

遺伝子系統解析Webシステム

demo ver. 0.1 操作マニュアル

10/09/30 Sh.Onda

1. はじめに

遺伝子系統は進化遺伝学の研究のみでなく、未知の遺伝子の機能の推定や細菌・ウィルスの検査など、医療・薬学・農学・環境関連等にも応用されつつあります。例えば細菌やウィルスのような微生物は検出と種類の判別が難しいので、遺伝子DNA塩基配列を利用し、生物種類の推定を行う場合があります。

このシステムは、生物サンプルから分子生物学的手法によって求めたDNA配列のデータを入力して、既存のデータベースに登録された生物のDNA配列と比較して、生物種類や遺伝子機能の推定等を行うことを目的としています。

○商標などの表記について

本マニュアルでは(R)(C)(TM)などの商標は省略しております。ご了承ください。

本文に登場する会社名、製品名は各社の登録商標または商標です。

2. 動作環境

2.1. ソフトウェア環境

このソフトウェアは、WebブラウザとAdobe **Flash Player** 上で動作します。

WebブラウザはMicrosoft Internet Explorer ver. 6.0 以上、Mozilla Firefox ver. 3.5 以上を奨励します。**Flash Player** は ver. 10 以上を奨励します。

2.2. ハードウェア環境

このソフトウェアは、Webブラウザ上で動作しますので**Flash Player** が動作するものであれば機種は問いませんが、CPU速度1GHz以上、メモリ512MB以上を奨励します。

2.3. 通信環境

このソフトウェアは、サーバとの通信を行うためインターネットに接続されていることを前提とします。回線速度は10Mbps以上を奨励します。

3. 基本機能

(1) DNAデータベース

解析に必要なDNAデータは予めサーバに保存してあります。

- ・生物種類の判定に必要なデータ

16SrDNA (リボゾーム RNA をコードする DNA)

(このバージョンでは、細菌の 16srDNA のみをサポートする。)

DDBJ Web ページの全データから定期的に更新

(2) DNA相同性検索

検査対象のDNAデータと既存のDNAデータの相同性を検索します。

Smith-Waterman アルゴリズム(Ssearch)または FastA による相同性検索

(Smith-Waterman アルゴリズム 参考文献 :

Smith, T.F., Waterman, M.S.: Identification of common molecular
Subsequences, J. Mol. Biol., 147, 195-197(1981))

この部分の基本ロジックは Linux 版の FastA パッケージを使っています。

(3) マルチプルアライメント

既存のDNAデータと検査対象の多重整列を行います。

アルゴリズムは ClustalW

(ClustalW 参考文献 :

Thompson, J. D., Higgins, D. G., T. J. Gibson: *CLUSTALW*,
Nucleic Acids Res. 22, 4673-4680 (1994))

この部分の基本ロジックは Linux 版の ClustalW パッケージを使用しています。

(4) 遺伝子系統樹解析と系統樹の作成

系統樹の作成機能

アルゴリズムは N-J 法とする。

(N-J 法 参考文献 :

Saitou, N., Nei, M. :Mol. Biol. Evol., 4, 406-425(1986))

この部分の基本ロジックは Linux 版の ClustalW パッケージを使用しています。

表示編集機能の詳細は、「遺伝子系統樹描画編集 Web システム」を
参照してください。

4. 基本操作

4.1. 初期画面

初期画面は下記の図のようになっています。

ユーザ登録がまだの方は、新規登録ボタンを押して登録してください。

ユーザ登録がお済みの方はログインボタンを押してログインしてください。

The screenshot shows the main interface of the "DNA系統解析Webシステム Demo ver. 0.1". At the top, there are navigation tabs for "DNA相同性検索", "マルチプルアライメント", and "遺伝子系統樹作成", along with "ログイン", "ログアウト", and "新規登録" buttons. The main content area includes a text input field for "検索配列データ" with a "ファイル読み込" button, a "検索配列名" field, and radio buttons for "検索対象データベース" (Bacteria 16SrDNA, DDB) and "検索アルゴリズム" (FastA, Ssearch). At the bottom, there are buttons for "入力内容の送信", "処理待ちリスト", and "処理結果参照". A status bar at the very bottom indicates "localhost からデータを転送しています..."

4.2. ユーザ登録

ユーザIDとパスワードを入力して登録ボタンを押してください。

ユーザIDとパスワードは半角英数字と・(ハイフン) _ (アンダーバー) が使えます。

大文字と小文字は区別します。(このバージョンではメールアドレス登録はできません)

The screenshot shows a "新規登録" (New Registration) dialog box. It contains the instruction "ユーザID、パスワードを記入して、登録ボタンを押してください。" and three input fields: "ユーザID" (containing "userId"), "パスワード" (containing "****"), and "メールアドレス" (with a message "このバージョンではメールアドレスは登録できません。"). At the bottom, there are "キャンセル" and "登録" buttons.

4.3. メイン画面

ユーザ登録後、またはログインするとメイン画面が使えるようになります。
左上のタブを選択して必要な機能画面を選択してください。
終了するときには右上のログアウトボタンを押してください。

4.4. DNA 相同性検索画面 (タブ)

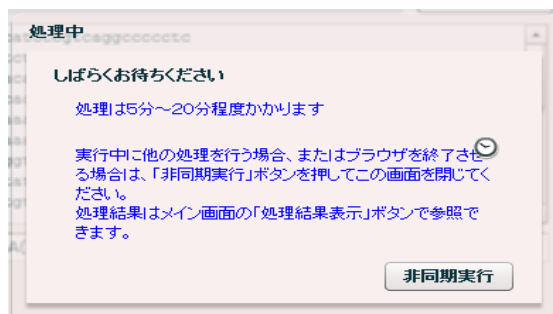
検索対象の塩基配列データは、ファイル読込またはコピー&ペーストで貼り付けることができます。ファイル読込では通常のテキスト形式ファイルと、F a s t Aフォーマットのファイルが使えます。

塩基配列名を入力してから「入力内容の送信」ボタンを押してください。
(塩基配列名は30文字まで、英数字スペース・_・()以外以外の文字は使えません。)

- 検索対象データベース：
このバージョンでは細菌の16SrDNAのみをサポートしています。
- 検索アルゴリズム：
FastA, またはSsearchから選択できます。

処理中には、以下のような画面が表示されます。処理待ち中に他の処理を行う場合は、「非同期実行」ボタンを押して他の処理を行ってください。実行中の処理は「処理待ちリスト」ボタン、結果は「処理結果」ボタンで参照できます。(詳細は、4.8.処理待ちリスト、

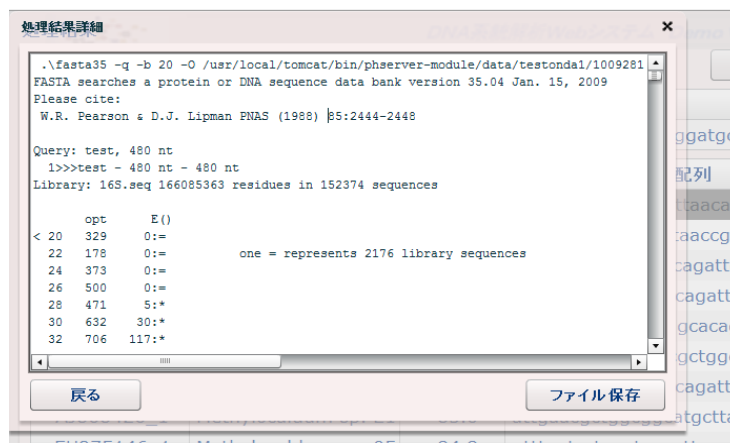
4.7.処理結果参照を参照。)



処理が終了すると、相同性の高い遺伝子の一覧が開きます。

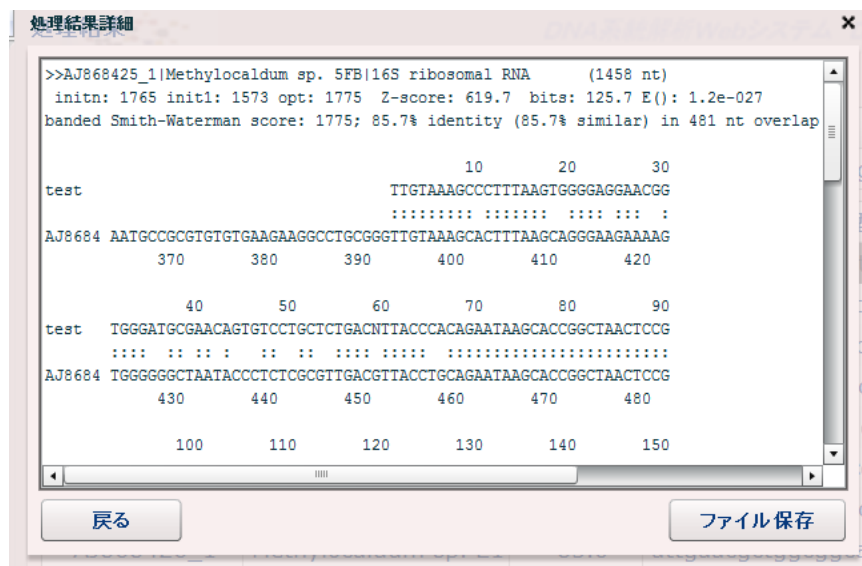


「全データ詳細」ボタンを押すとでは、検索モジュール (FastA, Ssearch) の出力結果全てが表示されます。「ファイル保存」ボタンでパソコンへ保存できます。



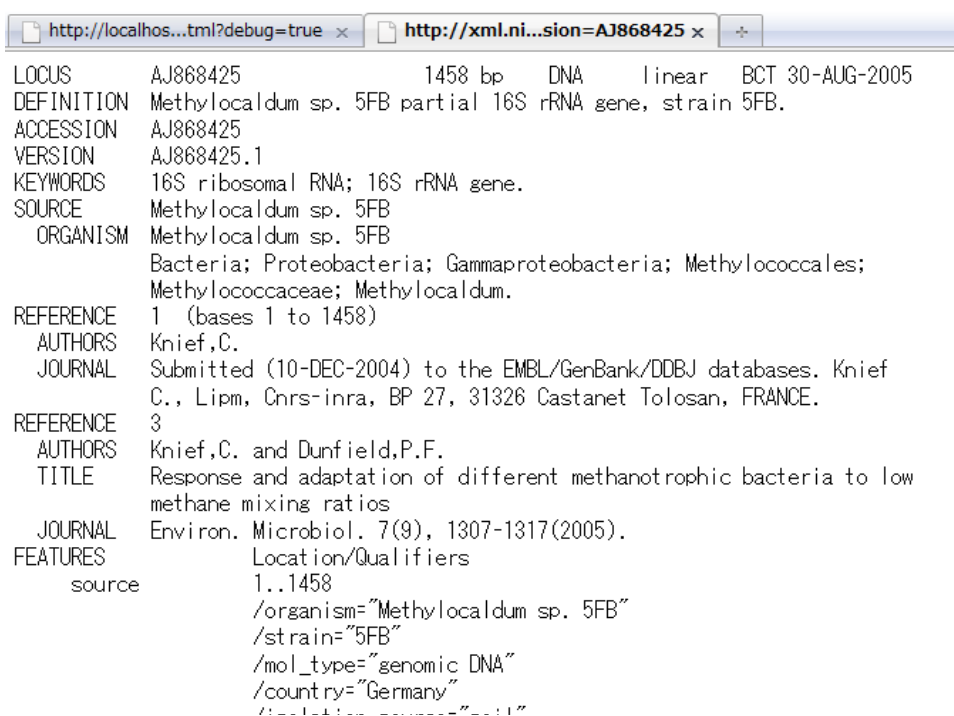
一覧から遺伝子を選んで「相同性の詳細」ボタンを押すと、その遺伝子と検索対象遺伝子の比較結果の詳細情報が表示されます。

「ファイル保存」ボタンでパソコンへ保存できます。



処理結果画面の一覧から遺伝子を選んで「情報の取得」を押すとその遺伝子の情報を DDB J データベースから検索して表示します。別のウィンドウ、または別のタブで開かれますのでこのページを開けたまま他の処理を行えます。

(注意：P o p UP表示を不可にしていると表示されません。)



「ファイル保存」ボタンで結果の一覧とDNA塩基配列が保存できます。（ここで保存できるデータは遺伝子の一覧と各遺伝子のDNAです。）ファイル形式は通常のテキスト形式とExcelなどで読めるCSV形式から選択できます。

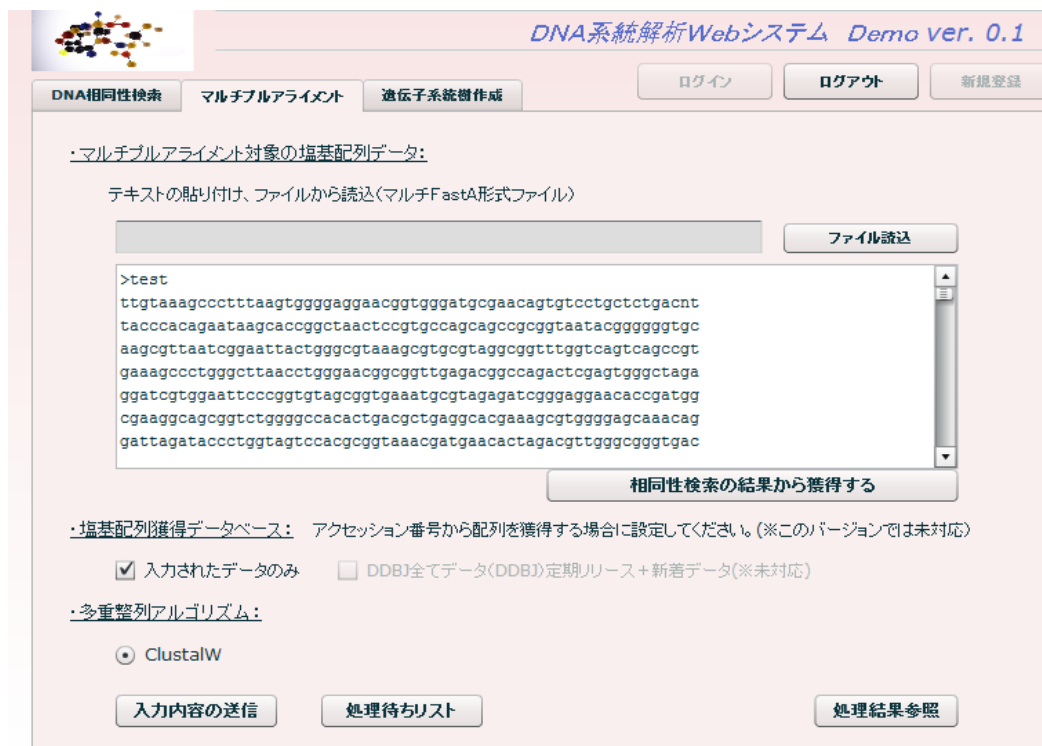
検索結果画面で、一覧のデータをマルチプルアライメント画面に反映させるには「マルチプルアライメントへ」ボタンを押します。（この場合は、マルチプルアライメントのClustalWパッケージの制約のため遺伝子名の空白や()が_に置き換わります。）

4.5. マルチプルアライメント画面（タブ）

塩基配列データは、ファイル読込またはコピー&ペーストで貼り付けることができます。ここではマルチFASTA形式(*)で入力してください。

（塩基配列名は30文字まで、英数字スペース - _ . 以外の文字は使えません。空白は入れないでください。）

実行するには「入力内容の送信」ボタンを押してください。



※マルチFASTA: 以下のような可変長テキスト形式

>生物名 説明など (半角 >で始まり、改行で終わる。)

塩基配列 (次の >までが塩基配列)

>生物名 説明など

塩基配列

塩基配列

例) >sampleNo.1

ATGCATGCATGCATGC

ATGCATGCATGCATGC

>sampleNo.2

ATGCATGCATGCATGC

ATGCATGCATGCATGC

「相同性検索の結果から獲得する」ボタンを押すと、以前に実行したDNA相同検索処理の結果から読み込むこともできます。一覧から結果を選んで「結果表示ボタン」を押してください。（以降の手順は、DNA 相同性検索画面と同じです。）

名称	開始日付	状態	処理選択
testonda1	10/09/27 07:..	END	SIM:FASTA
testonda1	10/09/27 07:..	END	SIM:FASTA
testonda1	10/09/27 07:..	END	SIM:FASTA
testonda1	10/09/27 07:..	END	SIM:FASTA
testonda1	10/09/27 07:..	END	SIM:FASTA
testonda1	10/09/27 07:..	END	SIM:FASTA
testonda1	10/09/27 07:..	END	SIM:FASTA
testonda1	10/09/27 07:..	END	SIM:FASTA

- 塩基配列獲得データベースはこのバージョンでは未対応です。
塩基配列を入力しないでアクセッション番号を指定してDNA塩基配列をデータベースから獲得する場合に設定します。
- 多重整列アルゴリズム：
このバージョンでは、ClustalW パッケージの多重整列機能を使用しています。

処理が終了するとマルチプルアライメント結果の画面が表示されます。
ここでは ClustalW 形式のマルチプルアライメント (※) で表示されています。



※ClustalW形式のマルチプルアライメントデータ

ClustalWなどのオープンソースパッケージからのマルチプルアライメント結果出力は以下のような形式となっています。

CLASTAL W xx.xx.xx

文字列1 スペース 塩基配列・・・
 文字列2 スペース 塩基配列・・・
 文字列3 スペース 塩基配列・・・

文字列1 スペース 塩基配列・・・
 文字列2 スペース 塩基配列・・・
 文字列3 スペース 塩基配列・・・

例)

CLUSTAL W (1.83) multiple sequence alignment

```
swine_H3N2__AY363575_HongKong      -----CTAACCGAGGTCGAAACGTACGTTCTTTCTATCATCCCGTC
swine_H3N2__AY363574_HongKong      -----CTAACCGAGGTCGAAACGTACGTTCTTTCTATCATCCCGTC
swine_H3N2__AY363580_HongKong      -----CTAACCGAGGTCGAAACGTACGTTCTTTCTATCATCCCGTC
```

「ファイル保存」ボタンでパソコンに保存できます。

「系統樹作成へ」ボタンでこのデータを系統樹作成画面へコピーして、系統樹作成画面へ移行します。

4.6. 遺伝子系統樹作成画面 (タブ)

塩基配列データは、ファイル読込またはコピー&ペーストで貼り付けることができます。ここでは ClustalW 形式のマルチブルアライメントデータ(※4.5. マルチブルアライメント画面の説明参照)が使えます。

(塩基配列名は 30 文字まで、英数字スペース - _ . 以外の文字は使えません。

空白は入れないでください。先頭の CLUSTAL . . . の文字は必要です。)

実行するには「入力内容の送信」ボタンを押してください。

DNA系統解析Webシステム Demo ver. 0.1

DNA相同性検索 マルチブルアライメント 遺伝子系統樹作成 ログイン ログアウト 新規登録

・多重整列データ:

テキストの貼り付け、ファイルから読込(CLUSTAL形式のテキストファイル)

ファイル読込

```
CLUSTAL 2.0.12 multiple sequence alignment

EU275146_1|Methylocaldum_sp_0  -----GTTTGATCCTGGCTCAGATTGAACGCTGGCGGCAIGCTT
U89298_1|Methylocaldum_gracile -----AGTTTGATCATGGCTCAGATTGAACGCTGGCGGCAIGCTT
AJ868425_1|Methylocaldum_sp_5  -----AACGCTGGCGGCAIGCTT
U89297_1|Methylocaldum_tepidum -----AGAGTTTGATCATGGCTCAGATTGAACGCTGGCGGCAIGCTT
```

マルチブルアライメントの結果から獲得する

・結果の獲得方法:

テキスト(Ph形式) 系統樹描画編集ツールを自動起動 ※別ウインドウ、または別タブで開きます。PopUpのを設定を許可してください。

・系統樹作成アルゴリズム:

ClustalW N-J法

入力内容の送信 処理待ちリスト 処理結果参照

アルゴリズムは ClustalW の NJ 法を用いています。

「マルチブルアライメントの結果から獲得する」ボタンを押すと、以前に実行したマルチブルアライメントの結果から読み込むこともできます。

過去の処理結果の一覧から選んで「結果表示ボタン」を押してください。

(以降の手順は、マルチブルアライメント画面と同じです。)

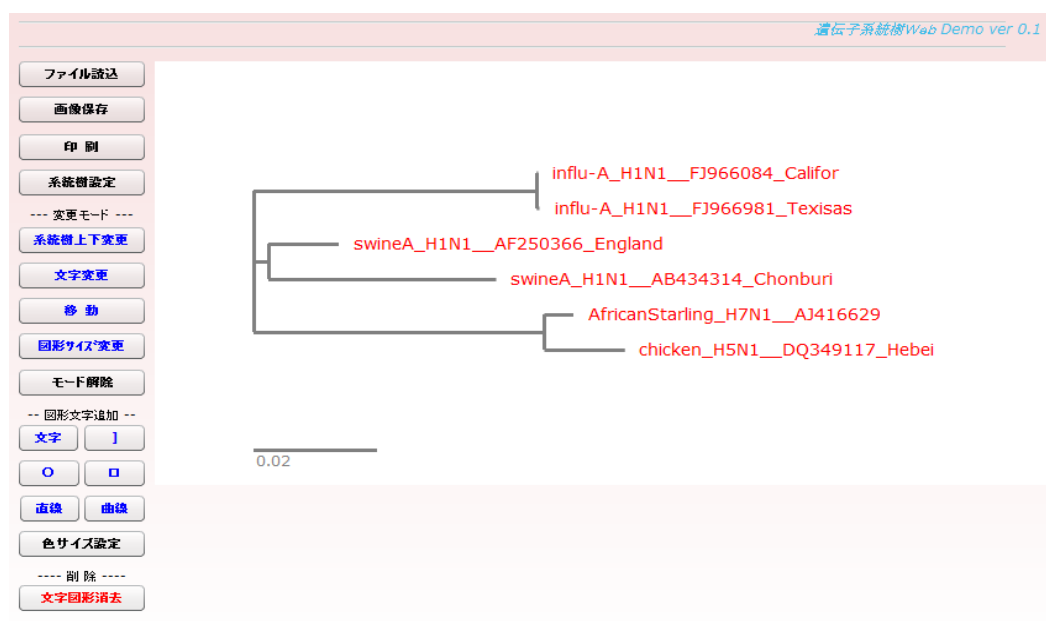
・結果の獲得方法：

テキスト形式（P h形式）のみと系統樹描画編集ツール起動を選択できます。
テキスト形式では結果は以下のように表示されます。



系統樹描画編集ツールは新規ウィンドウ（ブラウザによってはタブ）で開かれます。
以下のように結果をグラフィック表示します。

（注意：P o p U P表示を不可にしていると表示されません。）



詳細は、系統樹描画編集ツールのマニュアルを参照してください。

4.7. 処理結果参照

各画面の「処理結果参照」ボタンを押すとそれぞれの画面で行った過去の処理の一覧が表示されます。



名称	開始日付	状態	処理選択
testonda1	10/09/27 07:...	END	SIM:FASTA
testonda1	10/09/27 07:...	END	SIM:FASTA
testonda1	10/09/27 07:...	END	SIM:FASTA
testonda1	10/09/27 07:...	END	SIM:FASTA
testonda1	10/09/27 07:...	END	SIM:FASTA
testonda1	10/09/27 07:...	END	SIM:FASTA
testonda1	10/09/27 07:...	END	SIM:FASTA
testonda1	10/09/27 07:...	END	SIM:FASTA

名称：実行したユーザIDが表示されます。ログインした人のユーザIDのみ参照できます。

開始日付：処理を送信した日時が表示されます。

状態：END：終了を表します。異常ならERRとなります。

処理選択：SIM:FASTA FastAでの同源性検索
SIM:SSEARCH Ssearchでの同源性検索
ALIGNMENT ClustalWでのマルチプルアライメント
PHTREE ClustalWでの系統樹描画

「削除」ボタンを押すと、選択した過去の処理が削除されます。

「結果」表示ボタンを押すと、過去の実行結果が参照できます。

4.8. 処理待ちリスト

各画面の「処理待ちリスト」ボタンを押すと、現在実行中の全ての処理（他のユーザを含む）の一覧が表示されます。

（処理終了のものはここには表示されていません。各処理の**処理結果参照**ボタンで結果を参照できます。）

実行待ちリスト

名称	開始日付	状態	処理選択
testonda1	10/09/27 11:0	RUN	SIM:FASTA
testonda1	10/10/01 01:1	REQ	ALIGNMENT

戻る 削除

名称：ユーザ I D が表示されます。

開始日付：処理を送信した日時が表示されます。

状態：REQ：処理要求後、処理待ちを表します。RUN：実行中の処理です。

処理選択：SIM:FASTA FastA での相同性検索

SIM:SSEARCH Ssearch での相同性検索

ALIGNMENT ClustalW でのマルチプルアライメント

PHTREE ClustalW での系統樹描画

待ち状態で不要な処理をキャンセルする場合は「削除」ボタンを押してください。
 (実行中の処理は削除できません。)

===== END =====