

# GeneWebⅢ利用の手引き

2011/8/10

大阪大学微生物病研究所附属遺伝情報実験センター

# 目次

<b>1. Geneweb III を利用するためには</b> .....	<b>1</b>
1.1 ブラウザソフトについて.....	1
1.2 Java2 プラグインのインストール.....	2
1.3 コピー&ペーストを許可する設定方法.....	3
<b>2. 起動方法</b> .....	<b>6</b>
<b>3. 画面説明</b> .....	<b>7</b>
3.1 メインウィンドウ.....	7
3.2 メニュー.....	9
<b>4. 基本機能</b> .....	<b>12</b>
4.1 Length, Undo, Clear 機能.....	12
4.2 データフォーマット変換機能.....	13
4.3 データ文字変換機能.....	14
4.4 キーワード検索機能.....	15
<b>5. 核酸配列データ解析</b> .....	<b>16</b>
5.1 系統樹作成機能.....	16
5.2 ホモロジーマトリックス作成機能.....	19
5.3 相違度計算機能.....	21
5.4 2次構造予測機能.....	23
5.5 核酸配列マルチプルアライメント機能.....	25
5.6 含量計算機能.....	27
5.7 相同性検索機能(インタラクティブ処理).....	29
5.8 相同性検索機能(バッチ処理).....	36
5.9 Splice 機能.....	39
5.10 Codon 含有率計算機能.....	41
5.11 ヌクレオチド翻訳機能.....	43
5.12 Open Reading Frame 検索機能.....	46
5.13 制限酵素地図作成機能.....	49
5.14 ゲノムマッピング機能.....	52
<b>6. アミノ酸配列データ解析</b> .....	<b>56</b>
6.1 系統樹作成機能.....	56
6.2 ホモロジーマトリックス作成機能.....	59
6.3 相違度計算機能.....	61
6.4 アミノ酸配列マルチプルアライメント機能.....	63
6.5 含量計算機能.....	65
6.6 相同性検索機能(インタラクティブ処理).....	67
6.7 相同性検索機能(バッチ処理).....	74
6.8 モチーフ検索機能.....	77
6.9 高次構造予測機能.....	78

# 1. Geneweb III を利用するためには

---

## 1.1 ブラウザソフトについて

GeneWebIIIのクライアントソフトは、Web ブラウザ上で動く Java アプレットを利用していますので、GeneWebIIIを利用するためには、Webブラウザ以外に特別なソフトウェアは不要です。ただしブラウザに Java プラグイン (JDK1.4.0 以上) がインストールされている必要があります。

現在動作が確認されているブラウザ (2010 年 7 月)

- Windows Vista  
Internet Explorer 7.0  
FireFox 3.6.3
- Windows XP  
Internet Explorer 6.0
- Mac OS X ver.10.4.10  
Safari 2.04
- UNIX Solaris10  
Mozilla 1.7
- Linux Debian GNU/Linux4.0  
FireFox2.0.0.3

### ! 注意

- ポップアップウィンドウをブロックしている場合は、解除してからご利用下さい。
- Mac OSX をお使いで、入力エリアにコピーペーストができない場合、Safari を 1 回から数回リセットして下さい。

## 1. 2 Java2 プラグインのインストール

ブラウザにJavaプラグインがインストールされていない場合は、<http://java.sun.com/>からインストールして下さい。

### ！注意

- JDK1.4.0 以上をインストールして下さい。
  - MacOS X には標準で Java 環境が含まれています。
  - 2007 年 9 月現在の最新バージョンは、JDK6.0 です。
- Windows XP/Vista  
InternetExplore を立ち上げ、[ツール]→[インターネットオプション]を選択後、[詳細設定]を選択し、[<applet>に JRE1.X を使用(再起動が必要)]にチェックをし、[OK]ボタンをクリックして下さい。
  - Mac OS X  
Safari では、インストールされている Java プラグインの最新版を活用します。  
JDK1.5(5.0)以上のプラグインがインストールされている場合、  
/アプリケーション/ユーティリティ/Java/J2SE5.0/JavaPreference をクリックして、Java 環境設定ダイアログを立ち上げ、Java のバージョンを選択し、[保存]ボタンをクリックして下さい。
  - UNIX Solaris10  
以下のコマンドを実行して Mozilla に Java プラグインをインストールして下さい。  
%ln -s Javaインストールディレクトリ/plugin/sparc/ns7/libjavaplugin\_oji.so\* Mozillaインストールディレクトリ/plugins/

## 1.3 コピー&ペーストを許可する設定方法

GeneWebⅢは Java アプレットを用いて作られています。Java アプレットでは、セキュリティ上の問題からアプレット内とアプレット外とのコピー&ペーストがデフォルトでは出来ないようになりました。コピー&ペーストを出来るようにするためには、Java のポリシーファイルでコピー&ペーストを許可する設定を行う必要があります。

ここでは GeneWebⅢのメインウィンドウ、データ編集領域に配列等の入力データを貼り付けられるように、コピー&ペーストを許可する設定方法を説明します。

この設定は各マシン、各ユーザごとに必要です。設定前に Java を最新版へ Update してください。Java の最新版は <http://www.java.com/ja/download/> からダウンロード出来ます。

- Windows の場合  
ローカルディスク > Program files > Java > jre6 > bin > policytool
- Mac OSX の場合  
MacHD > ライブラリ > Java > Home > bin > policytool

- ① policy tool をダブルクリックして「Policy Tool」画面を開き、ポリシーエントリの追加を選択します。(図 1-1)



図 1-1 Policy Tool 画面

- ② 表示された「ポリシーエントリ」画面で CodeBase に「http://www.gen-info.osaka-u.ac.jp/」を入力し、アクセス権の追加を選択します。(図 1-2)

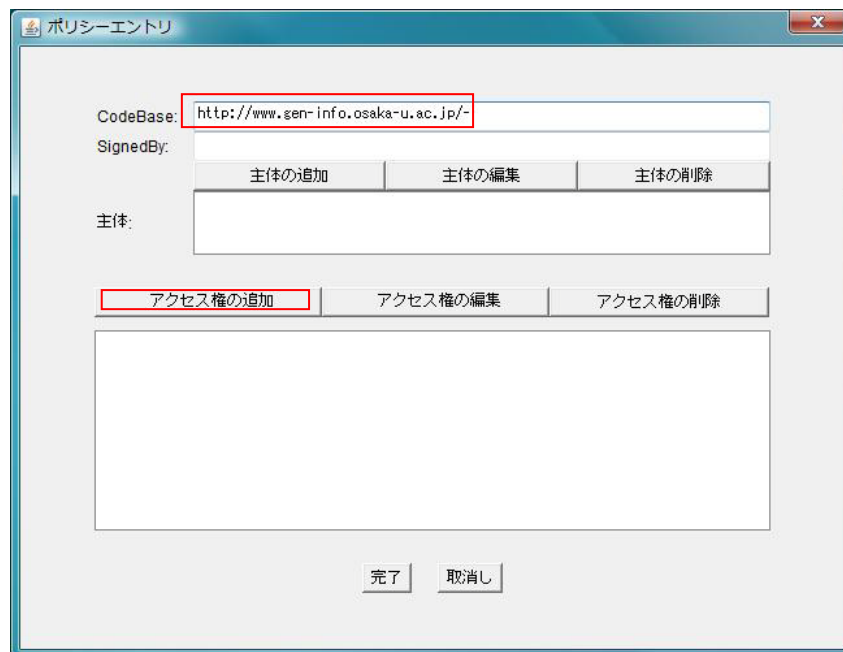


図 1-2 ポリシーエントリ画面

- ③ 「アクセス権」画面で下記を選択し、**了解** を選択します。(図 1-3)

アクセス権: AWTPermission  
ターゲット名: accessClipboard

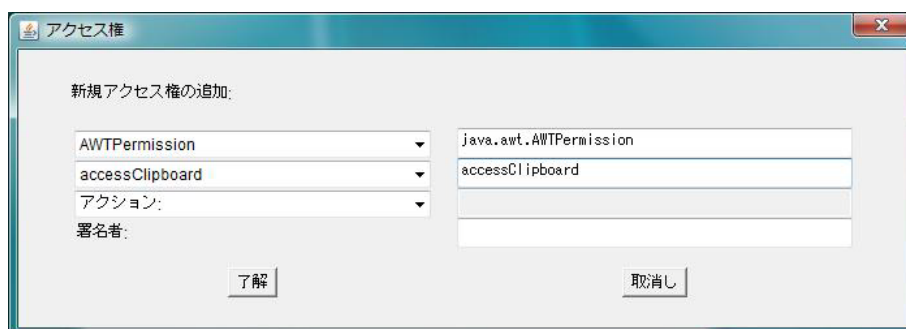


図 1-3 アクセス権設定画面

- ④ **完了** ボタンを選択した後、「ファイル」メニューから保存を選択します。各ユーザのホームディレクトリの下に `.java.policy` を付けて保存します。例：`:/Users/test user/.java.policy` (図 1-4)

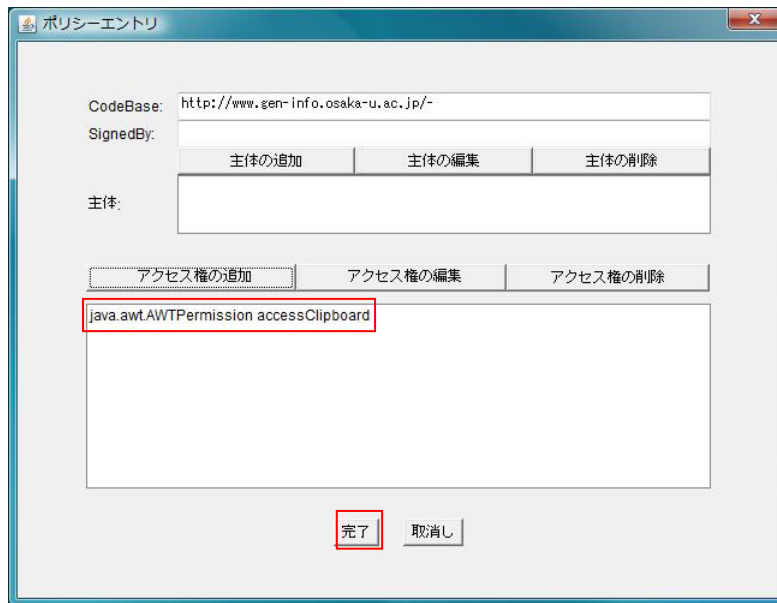


図 1-4 アクセス権設定後のポリシーエントリ画面

ホームディレクトリがわからなければ、コマンドプロンプトで確認します。(図 1-5)  
〈Mac の場合はターミナルで確認〉

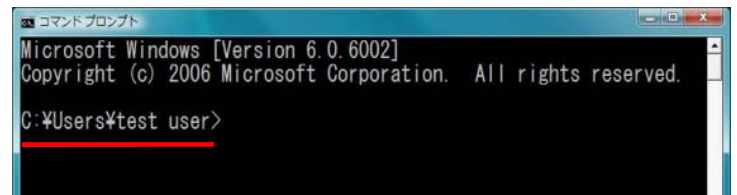


図 1-5 コマンドプロンプト

- ⑤ 「Policy Tool」画面のポリシーファイル欄に保存先が正しく入力されていることを確認して、閉じます。(図 1-6)

- ⑥ インターネットブラウザを再起動します。

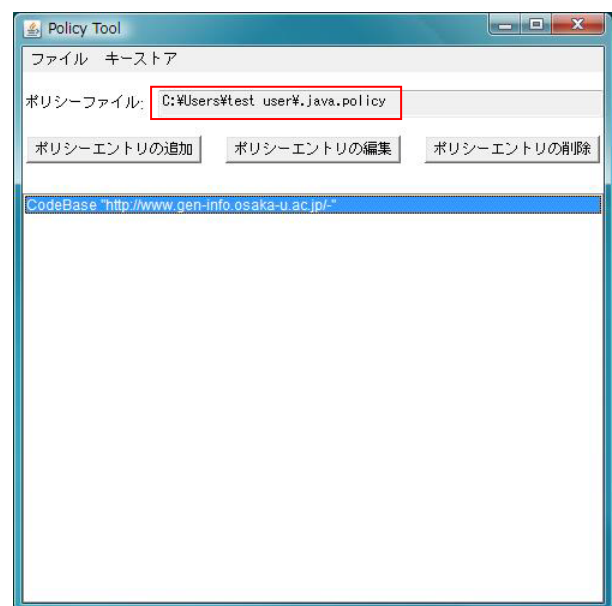


図 1-6 Policy Tool 設定後の画面

## 2. 起動方法

GeneWebⅢを利用するためには GeneWebⅢスタート画面から起動する必要があります。

### <起動方法>

- ① ブラウザを起動し、下記の URL アドレスを入力すると、以下のような GeneWebⅢスタート画面が表示されます。(図 2-1)

<http://www.gen-info.osaka-u.ac.jp/geneweb3/>



図 2-1 GeneWebⅢ スタート画面

- ② 「Start」ボタンをクリックします。別のブラウザ画面が立ち上がり、GeneWebⅢアプレットが読み込まれた作業画面が表示されます。(図 3-1)



# 3. 画面説明

## 3.1 メインウィンドウ

起動した直後、以下のようなメインウィンドウが表示されます。(図 3-1)

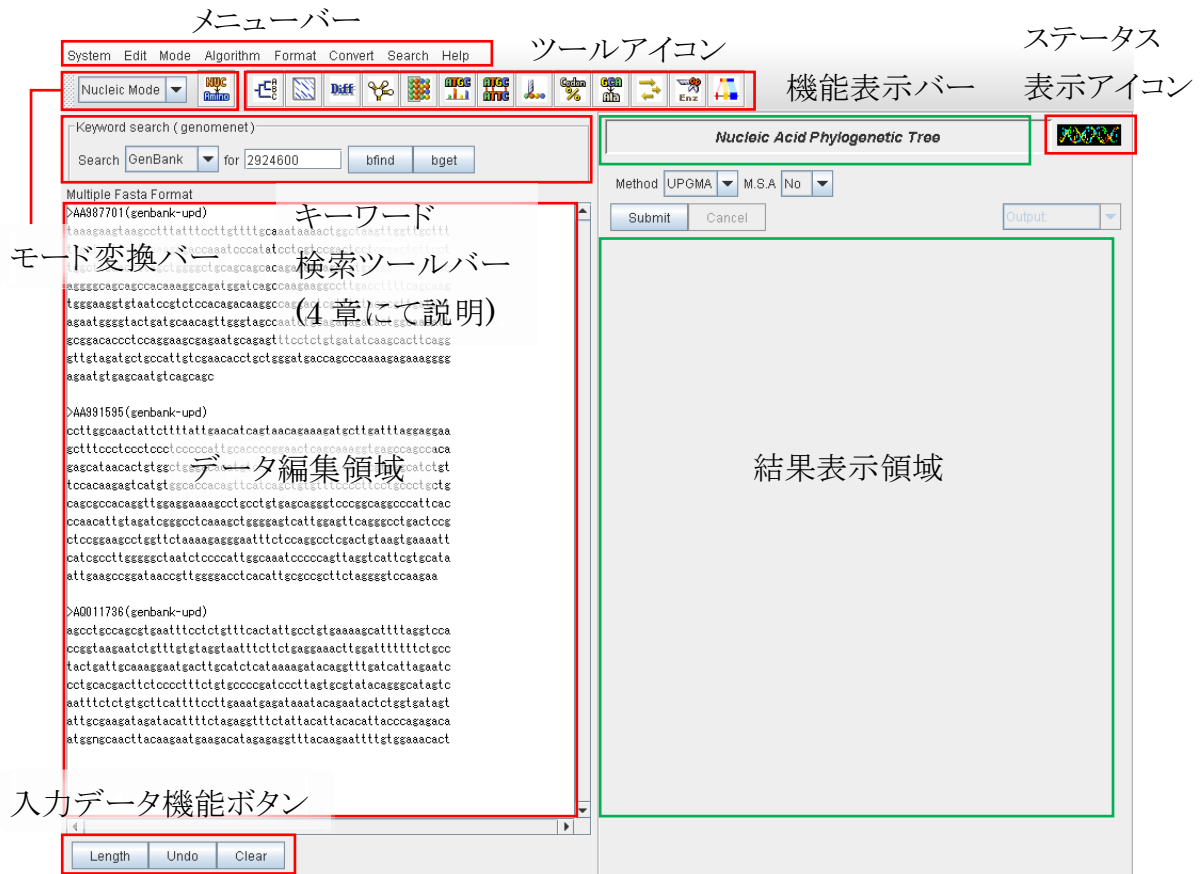


図 3-1 メインウィンドウ

## <全体画面の説明>

### 共通領域（赤枠で囲った部分）

画面上側の部分は、メニューバー、モード変換バー、ツールアイコン、画面左側の部分は、キーワード検索ツールバー、データ編集領域、入力データ機能ボタンからなり、常に表示されます。

データ編集領域にインターネットブラウザに表示された配列データを貼り付けることも可能です。

### 実行領域（緑枠で囲った部分）

画面右側の部分は、機能表示バー、結果表示領域からなり、メニューに従って切り替わり、解析処理のパラメータ指定と結果表示が行われます。

### メニュー

「System」、「Edit」、「Mode」、「Algorithm」、「Format」、「Convert」、「Search」、「Help」があります。(3.2 参照)

### ステータス表示アイコン

サーバとの通信の状態を表示します。通信中はアイコンが回転します。

## 3.2 メニュー

### <System>

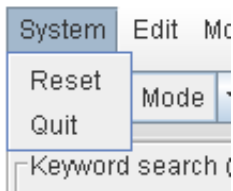


図 3-2-1

- Reset** 起動時の状態に戻します。
- Quit** サーバとの接続を切り、ウィンドウを閉じます。

### <Edit>

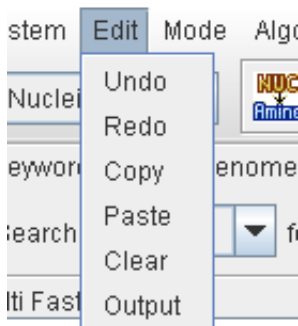


図 3-2-2

- Undo** 編集領域の一つ前の操作を取り消します。(4.1 参照)
- Redo** Undo の操作を取り消します。
- Copy** 編集領域の選択範囲をコピーします。
- Paste** 編集領域の選択範囲をペーストします。
- Clear** 編集領域をクリアします。(4.1 参照)
- Output** 編集領域のデータをブラウザ画面に出力します。ブラウザ画面にて「名前を付けて保存」を選択するなどして、変換したデータを保存することが可能です。

### <Mode>

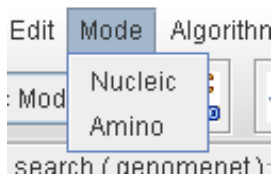


図 3-2-3

- Nucleic** 「Nucleic」モードに切り替えます。このモードでは、塩基配列の解析を行えます。
- Amino** 「Amino」モードに切り替えます。このモードでは、アミノ酸配列の解析を行えます。

選択ボックス      アイコンボタン

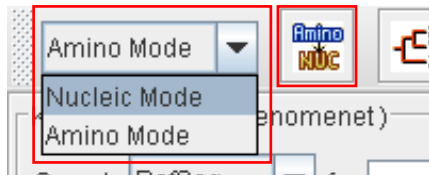


図 3-2-4

モード変換ツールバーの選択ボックス、またはアイコンボタンをクリックしても、同じようにメニューの切り替えができます。(図 3-2-4)

<Algorithm メニュー> -Nucleic mode-

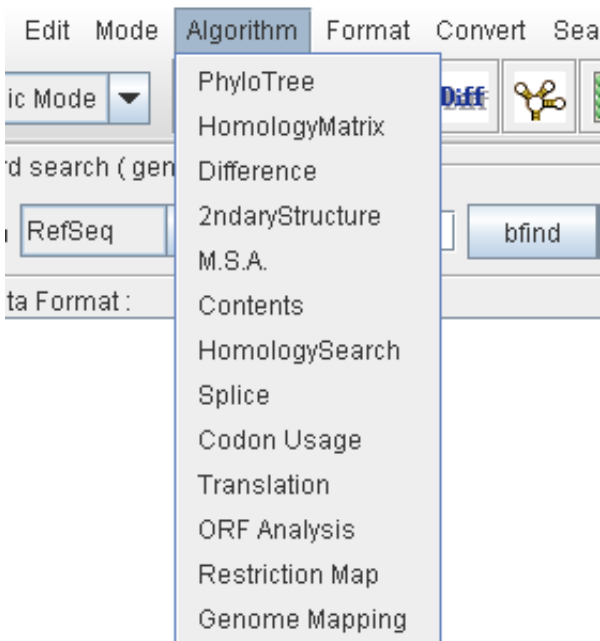


図 3-2-5

\* 詳細は、5章にてご確認ください。

- 5.1 系統樹作成機能
- 5.2 ホモロジーマトリックス作成機能
- 5.3 相違度計算機能
- 5.4 2次構造予測機能
- 5.5 アライメント機能
- 5.6 含量計算機能
- 5.7,8 相同性検索機能
- 5.9 Splice 機能
- 5.10 Codon 含有率計算機能
- 5.11 スクレオチド翻訳機能
- 5.12 Open Reading Frame 検索機能
- 5.13 制限酵素地図作成機能
- 5.14 ゲノムマッピング機能



<Algorithm メニュー> -Amino mode-

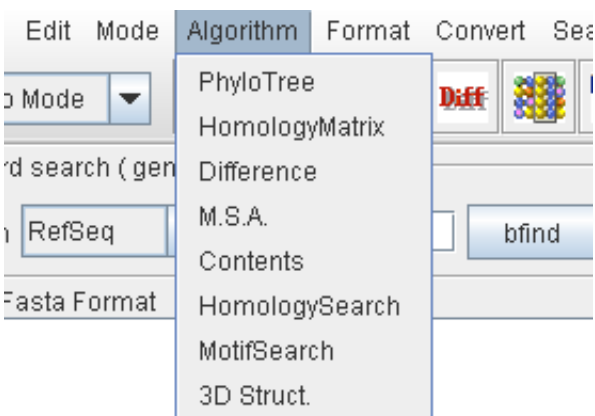
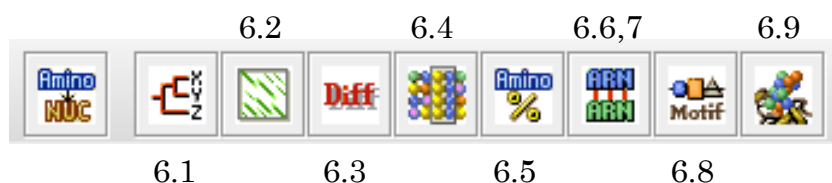


図 3-2-6

\* 詳細は、6章にてご確認ください。

- 6.1 系統樹作成機能
- 6.2 ホモロジーマトリックス作成機能
- 6.3 相違度計算機能
- 6.4 アライメント機能
- 6.5 含量計算機能
- 6.6,7 相同性検索機能
- 6.8 モチーフ検索機能
- 6.9 高次構造予測機能



\* Format メニュー、Convert メニュー、Search メニューは 4 章にてご確認  
 ください。

<help メニュー>

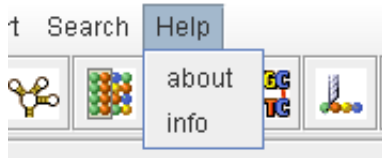


図 3-2-7

about ソフトウェアのバージョン情報を出力します。  
 (図 3-2-8)

info 現在使用しているシステムプロパティを出力します。  
 (図 3-2-9)

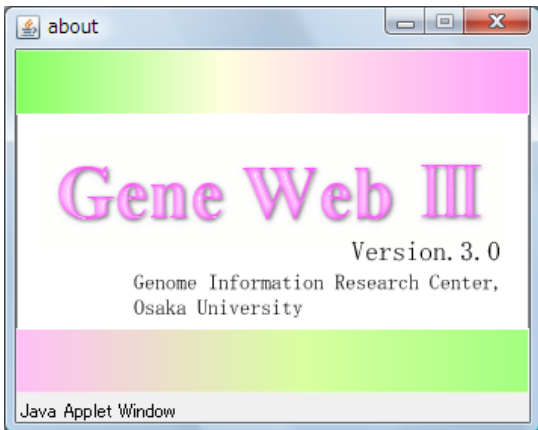


図 3-2-8 バージョン情報

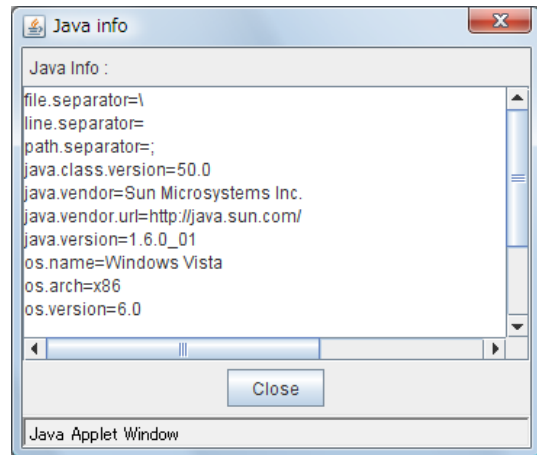


図 3-2-9 システムプロパティ

<補足>

**fasta 形式について**

GenewebIII では、データフォーマット変換機能以外の処理に対する核酸(アミノ酸)配列データは、**fasta 形式で入力**する必要があります。

fasta 形式以外の配列データを入力する場合は、事前に fasta 形式に変換してから処理を行ってください。GenewebIII には、各種配列フォーマットを fasta 形式に変換する機能があります。(4.2 参照)

fasta 形式は、以下のようなフォーマットです。

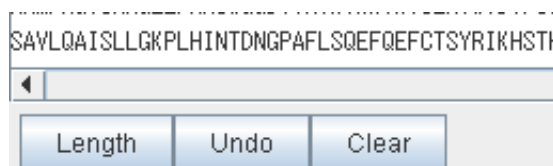
```
>AA987701 (genbank-upd)           ← 先頭行は、“>” で始まるコメント
taaagaagtaagcctttatttccttgttttgca ← 2行目以降が配列データ
tggtttcaaccttagctggggctgcagcagcac
```

```
>AA987701 (genbank-upd)           ← 複数の配列の場合は、続けて記入
taaagaagtaagcctttatttccttgttttgca
tggtttcaaccttagctggggctgcagcagcac
```

## 4. 基本機能

### 4. 1 Length, Undo, Clear 機能

データ編集領域を、1つ前の状態に戻したり、クリアしたり、配列の長さを計算することができます。(図 4-1-1)



① ② ③ 図 4-1-1 機能ボタン

#### <Length>

- ① 「Length」ボタンをクリックすると、データ編集領域に入力された配列データの長さが以下のように別ウィンドウに表示されます。(図 4-1-2)

#### <Undo>

- ② 「Undo」ボタンをクリックすると、データ編集領域を一つ前の状態まで戻ります。

#### <Clear>

- ③ 「Clear」ボタンをクリックすると、データ編集領域をクリアします。

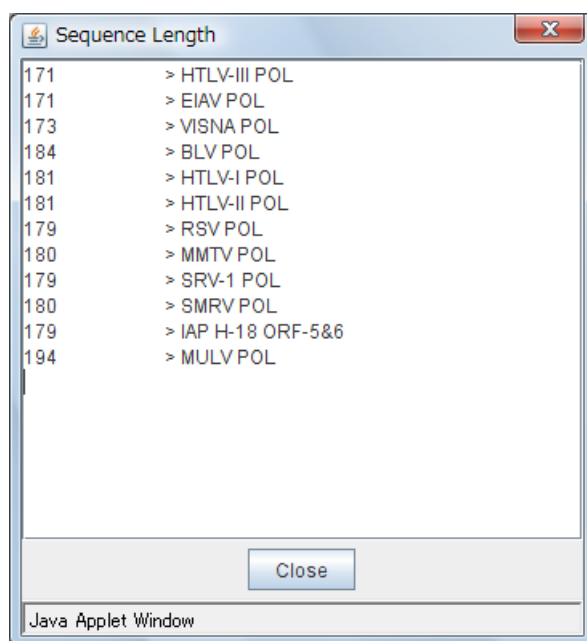


図 4-1-2 配列の長さ

## 4.2 データフォーマット変換機能

入力された核酸・アミノ酸配列データのフォーマットを、指定されたフォーマットへ変換することができます。サーバ内での処理は readseq Ver. 30 Dec92 を使用しています。

(<http://iubio.bio.indiana.edu/soft/molbio/readseq/classic/Readme>)

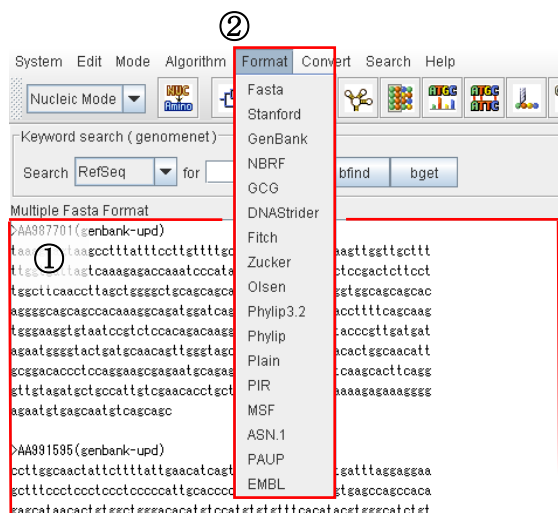


図 4-2-1 Format 変換

① データ編集領域に、配列データを入力します。(図 4-2-1)

② 「Format」メニューでフォーマットを選択すると、指定されたフォーマットにデータが変換されて、以下のように別ウィンドウに出力されます。(図 4-2-2)

③ 変換結果ウィンドウの「Copy」ボタンをクリックすると、変換されたデータが、データ編集領域にコピーされます。

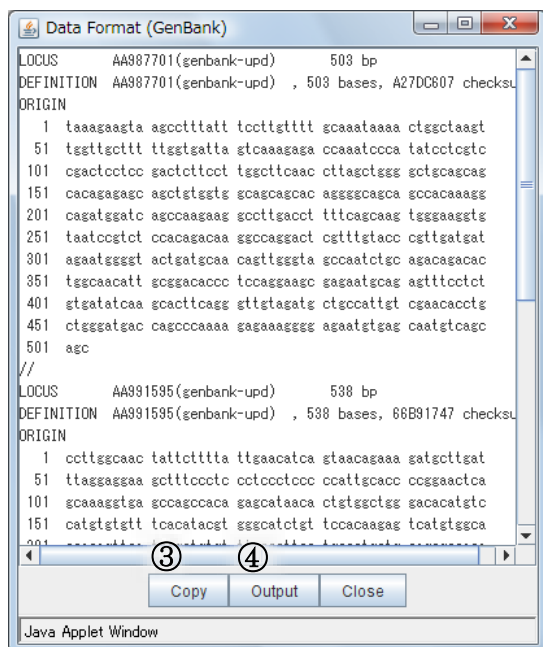


図 4-2-2 Format 変換結果

④ 変換結果ウィンドウの「Output」ボタンをクリックすると、変換されたデータが、別のブラウザ画面に表示されます。ブラウザ画面にて「名前を付けて保存」を選択するなどにより、変換したデータを保存することが可能です。

## 4.3 データ文字変換機能

入力された核酸・アミノ酸配列データに対して大文字変換、小文字変換、Reverse、相補鎖等の変換を行うことができます。



図 4-3-1 データ文字変換

① データ編集領域に、配列データを fasta 形式で入力します。(図 4-3-1)

② 「Convert」メニューで変換方法を選択すると、指定された変換方法に変換されて、以下のように別ウィンドウに出力されます。(図 4-3-2)

### <変換の種類>

Uppercase	大文字へ変換
Lowercase	小文字へ変換
R-Complement	相補鎖変換
Reverse	リバース変換
DNA→RNA	DNA→RNA 変換
RNA→DNA	RNA→DNA 変換

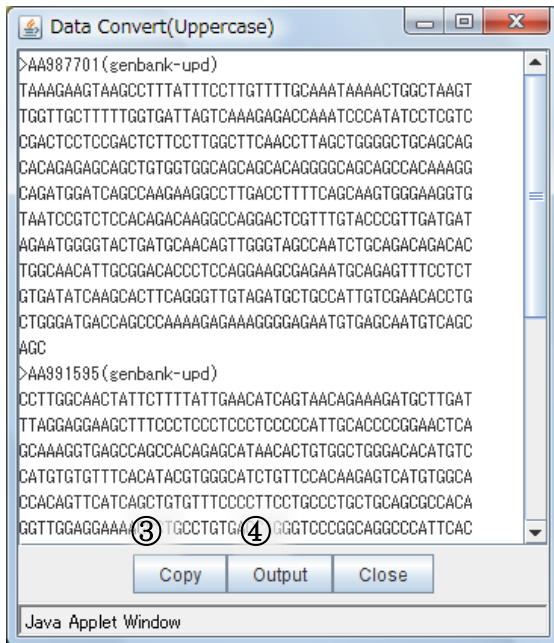


図 4-3-2 変換結果(大文字変換)

③ 変換結果ウィンドウの「Copy」ボタンをクリックすると、変換されたデータが、データ編集領域にコピーされます。

④ 変換結果ウィンドウの「Output」ボタンをクリックすると、変換されたデータが、別のブラウザ画面に表示されます。ブラウザ画面にて「名前を付けて保存」を選択するなどにより、変換したデータを保存することが可能です。



## 4.4 キーワード検索機能

指定されたデータベースに対して、キーワード検索を行います。

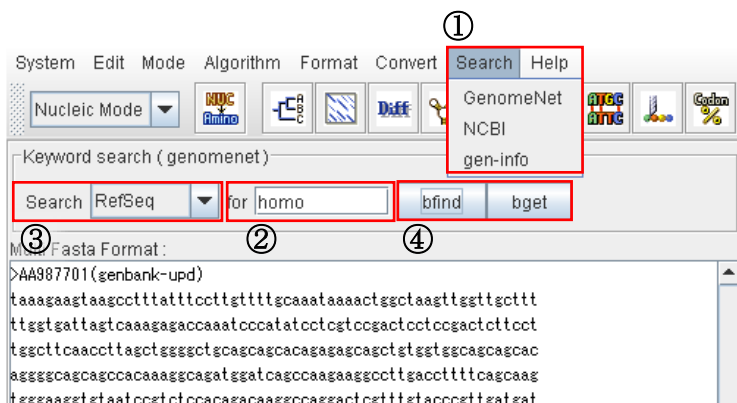


図 4-4 キーワード検索

- ① メニューバーの「Search」メニューから、キーワード検索サイトを選択します。(表 4-4)
- ② キーワードを入力します。
- ③ データベースを選択します。(表 4-4)
- ④ GenomeNet の場合「bfind」ボタン、もしくは「bget」ボタン、それ以外のサイトの場合、「Go」ボタンをクリックすると、結果が別画面に表示されます。

### ！注意

- 「bfind」は、いわゆるキーワード検索で、入力されたキーワードを含む全ての情報が検索されます。
- 「bget」は、遺伝子登録名のみを検索するので、エントリー名、LOCUS、Accession 番号が既知の場合に利用します。  
詳細は、[http://www.genome.jp/dbget/dbget\\_manual.html](http://www.genome.jp/dbget/dbget_manual.html) のマニュアルを参照してください。
- gen-info (遺伝情報実験センター) の dbext 検索システムの場合、エントリー名、LOCUS、Accession 番号により検索を行います。高速な検索が可能です。

表 4-4 キーワード検索データベース一覧

検索サイト	検索システム	データベース
GenomeNet	DBGET 検索システム	Refseq,GenBank,EMBL,UniProt, Swiss-Prot,PIR,PDBSTR,PDB
NCBI	Entrez 検索システム	Nucleotide,Protein
gen-info	dbext 検索システム	GenBank,EMBL,DDBJ,UniProt

# 5. 核酸配列データ解析

## 5.1 系統樹作成機能

複数の核酸配列の系統樹が作成され、結果が表示されます。サーバ内での処理は phylip ver. 3.66 を使用しています。

(<http://evolution.genetics.washington.edu/phylip.html>)

<系統樹作成の方法>

- ① 「Mode」メニューで「Nucleic」を選択し、「Algorithm」メニュー、またはツールアイコンで「Phylo Tree」を選択します。(図 5-1-1)
- ② データ編集領域に核酸配列データを multi fasta 形式で入力します。複数の配列を入力する必要があります。

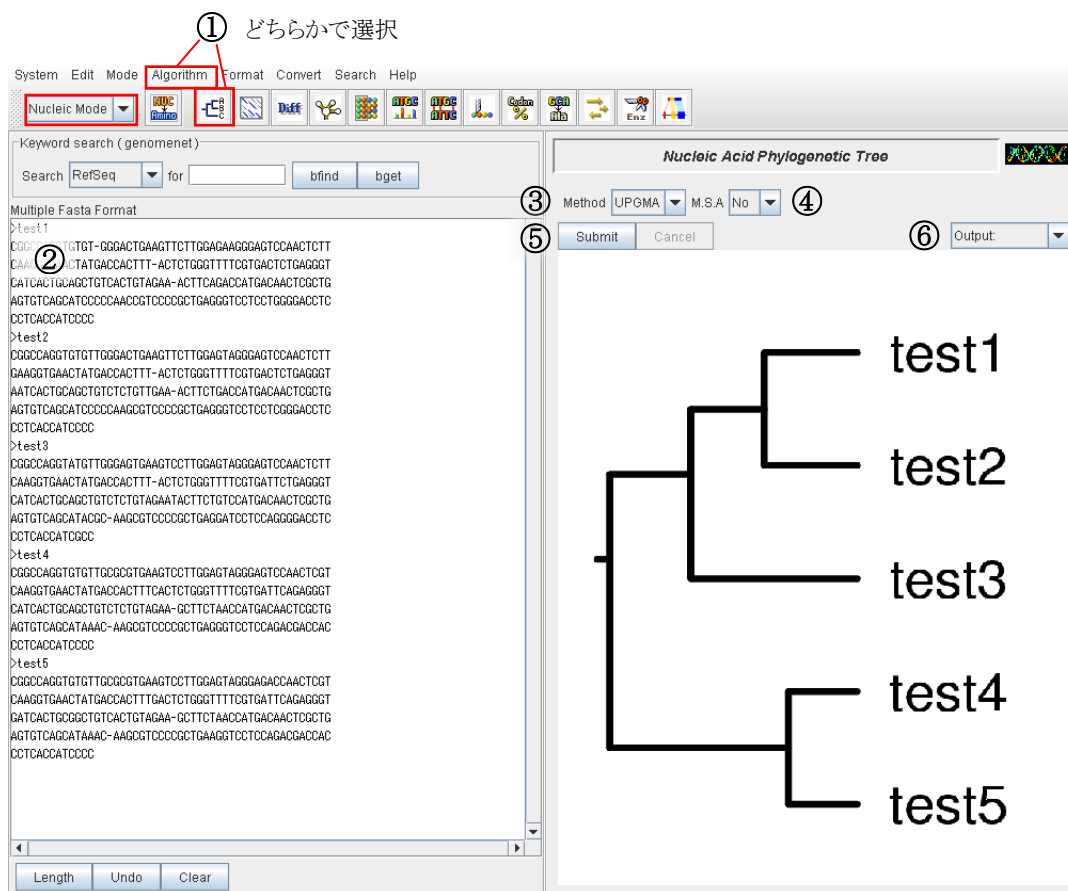


図 5-1-1 系統樹作成画面

③ 系統樹作成方法を選択します。(図 5-1-2)

UPGMA(平均距離)法と NJ(近隣結合)法が選択可能です。

\*UPGMA 法、NJ 法は配列対間の遺伝距離に基づく方法です。

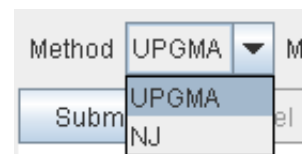


図 5-1-2

④ あらかじめマルチプルシーケンスアライメントを実行するか選択します。(図 5-1-3)

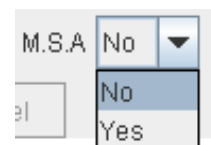


図 5-1-3

**! 注意**

- 入力する核酸配列がアライメントされていない場合、必ず「Yes」を指定してください。
- 各配列の長さが違い、MSA が No の場合、警告メッセージが出力されます。
- MSA が Yes の場合、サーバ内で ClustalW を用いてアライメントした後、系統樹が作成されます。

⑤ 「Submit」ボタンをクリックすると、系統樹が結果表示領域に表示されます。

ここで表示される系統樹は、入力された配列から作成した推定系統樹です。

⑥ プルダウンメニューから以下の出力形式を選択すると、別画面に結果が表示されます。(図 5-1-4)

gif	一般的な画像形式のひとつ
ps	高品質な印刷が可能な画像形式
distance matrix	配列間の進化的距離の行列
tree	系統樹の Newick 形式のテキスト表記

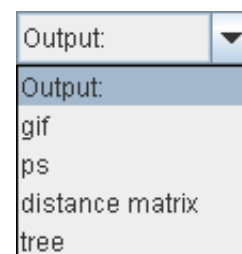


図 5-1-4

< distance matrix 形式 >

distance matrix で表示されている値は、2つの配列で異なるサイトの割合(P)に対して、Jukes-Cantor モデルに従って補正を行った補正距離(D)です。

5					
test1	0.000000	0.039097	0.070168	0.102373	0.119010
test2	0.039097	0.000000	0.069818	0.101851	0.123993
test3	0.070168	0.069818	0.000000	0.075063	0.101851
test4	0.102373	0.101851	0.075063	0.000000	0.028989
test5	0.119010	0.123993	0.101851	0.028989	0.000000

$P = m/n$	$P$ : 2つの配列で異なるサイトの割合
	$m$ : 2つの配列の間で異なる塩基の数
	$n$ : 配列の長さ
$D = -3/4 \log(1 - 4/3P)$	

<tree 形式>

tree 形式で表示されている値は、各アルゴリズムに基づいて計算された枝長です。

```
((test1:0.01955,test2:0.01955):0.01545,test3:0.03500):0.01702,  
(test4:0.01449,test5:0.01449):0.03752);
```

## 5.2 ホモロジーマトリックス作成機能

2つの核酸配列の相同性を、ホモロジーマトリックスで表示します。  
サーバ内では blast ver 2.2.16 を使用して、相同性を検出しています。

<ホモロジーマトリックス作成方法>

- ① 「Mode」メニューで「Nucleic」を選択し、「Algorithm」メニュー、またはツールアイコンで「Homology Matrix」を選択します。(図 5-2)
- ② データ編集領域に核酸配列データを fasta 形式で入力します。  
Sequence1 と Sequence2 に、それぞれひとつずつ配列を入力する必要があります。
- ③ 相同性の検出には、blast を使用していますので、blast のオプションである E-value と Word size の値を入力します。

### ！注意

- Word size の値は 7 以上の整数値とします。
- E-value は実数値とし、指数表記での指定もできます。例) 1E-2 (0.01)

① どちらかで選択

System Edit Mode Algorithm Format Convert Search Help

Nucleic Mode

Keyword search (genomenet)

Search RefSeq for bfind bget

Sequence 1:  
>Homo sapiens aggrecan 1 gi=693593  
CDDGTTGTTGGGACTGAAGTCTTTGGAGAAGGGAGTCCAACCTCTCAAGGTGAACATGACCACTT  
TACCTTTTTCGCTACTCTGAGGGTCATCACTGCAGCTGTCACTGATAGAACTTCAGACCATGACAA  
CTGCTGAGTCTCAGCATCCCAACCGTCCCGCTGAGGGTCTCTCTGGGACCTCCCTCACCATCCCC  
TGCTATTTTCATCAGCCCATCCACCTGTGACCAACCGCCCTTCTACCGCCCACTGGCCCAAGAATCA  
AGTGGAGCCGTGCTCCAGAGCAGAAGGAGGTAGTCTGCTGCTGGGCCACTGAAGGCCCTGGGGTCAA  
DAGTGCCTATCAGGAAAGTCTCACTGCCAACTACCCGGCCATCCCCAGTGACGCCACCTGGAAGTC  
CAGAGCCTGGGCTCAAATGACTCTGGGGTCTACCGCTGGGAGGTGATGCAATGGCATGGAGGACAGCGAGG  
CCACCCCTGGAAAGTCTGGTGAAGGACATGCTGTTCCATACAGAGCCATCTCTACACGCTACACCCCTGA  
CITTTGACAGGGCCAGCGGGCCCTGCCAGAACAGTCCCATCATTGCCACCCCTGAGCAGCTCCAGGCC  
GCCTACAGAGAGGCTTCCACCAGTGTGACCGGGCTGGCTGACAGACAGCTGCAGATACCCGATCC  
ACACTCCCGGGAAGGCTCTATGAGACAGAGTGAAGTTCCCTGCTGTGAGGAGCTATGGCATCCGAGA  
CACCAACGAGACCTATGATGCTGCTTCCCGAGGAGATGGAGGTTGAGGCTTTTATGCAACATCT  
CCAGAGAAATTCACCTTCCAGGAGCAGCCAAATGAGTGGCGGGCTGGGTGGCCGGCTGGCCACCAAGG  
GCCACGCTACCTGGCTGGCAGCTGGCATGGACATGTGACAGCGCGCTGCTGGCCAGCCGACGCT

Sequence 2:  
>Mus musculus aggrecan gi=878541  
CDDDDGGGCAAGCATCCCTGCTGAGGCGAGCAAGACCGGGGCTGGCAGCCCAAGAAATCCCTAGCT  
GCTTAAAGGACTGAAGTCTTTGGAGGAGCGAGTCCAACCTCTCAAGCTGAACATGACCACT  
TTTCTTTGCTACTCTGAGGGTCACTGCTCAGTGTGATCAGAAAGTTCAGACCATGAC  
AACTCACTGAGCGTGAACATCCCTCAACCATCCCATTTGAAGTCTCTTAGGGTCTTCCCTCACCATCC  
CTGCTACTTTCATCGACCCATGATCCTGTCAGCACTGGCCCTCCACTGGCCCTCCACCCCAAGAAAT  
CAAGTGGAGCGCTTTTCCAGGAAAGGAGGTTGACTGCTGGTGGCCACTGAAGGACAGGTTGAGTCA  
AACAGCTCTACCAAGCAAGGTTGCTGCCCAACTACAGCCATCCCAAGGATGCTACTACCTTTGGAGA  
TCCAGAACCTTGGCTCCATGACTCTGGGATCTACCGCTGGAAGTATGACAGCCATCCAGGACAGCGA  
AGCCACCCCTGGAGTCAATAGTGAAGGATTTGTTTCCACTACAGGCTATTTCCACAGCTACACCCCTG  
GACTTTGATGAGCAGCAGCGCTTGGCTACAGAACAGCCCATCATGGCCACCCAGAACACTGCAGG  
CTGGCTATGAGGATGGCTTCCACAGTGGATGAGGCTGGCTGGCTGACAGACAGCTCAGATACCCCAT  
CCACACGCCCGGGAAGGTTGCTATGTTGACAAAGGAGGTTCCCTGGAGTGAGAACTACGAAATCCGG  
GACACCAATGAGACATGATGTTGCTGCTTCCCTGAGGAGATGGAGGTTGAGCTTTTATGGGACAT  
CCDDCAGAAATTCACCTTCCAGGAGCCAGCCAGTGGCGAGGCTGGGGGACCGCTTGGCCACAC

Length Undo Clear

Length Undo Clear

Nucleic Acid Homology Matrix

BlastParams: E-Value 10 Word Size 11

Submit Cancel Default ps gif

Homo (1-7434)

Mus (1-7382)

図 5-2 ホモロジーマトリックス表示画面

- ④ 「Submit」ボタンをクリックすると、ホモロジーマトリックスが結果表示領域に表示されます。
- ⑤ 「ps」ボタン、または「gif」ボタンをクリックすると、別画面に結果が表示されます。

## 5.3 相違度計算機能

複数の核酸配列の相違度が計算され、結果が表示されます。サーバ内での処理は、sqdifを使用しています。(Miyata,T & Yasunaga,T.(1980), J.Mol.Eval, 16, 23-26)

<相違度計算の方法>

- ① 「Mode」メニューで「Nucleic」を選択し、「Algorithm」メニュー、またはツールアイコンで「Difference」を選択します。(図 5-3)
- ② データ編集領域に、アライメントされている核酸配列データを fasta 形式で入力します。複数の配列を入力する必要があります。
- ③ 必要であれば、Coding Region を入力します。開始位置数と終了位置数の間に「-」を入れ、空白もしくはタブの後ろにコメントを入力します。複数入力も可能です。(コメントは省略可能です。)

例) 57-500 comment1  
521-610 comment2

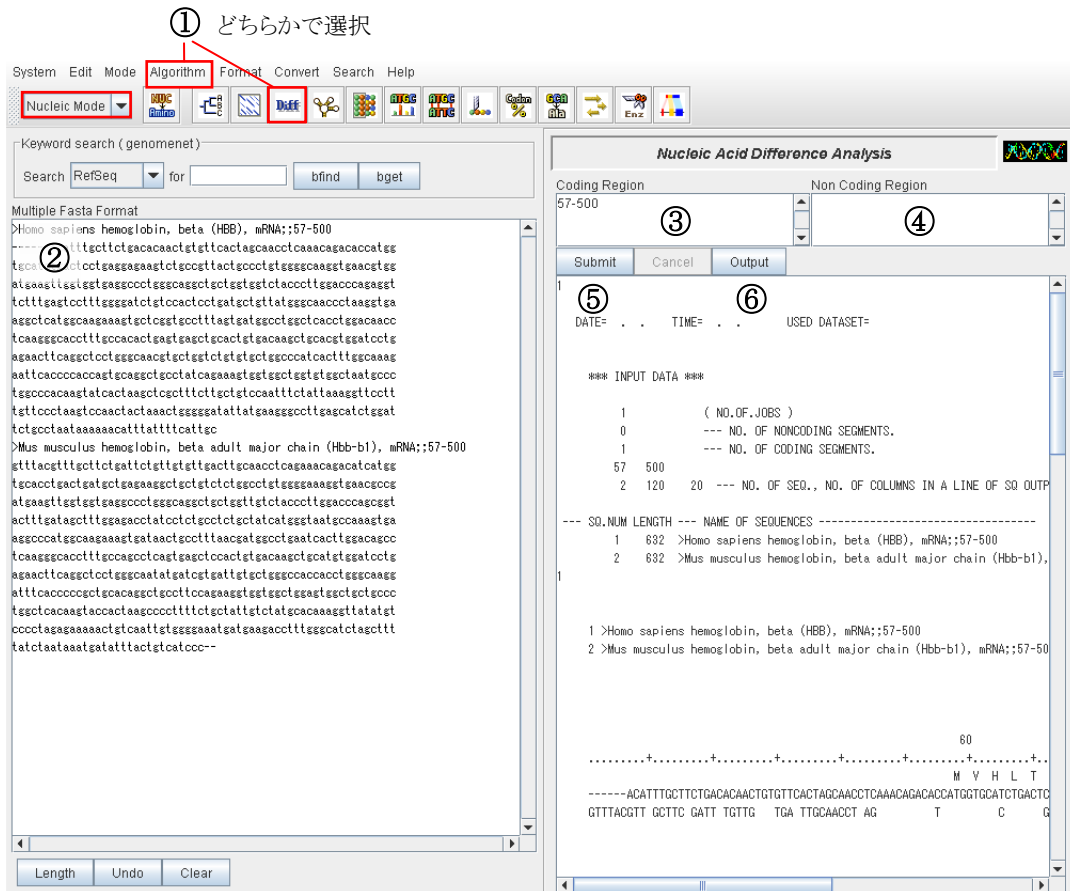


図 5-3 相違度計算機能画面

- ④ 必要であれば、**Non-coding Region** を入力します。  
(**Coding Region** の入力方法と同様です。)
- ⑤ 「**Submit**」ボタンをクリックすると、相違度が結果表示領域に表示されます。
- ⑥ 「**Output**」ボタンをクリックすると、結果が別ブラウザ画面に表示されます。  
ブラウザ画面にて「名前を付けて保存」を選択するなどにより、出力結果を保存することが可能です。



## 5. 4 2次構造予測機能

核酸配列が RNA に変換され、2次構造の予測結果が表示されます。サーバ内での処理は、RNA fold ver. 1.4 を使用しています。

(<http://www.tbi.univie.ac.at/~ivo/RNA/RNAfold.html>)

<2次構造の予測方法>

- ① 「Mode」メニューで「Nucleic」を選択し、「Algorithm」メニュー、またはツールアイコンで「2ndary Structure」を選択します。(図 5-4-1)
- ② データ編集領域に核酸配列データを single fasta 形式で入力します。複数の配列を入力することはできません。
- ③ Temp (温度°C) を整数で入力します。

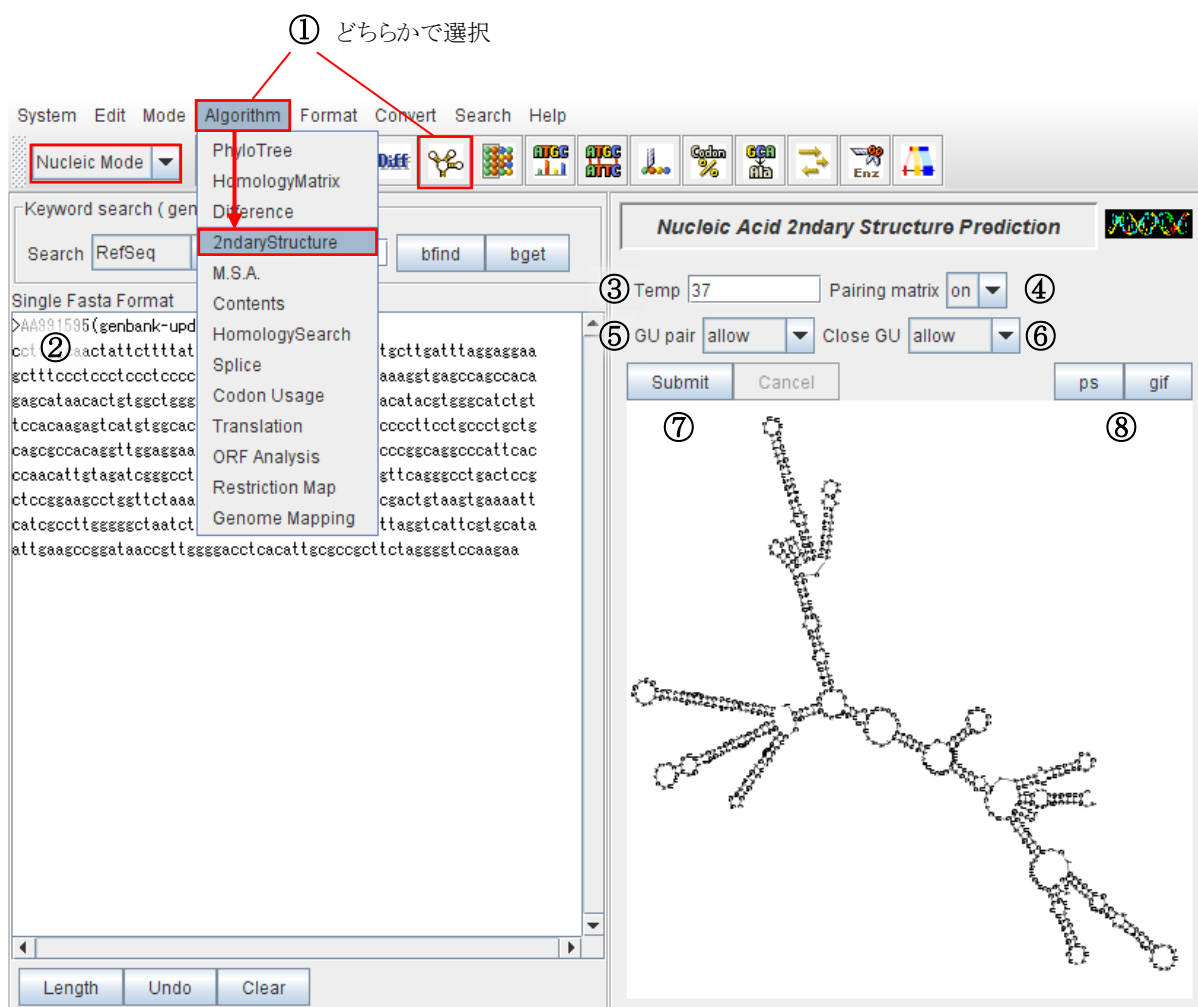


図 5-4-1 2次構造予測機能画面

④ Pairing matrix を指定します。(図 5-4-2)

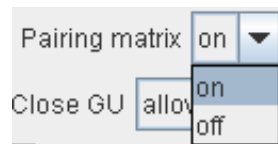


図 5-4-2

⑤ GU pair を許可するかを指定します。(図 5-4-3)

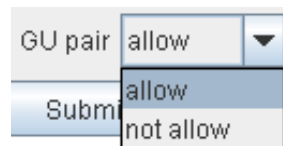


図 5-4-3

⑥ Close GU を許可するかを指定します。(図 5-4-4)

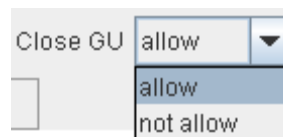


図 5-4-4

⑦ 「Submit」ボタンをクリックすると、結果表示領域に2次構造のイメージが表示されます。

⑧ 「ps」ボタン、または「gif」ボタンをクリックすると、別画面に結果が表示されます。

## 5.5 核酸配列マルチプルアライメント機能

複数の核酸配列データをアライメントし、結果を表示します。サーバ内での処理は ClustalW ver. 1.81 (<http://www.ebi.ac.uk/clustalw>) か、MAFFT Ver. 6.705b (<http://align.bmr.kyushu-u.ac.jp/mafft/software>) を使用しています。

<アライメントの方法>

- ① 「Mode」メニューで「Nucleic」を選択し、「Algorithm」メニュー、またはツールアイコンで「M.S.A.」を選択します。(図 5-5-1)
- ② データ編集領域に、核酸配列データを multi fasta 形式で入力します。複数の配列を入力する必要があります。

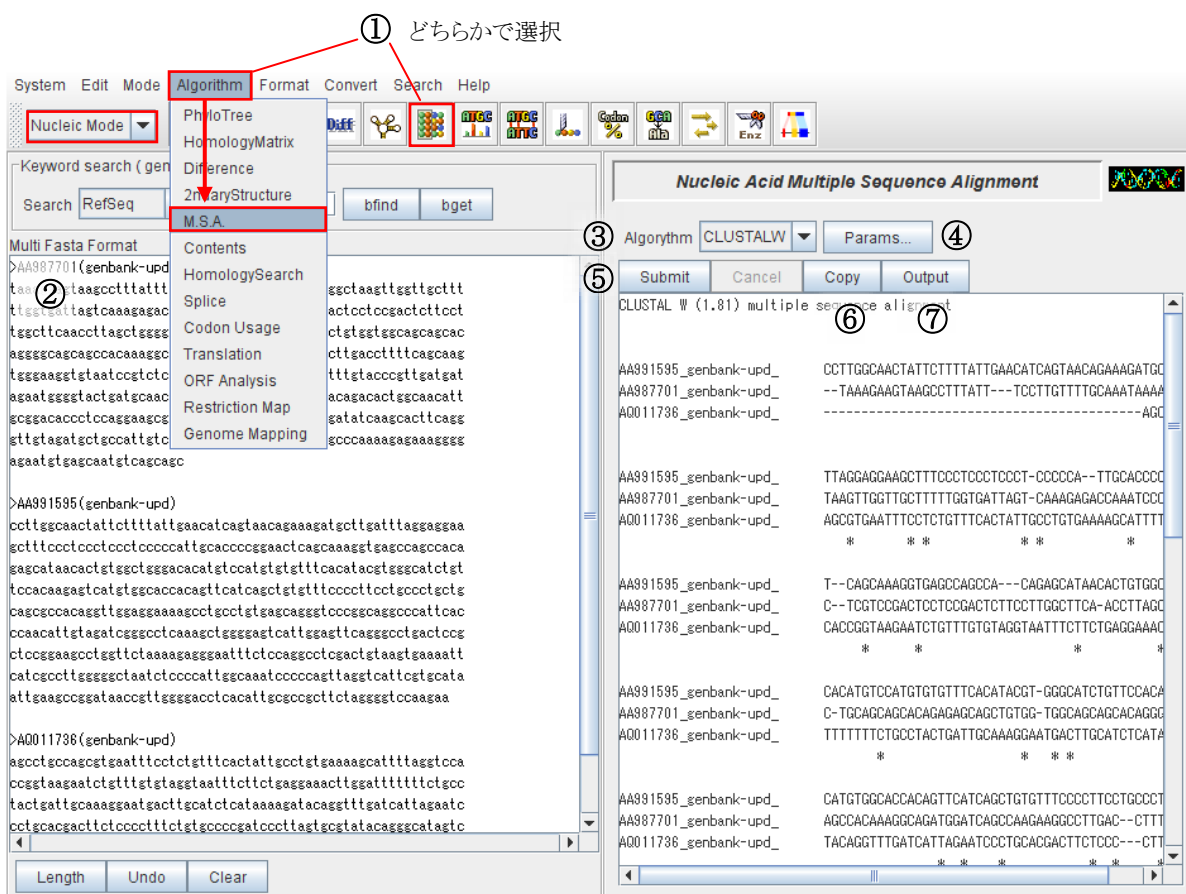


図 5-5-1 アライメント機能画面

- ③ アルゴリズムを選択します。CLUSTALW か MAFFT を選択します。(図 5-5-2)

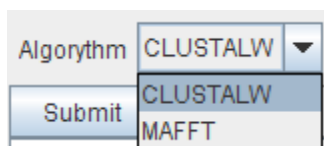


図 5-5-2 アルゴリズム選択画面

- ④ 「Params…」ボタンをクリックし、パラメータを設定します。(図 5-5-3、図 5-5-4)

ClustalW の場合、%identify の値は整数、それ以外の値は実数とし、実数(整数)以外が入力されるとデフォルト値に戻ります。

MAFFT の場合、一番下の項目「Max num of ...」の値は 1 以上の整数、それ以外の値は実数とし、実数(整数)以外が入力されるとデフォルト値に戻ります。

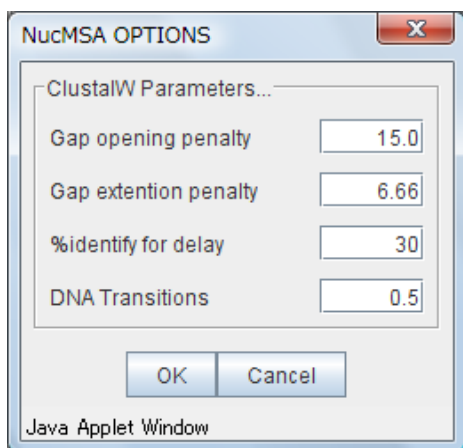


図 5-5-3 ClustalW パラメータダイアログ

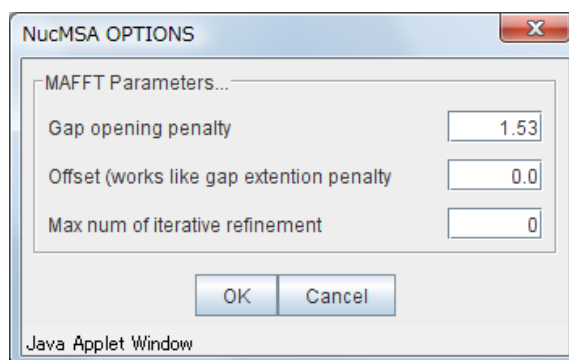


図 5-5-4 MAFFT パラメータダイアログ

- ⑤ 「Submit」ボタンをクリックすると、アライメント結果が結果表示領域に表示されます。

#### < 共通領域へのコピーの方法 >

- ⑥ 「Copy」ボタンをクリックすると、アライメント結果が fasta 形式に変換されてデータ編集領域に表示されます。
- ⑦ 「Output」ボタンをクリックすると、アライメント結果が別ブラウザ画面に表示されます。ブラウザ画面にて「名前を付けて保存」を選択するなどにより、出力結果を保存することが可能です。

## 5.6 含量計算機能

核酸配列の各塩基の含有量(率)が計算され、結果がテーブルとグラフで表示されます。

<含量計算の方法>

- ① 「Mode」メニューで「Nucleic」を選択し、「Algorithm」メニュー、またはツールアイコンで「Contents」を選択して機能を切り替えます。(図 5-6-1)
- ② データ編集領域に核酸配列データを fasta 形式で入力します。複数の配列を入力することも可能です。
- ③ ウィンドウサイズ (Window Size)、シフト値 (Shift) の設定をします。例えば、Window Size = 20, Shift = 1 の場合、まず 1 bp~20 bp の含量を計算し、次に 2 bp~21 bp の含量を計算というように、Shift 値分ずらしながら Window Size 分の計算を行います。

① どちらかで選択

配列全体の  
含量テーブル

System Edit Mode **Algorithm** Format Convert Search Help

Nucleic Mode ▼

- PhyloTree
- HomologyMatrix
- Difference
- 2ndaryStructure
- MVA
- Contents**
- HomologySearch
- Splice
- Codon Usage
- Translation
- ORF Analysis
- Restriction Map
- Genome Mapping

Keyword search (gen) Search RefSeq

Multi Fasta Format

```
>AA987701(zenbank-upd)
taagcctttattt
ttggtgatagcaagagac
tgcttcaaccttagctggg
agggcagccagcccaaaagc
tggaaagtgtaactcctc
agaatgggtactgatcaac
ggacacccctccagaaagc
gttgatgctgacctgtc
agaatgtagcaatgtagcagc

>AA991595(zenbank-upd)
ccttggcaactattctttatgaacatcagtaacagaagatgctgatttagaggaa
gcttccctccctccctccccattgcacccggaaactcagcaagtgagccagccaca
gagcataaactgtgctggacacatgcatgtgtttcacatagctggcatctgt
tcacaagagtcagtgccaccagttcatcagctgtttccocctccctccctcctg
cagccccagcttggagaaaagcctcctctgagcagggcccccagccattcac
ccaacattgtagatcgccctcaagctggggagtcattggagttcaggccctgactcc
ctccggaagcctgcttctaaaagaggaatttctccagccctgactgtaagtgaatt
catgccttggggcctaatctccccattggcaaatccccagttagctcattcgtgata
attgagccggataaccgttgggacctcacattgcgcctcttaggggtccaagaa

>A0011736(zenbank-upd)
agcctgccagcgtgaatttccctctgtttcactattccctgtaaaagcatttttagtcca
ccgtaagaactctgtttgttagtaatttctctgagaaactggattttttctgccc
tactgattgcaaaaggaatgacttgcactctcataaaaagatacaggttgatcattagaatc
cctgcagactctccocctttctgtgcccogacocctagtcgctatacagggcactagtc
aatttctctgctcatttttccctgaaatgagataaatacagaatactcgtgtagatg
attgcaagataagatacattttctagaggtttctattacattacacattacacagagaca
atgngcaacttacaagaatgaagacatagagaggtttacaagaattttggaacact
```

Length Undo Clear

**Nucleic Acid Contents**

Name	All	A	T(U)	G	C	A-T	G-C
AA987...	503	137	113	135	118	250	253
	[%]	27.2	22.5	26.8	23.5	49.7	50.3
AA991...	538	124	127	135	152	251	287

⑦ AA987701 Window Size 20 Shift 1

Submit Redraw Cancel Graph Output

Pos	A (%)	T (%)	G (%)	C (%)	A-T (%)	G-C (%)
1-20	8 (40.0)	7 (35.0)	3 (15.0)	2 (10.0)	15 (75.0)	5 (25.0)
2-21	8 (40.0)	7 (35.0)	3 (15.0)	2 (10.0)	15 (75.0)	5 (25.0)
3-22	7 (35.0)	7 (35.0)	3 (15.0)	3 (15.0)	14 (70.0)	6 (30.0)
4-23	6 (30.0)	7 (35.0)	3 (15.0)	4 (20.0)	13 (65.0)	7 (35.0)
5-24	5 (25.0)	8 (40.0)	3 (15.0)	4 (20.0)	13 (65.0)	7 (35.0)
6-25	5 (25.0)	9 (45.0)	2 (10.0)	4 (20.0)	14 (70.0)	6 (30.0)
7-26	4 (20.0)	9 (45.0)	3 (15.0)	4 (20.0)	13 (65.0)	7 (35.0)
8-27	3 (15.0)	10 (50.0)	3 (15.0)	4 (20.0)	13 (65.0)	7 (35.0)
9-28	3 (15.0)	11 (55.0)	2 (10.0)	4 (20.0)	14 (70.0)	6 (30.0)
10-29	3 (15.0)	11 (55.0)	2 (10.0)	4 (20.0)	14 (70.0)	6 (30.0)
11-30	2 (10.0)	12 (60.0)	2 (10.0)	4 (20.0)	14 (70.0)	6 (30.0)
12-31	1 (5.0)	12 (60.0)	3 (15.0)	4 (20.0)	13 (65.0)	7 (35.0)
13-32	1 (5.0)	12 (60.0)	2 (10.0)	5 (25.0)	13 (65.0)	7 (35.0)
14-33	2 (10.0)	12 (60.0)	2 (10.0)	4 (20.0)	14 (70.0)	6 (30.0)
15-34	3 (15.0)	12 (60.0)	2 (10.0)	3 (15.0)	15 (75.0)	5 (25.0)
16-35	4 (20.0)	11 (55.0)	2 (10.0)	3 (15.0)	15 (75.0)	5 (25.0)
17-36	4 (20.0)	11 (55.0)	2 (10.0)	3 (15.0)	15 (75.0)	5 (25.0)
18-37	5 (25.0)	10 (50.0)	2 (10.0)	3 (15.0)	15 (75.0)	5 (25.0)
19-38	5 (25.0)	10 (50.0)	2 (10.0)	3 (15.0)	15 (75.0)	5 (25.0)
20-39	6 (30.0)	9 (45.0)	2 (10.0)	3 (15.0)	15 (75.0)	5 (25.0)
21-40	7 (35.0)	8 (40.0)	2 (10.0)	3 (15.0)	15 (75.0)	5 (25.0)
22-41	7 (35.0)	7 (35.0)	2 (10.0)	4 (20.0)	14 (70.0)	6 (30.0)
23-42	7 (35.0)	8 (40.0)	2 (10.0)	3 (15.0)	15 (75.0)	5 (25.0)
24-43	7 (35.0)	8 (40.0)	2 (10.0)	3 (15.0)	15 (75.0)	5 (25.0)
25-44	7 (35.0)	7 (35.0)	2 (10.0)	4 (20.0)	14 (70.0)	6 (30.0)

図 5-6-1 含量計算画面

計算結果テーブル

- ④ 「Submit」ボタンをクリックすると、実行領域上部に配列全体の含量表が表示されます。そして結果表示領域にデフォルトのテーブル、別ウィンドウにデフォルトの含量グラフが表示されます。(図 5-6-2)

### ！注意

- デフォルトでは、先頭配列の GC グラフと計算結果テーブルが表示されます。
- 各塩基の含量率の値は、小数点第 2 位で四捨五入しています。
- 含量グラフウィンドウの右のチェックボックス、color ボタンをクリックすると、描画する塩基と色を変更することができます。
- 1度に1つの配列の含量グラフのみ表示されます。

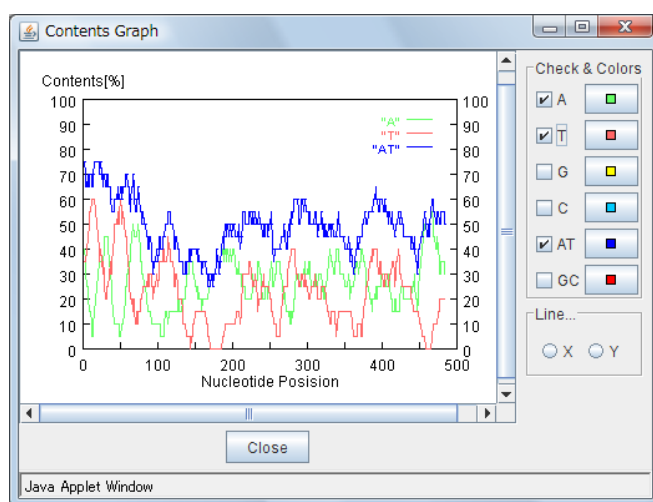


図 5-6-2 含量グラフ

- ⑤ 含量グラフを閉じた後に再表示したい場合、「Graph」ボタンをクリックします。(図 5-6-1)
- ⑥ 「Output」ボタンをクリックすると、配列全体のコンテンツ表が別画面に表示されます。(図 5-6-1) ブラウザ画面にて「名前を付けて保存」を選択するなどにより、出力結果を保存することが可能です。

### <再計算し、グラフを再描画する場合>

- ⑦ グラフに描画する核酸配列名を指定します。(図 5-6-3)
- ⑧ ウィンドウサイズ (Window Size)、シフト値 (Shift) を変更したい場合は設定します。(図 5-6-1)

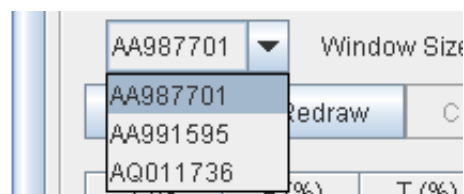


図 5-6-3

- ⑨ 「Redraw」ボタンをクリックすると、再計算を行い、テーブルの更新、グラフが再描画されます。(図 5-6-1)

## 5.7 相同性検索機能(インタラクティブ処理)

指定された核酸配列データベースの中から相同性の高い配列を検索し、結果を表示します。相同性検索には、検索が終了するまで待って結果を表示するインタラクティブ処理と、検索要求のみを行って結果を後で見るバッチ処理があります。

時間を要する検索結果を行う場合は、バッチ処理を行います。(5.8 参照)

<相同性検索の方法>

- ① 「Mode」メニューで「Nucleic」を選択し、「Algorithm」メニュー、またはツールアイコンで「Homology Search」を選択します。(図 5-7-1)
- ② データ編集領域に核酸配列データを single fasta 形式で入力します。複数の配列を入力することはできません。

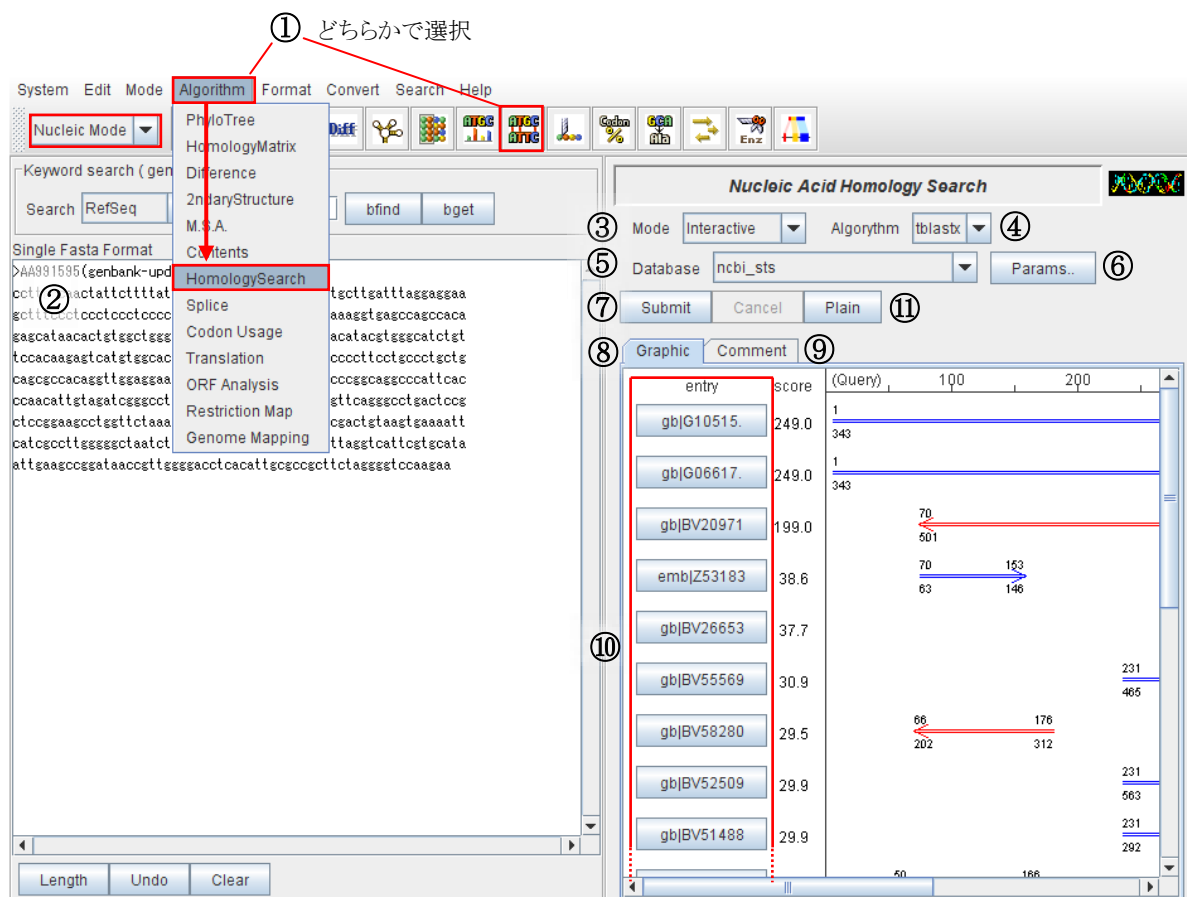


図 5-7-1 相同性検索画面 (Interactive)

③ 「Interactive」を選択します。(図 5-7-2)

BatchSearch、BatchView については 5.8 にて説明します。

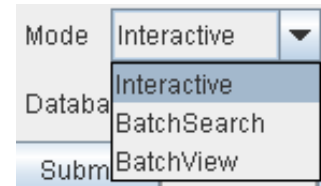


図 5-7-2 Mode 切替画面

④ アルゴリズムを選択します。(図 5-7-3)

選択可能なアルゴリズムは、blastn、blastx、tblastx、fasta3、fastx3 の 5 種類です。(表 5-7 参照)

データベースがアルゴリズムに対応するものに切り替わります。

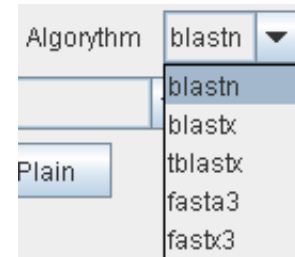


図 5-7-3 アルゴリズム選択画面

表 5-7 アルゴリズムの説明

アルゴリズム	query 配列	データベース	検索結果	備考
blastn	核酸	核酸	核酸	
blastx	核酸	タンパク質	タンパク質	query 配列を翻訳して検索
tblastx	核酸	核酸	タンパク質	両方とも翻訳しながら検索
fasta3	核酸	核酸	核酸	
fastx3	核酸	タンパク質	タンパク質	query 配列を翻訳して検索

⑤ データベースを選択します。(図 5-7-4)

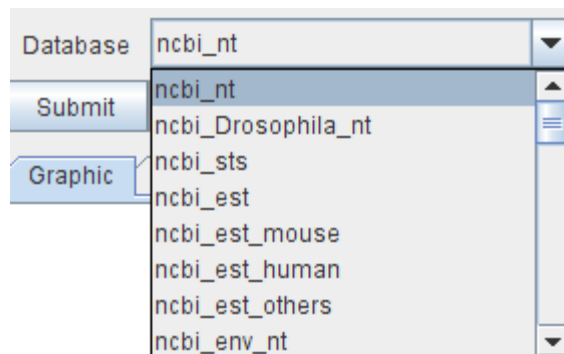


図 5-7-4 データベース選択画面

⑥ 「Params…」ボタンをクリックすると、各アルゴリズムに対応したパラメータ設定パネルが表示され、パラメータの設定を行うことができます。

⑦ 「Submit」ボタンをクリックすると、相同性検索を行い、結果表示領域にリストが表示されます。



### <グラフィック表示の方法>

- ⑧ 「Graphic」タブをクリックすると、グラフィック表示に切り替わります。(図 5-7-5)

矢印の向きは、検索結果の query 配列の向きです。

矢印の上の数值は query 配列の start、stop 位置で、矢印の下の数值はデータベース配列の start、stop 位置です。

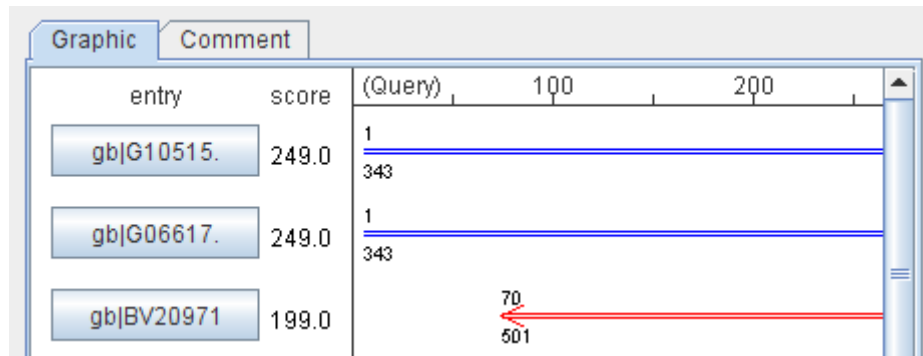


図 5-7-5 Graphic タブ選択時表示画面

### <コメント表示の方法>

- ⑨ 「Comment」タブをクリックすると、コメント表示に切り替わります。(図 5-7-6)

entry	score	Definition
gb G10515.	249.0	human STS CHLC.UTR_00851_X14618.P371
gb G06617.	249.0	human STS WI-7608, sequence tagged site
gb BV20971	199.0	ACP5_4244 Rhesus macaque genomic DNA

図 5-7-6 Comment タブ選択時表示画面

## < 詳細情報の表示 >

- ⑩ entry 名をクリックすると、相同性検索結果の詳細情報が別ウィンドウに表示されます。  
(図 5-7-7)

同時にクリックされた entry 名をキーワードとして、遺伝情報実験センターのデータベースへ検索を行います。見つかった場合、別画面に entry 名に関する情報が表示され  
(図 5-7-8)、見つからなかった場合、“No hits”と表示されたダイアログが出力されます。

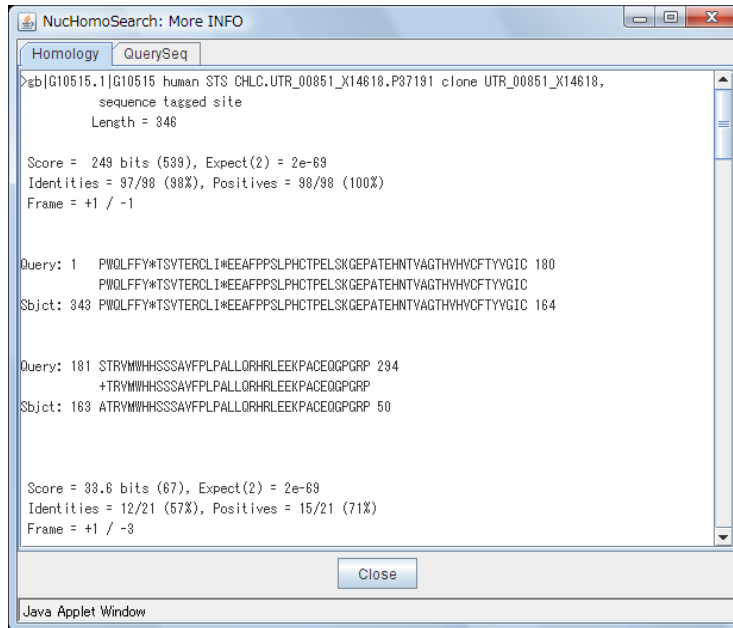


図 5-7-7 相同性検索詳細情報

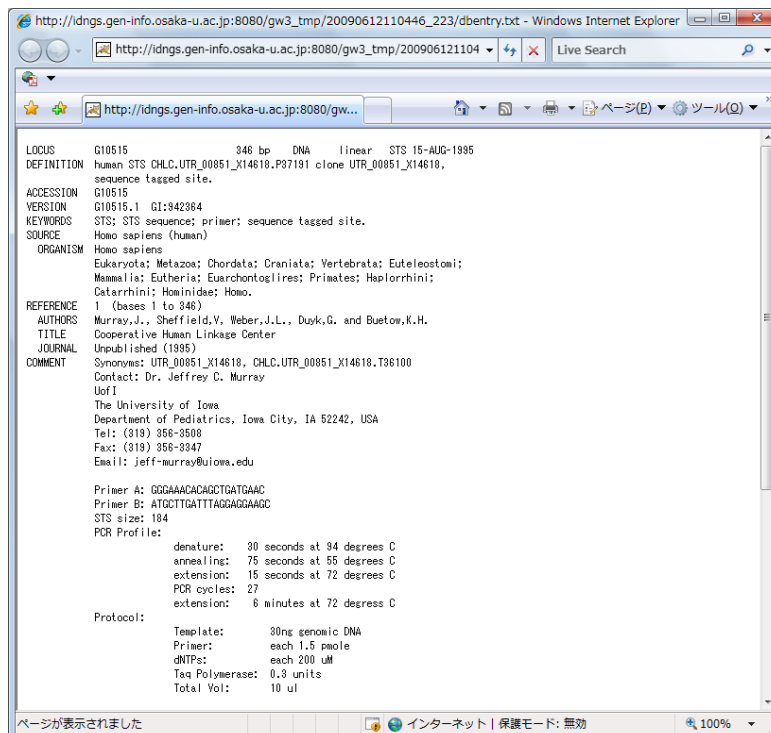


図 5-7-8 entry 名検索結果

## < 検索結果の表示 >

- ① 「Plain」ボタンをクリックすると、別ウィンドウに相同性検索結果が表示されます。「Plain」タブをクリックすると Plain 形式 (図 5-7-9)、「table」タブをクリックすると table 形式 (図 5-7-10) で検索結果が表示されます。「Output」ボタンをクリックすると、検索結果が別ブラウザ画面に表示され、ブラウザ画面にて「名前を付けて保存」を選択するなどにより、出力結果を保存することが可能です。

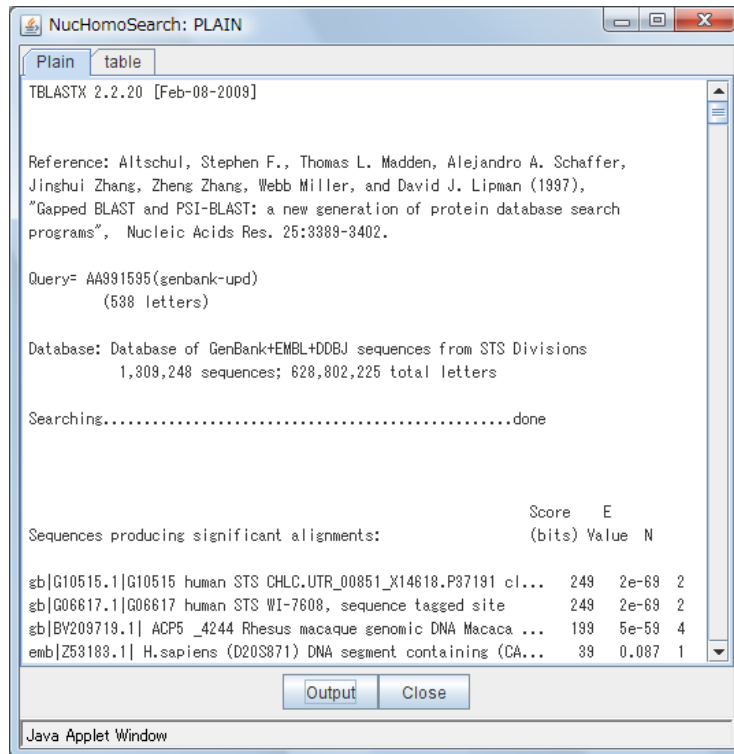


図 5-7-9 Plain 形式

q_na...	db_n...	(%).id...	align...	mism...	gap	q_start	q_stop	db_st...	db_st...	e-value	score
AA99...	gb G...	98	98	1	0	1	294	343	50	2e-69	249.0
AA99...	gb G...	98	98	1	0	1	294	343	50	2e-69	249.0
AA99...	gb B...	83	98	16	0	363	70	208	501	5e-59	199.0
AA99...	emb ...	50	28	14	0	70	153	63	146	0.087	38.6
AA99...	gb B...	45	35	19	0	280	384	140	244	0.16	37.7
AA99...	gb B...	35	40	26	0	231	350	465	346	0.16	30.9
AA99...	gb B...	35	37	24	0	176	66	312	202	0.26	29.5
AA99...	gb B...	34	35	23	0	231	335	563	667	0.30	29.9
AA99...	gb B...	34	35	23	0	231	335	292	188	0.30	29.9
AA99...	gb B...	35	39	25	0	50	166	220	104	0.31	36.8
AA99...	gb G...	40	44	26	0	534	403	44	175	0.42	36.3
AA99...	gb B...	35	40	26	0	65	184	550	431	0.80	35.4
AA99...	gb B...	44	25	14	0	92	166	75	1	0.80	35.4
AA99...	gb B...	45	31	17	0	55	147	439	347	0.80	35.4
AA99...	gb B...	35	39	25	0	85	201	328	212	0.80	35.4
AA99...	gb B...	42	35	20	0	144	248	257	361	1.1	35.0
AA99...	gb B...	47	19	10	0	148	92	582	638	1.3	29.9
AA99...	gb B...	30	53	37	0	296	138	2645	2487	1.5	34.5
AA99...	gb G...	41	29	17	0	103	17	326	240	1.5	34.5
AA99...	gb B...	50	24	12	0	527	456	3	74	1.8	27.2

図 5-7-10 table 形式

<blastn 用のパラメータの指定方法>

- ⑫ 期待値 (Expectation Value) を実数で入力します。
- ⑬ アライメントを表示する数 (Alignment) を整数で入力します。(上限は 100)

**！注意** 数値以外が入力された場合、デフォルト値に戻ります。

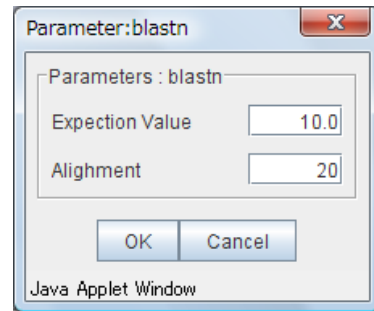


図 5-7-11 blastn パラメータ

<blastx 用のパラメータの指定方法>

Expectation Value、Alignment は

⑫～⑬と同様。

- ⑭ マトリックスを選択します。

選択肢：  
PAM30  
PAM70  
BLOSUM80  
BLOSUM62  
BLOSUM45

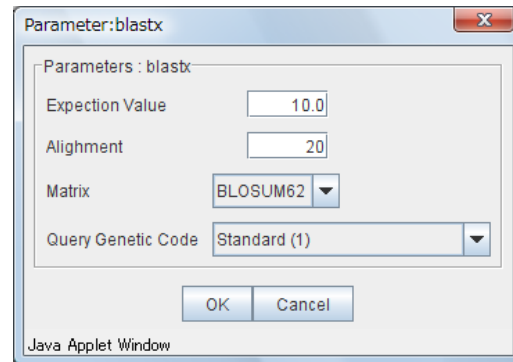


図 5-7-12 blastx パラメータ

- ⑮ query 配列の Genetic Code を選択します。

選択肢：  
Standard  
Vertebrate Mitochondrial  
Yeast Mitochondria  
Mold Mitochondrial  
Invertebrate Mitochondrial  
Ciliate Nuclear  
Echinoderm Mitochondrial  
Euplotid Nuclear  
Bacterial  
Alternative Yeast Nuclear  
Ascidian Mitochondrial  
Flatworm Mitochondrial  
Blepharisma Macronuclear

<tblastx 用のパラメータの指定方法>

Expectation Value、Alignment、Matrix、Query Genetic Code は、⑫～⑮と同様。

- ⑯ DataBase の Genetic Code を選択します。  
(選択肢は、⑮同様)

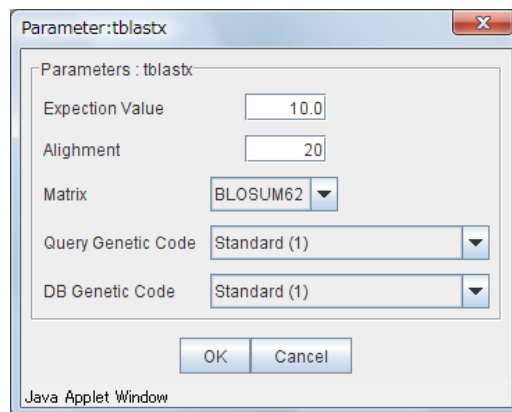


図 5-7-13 tblastx パラメータ

<fasta3, fastx3 用のパラメータの指定方法>

- ⑰ KTUP を選択します。  
選択肢: fasta3 の場合、1, 2, 3, 4, 5, 6  
fastx3 の場合、1, 2

Expectation Value、Alignment は、⑫～⑬と同様。

- ⑱ Reverse Complement (相補鎖) を行うかを選択します。

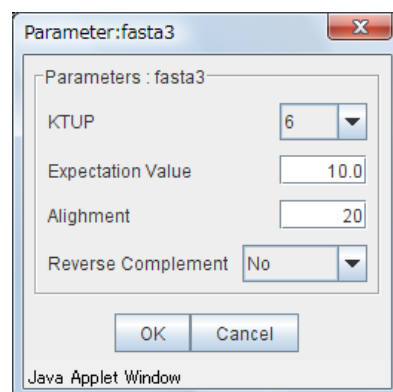


図 5-7-14 fasta3 パラメータ

## 5. 8 相同性検索機能(バッチ処理)

5. 7の核酸配列データ相同性検索で処理時間が要する場合は、バッチ処理により検索と結果の表示を分けて行うことができます。

- ① 「Mode」メニューで「Nucleic」を選択し、「Algorithm」メニュー、またはツールアイコンで「Homology Search」を選択します。(図 5-8-1)

<相同性検索の方法>

- ② データ編集領域に核酸配列データを single fasta 形式で入力します。複数の配列を入力することはできません。
- ③ 「Mode」メニューで「BatchSearch」を選択します。
- ④ Algorithm、Database、Params..の設定をします。(インタラクティブ処理 p27-32 参照)

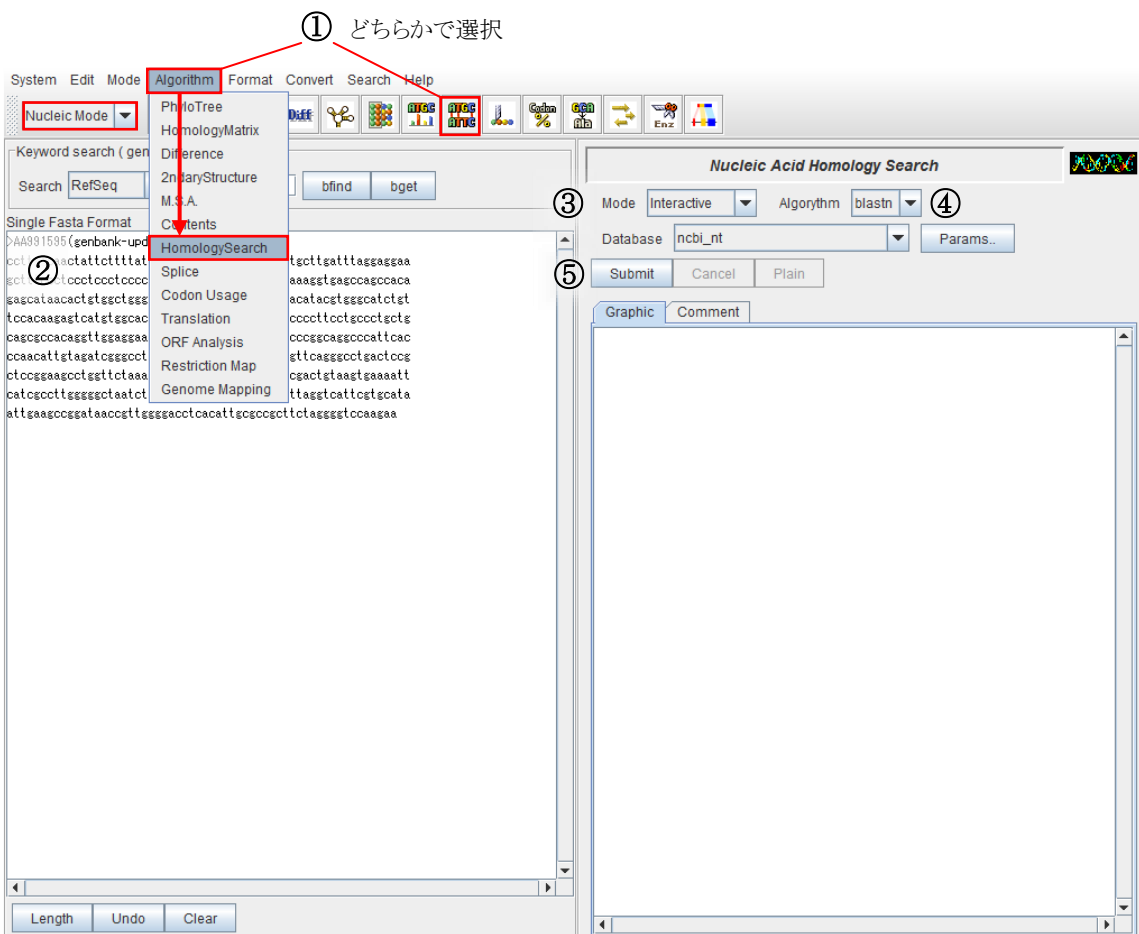


図 5-8-1 相同性検索画面 (BatchSearch)

- ⑤ 「Submit」ボタンをクリックすると、相同性検索を行い、結果を問い合わせるための Request ID のダイアログが表示されます。(図 5-8-2、次ページの注意参照)

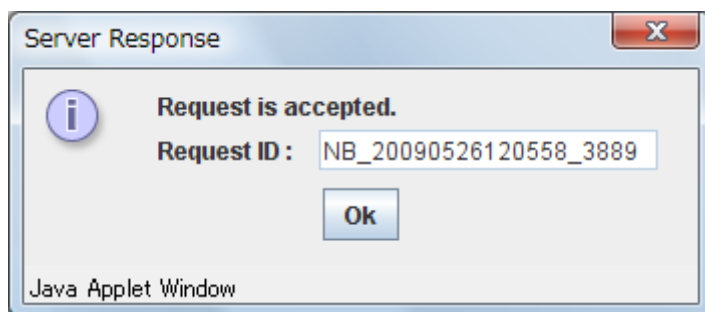


図 5-8-2 RequestID ダイアログ

### <結果表示の方法>

- ⑥ 「Mode」メニューで「BatchView」を選択します。(図 5-8-3)

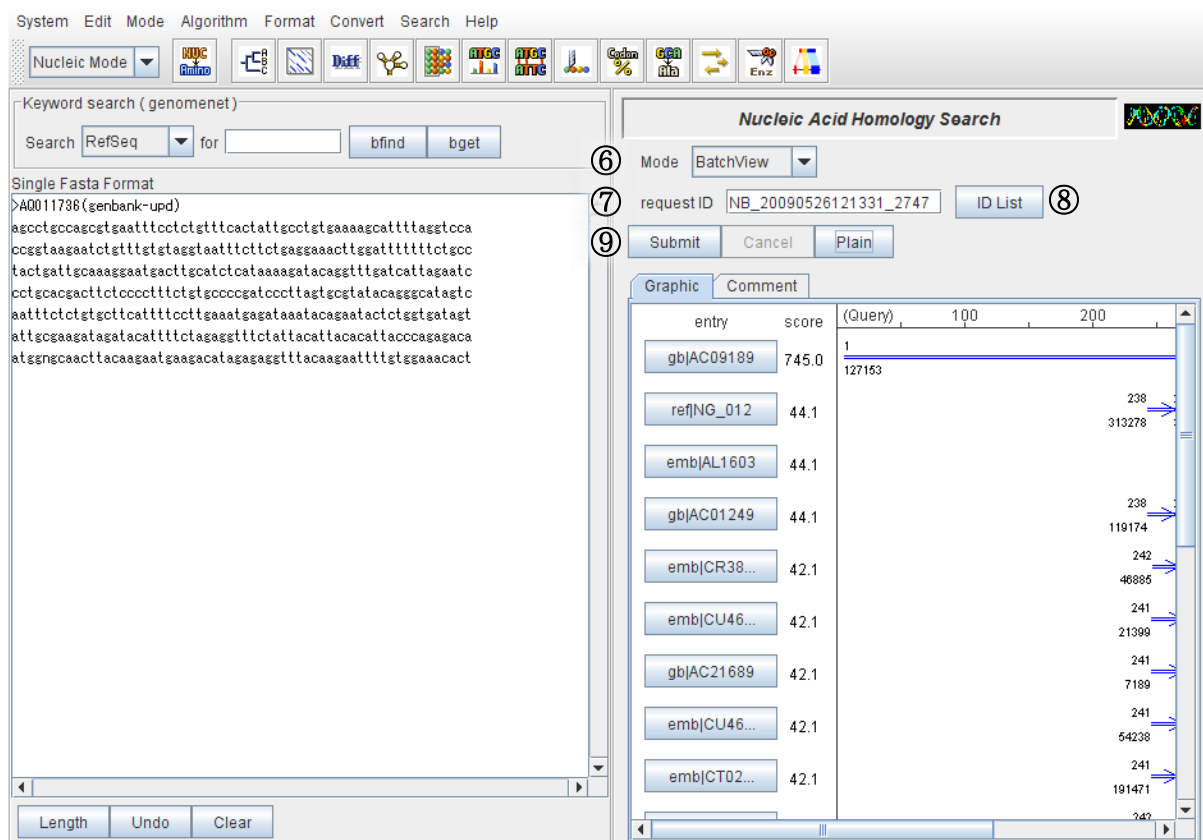
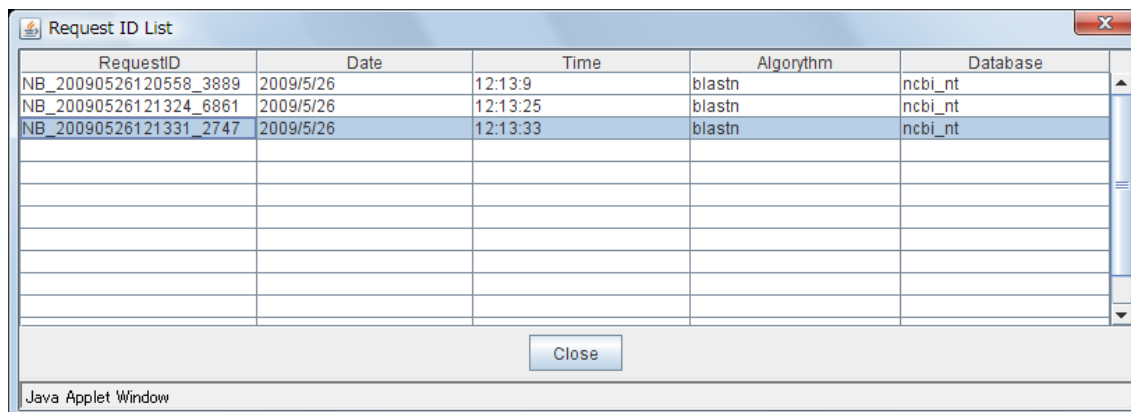


図 5-8-3 相同性検索画面 (BatchView)

- ⑦ 「BatchSearch」処理で受け取った Request ID を入力します。
- ⑧ 「ID List」ボタンをクリックすると、GeneWebⅢ起動後の BatchSearch 処理リクエストに対する Request ID の一覧が表示され、詳細の確認ができます。(図 5-8-4)



RequestID	Date	Time	Algorithhm	Database
NB_20090526120558_3889	2009/5/26	12:13:9	blastn	ncbi_nt
NB_20090526121324_6861	2009/5/26	12:13:25	blastn	ncbi_nt
NB_20090526121331_2747	2009/5/26	12:13:33	blastn	ncbi_nt

Close

Java Applet Window

図 5-8-4 リクエスト ID 一覧  
(選択中の Request ID は色付背景で表示される)

このリクエスト ID 一覧画面において、Request ID をクリックすると、背景色が変わり、⑦の request ID 入力ボックスに自動入力されます。

### ！注意

GeneWebⅢ起動以前のリクエストに対する Request ID は保存されていません。  
GeneWebⅢを一度終了させ、後で結果を表示させたい場合は、Request ID をメモしておいてください。

この ID を入力することにより、以前実行させた結果を表示することができます。  
ただし、結果の保存期間は 1 ヶ月です。

- ⑨ 「Submit」をクリックすると、結果表示領域に結果が表示されます。



## 5.9 Splice 機能

核酸配列の指定した範囲を切り出します。

### <Splice の方法>

- ① 「Mode」メニューで「Nucleic」を選択し、「Algorithm」メニュー、またはツールアイコンで「Splice」を選択します。(図 5-9)
- ② データ編集領域に核酸配列データを fasta 形式で入力します。複数の配列を入力することも可能です。
- ③ 切り出す範囲を指定します。範囲の指定する方法には、2通りがあります。(次頁参照)
- ④ 「Submit」ボタンをクリックすると、範囲指定して切り出された核酸配列データが結果表示領域に表示されます。

**！注意** この処理は、入力がアミノ酸の場合にも利用できます。

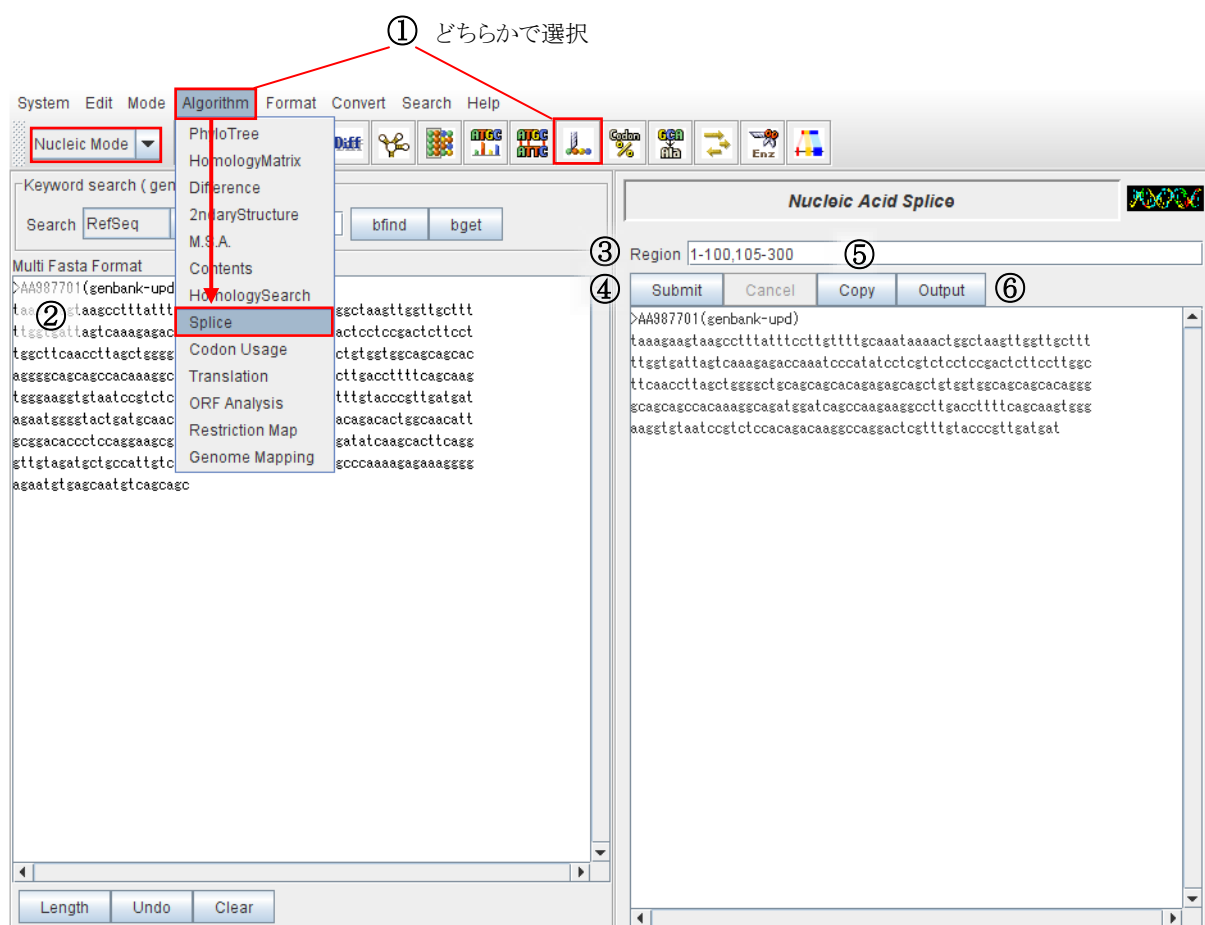


図 5-9 Splice 機能画面

### <切り出す範囲の指定方法>

#### 指定方法 1

Region に開始位置数と終了位置数の間に「-」を入れ入力します。範囲を複数指定する場合は、「,」(カンマ)で区切ります。 例) 1-24, 40-102

#### 指定方法 2

データ編集領域に入力した核酸配列データ中の切り出したい範囲を括弧[ ]でくくりま  
す。括弧は、2組以上指定することも可能です。

例) actgatgcatcgatc[gacactaggcta]gcgagctacgagc

**！注意** 両方の指定がされた場合、指定方法2が優先されます。

### <共通領域へのコピーの方法>

- ⑤ 結果表示領域に表示されている核酸配列データを、データ編集領域に表示させるため  
には、「Copy」ボタンをクリックします。
  
- ⑥ 「Output」ボタンをクリックすると、結果が別ブラウザ画面に表示されます。ブラウザ画面  
にて「名前を付けて保存」を選択するなどにより、出力結果を保存することができます。

## 5. 10 Codon 含有率計算機能

核酸配列に含まれる Codon の含有率の計算がされ、表示されます。

### < Codon 含有率の計算方法 >

- ① 「Mode」メニューで「Nucleic」を選択し、「Algorithm」メニュー、またはツールアイコンで「Codon Usage」を選択します。(図 5-10-1)
- ② データ編集領域に、核酸配列データを fasta 形式で入力します。複数の配列を入力することも可能です。
- ③ 計算する範囲を指定します。Region に開始位置と終了位置の間に「-」を入れ、入力します。範囲を複数指定する場合は、「,」で区切ります。何も指定していなければ、すべての範囲が対象として処理されます。

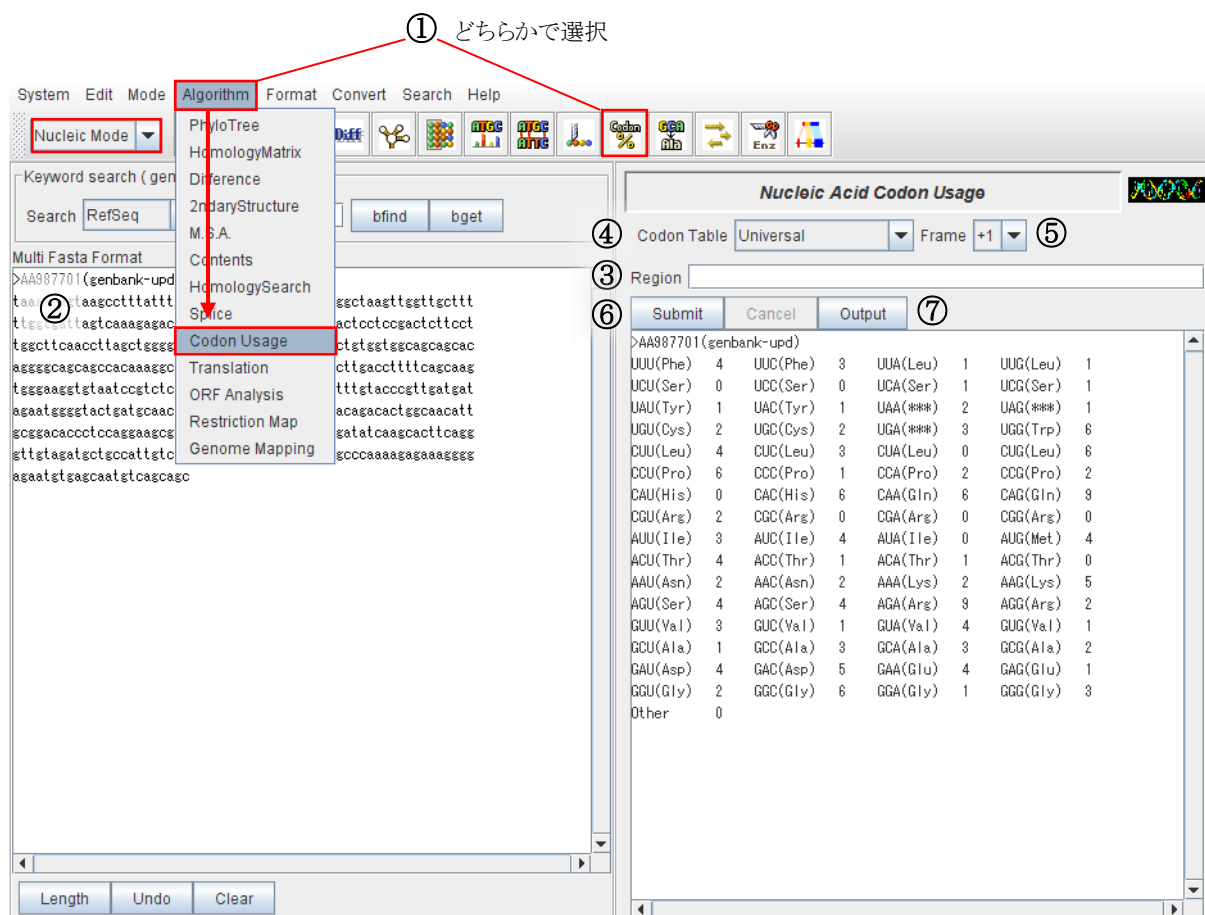


図 5-10-1 Codon Usage 画面

- ④ 生物種別コドン表を選択します。(図 5-10-2)

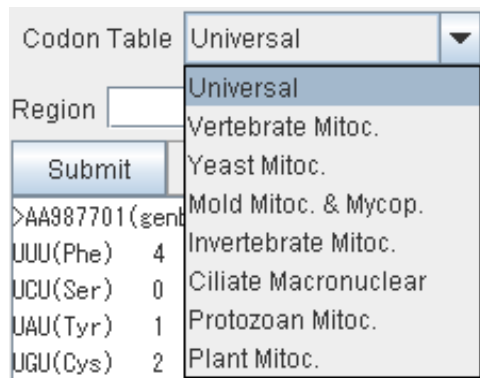


図 5-10-2 生物種別コドン表

- ⑤ 計算を開始するフレーム位置を選択します。(図 5-10-3)

- + $n$ : 先頭の  $n$  文字目から順方向に計算します。
- $n$ : 末尾の  $n$  文字目から逆方向に相補的な塩基について (Reverse Complement) 計算します。

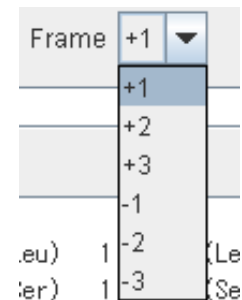


図 5-10-3 フレーム位置選択

- ⑥ 「Submit」ボタンをクリックすると含有率が結果表示領域に表示されます。範囲の指定方法が間違っていたり、指定された範囲が配列の長さより大きかったり 0 以下の場合、エラーメッセージダイアログが出力されます。
- ⑦ 「Output」ボタンをクリックすると、結果が別ブラウザ画面に表示されます。ブラウザ画面にて「名前を付けて保存」を選択するなどにより、出力結果を保存することができます。

## 5. 11 ヌクレオチド翻訳機能

核酸配列をアミノ酸に翻訳します。

<ヌクレオチド翻訳方法>

- ① 「Mode」メニューで「Nucleic」を選択し、「Algorithm」メニュー、またはツールアイコンで「Translation」を選択します。(図 5-11-1)
- ② データ編集領域に核酸配列データを fasta 形式で入力します。複数の配列を入力することも可能です。
- ③ 翻訳する範囲を指定します。Region に開始位置数と終了位置数の間に「-」を入れ、入力します。範囲を複数指定する場合は、「,」で区切ります。何も指定していなければ、すべての範囲を対象として処理されます。

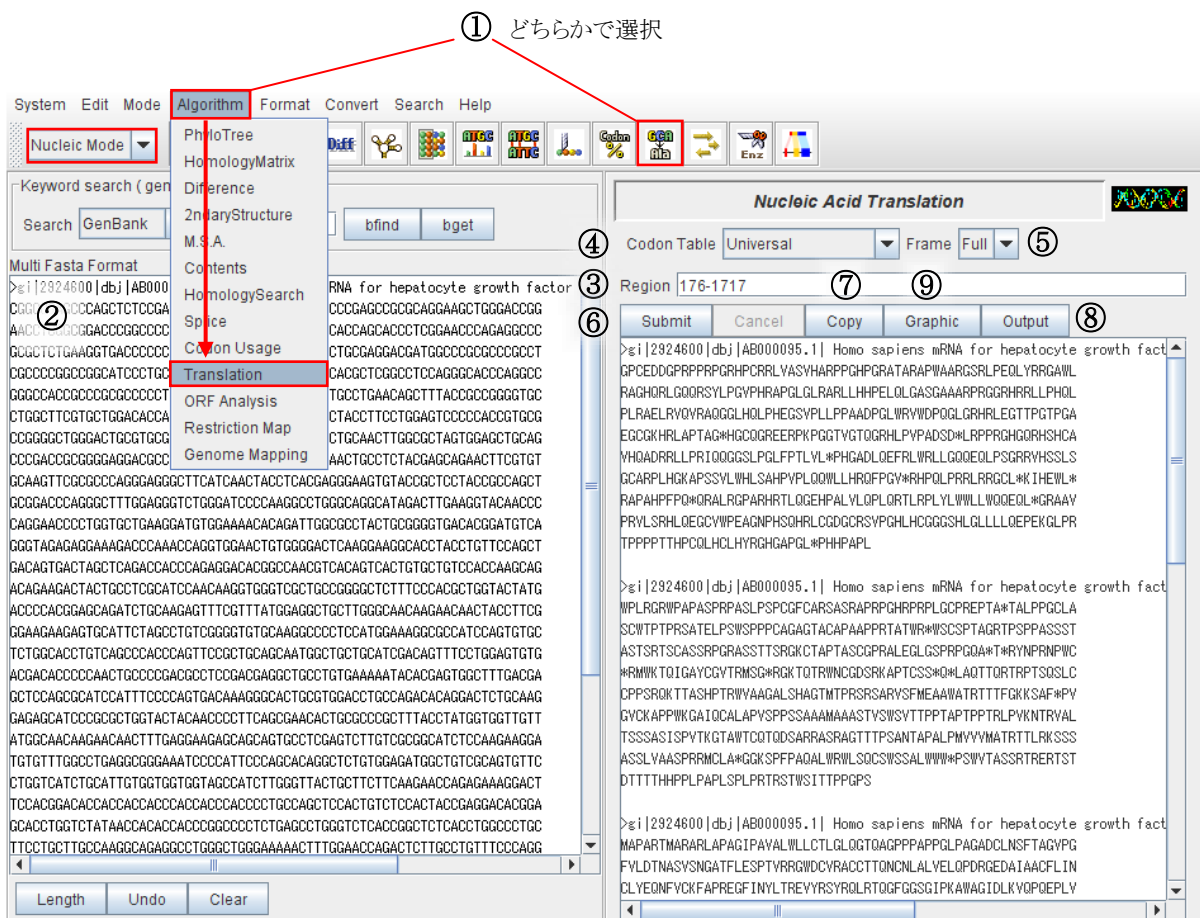


図 5-11-1 Translation 機能画面

- ④ 生物種別コドン表を選択します。(図 5-11-2)

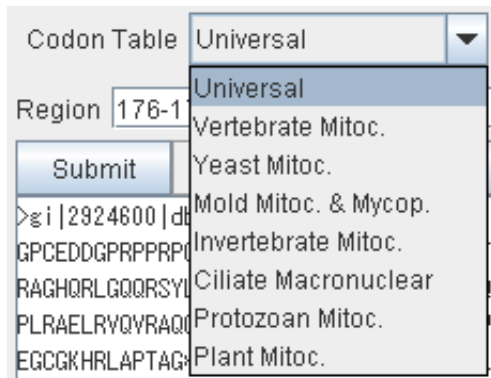


図 5-11-2 生物種別コドン表選択

- ⑤ 計算を開始するフレーム位置を選択します。(図 5-11-3)

「Full」を指定すると、全6通りの変換が行われます。

+*n*: 先頭の *n* 文字目から順方向に翻訳します。

-*n*: 末尾の *n* 文字目から逆方向に相補的な塩基について  
(Reverse Complement) 翻訳します。

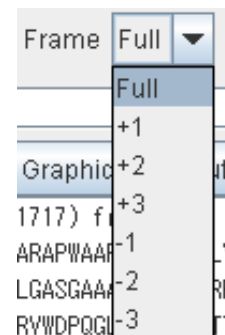


図 5-11-3 フレーム位置選択

- ⑥ 「Submit」ボタンをクリックすると、核酸からアミノ酸に変換された結果が結果表示領域に表示されます。範囲の指定方法の間違い、指定された範囲が配列の長さより大きい、また 0 以下の場合、エラーメッセージダイアログが出力されます。
- ⑦ 結果表示領域に表示されている核酸配列をデータ編集領域に表示させるためには、「Copy」ボタンをクリックします。
- ⑧ 「Output」ボタンをクリックすると、結果が別ブラウザ画面に表示されます。ブラウザ画面にて「名前を付けて保存」を選択するなどにより、出力結果を保存することができます。

- ⑨ 「Graphic」ボタンをクリックすると、別ウィンドウに塩基配列データとアミノ酸翻訳データがグラフィカルに表示されます。(図 5-11-4)

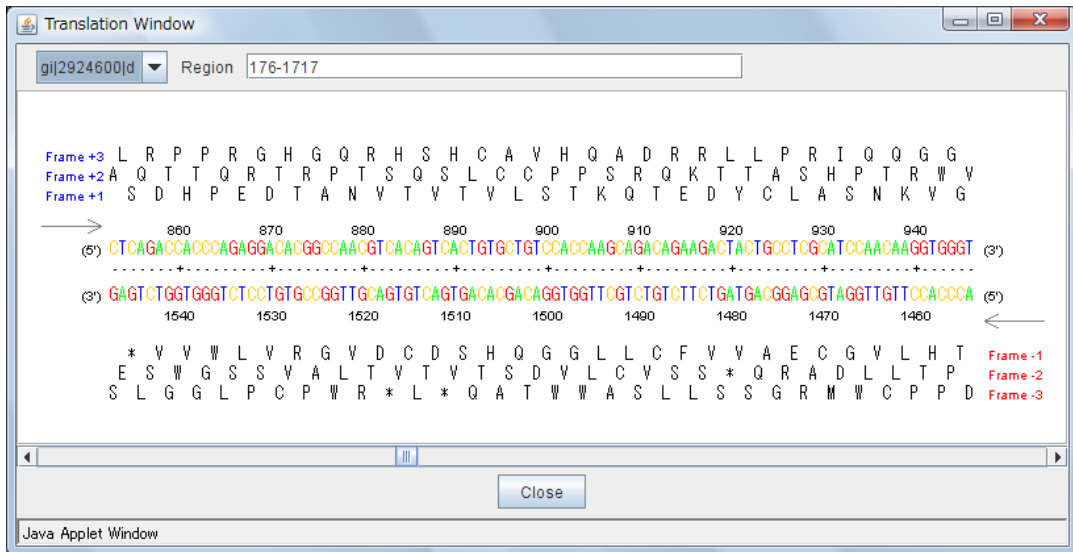


図 5-11-4 Translation グラフィカル表示

## 5.12 Open Reading Frame 検索機能

核酸配列の中に含まれている Open Reading Frame (ORF) を検索し、その表示を行います。

### <Open Reading Frame 検索方法>

- ① 「Mode」メニューで「Nucleic」を選択し、「Algorithm」メニュー、またはツールアイコンで「ORF Analysis」を選択します。(図 5-12-1)
- ② データ編集領域に核酸配列データを single fasta 形式で入力します。複数の配列を入力することはできません。
- ③ Cut Off 値を入力します。

#### ！注意

- Cut Off 値は、0 以上 100 以下の整数とし、その範囲以外が入力されるとデフォルト値 (50) に戻ります。
- アミノ酸配列に翻訳したとき、Cut Off 値で指定された値以上の長さとなる ORF が抽出されます。

① どちらかで選択

②

③

④

⑤

⑥

⑦

図 5-12-1 ORF 検索画面



- ④ 生物種別コドン表を選択します。

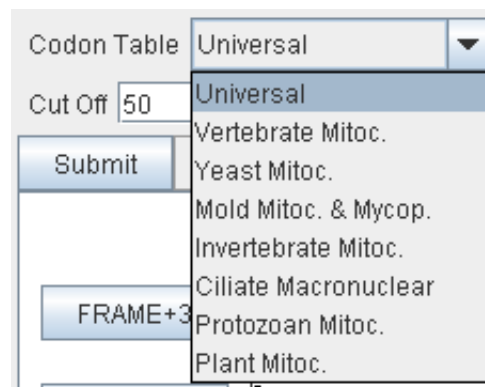


図 5-12-2 生物種別コドン表選択画面

- ⑤ 「Submit」ボタンをクリックすると、Open Reading Frame の検索結果が結果表示領域に表示されます。
- ⑥ 「Output」ボタンをクリックすると、結果が別ブラウザ画面に表示されます。ブラウザ画面にて「名前を付けて保存」を選択するなどにより、出力結果を保存することができます。

#### ＜フレームのアミノ酸翻訳データの表示方法＞

- ⑦ 結果表示領域の Open Reading Frame 解析結果の Frame 番号をクリックすると、そのフレームでのアミノ酸翻訳データが別ウィンドウに表示されます。(図 5-12-3)

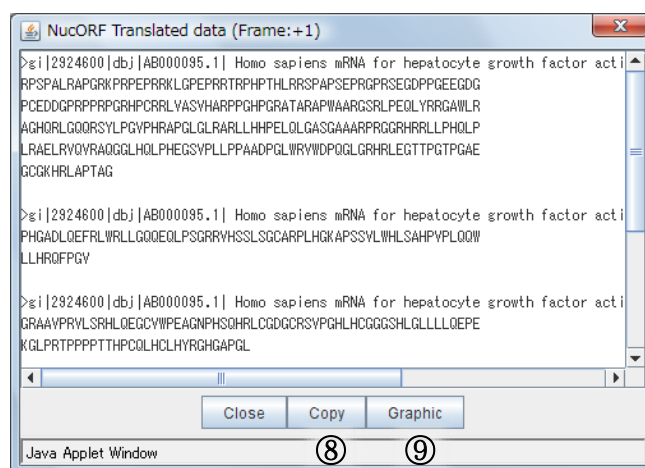


図 5-12-3 アミノ酸翻訳データ

- ⑧ 「Copy」ボタンをクリックすると、アミノ酸翻訳データがデータ編集領域にコピーされます。

- ⑨ 「Graphic」ボタンをクリックすると、塩基配列データとアミノ酸翻訳データがグラフィカルに表示されます。(図 5-12-4)

左側の ORF 領域をクリックすると、その領域が表示されます。

### ORF 領域のリスト

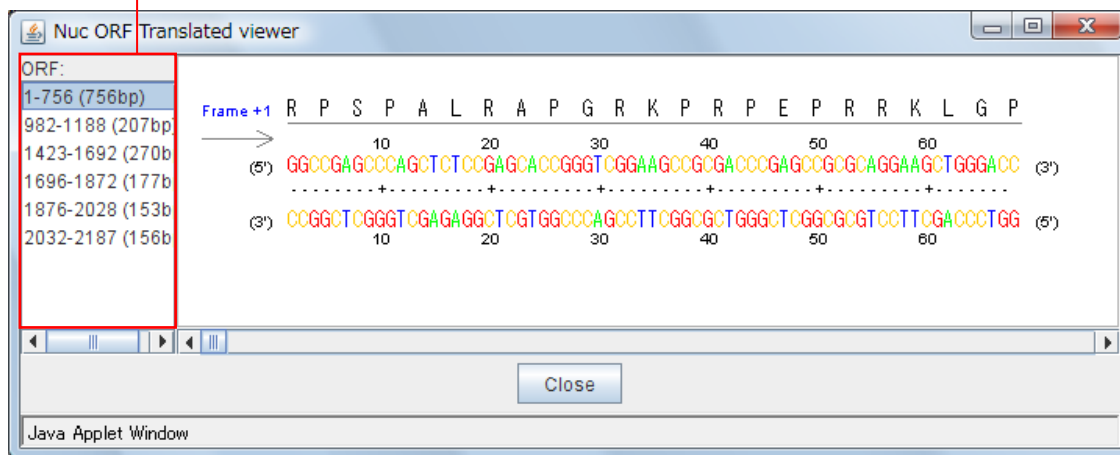


図 5-12-4 ORF グラフィカル表示

## 5.13 制限酵素地図作成機能

指定した制限酵素により切断される核酸配列データ中の位置情報が、地図とテキスト形式によって表示されます。

サーバ内での処理はtacg ver. 2.35 を使用しています。(http://tacg.sourceforge.net/)

### <制限酵素地図作成方法>

- ① 「Mode」メニューで「Nucleic」を選択し、「Algorithm」メニュー、またはツールアイコンで「Restriction Map」を選択します。(図 5-13-1)
- ② データ編集領域に核酸配列データを single fasta 形式で入力します。複数の配列を入力することはできません。
- ③ 利用したい制限酵素を酵素セットリスト内に作成します。酵素を追加・削除する方法は次々項を参照してください。

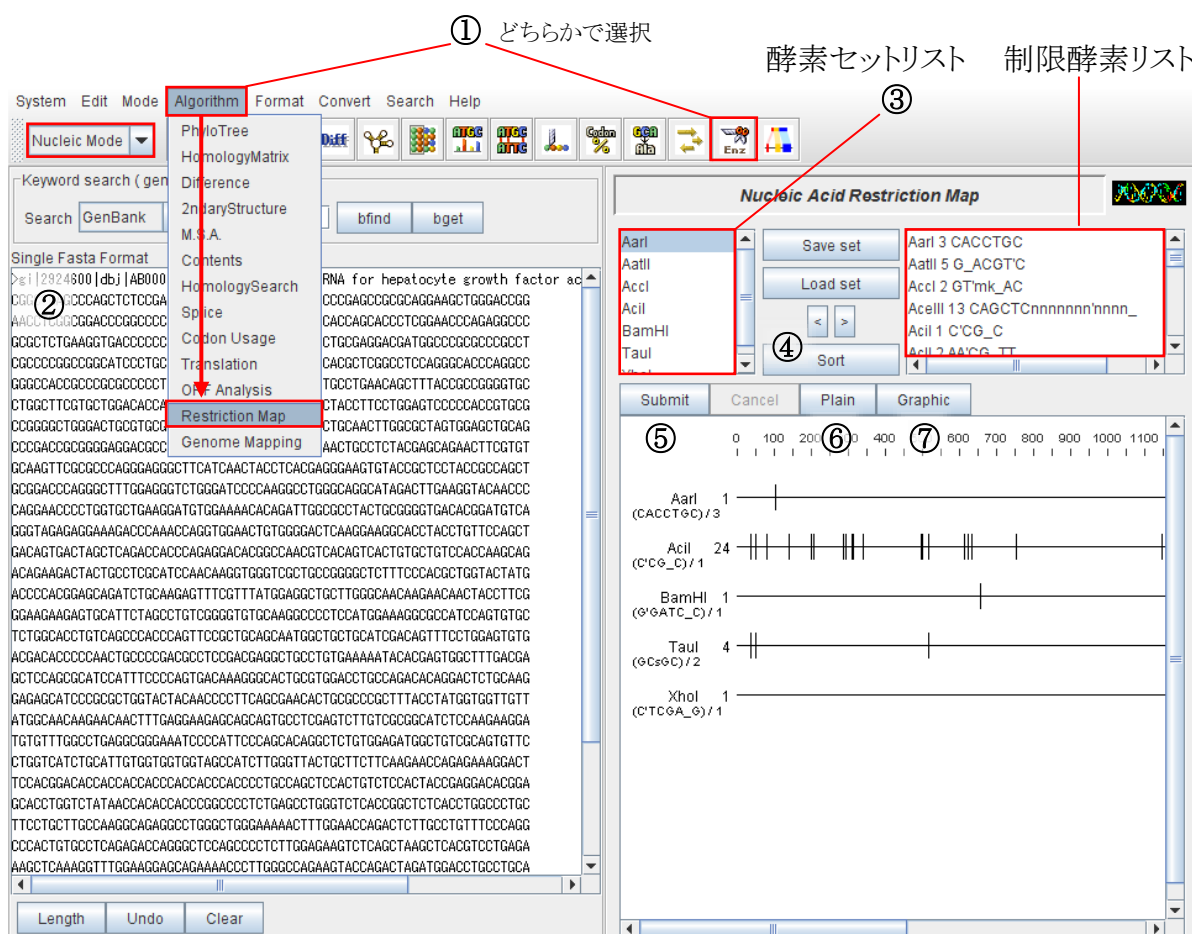


図 5-13-1 制限酵素地図作成画面

- ④ 「Sort」ボタンをクリックして、酵素セットリストを酵素名順（アルファベット順）にソートします。
- ⑤ 「Submit」ボタンをクリックすると、指定した制限酵素によって、核酸配列データの切断される位置情報の地図が結果表示領域に表示されます。
- ⑥ 「Plain」ボタンをクリックすると、表示された地図の結果データが別ウィンドウにテキスト形式で表示されます。(図 5-13-2) 「Output」ボタンをクリックすると、結果が別ブラウザ画面に表示されます。ブラウザ画面にて「名前を付けて保存」を選択するなどにより、出力結果を保存することが可能です。

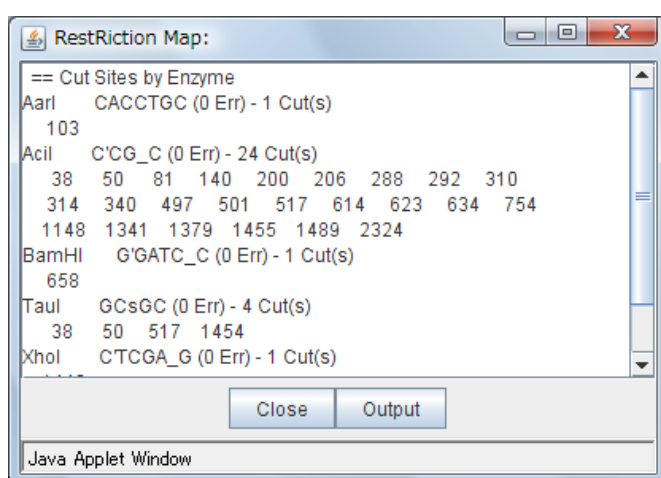


図 5-13-2 制限酵素地図 (テキスト形式)

- ⑦ 「Graphic」ボタンをクリックすると、別ウィンドウに塩基配列と切断される位置情報がグラフィカルに表示されます。(図 5-13-3)  
左側のウィンドウ内の数字をクリックすると、その領域が表示されます。

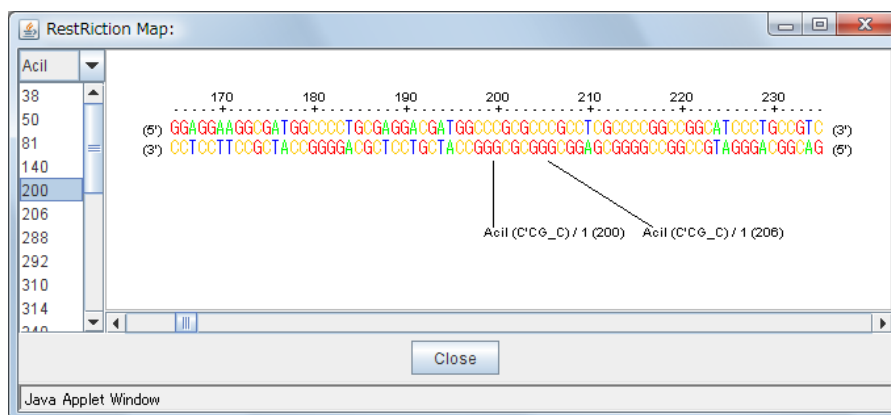


図 5-13-3 制限酵素地図 (グラフィカル表示)

### <制限酵素の追加方法>

利用したい制限酵素が酵素セットに含まれない場合、制限酵素リストから追加して利用することができます。右側の制限酵素リストから利用する制限酵素をクリックし、「<」ボタンをクリックすると、選択されている酵素セットにその制限酵素が追加されます。(図 5-13-1)

### <制限酵素の削除方法>

酵素セットに表示されている制限酵素を削除するには、削除したい制限酵素をクリックし、「>」ボタンをクリックします。(図 5-13-1)

### <酵素セットの保存方法>

追加した制限酵素を酵素セットとして保存するには、「Save Set」をクリックすると、サーバ内に保存され、Request ID ダイアログが表示されます。(図 5-13-4)

\* 必要であれば、Request ID を手元にメモ等で保存してください。

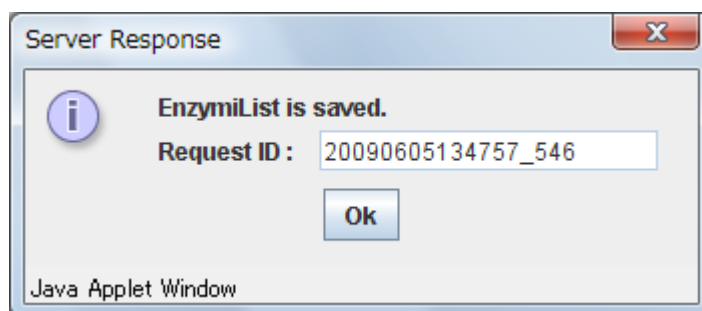


図 5-13-4 Request ID ダイアログ

### <保存した酵素セットの読み込み方法>

保存した酵素セットを読み込むには、「Load set」をクリックします。

入力ダイアログが表示されるので、保存の時に受け取った Request ID を入力し、「了解」ボタンをクリックします。

サーバ内に保存された酵素セットが酵素セットリストに読み込まれます。

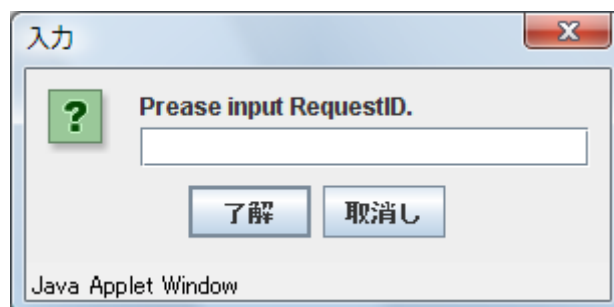


図 5-13-5 Request ID 入力ダイアログ

## 5. 14 ゲノムマッピング機能

核酸配列を指定された核酸配列ゲノムデータベースにマッピングします。サーバ内では、blast ver 2.2.16, sim4、または blat を使用しています。

<ゲノムマッピングの方法>

- ① 「Mode」メニューで「Nucleic」を選択し、「Algorithm」メニュー、またはツールアイコンで「Genome Mapping」を選択します。(図 5-14-1)
- ② データ編集領域に核酸配列データを fasta 形式で入力します。複数の配列を入力することも可能です。
- ③ 相同性の検出には blast を使用していますので、blast のオプションである E-value の値を入力します。

**！注意** E-value は実数値とし、指数表記での指定もできます。例) 1E-2 (0.01)

① どちらかで選択

②

③

④

⑤

⑥

⑦

⑧

⑨

query	chr.	identity	match	query st...	query st...	chr.start	chr.stop
gi 1445...	chr16	100	132	1	132	166679	166810
gi 1445...	chr16	100	205	133	337	166928	167132
gi 1445...	chr16	100	239	338	576	167282	167520

⑩

⑪

genome database : **Human (UCSC hg18)**

chr.name : **chr16** size : 88827254 bp

(display : 165873bp-168326bp)

query : gi|14456711|ref|NM\_000558.3| size : 576 bp

166500 166700 166900 167100

図 5-14-1 ゲノムマッピング画面

- ④ フィルタリングを行うかを指定します。
- ⑤ gap を考慮するかを指定します。
- ⑥ mega blast を使用するかを指定します。
- ⑦ ゲノムデータベースを指定します。(図 5-14-2)

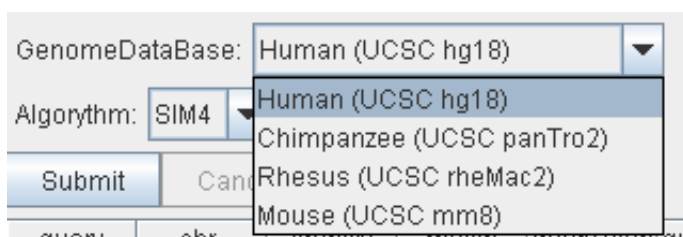


図 5-14-2 ゲノムデータベース選択画面

- ⑧ アルゴリズムを指定します。(図 5-14-3)

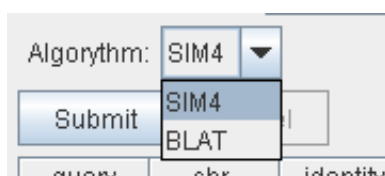


図 5-14-3 アルゴリズム選択画面

- ⑨ 「Submit」ボタンをクリックすると、マッピング結果が結果表示領域上部にテーブル形式(図 5-14-4)、結果表示領域下部にグラフィック形式で表示されます(図 5-14-5)。

query	chr.	identity	match	query start	query stop	chr.start	chr.stop
gi 14456...	chr16	100	132	1	132	166679	166810
gi 14456...	chr16	100	205	133	337	166928	167132
gi 14456...	chr16	100	239	338	576	167282	167520

図 5-14-4 マッピング結果テーブル

- ⑩ マッピング結果テーブルの 1 行を選択すると、その領域が Mapping View ウィンドウに描画され(図 5-14-5)、その領域の詳細情報が別ウィンドウに表示されます。(図 5-14-6)

- ⑪ エクソン領域グラフィック表示は、順方向にマッピングされた場合、その領域が青色、逆方向の相補鎖にマッピングされた場合、その領域が赤色、マッピング結果テーブルで選択されている領域は、少し薄い青、または橙色で表示されます。エクソン領域をクリックすると、その領域の詳細情報が別ウィンドウに表示されます。(図 5-14-6)

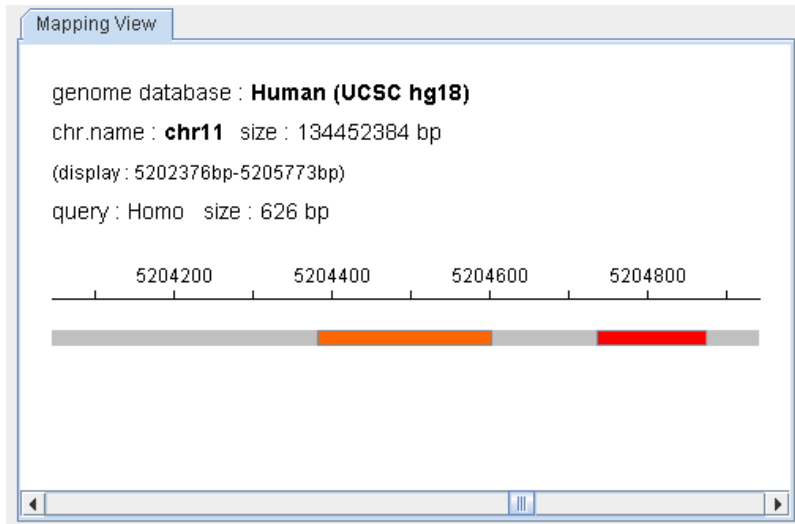


図 5-14-5 エクソン領域グラフィック表示

<エクソン領域詳細表示ウィンドウ>

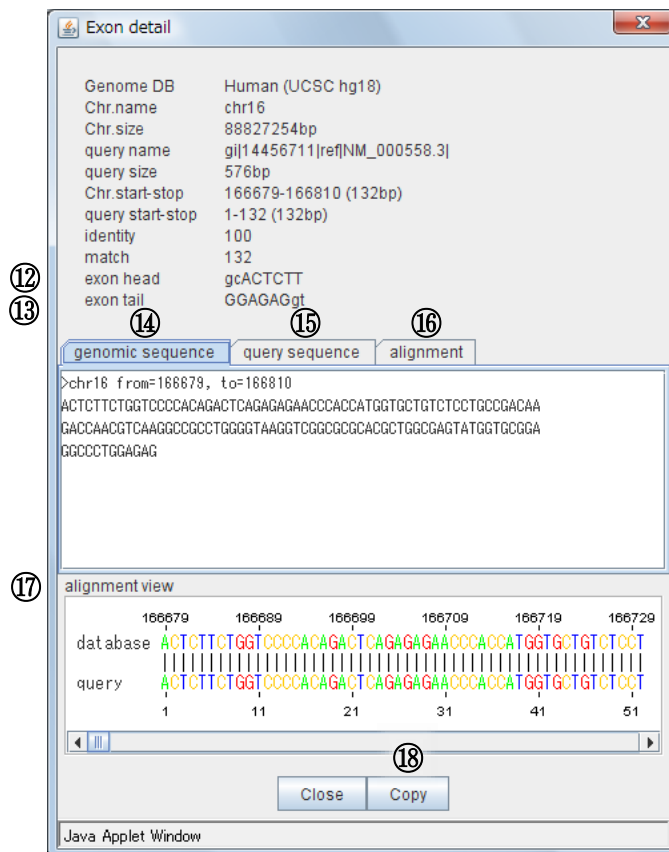


図 5-14-6 エクソン領域詳細ウィンドウ



- ⑫ exon head は、エクソンの開始位置の 2bp 前からエクソンの先頭 6bp までの配列情報です。
- ⑬ exon tail は、エクソンの終了位置の 6bp 前からエクソンの終了位置の後ろ 2bp までの配列情報です。
- ⑭ genomic sequence タブをクリックすると、マッピングされたゲノムデータベースの領域が fasta 形式で表示されます。
- ⑮ query sequence タブをクリックすると、マッピングされた query 配列の領域が fasta 形式で表示されます。
- ⑯ alignment タブをクリックすると、マッピングされたゲノムデータベースの領域と query 配列の領域とを ClustalW を用いてアライメントした結果がテキスト形式で表示されます。ClustalW の実行の際のパラメータは、デフォルトの値を使用しています。
- ⑰ alignment view は、マッピングされたゲノム配列の領域と query 配列の領域とを ClustalW を用いてアライメントした結果をグラフィカルに表示しています。
- ⑱ 「Copy」ボタンをクリックすると、選択されているタグに表示されている配列情報がデータ編集領域に表示されます。alignment タグがクリックされている場合は、アライメント結果が fasta 形式に変換されて、データ編集領域に表示されます。

# 6. アミノ酸配列データ解析

## 6.1 系統樹作成機能

複数のアミノ酸配列の系統樹が作成され、結果が表示されます。サーバ内での処理は、*phylip ver. 3.66* を使用しています。

(<http://evolution.genetics.washington.edu/phylip.html>)

### <系統樹作成方法>

- ① 「Mode」メニューで「Amino」を選択し、「Algorithm」メニュー、またはツールアイコンで「PhyloTree」を選択します。(図 6-1-1)
- ② データ編集領域にアミノ酸配列データを multi fasta 形式で入力します。  
複数の配列データを入力する必要があります。

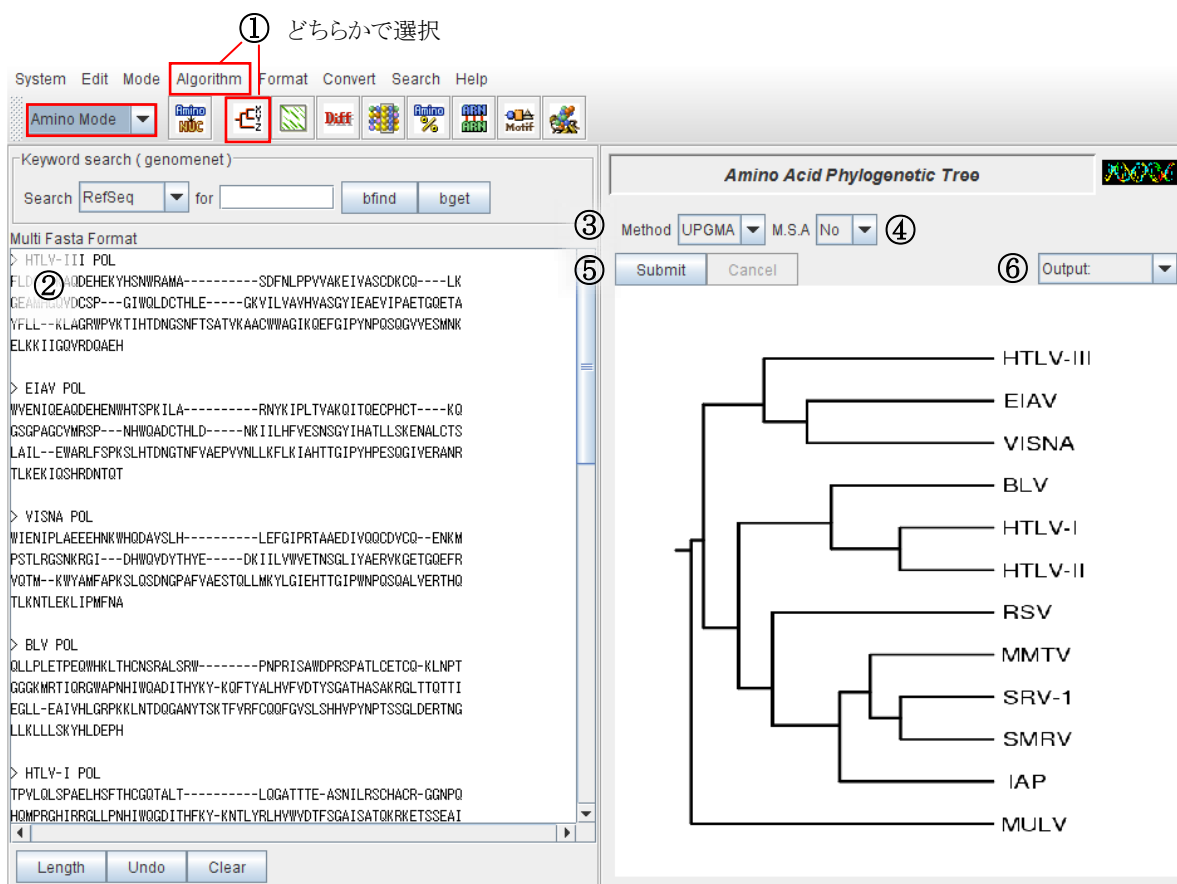


図 6-1-1 系統樹作成画面

③ 系統樹作成方法を選択します。(図 6-1-2)

UPGMA (平均距離) 法と NJ (近隣結合)法が選択可能です。

\*UPGMA 法、NJ 法は配列対間の遺伝距離に基づく方法です。

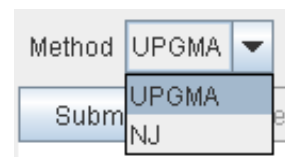


図 6-1-2 系統樹作成方法選択画面

④ あらかじめマルチプルシーケンスアライメントを実行するか選択します。(図 6-1-3)

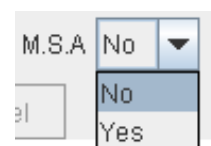


図 6-1-3 マルチプルシーケンスアライメント実行選択画面

### ！注意

- 入力する核酸配列がアライメントされていない場合、必ず「Yes」を指定してください。
- 各配列の長さが違い、MSA が No の場合、警告メッセージが出力されます。
- MSA が Yes の場合、サーバ内で ClustalW を用いてアライメントした後、系統樹が作成されます。

⑤ 「Submit」ボタンをクリックすると、系統樹が結果表示領域に表示されます。

ここで表示される系統樹は、入力された配列から作成した推定系統樹です。

⑥ プルダウンメニューから以下の出力形式を選択すると、別画面に結果が表示されます。  
(図 6-1-4)

gif	一般的な画像形式のひとつ。
ps	高品質な印刷が可能な画像形式。
distance matrix	配列間の進化的距離の行列。
tree	系統樹の Newick 形式のテキスト表記。

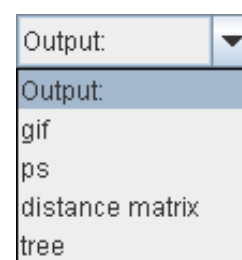


図 6-1-4 出力形式選択画面

< distance matrix 形式 >

distance matrix で表示されている値は、2 つの配列で異なるサイトの割合(P)に対して、Kimura モデルに従って補正を行った補正距離(D)です。

12						
HTLV-III	0.000000	1.146224	1.408767	1.773803	1.806792	1.599488
1.688766	1.867104	1.429310	1.610670	1.943057	1.956716	
EIAV	1.146224	0.000000	1.038411	1.970407	1.677109	1.718419
1.538489	1.819324	1.460988	1.610670	1.665458	1.637510	
VISNA	1.408767	1.038411	0.000000	1.943987	1.549259	1.418863
1.610670	1.277748	1.169388	1.277748	1.368978	1.621852	
BLV	1.773803	1.970407	1.943987	0.000000	0.950712	0.851625
1.661650	1.599488	1.458292	1.599488	1.544168	1.786274	
HTLV-I	1.806792	1.677109	1.549259	0.950712	0.000000	0.517379
1.544168	1.428267	1.208549	1.241291	1.351964	1.564919	
HTLV-II	1.599488	1.718419	1.418863	0.851625	0.517379	0.000000
1.479262	1.370939	1.162691	1.217389	1.438507	1.468174	
RSV	1.688766	1.538489	1.610670	1.661650	1.544168	1.479262
0.000000	1.171355	1.208549	1.194087	1.351964	1.957931	
MMTV	1.867104	1.819324	1.277748	1.599488	1.428267	1.370939
1.171355	0.000000	0.629565	0.741145	0.893250	1.624595	
SRV-1	1.429310	1.460988	1.169388	1.458292	1.208549	1.162691
1.208549	0.629565	0.000000	0.514703	0.798508	1.418521	
SMRV	1.610670	1.610670	1.277748	1.599488	1.241291	1.217389
1.194087	0.741145	0.514703	0.000000	0.876747	1.700199	
IAP	1.943057	1.665458	1.368978	1.544168	1.351964	1.438507
1.351964	0.893250	0.798508	0.876747	0.000000	1.758849	
MULV	1.956716	1.637510	1.621852	1.786274	1.564919	1.468174
1.957931	1.624595	1.418521	1.700199	1.758849	0.000000	

$$P = m/n$$

P : 2つの配列で異なるサイトの割合

m : 2つの配列の間で異なる塩基の数

n : 配列の長さ

$$D = -\log(1 - P - 0.2 \times P^2)$$

< tree 形式 >

tree 形式で表示されている値は、各アルゴリズムに基づいて計算された枝長です。

((HTLV-III:0.63875,(EIAV:0.51921,VISNA:0.51921):0.11954):0.16951,  
 ((BLV:0.45058,(HTLV-I:0.25869,HTLV-II:0.25869):0.19189):0.25962,  
 (RSV:0.61574,((MMTV:0.34268,(SRV-1:0.25735,SMRV:0.25735):0.08533):0.08541,  
 IAP:0.42808):0.18766):0.09446):0.09806):0.03245,MULV:0.84071);

## 6.2 ホモロジーマトリックス作成機能

2つのアミノ酸の相同性をホモロジーマトリックスで表示します。  
サーバ内では blast ver. 2.2.16 を使用して相同性を検出しています。

<ホモロジーマトリックス作成方法>

- ① 「Mode」メニューで「Amino」を選択し、「Algorithm」メニュー、またはツールアイコンで「Homology Matrix」を選択します。(図 6-2)
- ② データ編集領域にアミノ酸配列データを fasta 形式で入力します。Sequence1 と Sequence2 に、それぞれひとつずつ配列を入力する必要があります。
- ③ 相同性の検出には、blast を使用していますので、blast のオプションである E-Value と Word Size の値を入力します。「Default」 ボタンをクリックすると、初期値 (E-Value: 10、Word Size: 3 )に戻ります。

### ! 注意

- Word Size は、Word Size = 2 or 3
- E-Value は実数値とし、指数表記での指定もできます。 例) 1E-2(0.01)

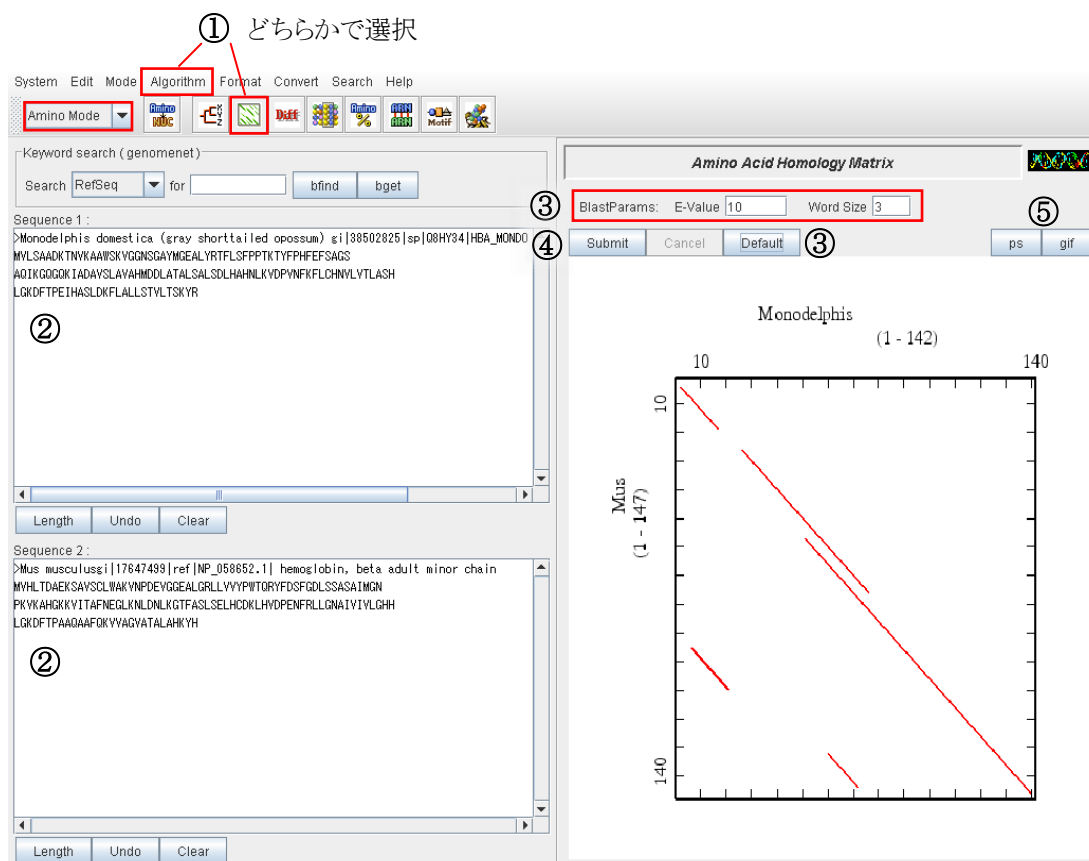


図 6-2 ホモロジーマトリックス表示画面

- ④ 「Submit」ボタンをクリックすると、ホモロジーマトリックスが結果表示領域に表示されます。
- ⑤ 「ps」ボタン、または「gif」ボタンをクリックすると、別画面に結果が表示されます。

## 6.3 相違度計算機能

複数のアミノ酸配列の相違度が計算され、結果が表示されます。  
サーバ内での処理は、prodifを使用しています。

<相違度計算の方法>

- ① 「Mode」メニューで「Amino」を選択し、「Algorithm」メニュー、またはツールアイコンで「Difference」を選択します。(図 6-3)
- ② データ編集領域に、アライメントされているアミノ酸配列データを fasta 形式で入力します。複数の配列データを入力する必要があります。
- ③ **Region** に開始位置と終了位置を入力します。何も指定していなければ、すべての範囲が対象として処理されます。

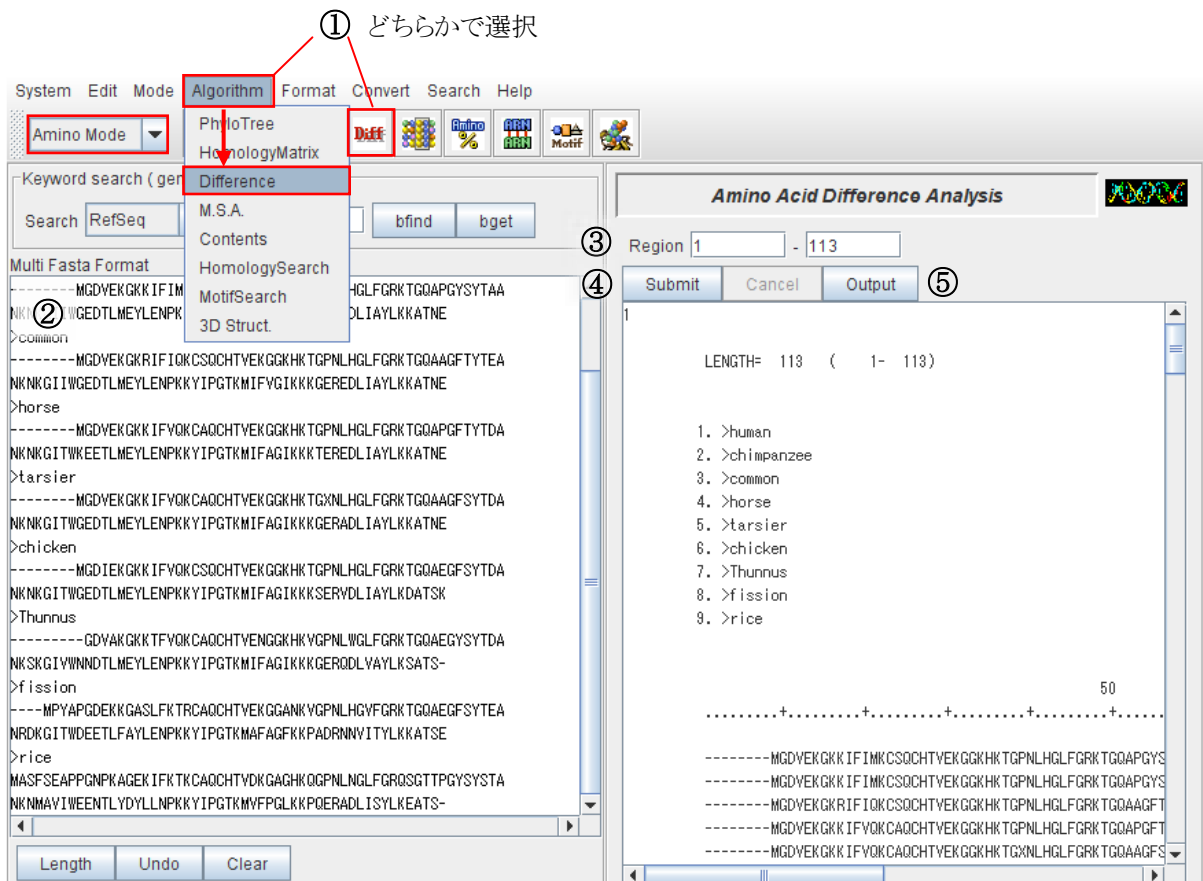


図 6-3 相違度計算機能画面

- ④ 「Submit」ボタンをクリックすると、計算結果が結果表示領域に表示されます。
- ⑤ 「Output」ボタンをクリックすると、結果が別ブラウザ画面に表示されます。ブラウザ画面にて「名前を付けて保存」を選択するなどにより、出力結果を保存することができます。



## 6. 4 アミノ酸配列マルチプルアライメント機能

複数のアミノ酸配列データをアライメントし、結果を表示します。サーバ内での処理は ClustalW ver. 81 (<http://www.ebi.ac.uk/clustalw>) か、MAFFT Ver. 6.705b (<http://align.bmr.kyushu-u.ac.jp/mafft/software>) を使用しています。

<アライメントの方法>

- ① 「Mode」メニューで「Amino」を選択し、「Algorithm」メニュー、またはツールアイコンで「M.S.A.」を選択します。(図 6-4-1)
- ② データ編集領域にアミノ酸配列データを multi fasta 形式で入力します。複数の配列を入力する必要があります。

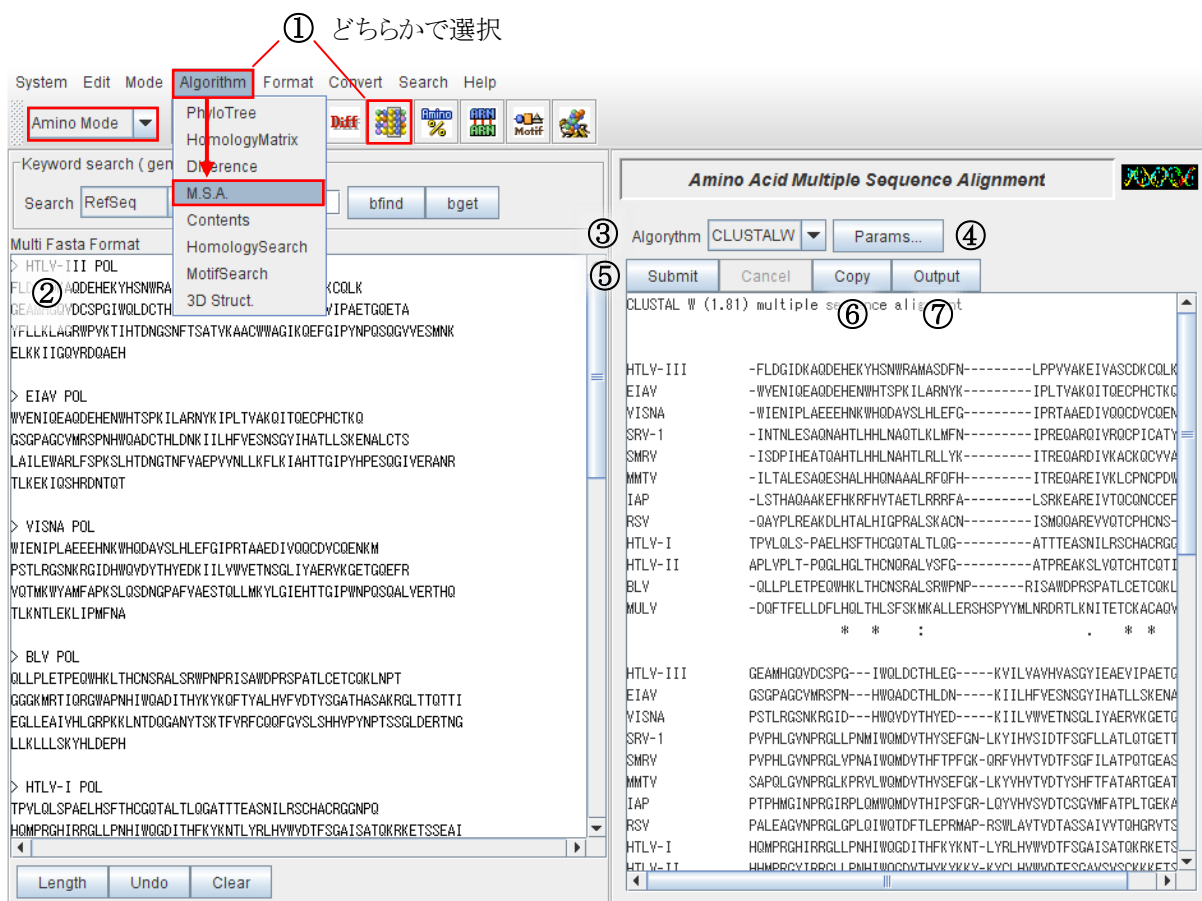


図 6-4-1 MSA 機能画面

- ③ アルゴリズムを選択します。CLUSTALW か MAFFT を選択します。(図 6-4-2)

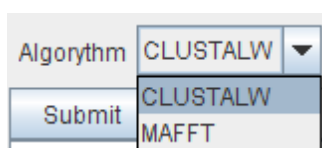


図 6-4-2 アルゴリズム選択画面

- ④ 「Params…」ボタンをクリックし、パラメータを設定します。(図 6-4-3、図 6-4-4)
- ClustalW の場合、「%identify」、「Gap Separation」の値は整数、それ以外の値は実数とし、実数(整数)以外が入力されるとデフォルト値に戻ります。
- MAFFT の場合、一番下の項目「Max num of ...」の値は 1 以上の整数、それ以外の値は実数とし、実数(整数)以外が入力されるとデフォルト値に戻ります。

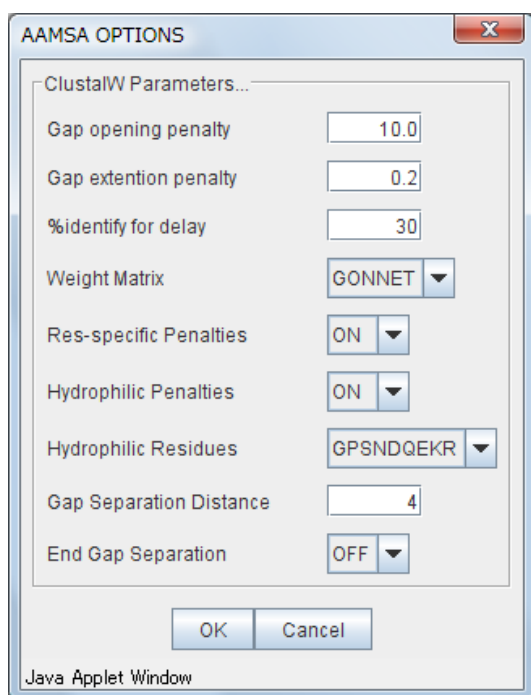


図 6-4-3 ClustalW パラメータダイアログ

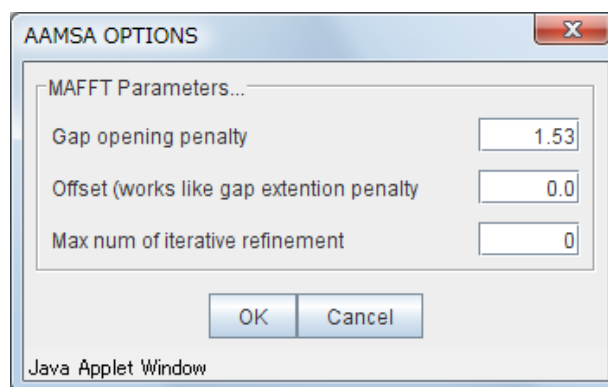


図 6-4-4 MAFFT パラメータダイアログ

- ⑤ 「Submit」ボタンをクリックすると、アライメント結果が結果表示領域に表示されます。

#### < 共通領域へのコピー方法 >

- ⑥ 「Copy」ボタンをクリックすると、アライメント結果が fasta 形式に変換されてデータ編集領域に表示されます。
- ⑦ 「Output」ボタンをクリックすると、結果が別ブラウザ画面に表示されます。ブラウザ画面にて「名前を付けて保存」を選択するなどにより、出力結果を保存することができます。

## 6.5 含量計算機能

アミノ酸配列の各アミノ酸の含有量(率)の計算を行い、結果が表示されます。

### <含量計算の方法>

- ① 「Mode」メニューで「Amino」モードに切り替え、「Algorithm」メニュー、またはツールアイコンで「Contents」を選択し機能を切り替えます。(図 6-5)
- ② データ編集領域にアミノ酸配列データを fasta 形式で入力します。複数の配列を入力することも可能です。
- ③ 計算する範囲を指定します。Region に開始位置と終了位置の間に「-」を入れ、入力します。範囲を複数指定する場合は、「,」で区切ります。何も指定していなければ、すべての範囲が対象として処理されます。
- ④ 「Submit」ボタンをクリックすると、含量計算結果が結果表示領域に表示されます。
- ⑤ 「Output」ボタンをクリックすると、結果が別ブラウザ画面に表示されます。ブラウザ画面にて「名前を付けて保存」を選択するなどにより、出力結果を保存することができます。

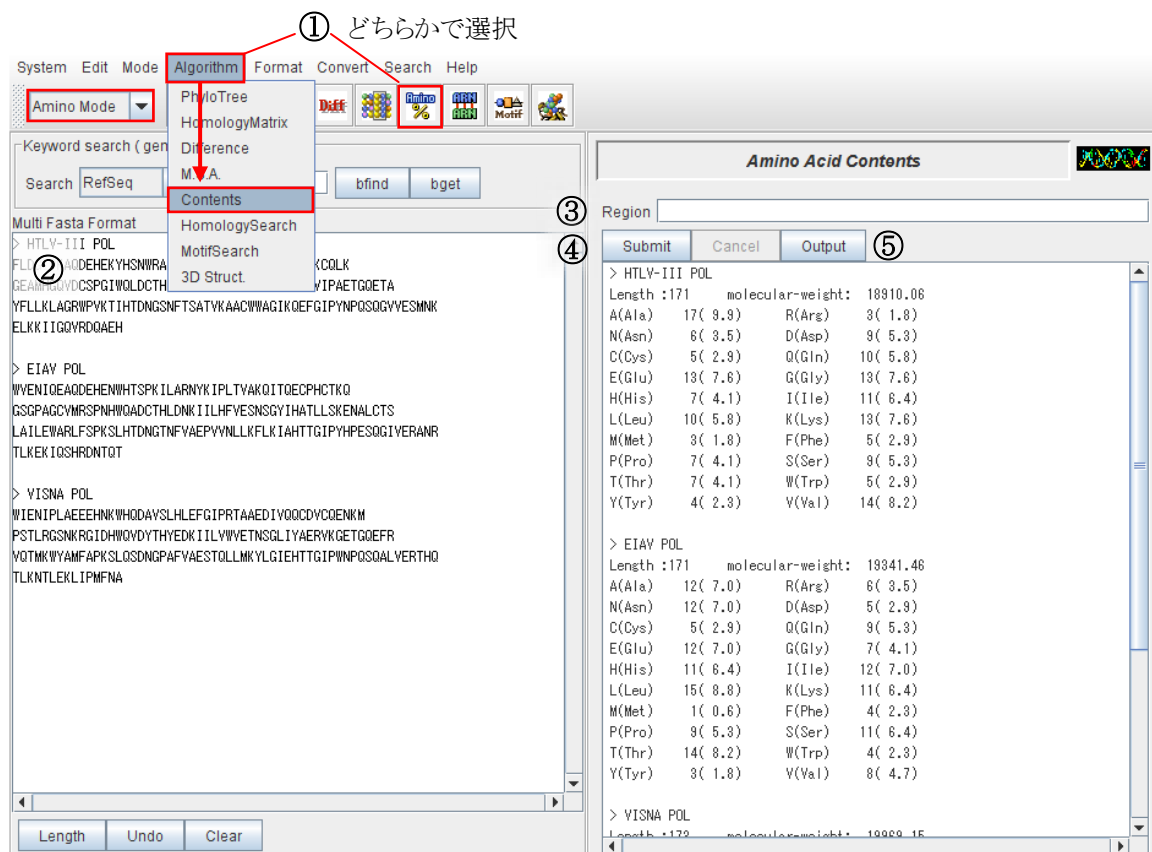


図 6-5 含量計算画面

<計算結果の説明>

⑥ 各アミノ酸の分子量は、下記のとおりです。(表 6-5)

表 6-5 アミノ酸の分子量

A	89.10	Q	146.15	L	131.17	S	105.09
R	174.21	E	147.13	K	146.19	T	119.12
N	132.12	G	75.07	M	149.21	W	204.21
D	133.10	H	155.16	F	165.19	Y	181.19
C	121.16	I	131.17	P	115.13	V	117.15

⑦ 各配列における molecular-weight の値は、(配列中に含まれるアミノ酸の数×アミノ酸の分子量) の値を 20 種類のアミノ酸分すべて加えて、最後に  $18 \times (\text{配列の長さ} - 1)$  を引きます。

**! 注意**

- molecular-weight の値は小数点第 3 位で四捨五入しています。
- 入力されたアミノ酸配列中に、20 種類のアミノ酸以外の文字が入っていた場合は、20 種類のアミノ酸の分子量の平均値で計算されます。(平均値 136.901 を 3 桁目四捨五入して 136.90) その場合、出力結果の molecular-weight の値の後に「？」がつけられます。

⑧ 各アミノ酸の結果は、アミノ酸 1 文字表記、(アミノ酸 3 文字表記)、アミノ酸の数、(割合) の順で出力されます。

**! 注意**

割合は (アミノ酸の数 / 配列の長さ  $\times 100$ ) の値を小数点第 2 位で四捨五入します。

## 6. 6 相同性検索機能(インタラクティブ処理)

指定されたアミノ酸配列データベースの中から相同性の高い配列を検索し、結果を表示します。相同性検索には、検索が終了するまで待って結果を表示するインタラクティブ処理と、検索要求のみを行って結果を後で見るバッチ処理があります。

時間を要する検索結果を行う場合は、バッチ処理を行います。(6. 7参照)

### <相同性検索の方法>

- ① 「Mode」メニューで「Amino」を選択し、「Algorithm」メニュー、またはツールアイコンで「Homology Search」を選択します。(図 6-6-1)
- ② データ編集領域にアミノ酸配列データを single fasta 形式で入力します。複数の配列を入力することはできません。

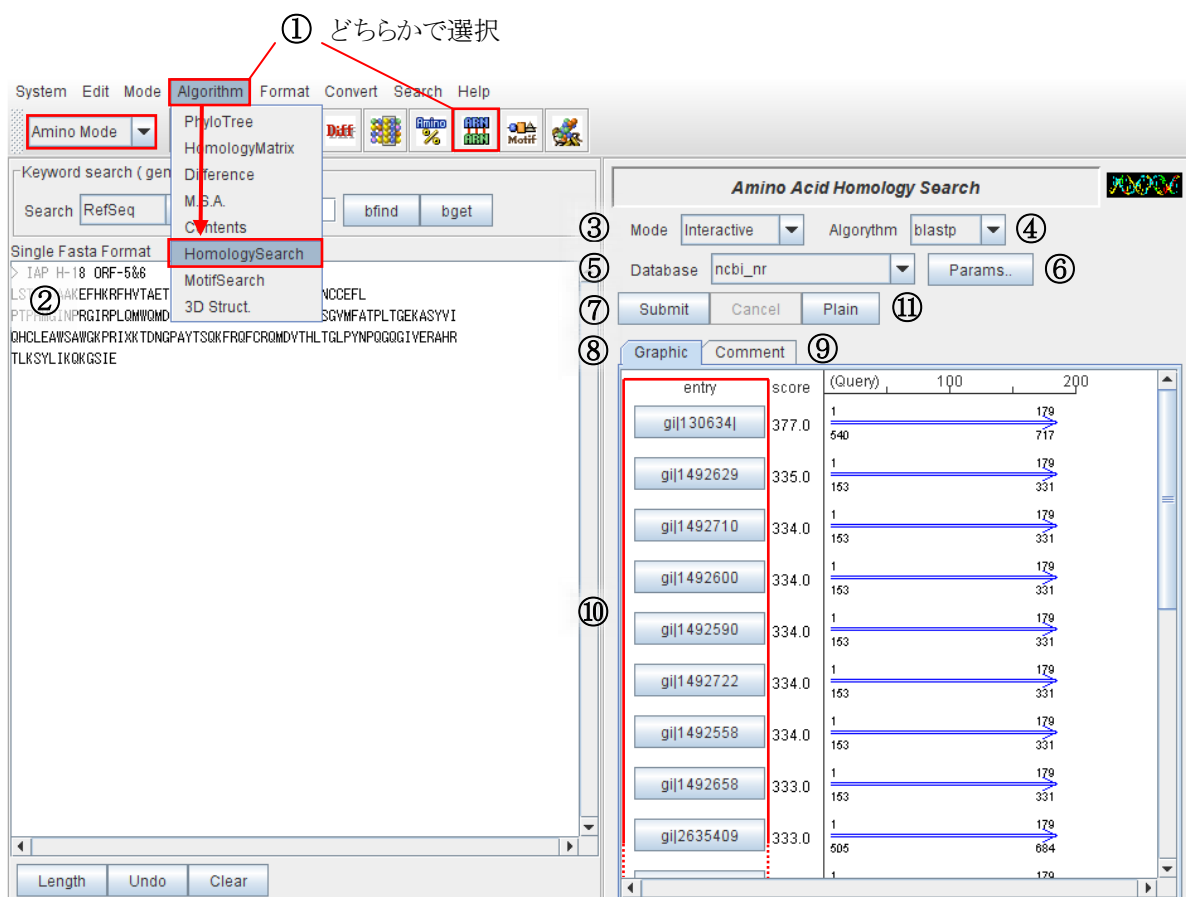


図 6-6-1 相同性検索画面 (Interactive)

③ 「Interactive」を選択します。(図 6-6-2)

BatchSearch、BatchView については 6-7 章にて説明します。

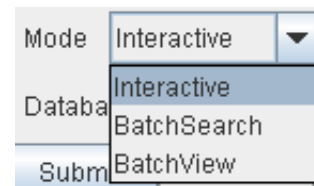


図 6-6-2 Mode 切替画面

④ アルゴリズムを選択します。(図 6-6-3)

選択可能なアルゴリズムは、blastp、tblastn、fasta3、tfasta3、tfastx3、pfasta の 6 種類です。(表 6-6 参照)

Database がアルゴリズムに対応するものに切り替わります。

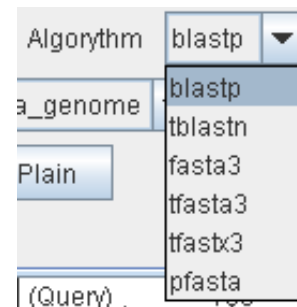


図 6-6-3 アルゴリズム選択画面

表 6-6 アルゴリズムの説明

プログラム	問い合わせ配列	データベース	検索結果	備考
blastp	タンパク質	タンパク質	タンパク質	
tblastn	タンパク質	核酸	タンパク質	データベースを翻訳しながら検索
fasta3	タンパク質	タンパク質	タンパク質	
tfasta3	タンパク質	核酸	タンパク質	データベースを翻訳しながら検索
tfastx3	タンパク質	核酸	タンパク質	データベースを翻訳しながら検索 (フレームシフト突然変異を考慮)
pfasta	タンパク質	nr-aa	タンパク質	遺伝情報ワークステーションによる並列計算システム (fasta3を使用)

⑤ データベースを選択します。(図 6-6-4)

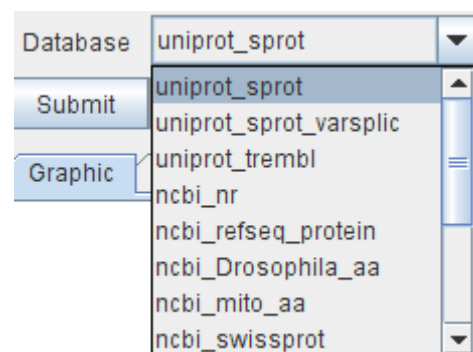


図 6-6-4 データベース選択画面

- ⑥ 「Params…」ボタンをクリックと、各アルゴリズムに対応したパラメータ設定パネルが表示され、パラメータの設定を行うことができます。
- ⑦ 「Submit」ボタンをクリックすると、相同性検索を行い、結果表示領域にリストが表示されます。

#### <グラフィック表示の方法>

- ⑧ 「Graphic」タブをクリックすると、グラフィック表示に切り替わります。(図 6-6-5)  
 矢印の向きは、検索結果の query 配列の向きです。  
 矢印の上の数值は、query 配列の start、stop 位置で、矢印の下の数值は、データベース配列の start、stop 位置です。

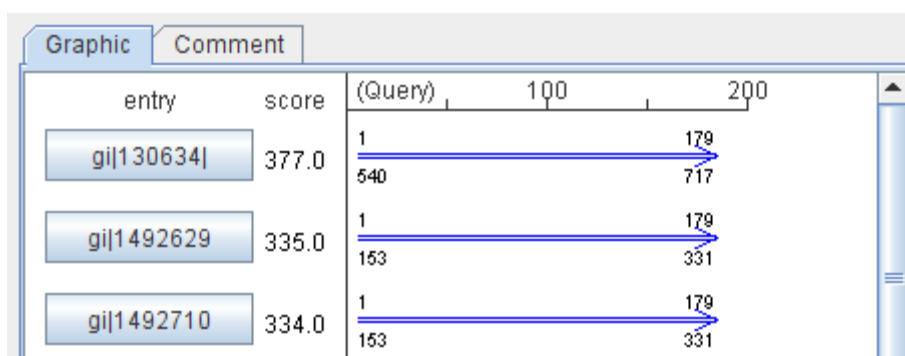


図 6-6-5 Graphic タブ選択時表示画面

#### <コメント表示の方法>

- ⑨ 「Comment」タブをクリックすると、コメント表示に切り替わります。(図 6-6-6)

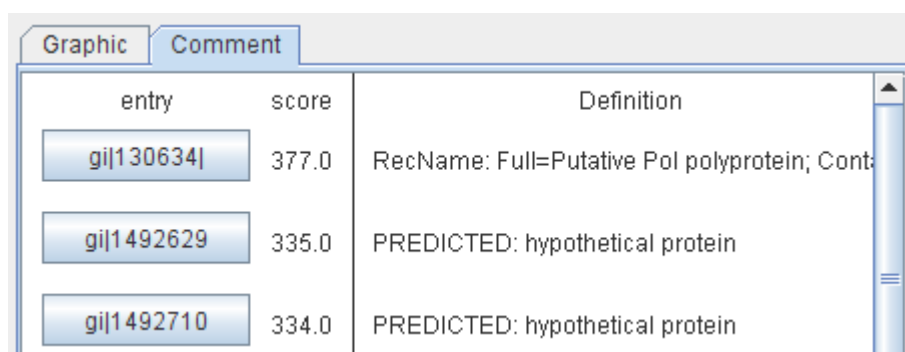


図 6-6-6 Comment タブ選択時表示画面

## <詳細情報の表示>

- ⑩ entry 名をクリックすると、相同性検索結果の詳細情報が別ウィンドウに表示されます。  
(図 6-6-7)

同時にクリックされた entry 名をキーワードとして、遺伝情報実験センターのデータベースに検索を行います。見つかった場合、別画面に結果が表示され(図 6-6-8)、見つからなかった場合、”No hits”と表示されたダイアログが出力されます。

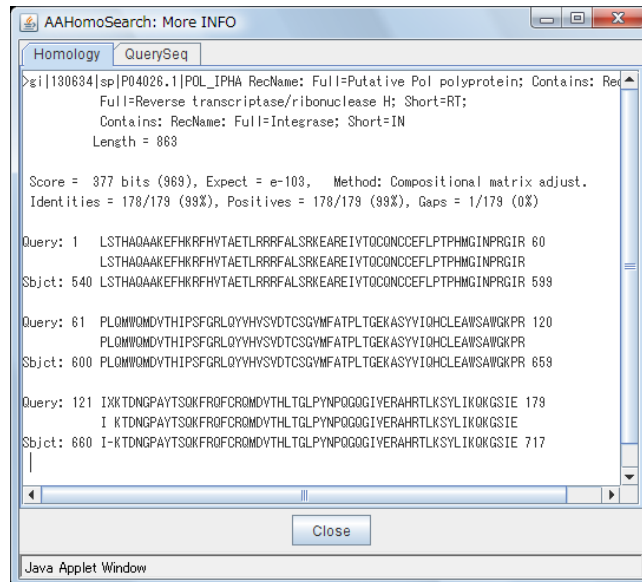


図 6-6-7 相同性検索詳細情報

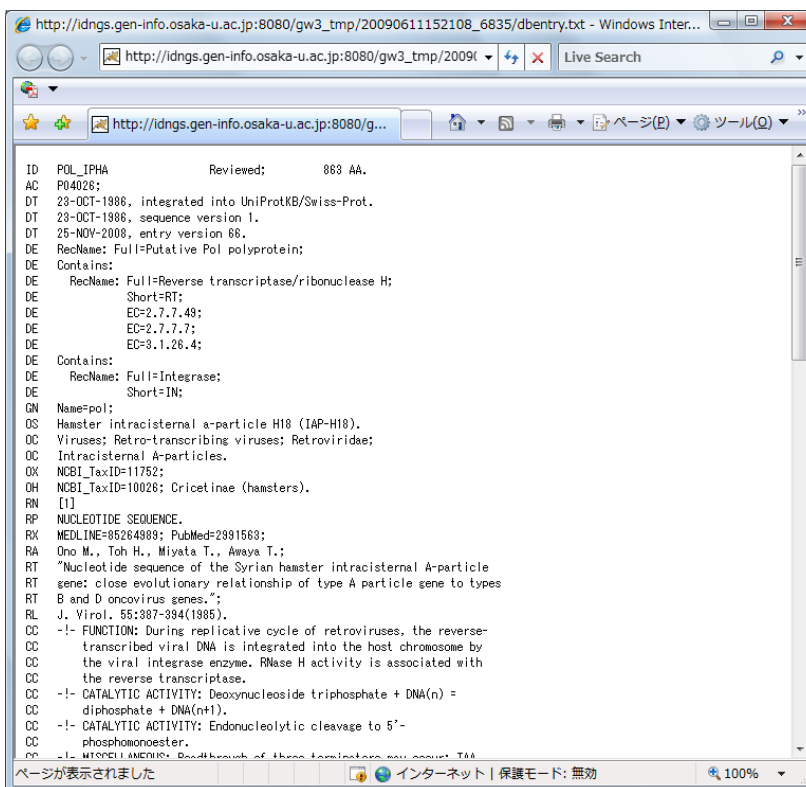


図 6-6-8 entry 名検索結果画面



- ⑪ 「Plain」ボタンをクリックすると、別ウィンドウに検索結果が表示されます。「Plain」タブをクリックすると Plain 形式 (図 6-6-9)、「table」タブをクリックするとテーブル形式 (図 6-6-10) で検索結果が表示されます。「Output」ボタンをクリックすると、検索結果が別ブラウザ画面に表示され、ブラウザ画面にて「名前を付けて保存」を選択するなどにより、出力結果を保存することが可能です。

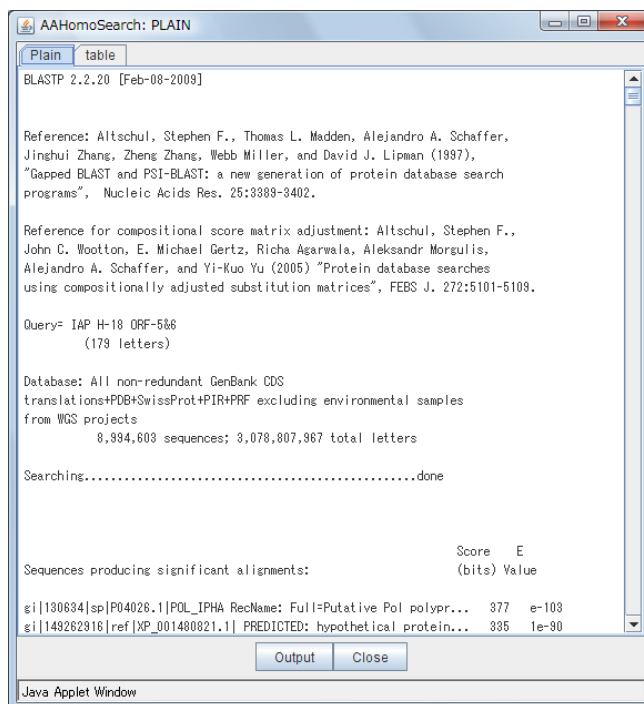


図 6-6-9 Plain 形式

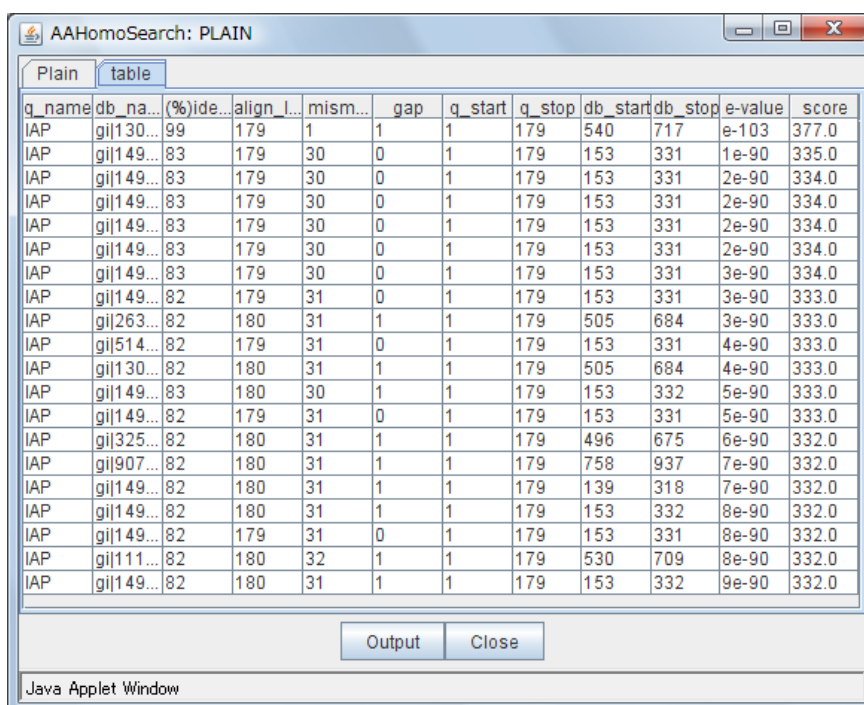


図 6-6-10 table 形式

### <blastp 用パラメータの指定方法>

- ⑫ 期待値 (Expectation Value) を実数で入力します。
- ⑬ アライメントを表示する数 (Alignment) を整数で入力します。(上限は 100)

**！注意** 数値以外が入力された場合、デフォルト値に戻ります。

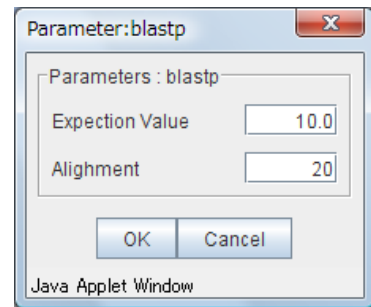


図 6-6-11 blastp パラメータ

### <tblastn 用のパラメータの指定方法>

Expectation Value、Alignment は、

⑫～⑬と同様。

- ⑭ マトリックスを選択します。

選択肢：  
PAM30  
PAM70  
BLOSUM80  
BLOSUM62  
BLOSUM45

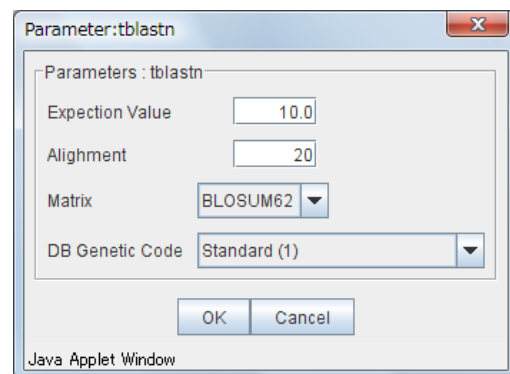


図 6-6-12 tblastn パラメータ

- ⑮ Database の Genetic Code を選択します。

選択肢：  
Standard  
Vertebrate Mitochondrial  
Yeast Mitochondria  
Mold Mitochondrial  
Invertebrate Mitochondrial  
Ciliate Nuclear  
Echinoderm Mitochondrial  
Euplotid Nuclear  
Bacterial  
Alternative Yeast Nuclear  
Ascidian Mitochondrial  
Flatworm Mitochondrial  
Blepharisma Macronuclear

< fasta3、tfasta3、tfastx3 用のパラメータの指定方法 >

- ⑩ KTUP を選択します。

選択肢 : 1,2

Expectation Value、Alignment は、

⑫～⑬と同様。

- ⑰ Reverse complement (相補鎖) を行うか選択します。

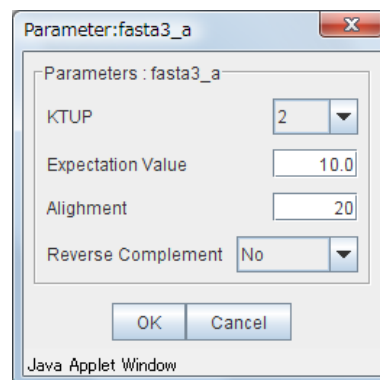


図 6-6-13 fasta3 パラメータ

< pfasta 用パラメータの指定方法 >

- ⑱ Alignment は、⑬と同様。

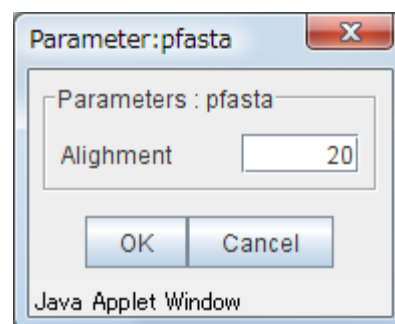


図 6-6-14 pfasta パラメータ

## 6. 7 相同性検索機能(バッチ処理)

6. 6の amino 酸配列データ相同性検索で処理時間が要する場合は、バッチ処理により、検索と結果の表示を分けて行うことができます。

- ① 「Mode」メニューで「Amino」を選択し、「Algorithm」メニュー、またはツールアイコンで「HomologySearch」を選択します。(図 6-7-1)

### < 検索の方法 >

- ② データ編集領域に amino 酸配列データを single fasta 形式で入力します。複数の配列データを入力することはできません。
- ③ 「Mode」メニューで「BatchSearch」を選択します。
- ④ Algorithm、Database、Params.. の設定をします。(インタラクティブ処理 p65-70 参照)

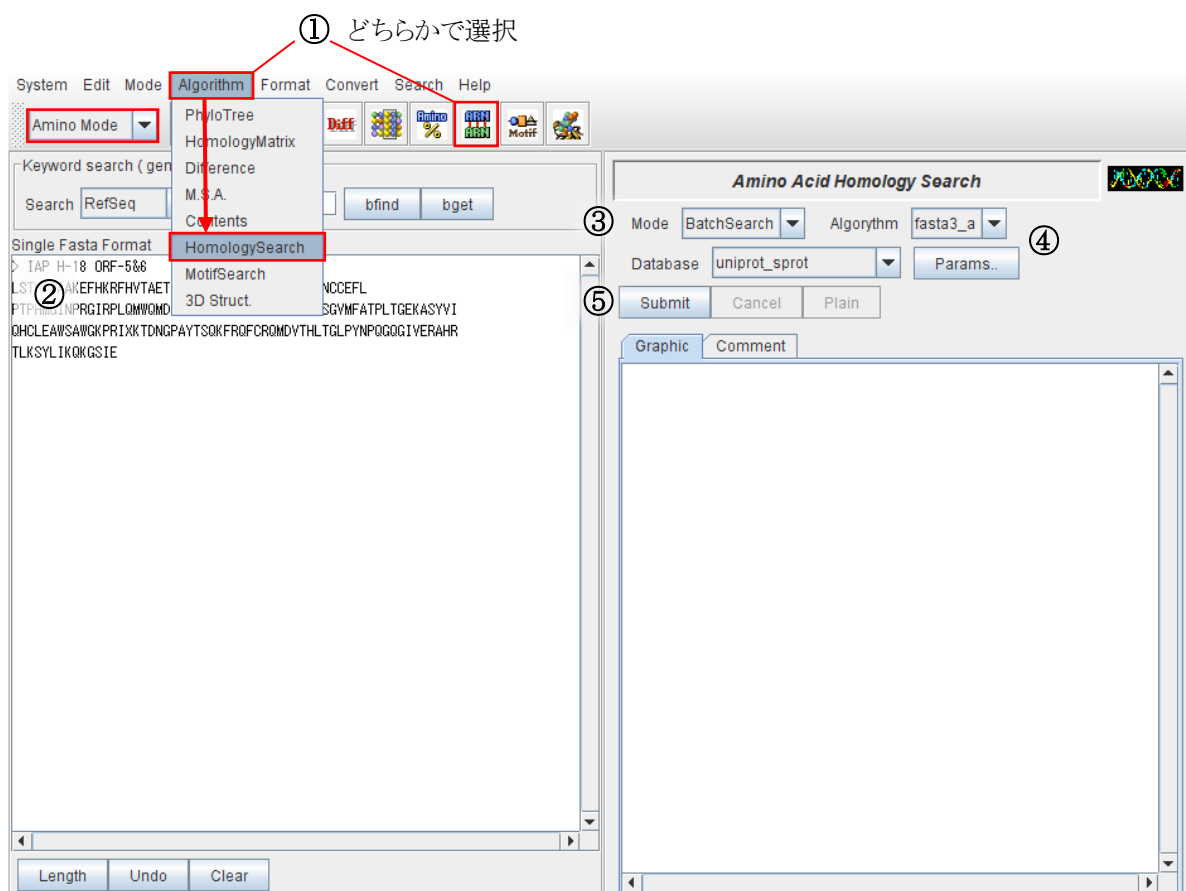


図 6-7-1 相同性検索画面 (BatchSearch)

- ⑤ 「Submit」ボタンをクリックすると、相同性検索を行い、結果を問い合わせるための Request ID のダイアログが表示されます。(図 6-7-2)

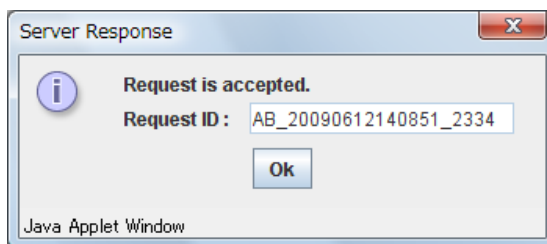


図 6-7-2 Request ID ダイアログ

### <結果表示の方法>

- ⑥ 「Mode」メニューで「BatchView」を選択します。(図 6-7-3)

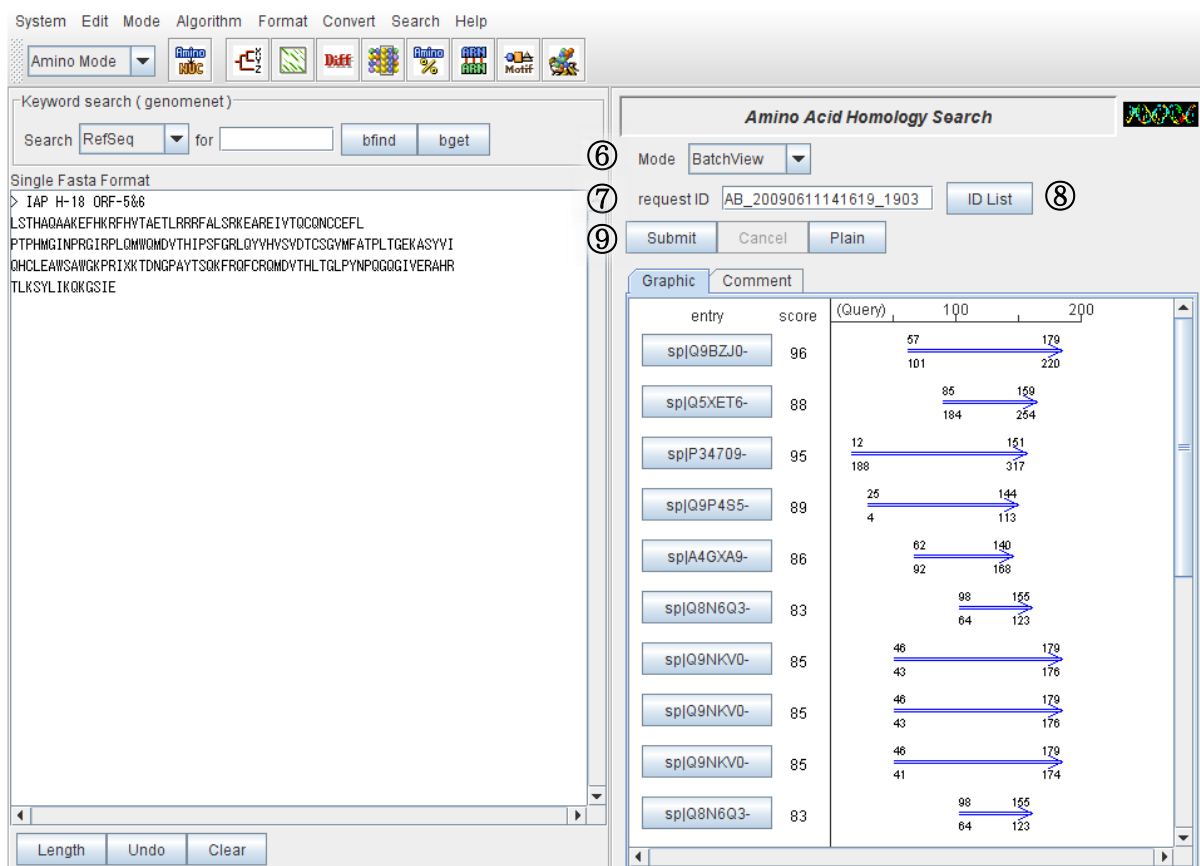


図 6-7-3 相同性検索画面 (BatchView)

- ⑦ 「BatchSearch」処理で受け取った Request ID を入力します。
- ⑧ 「ID List」ボタンをクリックすると、GeneWebIII 起動後の BatchSearch 処理リクエストに対する Request ID の一覧が表示され、詳細の確認ができます。(図 6-7-4)

RequestID	Date	Time	Algorithm	Database
AB_20090611141549_10...	2009/6/11	14:15:51	tblastn	ncbi_est_mouse
AB_20090611141554_18...	2009/6/11	14:15:55	tblastn	ncbi_est_human
AB_20090611141557_44...	2009/6/11	14:15:59	tblastn	ncbi_est_others
AB_20090611141601_27...	2009/6/11	14:16:3	tblastn	ncbi_env_nt
AB_20090611141615_89...	2009/6/11	14:16:17	fasta3_a	uniprot_sprot
AB_20090611141619_19...	2009/6/11	14:16:20	fasta3_a	uniprot_sprot_varsplc
AB_20090611141622_60...	2009/6/11	14:16:23	fasta3_a	uniprot_trembl
AB_20090611141626_30...	2009/6/11	14:16:27	fasta3_a	ncbi_nr
AB_20090611141630_58...	2009/6/11	14:16:31	fasta3_a	ncbi_refseq_protein
AB_20090611141634_57...	2009/6/11	14:16:35	fasta3_a	ncbi_Drosophila_aa
AB_20090611141637_820	2009/6/11	14:16:38	fasta3_a	ncbi_mito_aa

図 6-7-4 リクエスト ID 一覧  
(選択中の Request ID は色付背景で表示される)

このリクエスト ID 一覧画面において、RequestID をクリックすると、背景色が変わり、⑦の requestID 入力ボックスに自動入力されます。

### ！注意

GeneWebIII 起動以前のリクエストに対する Request ID は保存されていません。GeneWebIII を一度終了させ、後で結果を表示させたい場合は、Request ID をメモしておいてください。

この ID を入力することにより、以前実行させた結果を表示することができます。ただし、結果の保存期間は 1 ヶ月です。

- ⑨ 「Submit」をクリックすると、結果表示領域に結果が表示されます。

## 6.8 モチーフ検索機能

アミノ酸配列のモチーフ検索のページが、別画面に表示されます。

<モチーフ検索の方法>

- ① 「Mode」メニューで「Amino」を選択し、「Algorithm」メニュー、またはツールアイコンで「Motif Search」を選択します。(図 6-8-1)

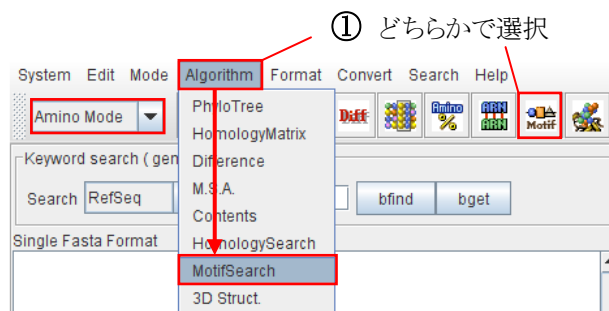


図 6-8-1 モチーフ検索選択画面

- ② 別ブラウザ画面が表示され、モチーフ検索のページに表示されます。(図 6-8-2)  
URLは <http://motif.genome.jp/> です。

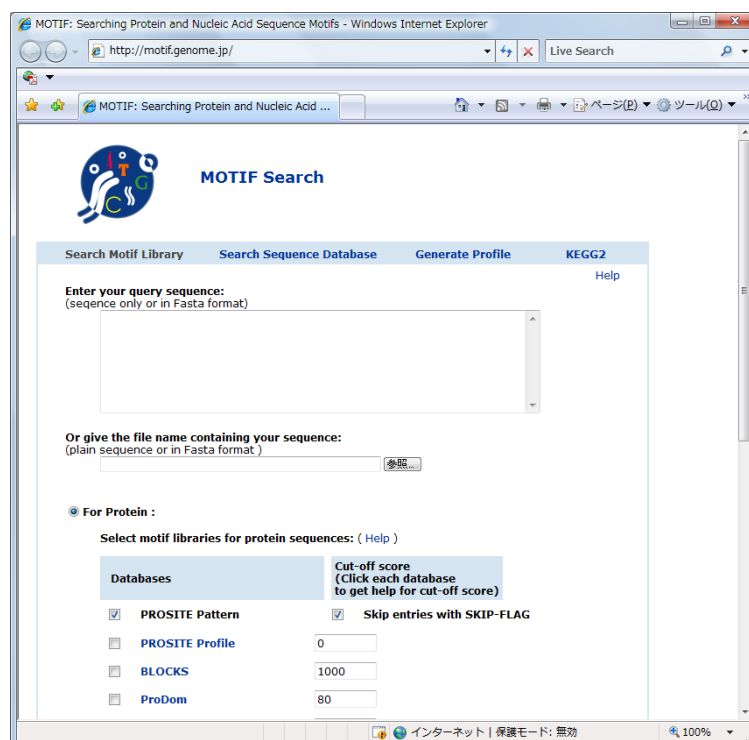


図 6-8-2 モチーフ検索ページ

- ③ モチーフ検索が、利用可能になります。  
詳細は、ホームページ内のマニュアルをご参照ください。

## 6.9 高次構造予測機能

立体構造既知のアミノ酸配列データベースとの相同性を利用して、新規のアミノ酸配列データの立体構造が予測されます。立体構造の予測には時間を要するため、問い合わせと結果の表示が別々に行われます。

サーバ内での処理は、ModellerRelease4 を使用しています。(A.Sali and T.L.Blundell. J.Mol.Biol., 234, 779-815, 1993)

- ① 「Mode」メニューで「Amino」を選択し、「Algorithm」メニュー、またはツールアイコンで「3D Struct.」を選択します。(図 6-9-1)

<問い合わせの方法>

- ② データ編集領域にアミノ酸配列データを single fasta 形式で入力します。複数の配列データを入力することはできません。

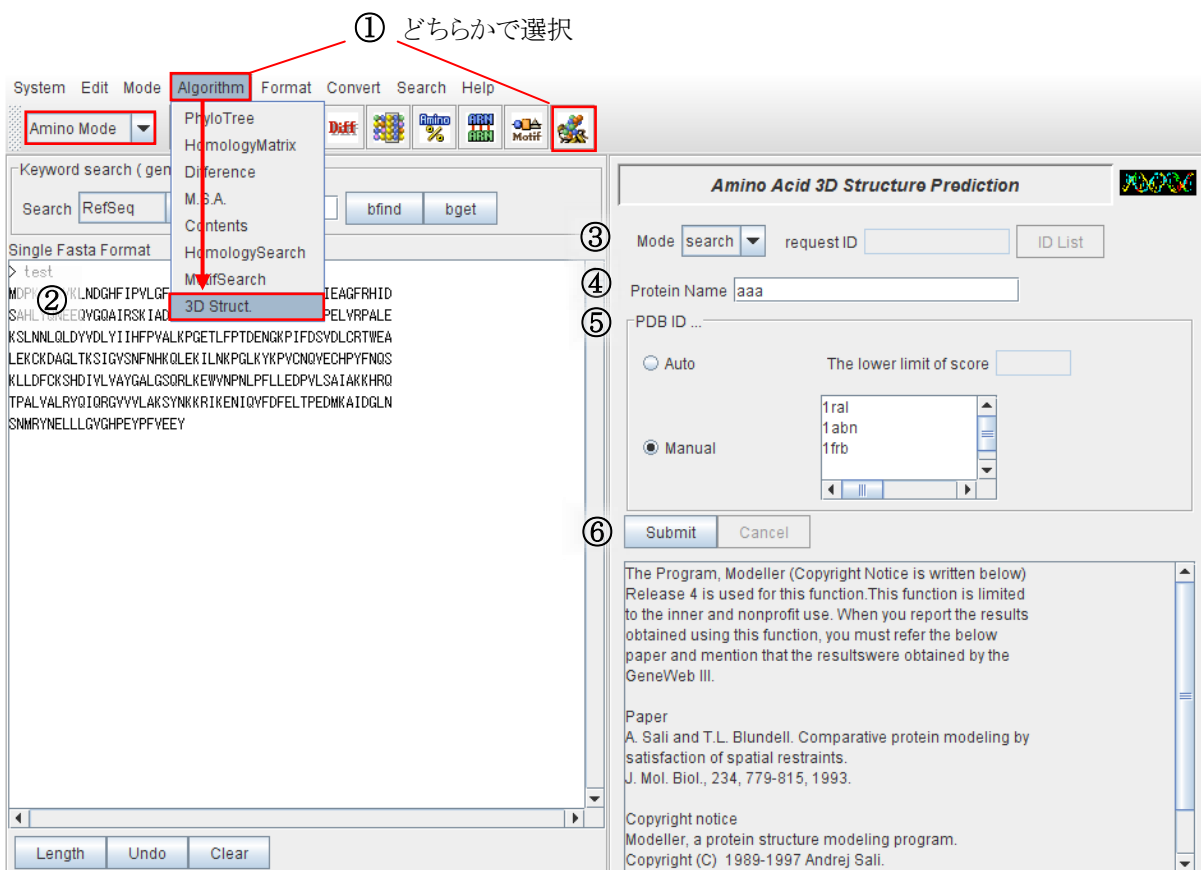


図 6-9-1 高次構造予測機能画面 (Search)



③ Mode で「Search」を指定します。(図 6-9-2)

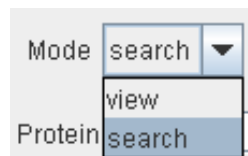


図 6-9-2 Mode 選択画面

④ Protein Name を入力します。(空白を含まない半角英数字で、最大 8 文字とします。)

⑤ Auto または Manual で、PDB ID の指定方法を選択します。

### ！注意

- Auto の時は Score の下限値を実数で入力します。サーバ内で pdb データベースに対して blast 検索を行い、検索結果の中で入力された Score 値以上の PDB ID を使用します。
- Manual の時は、PDB ID を最大 20 個まで指定できます。
- PDB ID の指定は、大文字、小文字のどちらでも入力できます。

⑥ 「Submit」ボタンをクリックすると、処理が開始され、結果を問い合わせるための Request ID のダイアログが表示されます。(図 6-9-3)

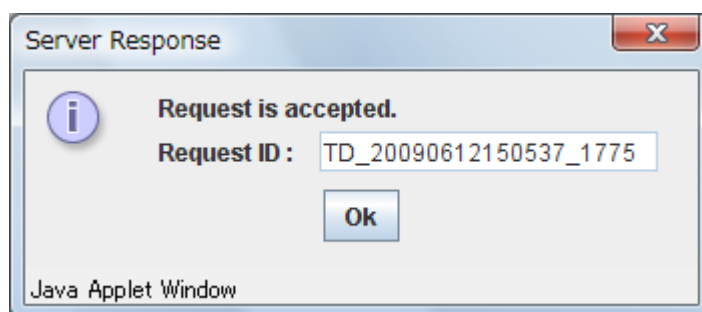


図 6-9-3 Request ID ダイアログ

### ！注意

サーバ内で実行されるソフトウェア (ModellerRelease4) のライセンスの制限により、大阪大学内のドメイン (133.1.xxx.xxx) 以外からのアクセスの場合、エラーメッセージダイアログが出力され、使用することができません。

<結果を表示する場合>

- ⑦ Mode で「View」を選択します。(図 6-9-4)

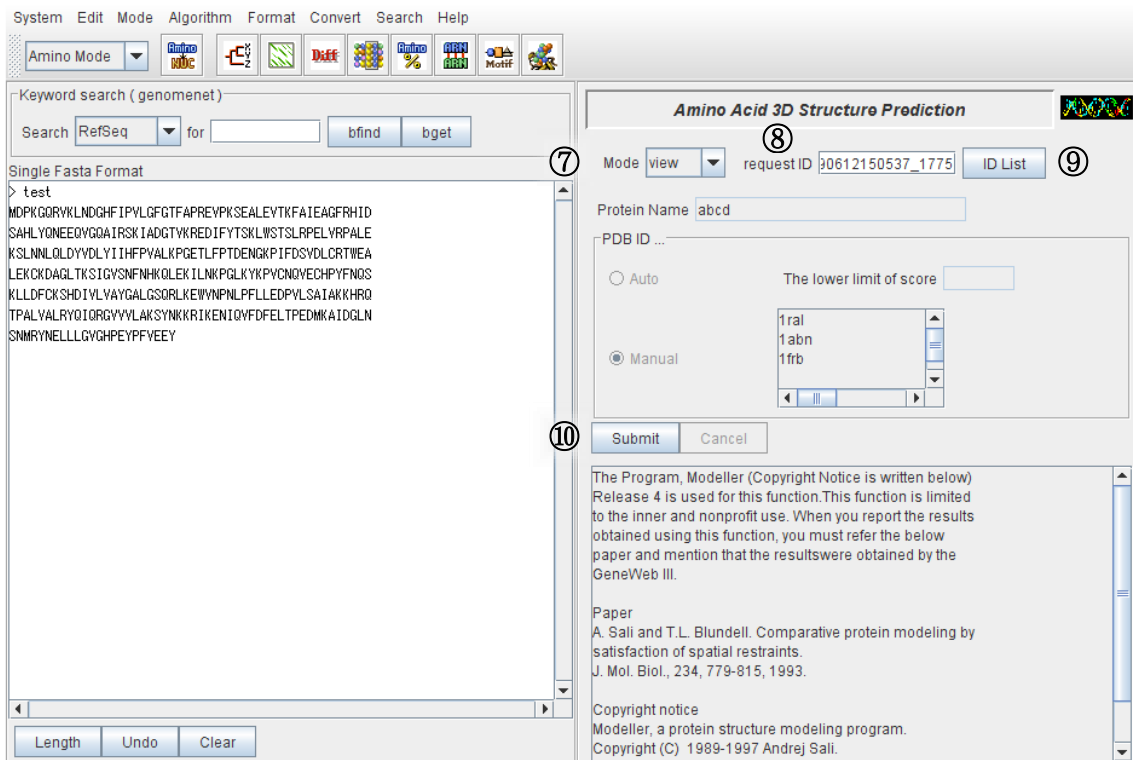


図 6-9-4 高次構造予測機能画面 (View)

- ⑧ 「Search」処理で受け取った Request ID を入力します。
- ⑨ 「ID List」ボタンをクリックすると、GeneWebIII 起動後の 3D Structure Prediction 処理リクエストに対する問い合わせ番号の一覧が表示され、Request ID を確認できます。(図 6-9-5)

RequestID	Date	Time	ProteinName
TD_20090612150537_1775	2009/6/12	15:6:18	aaa
TD_20090612150738_493	2009/6/12	15:7:41	abcd

図 6-9-5 リクエスト ID 一覧  
(選択中の Request ID は色付背景で表示される)

このリクエスト ID 一覧画面において、RequestID をクリックすると、背景色が変わり、⑧の requestID 入力ボックスに自動入力されます。

**！注意**

GeneWebIII 起動以前のリクエストに対する Request ID は保存されていません。

GeneWebIII を一度終了させ、後で結果を表示させたい場合は、Request ID をメモしておいてください。

この ID を入力することにより、以前実行させた結果を表示することができます。

ただし、結果の保存期間は 1 ヶ月です。

- ⑩ 「Submit」ボタンをクリックすると、結果 pdb 形式ファイルが別ブラウザ画面に表示されます。
- ⑪ 結果 pdb 形式ファイルを「ファイル名.pdb」として、名前付けて保存（ブラウザによってはテキスト形式を選択）を行い、RasMol と呼ばれるフリーソフトを用いて開けると、立体表示することができます。

## <RasMol のダウンロードの方法>

pdb ファイルの内容を表示するためには、RasMol と呼ばれるフリーソフトが必要です。RasMol を使用する前に現在使用しているパソコンに、「RasMol」をダウンロードしなければ、使用できません。

⑫ ブラウザで下記の URL アドレスを入力します。(図 6-9-6)

<http://openrasmol.org/>

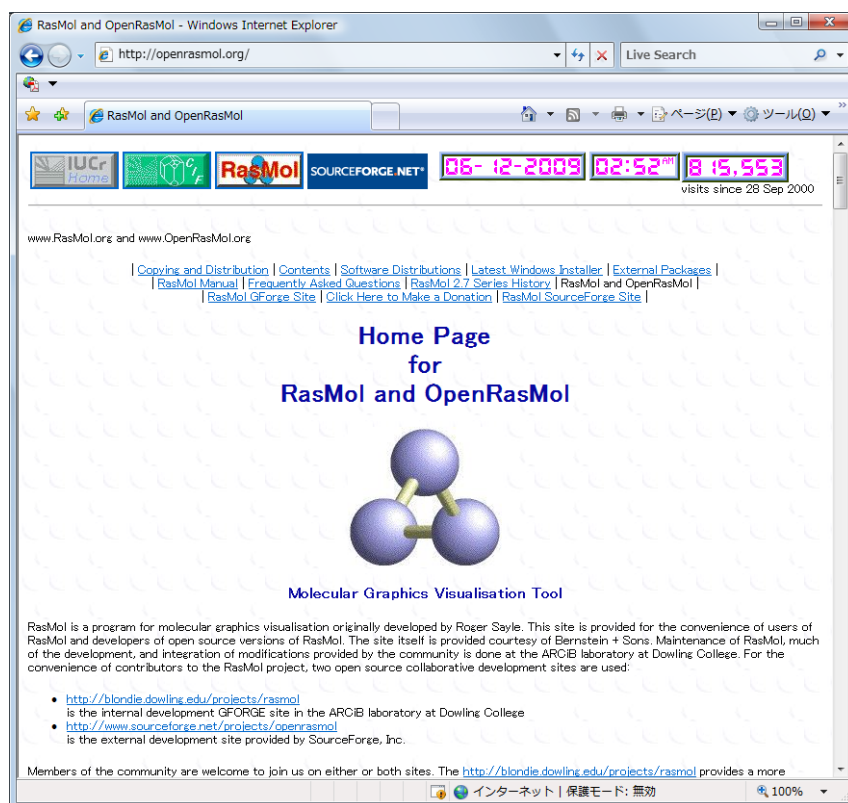


図 6-9-6 RasMol ホームページ

- ⑬ RasMol ホームページの「Software Distributions」からダウンロードするバージョンを選択します。(図 6-9-7)



図 6-9-7 Software Distributions

- ⑭ 使用しているコンピュータに合わせてバージョンを選び、ダウンロードします。(図 6-9-8)

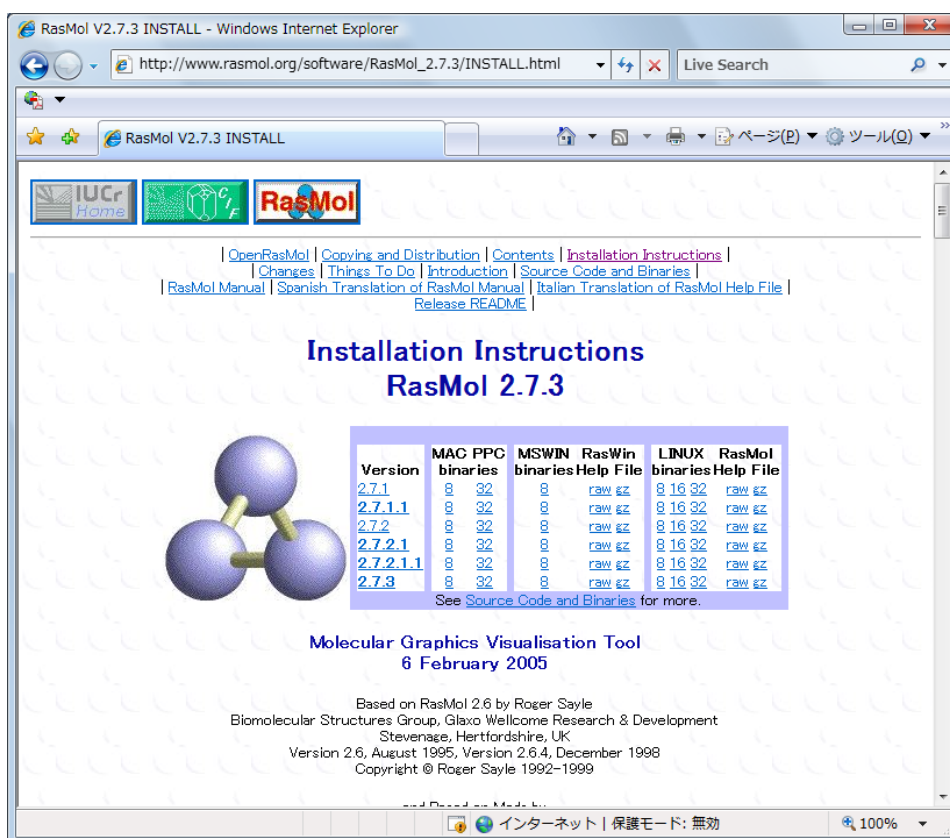


図 6-9-8 RasMol ダウンロード一覧