Class11_Part 4_from_Class_10

Dhruv

Part 4 (11/6/24)

Comparative Structural Analysis

```
library(bio3d)
```

```
id <- "1ake_A"
```

```
aa <- get.seq(id)
```

Warning in get.seq(id): Removing existing file: seqs.fasta

Fetching... Please wait. Done.

```
aa
```

```
             1        .        .        .        .        .          60
pdb|1AKE|A    MRIILLGAPGAGKGTQAQFIMEKYGIPQISTGDMLRAAVKSGSELGKQAKDIMDAGKLVT
             1        .        .        .        .        .          60

             61       .        .        .        .        .         120
pdb|1AKE|A    DELVIALVKERIAQEDCRNGFLLDGFPRTIPQADAMKEAGINVDYVLEFDVPDELIVDRI
             61       .        .        .        .        .         120

             121      .        .        .        .        .         180
pdb|1AKE|A    VGRRVHAPSGRVYHVKFNPPKVEGKDDVTGEELTTRKDDQEETVRKRLVEYHQMTAPLIG
             121      .        .        .        .        .         180

             181      .        .        .  214
pdb|1AKE|A    YYSKEAEAGNTKYAKVDGTKPVAEVRADLEKILG
```

```
            181          .        .           .  214
```

Call:
  read.fasta(file = outfile)

Class:
  fasta

Alignment dimensions:
  1 sequence rows; 214 position columns (214 non-gap, 0 gap)

+ attr: id, ali, call

```
b <- blast.pdb(aa)
```

```
 Searching ... please wait (updates every 5 seconds) RID = JU4H3RXM016
 ...
 Reporting 85 hits
```
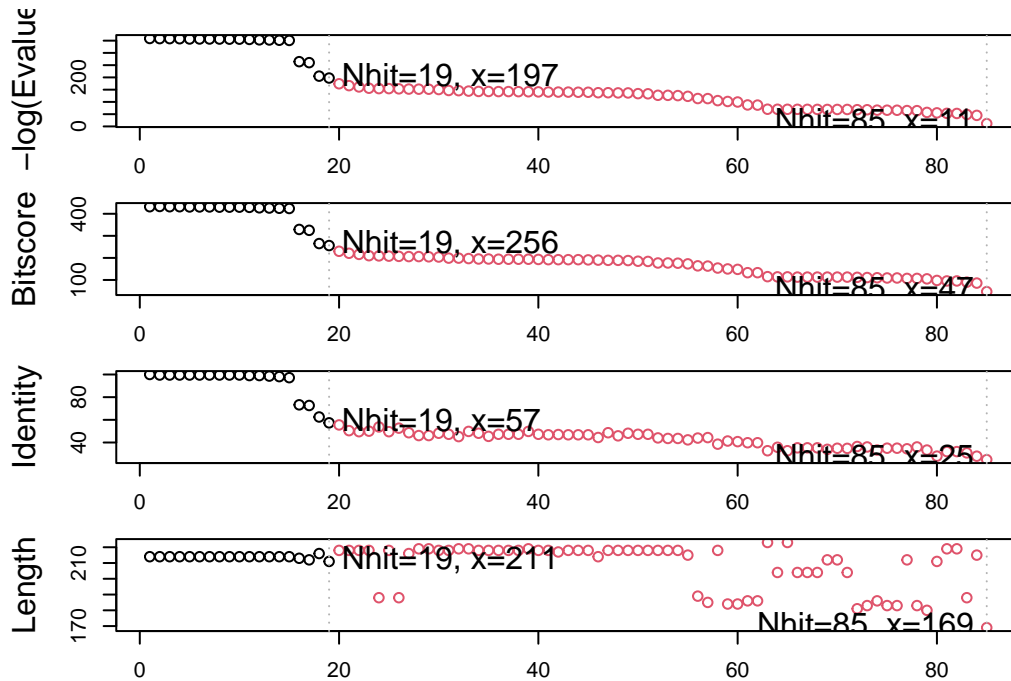
```
head(b$hit.tbl)
```

```
         queryid subjectids identity alignmentlength mismatches gapopens q.start
1 Query_4414991     1AKE_A  100.000             214          0        0       1
2 Query_4414991     8BQF_A   99.533             214          1        0       1
3 Query_4414991     4X8M_A   99.533             214          1        0       1
4 Query_4414991     6S36_A   99.533             214          1        0       1
5 Query_4414991     8Q2B_A   99.533             214          1        0       1
6 Query_4414991     8RJ9_A   99.533             214          1        0       1
  q.end s.start s.end    evalue bitscore positives mlog.evalue pdb.id    acc
1   214       1   214 1.58e-156      432    100.00    358.7458 1AKE_A 1AKE_A
2   214      21   234 2.58e-156      433    100.00    358.2555 8BQF_A 8BQF_A
3   214       1   214 2.82e-156      432    100.00    358.1665 4X8M_A 4X8M_A
4   214       1   214 4.14e-156      432    100.00    357.7826 6S36_A 6S36_A
5   214       1   214 1.10e-155      431     99.53    356.8054 8Q2B_A 8Q2B_A
6   214       1   214 1.10e-155      431     99.53    356.8054 8RJ9_A 8RJ9_A
```

```
hits <- plot(b)
```

```
  * Possible cutoff values:    197 11
          Yielding Nhits:    19 85
```

```
* Chosen cutoff value of:     197
        Yielding Nhits:       19
```


What is inside "hits"

```
attributes(hits)
```

```
$names
[1] "hits"   "pdb.id" "acc"    "inds"

$class
[1] "blast"
```

What are the top hits form our blast results

```
hits$pdb.id
```

```
 [1] "1AKE_A" "8BQF_A" "4X8M_A" "6S36_A" "8Q2B_A" "8RJ9_A" "6RZE_A" "4X8H_A"
 [9] "3HPR_A" "1E4V_A" "5EJE_A" "1E4Y_A" "3X2S_A" "6HAP_A" "6HAM_A" "4K46_A"
[17] "4NP6_A" "3GMT_A" "4PZL_A"
```

Download PDB files

```r
files <- get.pdb(hits$pdb.id, path = "pdbs", split = TRUE, gzip = TRUE)
```

Warning in get.pdb(hits$pdb.id, path = "pdbs", split = TRUE, gzip = TRUE):
pdbs/1AKE.pdb.gz exists. Skipping download

Warning in get.pdb(hits$pdb.id, path = "pdbs", split = TRUE, gzip = TRUE):
pdbs/8BQF.pdb.gz exists. Skipping download

Warning in get.pdb(hits$pdb.id, path = "pdbs", split = TRUE, gzip = TRUE):
pdbs/4X8M.pdb.gz exists. Skipping download

Warning in get.pdb(hits$pdb.id, path = "pdbs", split = TRUE, gzip = TRUE):
pdbs/6S36.pdb.gz exists. Skipping download

Warning in get.pdb(hits$pdb.id, path = "pdbs", split = TRUE, gzip = TRUE):
pdbs/8Q2B.pdb.gz exists. Skipping download

Warning in get.pdb(hits$pdb.id, path = "pdbs", split = TRUE, gzip = TRUE):
pdbs/8RJ9.pdb.gz exists. Skipping download

Warning in get.pdb(hits$pdb.id, path = "pdbs", split = TRUE, gzip = TRUE):
pdbs/6RZE.pdb.gz exists. Skipping download

Warning in get.pdb(hits$pdb.id, path = "pdbs", split = TRUE, gzip = TRUE):
pdbs/4X8H.pdb.gz exists. Skipping download

Warning in get.pdb(hits$pdb.id, path = "pdbs", split = TRUE, gzip = TRUE):
pdbs/3HPR.pdb.gz exists. Skipping download

Warning in get.pdb(hits$pdb.id, path = "pdbs", split = TRUE, gzip = TRUE):
pdbs/1E4V.pdb.gz exists. Skipping download

Warning in get.pdb(hits$pdb.id, path = "pdbs", split = TRUE, gzip = TRUE):
pdbs/5EJE.pdb.gz exists. Skipping download

Warning in get.pdb(hits$pdb.id, path = "pdbs", split = TRUE, gzip = TRUE):
pdbs/1E4Y.pdb.gz exists. Skipping download

```
Warning in get.pdb(hits$pdb.id, path = "pdbs", split = TRUE, gzip = TRUE):
pdbs/3X2S.pdb.gz exists. Skipping download

Warning in get.pdb(hits$pdb.id, path = "pdbs", split = TRUE, gzip = TRUE):
pdbs/6HAP.pdb.gz exists. Skipping download

Warning in get.pdb(hits$pdb.id, path = "pdbs", split = TRUE, gzip = TRUE):
pdbs/6HAM.pdb.gz exists. Skipping download

Warning in get.pdb(hits$pdb.id, path = "pdbs", split = TRUE, gzip = TRUE):
pdbs/4K46.pdb.gz exists. Skipping download

Warning in get.pdb(hits$pdb.id, path = "pdbs", split = TRUE, gzip = TRUE):
pdbs/4NP6.pdb.gz exists. Skipping download

Warning in get.pdb(hits$pdb.id, path = "pdbs", split = TRUE, gzip = TRUE):
pdbs/3GMT.pdb.gz exists. Skipping download

Warning in get.pdb(hits$pdb.id, path = "pdbs", split = TRUE, gzip = TRUE):
pdbs/4PZL.pdb.gz exists. Skipping download


  |
  |                                                                      |   0%
  |
  |====                                                                  |   5%
  |
  |=======                                                               |  11%
  |
  |==========                                                            |  16%
  |
  |==============                                                        |  21%
  |
  |=================                                                     |  26%
  |
  |=====================                                                 |  32%
  |
  |========================                                              |  37%
  |
  |===========================                                           |  42%
```

```
  |
  |================================                                    |  47%
  |
  |===================================                                 |  53%
  |
  |=======================================                             |  58%
  |
  |==========================================                          |  63%
  |
  |=============================================                       |  68%
  |
  |=================================================                   |  74%
  |
  |====================================================                |  79%
  |
  |========================================================            |  84%
  |
  |============================================================        |  89%
  |
  |=================================================================   |  95%
  |
  |====================================================================| 100%
```

I have now found and downloaded all ADK structures in the PDB database. Viewing them is difficult and needs to be alligned and then superimposed. This is evident when we use MolViewer to load upp all the structures mentioned before.

Install BiocManager package from CRAN. use `BiocManager::install()` to install any package from bioconductor.

Align releated PDBs

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

```
Reading PDB files:
pdbs/split_chain/1AKE_A.pdb
pdbs/split_chain/8BQF_A.pdb
pdbs/split_chain/4X8M_A.pdb
pdbs/split_chain/6S36_A.pdb
pdbs/split_chain/8Q2B_A.pdb
pdbs/split_chain/8RJ9_A.pdb
pdbs/split_chain/6RZE_A.pdb
pdbs/split_chain/4X8H_A.pdb
```

```
pdbs/split_chain/3HPR_A.pdb
pdbs/split_chain/1E4V_A.pdb
pdbs/split_chain/5EJE_A.pdb
pdbs/split_chain/1E4Y_A.pdb
pdbs/split_chain/3X2S_A.pdb
pdbs/split_chain/6HAP_A.pdb
pdbs/split_chain/6HAM_A.pdb
pdbs/split_chain/4K46_A.pdb
pdbs/split_chain/4NP6_A.pdb
pdbs/split_chain/3GMT_A.pdb
pdbs/split_chain/4PZL_A.pdb
    PDB has ALT records, taking A only, rm.alt=TRUE
.   PDB has ALT records, taking A only, rm.alt=TRUE
..   PDB has ALT records, taking A only, rm.alt=TRUE
.   PDB has ALT records, taking A only, rm.alt=TRUE
.   PDB has ALT records, taking A only, rm.alt=TRUE
.   PDB has ALT records, taking A only, rm.alt=TRUE
..   PDB has ALT records, taking A only, rm.alt=TRUE
..   PDB has ALT records, taking A only, rm.alt=TRUE
....    PDB has ALT records, taking A only, rm.alt=TRUE
.   PDB has ALT records, taking A only, rm.alt=TRUE
....


Extracting sequences

pdb/seq: 1   name: pdbs/split_chain/1AKE_A.pdb
    PDB has ALT records, taking A only, rm.alt=TRUE
pdb/seq: 2   name: pdbs/split_chain/8BQF_A.pdb
    PDB has ALT records, taking A only, rm.alt=TRUE
pdb/seq: 3   name: pdbs/split_chain/4X8M_A.pdb
pdb/seq: 4   name: pdbs/split_chain/6S36_A.pdb
    PDB has ALT records, taking A only, rm.alt=TRUE
pdb/seq: 5   name: pdbs/split_chain/8Q2B_A.pdb
    PDB has ALT records, taking A only, rm.alt=TRUE
pdb/seq: 6   name: pdbs/split_chain/8RJ9_A.pdb
    PDB has ALT records, taking A only, rm.alt=TRUE
pdb/seq: 7   name: pdbs/split_chain/6RZE_A.pdb
    PDB has ALT records, taking A only, rm.alt=TRUE
pdb/seq: 8   name: pdbs/split_chain/4X8H_A.pdb
pdb/seq: 9   name: pdbs/split_chain/3HPR_A.pdb
    PDB has ALT records, taking A only, rm.alt=TRUE
pdb/seq: 10   name: pdbs/split_chain/1E4V_A.pdb
pdb/seq: 11   name: pdbs/split_chain/5EJE_A.pdb
```

```
    PDB has ALT records, taking A only, rm.alt=TRUE
pdb/seq: 12   name: pdbs/split_chain/1E4Y_A.pdb
pdb/seq: 13   name: pdbs/split_chain/3X2S_A.pdb
pdb/seq: 14   name: pdbs/split_chain/6HAP_A.pdb
pdb/seq: 15   name: pdbs/split_chain/6HAM_A.pdb
    PDB has ALT records, taking A only, rm.alt=TRUE
pdb/seq: 16   name: pdbs/split_chain/4K46_A.pdb
    PDB has ALT records, taking A only, rm.alt=TRUE
pdb/seq: 17   name: pdbs/split_chain/4NP6_A.pdb
pdb/seq: 18   name: pdbs/split_chain/3GMT_A.pdb
pdb/seq: 19   name: pdbs/split_chain/4PZL_A.pdb
```

pdbs

```
                                1         .         .         .        40
[Truncated_Name:1]1AKE_A.pdb    ----------MRIILLGAPGAGKGTQAQFIMEKYGIPQIS
[Truncated_Name:2]8BQF_A.pdb    ----------MRIILLGAPGAGKGTQAQFIMEKYGIPQIS
[Truncated_Name:3]4X8M_A.pdb    ----------MRIILLGAPGAGKGTQAQFIMEKYGIPQIS
[Truncated_Name:4]6S36_A.pdb    ----------MRIILLGAPGAGKGTQAQFIMEKYGIPQIS
[Truncated_Name:5]8Q2B_A.pdb    ----------MRIILLGAPGAGKGTQAQFIMEKYGIPQIS
[Truncated_Name:6]8RJ9_A.pdb    ----------MRIILLGAPGAGKGTQAQFIMEKYGIPQIS
[Truncated_Name:7]6RZE_A.pdb    ----------MRIILLGAPGAGKGTQAQFIMEKYGIPQIS
[Truncated_Name:8]4X8H_A.pdb    ----------MRIILLGAPGAGKGTQAQFIMEKYGIPQIS
[Truncated_Name:9]3HPR_A.pdb    ----------MRIILLGAPGAGKGTQAQFIMEKYGIPQIS
[Truncated_Name:10]1E4V_A.pdb   ----------MRIILLGAPVAGKGTQAQFIMEKYGIPQIS
[Truncated_Name:11]5EJE_A.pdb   ----------MRIILLGAPGAGKGTQAQFIMEKYGIPQIS
[Truncated_Name:12]1E4Y_A.pdb   ----------MRIILLGALVAGKGTQAQFIMEKYGIPQIS
[Truncated_Name:13]3X2S_A.pdb   ----------MRIILLGAPGAGKGTQAQFIMEKYGIPQIS
[Truncated_Name:14]6HAP_A.pdb   ----------MRIILLGAPGAGKGTQAQFIMEKYGIPQIS
[Truncated_Name:15]6HAM_A.pdb   ----------MRIILLGAPGAGKGTQAQFIMEKYGIPQIS
[Truncated_Name:16]4K46_A.pdb   ----------MRIILLGAPGAGKGTQAQFIMAKFGIPQIS
[Truncated_Name:17]4NP6_A.pdb   --------NAMRIILLGAPGAGKGTQAQFIMEKFGIPQIS
[Truncated_Name:18]3GMT_A.pdb   ----------MRLILLGAPGAGKGTQANFIKEKFGIPQIS
[Truncated_Name:19]4PZL_A.pdb   TENLYFQSNAMRIILLGAPGAGKGTQAKIIEQKYNIAHIS
                                **^***** ******* * *^ * **
                                1         .         .         .        40

                                41        .         .         .        80
[Truncated_Name:1]1AKE_A.pdb    TGDMLRAAVKSGSELGKQAKDIMDAGKLVTDELVIALVKE
[Truncated_Name:2]8BQF_A.pdb    TGDMLRAAVKSGSELGKQAKDIMDAGKLVTDELVIALVKE
[Truncated_Name:3]4X8M_A.pdb    TGDMLRAAVKSGSELGKQAKDIMDAGKLVTDELVIALVKE
[Truncated_Name:4]6S36_A.pdb    TGDMLRAAVKSGSELGKQAKDIMDAGKLVTDELVIALVKE
```

8

```
[Truncated_Name:5]8Q2B_A.pdb     TGDMLRAAVKSGSELGKQAKDIMDAGKLVTDELVIALVKE
[Truncated_Name:6]8RJ9_A.pdb     TGDMLRAAVKSGSELGKQAKDIMDAGKLVTDELVIALVKE
[Truncated_Name:7]6RZE_A.pdb     TGDMLRAAVKSGSELGKQAKDIMDAGKLVTDELVIALVKE
[Truncated_Name:8]4X8H_A.pdb     TGDMLRAAVKSGSELGKQAKDIMDAGKLVTDELVIALVKE
[Truncated_Name:9]3HPR_A.pdb     TGDMLRAAVKSGSELGKQAKDIMDAGKLVTDELVIALVKE
[Truncated_Name:10]1E4V_A.pdb    TGDMLRAAVKSGSELGKQAKDIMDAGKLVTDELVIALVKE
[Truncated_Name:11]5EJE_A.pdb    TGDMLRAAVKSGSELGKQAKDIMDACKLVTDELVIALVKE
[Truncated_Name:12]1E4Y_A.pdb    TGDMLRAAVKSGSELGKQAKDIMDAGKLVTDELVIALVKE
[Truncated_Name:13]3X2S_A.pdb    TGDMLRAAVKSGSELGKQAKDIMDCGKLVTDELVIALVKE
[Truncated_Name:14]6HAP_A.pdb    TGDMLRAAVKSGSELGKQAKDIMDAGKLVTDELVIALVRE
[Truncated_Name:15]6HAM_A.pdb    TGDMLRAAIKSGSELGKQAKDIMDAGKLVTDEIIIALVKE
[Truncated_Name:16]4K46_A.pdb    TGDMLRAAIKAGTELGKQAKSVIDAGQLVSDDIILGLVKE
[Truncated_Name:17]4NP6_A.pdb    TGDMLRAAIKAGTELGKQAKAVIDAGQLVSDDIILGLIKE
[Truncated_Name:18]3GMT_A.pdb    TGDMLRAAVKAGTPLGVEAKTYMDEGKLVPDSLIIGLVKE
[Truncated_Name:19]4PZL_A.pdb    TGDMIRETIKSGSALGQELKKVLDAGELVSDEFIIKIVKD
                                 ****^*  ^* *^ **   * ^*   ** *  ^^ ^^^^
                             41           .         .         .        80


                             81           .         .         .       120
[Truncated_Name:1]1AKE_A.pdb     RIAQEDCRNGFLLDGFPRTIPQADAMKEAGINVDYVLEFD
[Truncated_Name:2]8BQF_A.pdb     RIAQE----GFLLDGFPRTIPQADAMKEAGINVDYVIEFD
[Truncated_Name:3]4X8M_A.pdb     RIAQEDCRNGFLLDGFPRTIPQADAMKEAGINVDYVLEFD
[Truncated_Name:4]6S36_A.pdb     RIAQEDCRNGFLLDGFPRTIPQADAMKEAGINVDYVLEFD
[Truncated_Name:5]8Q2B_A.pdb     RIAQEDCRNGFLLDGFPRTIPQADAMKEAGINVDYVLEFD
[Truncated_Name:6]8RJ9_A.pdb     RIAQEDCRNGFLLAGFPRTIPQADAMKEAGINVDYVLEFD
[Truncated_Name:7]6RZE_A.pdb     RIAQEDCRNGFLLDGFPRTIPQADAMKEAGINVDYVLEFD
[Truncated_Name:8]4X8H_A.pdb     RIAQEDCRNGFLLDGFPRTIPQADAMKEAGINVDYVLEFD
[Truncated_Name:9]3HPR_A.pdb     RIAQEDCRNGFLLDGFPRTIPQADAMKEAGINVDYVLEFD
[Truncated_Name:10]1E4V_A.pdb    RIAQEDCRNGFLLDGFPRTIPQADAMKEAGINVDYVLEFD
[Truncated_Name:11]5EJE_A.pdb    RIAQEDCRNGFLLDGFPRTIPQADAMKEAGINVDYVLEFD
[Truncated_Name:12]1E4Y_A.pdb    RIAQEDCRNGFLLDGFPRTIPQADAMKEAGINVDYVLEFD
[Truncated_Name:13]3X2S_A.pdb    RIAQEDSRNGFLLDGFPRTIPQADAMKEAGINVDYVLEFD
[Truncated_Name:14]6HAP_A.pdb    RICQEDSRNGFLLDGFPRTIPQADAMKEAGINVDYVLEFD
[Truncated_Name:15]6HAM_A.pdb    RICQEDSRNGFLLDGFPRTIPQADAMKEAGINVDYVLEFD
[Truncated_Name:16]4K46_A.pdb    RIAQDDCAKGFLLDGFPRTIPQADGLKEVGVVVDYVIEFD
[Truncated_Name:17]4NP6_A.pdb    RIAQADCEKGFLLDGFPRTIPQADGLKEMGINVDYVIEFD
[Truncated_Name:18]3GMT_A.pdb    RLKEADCANGYLFDGFPRTIAQADAMKEAGVAIDYVLEID
[Truncated_Name:19]4PZL_A.pdb    RISKNDCNNGFLLDGVPRTIPQAQELDKLGVNIDYIVEVD
                                 *^       *^*  * **** **  ^    *^ ^**^^* *
                             81           .         .         .       120


                             121          .         .         .       160
[Truncated_Name:1]1AKE_A.pdb     VPDELIVDRIVGRRVHAPSGRVYHVKFNPPKVEGKDDVTG
```

9

```
[Truncated_Name:2]8BQF_A.pdb     VPDELIVDRIVGRRVHAPSGRVYHVKFNPPKVEGKDDVTG
[Truncated_Name:3]4X8M_A.pdb     VPDELIVDRIVGRRVHAPSGRVYHVKFNPPKVEGKDDVTG
[Truncated_Name:4]6S36_A.pdb     VPDELIVDKIVGRRVHAPSGRVYHVKFNPPKVEGKDDVTG
[Truncated_Name:5]8Q2B_A.pdb     VPDELIVDRIVGRRVHAPSGRVYHVKFNPPKVEGKDDVTG
[Truncated_Name:6]8RJ9_A.pdb     VPDELIVDRIVGRRVHAPSGRVYHVKFNPPKVEGKDDVTG
[Truncated_Name:7]6RZE_A.pdb     VPDELIVDAIVGRRVHAPSGRVYHVKFNPPKVEGKDDVTG
[Truncated_Name:8]4X8H_A.pdb     VPDELIVDRIVGRRVHAPSGRVYHVKFNPPKVEGKDDVTG
[Truncated_Name:9]3HPR_A.pdb     VPDELIVDRIVGRRVHAPSGRVYHVKFNPPKVEGKDDGTG
[Truncated_Name:10]1E4V_A.pdb    VPDELIVDRIVGRRVHAPSGRVYHVKFNPPKVEGKDDVTG
[Truncated_Name:11]5EJE_A.pdb    VPDELIVDRIVGRRVHAPSGRVYHVKFNPPKVEGKDDVTG
[Truncated_Name:12]1E4Y_A.pdb    VPDELIVDRIVGRRVHAPSGRVYHVKFNPPKVEGKDDVTG
[Truncated_Name:13]3X2S_A.pdb    VPDELIVDRIVGRRVHAPSGRVYHVKFNPPKVEGKDDVTG
[Truncated_Name:14]6HAP_A.pdb    VPDELIVDRIVGRRVHAPSGRVYHVKFNPPKVEGKDDVTG
[Truncated_Name:15]6HAM_A.pdb    VPDELIVDRIVGRRVHAPSGRVYHVKFNPPKVEGKDDVTG
[Truncated_Name:16]4K46_A.pdb    VADSVIVERMAGRRAHLASGRTYHNVYNPPKVEGKDDVTG
[Truncated_Name:17]4NP6_A.pdb    VADDVIVERMAGRRAHLPSGRTYHVVYNPPKVEGKDDVTG
[Truncated_Name:18]3GMT_A.pdb    VPFSEIIERMSGRRTHPASGRTYHVKFNPPKVEGKDDVTG
[Truncated_Name:19]4PZL_A.pdb    VADNLLIERITGRRIHPASGRTYHTKFNPPKVADKDDVTG
                                 *      ^^^ ^ *** *  *** ** ^*****  *** **
                             121          .         .         .         160


                             161          .         .         .         200
[Truncated_Name:1]1AKE_A.pdb     EELTTRKDDQEETVRKRLVEYHQMTAPLIGYYSKEAEAGN
[Truncated_Name:2]8BQF_A.pdb     EELTTRKDDQEETVRKRLVEYHQMTAPLIGYYSKEAEAGN
[Truncated_Name:3]4X8M_A.pdb     EELTTRKDDQEETVRKRLVEWHQMTAPLIGYYSKEAEAGN
[Truncated_Name:4]6S36_A.pdb     EELTTRKDDQEETVRKRLVEYHQMTAPLIGYYSKEAEAGN
[Truncated_Name:5]8Q2B_A.pdb     EELTTRKADQEETVRKRLVEYHQMTAPLIGYYSKEAEAGN
[Truncated_Name:6]8RJ9_A.pdb     EELTTRKDDQEETVRKRLVEYHQMTAPLIGYYSKEAEAGN
[Truncated_Name:7]6RZE_A.pdb     EELTTRKDDQEETVRKRLVEYHQMTAPLIGYYSKEAEAGN
[Truncated_Name:8]4X8H_A.pdb     EELTTRKDDQEETVRKRLVEYHQMTAALIGYYSKEAEAGN
[Truncated_Name:9]3HPR_A.pdb     EELTTRKDDQEETVRKRLVEYHQMTAPLIGYYSKEAEAGN
[Truncated_Name:10]1E4V_A.pdb    EELTTRKDDQEETVRKRLVEYHQMTAPLIGYYSKEAEAGN
[Truncated_Name:11]5EJE_A.pdb    EELTTRKDDQEECVRKRLVEYHQMTAPLIGYYSKEAEAGN
[Truncated_Name:12]1E4Y_A.pdb    EELTTRKDDQEETVRKRLVEYHQMTAPLIGYYSKEAEAGN
[Truncated_Name:13]3X2S_A.pdb    EELTTRKDDQEETVRKRLCEYHQMTAPLIGYYSKEAEAGN
[Truncated_Name:14]6HAP_A.pdb    EELTTRKDDQEETVRKRLVEYHQMTAPLIGYYSKEAEAGN
[Truncated_Name:15]6HAM_A.pdb    EELTTRKDDQEETVRKRLVEYHQMTAPLIGYYSKEAEAGN
[Truncated_Name:16]4K46_A.pdb    EDLVIREDDKEETVLARLGVYHNQTAPLIAYYGKEAEAGN
[Truncated_Name:17]4NP6_A.pdb    EDLVIREDDKEETVRARLNVYHTQTAPLIEYYGKEAAAGK
[Truncated_Name:18]3GMT_A.pdb    EPLVQRDDDKEETVKKRLDVYEAQTKPLITYYGDWARRGA
[Truncated_Name:19]4PZL_A.pdb    EPLITRTDDNEDTVKQRLSVYHAQTAKLIDFYRNFSSTNT
                                 * *  *   * *^ *  **   ^    *  ** ^*
                             161          .         .         .         200
```

10

```
                                   201        .         .       227
[Truncated_Name:1]1AKE_A.pdb      T--KYAKVDGTKPVAEVRADLEKILG-
[Truncated_Name:2]8BQF_A.pdb      T--KYAKVDGTKPVAEVRADLEKIL--
[Truncated_Name:3]4X8M_A.pdb      T--KYAKVDGTKPVAEVRADLEKILG-
[Truncated_Name:4]6S36_A.pdb      T--KYAKVDGTKPVAEVRADLEKILG-
[Truncated_Name:5]8Q2B_A.pdb      T--KYAKVDGTKPVAEVRADLEKILG-
[Truncated_Name:6]8RJ9_A.pdb      T--KYAKVDGTKPVAEVRADLEKILG-
[Truncated_Name:7]6RZE_A.pdb      T--KYAKVDGTKPVAEVRADLEKILG-
[Truncated_Name:8]4X8H_A.pdb      T--KYAKVDGTKPVAEVRADLEKILG-
[Truncated_Name:9]3HPR_A.pdb      T--KYAKVDGTKPVAEVRADLEKILG-
[Truncated_Name:10]1E4V_A.pdb     T--KYAKVDGTKPVAEVRADLEKILG-
[Truncated_Name:11]5EJE_A.pdb     T--KYAKVDGTKPVAEVRADLEKILG-
[Truncated_Name:12]1E4Y_A.pdb     T--KYAKVDGTKPVAEVRADLEKILG-
[Truncated_Name:13]3X2S_A.pdb     T--KYAKVDGTKPVAEVRADLEKILG-
[Truncated_Name:14]6HAP_A.pdb     T--KYAKVDGTKPVCEVRADLEKILG-
[Truncated_Name:15]6HAM_A.pdb     T--KYAKVDGTKPVCEVRADLEKILG-
[Truncated_Name:16]4K46_A.pdb     T--QYLKFDGTKAVAEVSAELEKALA-
[Truncated_Name:17]4NP6_A.pdb     T--QYLKFDGTKQVSEVSADIAKALA-
[Truncated_Name:18]3GMT_A.pdb     E-------NGLKAPA-----YRKISG-
[Truncated_Name:19]4PZL_A.pdb     KIPKYIKINGDQAVEKVSQDIFDQLNK
                                              *
                                   201        .         .       227

Call:
  pdbaln(files = files, fit = TRUE, exefile = "msa")

Class:
  pdbs, fasta

Alignment dimensions:
  19 sequence rows; 227 position columns (199 non-gap, 28 gap)

+ attr: xyz, resno, b, chain, id, ali, resid, sse, call
```
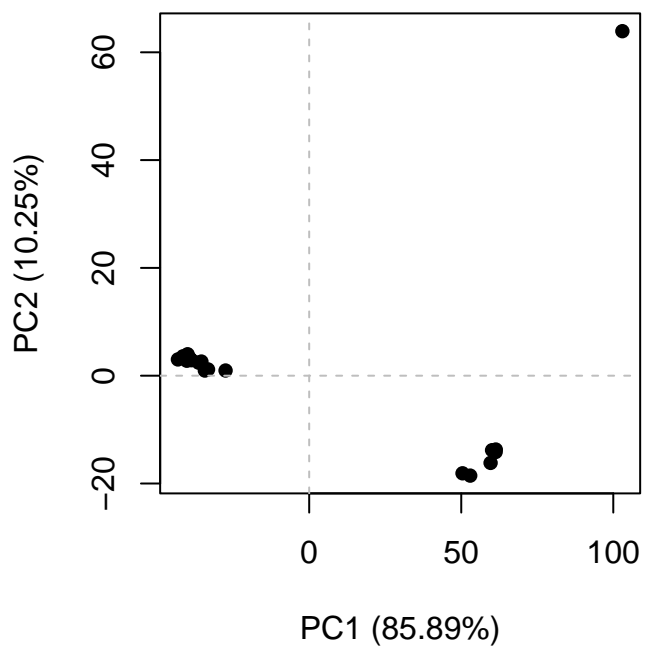
Now that we have our sequences aligned, we can now do a PCA

```
pc <- pca(pdbs)
plot(pc, pc.axes = c(1:2))
```

To examine what the PC is capturing, we can plot the loadings or make a movie (trajector) of moving along PC1

```
mktrj(pc, pc=1, file="pc1.pdb")
```