Class 11: AlphaFold

Ava Limtiaco (PID: A18007672)

Here we read the results from AlphaFold and try to interpret all the models and quality score metrics.

```
library(bio3d)
pth <- "dimer_23119/"
pdb.files <- list.files(path=pth, full.names=TRUE, pattern=".pdb")
```

Align and supperpose all these models:

```
file.exists(pdb.files)
```

```
[1] TRUE TRUE TRUE TRUE TRUE
```

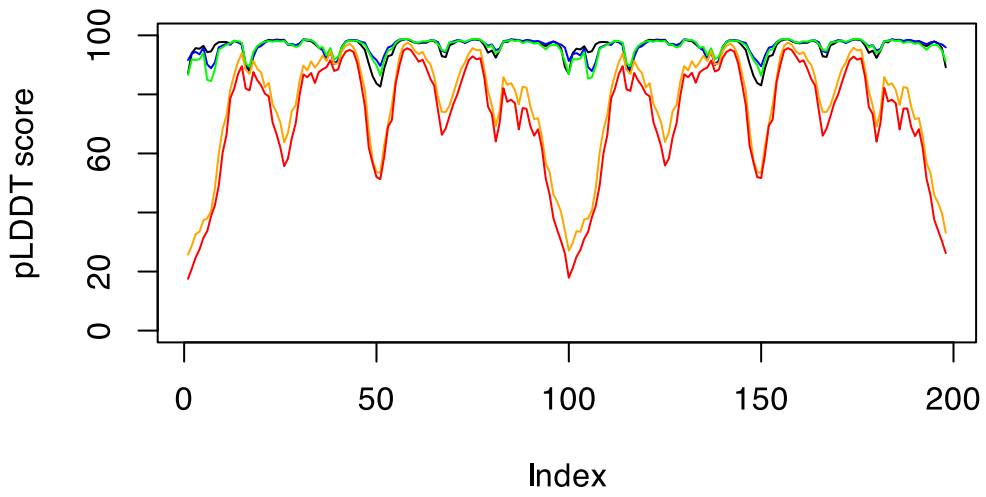
```
pdbs <- pdbaln(pdb.files, fit=TRUE, exefile="msa")
```

```
Reading PDB files:
dimer_23119//
dimer_23119_unrelaxed_rank_001_alphafold2_multimer_v3_model_5_seed_000.pdb
dimer_23119//
dimer_23119_unrelaxed_rank_002_alphafold2_multimer_v3_model_4_seed_000.pdb
dimer_23119//
dimer_23119_unrelaxed_rank_003_alphafold2_multimer_v3_model_1_seed_000.pdb
dimer_23119//
dimer_23119_unrelaxed_rank_004_alphafold2_multimer_v3_model_2_seed_000.pdb
dimer_23119//
dimer_23119_unrelaxed_rank_005_alphafold2_multimer_v3_model_3_seed_000.pdb
.....

Extracting sequences

pdb/seq:         1                              name:         dimer_23119//
dimer_23119_unrelaxed_rank_001_alphafold2_multimer_v3_model_5_seed_000.pdb
pdb/seq:         2                              name:         dimer_23119//
dimer_23119_unrelaxed_rank_002_alphafold2_multimer_v3_model_4_seed_000.pdb
pdb/seq:         3                              name:         dimer_23119//
dimer_23119_unrelaxed_rank_003_alphafold2_multimer_v3_model_1_seed_000.pdb
pdb/seq:         4                              name:         dimer_23119//
```

```
dimer_23119_unrelaxed_rank_004_alphafold2_multimer_v3_model_2_seed_000.pdb
pdb/seq:         5                              name:         dimer_23119//
dimer_23119_unrelaxed_rank_005_alphafold2_multimer_v3_model_3_seed_000.pdb
```

```
#library(bio3dview)
#view.pdbs(pdbs)
```

```
plot(pdbs$b[1,], typ="l", ylim=c(0,100), ylab="pLDDT score")
lines(pdbs$b[2,], typ="l", col="blue")
lines(pdbs$b[3, ], typ="l", col="green")
lines(pdbs$b[4, ], typ="l", col="orange")
lines(pdbs$b[5, ], typ="l", col="red")
```




```
pdbs$sse[1,]
```

```
NULL
```

AlphaFold returns its large alignemnt file used for analysis. Here we read this file and score conservation per position.

For homework: input all the rest of the page do find a gene project on alpha fold then predict.

```
```

```
library(bio3d)
mygene <- "AlphaFold2.ipynb/"
mygene.files <- list.files(path=pth, full.names=TRUE, pattern=".pdb")
```

```
file.exists(mygene.files)
```

```
[1] TRUE TRUE TRUE TRUE TRUE
```

```
pdbg <- pdbaln(mygene.files, fit=TRUE, exefile="msa")
```

```
Reading PDB files:
dimer_23119//
dimer_23119_unrelaxed_rank_001_alphafold2_multimer_v3_model_5_seed_000.pdb
dimer_23119//
dimer_23119_unrelaxed_rank_002_alphafold2_multimer_v3_model_4_seed_000.pdb
dimer_23119//
dimer_23119_unrelaxed_rank_003_alphafold2_multimer_v3_model_1_seed_000.pdb
dimer_23119//
dimer_23119_unrelaxed_rank_004_alphafold2_multimer_v3_model_2_seed_000.pdb
dimer_23119//
dimer_23119_unrelaxed_rank_005_alphafold2_multimer_v3_model_3_seed_000.pdb
.....

Extracting sequences

pdb/seq:         1                              name:         dimer_23119//
dimer_23119_unrelaxed_rank_001_alphafold2_multimer_v3_model_5_seed_000.pdb
pdb/seq:         2                              name:         dimer_23119//
dimer_23119_unrelaxed_rank_002_alphafold2_multimer_v3_model_4_seed_000.pdb
pdb/seq:         3                              name:         dimer_23119//
dimer_23119_unrelaxed_rank_003_alphafold2_multimer_v3_model_1_seed_000.pdb
pdb/seq:         4                              name:         dimer_23119//
dimer_23119_unrelaxed_rank_004_alphafold2_multimer_v3_model_2_seed_000.pdb
pdb/seq:         5                              name:         dimer_23119//
dimer_23119_unrelaxed_rank_005_alphafold2_multimer_v3_model_3_seed_000.pdb
```

Image of My Find a Gene : albumin.

```
#library(bio3dview)
#view.pdbs(pdbg)
```

```
rd <- rmsd(pdbg, fit=TRUE)
```

```
Warning in rmsd(pdbg, fit = TRUE): No indices provided, using the 198 non
NA positions
```

Drawing a heatmap fo Albumin:

```
library(pheatmap)
colnames(rd) <- paste0("m", 1:5)
rownames(rd) <- paste0("m", 1:5)
pheatmap(rd)
```