

AlphaFold 2.1.1 (2021/11/6)

ウェブページ(コード)

<https://github.com/deepmind/alphafold>

バージョン

2.1.1 (2021/11/6; データベースは 2021/11/5 に取得)

リファレンス

- RCCS での初期導入時のメモ(2021/7/20): <https://ccportal.ims.ac.jp/node/2945>
- 2021/8/19 更新時のメモ: <https://ccportal.ims.ac.jp/node/2971>
- 論文引用情報(公式ページ中): <https://github.com/deepmind/alphafold#citing-this-work>
 - multimer についての文献が一つ追加されています。

導入に関するメモ

- alphafold 本体の導入については以前と同じ手続きで実行しています。
 - alphafold/common/residue_constants.py の修正は不要となっています。

python 環境更新

2021/7 の初期導入時に準備した python 環境を再利用しています。

ただし、このバージョンでは pandas が入っていなかったため、今回追加で pandas を導入しています。すでに conda と pip のパッケージが混在した環境ではありましたが、あまり無理なく導入できる状況だったため、別環境を作ったりはせず、そのまま導入しています。既存の alphafold への影響も無いと想定しています。

データベースの更新

以下の 4 つを除くデータベースについては 2021/8/19 のものを流用しています。

- alphafold model parameters (multimer 用のモデル追加)
- PDB (mmCIF format)
- PDB seqres
 - これら 2 つは 2021/11/5 に[日本のミラーサイト](#)よりダウンロードしています
 - (同日のものを使う必要がある、との指示が公式ページにあります)
 - 2021/10/30 あたりまでのデータとなっているようです
- UniProt (新規導入)

なお、uniclust30 については、[2021/8/19 のバージョン](#)と同じく、少し新しい 2021/6 のデータベースを使っています。ご注意ください。

uniprot データベースに関するディスク負荷軽減対応(Lustre ファイルシステム限定の処理です)

```
[root@ccfep4 uniprot]$ lfs migrate -c 20 uniprot.fasta
```

実行スクリプトの更新

alphafold 本体側でオプション指定に変更があったため、実行スクリプトも更新しています。

[run_alphafold21x_rccs.sh](#)

実行サンプル(multimer)

multimer の場合には、-f で与えるファイルは multi-FASTA 形式で複数の配列を含む必要があります。

また、model_preset(-m) に multimer の指定も必要です。

```
#!/bin/sh
#PBS -l select=1:ncpus=12:mpiprocs=1:ompthreads=12:jobtype=core
#PBS -l walltime=72:00:00

# NOTE: sample for alphafold 2.1.x (downloaded Nov 2021)
# this may be incompatible with previous versions!

# at least 8 cpu cores will be requested internally.
# in this sample, we employ 12 cores to get enough amount of memory.
# not sure how much is necessary/required, though.
```

```

# note about available memory:
# Available memory amount is proportional to ncpus value.
# If you need more memory, please increase ncpus in the header.

if [ ! -z "${PBS_O_WORKDIR}" ]; then
    cd "${PBS_O_WORKDIR}"
fi

AF2ROOT=/local/apl/lx/alphafold2
RUNAF2=${AF2ROOT}/run_alphafold21x_rccs.sh

# pass "-a $AF2DIR" to $RUNAF2 if you want to change alphafold version
# AF2DIR=/local/apl/lx/alphafold2/alphafold-20211106

# load miniconda environment (where necessary binaries reside)
. ${AF2ROOT}/conda_init.sh

# Required:
# -o [output directory]
# -f [sequence file (FASTA)]

# Optional arguments:
# -a [alphafold code root]
#   (default: /local/apl/lx/alphafold2/alphafold-20211106)
# -d [database root]
#   (default: /local/apl/lx/alphafold2/databases-20211105)
# -t [max template date] (default: 2021-11-05)
#   in the original example 2020-05-14 was used (CASP14 target case?)
# -m [model_preset] (default: monomer)
#   model control. available values: monomer, monomer_casp14, monomer_ptm, multimer
# -p [db_preset] (default: full_dbs)
#   database control. available values: full_dbs, reduced_dbs
# -Q
#   change model names to _ptm version; additional scores will be available
#   in output file. pLDDT score will be in B-factor column of output pdbs.
#   (same as "-p monomer_ptm")
${RUNAF2} \
-o ./multimer_test/ \
-m multimer \
-f complex.fasta

```