n fragment ions in the 'Margin Prc' and 'Margin MSn' fields, respectively. These values are used for comparing the spectra to those in FsDatabase.



Click the 'Calc' button to search. The results will be displayed in the main table at the

bottom of the tool. The information of the input spectrum is displayed in the center panel and the table at the right.



Search results

In the main table, candidates of flavonoid aglycones are listed in the order of the hit score (Jaccard index). The score shows the similarity of *m/z* values between the query fragments and the predicted fragments in FsDatabase. The score is calculated as the following equation.

Score (Jaccard) = Number of query fragments shared with predicted fragments / Number of unique fragments in query and predicted fragments

* Intensity of the fragment ion is not used in FlavonoidSearch system.

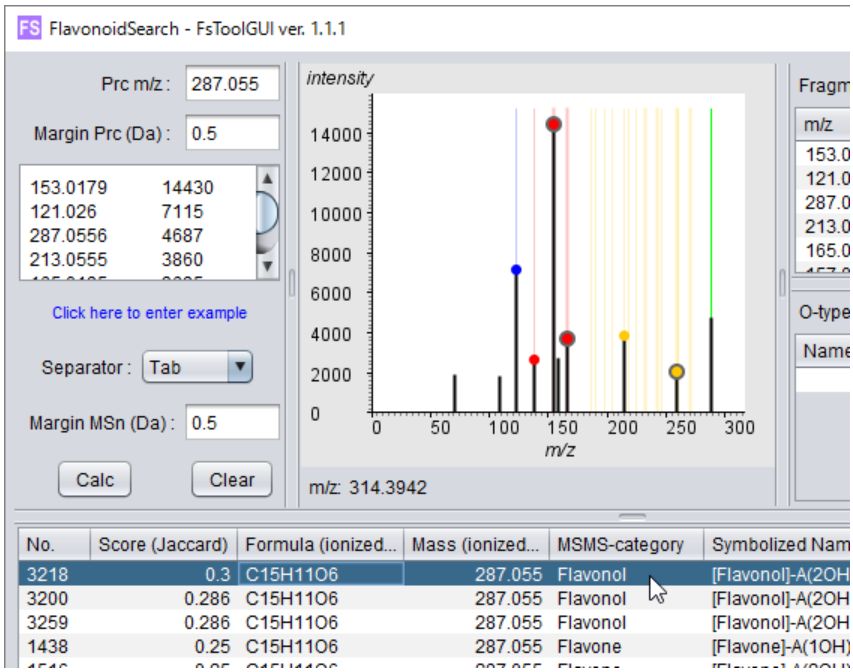
The table columns show following information:

No	The No. assigned for the flavonoids in FlavonoidSearch
Score (Jaccard)	The similarity value (0-1) of the *m/z*. A higher value

	means a higher similarity.
Formula (ionized aglycone)	The formula of the ionized aglycone
Mass (ionized aglycone)	The mass value of the ionized aglycone
MSMS-category	The class name of the aglycone assigned in FlavonoidSearch
Symbolized Name	The name assigned to the aglycone to facilitate understanding the structural similarities among the flavonoids. The symbolized names include information about the MSMS-category, types and number of substituents on A- and B-rings.
MSMS-aglycone	The name assigned to the structural unit that fragments with the aglycone moiety including substituents that have low degrees of dissociation in the MS/MS analysis.

* An additional information including ID for MSMS-category is available with the command line FsTool.

Select a row in the main table. The predicted fragments and the matched query fragments will be represented in the center panel.

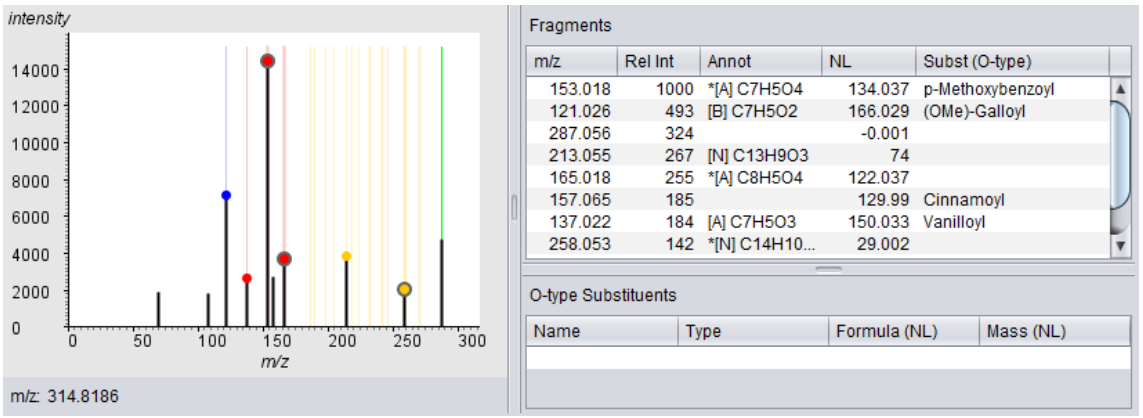


The colour of the lines and symbols are as follows:

Black line	The query fragments
Green line	The *m/z* value of queried precursor ion
Red, blue and yellow lines	The *m/z* values of predicted fragments in FsDatabase. The colour means as follows: Red: Fragments derived from A-ring Blue: Fragments derived from B-ring Yellow: Neutral loss fragments The thick lines are fragments expected to be frequently observed (Essential Fragment[1])
Circle	The query fragments matched to the predicted fragments. The colours mean as same as the red, blue and yellow lines. The query fragments matched to the essential fragments are represented as black borders.

Information of the fragments

The details of the fragment ions are displayed in the 'Fragments' table.



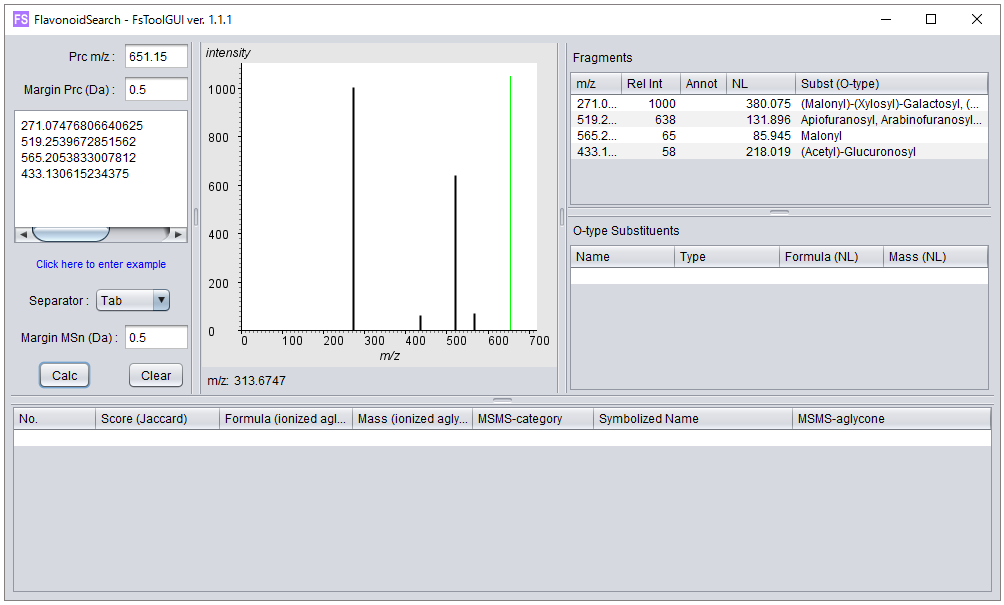
The following information is available:

m/z	The *m/z* value of the query fragment ion
Rel Int	The relative intensity of the fragment ion when the intensity of the most intense ion in the spectrum (base peak ion) is assumed to 1000.
Annot	Information of the predicted ions. The letters in the parentheses [ ]

	denote as follows: [A]: Fragments derived from A-ring [B]: Fragments derived from B-ring [N]: Neutral loss fragments The formula of the fragment ion is shown without their charge. The essential fragments are represented as asterisks (*).
NL	The mass of the neutral loss (differences of the mass between the precursor and fragment ions)
Subst (O-type)	The known *O*-type substituents (Supplementary Table S8[1]) matched to the neutral loss value.

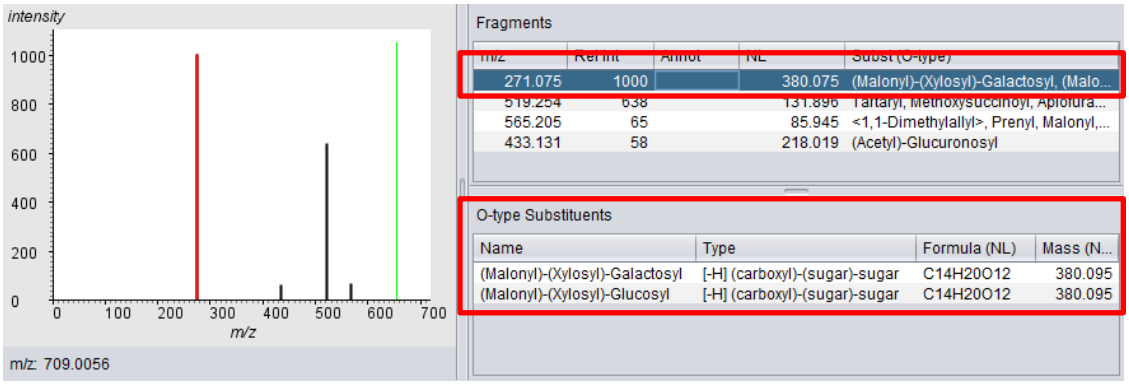
Information of possible *O*-type substituents

If the neutral loss value matches to the masses of the *O*-type substituents of known flavonoids (Supplementary Table S8[1]), the information of the substituents is displayed in the 'Subst (Otype)' column in the 'Fragments' table.



* In this example, no aglycones were found.

Click the row of the 'Fragments' table. The candidates of the *O*-type substituents matched to the neutral loss value are listed in the 'O-type Substituents' table.



'O-type Substituents' table shows following information:

Name	Name of the substituent. When several structural units are included, the names are concatenated with hyphens. A pair of parentheses '( )' means a block of structural unit(s). The block binds to the aglycone or another block via the unit of the name at the tail of the block.
Type	To facilitate understanding the types of the substituents, the general name in the 'Name' coulmn was replaced with class names as follows: carboxyl, hydroxyl, OMe, sugar and sulfo. [-H] and [OH] are formulae that are removed from (H) or added to (OH) the structure when the neutral loss fragment is dissociated from the aglycone. Only [-H] form is displayed here for the substituents whose number of substituents units (Supplementary Table S8[1]) is >2 and which bind to the aglycone via sugar.
Formula (NL)	The formula of the substituent as neutral loss
Mass (NL)	The mass value of the substituent as neutral loss

In the example above, a neutral loss value of 380.075 for the fragment ion 271.075 is suggested as a dissociation of a substituent ($C_{14}H_{20}O_{12}$) in which malonyl and xylosyl residues are conjugated to a hexose (galactose or glucose).

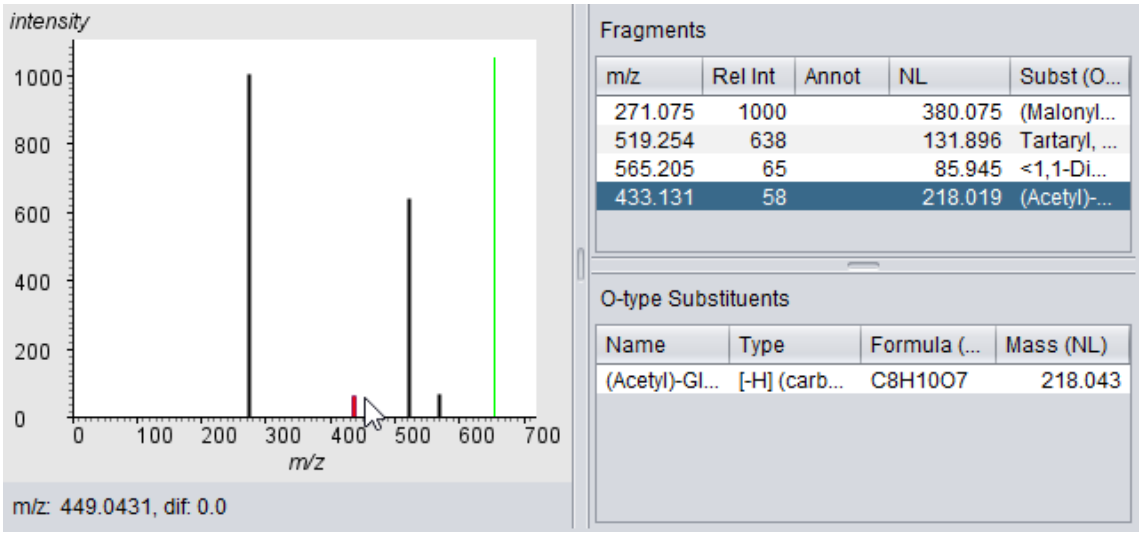
Mouse operations on the spectrum panel

Wheel rotation	Zooming in/out. The direction of zooming can be fixed using CTRL and SHIFT keys. - Operations with the CTRL key restricts the zooming only to y-direction.

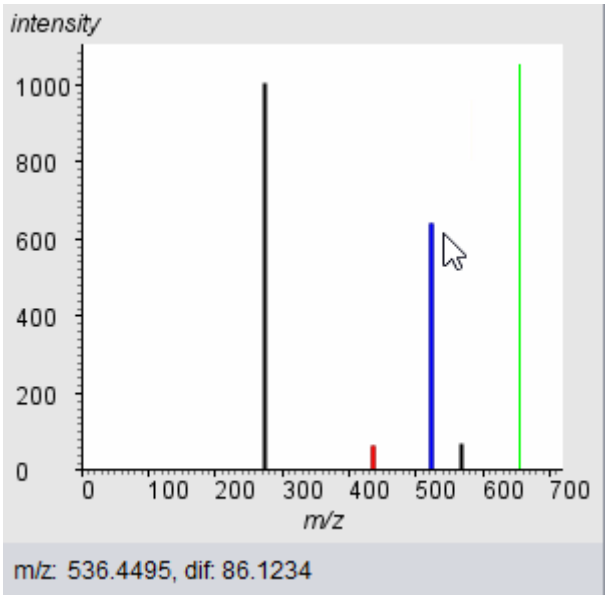
	- Operations with the SHIFT key restricts the zooming only to x-direction.
Left button drag	Moving
Right button double click	Zooming full out
Right button drag	Zooming in the selected area
Left button click	Selecting the nearest ion (see next section)

Selection of the fragment ion

Click on the fragment ion in the spectrum panel or select a row in the 'Fragments' table. The selected query fragment ion will be represented in red.



Move the mouse cursor on the spectrum panel. The nearest ion will be represented in blue and a mass difference between the red and blue ions is displayed next to the 'dif' label at the bottom of the panel.



Use as a command line tool

Users can use FsTool as a command line tool to query a spectrum data without GUI operations. This is useful for searching a large number of spectra on servers. The search results contain additional information such as ID of MSMS-category that are not displayed in FsTool GUI.

Preparation of the spectrum data

Prepare query data in a text file in NIST-like format as follow:

PrecursorMZ: The mass value of the precursor ion.
Num peaks: Number of the MS^n fragments given in the following lines as (mass intensity)

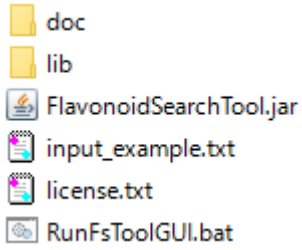
ex)
PrecursorMZ: 287.0550145047
Num peaks: 10

153.0179 14430
121.026 7115
287.0556 4687
213.0555 3860
165.0185 3685
157.0646 2671
137.022 2665
258.053 2058
68.9872 1822
107.0454 1744

Please see the example file named ‘input_example.txt’.

Preparation of FsTool

Decompress the zip file downloaded using a decompressing tool such as 7zip. The following files will be generated.



Execution of search

Run a terminal software, and go to the directory where the decompressed files exist. Execute the following command:

```
java -jar FlavonoidSearchTool.jar -i INFILE
```

where INFILE is the path to the query file prepared above.

As an option, the mass tolerances for the precursor ion and the fragment ions can be defined using the following command:

```
java -jar FlavonoidSearchTool.jar -i INFILE -p MARGIN_PRE -m MARGIN_MSN
```

where MARGIN_PRE is the mass tolerance for the precursor ion (default = 0.01), and MARGIN_MSN is the mass tolerance for the fragment ions (default = 0.5) in daltons (Da).

ex) Using the sample file 'input_example.txt' included in the downloaded file set, execute the following command:

```
java -jar FlavonoidSearchTool.jar -i input_example.txt -p 0.2
```

The following results will be displayed:

ID JaccardScore IonizedAglyconeFormula IonizedAglyconeMass MSMS-CategoryID MSMS-CategoryName SymbolizedAglyconeID SymbolizedAglyconeName MSMS-AglyconeID MSMS-AglyconeName RelatedID RelatedCAS

3218 0.3 C15H11O6 287.0550145 C059 Flavonol S1441 [Flavonol]-A(2OH)-B(1OH) A2050 Datiscetin FL5FA8NS0001 480-15-9

3200 0.2857142857142857 C15H11O6 287.0550145 C059 Flavonol S1441 [Flavonol]-A(2OH)-B(1OH) A2033 3,7,8,4'-Tetrahydroxyflavone FL5F3ANS0001 1429-28-3

3259 0.2857142857142857 C15H11O6 287.0550145 C059 Flavonol S1441 [Flavonol]-A(2OH)-B(1OH) A2078 Kaempferol FL5FAANS0001 520-18-3

1438 0.25 C15H11O6 287.0550145 C038 Flavone S0853 [Flavone]-A(1OH)-B(3OH) A1110 7,3',4',5'-Tetrahydroxyflavone FL3F1GNS0001 67858-31-5

...

RelatedID: An ID of the flavonoid database in metabolomics.jp (http://metabolomics.jp/wiki/Category:FL) is shown as an example of the flavonoid that includes the aglycone.

RelatedCAS: The CAS number of the flavonoid exemplified in the RelatedID.

The results are saved in a text file using a redirection command as follows:

```
java -jar FlavonoidSearchTool.jar -i input_example.txt -p 0.2 > output.txt
```

A brief summary of the commands can be displayed by entering '-h' as below:

```
java -jar FlavonoidSearchTool.jar -h
```

To run the FsTool GUI, execute the FlavonoidSearchTool.jar without any options as below:

```
java -jar FlavonoidSearchTool.jar
```

About us

Department of Technology Development, Metabolomics Team,
Kazusa DNA Research Institute, Chiba 292-0818, Japan.

Developer:
Nozomu Sakurai
E-mail: sakurai AT kazusa.or.jp (replace AT with @)

References

1. Akimoto, N. et al. FlavonoidSearch: A system for comprehensive flavonoid annotation by mass spectrometry. *Sci Rep* **7**, 1243 (2017).