

SwissProt 手動更新方法

以下、SwissProt データベースを手動で更新する方法についてご案内します。

- 1 ファイル取得
- 2 取得ファイルの解凍・展開
- 3 reldate.txt ファイルの記載内容確認と各ファイルの改名
- 4 ファイルの移動
- 5 Database Status の確認

■1 ファイル取得

以下のファイルを取得します。**ブラウザ**、FTP や SCP クライアントソフトウェアなどが利用できます。ダウンロードしたファイルは **C:¥temp¥download** フォルダに直接ダウンロードするか取得後そちらに移動してください。

A. FASTA ファイル

[http://ftp.ebi.ac.uk/pub/databases/uniprot/current release/knowledgebase/complete/uniprot sprot.fasta.gz](http://ftp.ebi.ac.uk/pub/databases/uniprot/current%20release/knowledgebase/complete/uniprot%20sprot.fasta.gz)

B. Reference ファイル

[http://ftp.ebi.ac.uk/pub/databases/uniprot/current release/knowledgebase/complete/uniprot sprot.dat.gz](http://ftp.ebi.ac.uk/pub/databases/uniprot/current%20release/knowledgebase/complete/uniprot%20sprot.dat.gz)

C. Version ファイル

[http://ftp.ebi.ac.uk/pub/databases/uniprot/current release/knowledgebase/complete/reldate.txt](http://ftp.ebi.ac.uk/pub/databases/uniprot/current%20release/knowledgebase/complete/reldate.txt)

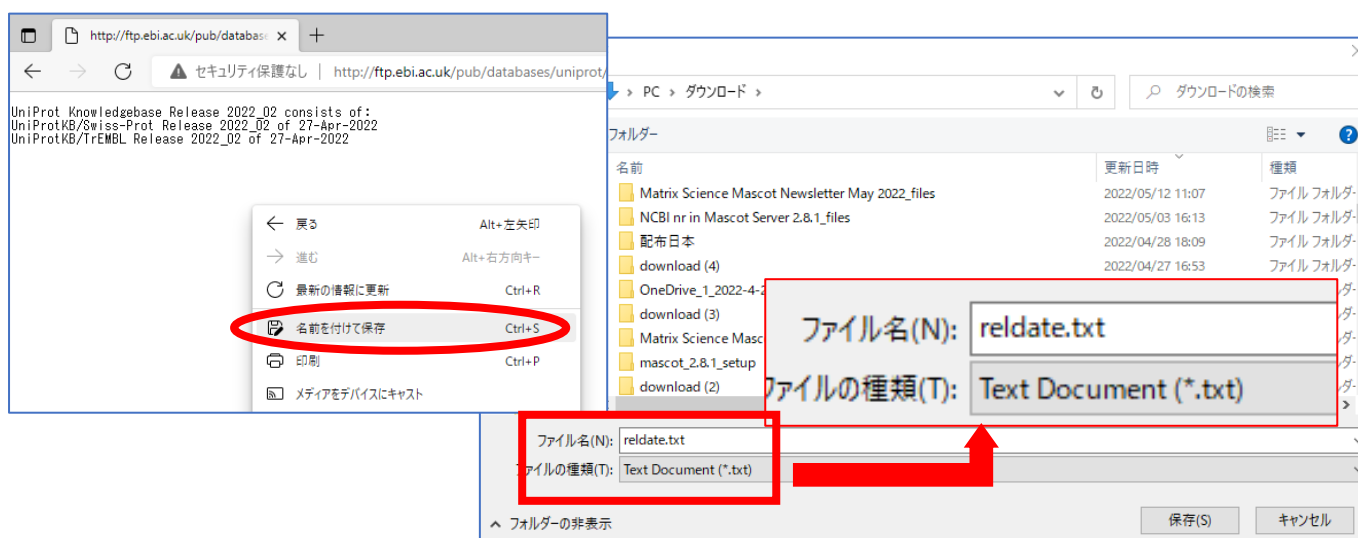
D. taxonomy ファイル（複数）

<https://ftp.ncbi.nlm.nih.gov/pub/taxonomy/taxdump.tar.gz>

E. speclist ファイル

<http://ftp.ebi.ac.uk/pub/databases/uniprot/knowledgebase/docs/speclist.txt>

C と E については、ブラウザで URL を入力しエンターを押すとページが開いてしまいます。以下のように右クリックでファイルを保存する選択肢を選び、C「reldate.txt」、E「speclist.txt」の名称でそのまま保存してください。



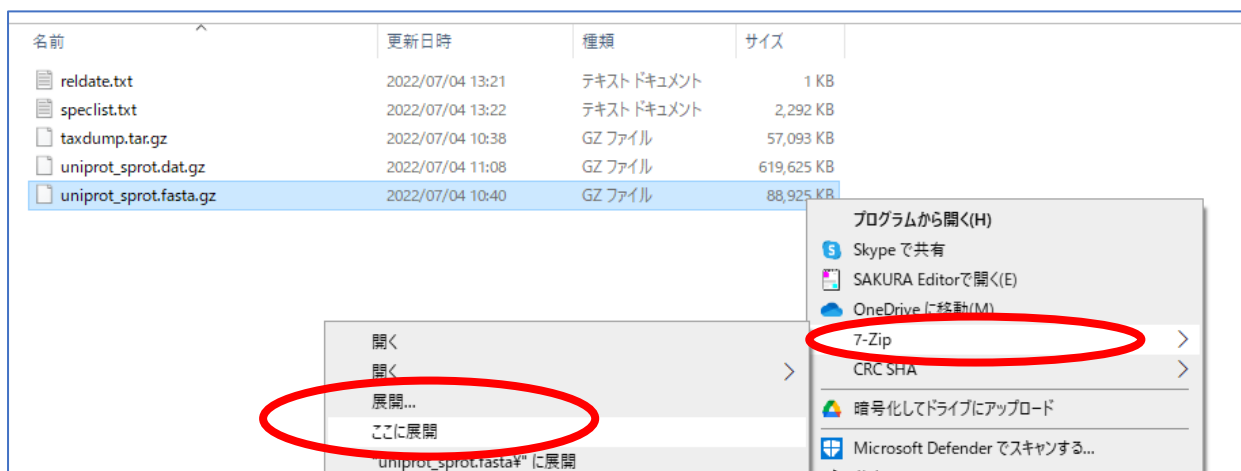
A~E すべてのファイルを取得し 1 か所に保存すると以下の図のようになるはずですが。

名前	更新日時	種類	サイズ
reldate.txt	2022/07/04 13:21	テキストドキュメント	1 KB
speclist.txt	2022/07/04 13:22	テキストドキュメント	2,292 KB
taxdump.tar.gz	2022/07/04 10:38	GZ ファイル	57,093 KB
uniprot_sprot.dat.gz	2022/07/04 11:08	GZ ファイル	619,625 KB
uniprot_sprot.fasta.gz	2022/07/04 10:40	GZ ファイル	88,925 KB

■2 取得ファイルの解凍・展開

取得したファイルのうち、「A.FASTA ファイル(uniprot_sprot.fasta.gz)」「B.Reference ファイル(uniprot_sprot.dat.gz)」「D.taxonomy ファイル(taxdump.tar.gz)」を**解凍・展開**します。取得したファイルの圧縮形式「.gz」や「.tar」は windows の標準で解凍や展開をする事ができません。「7zip」(<https://sevenzip.osdn.jp/>)などのフリーソフトウェアや各種有料ソフトウェアをインストールする必要があります。

なお taxonomy ファイルについては、gz を解凍後「taxdump.tar」ファイルになった場合、さらに展開を実行して複数のファイルを得ます。展開方法によっては「taxdump」フォルダができてその中にファイルがすべて展開される事もありますし、あるいはフォルダができずに taxdump.tar ファイルと同じ階層に関連ファイルがすべて展開される事もあります。

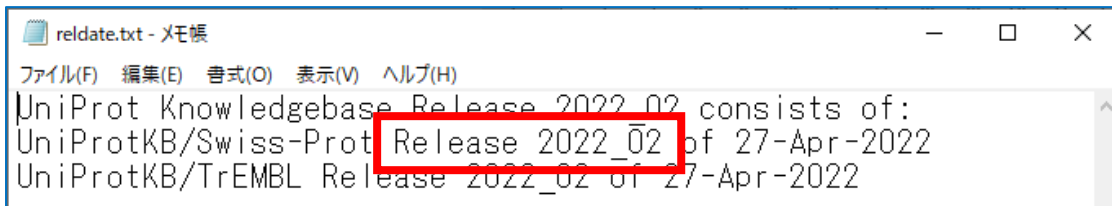


3つのファイルを解凍・展開すると以下のようになるはずですが(その場に解凍・展開した場合と一旦フォルダに解凍された場合で状況が多少異なります。)

名前	更新日時	種類	サイズ
taxdump	2022/07/04 13:26	ファイル フォルダー	
reldate.txt	2022/07/04 13:21	テキストドキュメント	1 KB
speclist.txt	2022/07/04 13:22	テキストドキュメント	2,292 KB
taxdump.tar	2022/07/04 10:28	TAR ファイル	402,690 KB
taxdump.tar.gz	2022/07/04 10:38	GZ ファイル	57,093 KB
uniprot_sprot.dat	2022/03/11 9:12	DAT ファイル	3,627,649 KB
uniprot_sprot.dat.gz	2022/07/04 11:08	GZ ファイル	619,625 KB
uniprot_sprot.fasta	2022/03/11 9:20	FASTA ファイル	275,330 KB
uniprot_sprot.fasta.gz	2022/07/04 10:40	GZ ファイル	88,925 KB

■3 reldate.txt ファイルの記載内容確認とファイルの改名

reldate.txt の中身を確認し、swissprot のバージョンを確認します。以下の例の場合、「2022_02」です。以降、バージョン番号を「YYYY_VV」と表記します。



```
reldate.txt - メモ帳
ファイル(F) 編集(E) 書式(O) 表示(V) ヘルプ(H)
UniProt Knowledgebase Release 2022_02 consists of:
UniProtKB/Swiss-Prot Release 2022_02 of 27-Apr-2022
UniProtKB/TrEMBL Release 2022_02 of 27-Apr-2022
```

reldate.txt を閉じ、A～C の 3 ファイルについて以下のように名称を変更します。

***大文字小文字、「_」をデータベース名に揃えるようご注意ください。**

A. FASTA ファイル

uniprot_sprot.fasta → [SwissProt_YYYY_VV.fasta](#)

B. Reference ファイル

uniprot_sprot.dat → [SwissProt_YYYY_VV.dat](#)

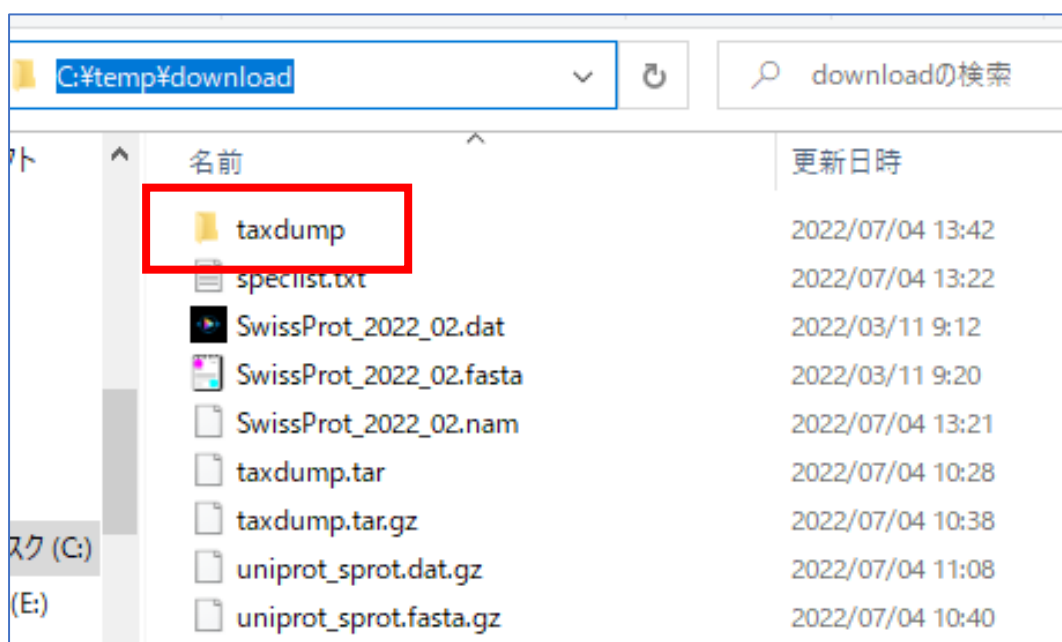
C. Version ファイル

reldate.txt → [SwissProt_YYYY_VV.nam](#) * ファイル名も拡張子も変わるのでご注意ください。

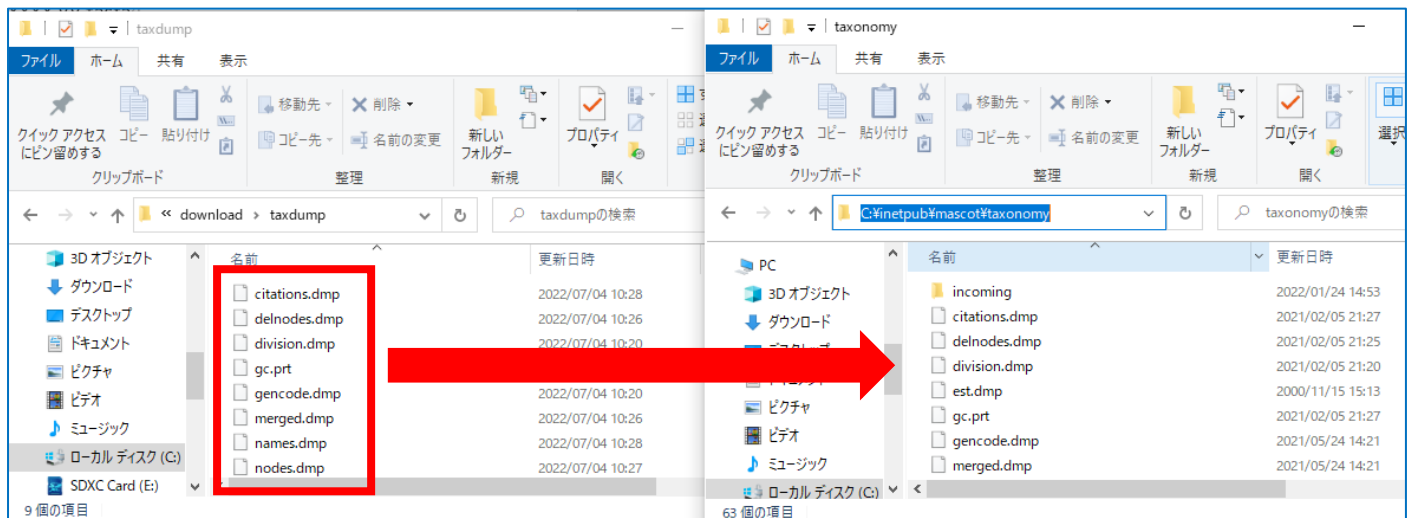
■4 ファイルの移動

解凍や展開、改名したファイルを移動します。まず、taxonomy 関連である D,E のファイルを移動して、その後 A～Cのファイルを**同時に**ドラッグ&ドロップ(またはカット&ペースト)で移動させます。

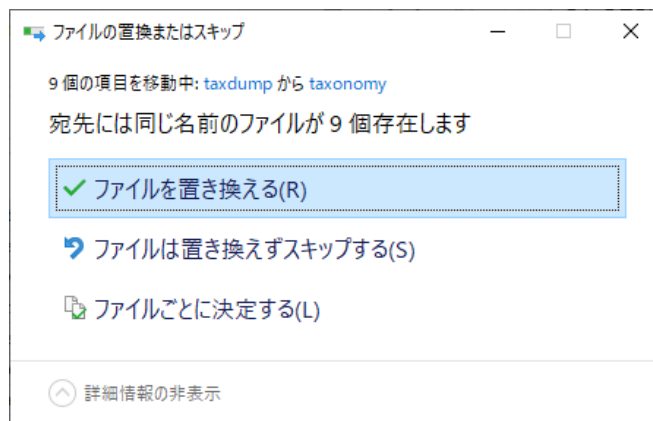
D. taxonomy フォルダ内のすべてのファイル→C:¥inetpub¥mascot¥taxonomy フォルダへ移動させます。まず、taxdump.tar.gz ファイルを解凍・展開して得られた taxdump フォルダを開きます。



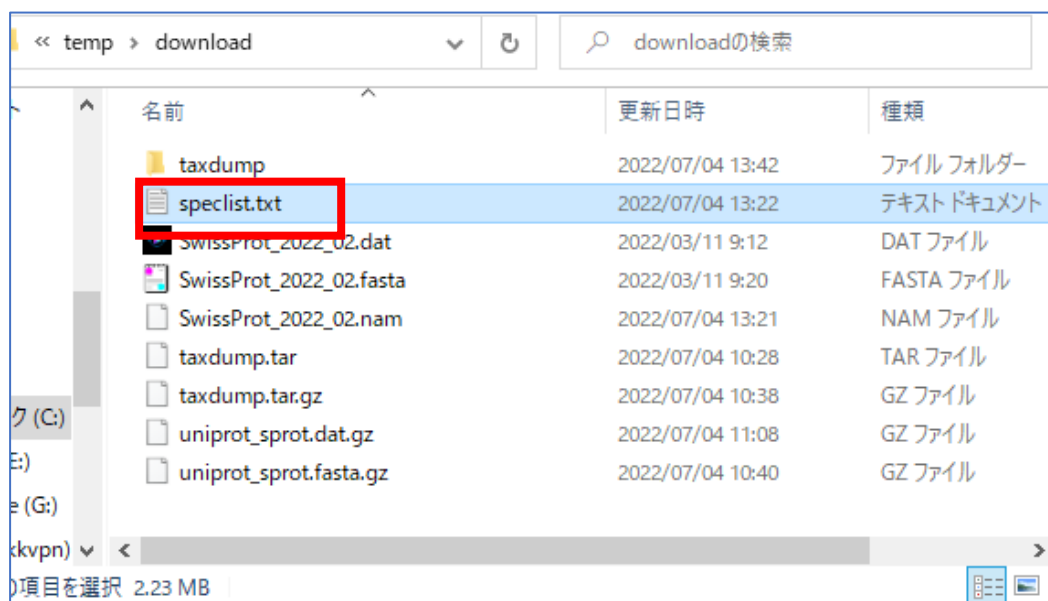
taxdump フォルダ内にあったファイル全てを、C:\inetpub\mascot\taxonomy に移動(ドラッグ&ドロップ、またはカット&ペースト)します。



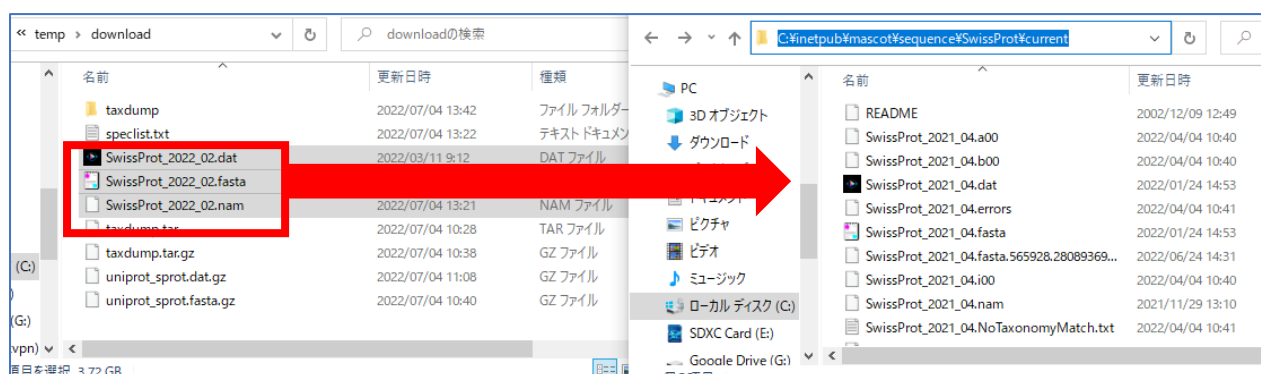
9 ファイルの置換を許可します。



E の speclist.txt ファイルも同様に C:\inetpub\mascot\taxonomy フォルダへ移動(ドラッグ&ドロップまたはカット&ペースト)します。



A～C の 3 ファイルについては必ず同時に [C:¥inetpub¥mascot¥sequence¥SwissProt¥current](#) フォルダに移動させてください。



■5 Database Statusの確認

Database Status 画面で SwissProt の Status 項目をご覧いただきながらデータベース再構築の進捗をご確認ください。

「step1.データベースの構築」→「step2.検索テスト」→「step3.使用可能 (In use)」という過程を経ます。Status の表示が「**Creating compressed files N%**」となっている場合、step1 のデータベース構築中です。また status 表示が「**Running 1st test**」の場合、step2 の検索テスト実行中です。これらの表記は特に問題ありません。次のステップに移行するまでお待ちください。

Status が「**Halted:Failed to map**」となっている場合、何かしらの理由でエラーになったことを示しています。まずは PC を再起動してもう一度ご確認ください。同じエラーが出た場合、status 項目のすぐ隣に現れているハイパーリンク「[compression warning](#)」「[Check error log](#)」をクリックし、エラー内容を確認してわかる範囲でご対処頂き、ご不明な点がございましたら弊社までご連絡ください。