

Mascot Server 2.4 配列データベースの手動更新

● はじめに

MASCOT Server でデータベースファイルの自動更新設定が正常に動作しない場合、ファイ ルを自身で取得し MASCOT Server の所定の位置にファイルを配置する事でデータベースファ イルを更新する事ができます(この操作の事を「データベースの手動更新」と呼んでいます)。 以下手動更新の具体的な方法について、NCBInr, SwissProt のデータベース別にご説明いたし ます。

● NCBInr 手動更新

ロ ファイル取得

始めに以下の3つのファイルを WEB ブラウザ等でダウンロードします。ダウンロード先はデ スクトップにせず、C:¥temp¥download 等のフォルダを作成しそこへ直接ダウンロードする ようにしてください。

[FASTA]

ftp://ftp.ncbi.nih.gov/blast/db/FASTA/
にある「nr.gz」ファイル
【taxonomy 1】
ftp://ftp.ncbi.nih.gov/pub/taxonomy/
にある、「gi_taxid_prot.dmp.gz」ファイル
【taxonomy 2】
ftp://ftp.ncbi.nih.gov/pub/taxonomy/
にある、「taxdump.tar.gz」ファイル

*各操作で一度でもデスクトップやマイドキュメント、各ユーザーの download フォルダなど、使用ユ ーザー固有のセキュリティ権限に設定されているフォルダにファイルを置いてしまうと、ファイルセキュ リティ権限の問題で MASCOT における検索が正常に動作しない事があります。ダウンロードの際も必ずデス



クトップ等を避け、 C:¥temp¥download や C:¥inetpub¥mascot¥mskk¥download など、セキュリティに問題 が出ないフォルダを指定するようにしてください。

【ブラウザ操作によるファイルダウンロード操作例】

後 ftp://ftp.ncbi.nih.gov/blast/db/FASTA/の一覧 - Mozilla Firefox ファイル(F) 寝生(F) 表示(V) 扇唇(S) ブックマーク(B) ツール(T)	アドレス部分を入力し…		
Ftp://ftp.ncbi.nih.gov/bloct/db/rA51		_	8
♦ ⇒ ↑ (2) ftp://ftp.ncbi.nih.gov/blast/db/FASTA/		ු ⊽ C 🚼 - Goo	ogle 🔎
🖻 よく見るページ 🥹 Firefox を使いこなそう <u>Ы</u> 最新ニュース			
month.nt.gz	118962 KB	2012/12/12	19:03:00 ^
month.nt.gz md5	1 / P	2012/12/12	11:44:00
nr.gz 対象ファ1	イルを選んで右クリック	2012/12/12	19:06:00
リンクを新しいタブで開く(江) で開<(<u>W</u>) 1 KB	2012/12/12	4:06:00
nt.gz ろりとクをすいりマーク(1484055 KB	2012/12/12	19:12:00
			5:07:00
	24103004 KD	2012/12/12	19:23:00
other_genomic.gz.md	d5 1 KB	2012/12/12	9:08:00
notos az	162060 MD	2012/12/12	10.27.00 -

🥹 名前を付けて保存			×
🚱 🤇 🦞 « temp 🖡 download	-	6 downloadの検索	Q
整理 ▼ 新しいフォルダー			···· •
 ☆ お気に入り ♪ ダウンロード ■ デスクトップ 10 11 12 13 14 <	~ デスクトップや「ダウンロ ダ以外の場所に保存します。	更新日時 ード」フォル ん。	種類
⇒ イブラリ ドキュメント ビクチャ マ			Þ
ファイル名(N): nr.gz ファイルの種類(T): すべてのファイル	(*.*)		
 フォルダーの非表示 		保存(S)	キャンセル



□ ファイル解凍

ダウンロードしたファイルを解凍します。 圧縮形式「gz」は Windows の標準機能で解凍する 事ができません。 解凍を行うソフトウェアをお持ちの場合それをお試しください。 弊社も手軽に 使用できる解凍ソフトを準備しておりますので、以下そのソフトウェアを使った操作例を記しま す。

MASCOT の mskk フォルダ、または インストール DVD 内に入っている解凍プログラム 「**ERANGE.EXE**」(インストール不要、インターネットから取得する場合 <u>http://www.vector.co.jp/soft/win95/util/se094501.html</u>)をダブルクリックして、「書庫 名と同名のフォルダを作る」にチェックを入れてください(例:下図)。

	(C:) ▶ inetpub ▶ mascot ▶ mskk ▶	utilities	• \$9
整理 ▼ ライブラリに追加 ▼	共有 ▼ 書き込む 新しいフォルタ	ř—	
🔶 お気に入り 🔷 名前	ăŬ	更新日時	種類
 ダウンロード デスクトップ アスクトップ アスクトップ 	ERANGE.EXE	2002/06/05 1:41 2012/05/11 12:27	アプ! アプ!
	²⁵ (→ sR 定	○ フォレタを指定(E) 参照(B) , i	

圧縮ファイルを「ERANGE.EXE」にドラッグ&ドロップで重ねるだけで解凍が開始し、ファ イルとと同名のフォルダが作成されます。フォルダ内に解凍ファイルが作成されます。



□ ファイル名の変更

解凍して得られたファイルのうち、「nr」フォルダ内の「nr」ファイルの名称を nr から 「NCBInr_yyyymmdd.fasta」 に変更します。

* yyyymmdd 部分はダウンロードした日付です。例)2012 年 12 月 17 日→20121217



□ ファイルの移動

展開して得られたファイルを以下のフォルダへ移動します。

ファイル	移動先のフォルダ
「taxdump」フォルダ内複数ファイルすべてと	C:¥inetpub¥mascot¥ taxonomy
「gi_taxid_prot.dmp」フォルダ内の	
「gi_taxid_prot.dmp」ファイル	
NCBInr_yyyymmdd.fasta	C:¥inetpub¥mascot¥ <mark>sequence¥NCBInr</mark> ¥current

taxonomy フォルダへカット&ペーストしたファイルは上書き保存するか尋ねられると思いますが、「はい」を選択し上書きしてください。



□ Database Status での確認

Web browser で Home \rightarrow Database Status を開きます。

		x	
(⇐) (⇐) (⇐) http://localhost/mascot/ (⇐) (⇐) (⇐) http://localhost/mascot/ (⇐) (⇐) (⇐) (⇐) (⇐) (⇐) (⇐) (⇐) (⇐) (⇐)	î î	7 🛱	
× Google →	ログイン	、 設定	
👍 🥙 Matrix Science - Home 🐄 Yahoo! JAPAN 💿 私のPogoplug 실 国土交通省道路局 渋滞 🎭 地震情報 - 日本3	氢象協会		
Mascot Utilities		-	
Your system administrator may have restricted access to these utilities			
Database Status View the status of all the sequence databases. Links to database statistics, search status, log files, etc.			
Search Log A tabular view of the search log. Can be filtered to find specific search results.			
Configuration Editor Browser-based configuration editors, including security administration and sequence database management.			
Mascot Daemon			
Mascot Daemon is a client application which automates the submission of data files to a Mascot server. It under Microsoft Windows 2000 and later. Install Upgrade Troubleshoot	runs	-	
http://localhost/mascot/x-cgi/ms-status.exe	۹ 100%	•	

以下のように、一連の作業で current フォルダに置いた NCBInr のデータベースファイルが MASCOT に認識され、データベースの構築中(Creating compressed files N% complete)になっていればOKです。「Status」項目が "In Use"になれば構築が完了です。In Use になるまでの時間は環境により変わりますが、30 分~1 時間 程度とお考えください。

Mem mapped = NO Request to mem map = NO Request unmap = NO Mem locked = NO Number of threads = 4 Current = YES
Name = <u>IPI human</u> Family = C:/inetpub/mascot/sequence/IPI_human/current/IPI_human_*.fasta Filename = IPI_human_20120825.fasta Pathname = C:/inetpub/mascot/sequence/IPI_human/current/IPI_human_20120825.fasta Status = In use <u>Statistics</u> State Time = Mon Dec 10 16:20:19 # searches = 0
Mem mapped = YES Request to mem map = YES Request unmap = NO Mem locked = NO Number of threads = 4 Current = YES
Name = <u>NCBInr</u> Family = C:/inetpub/mascot/sequence/NCBInr/current/NCBInr_*.fasta Filename = NCBInr_20121121_fasta Filename = Creating compressed files 1% complete Status = Creating compressed files 1% complete Status Fire = Two Dec 13 15:24:49 H searches = 0
Mem mapped = NO Request to mem map = YES Request unmap = NO Mem locked = NO Number of threads = 4 Current = NO
Name = NCBInr Family = C:/inetpub/mascot/sequence/NCBInr/current/NCBInr_*.fasta Filename = NCBInr_20120910.fasta Pathname = C:/inetpub/mascot/sequence/NCBInr/current/NCBInr_20120910.fasta Status = In use Statistics Compression warnings Unidentified taxonomy State Time = Thu Dec 13 15:20:40 # searches = 0 Mem mapped = YES Request to mem map = YES Request unmap = NO Mem locked = YES Number of threads = 4 Current = YES Current = YES Compression warnings Compression warni
Name = <u>&wissProt</u> Family = C:/inetpub/mascot/sequence/SwissProt/current/SwissProt_*.fasta Filename = SwissProt_2012_10.fasta Pathname = C:/inetpub/mascot/sequence/SwissProt/current/SwissProt_2012_10.fasta



SwissProt 手動更新

□ ファイル取得

始めに以下の5つのファイルをWEBブラウザ等でダウンロードします。ダウンロード先はデ スクトップにせず、C:¥temp¥download 等のフォルダを作成しそこへ直接ダウンロードする ようにしてください。

[FASTA]

ftp://ftp.ebi.ac.uk/pub/databases/uniprot/current_release/knowledgebase/complete/

にある「uniprot_sprot.fasta.gz」ファイル

[dat]

ftp://ftp.ebi.ac.uk/pub/databases/uniprot/current_release/knowledgebase/complete/

にある、「uniprot_sprot.dat.gz」ファイル

[nam]

ftp://ftp.ebi.ac.uk/pub/databases/uniprot/current_release/knowledgebase/complete/

にある、「reldate.txt」ファイル

[taxonomy 1]

ftp://ftp.ncbi.nih.gov/pub/taxonomy/

にある、「taxdump.tar.gz」ファイル

[taxonomy 2]

http://www.matrixscience.com/downloads/speclist.txt

にある、「**speclist.txt**」ファイル

* 上記 URL をそのまま入力し、web browser で開いたページをファイル保存します。

*各操作で一度でもデスクトップやマイドキュメント、各ユーザーの download フォルダなど、使用ユ ーザー固有のセキュリティ権限に設定されているフォルダにファイルを置いてしまうと、ファイルセキュ リティ権限の問題で MASCOT における検索が正常に動作しない事があります。ダウンロードの際も必ずデス クトップ等を避け、 C:¥temp¥download や C:¥inetpub¥mascot¥mskk¥download など、セキュリティに問題 が出ないフォルダを指定するようにしてください。



【ブラウザ操作によるファイルダウンロード例】



保存ファイル名を入力してください…	the second station	
 ・ temp → download 整理 ▼ 新しいフォルダー 	くし、dewstandの検索 デスクトップや「ダウンロード」以外の 場所に保存します	< € • ==
名前 ペ ホームグループ	検索条件に一致する項目はありません。	種類
1 コンピューター		
♣ \$\u00ed\$ \u00ed\$ \u00ed\$	III	4
ファイル名(N): uniprot_sprot.fasta.g	z	•
ファイルの種類(T): gzip (*.gz)		•
🗻 フォルダーの非表示	保存(S)	キャンセル



□ ファイル解凍

ダウンロードしたファイルを解凍します。ただし解凍が必要なのは、5 ファイルのうち 「uniprot_sprot.fsata.gz」、「uniprot_sprot.dat.gz」「taxdump.tar.gz」の3つで、後の こつは解凍する必要はありません。また、圧縮形式「gz」はWindowsの標準機能で解凍する事 ができません。解凍を行うソフトウェアをお持ちの場合それをお試しください。弊社も手軽に使 用できる解凍ソフトを準備しておりますので、以下そのソフトウェアを使った操作例を記します。

MASCOT の mskk フォルダ、または インストール DVD 内に入っている解凍プログラム 「**ERANGE.EXE**」(インストール不要、インターネットから取得する場合 <u>http://www.vector.co.jp/soft/win95/util/se094501.html</u>)をダブルクリックして、「書庫 名と同名のフォルダを作る」にチェックを入れてください(例:下図)。



圧縮ファイルを「ERANGE.EXE」にドラッグ&ドロップで重ねるだけで解凍が開始し、ファ イルと同じフォルダに回答ファイルと同名のフォルダが作成されます。



□ ファイル名の変更

ファイルのうち、以下の3つのファイルの名称を変更してください。

- ・「uniprot_sprot.fsata」フォルダ内の「**uniprot_sprot.fsata**」ファイル → SwissProt_YYYY_MM.fsata
 - * YYYY_MM 部分はダウンロードした年と月。例) 2012_12 →2012 年 8 月
- ・「uniprot_sprot.dat」フォルダ内の「**uniprot_sprot.dat**」ファイル
 - \rightarrow SwissProt_YYYY_MM.dat
- ・「**reldate.txt**」ファイル
 - → SwissProt_YYYY_MM.nam





□ ファイルの移動

展開して得られたファイルを以下のフォルダへ移動します。taxonomy フォルダへカット& ペーストしたファイルは上書き保存するか尋ねられると思いますが、「はい」を選択し上書きしてください。

ファイル	移動先のフォルダ
「taxdump」フォルダ内複数ファイルすべて	C:¥inetpub¥mascot <mark>¥taxonomy</mark>
と「speclist.txt」ファイル	
SwissProt_YYYY_MM.fsata	C:¥inetpub¥mascot¥ <mark>sequence</mark> ¥ <mark>SwissProt</mark> ¥current
SwissProt_YYYY_MM.dat	
SwissProt_YYYY_MM. nam	

□ Database Status での確認

Web browser で Home → Database Status を開きます。





以下のように、一連の作業で current フォルダに置いた SwissProt のデータベースファイル が MASCOT に認識され、データベースの構築中(Creating compressed files N% complete) になっていればOKです。「Status」項目が "In Use"になれば構築が完了です。In Use にな るまでの時間は環境により変わりますが、10分~25分 程度とお考えください。



● 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