

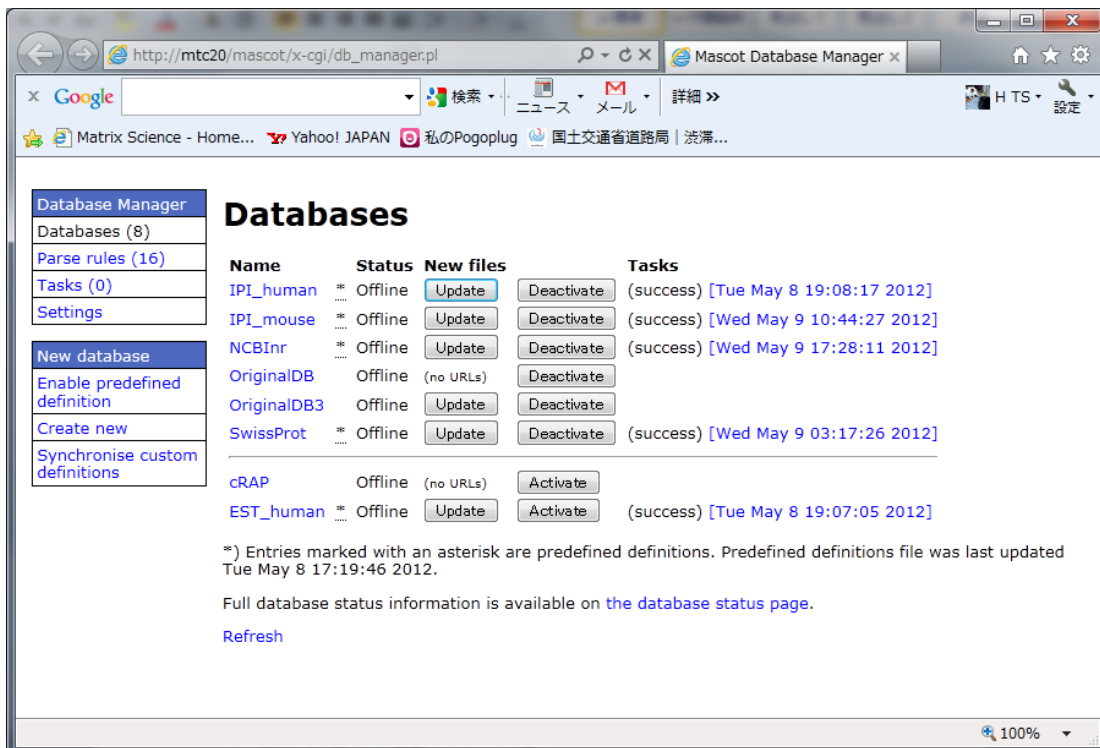
配列データベース管理マニュアル

● はじめに

これまでのバージョンでConfiguration Editor の「Database maintenance」 ページから管理していたMASCOT Server の設定内容が2つに分割され、主にデータベースの設定を中心とした「Database Manager」とMASCOT Serverの各種パラメータを設定する「Configuration Options」において管理する形式に変わりました。

「Database Manager」については、これまでのデータベース設定に加え、データベースファイルの自動更新の設定や最適なParse Rule の選別、さらには公開データベースのフォーマット変更にも対応し自動的に変更するなど、様々な機能が追加されました。

本資料では、インストール/バージョンアップ直後の初期設定方法と、「Database Manager」で使用可能な各機能について説明します。



● 本資料で取り扱う内容と掲載ページ

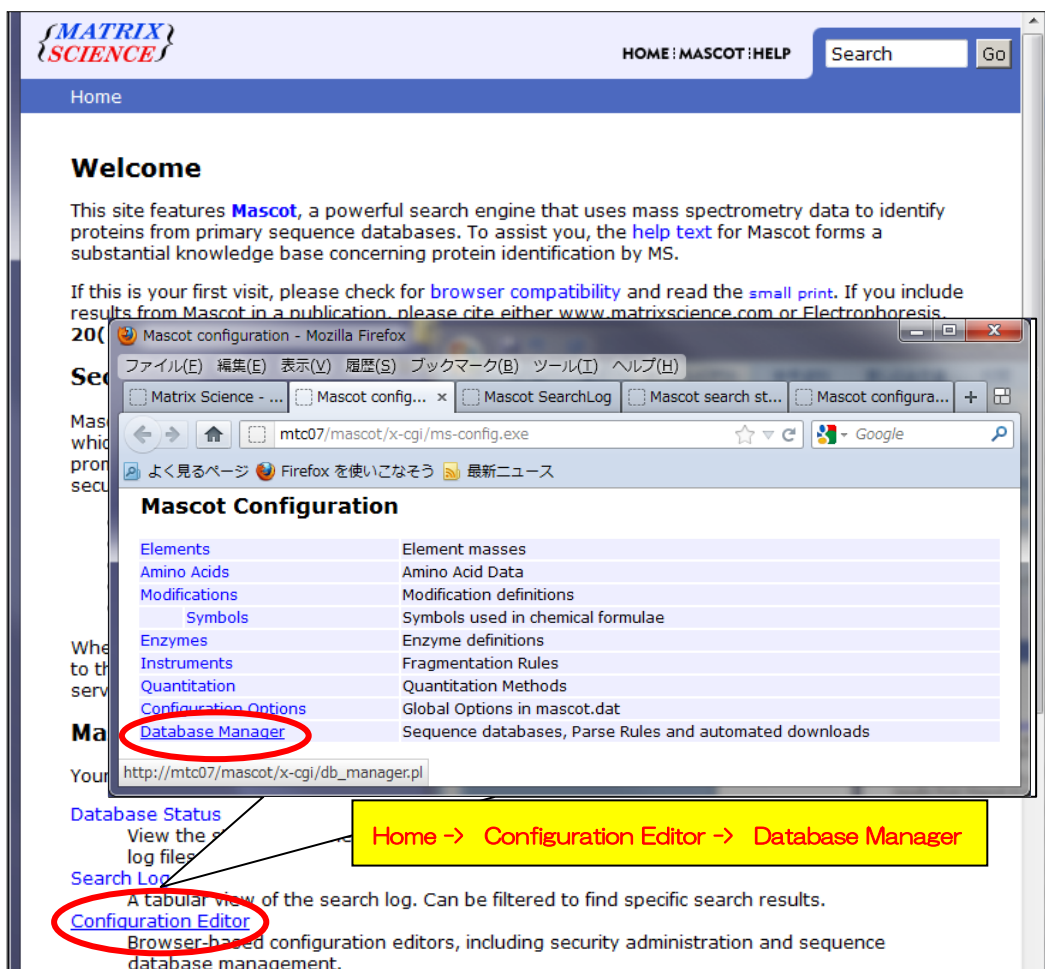
□ Database Manager 初期設定	P.3		
▼ Mascot ver. 2.3 からのアップデート時	P.3		
▼ Mascot ver. 2.4 新規インストール時	P.8		
□ Section 別 設定内容 :: Settings Section	P.11		
▼ MASCOT Server のInternet 接続の可否	P.13		
▼ MASCOT Server の 外部公開プログラムへの接続の可否	P.13		
▼ プロキシ設定	P.14		
▼ 配列データベース設置場所の変更	P.16		
□ Section 別 設定内容 :: Database Manager Section	P.17		
▼ 使用データベースのON/OFF, データベースファイルの取得 (更新)	P.18		
▼ データベースファイルの定期的な自動取得・更新設定	P.21		
▼ 新規データベースの追加	P.23		
- predefined database definition	P.24	- template	P.28
- copy	P.34	-custom	P.36
□ Section 別 設定内容 :: Tasks Section	P.45		
▼ データベース更新進捗の確認	P.45		

● Database Manager 初期設定

□ ver.2.3からのバージョンアップ時に行う初期設定

Database Manager 機能の実装に伴い、ver. 2.3 から ver. 2.4 へのアップグレード時に改めてデータベースの設定移行作業を行う必要があります。

データベースの設定移行作業を行う画面へアクセスするためには MASCOT のライセンス登録後 Web ブラウザで Home → Configuration Editor → **Database Manager** を開きます。



バージョンアップ後最初に Database Manager を開くと、次頁図のように初期設定画面が現れます。ver. 2.4 では、公開データベースの設定に関して Parse Rule をはじめとする各種項目が XML ファイル **configuration.xml** (ローカルサーバーの config フォルダ内) にまとめられており、弊社公開サイト www.matrixscience.com においてもそのファイルを公開しています。

Database Manager の最初の設定では、各データベースの設定に関して、これまで利用していた各設定項目をそのまま継続して利用する(Keep as Custom)か、それとも configuration.xml ファイルの設定を適用し、以降も定期的に変更がないかチェックし続ける(Synchronise)かを選択します。**通常は最適設定が自動的に選ばれており変更する必要がない場合が多く、そのままパラメータを変更することなく「Import」を押します。**

Mascot Database Manager Configuration Import

Note:
Please disable other programs and scripts that configure or update sequence databases before continuing.

Database Manager can keep your sequence database configuration up to date and schedule the download of new files. If you choose to use Database Manager, you should not edit mascot.dat in future, because your changes will be lost. If you do not wish to use Database Manager now, please [return to the Configuration Editor](#).

Database Manager has two kinds of database definitions:

1. *Predefined database definition:* Automatic updates from the Matrix Science website.
2. *Custom database definition:* Manually maintained.

You can choose to keep the existing definition (Keep as Custom) or update it (Synchronise). Synchronisation means the predefined definitions (FASTA, reference, taxonomy) will be updated.

If you do not want to update the files now, choose Synchronise Custom Definitions from the Database menu.

Existing	Predefined	Match
SwissProt	SwissProt	perfect Synchronise
NCBIInr	NCBIInr	perfect Synchronise
IPI_human	IPI_human	perfect Synchronise
EST_human	EST_human	perfect Synchronise
cRAP	(none)	none Keep as Custom
OriginalDB	(none)	none Synchronise
		none Keep as Custom

Import database configuration by choosing Import.

【選択項目の目安】
 ●公開データベースで mascot デフォルト設定準備されている
 → Synchronise
 ●カスタムデータベース、公開DBだが抜き出しルール変更時など
 → Keep as Custom

通常はパラメータを変更することなく「Import」

設定の目安ですが、SwissProtやNCBIInrなど、弊社でもデフォルト設定を準備している公開データベースの場合「Synchronise」を選択し、弊社でデフォルト設定を準備していない他の公開データベースや、完全にオリジナルなカスタムデータベースの場合は「Keep as Custom」が選択する事をお勧めします。

Existing	Predefined	Match	
SwissProt	SwissProt	perfect	Synchronise ▼
NCBIInr	NCBIInr	perfect	Synchronise ▼
IPI_human	IPI_human	perfect	Synchronise ▼
EST_human	EST_human	perfect	Synchronise ▼
cRAP	(none)	none	Keep as Custom
OriginalDB	(none)	none	Synchronise
			Keep as Custom ▼

Import database configuration by choosing Import.

Import

Database Manager 画面:データベースと選択項目について

表内の各項目の意味は以下の通りです。

項目名	説明
Existing	これまでのバージョンで利用していたデータベースの名称
Predefined	configuration.xml 内で定義されている設定の名称
Match	Predefined の設定内容と、既存設定との一致度。 Perfect なら問題ない。
Synchronise or Keep as Custom	Synchronise …configuration.xml の定義内容と入れ替え Keep as Custom … これまでの定義内容をそのまま利用

すべて定義後、画面下の「**Import**」ボタンを押します。Importボタンを押すと画面が切り替わります。左側の「Database Manager」->「Databases」に該当する項目の画面です（下図）。

Database Manager

Databases (6)

Parse rules (15)

Tasks (4)

Settings

Tasks (4)

New database

Enable predefined definition

Create new

Synchronise custom definitions

Databases

Name	Status	New files	Tasks
IPI_human *	In use	Update	Deactivate 1 task in queue
NCBIInr *	In use	Update	Deactivate 1 task in queue
OriginalDB	In use	(no URLs)	Deactivate
SwissProt *	In use	Update	Deactivate 1 task in queue
cRAP	Offline	(no URLs)	Activate
EST_human *	Offline	Download	No local FASTA file 1 task in queue

*) Entries marked with an asterisk are predefined definitions. Predefined definitions file was last updated Tue May 8 17:19:46 2012.

Full database status information is available on [the database status page](#).

[Refresh](#)

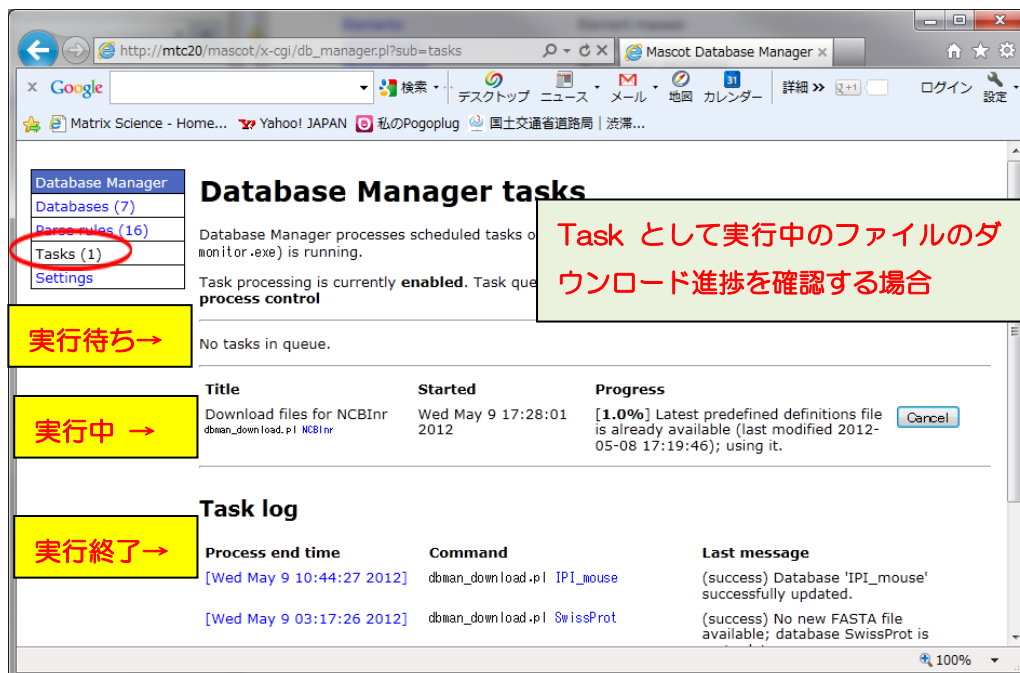
Import ボタンクリック後自動的に表示される「Databases」画面

Importボタンを押すと、既存のファイルより公開されているデータベースが新しいかどうかを判定した上で、自動的にデータベースの更新を試みます。この段階で既に既存データベースに対するMASCOT検索ができる状態になっており、**データベースの更新を放置してご利用頂いても結構です。更新状況を確認したい場合、下記に記す方法で行ってください。**

【データベースファイルの自動取得とデータベース更新の進捗を確認する方法】

ver.2.4 よりデータベース更新プログラムは既存のバージョンで動作していたプログラムとは異なる仕組みで動き、これまでのプログラムで使用していたプログラム本体と設定は全く使用されなくなります。

新たに実行された各データベースの更新プロセスは「Task」として扱われます。Taskの進捗度合いは下図のように「Tasks」セクションにて確認する事ができます。左側の「Database Manager」->「**Tasks**」をクリックすると、現在実行・実行待ちとなっているTaskが一覧で表示されます。Taskは「実行待ち」「実行中」「実行終了」の各セクションに分かれて表示されています。



Database Manager : Tasks 画面

ファイルダウンロードが終了後、プログラム ms-monitor.exeがダウンロードしたファイルをMASCOT で検索できるようファイル変換を行います。変換状況は、**Database Status** (Home -> Database Status)で各データベースの更新状況が確認できます。

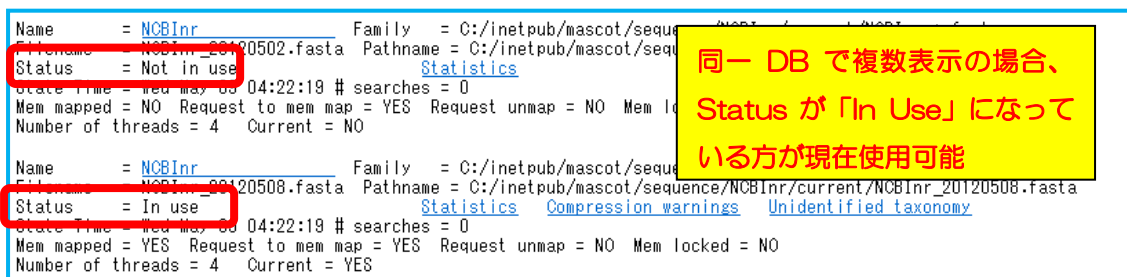
各データベースの「Filename」項目に、(Database名)_(バージョンまたは日付).fasta と表示されています。後部の(バージョンまたは日付)でデータベースの新しさを管理してください。



Database Status 画面: 構築中のデータベース

図例のNCBIInrのように、2つの日付のデータベースが表示されているケースがあります。

「Filename」項目と「Status」項目を見比べてください。例図では、日付が20120502 のStatus項目は「Not in use」となっています。一方より新しい20120508の方は「In use」となっています。これは20120502の方は既に使用されておらず、新しい20120508の方が現在使用可能となっている事を表しています。



【ネットワークご利用時Proxyサーバーをご利用のお客様のみ:Proxyサーバーの設定】

自動更新プログラムの設定には、Proxy サーバーの設定が必要です。詳細は Section 別設定内容:: Settings Section の「プロキシ設定」(P. 14) をご覧ください。

□ Database Manager 最初の設定 【ver. 2.4 新規インストール時】

インストール後、使用するデータベースについて、初期登録されている **SwissProt** の設定を確認し、**MASCOT** 側で設定が定義されているデータベースから使用したいものを選ぶ操作が必要となります。データベースの環境設定を行う画面へアクセスするためには MASCOT のライセンス登録後 Web ブラウザで Home → Configuration Editor → **Database Manager** を開きます。

The screenshot shows the Mascot Configuration page in a Mozilla Firefox browser. The page title is "Mascot configuration - Mozilla Firefox". The browser address bar shows "mtc07/mascot/x-cgi/ms-config.exe". The page content includes a "Welcome" message and a list of configuration options. The "Database Manager" link is circled in red. A yellow box highlights the navigation path: "Home -> Configuration Editor -> Database Manager". The "Database Manager" link is also circled in red.

Elements	Element masses
Amino Acids	Amino Acid Data
Modifications	Modification definitions
Symbols	Symbols used in chemical formulae
Enzymes	Enzyme definitions
Instruments	Fragmentation Rules
Quantitation	Quantitation Methods
Configuration Options	Global Options in mascot.dat
Database Manager	Sequence databases, Parse Rules and automated downloads

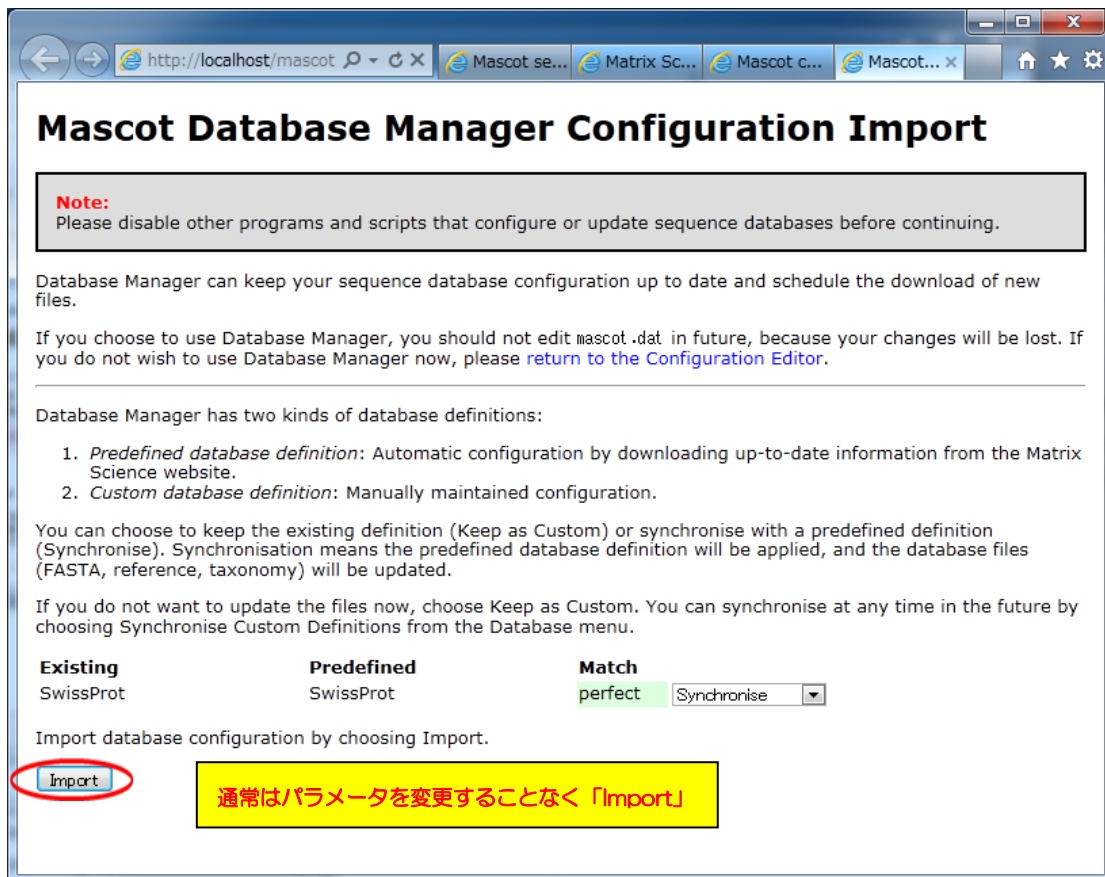
Database Status
View the status of the sequence databases. Links to database statistics, search status, log files.

Search Log
A tabular view of the search results.

Configuration Editor
Browser-based configuration editors, including security administration and sequence database management.

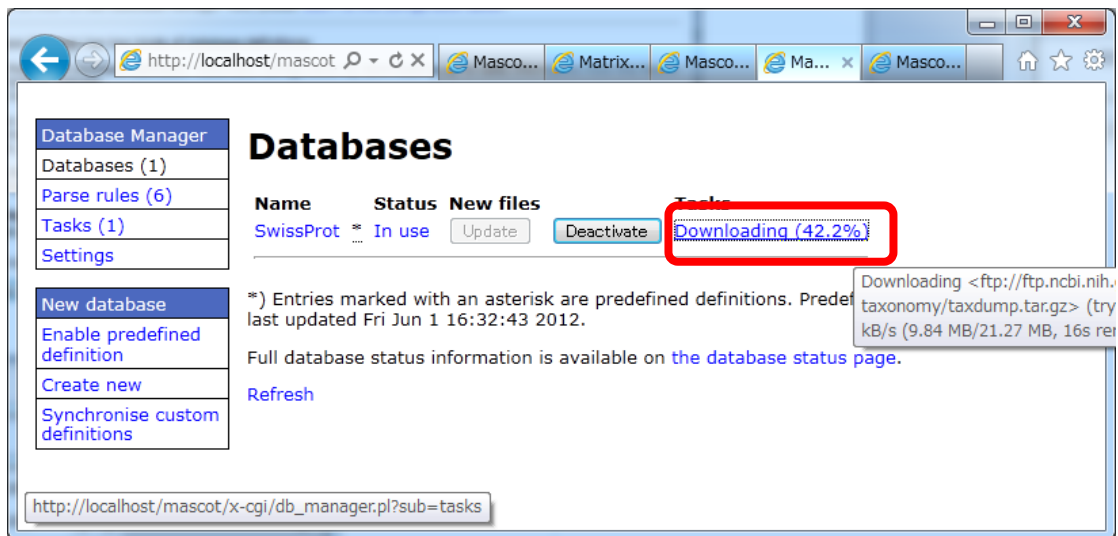
Home -> Configuration Editor -> Database Manager

インストール後最初にDatabase Managerを開くと、以下のように初期設定画面が現れます。ver. 2.4 では、公開データベースの設定に関してParse Rule をはじめとする各種項目がXML ファイル **configuration.xml** (ローカルサーバーのconfigフォルダ内)にまとめられており、弊社公開サイト www.matrixscience.com においてもそのファイルを公開しています。Database Manager の最初の設定では、各データベースの設定に関して、これまで利用していた各設定項目をそのまま継続して利用する (**Keep as Custom**) か、それとも configuration.xml ファイルの設定を適用し、以降も定期的に変更がないかチェックし続ける (**Synchronise**)かを選択します。**SwissProtに関しては最適設定が自動的に選ばれているため、そのままパラメータを変更することなく「Import」を押します。**(但しID部分を変更したい場合はその限りではありません)



Database Manager 初期設定画面

確認後、画面下の「**Import**」ボタンを押します。Importボタンを押すと画面が切り替わります。左側の「Database Manager」->「Databases」に該当する項目の画面です(次頁図)。



Importボタンを押すと、**既存のファイル(DVD内)よりインターネットに公開されているデータベースの方が新しいかどうかを判定し、自動的にデータベースの更新を試みます。**この段階で既に既存データベースに対するMASCOT検索ができる状態になっていますので、**データベースの更新は放置して以降の動作確認のステップへ移行してください。**更新状況は後述する確認方法で進捗をチェックしてください。

【ネットワークご利用時Proxyサーバーをご利用のお客様のみ:Proxyサーバーの設定】

自動更新プログラムの設定には、Proxy サーバーの設定が必要です。詳細は **Section 別設定内容:: Settings Section の「プロキシ設定」(P. 14)** をご覧ください。

【SwissProt以外のデータベースファイルを使用可能にする方法】

DVD 内に準備されている検索用データベースは SwissProt のみですが、MASCOT 用のデータベースとしてはその他にも公開されているデータベースや、ご自身で作成されたデータベースを使用する事ができます。

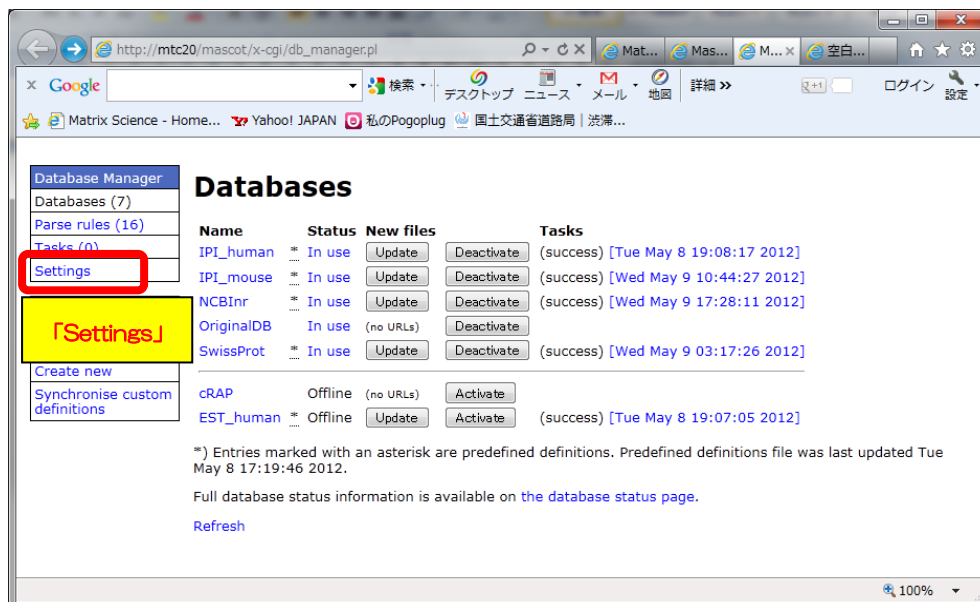
データベースを追加・使用する方法については、**Section 別設定内容:: Database Manager Section(P. 22)**をご覧ください。

● Section 別設定内容 :: Settings Section

Settings 項目では大きく分けて以下の4項目の設定変更が可能です。

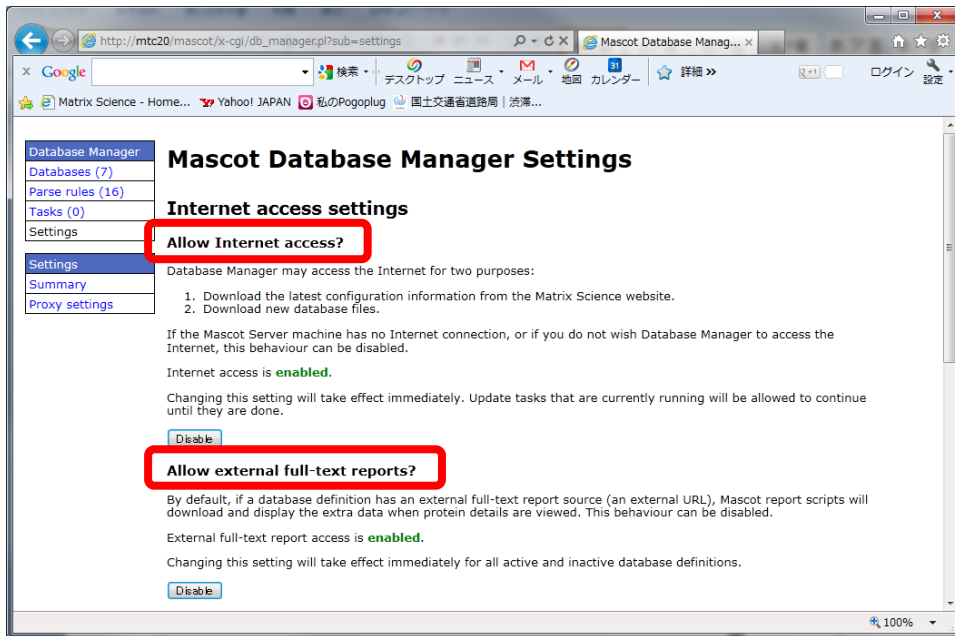
- MASCOT Server のInternet 接続の可否
- MASCOT Server の 外部プログラムへの接続の可否
- プロキシ設定
- 配列データベースの格納位置の設定

Settings セクション画面へは、画面左側の「Database Manager」 - 「Settings」をクリックします。



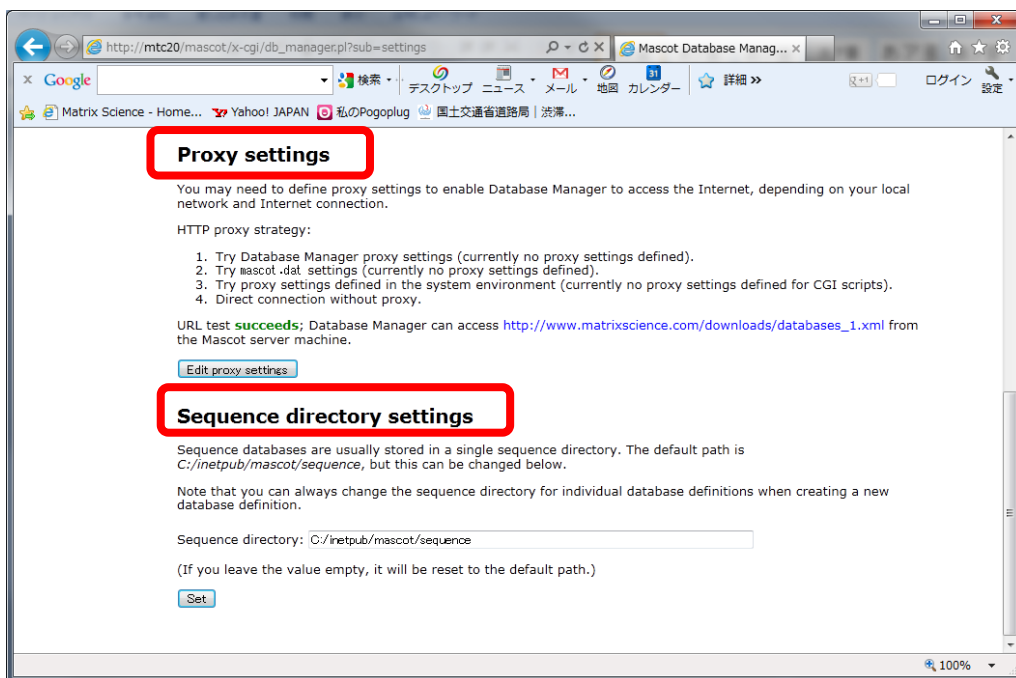
Settingsの画面は、上から順に、

「MASCOT Server のInternet 接続の可否」「MASCOT Server の 外部プログラムへの接続の可否」、



Settings 画面上部

並びに「プロキシ設定」「配列データベース格納位置の設定」に関する項目やリンクのボタンが並んでいます。



Settings 画面下部

□ MASCOTサーバーのインターネットへの接続許可【Internet access settings】

MASCOT では 主にデータベースファイル取得や、databaseの定義ファイルのチェックのためインターネットに接続します。インターネットへの接続を希望しない場合、【Internet access settings】項目で設定を変更し切り替える事ができます。

Internet access settings

Allow Internet access?

Database Manager may access the Internet for two purposes:

1. Download the latest configuration information from the Matrix Science website.
2. Download new database files.

If the Mascot Server machine has no Internet connection, or if you do not wish Database Manager to access the Internet, this behaviour can be disabled.

Internet access is **enabled.** ←現在の設定内容

Changing this setting will take effect immediately. Update tasks that are currently running will be allowed to continue until they are done.

←設定の切り替え

□ MASCOT Server の 外部プログラムへの接続の可否【Allow external full-text reports?】

MASCOT の結果画面の中で、タンパク質の詳細情報を表示する箇所があります。詳細情報はローカルコンピュータにダウンロードされたファイルから取得する場合と、外部のサーバーから取得する場合があります。詳細情報を得るためにヒットタンパク質の情報を外部サーバーへ投げる事を好まない場合、【Allow external full-text reports?】項目で設定を変更する事ができます。

Allow external full-text reports?

By default, if a database definition has an external full-text report source (an external URL), Mascot report scripts will download and display the extra data when protein details are viewed. This behaviour can be disabled.

External full-text report access is **enabled.** ←現在の設定内容

Changing this setting will take effect immediately for all active and inactive database definitions.

←設定の切り替え

□ Proxyサーバー設定【Proxy settings】

前述の通り、MASCOTではデータベースファイル取得や、databaseの定義ファイルのチェックのためインターネットに接続しますが、使用しているネットワーク環境によってはProxyサーバー経由でないとインターネットへ接続できないことがあります。【Proxy settings】項目ではProxyに関する設定を行います。

設定画面には、【Proxy settings】項目内の、「Edit proxy settings」を選択します。

Proxy settings

You may need to define proxy settings to enable Database Manager to access the Internet, depending on your local network and Internet connection.

HTTP proxy strategy:

1. Try Database Manager proxy settings (currently no proxy settings defined).
2. Try mascot.dat settings (currently no proxy settings defined).
3. Try proxy settings defined in the system environment (currently no proxy settings defined for CGI scripts).
4. Direct connection without proxy.

URL test **succeeds**; Database Manager can access http://www.matrixscience.com/downloads/databases_1.xml from the Mascot server machine.

プロキシ設定は「Edit proxy settings」

プロキシサーバーの設定画面へ移動します（次頁図）。画面が開く際に既に接続テストが行われており、設定が適切でファイル

http://www.matrixscience.com/downloads/databases_1.xml

にアクセスできると「**Succeeds**」と表示します。

設定画面では Proxy サーバーのアクセス先、ポートに加え、必要があればユーザー名やパスワードを入力する箇所があります。また、HTTPだけでなくFTP についても設定する箇所があります。

The screenshot shows the 'Mascot Database Manager Settings' page. The left sidebar contains a navigation menu with 'Database Manager', 'Databases (8)', 'Parse rules (16)', 'Tasks (0)', 'Settings', 'Settings', 'Summary', and 'Proxy settings'. The main content area is titled 'Proxy settings' and includes a success message: 'Proxy test **succeeds** Database Manager can access http://www.matrixscience.com/downloads/databases_1.xml from the Mascot server machine.' Below this, there is explanatory text: 'HTTP is used for downloading configuration data from the Matrix Science website. HTTP and FTP are used for downloading sequence database files from specified locations. Native FTP proxy servers are not supported.' The form contains four input fields: 'Host (HTTP proxy) (?)', 'Port (HTTP proxy) (?)', 'User (HTTP proxy) (?)', and 'Password (HTTP proxy) (?)'. A 'Test and save' button is highlighted with a red box, and a 'Back' button is also visible. Several yellow callout boxes provide instructions: 'Proxy サーバアドレス。例: http://proxy.mascot.jp/proxy', 'Proxy ポート。例: 8080', '認証ユーザー名とパスワード。使用していない場合、空欄で', and '設定後、適用と同時に「Test」実施'.

Proxy Settings 画面

□ Sequence フォルダデフォルト設置箇所設定【Sequence directory settings】

MASCOT で使用するデータベースのファイルはファイル容量が大きく、しばしばファイルの置き場所に問題が生じる事があります。ファイルの置き場所を変更するには、【Sequence directory settings】の項目を書き換え、設定を適用するため「Set」ボタンを押してください(下図)。

Sequence directory settings

Sequence databases are usually stored in a single sequence directory. The default path is *C:/inetpub/mascot/sequence*, but this can be changed below.

Note that you can always change the sequence directory for individual database definitions when creating a new database definition.

Sequence directory: ←現在の設定内容、必要に応じて書き換える

(If you leave the value empty, it will be reset to the default path.)

←設定更新の適用

● Section 別 設定内容 :: Database Manager Section

databases セクションでは、以下の内容の設定変更を行うことができます。

- ▼ 使用データベースのON/OFF
- ▼ データベースファイルの取得
- ▼ データベースファイルの定期的な自動取得・更新設定
- ▼ 新規データベースの追加

【新規データベース追加作業におけるこれまでのバージョンとの変更点】

ver.2.3までは、新規データベースの構築時に必要な、FASTAファイル内先頭行から特定のルールでID部分を抜き出す「抜き出しルール」の記述がコンピューターに慣れていない方にとって難しく、この件で多くのお問い合わせがありました。また、公開データベースのファイルフォーマットやファイルの保存場所が不定期に変わる事があり、トラブルの要因の1つでした。

新バージョンではこれらの点に関する改善が行われました。まず、IDの抜き出しルール選択について、選択時にMASCOT側が候補となるルールを提示して、ユーザーが最適な項目を選ぶだけで良いようにしました。また、弊社が管理するサーバー上でデータベースの設定ファイルを公開し、ユーザー側のMASCOTが定期的に弊社サーバーの設定ファイルの中身をチェックする仕組みを採用しました。これによりフォーマットやファイルの置き場所が変わり設定の変更の必要が生じた場合、適切に書き換えられた設定内容をユーザー側のMASCOTにも自動的に対応させるようになりました。さらに、これまでWindowsのTask機能を使って実施していたデータベースの自動更新を、このセクションの中で設定・実行できるようにいたしました。

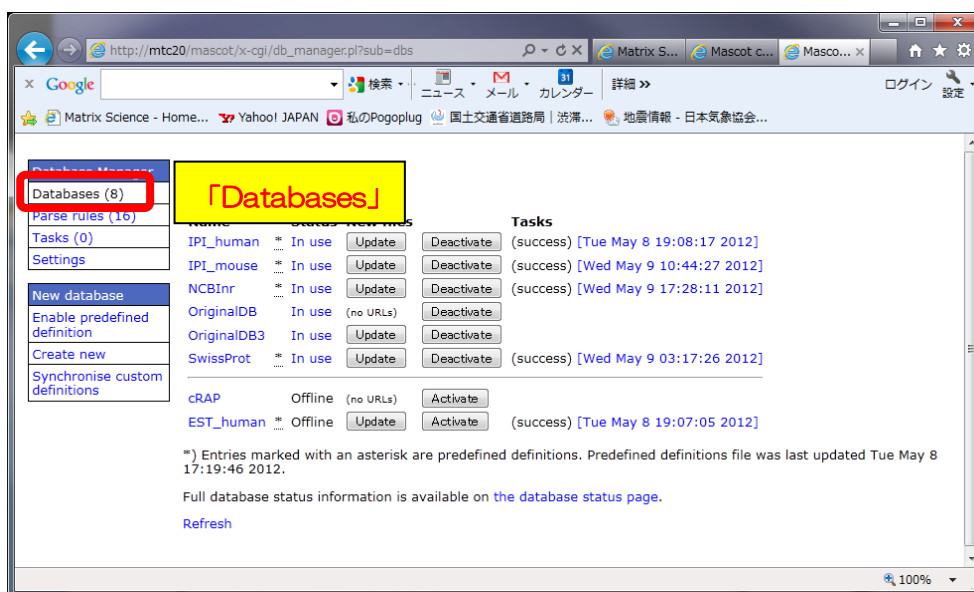
【データベース追加作業の変更点・まとめ】

- MASCOTが自動的に最適な抜き出しルールの候補を挙げてくれる。ユーザーは候補から選択するだけでよい。
- データベースの抜き出しルール・ダウンロード先に変更がないか自動的・定期的に確認をし、必要に応じて自動的に変更させる事ができる。
- データベース自動更新の日時・頻度を設定できる

以下、Databases Manager Sectionで行う事のできる設定について、小セクションごとにまとめています。

□ 使用データベースのON/OFF、データベースファイルの取得（更新） 【Databases section】

Database Manager の最初の画面、または左フレームの一番上、「Database Manager」 - 「**Databases**」をクリックすると現れる画面で、現在ピックアップ中のデータベースの一覧が表示されます。



Databases セクションでは、使用するデータベースの選択やデータベースファイルの更新を行う事ができます。

使用状況

Databases

DB更新

Active / Deactive 切替

Name	Status	New files	Tasks
IPI_human	* In use	Update	Deactivate (success) [Tue May 8 19:08:17 2012]
IPI_mouse	* In use	Update	Deactivate (success) [Wed May 9 10:44:27 2012]
NCBIInr	* In use	Update	Deactivate (success) [Wed May 9 17:28:11 2012]
OriginalDB	In use	(no URLs)	Deactivate
OriginalDB3	In use	Update	Deactivate
SwissProt	* In use	Update	Deactivate (success) [Wed May 9 03:17:26 2012]
cRAP	Offline	(no URLs)	Activate
EST_human	* Offline	Update	Activate (success) [Tue May 8 19:07:05 2012]

*) Entries marked with an asterisk are predefined databases. The last update of the database file was last updated Tue May 8 17:19:46 2012.

Full database status information is available on [the database status page](#).

[Refresh](#)

線上部は "Active"

線下部は "Deactive"

***は MASCOT 側が設定を準備している DB "Predefined"**

Name	Status
IPI_human	* In use
IPI_mouse	* In use
NCBIInr	* In use
OriginalDB	In use
OriginalDB3	In use
SwissProt	* In use

各データベースの名称部分のハイパーリンクをクリックすると、該当データベースの設定やバージョン、ファイル取得の状況についてより詳しく見る事ができます（次頁図）。

Mascot Database Manager: SwissProt - Mozilla Firefox

ファイル(E) 編集(E) 表示(V) 履歴(S) ブックマーク(B) ツール(I) ヘルプ(H)

Matrix Science - Home x Mascot configuration x Mascot Database Ma... x

mtc20/mascot/x-cgi/db_manager.pl?sub=d

よく見るページ Firefox を使ってみよう 最新ニュース

mtc20/mascot/x-cgi/db_ma...r.pl?sub=dbs%3ASwissProt

Database: SwissProt

Database Manager

- Databases (8)
- Parse rules (16)
- Tasks (0)
- Settings

SwissProt

- Copy
- Delete

Name
SwissProt

設定ファイル configuration.xml 内での名称

Using predefined definition
SwissProt_ID

Configuration last updated Tue May 8 17:19:46 2012.

Database type
Amino acid (protein database)

Database directory
C:/inetpub/mascot/sequence/SwissProt/current

Filename pattern
SwissProt_*.fasta

Database FASTA ファイルが置かれている場所

Matching files in directory

- SwissProt_2012_03.fasta (243 MB) (current)
- SwissProt_2012_03.dat (2.4 GB) (current)

Database status

In use

Deactivate

Active/Inactive 切り替え

Scheduled updates
(no schedules defined)

Edit schedule

自動更新スケジュール編集

Most recent finished task
[Wed May 9 03:17:26 2012] (success) No new FASTA file available; database SwissProt is up to date.

Update database now

データベースファイル更新

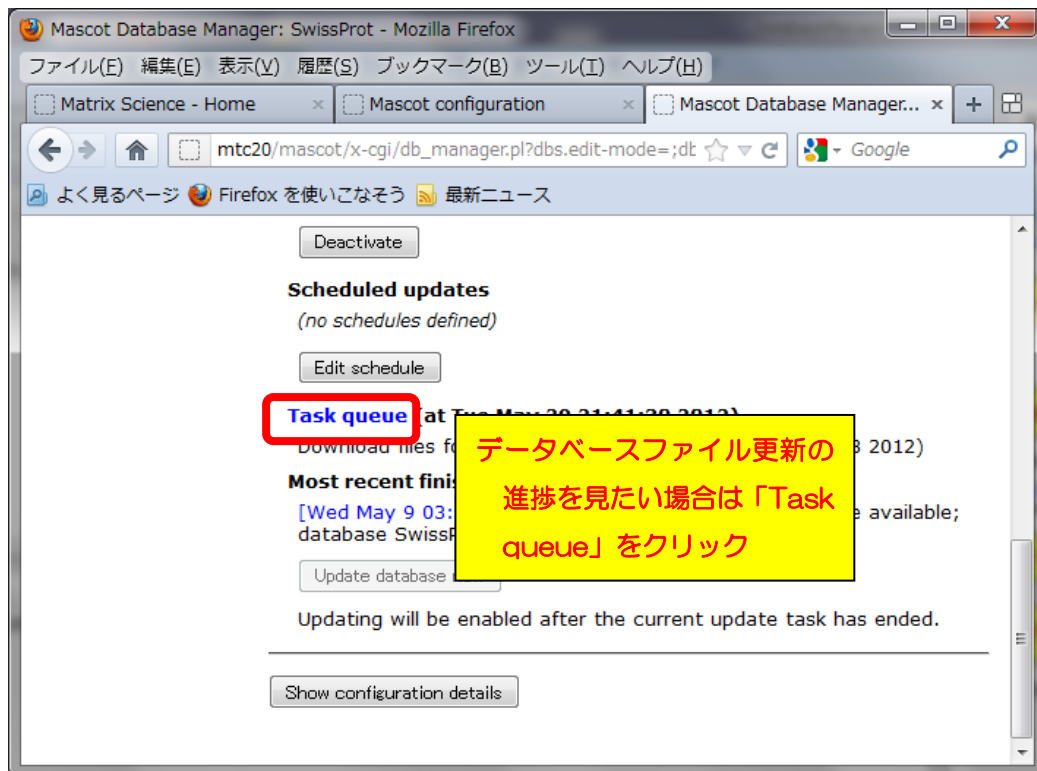
Show configuration details

さらに詳細な設定内容が見たい場合

Databases 個別データベース設定画面

各種ページにある「**Update**」または「**Update database now**」 ボタンを押すと、データベースファイルの更新がその場でスタートします。スタートすると「Update」のボタンが押せなくなり、Taskとして進捗度合いが画面に表示されます。

- * ネットワークにて Proxy サーバーをご利用されている場合、Database Manager ページで Proxy サーバーの設定が必要です。設定方法について詳しくは Section 別設定内容:: Settings Section の「プロキシ設定」(P.14) をご覧ください。



□ データベースファイルの定期的な自動取得・更新設定【**databases section**】

前述の、各データベースの設定確認画面にある「**Edit Schedule**」を押すと、該当データベースの取得を定期的に行うための設定画面が現れます。

設定は、毎日<Daily>、毎週<Weekly>、毎月<Monthly>の3種類があります（次頁図）

The image shows a sequence of three screenshots from a web browser displaying the Mascot Database Manager interface.

Top Screenshot: Databases List
 The browser address bar shows `http://mtc20/mascot/x-cgi/db_manager.pl?sub=dbs`. The page title is "Databases". A table lists several databases:

Name	Status	New files	Tasks
IPI_human	In use	[Update] [Deactivate]	(success) [Tue May 8 19:08:17 2012]
IPI_mouse	In use	[Update] [Deactivate]	(success) [Wed May 9 10:44:27 2012]
NCBINr	In use	[Update] [Deactivate]	(success) [Wed May 9 17:28:11 2012]
OriginalDB	In use	[Update] [Deactivate]	(success) [Wed May 9 03:17:26 2012]
SwissProt	In use	[Update] [Deactivate]	(success) [Wed May 9 03:17:26 2012]

Annotations: A yellow box labeled "「databases」" points to the table header. A red box highlights the "Databases (8)" link in the left sidebar. A yellow box labeled "database 選択" points to the "SwissProt" row.

Middle Screenshot: Edit Schedule
 The browser address bar shows `mtc20/mascot/x-cgi/db_manager.pl?dbs.edit-mode=;dt`. The page title is "Mascot Database Manager: SwissProt". The "Edit schedule" button is highlighted with a red box. A yellow box labeled "「Edit schedule」" points to the button.

Bottom Screenshot: Database configuration: SwissProt
 The browser address bar shows `mtc20/mascot/x-cgi/db_manager.pl?dbs.edit-mode=cron;dbs`. The page title is "Database configuration: SwissProt". The "Scheduled updates" section has radio buttons for "None", "Daily at", "Weekly on", and "Monthly on the". The "Save" button is highlighted with a red box. A yellow box labeled "定期更新のスケジュール設定後「Save」" points to the button.

Databases 個別データベース Edit Schedule 画面

□ 新規データベースの追加【databases section】

MASCOTの検索対象となるデータベースを追加する方法には、以下の4つの方法があります。

【新規データベースを追加する4つの方法と主な対象データベース】

1. すべての設定がひとまとめになったセット【predefined database definition】を、ファイル設置場所以外変更せずそのまま使用（データベースの名前も変更できない）

→ SwissProt, NCBItr, IPI の他、NCBI の EST, EMBL の EST や IPI, trembl などの利用を想定

2. 上記項目1の「predefined definition」を使用するが、名前のみ編集して利用【template 利用】

→ 1 でリストアップされたデータベースと種類は同じだが、既成のものと同名前を変えたり、(taxonomy 抜き出しなど)データベース側に何らかの選別処理を施した後に利用したい場合などを想定

3. 既に使用しているデータベースのフォーマットを使用し、名前のみ編集して利用【copy 利用】

→ predefined のリストからでなく、既に使用中のデータベースから選択。既存データベースのバージョン違いを作成する場合などを想定

4. 完全に自分で各設定を記入・選択【Custom】

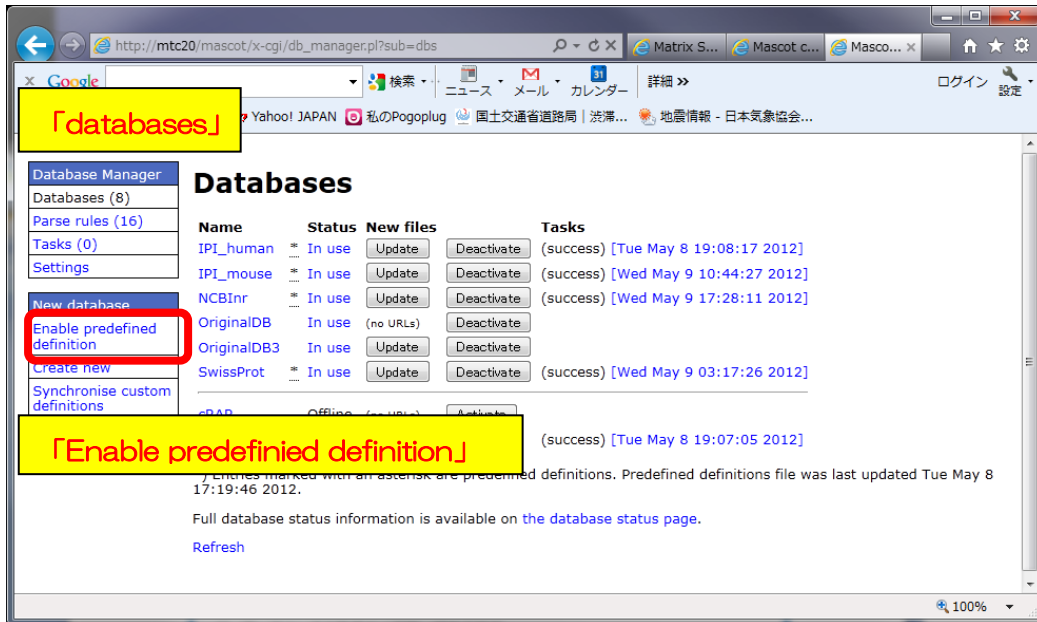
→ 1～3以外、自分で作成した FASTA ファイルなどの利用を想定

【1.すべての設定がひとまとめになった組み合わせ【predefined database definition】を使用】

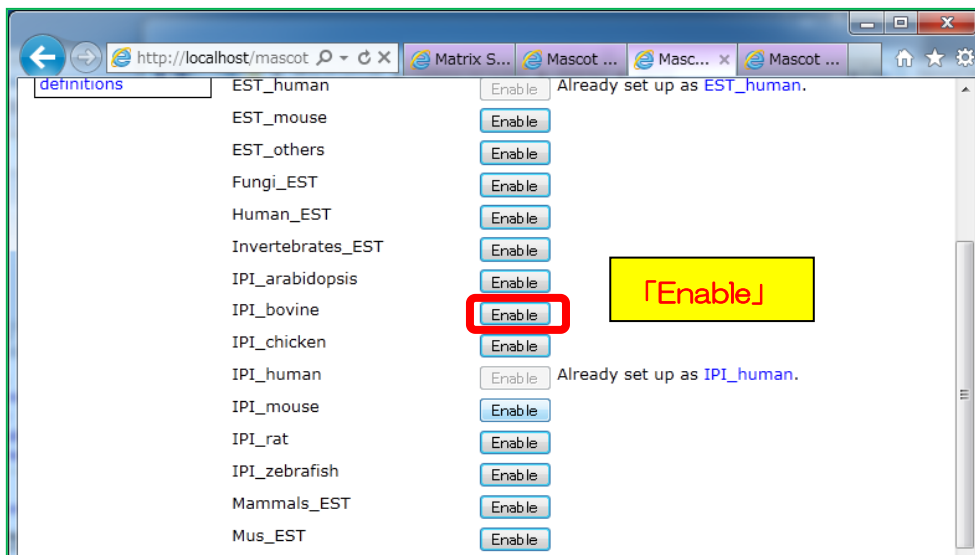
MASCOT 側で既に設定方法を準備しているいくつかのデータベースについて、面倒な設定を行わずそのまま利用する時には Predefinined database definition を利用します。predefined definition が準備されているデータベースは以下の通りです。

-
- **contaminants / cRAP** … コンタミ、または生体内に頻出するタンパク質。
 - **XXXX_EST** EMBL にあるEST データベース。詳細は http://www.matrixscience.com/help/seq_db_setup_EMBLEST.html をご覧ください。
<リスト> Environmental_EST , Fungi_EST, Invertebrates_EST, Mammals_EST, Mus_EST, Plants_EST, Prokaryotes_EST, Rodents_EST, Unclassified_EST, Vertebrates_EST
 - **EST_YYYY** NCBI にあるEST データベース。詳細は http://www.matrixscience.com/help/seq_db_setup_EST.html をご覧ください。
<リスト> EST_human, EST_mouse, EST_others
 - **IPI_ZZZZ** IPIデータベース。(2011年で更新が止まっています) **<リスト>**
IPI_arabidopsis, IPI_bovine, IPI_chicken , IPI_human, IPI_mouse, IPI_rat, IPI_zebrafish
 - **NCBIInr**
 - **Trembl**
 - **SwissProt**
 - **UniRef100**
-

「Predefined database definition」を使ってデータベースを追加する設定画面へアクセスするには、左フレームの「New database」 - 「**Enable predefined definition**」をクリックします（下図）。



使用したいデータベースを選んで、「**Enable**」ボタンを押します。

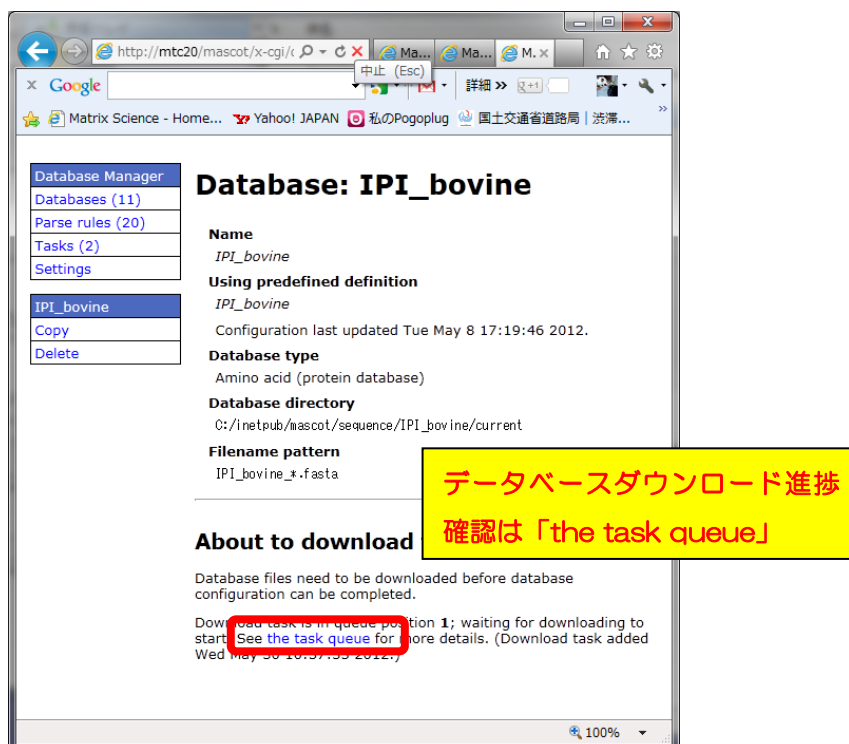


Enable predefined definition 画面

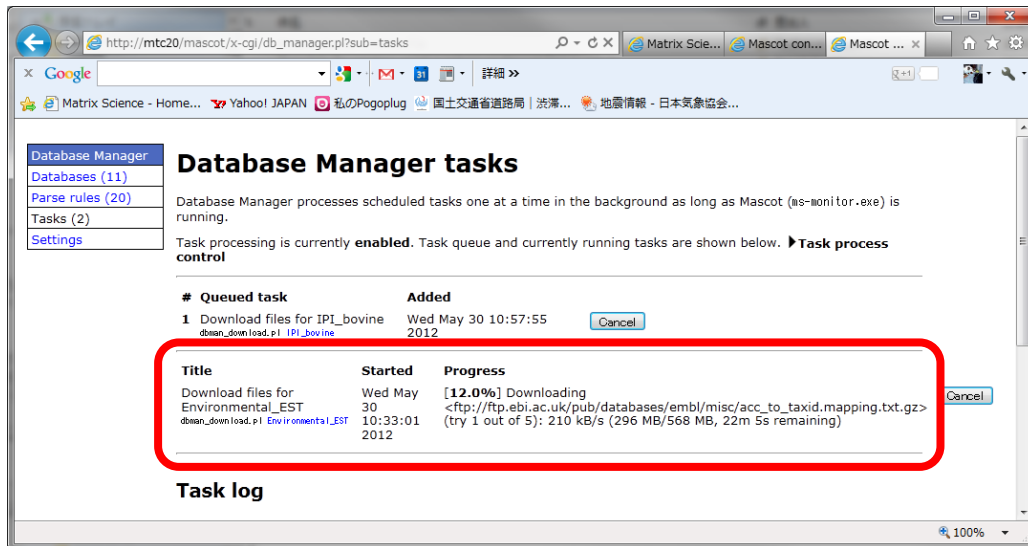
配列データベースファイルの置き場所を「Sequence directory」で確認し、必要があれば変更します。確認後、「**Create**」ボタンを押します。



データベースの定義に関するSummaryが表示されます。この段階で既に該当データベースがMASCOTにて使用対象のデータベースとして認識されると同時に、データベースファイルの取得を試みています。データベースの取得状況を確認したい場合は、ハイパーリンクの「**the task queue**」をクリックします。



データベースダウンロードの進捗状況が表示されます。



Database Status に表示され、Status が **“In Use”**となれば終了です(下図例は IPI_mouse)。



Database Status 画面

【2. predefined definitionを利用するが、名前のみ編集して利用【template利用】】

template 利用は、predefined definitionと全く同じ抜き出しルールなどのフォーマットを使い、名前だけ変えたい場合に利用する方法です。例えば2011年で更新が止まったIPIデータベースの代わりに、既存のIPIも含まれるuniprot データベースから特定生物種のみを抜き出したデータベースを作成して代わりに使う、といったケース等を想定しています。

例では、uniprot から生物種を抜き出して利用する方法を例に設定をご紹介します。

◇ uniprotからtaxonomy;humanを抜き出してFASTAファイルにする方法

データベース設定から少し話がそれますが、ユーザーからのニーズが高い、uniprotから特定生物種のエントリーだけを抜き出してFASTAファイルにする方法も併せてご紹介します。

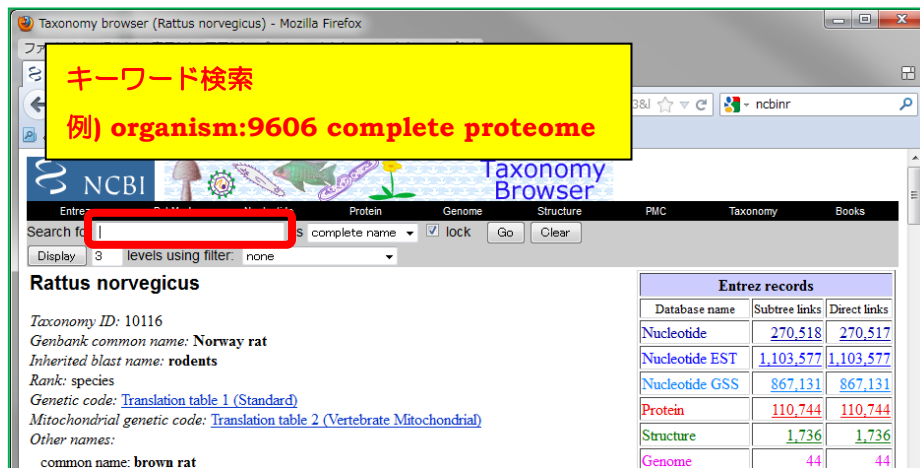
- ・ 抜き出したい生物種のTax ID を確認。Tax ID はMASCOTの生物種設定でも採用しているNCBIのtaxonomy ID です。調べたい場合は以下のサイトでフリーワード検索をして番号をお確かめください。

<http://www.ncbi.nlm.nih.gov/taxonomy>

Tax ID 例) Mus musculus : 10090, rattus norvegicus : 10116 など。

- ・ uniprot のサイト www.uniprot.orgへアクセス
- ・ キーワード **organism:TAX ID** と “**complete proteome**” で検索 (次頁図)

例) organism:9606 complete proteome



NCBI, Taonomy Browser <http://www.ncbi.nlm.nih.gov/taxonomy>

検索後、2012年5月30日時点で70663件のエントリーがヒットします。うちアノテーションがついているSwissProtのエントリーが 20240件です（下図）。FASTAファイルを入手するには、右上の「Download」をクリックします。

ダウンロードボタン

Download

Download

SwissProt データ数

トータルヒット数

Protein names	Gene names	Organism	Length
Calcineurin-like phosphoesterase domain-conta...	CPPED1 CSTP1	Homo sapiens (Human)	314
Erythrocyte band 7 integral membrane protein	STOM BAND7 EPB72	Homo sapiens (Human)	288
Minor histocompatibility protein HB-1	HMB1	Homo sapiens (Human)	41

revised (20,240) (UniProtKB/Swiss-Prot) or unrevised (50,427) (UniProtKB/TrEMBL) entries

Restrict term "complete" to keyword (70,663), protein name (1)

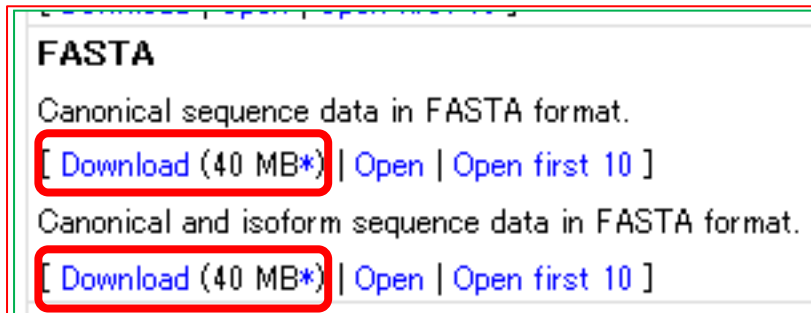
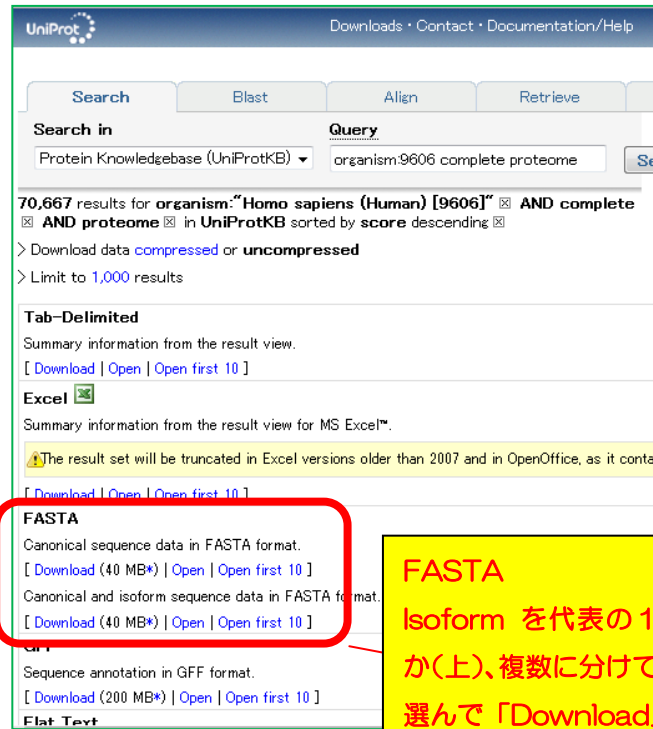
Restrict term "proteome" to keyword (70,663)

Expand search to "Homo sapiens (Human) [9606]" to include lower taxonomic ranks

Show only entries from a complete proteome set (70,663)

Show only entries from a reference proteome set (70,663)

FASTA の項目から、「Download」ボタンを押します。Isoform の扱いをどうするかによりクリックするハイパーリンクが異なるので注意してください。

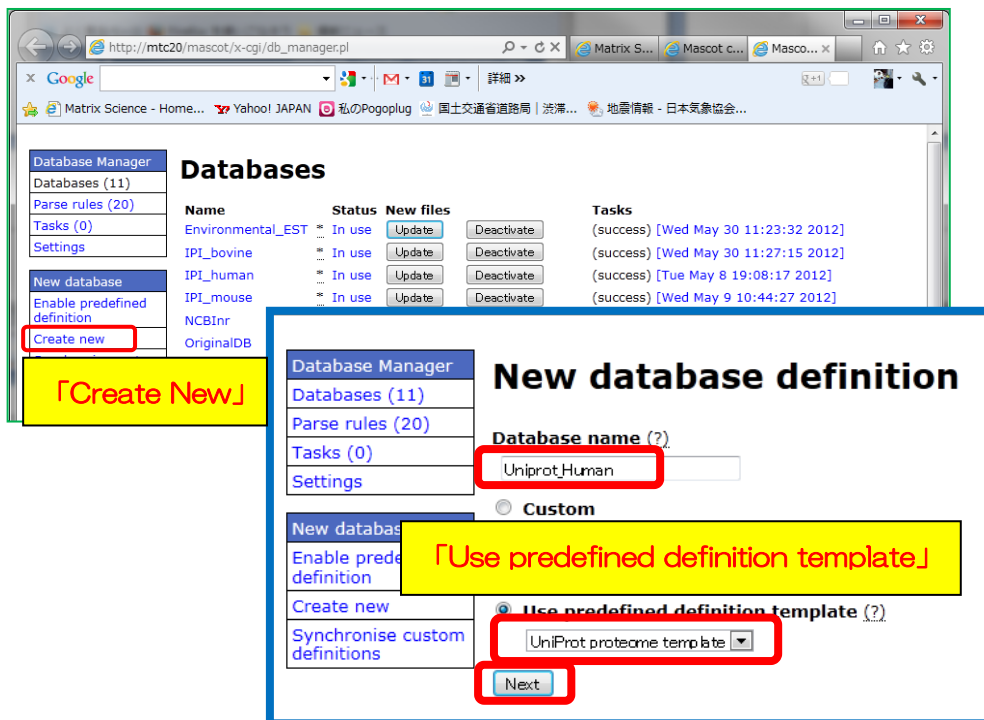


取得したファイルを MASCOT の検索対象データベースファイルとして利用します。

<uniprotからtaxonomy;humanを抜き出してFASTAファイルにする方法:おわり>

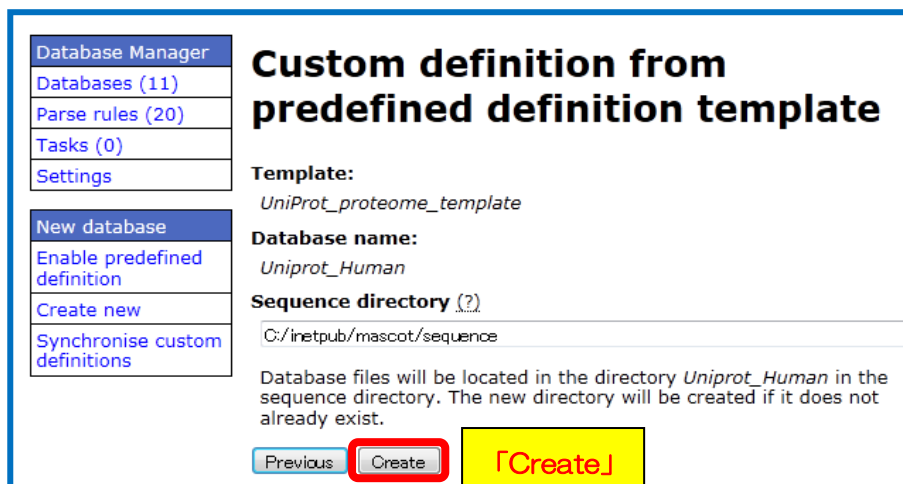
◇ **template** 利用新規データベース作成例：uniprot_human データベース作成

Database manager の、左フレーム「New database」 - 「**Create new**」を選びます。現れた画面で登録するデータベースの名称を「Database name」に入力し、「**Use predefined definition template**」で「Uniprot proteome template」を選びます。選択後、「Next」を押します。

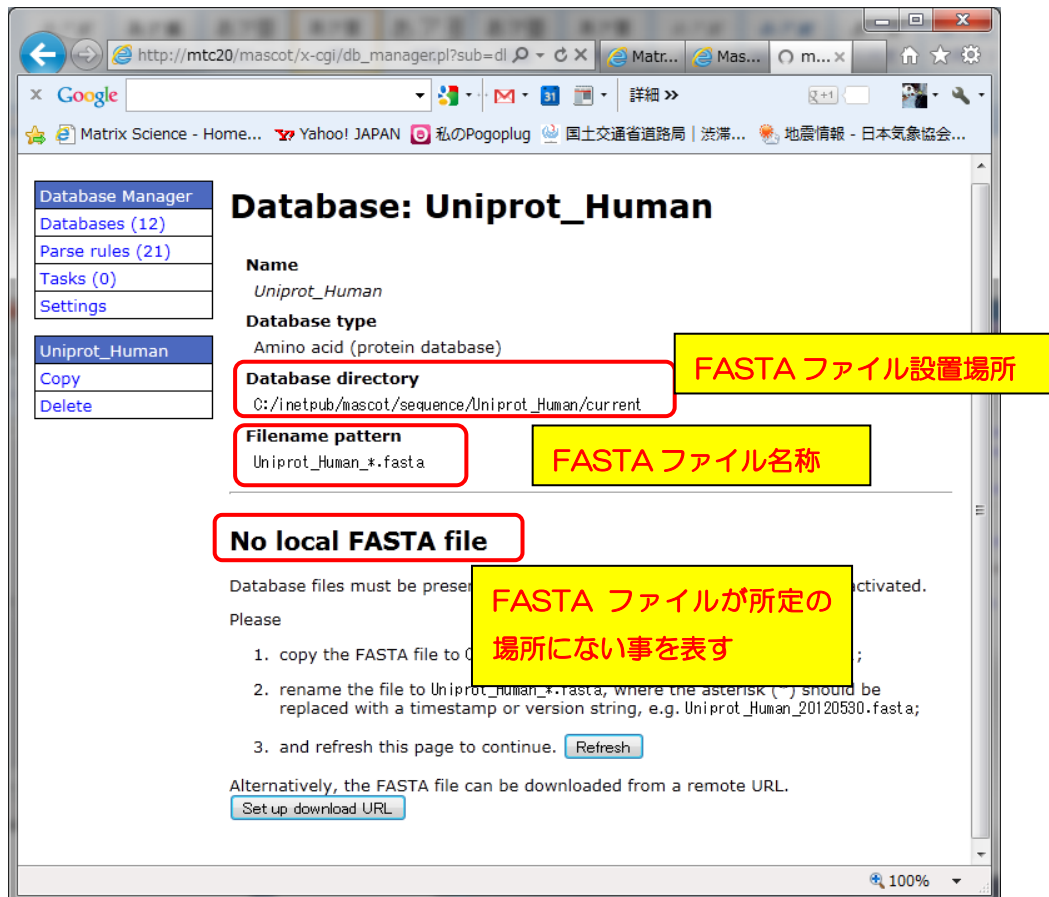


Databases -> Create New 画面

データベースファイルの置き場所を決め、「Create」を押します。



この時点ではFASTAファイルを置いておらず、Inactive なデータベースとしてMASCOT側に登録されます。使用するためには、別途取得したファイルを指定のフォルダに指定のファイル名に変えてからコピーする必要があります。



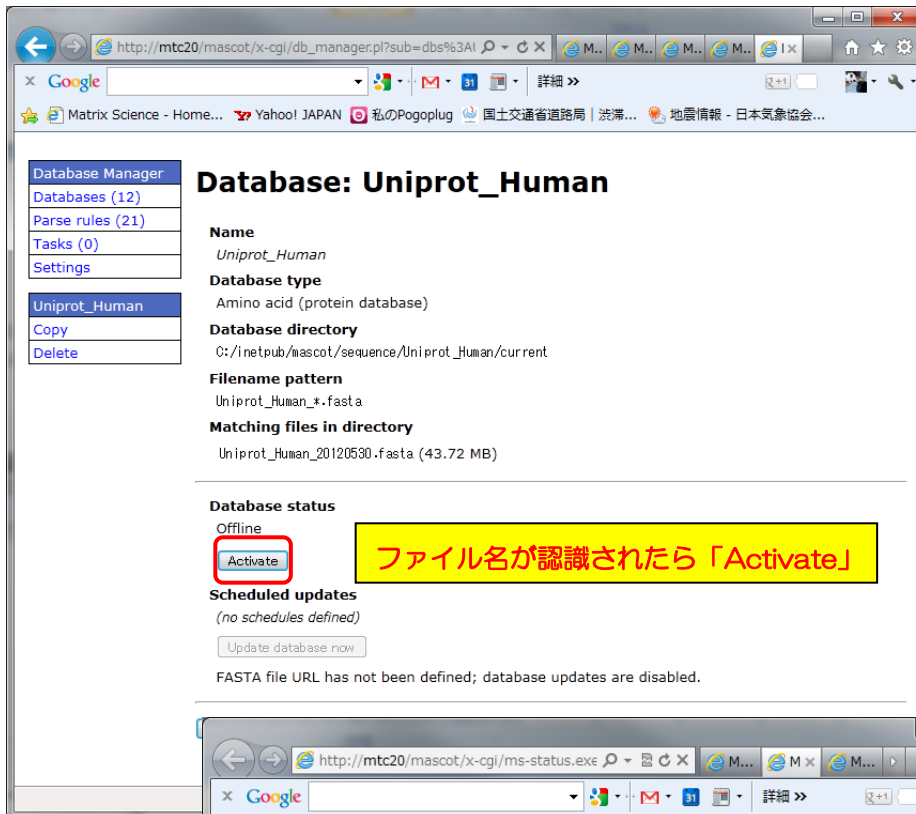
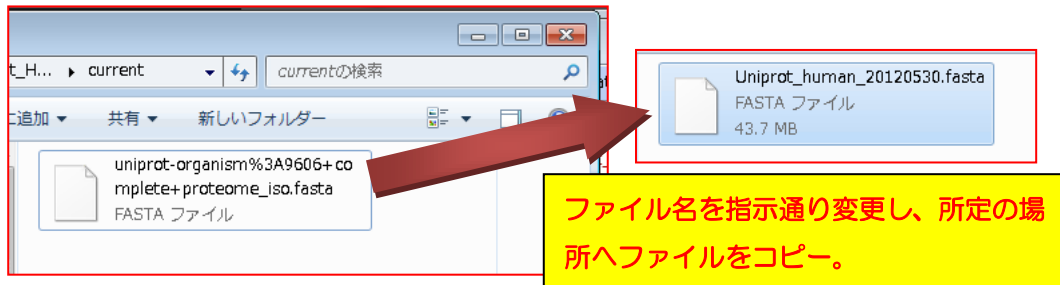
上図例の場合、「Database directory」で指定されているフォルダ、つまり

C:\inetpub\mascot\sequence\Uniprot_Human\current

に、ダウンロードしたファイルをコピーし、名称を「Filename pattern」で指定された形に書き換えます。但し名称中の「*」部分は任意の文字列、例えばダウンロード日などに変更します。

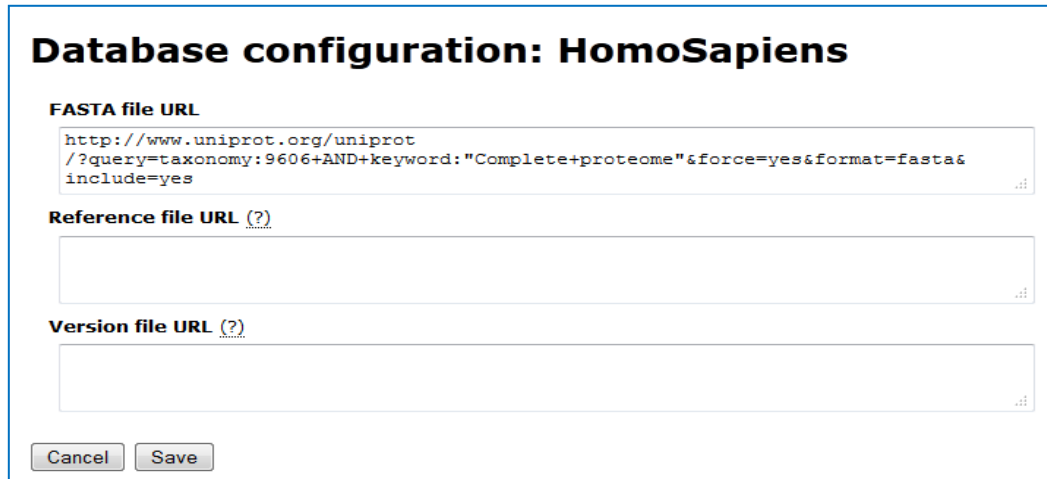
ファイル名変更例) Uniprot_Human_20120608.fasta (*はファイル名に書かない)

ファイル名を変更すると、FASTAファイルが自動的に認識され、「No local FASTA file」の表記がなくなります。「**Activate**」ボタンを押せば、データベース構築が開始します。



<参考>

File取得なども含めuniprotからtaxonomy:human のファイルを今後も定期的に自動取得する場合、FASTA file のURLに以下のように記入すればOKです（2012年6月現在）。但し今後も同じ操作でファイルが取得できるとは限りません。



Database configuration: HomoSapiens

FASTA file URL

```
http://www.uniprot.org/uniprot  
/?query=taxonomy:9606+AND+keyword:"Complete+proteome"&force=yes&format=fasta&  
include=yes
```

Reference file URL (?)

Version file URL (?)

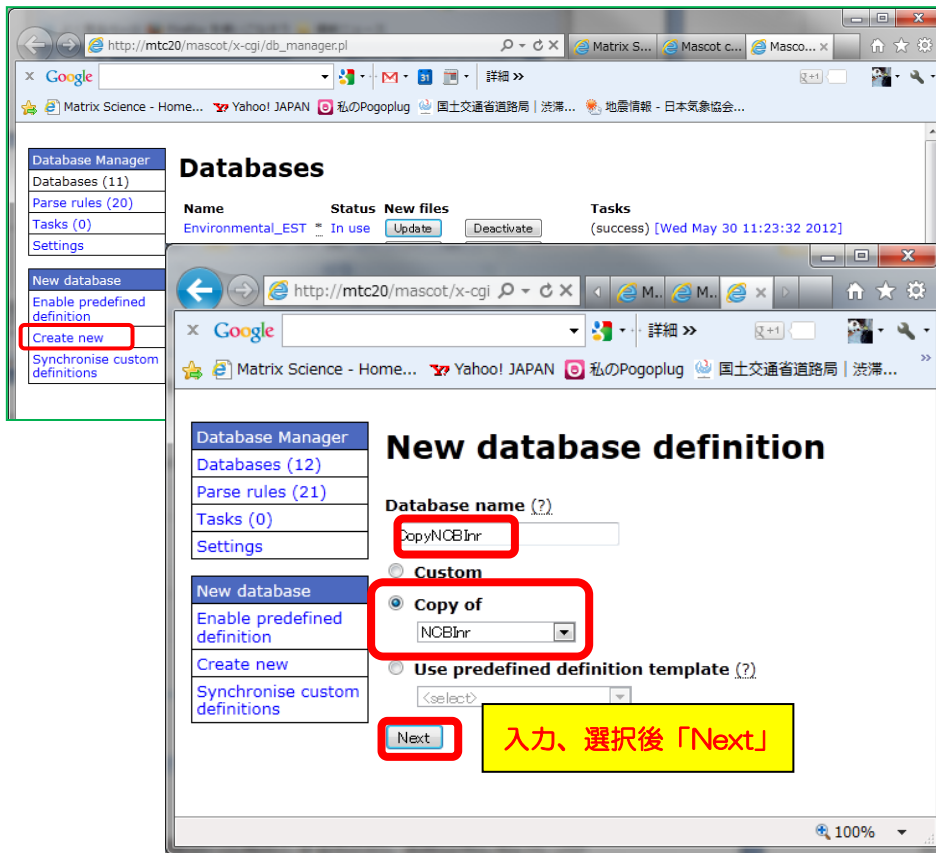
Cancel Save

【3. 既に使用しているデータベース設定をベースに、必要に応じ一部改変して使用【copy 利用】】

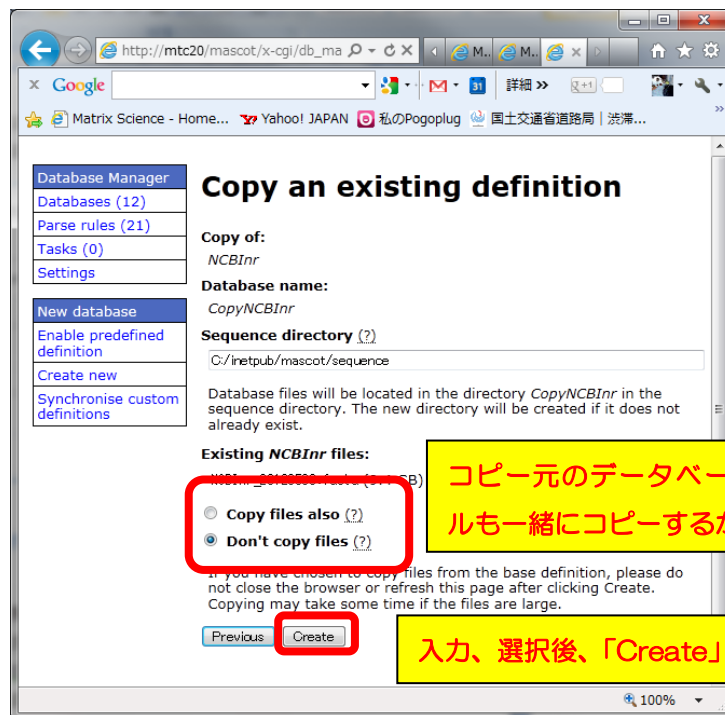
copy 利用は、設定ファイル中のテンプレートからでなく、既に利用しているデータベースの中から設定を選び、名前（と必要に応じてファイル設置場所）を変えてデータベース構築を行います。例えば既存データベースのバージョン違いを管理したい場合の利用等を想定しています。

copyを利用してデータベースを新規作成するには、Database Manager において、「New」 Database manager の、左フレーム「New database」-「**Create new**」を選びます（次頁図）。例では、NCBIInr と同じフォーマットで古いファイルを持ち続ける事を想定したデータベース「CopyNCBIInr」を作成します。現れた画面で「Database name」に名称を入力し、「**copy of**」で「**NCBIInr**」を選びます。記入・選択後、画面下の「Next」ボタンを押します。

copy利用の場合、ファイルそのものもフォルダにコピーして名前を変えて利用するかユーザーに訪ねてきます。コピーの必要がない場合は、「don't copy files」を選択し、「Create」ボタンを押します。



Databases -> Create New 画面

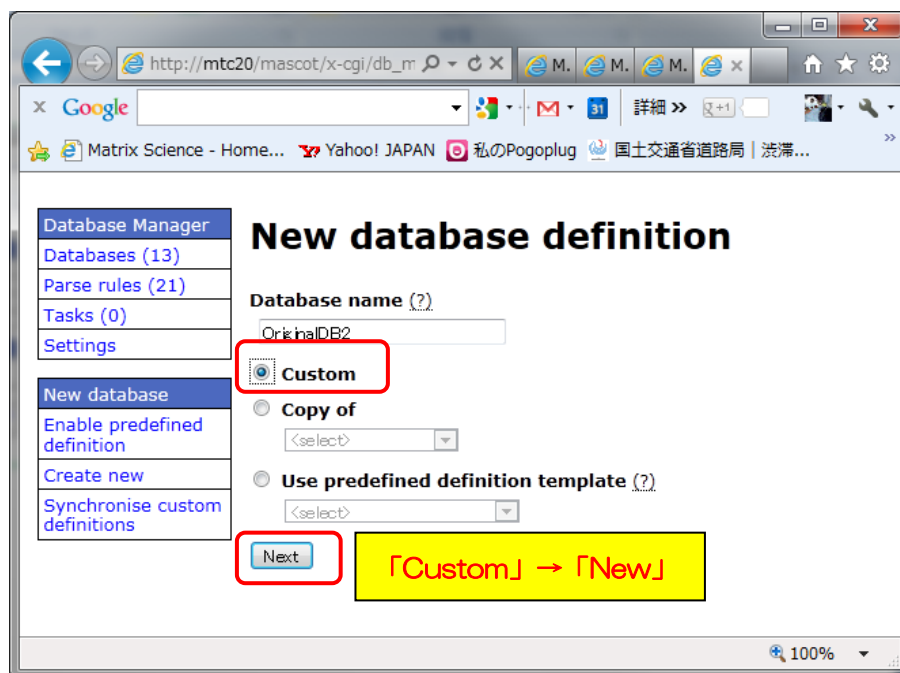
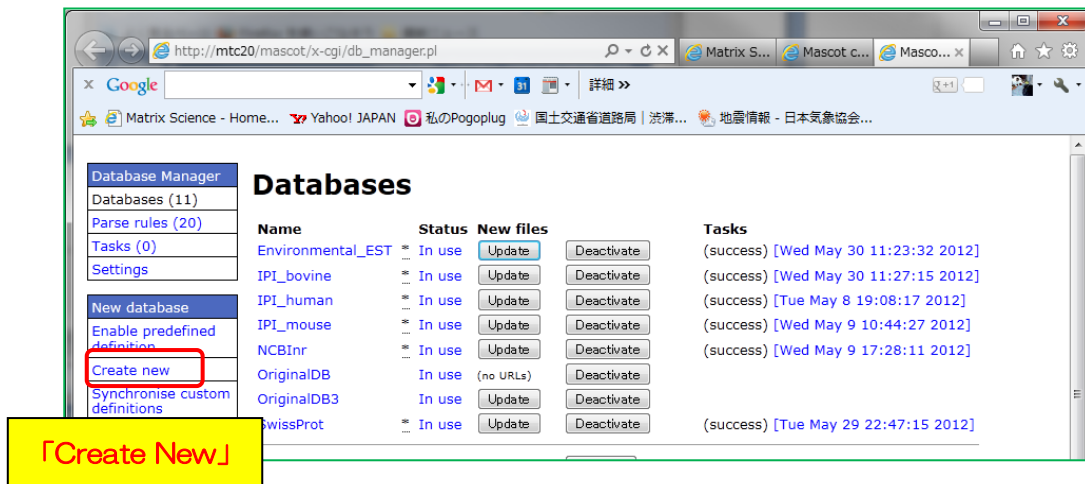


入力、選択後、「Create」

【4. 完全に自分で各設定を記入・選択 【Custom】】

1～3 で挙げた設定方法は、名前とファイルの設置場所以外、基本的に変更する事ができません。既存のデータベースのルールで当てはまらない抜き出しルールを適用する必要がある場合、本項目「Custom」を利用します。例では自らFASTAファイル「OriginalDB2」を作成しMASCOT用の検索データベースとして使用するケースを使って紹介します。

Database manager の、左フレーム「New database」- 「**Create new**」を選びます（下図）。現れた画面で「Database name」にデータベースの名称を入力し、「Use predefined definition template」で「**Custom**」を選びます。選択後、「Next」を押します。



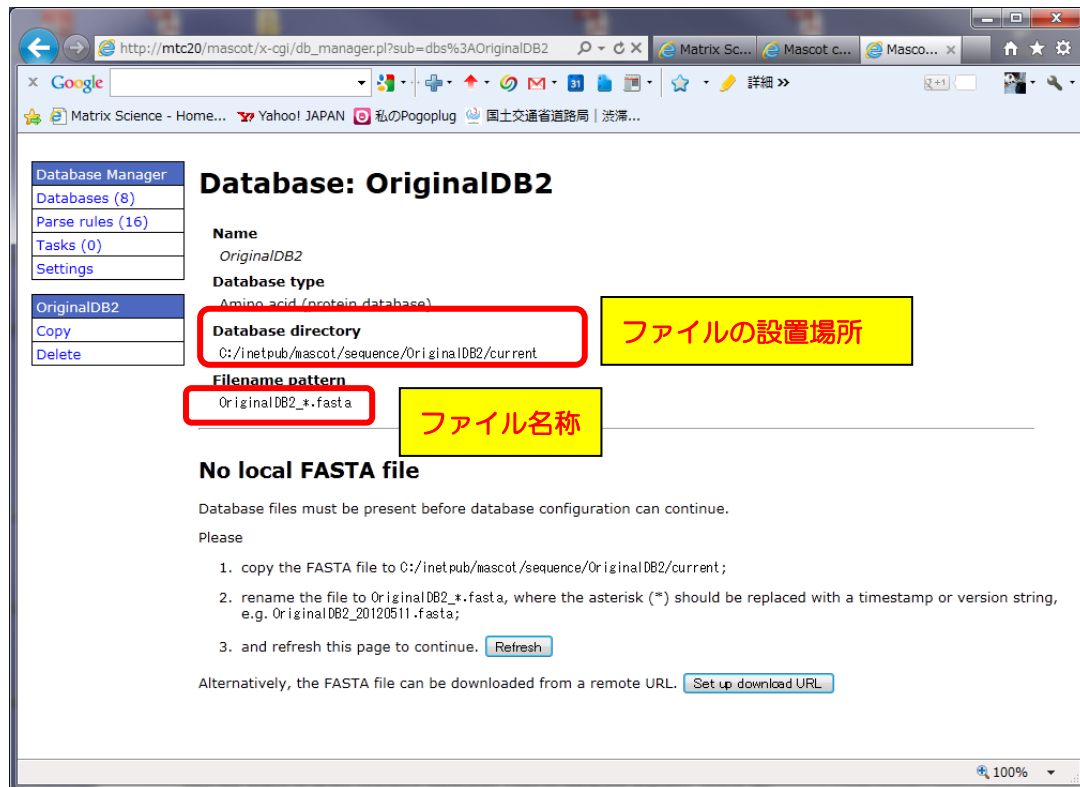
Databases -> Create New 画面

データベースの設定画面に移行します。データベースの配列の種類のほか、インターネット上にあるfastaファイルを自動で取得する設定にしたい場合、「Download database files」を選び、FASTAファイル、versionファイル、reference ファイルの取得先を記入します。自動取得でなく、手動で更新または更新する意思がない場合、「I will copy the files to the database directory」を選びます。各項目の選択後、「Create」ボタンを押します（下図）。

The screenshot shows the 'Custom definition' page of the Mascot Database Manager. The page includes a sidebar with navigation links, a main form area, and several yellow callout boxes with red text explaining specific options:

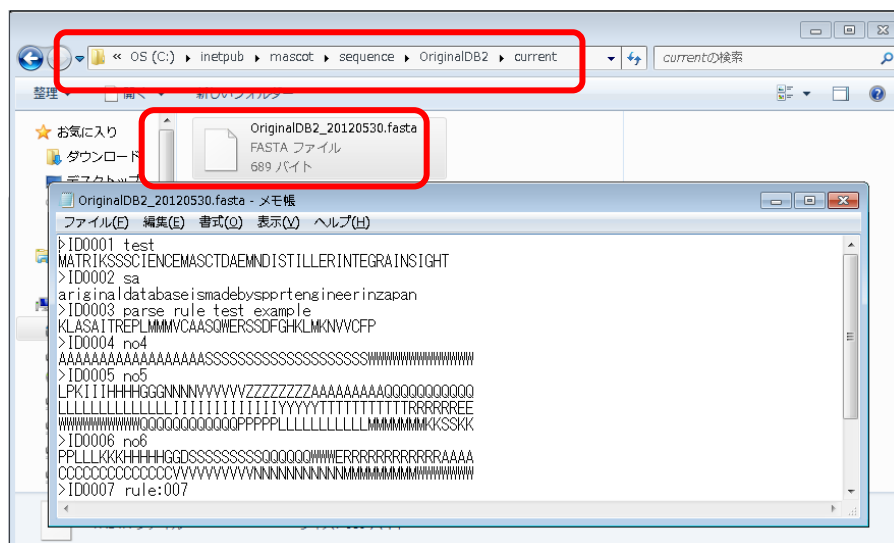
- Database name:** OriginalDB2
- Database type:** Amino acid (protein database) and Nucleic acid (DNA database). Callout: 「Database type」 アミノ酸か核酸か
- Sequence directory (?):** C:/inetpub/mascot/sequence. Callout: 「Sequence directory」 配列データベースファイルの置き場所
- Download database files (?):** Includes fields for FASTA file URL, Version file URL (optional), and Reference file URL (optional). Callout: 「Download database files」 自動更新プログラムにおけるファイルの取得先
- I will copy the files to the database directory (?):** . Callout: 「I will copy the files to the database directory」 自動ダウンロードしない場合に選択
- Buttons:** Previous and Create. Callout: 入力、選択後「Create」

まだファイルを所定の場所に置いていないので、ページ内に「No local FASTA file」と表示されます。検索対象のFASTAファイルを、「ファイル設置場所」に、ファイル名を「ファイル名称」に合わせた形に変更してコピーします。

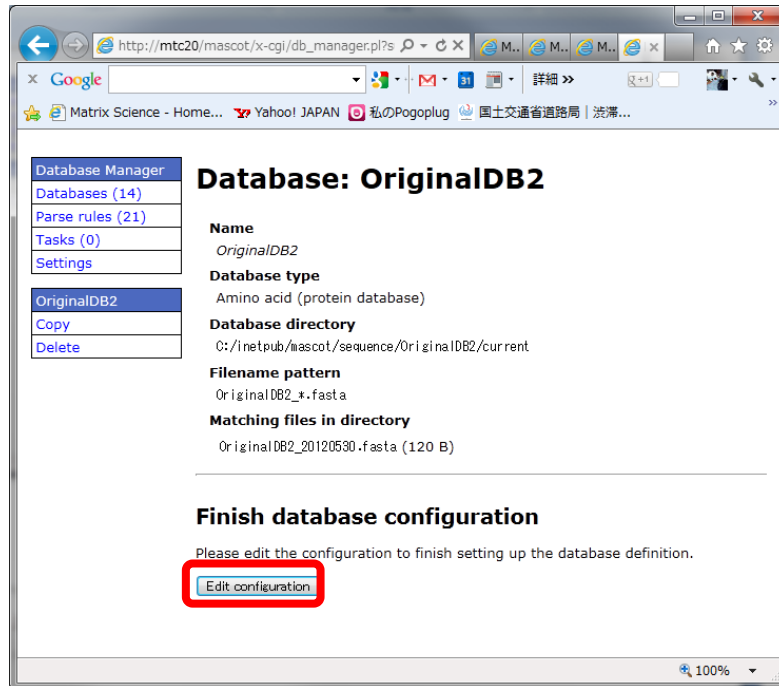


例) ファイル名: OriginalDB2_20120530.fastaを、

C:\inetpub\mascot\sequence\OriginalDB2\current フォルダへ



FASTAファイルを正しい名称にして所定の場所に正しく置くと、Database Manager 画面で認識されます。設定をさらに進めるため、「**Edit Configuration**」ボタンを押します。



続いて、IDとDescription の抜き出しルール設定を行います。各々の抜き出しルール設定画面へ移行すると、MASCOT 側で候補となるルールを最大5つ選択しユーザーに提示します（次頁図）。ユーザーはFASTAファイル内の各行の記述と抜き出された内容を見比べた上で、最適な抜き出し方を行っているルールを選択し、「Choose」ボタンを押します。続いて現れる画面で最終的にルールを適用した場合のIDが5例表示されますので、問題がなければ「**Confirm**」ボタンを押します。

同様の操作を「Description」でも実行します（次々頁図）

Accession の parse rule 設定

Database configuration: OriginalDB2 (step 1/2)

FASTA file

Matching files in directory
OriginalDB2_20120530.fasta (120 B)

Accession parse rule (?)
(none)

Choose Accession の「Choose」

Description parse rule (?)
(none)

Choose

Cancel **Next**

Accession and description parse rule must be selected before you can continue to the next step.

FASTA file accession parse rule for OriginalDB2

Raw test data (5 samples)

1. >ID0001 test

2. >ID0002 sa

3. >ID0003 parse rule test example

4. >ID0004 test

FASTA1 行目

Parse rules (5 suitable)

Parse rule	Match	Extracted data
<input type="radio"/> >["]* %(.*)	5/5	1. test ▶ 4 more matches
<input type="radio"/> >%(.*)	5/5	1. ID0001 test ▶ 4 more matches
<input checked="" type="radio"/> >%(" [^"]*")%	5/5	1. ID0001 ▶ 4 more matches
<input type="radio"/> >%("[^"]*")%	5/5	1. ID0001 test re matches re matches
<input type="radio"/> (new parse rule)		

**左) ルール
右) 抜き出し例
最適なものを選ぶ**

Match

1. ID0001
2. ID0002
3. ID0003
4. ID0004
5. ID0005

Previous **Confirm**

選択後「Choose」

Description の parse rule 設定

Database configuration: OriginalDB2 (step 1/2)

FASTA file
Matching files in directory
OriginalDB2_20120530.fasta (120 B)

Accession parse rule (?)
(none)

Description parse rule (?)
(none)
 Description の「Choose」

Accession and description parse rule are used in the next step.

FASTA file description parse rule for OriginalDB2

Raw test data (5 samples)

2. >ID0002 sa
3. >ID0003 parse rule test example

FASTA1 行目

Parse rules (5 suitable)

Parse rule	Match	Extracted data
<input checked="" type="radio"/> >[^]* \# (.*)	5/5	1. test ▶ 4 more matches
<input type="radio"/> >\# (.*)	5/5	1. ID0001 test ▶ 4 more matches
<input type="radio"/> >\# ([^]*)	5/5	1. ID0001 ▶ 4 more matches

左) ルール
右) 抜き出し例
最適なものを選ぶ

選択後「Choose」

FASTA file description parse rule for OriginalDB2

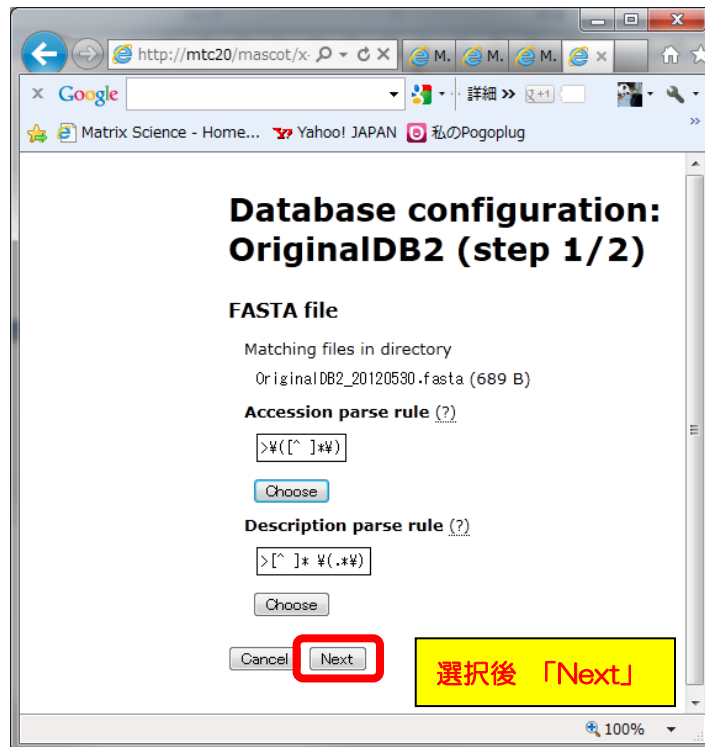
▶ Edit parse rule

Match

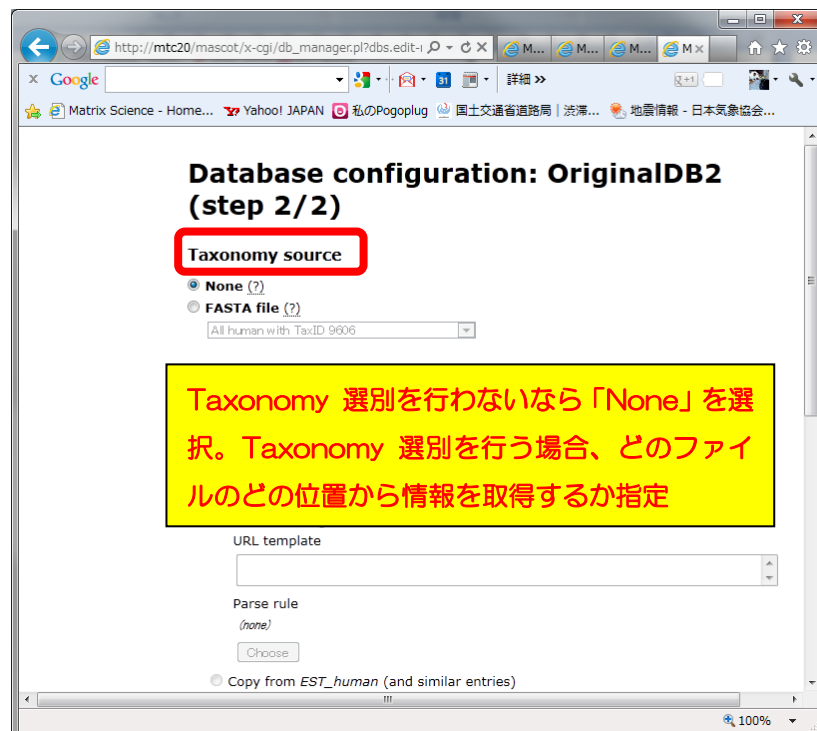
- test
- sa
- parse rule test example
- no4
- no5

他のエントリーでも問題なければ「Confirm」

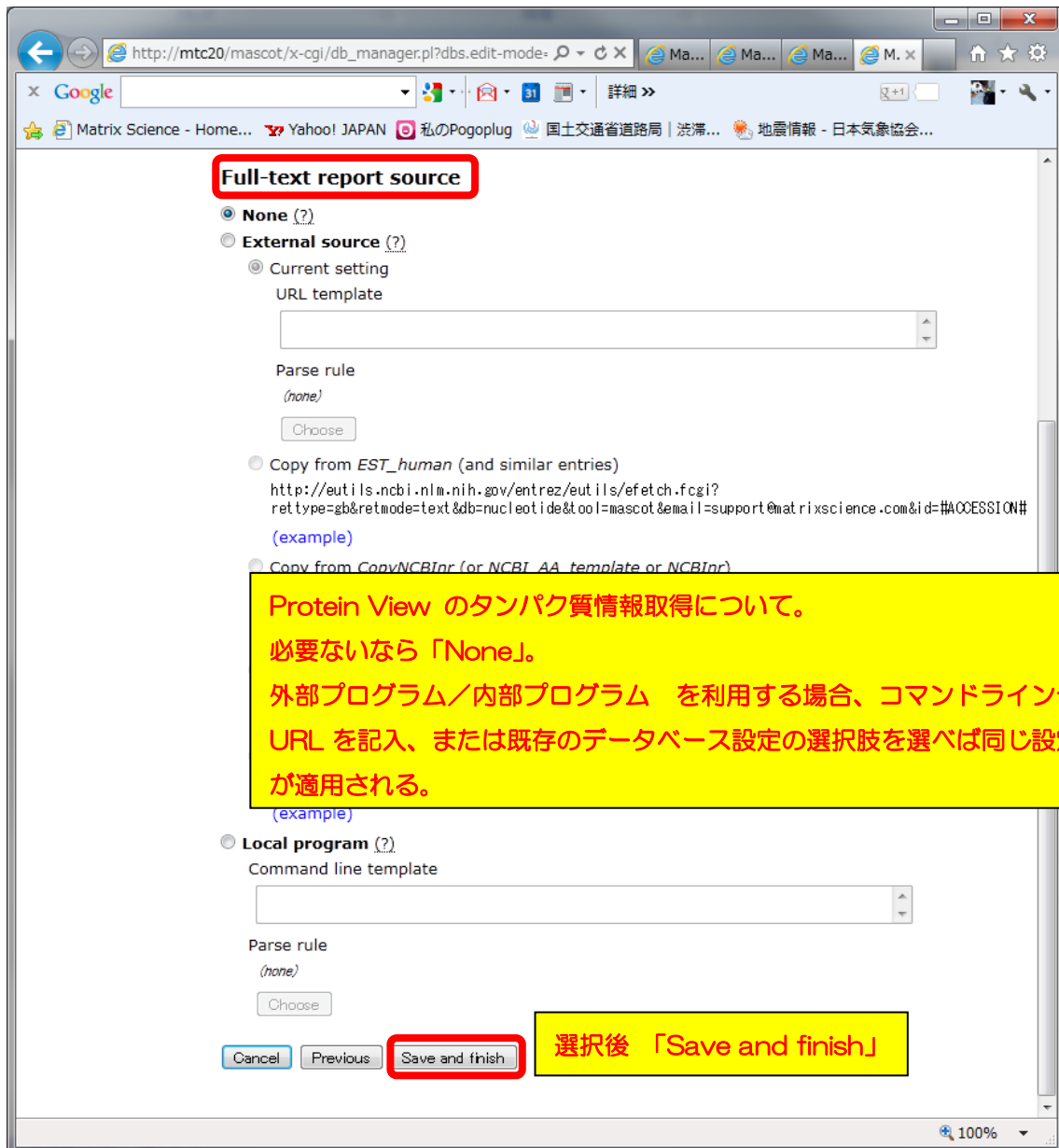
抜き出しルールを選んだら、「Next」ボタンを押します。



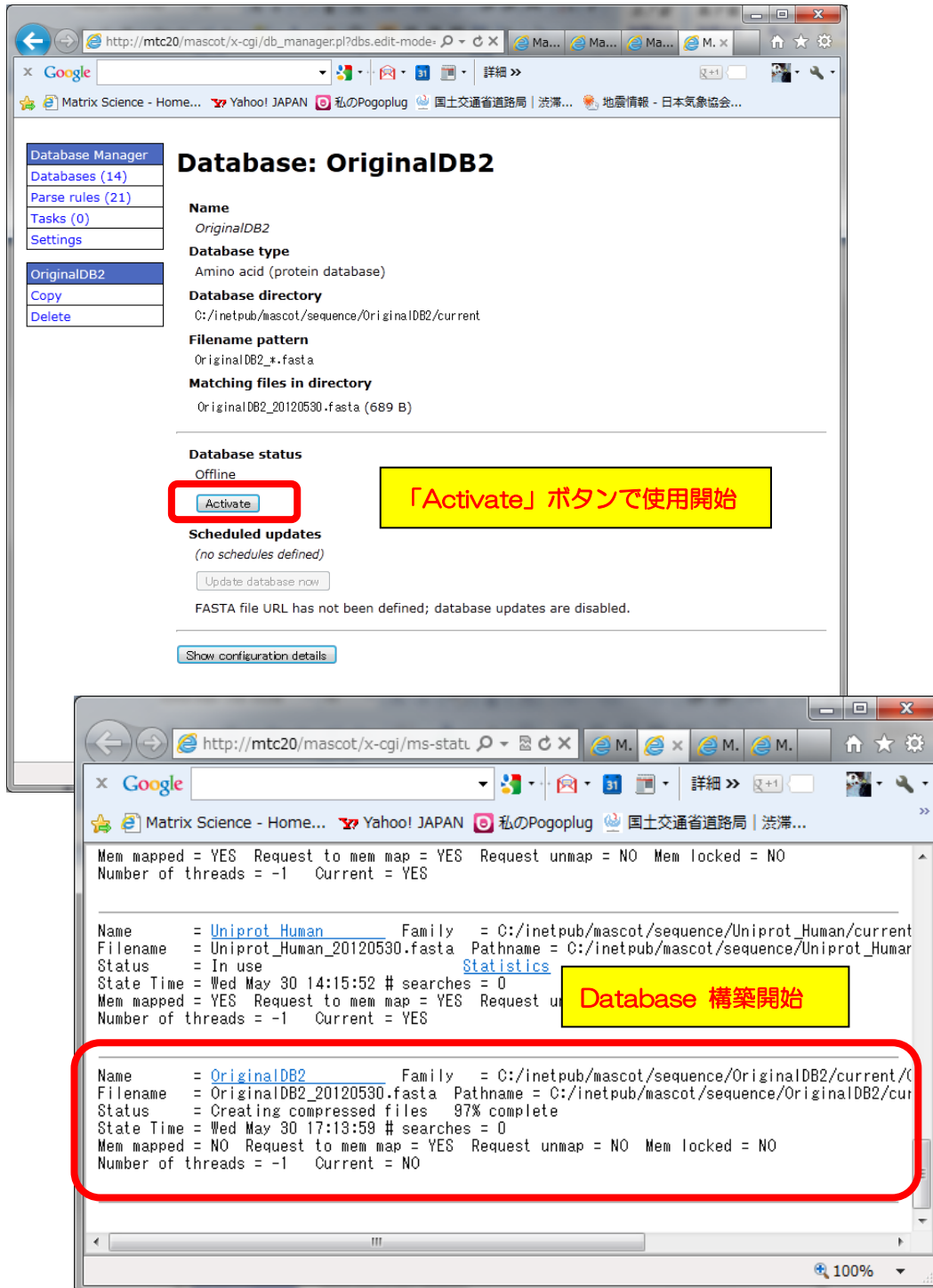
今度はTaxonomyとタンパク質の詳細情報取得先に関する設定画面となります。画面上部ではTaxonomy選別用のルールを選択します（下図）。



また、結果画面のProtein View 画面で表示されるタンパク質の詳細情報に関する設定も行うことができます。すべて選択を終えた後、画面最下部の「Save and finish」ボタンを押します。



個別のデータベース設定画面に戻ります。Save直後では Active状態になっていませんので、「Activate」ボタンを押せば、データベースの構築が開始され、検索に利用する事ができるようになります。



● Section 別 設定内容 :: Tasks Section

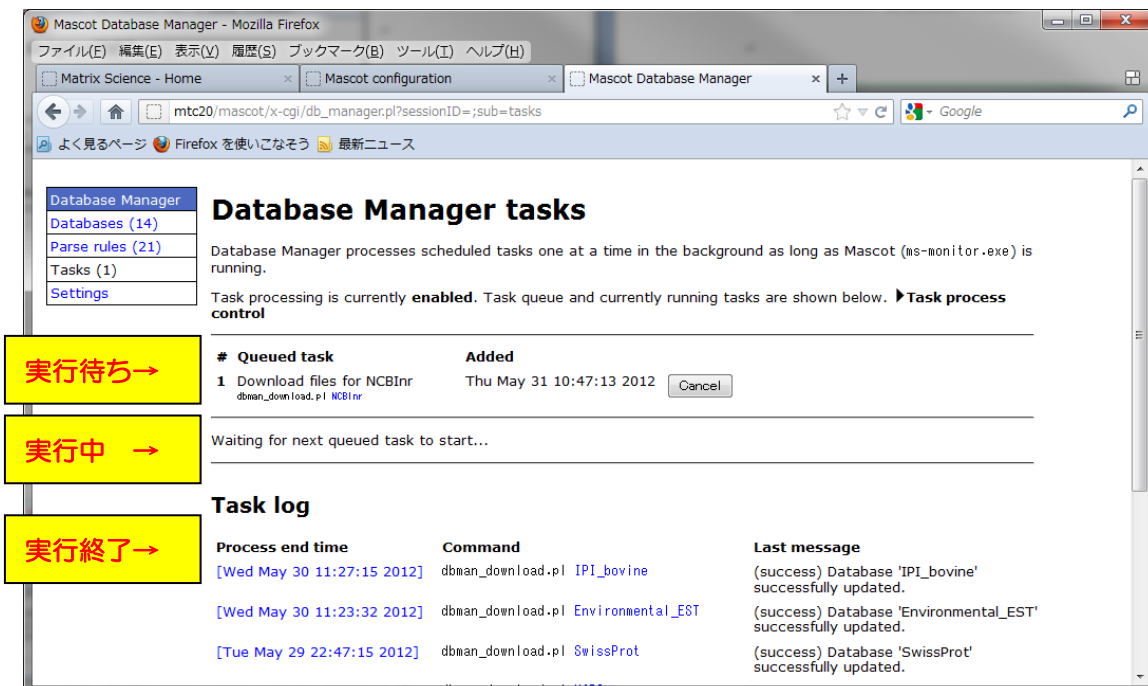
□ database 更新進捗の確認

データベース更新プログラムのプロセスは、「Task」として管理されます。更新の進捗を確認したい場合、「Tasks」セクションへアクセスしてください。

Database manager の、左フレーム「Database Manager」 - 「Tasks」を選びます。

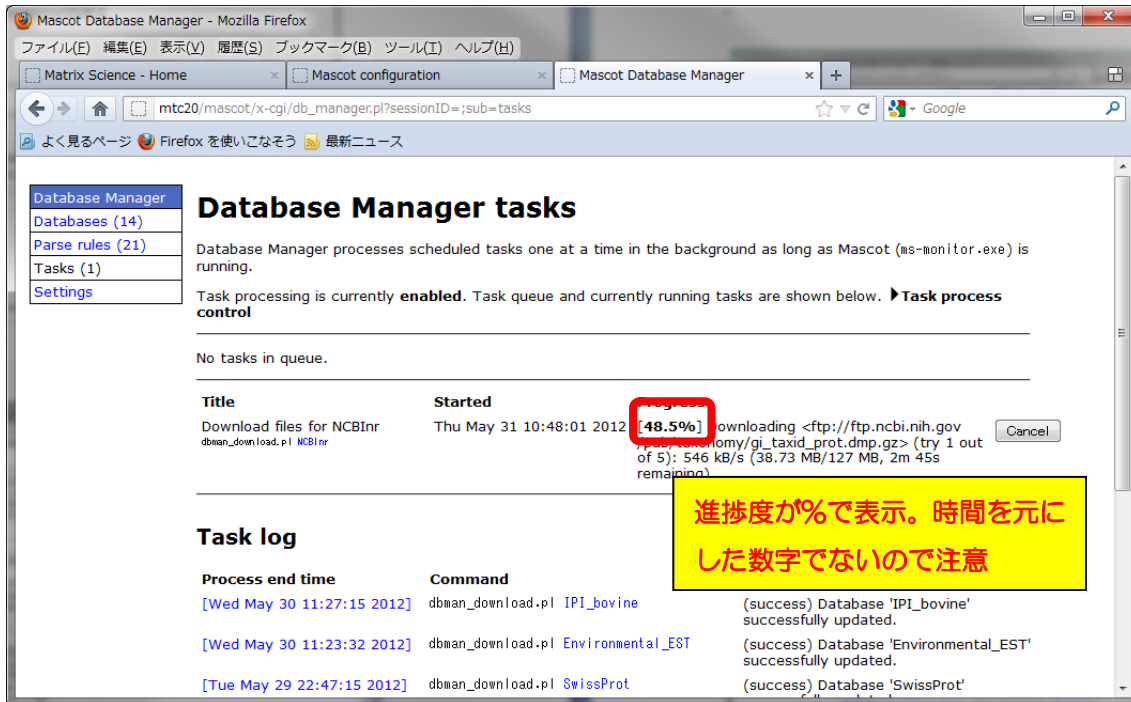


Task は 実行待ち/実行中/実行終了 の状態によって3つのセクションに分けて表示されます。



Tasks 画面

実行中のtaskについてはその進捗度が%表示で表され、ダウンロードするファイルのサイズと現段階でダウンロードしたファイルサイズ、並びに通信速度も併せて表示されます。



The screenshot shows the Mascot Database Manager web interface in a Mozilla Firefox browser. The page title is "Database Manager tasks". The main content area displays a task in progress: "Download files for NCBItr" with a progress bar at 48.5%. The task details include the start time "Thu May 31 10:48:01 2012" and the command "dbman_download.pl NCBItr". A yellow callout box with red text points to the progress bar, stating: "進捗度が%で表示。時間を元にした数字でないので注意" (Progress is shown as a percentage. Note that the number is based on time, not file size).

Title	Started	Progress	Details
Download files for NCBItr <small>dbman_download.pl NCBItr</small>	Thu May 31 10:48:01 2012	48.5%	downloading <ftp://ftp.ncbi.nih.gov/pub/blast/db/FASTA/gi_taxid_prot.dmp.gz> (try 1 out of 5): 546 kB/s (38.73 MB/127 MB, 2m 45s remaining)

Process end time	Command	Result
[Wed May 30 11:27:15 2012]	dbman_download.pl IPI_bovine	(success) Database 'IPI_bovine' successfully updated.
[Wed May 30 11:23:32 2012]	dbman_download.pl Environmental_EST	(success) Database 'Environmental_EST' successfully updated.
[Tue May 29 22:47:15 2012]	dbman_download.pl SwissProt	(success) Database 'SwissProt' successfully updated.

● 技術サポート

アップグレードに関してご質問等ありましたら弊社技術サポートにご連絡ください。

電子メール : support-jp@matrixscience.com

電話 : 03-5807-7897 ファックス : 03-5807-7896