

## MFSearcher ツールマニュアル

バージョン 1.6.0

2019 年 8 月 28 日

### ライセンス

このソフトウェアは学術目的においてフリーでご使用になれます。

このソフトウェアでは、以下の既存データベースの検索が可能です。検索された化合物データの取扱いについては、それぞれ元のデータベースのライセンス条項をご参照ください。

KEGG	<a href="http://www.genome.jp/kegg/">http://www.genome.jp/kegg/</a>
KNApSAcK	<a href="http://kanaya.naist.jp/KNApSAcK/">http://kanaya.naist.jp/KNApSAcK/</a>
FlavonoidViewer	<a href="http://metabolomics.jp/wiki/Category:FL">http://metabolomics.jp/wiki/Category:FL</a>
HMDB	<a href="http://www.hmdb.ca/">http://www.hmdb.ca/</a>
LipidMAPS	<a href="http://www.lipidmaps.org/">http://www.lipidmaps.org/</a>
UNPD	<a href="http://pkuxxj.pku.edu.cn/UNPD/">http://pkuxxj.pku.edu.cn/UNPD/</a>
PubChem	<a href="http://pubchem.ncbi.nlm.nih.gov/">http://pubchem.ncbi.nlm.nih.gov/</a>

### 動作環境

Java ランタイム環境 (1.7 以上) がインストールされ、インターネットに接続された PC (Windows 推奨) が必要です。

本ソフトウェアは、データ検索時に MFSearcher ウェブサービス (<http://webs2.kazusa.or.jp/mfsearcher/>) にアクセスします。また検索後、選択された化合物について、オリジナルのデータベースへリンクして、その情報をブラウザで表示することができます。これらの動作で、インターネットへの接続を行います。

プロキシサーバーを介してインターネットに接続している場合は、「プロキシサーバーの設定」に従って設定を行って下さい。

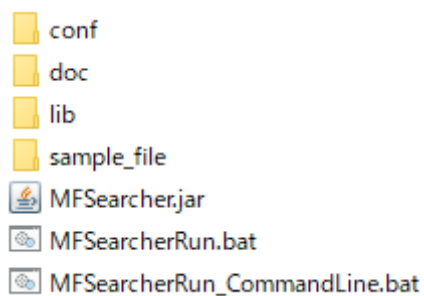
Java のインストールがされていない場合は、下記 URL に従ってインストールを行ってください。

[https://www.java.com/ja/download/help/download\\_options.xml](https://www.java.com/ja/download/help/download_options.xml)

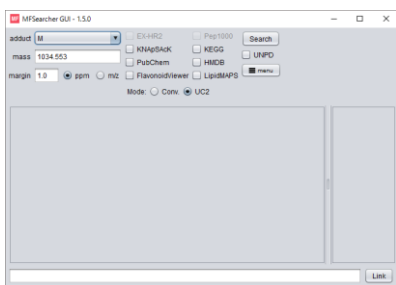
## 起動・終了

### ● GUIでの起動

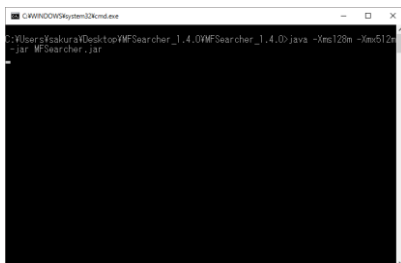
1) ダウンロードした zip ファイルを、ファイル解凍ソフトウェア (7zip 等) で解凍します。下記のファイルが生成されます。



2) MFSearcherRun.bat をダブルクリックすると、ソフトウェアが起動します。MFSearcher のメイン画面が表示されます。



※この際、下記のようなコンソール画面も同時に表示されます。この画面は、MFSearcher 起動中に閉じないようにしてください (最小化は可能です)。右上の「×」ボタンを押して画面を閉じると、MFSearcher 自体も終了してしまいます。



3) MFSearcher メイン画面の右上の「×」ボタンを押すことで、ソフトウェアが終了します。



## ● コマンドラインでの起動

上記1) で解凍された `MFSearcher_CommandLine.bat` をダブルクリックすると、コマンドラインでの使用方法が表示されます。検索したい質量値が多数あるとき、それらをファイルに記載しておくことで、全件について組成式推定およびデータベース検索を実施し、結果がテキストファイルに保存されます。

詳しい使い方は、「コマンドラインモード」の項で解説します。

## 基本的な使い方

1) 以下の設定を記載します。

**adduct** 欄：想定されるアダクトを選択します。検索時には、アダクトの差分が差し引かれたもとの中性分子の質量値が使用されます。

**mass** 欄：調べたい **mass** 値を入力します。

**margin** 欄：許容する質量誤差範囲。ppm か m/z を設定します。

The screenshot shows the MFSearcher interface with the following fields highlighted in a red box:

- adduct: [M+H]<sup>+</sup>
- mass: 1034.553
- margin: 1.0 (ppm selected)

Other visible options include:

- EX-HR2 (unchecked)
- KNAPSAcK (checked)
- PubChem (unchecked)
- FlavonoidViewer (checked)
- Pep1000 (unchecked)
- KEGG (checked)
- HMDB (checked)
- LipidMAPS (checked)
- UNPD (unchecked)
- Mode: Conv. (selected), UC2 (unchecked)

2) チェックボックスで、検索対象とするデータベースを選択します。

The screenshot shows the MFSearcher interface with the database selection checkboxes highlighted in a red box:

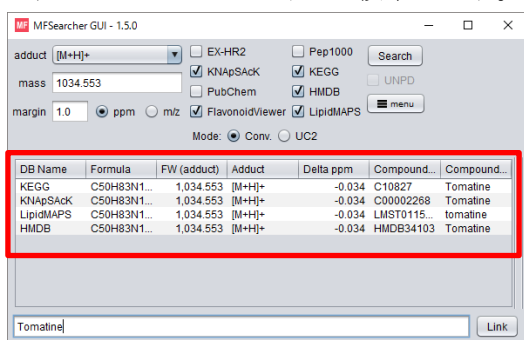
- EX-HR2 (unchecked)
- KNAPSAcK (checked)
- PubChem (unchecked)
- FlavonoidViewer (checked)
- Pep1000 (unchecked)
- KEGG (checked)
- HMDB (checked)
- LipidMAPS (checked)
- UNPD (unchecked)

使用できるデータベースは以下です。詳細は、**MFSearcher** のウェブサイト

(<http://webs2.kazusa.or.jp/mfsearcher>) をご覧下さい。

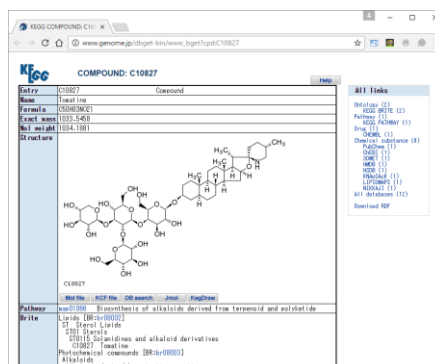
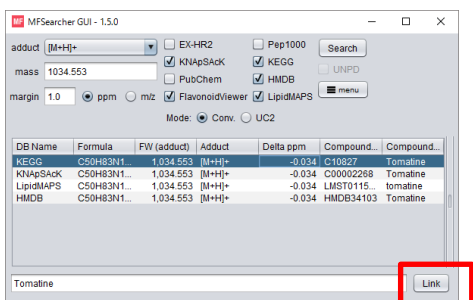
EX-HR2	化合物として価数の矛盾がない組成式のデータベース
Pep1000	分子量 1000 までの直鎖ペプチドのデータベース
KNApSAcK	天然物（主に植物）のデータベース <a href="http://kanaya.naist.jp/KNApSAcK/">http://kanaya.naist.jp/KNApSAcK/</a>
KEGG	KEGG 代謝マップに記載されている化合物 <a href="http://www.genome.jp/kegg/">http://www.genome.jp/kegg/</a>
PubChem	NCBI が構築する巨大な化合物データベース <a href="https://pubchem.ncbi.nlm.nih.gov/">https://pubchem.ncbi.nlm.nih.gov/</a>
HMDB	ヒトで検出、あるいは存在が推定される化合物のデータベース <a href="http://www.hmdb.ca/">http://www.hmdb.ca/</a>
FlavonoidViewer	植物二次代謝産物のフラボノイドを集めたデータベース <a href="http://metabolomics.jp/wiki/Category:FL">http://metabolomics.jp/wiki/Category:FL</a>
LipidMAPS	脂質・ポリフェノールなどのデータベース <a href="http://www.lipidmaps.org/">http://www.lipidmaps.org/</a>

3) Search ボタンを押して検索します。結果が一覧されます。



表の中の列をクリックすると、下部にある欄に化合物名が表示されます。長い化合物名などはこのようにご確認ください。

この状態で「Link」ボタンを押すと、ブラウザが開き、オリジナルの化合物データベースで詳しい情報を見ることができます。



## UC2 検索

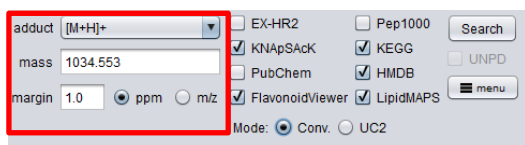
UC2 検索は、複数のデータベースで同じ化合物が登録されていた場合に、それを一つにまとめて結果表示してくれるという機能です。その他、以下の特徴があります。

- データベース中にチャージ分子として登録されている化合物（植物色素のシアニジンなど）についても、正しい価数と質量値で検索できる。
- データベース中で、塩などの複合体として登録されている化合物についても、塩を除去したものとして検索できる。

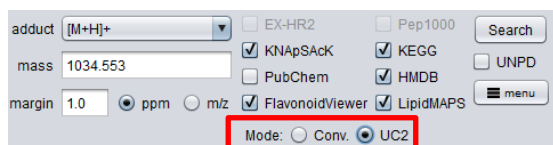
このため、塩やチャージ分子への誤ヒットをなくし、かつ検索結果の候補数を減らすことができます。

同じ化合物かどうかは、元素同士の結合様式が一致するかどうかを、InChIKeyの第一ブロック（14文字目）で判定することで実現されています。立体情報を考慮していないため、アミノ酸のD体L体などは、区別されずにひとまとめに表示されます。

1) 通常の検索と同様に、adduct、mass、marginを設定します。



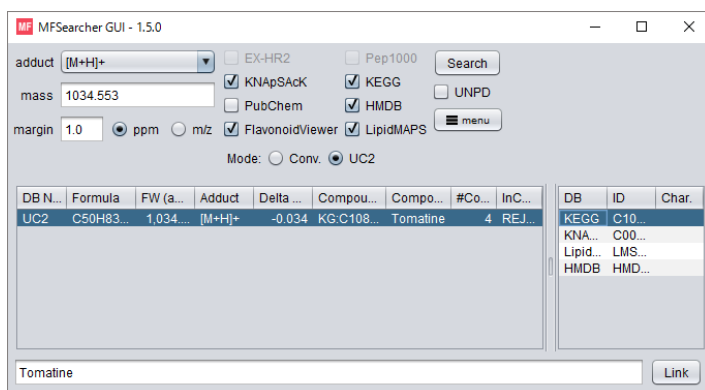
2) Mode欄でUC2にチェックを入れます。



UC2 モードになると、化学構造情報を持たない EX-HR2 と Pep1000 は選択できなくなり、選択対象から外れます。また、UNPD データベースが利用できるようになります。

UNPD	天然物の巨大なデータベース <a href="http://pkuxxj.pku.edu.cn/UNPD/">http://pkuxxj.pku.edu.cn/UNPD/</a>
------	--

3) Search ボタンを押して検索を実行します。検索結果が一覧されます。



UC2 モードでは、結果表示画面が二つ表示されます。左の画面には、同一化合物を一つにまとめた結果が表示されます。結果の一つをクリックすると、右の画面に、同一と判定された化合物の ID とデータベース名が表示されます。

右の画面で化合物 ID を一つ選択し、「Link」ボタンを押すと、ブラウザでオリジナルサイトが開き、詳細情報を確認できます。(UNPD のウェブサイトの仕様から、UNPD にはリンクできません)

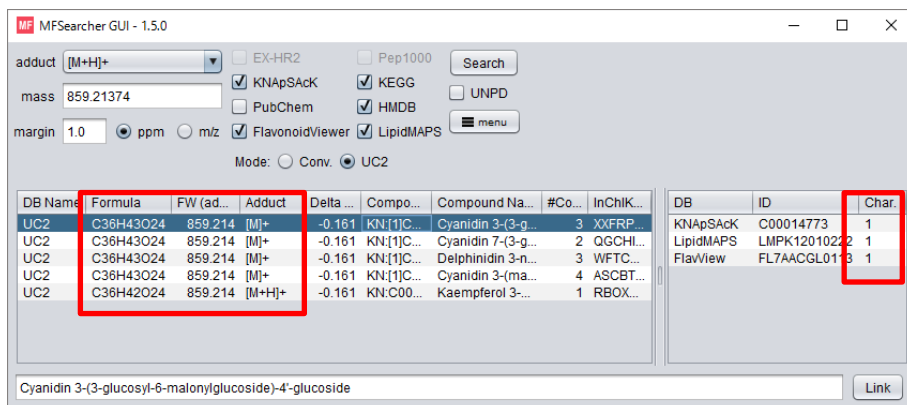
左の画面で化合物名は、複数のデータベースに登録されている化合物名の中で最も文字数が少なかったものが代表として一つ表示されています。

右の画面の Char. 欄には、その化合物が元のデータベースでどのように登録されていたかが表示されます。記載されたものは下記を意味します。

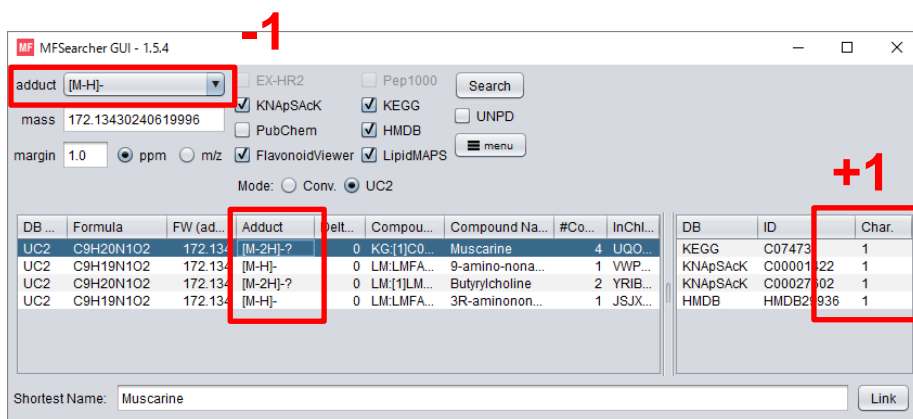
数字	チャージ分子として記載されていた
f	塩などの複合体として記載されていた
r	ラジカルとして記載されていた

UC2 検索では、検索された化合物のチャージを考慮して、明らかに分子イオンだと判断される場合には、Formula 欄、Adduct 欄、FW (molecule) 欄に、分子イオンとして計算されたものが表示されます。例えば [M+H]<sup>+</sup> で検索した際に、ヒットした化合物が全て +1 チャー

ジの分子イオンだった場合、[M]<sup>+</sup>のアダクトとして表示されます。

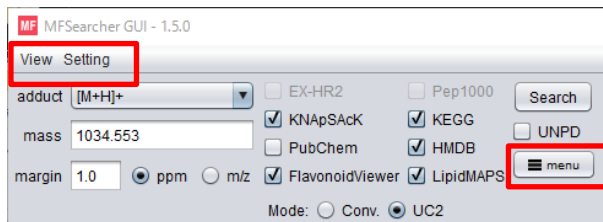


また、検索したチャージとデータベース内のチャージに矛盾がある場合には、「?」が付加された候補のアダクトがアダクト欄に表示されます。



## 高度な設定

「menu」ボタンを押すと、メニューが表示されます。



メニューには以下が含まれます。

View	Show/Hide Status Bar	画面の下部にステータスバーを表示/非表示しま
------	----------------------	------------------------

		す。
	Show DB Info.	データベースのレコード数と更新日時日時を示すウィンドウを表示します。 ※ここで表示されるのは、通常モード検索時に使われるデータベースです。UC2 モードで使われるデータは、これとは別に存在し、その更新日時は表示されません。
	Hide Menu	メニューを非表示にします。
Setting	Column Settings	検索結果一覧の表示内容を変更できます。
	Filter Settings	検索結果を元素数などで絞り込むためのフィルターを設定します。

## フィルター設定

Setting -> Filter Settings で、検索結果をさらに絞り込むためのフィルターを設定できます。

Filter Setting

Atom Constraints:

Constraint only for EX-HR2 DB

Assume P as phosphate

0 <=  C <= 100

0 <=  N <= 20

0 <=  S <= 5

0 <=  P <= 5

0 <=  < <= 10

\* Blanks are no limitation

13C / 12C Ratio:

Show only 13C in 0.0 +/- 5.0 %

Flavonoid Viewer Option:

Show only aglycon

### 1) 元素数での絞り込み

組成式に含まれる C, N, S, P の各元素の数で絞り込みます。各元素記号の左にあるチェックボックスにチェックを入れると、設定が有効になり、左にある下限値、右にある上限値の範囲にあるものだけを一覧に表示します。

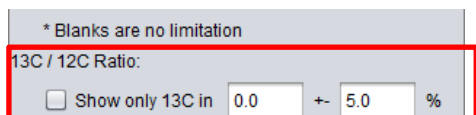
P の直下のテキストボックスには、任意の元素記号を一つだけ入れることができ、その元素の数でフィルタリングできます。



Constraint only for EX-HR2 DB にチェックを入れると、この元素数による制限を、EX-HR2 データベースのみに限定できます。

## 2) $^{13}\text{C}/^{12}\text{C}$ 比での絞り込み

$^{13}\text{C}_1$  の同位体イオンと、 $^{12}\text{C}$  のモノアイソトピックイオンの強度比を用いることで、候補の組成式を絞り込むことができます。設定方法は以下です。



\* Blanks are no limitation  
13C / 12C Ratio:  
 Show only 13C in 0.0 +- 5.0 %

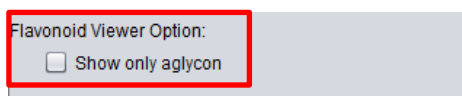
左のテキスト欄に、実測データなどから得られた基準となる  $^{13}\text{C}/^{12}\text{C}$  比を記入します。

右のテキスト欄に、許容誤差 (%) を設定します。

Show only 13C in...のチェックボックスをチェックします。

## 3) フラボノイドのアグリコンだけを検索対象にする

Flavonoid Viewer Option の Show only aglycon にチェックを入れると、Flavonoid Viewer で検索された候補のうち、アグリコン（非糖部）に相当する構造だけを表示します。



Flavonoid Viewer Option:  
 Show only aglycon

## コマンドラインモード

処理したい質量値（およびアダクト組み合わせ）が多数あるときに、コマンドラインモードを使うと、全件について、組成式推定と化合物データベース検索を自動的に行い、結果をテキストファイルとして入手できます。

### 1) 質量値リストの準備

処理する質量値（およびアダクトの組み合わせ）は、テキストファイルとして準備します。以下の3つのファイルのいずれかが使用可能です。

MFSearcher のプログラムセットの sample\_file というフォルダに、3つのファイル形式のサンプルがあります。

a) 質量値のみ (input\_text\_plain.txt)

質量値のみが各行に書かれたテキストファイルを準備します。  
この場合、記載した質量値がそのまま計算に使用されます。

b) 質量値とアダクトの組み合わせ (input\_text.txt)

検出された  $m/z$  値と、その判定されたアダクトが、タブ区切りで各行に記載したテキストファイルを準備します。

アダクトに記載する文字列は、「アダクトの設定」に記載した `adduct.ini` ファイル中で定義されている必要があります。

c) MassChroViewer CheckFile 形式 (input\_mcv\_checkFile.txt)

MassChroViewer というツールで使われているテキストファイルです。MassChroViewer や PowerGetBatch といったソフトウェアで生成されます。

MassChroViewer: <http://www.kazusa.or.jp/komics/software/MassChroViewer/ja>

PowerGetBatch: <http://www.kazusa.or.jp/komics/software/PowerGetBatch/ja>

2) コマンドプロンプト (ターミナル) を開きます。

3) 「`cd`」 コマンドを使い、MFSearcher のプログラムセットの中の、MFSearcher.jar が存在するフォルダへ移動します。

4) 以下のコマンドで、ヘルプを見ることができます。

```
java -jar MFSearcher.jar -h
```

5) 同様に、1) で準備した入力ファイル、出力先のディレクトリ、使用するデータベース、組成式推定での質量許容誤差、データベース検索での質量許容誤差を、コマンドオプションで指定することで、処理を実行します。

例) 質量値とアダクトのペアの入力ファイルの場合 (input\_text.txt)

```
java -jar MFSearcher.jar -i sample_file/input_text.txt -od outdir_plain -db
KG,KN,FL,HM,LM -fppms 1,3,5 -dppms 3,5,15
```

質量許容誤差は、ppm を単位として複数指定できます。小さい誤差範囲から順次検索をし、結果が返ったところで検索を終了します。

化合物データベース検索は、UC2 モードで実行されます。

6) -od で指定したフォルダに、結果ファイルが出力されます。

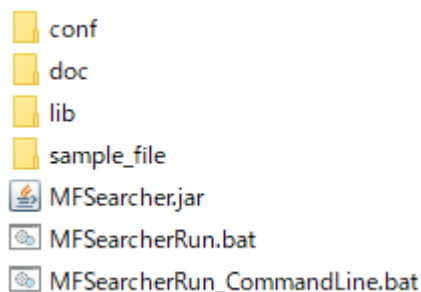
ResFormula\_[入力ファイル名]：組成式を推定した全結果。

ResDatabase\_[入力ファイル名]：データベース検索をした全結果。

ResSummary\_[入力ファイル名]：上記二つを集計したもの。

## アダクトの設定

MFSearcher のプログラムセットの中に conf フォルダがあります。



conf フォルダ中には、adduct.ini というファイルが存在します。

adduct.ini を編集することで、MFSearcher プログラム内で選択できるアダクトの種類やキャプションなどを変更することができます。

adduct.ini はタブ区切りのテキストフォーマットです。編集するにはテキストエディタ（メモ帳）等でファイルを開いてください。

ファイルの構成は以下です。

1)キャプション	2)量数	3)付加する組成式	4)差し引く組成式	5)質量値の補正	6)価数	7)不使用
M	1				0	TRUE
[M]+	1				1	TRUE
[M+H]+	1	H			1	TRUE
[M+NH <sub>4</sub> ]+	1	NH <sub>4</sub>			1	TRUE
[M+Na]+	1	Na			1	TRUE
[M+K]+	1	K			1	TRUE
[M-H <sub>2</sub> O+H]+	1	H	H <sub>2</sub> O		1	TRUE
[M-2(H <sub>2</sub> O)+H]+	1	H	H <sub>2</sub> OH <sub>2</sub> O		1	TRUE
[2M+H]+	2	H			1	TRUE
[2M+NH <sub>4</sub> ]+	2	NH <sub>4</sub>			1	TRUE
[2M+Na]+	2	Na			1	TRUE
[2M+K]+	2	K			1	TRUE
[2M+ACN+H]+	2	C <sub>2</sub> H <sub>3</sub> NH			1	TRUE
[M+2H] <sub>2</sub> <sup>+</sup>	1	H <sub>2</sub>			2	TRUE
[M+2Na] <sub>2</sub> <sup>+</sup>	1	Na <sub>2</sub>			2	TRUE
[M+Na+H] <sub>2</sub> <sup>+</sup>	1	NaH			2	TRUE
[M+3H] <sub>3</sub> <sup>+</sup>	1	H <sub>3</sub>			3	TRUE
[M+3Na] <sub>3</sub> <sup>+</sup>	1	Na <sub>3</sub>			3	TRUE
[M-H] <sup>-</sup>	1		H		-1	TRUE
[M+HCOO] <sup>-</sup>	1	HCOO			-1	TRUE
[M+K-2H] <sup>-</sup>	1	K	H <sub>2</sub>		-1	TRUE
[M+Na-2H] <sup>-</sup>	1	Na	H <sub>2</sub>		-1	TRUE
[M+2Na-3H] <sup>-</sup>	1	Na <sub>2</sub>	H <sub>3</sub>		-1	TRUE
[M+HCOO+Na-H] <sup>-</sup>	1	HCOONa	H		-1	TRUE
[M+HCOO+K-H] <sup>-</sup>	1	HCOOK	H		-1	TRUE
[M-2H] <sub>2</sub> <sup>-</sup>	1		H <sub>2</sub>		-2	TRUE
[M-3H] <sub>3</sub> <sup>-</sup>	1		H <sub>3</sub>		-3	TRUE

1) 任意の文字列を設定できます。

2) 量数とは、多量体になった時の倍数です。[2M+H]<sup>+</sup>の場合には2となります。

5) 組成式では表せない質量の差分を表現するための設定です。特になければ空欄で構いません。

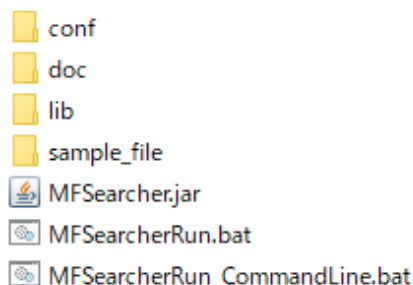
7) この列の設定は特に影響を与えません。

設定を反映させるには、編集後にファイルを上書き保存し、MFSearcher を再起動してください。

### プロキシサーバーの設定

プロキシサーバーを介してインターネットに接続している場合は、MFSearcher が正常に起動しないことがあります。下記の設定をしてから MFSearcher を起動してください。

MFSearcher のプログラムセットの中に conf フォルダがあります。



conf フォルダ中に、`system.ini` というファイルが存在します。`system.ini` ファイルを、テキストエディタ (メモ帳等) で開いてください。デフォルトでは下記のように記載されています。

```
# network proxy setting
network.proxy.enabled=false
network.proxy.server=your.proxy.server
network.proxy.port=8080
```

下記で始まる行の「=」の右側を、下記のように編集します。

<code>network.proxy.enabled</code>	<code>true</code>
<code>network.proxy.server</code>	ネットワーク管理者から指定されたプロキシサーバーの名前 または IP アドレス
<code>network.proxy.port</code>	ネットワーク管理者から指定されたプロキシサーバーのポート番号

編集例)

```
# network proxy setting
network.proxy.enabled=true
network.proxy.server=123.45.67.89
network.proxy.port=1234
```

`system.ini` ファイルを上書き保存し、MFSearcher を再起動してください。

## 公開論文

Sakurai N, Narise T, Sim JS, Lee CM, Ikeda C, Akimoto N and Kanaya S (2018) UC2 search: Using unique connectivity of uncharged compounds for metabolite annotation by database searching in mass spectrometry-based metabolomics. *Bioinformatics* 34:

698-700

[PMID: 29040459]

Sakurai N, Ara T, Kanaya S, Nakamura Y, Iijima Y, Enomoto M, Motegi T, Aoki K, Suzuki H and Shibata D (2012) Application of a relational database system for high-throughput prediction of elemental compositions from accurate mass values. *Bioinformatics* 29 (2): 290-291

[PMID: 23162084]

### お問い合わせ先

(開発元)

国立遺伝学研究所

生命情報・DDBJ センター

櫻井望 E-mail: sakurai AT nig.ac.jp (AT を半角@に変更してください)

本ソフトウェアの初期バージョンは、同開発者にて、公益財団法人かずさ DNA 研究所において開発されました。