

Recapitulation

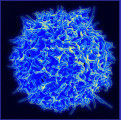
Benchmarking

Milestones

gliphR (turboGliph)

Monthly meeting

April 2023



Recapitulation

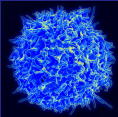
Benchmarking

Milestones

1 Recapitulation

2 Benchmarking

3 Milestones



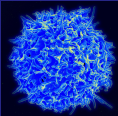
Status after last meeting

Recapitulation

Benchmarking

Milestones

- turboGliph refactored to gliphR (working title)



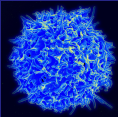
Status after last meeting

Recapitulation

Benchmarking

Milestones

- turboGliph refactored to gliphR (working title)
- 3 variants:



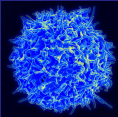
Status after last meeting

Recapitulation

Benchmarking

Milestones

- turboGliph refactored to gliphR (working title)
- 3 variants: gliph1,



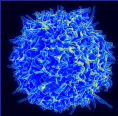
Status after last meeting

Recapitulation

Benchmarking

Milestones

- turboGliph refactored to gliphR (working title)
- 3 variants: gliph1, gliph2, and



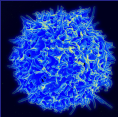
Status after last meeting

Recapitulation

Benchmarking

Milestones

- turboGliph refactored to gliphR (working title)
- 3 variants: gliph1, gliph2, and gliph3



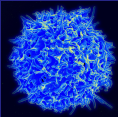
Status after last meeting

Recapitulation

Benchmarking

Milestones

- turboGliph refactored to gliphR (working title)
- 3 variants: gliph1, gliph2, and gliph3
- Possible focus on gliph3



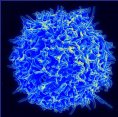
Status after last meeting

Recapitulation

Benchmarking

Milestones

- turboGliph refactored to gliphR (working title)
- 3 variants: gliph1, gliph2, and gliph3
- Possible focus on gliph3
- First to be submitted version clustering only



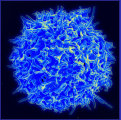
Status after last meeting

Recapitulation

Benchmarking

Milestones

- turboGliph refactored to gliphR (working title)
- 3 variants: gliph1, gliph2, and gliph3
- Possible focus on gliph3
- First to be submitted version clustering only
- Scoring & plotting in later iterations



Recapitulation

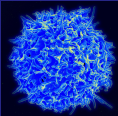
Benchmarking

Milestones

1 Recapitulation

2 **Benchmarking**

3 Milestones



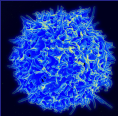
Testing gliphR

Recapitulation

Benchmarking

Milestones

- Tested gliphR with artificially enriched dataset (100fold HIV specific CDR3b)



Input data

Recapitulation

Benchmarking

Milestones

```
# set seed for reproducibility
set.seed(1234)

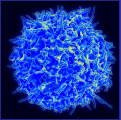
# HIV specific CDR3b sequence, taken from last research project
# TODO insert correct source / database
# CASSETGGTYEQYF TRBV6-1 TRBJ2-7
hiv_expansion <- c(CDR3b = "CASSETGGTYEQYF",
                  TRBV = "TRBV6-1",
                  TRBJ = "TRBJ2-7 ")

data("hs_CD8_ref")
data_ref <- hs_CD8_ref[, 1:3]

# insert single CDR3b sequence into reference repertoire
data_ref <- rbind(data_ref, hiv_expansion)

# sample 5000 sequences from reference repertoire
data_sample <- hs_CD8_ref[sample(x = 1:nrow(data_ref),
                                size = 5000, replace = FALSE), 1:3]

# create artificial expansion of 100 sequences
for(i in 0:99) {
  data_sample <- rbind(data_sample, hiv_expansion)
}
```

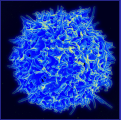


Recapitulation

Benchmarking

Milestones

- Tested gliphR with artificially enriched dataset (100fold HIV specific CDR3b)
- Tried to put in comparable input parameters for all versions (details see vignette, too many parameters)

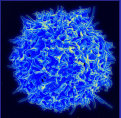


Recapitulation

Benchmarking

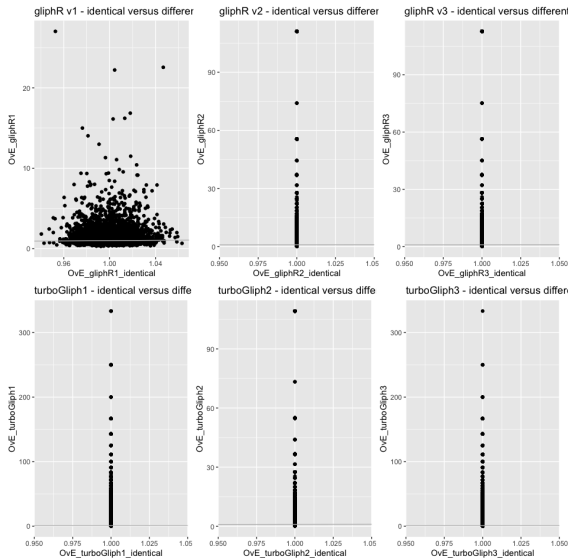
Milestones

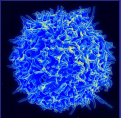
- Tested gliphR with artificially enriched dataset (100fold HIV specific CDR3b)
- Tried to put in comparable input parameters for all versions (details see vignette, too many parameters)
- Compared gliphR versions with turboGliph version by using motifs OvE as benchmark



Identical vs different input (sanity check)

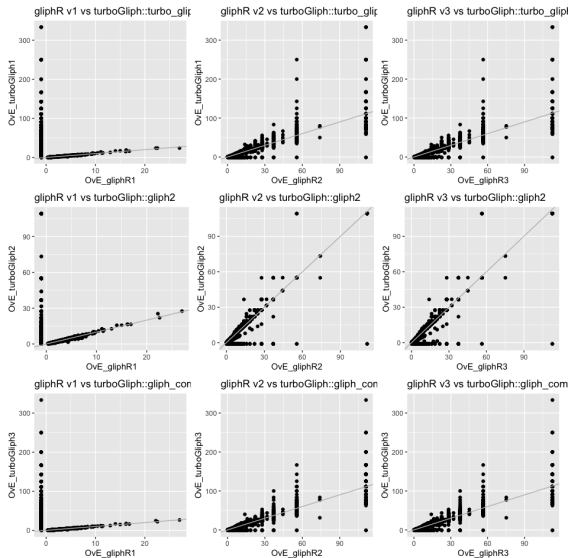
Recapitulation
Benchmarking
Milestones

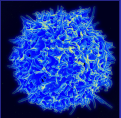




gliphR versions vs turboGliph versions

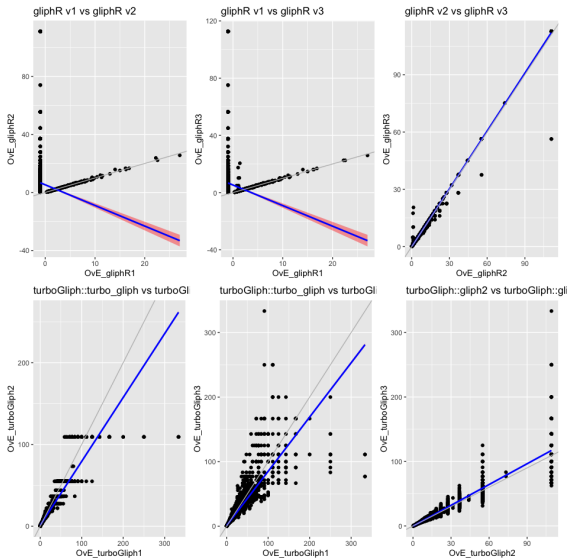
Recapitulation
Benchmarking
Milestones

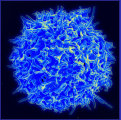




Internal version comparison

Recapitulation
Benchmarking
Milestones





Recapitulation

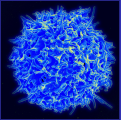
Benchmarking

Milestones

1 Recapitulation

2 Benchmarking

3 Milestones



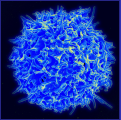
Pre-Submission

Recapitulation

Benchmarking

Milestones

- Iterate over benchmark / improve cycle until ready



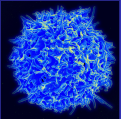
Pre-Submission

Recapitulation

Benchmarking

Milestones

- Iterate over benchmark / improve cycle until ready
- Reduce to gliph3 if possible



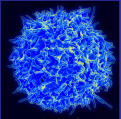
Pre-Submission

Recapitulation

Benchmarking

Milestones

- Iterate over benchmark / improve cycle until ready
- Reduce to glip3 if possible
- Bioconductor guidelines (unit tests, manuals)



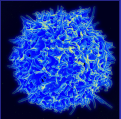
Pre-Submission

Recapitulation

Benchmarking

Milestones

- Iterate over benchmark / improve cycle until ready
- Reduce to glip3 if possible
- Bioconductor guidelines (unit tests, manuals)
- Hand in submission



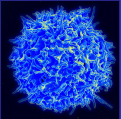
Pre-Release

Recapitulation

Benchmarking

Milestones

- Incorporate required changes to release



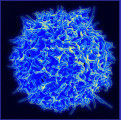
Pre-Release

Recapitulation

Benchmarking

Milestones

- Incorporate required changes to release
- Hopefully relatively short phase



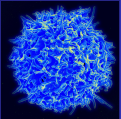
Post-Release

Recapitulation

Benchmarking

Milestones

- Maintain package, deal with potential issues



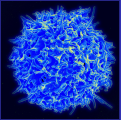
Post-Release

Recapitulation

Benchmarking

Milestones

- Maintain package, deal with potential issues
- Add scoring features



Post-Release

Recapitulation

Benchmarking

Milestones

- Maintain package, deal with potential issues
- Add scoring features
- Add plotting features