

Alpha fold analysis

Safiya Sayd (PID:A18027139)

Here we analysis our alpha fold structure prediction models. The input directory/folder comes from the colabfold server:

```
# Change this for YOUR results dir name
results_dir <- "hivpr_monomer_94b5b"
```

```
pdb_files <- list.files(path=results_dir,
                        pattern="*.pdb",
                        full.names = TRUE)

basename(pdb_files)
```

```
[1] "hivpr_monomer_94b5b_unrelaxed_rank_001_alphafold2_ptm_model_5_seed_000.pdb"
[2] "hivpr_monomer_94b5b_unrelaxed_rank_002_alphafold2_ptm_model_4_seed_000.pdb"
[3] "hivpr_monomer_94b5b_unrelaxed_rank_003_alphafold2_ptm_model_1_seed_000.pdb"
[4] "hivpr_monomer_94b5b_unrelaxed_rank_004_alphafold2_ptm_model_3_seed_000.pdb"
[5] "hivpr_monomer_94b5b_unrelaxed_rank_005_alphafold2_ptm_model_2_seed_000.pdb"
```

```
library(bio3d)
```

```
pdbbs <- pdbaln(pdb_files, fit=TRUE, exefile="msa")
```

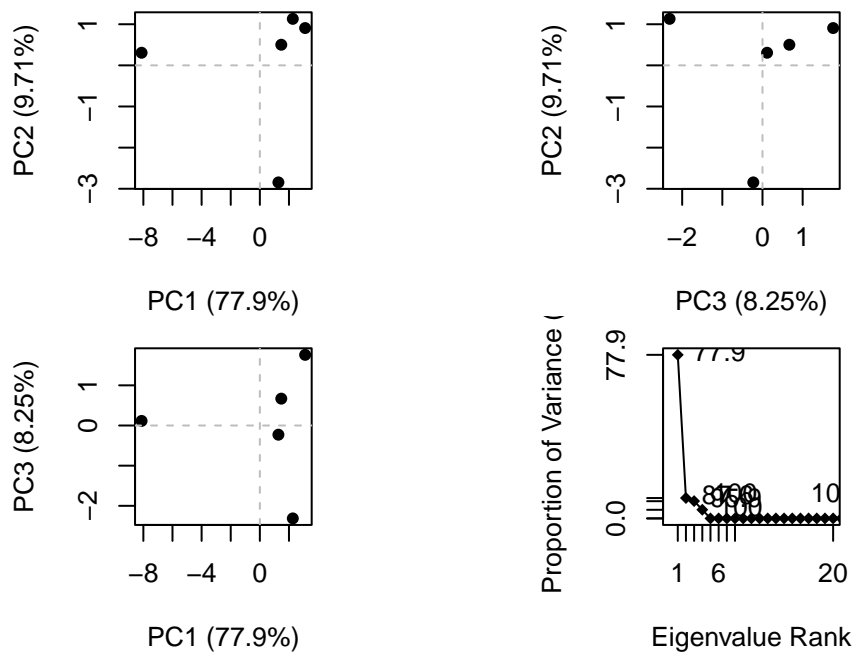
Reading PDB files:

```
hivpr_monomer_94b5b/hivpr_monomer_94b5b_unrelaxed_rank_001_alphafold2_ptm_model_5_seed_000.p
hivpr_monomer_94b5b/hivpr_monomer_94b5b_unrelaxed_rank_002_alphafold2_ptm_model_4_seed_000.p
hivpr_monomer_94b5b/hivpr_monomer_94b5b_unrelaxed_rank_003_alphafold2_ptm_model_1_seed_000.p
hivpr_monomer_94b5b/hivpr_monomer_94b5b_unrelaxed_rank_004_alphafold2_ptm_model_3_seed_000.p
hivpr_monomer_94b5b/hivpr_monomer_94b5b_unrelaxed_rank_005_alphafold2_ptm_model_2_seed_000.p
.....
```

Extracting sequences

```
pdb/seq: 1   name: hivpr_monomer_94b5b/hivpr_monomer_94b5b_unrelaxed_rank_001_alphafold2_ptm
pdb/seq: 2   name: hivpr_monomer_94b5b/hivpr_monomer_94b5b_unrelaxed_rank_002_alphafold2_ptm
pdb/seq: 3   name: hivpr_monomer_94b5b/hivpr_monomer_94b5b_unrelaxed_rank_003_alphafold2_ptm
pdb/seq: 4   name: hivpr_monomer_94b5b/hivpr_monomer_94b5b_unrelaxed_rank_004_alphafold2_ptm
pdb/seq: 5   name: hivpr_monomer_94b5b/hivpr_monomer_94b5b_unrelaxed_rank_005_alphafold2_ptm
```

```
pc<- pca(pdbbs)
plot(pc)
```



RMSD analysis

RMSD is a common measure of structural

```
rd <- rmsd(pdbbs, fit=T)
```

Warning in rmsd(pdbbs, fit = T): No indices provided, using the 99 non NA positions

```
rd
```

```
hivpr_monomer_94b5b_unrelaxed_rank_001_alphafold2_ptm_model_5_seed_000
hivpr_monomer_94b5b_unrelaxed_rank_002_alphafold2_ptm_model_4_seed_000
hivpr_monomer_94b5b_unrelaxed_rank_003_alphafold2_ptm_model_1_seed_000
hivpr_monomer_94b5b_unrelaxed_rank_004_alphafold2_ptm_model_3_seed_000
hivpr_monomer_94b5b_unrelaxed_rank_005_alphafold2_ptm_model_2_seed_000
hivpr_monomer_94b5b_unrelaxed_rank_001_alphafold2_ptm_model_5_seed_000
hivpr_monomer_94b5b_unrelaxed_rank_002_alphafold2_ptm_model_4_seed_000
hivpr_monomer_94b5b_unrelaxed_rank_003_alphafold2_ptm_model_1_seed_000
hivpr_monomer_94b5b_unrelaxed_rank_004_alphafold2_ptm_model_3_seed_000
hivpr_monomer_94b5b_unrelaxed_rank_005_alphafold2_ptm_model_2_seed_000
hivpr_monomer_94b5b_unrelaxed_rank_001_alphafold2_ptm_model_5_seed_000
hivpr_monomer_94b5b_unrelaxed_rank_002_alphafold2_ptm_model_4_seed_000
hivpr_monomer_94b5b_unrelaxed_rank_003_alphafold2_ptm_model_1_seed_000
hivpr_monomer_94b5b_unrelaxed_rank_004_alphafold2_ptm_model_3_seed_000
hivpr_monomer_94b5b_unrelaxed_rank_005_alphafold2_ptm_model_2_seed_000
hivpr_monomer_94b5b_unrelaxed_rank_001_alphafold2_ptm_model_5_seed_000
hivpr_monomer_94b5b_unrelaxed_rank_002_alphafold2_ptm_model_4_seed_000
hivpr_monomer_94b5b_unrelaxed_rank_003_alphafold2_ptm_model_1_seed_000
hivpr_monomer_94b5b_unrelaxed_rank_004_alphafold2_ptm_model_3_seed_000
hivpr_monomer_94b5b_unrelaxed_rank_005_alphafold2_ptm_model_2_seed_000
hivpr_monomer_94b5b_unrelaxed_rank_001_alphafold2_ptm_model_5_seed_000
hivpr_monomer_94b5b_unrelaxed_rank_002_alphafold2_ptm_model_4_seed_000
hivpr_monomer_94b5b_unrelaxed_rank_003_alphafold2_ptm_model_1_seed_000
hivpr_monomer_94b5b_unrelaxed_rank_004_alphafold2_ptm_model_3_seed_000
hivpr_monomer_94b5b_unrelaxed_rank_005_alphafold2_ptm_model_2_seed_000
```

```
library(pheatmap)
```

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

