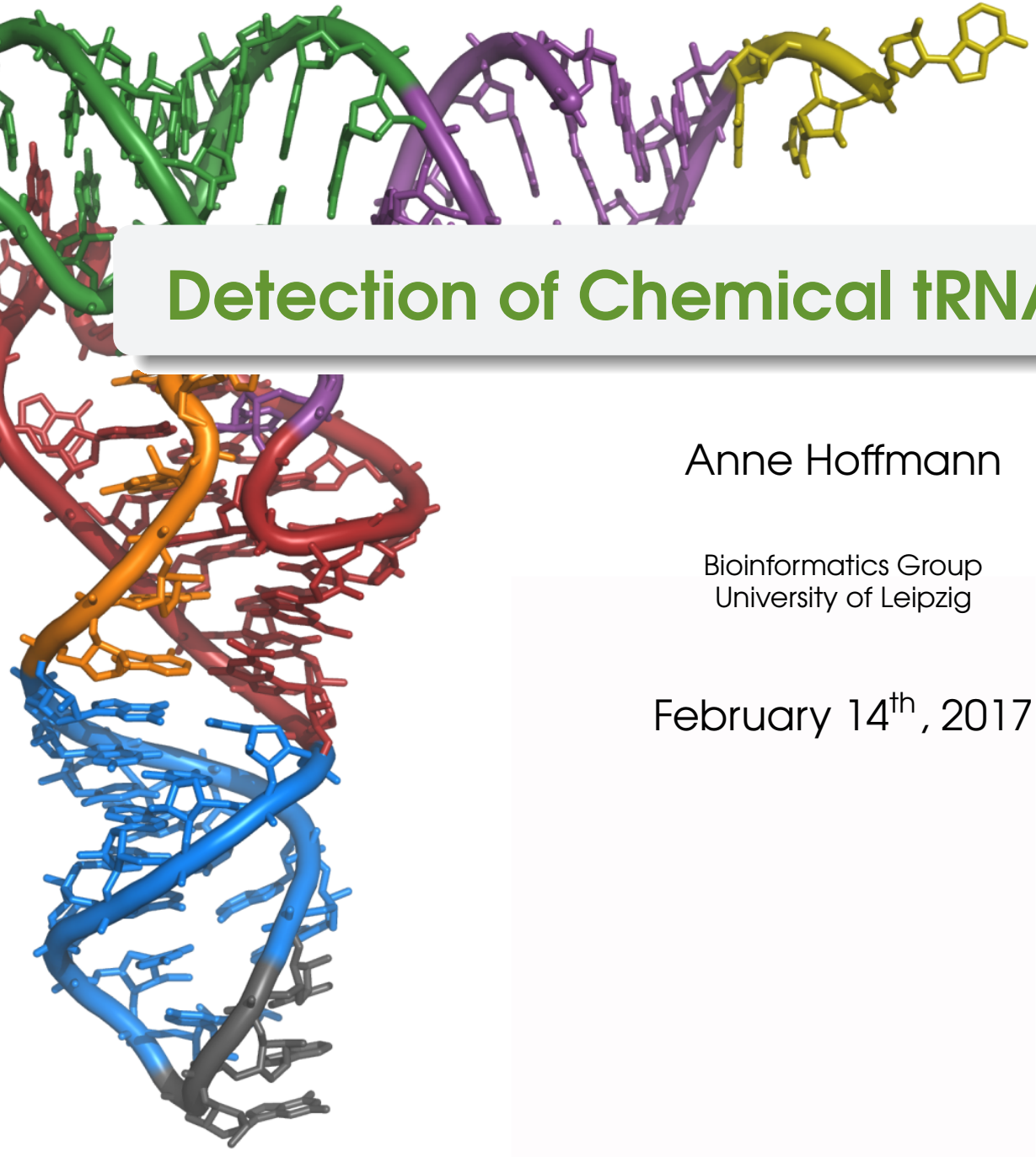
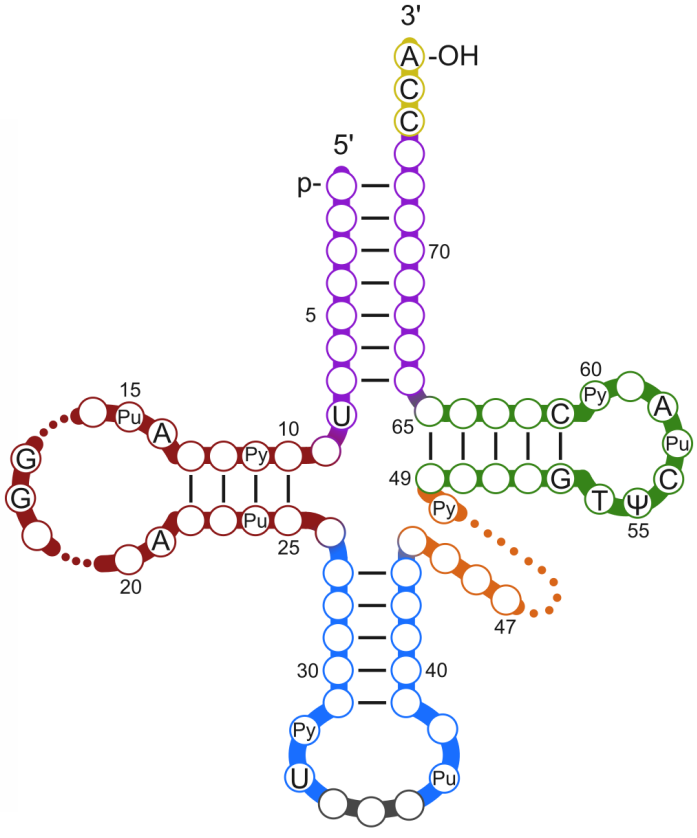


Detection of Chemical tRNA Modifications

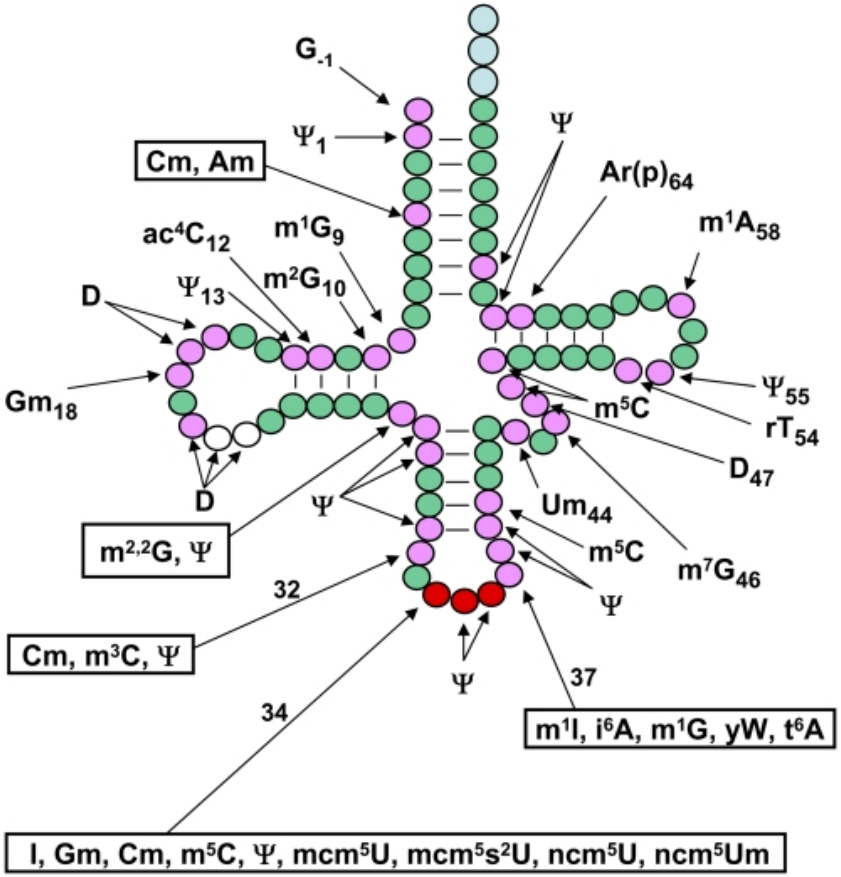
Anne Hoffmann

Bioinformatics Group
University of Leipzig

February 14th, 2017



tRNA Modifications



Phizicky et al., Genes Dev. 2010, 24: 1832-1860

- a large number exists
  - ↪ 105 are known
  - ↪ ∅ 8 modification per tRNA
  - ↪ in yeast: 25 modifications occur at 36 different positions
- **They are not well understood!!!**

tRNA Modifications affect the ...

... regulation of translation & growth (anticodon loop)

... catalytic function (various positions)

... structural stability in terms of flexibility & rigidity (main body)

tRNA Modifications affect the ...

... regulation of translation & growth (anticodon loop)

... catalytic function (various positions)

... structural stability in terms of flexibility & rigidity (main body)

- ↪ dihydrouridine in the D-arm promotes the C2'-endo sugar conformation
- ↪ are found in psychrophilic organism as a strategy for cold adaption
- ↪ playes a role in the temperature adaption of organism

Motivation

1. many modifications are recognizable in RNA-seq data as sequencing errors

Motivation

1. many modifications are recognizable in RNA-seq data as sequencing errors

   ↪ develop a bioinformatic pipeline

   ↪ understand the evolutional function and conservation

   ↪ analyze the variation across different cell types and tissues

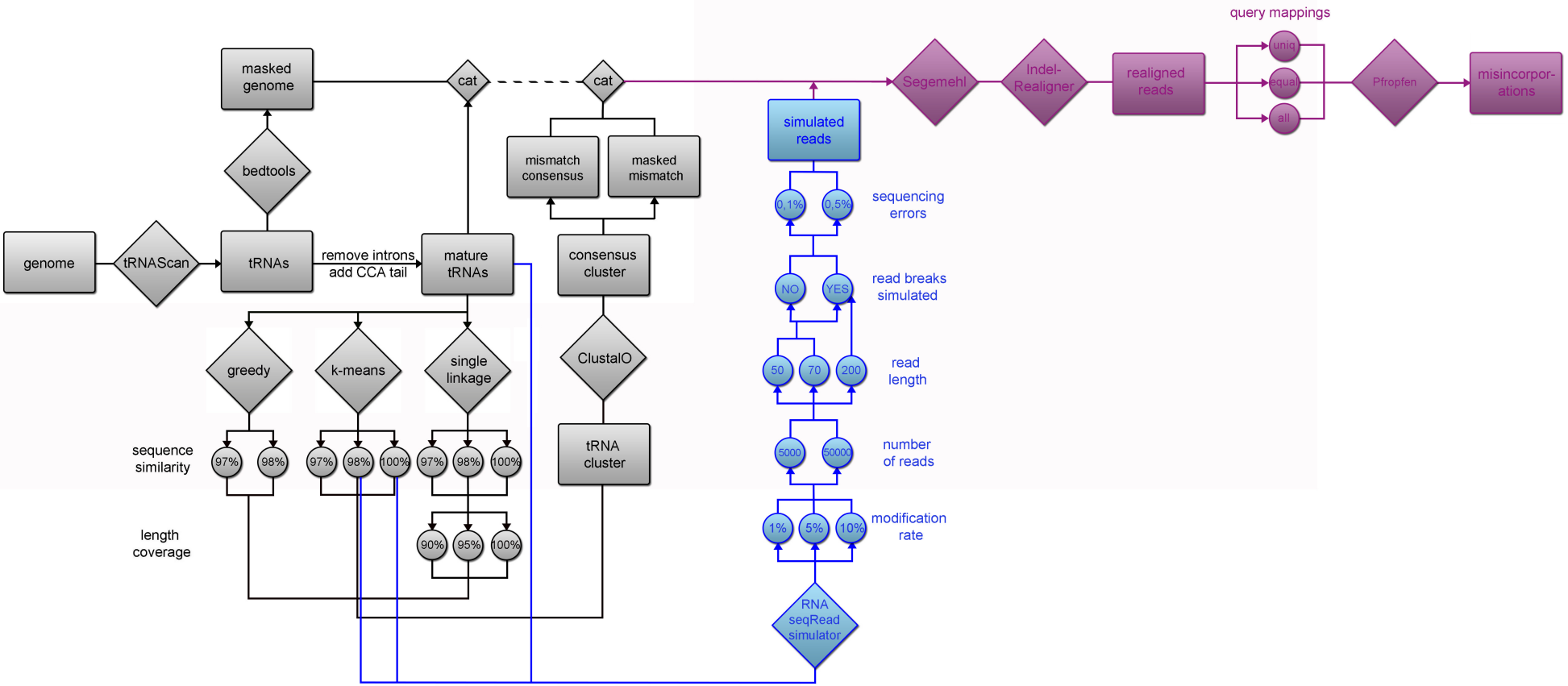
Motivation

1. many modifications are recognizable in RNA-seq data as sequencing errors
   - ↪ develop a bioinformatic pipeline
   - ↪ understand the evolutional function and conservation
   - ↪ analyze the variation across different cell types and tissues
2. many modifications modulate the tRNA structure

Motivation

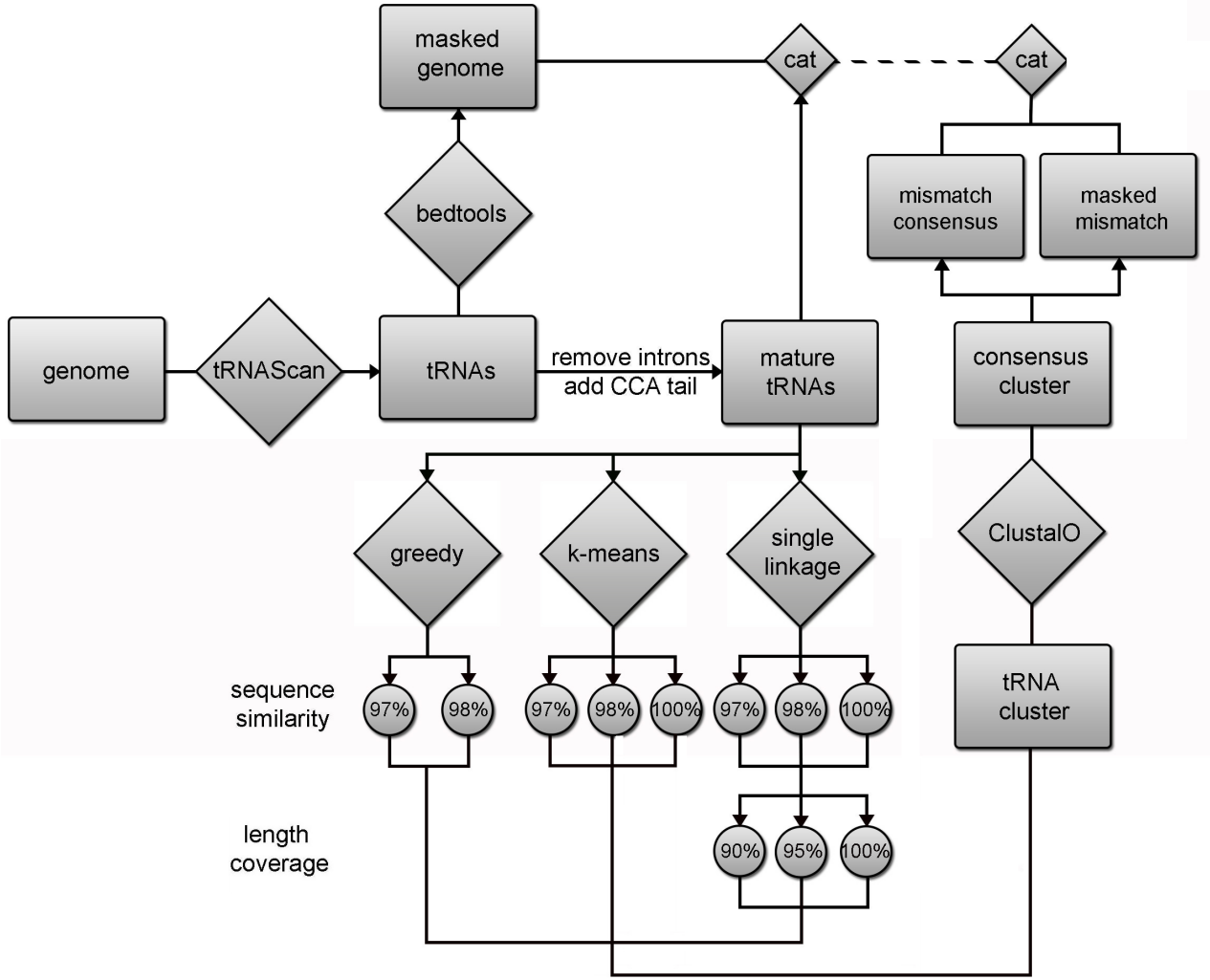
1. many modifications are recognizable in RNA-seq data as sequencing errors
   - ↪ develop a bioinformatic pipeline
   - ↪ understand the evolutional function and conservation
   - ↪ analyze the variation across different cell types and tissues
2. many modifications modulate the tRNA structure
   - ↪ understand the correlation with temperature adaption
   - ↪ analyze it in psychrophilic to thermophilic bacteria
   - ↪ analyze it in yeast
   - ↪ analyze treated vs. untreated RNA-seq data

Simulation Testing



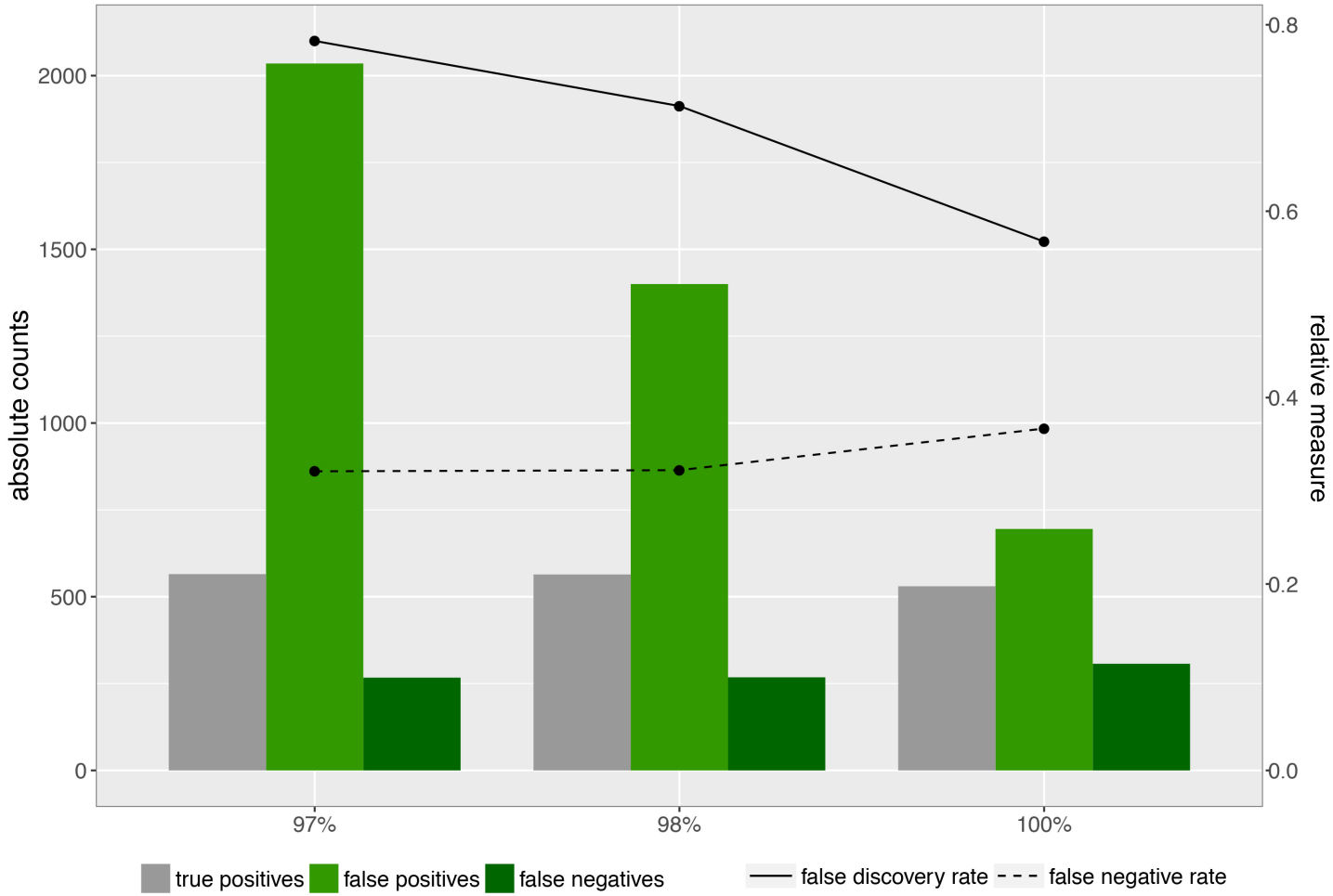
- adjust the parameters to improve the pipeline
- validate the pipeline by testing it with simulated NGS data

Simulation Testing - Genome Preparation

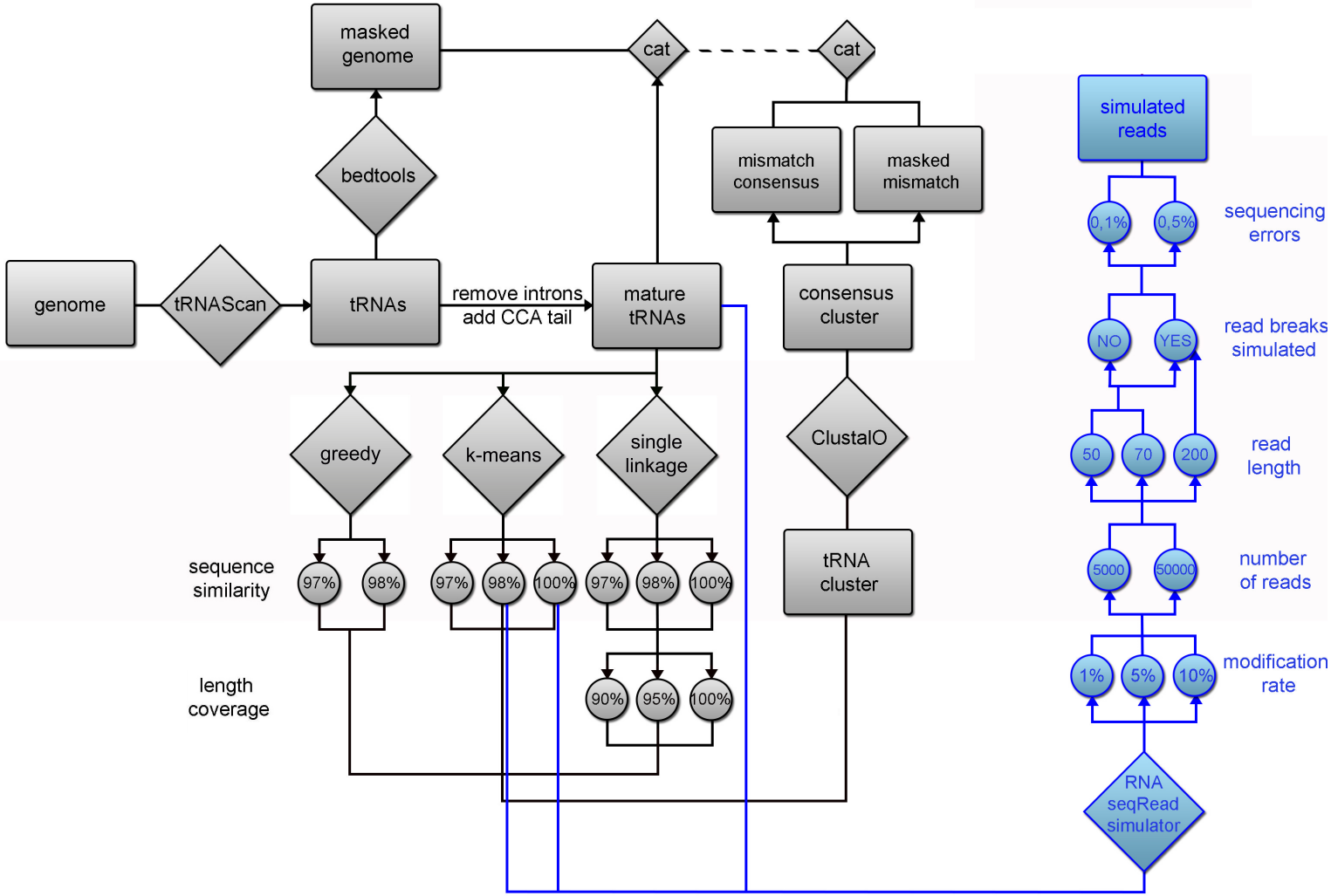


Simulation Testing - Genome Preparation

→ comparisons of different sequence similarities (greedy)

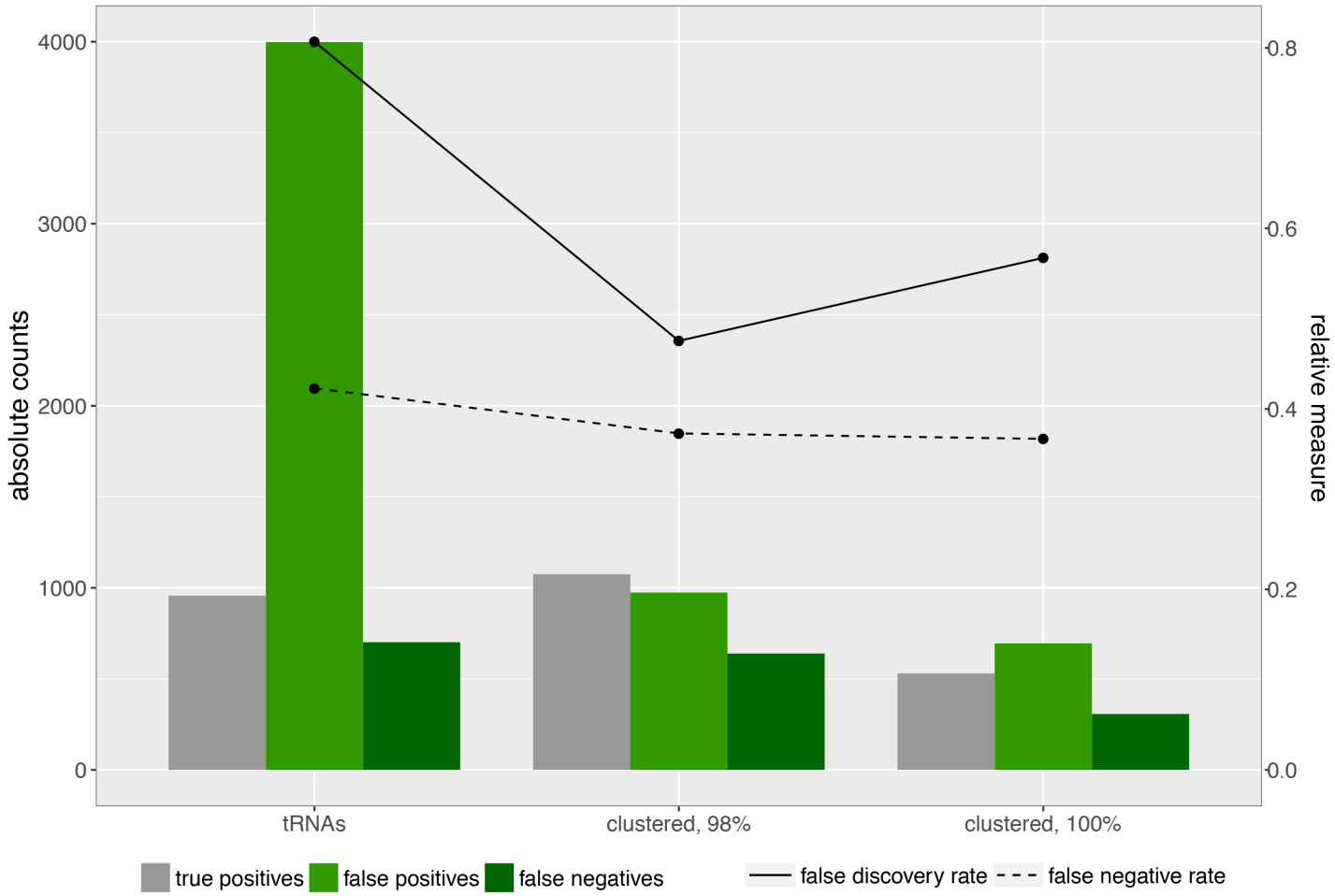


Simulation Testing - Read Simulation

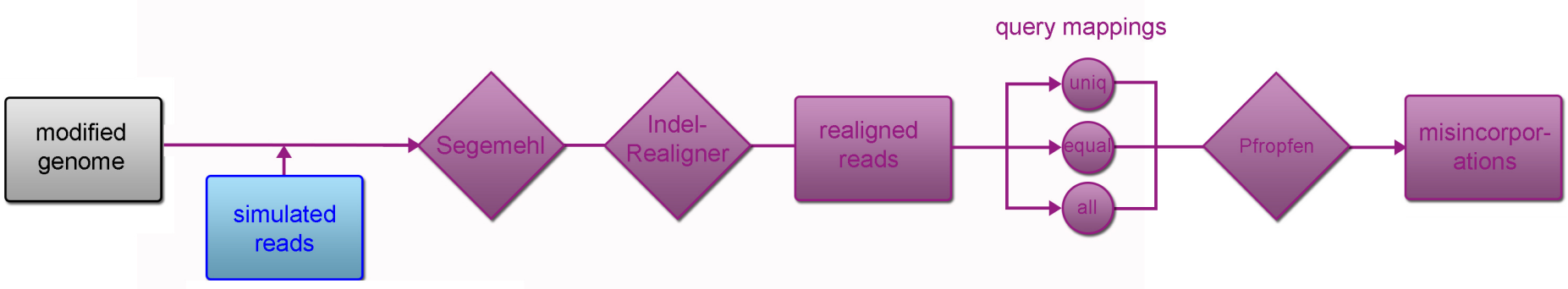


Simulation Testing - Read Simulation

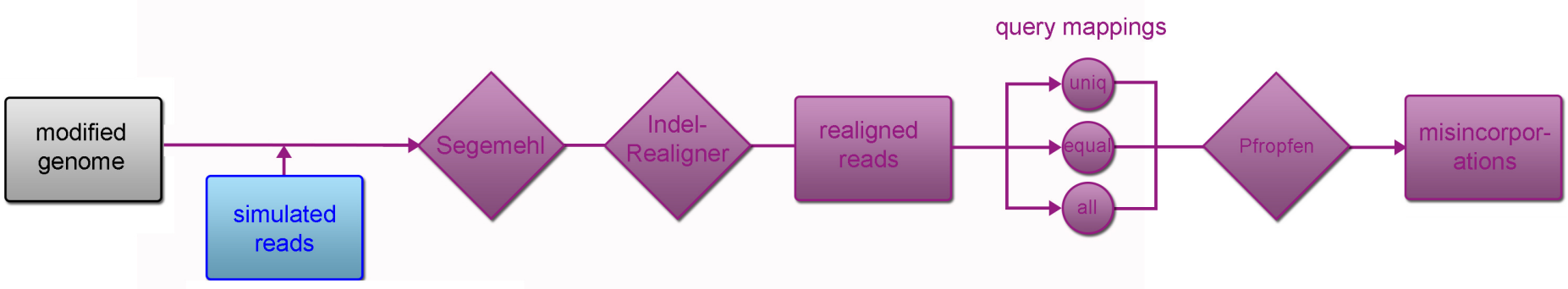
→ comparisons of different backgrounds (greedy, 100%)



Simulation Testing - Analysis



Simulation Testing - Analysis

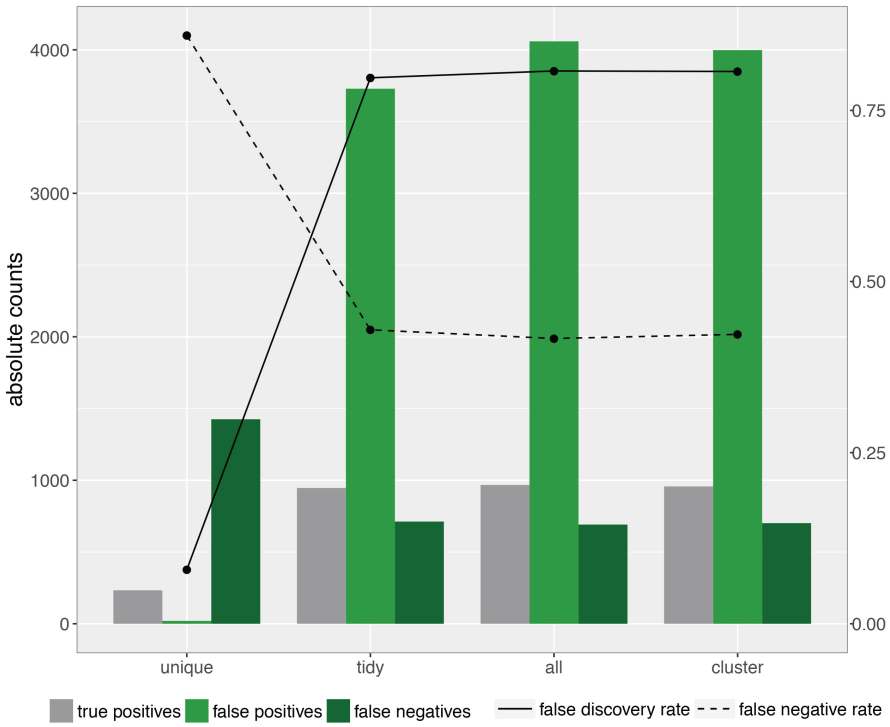


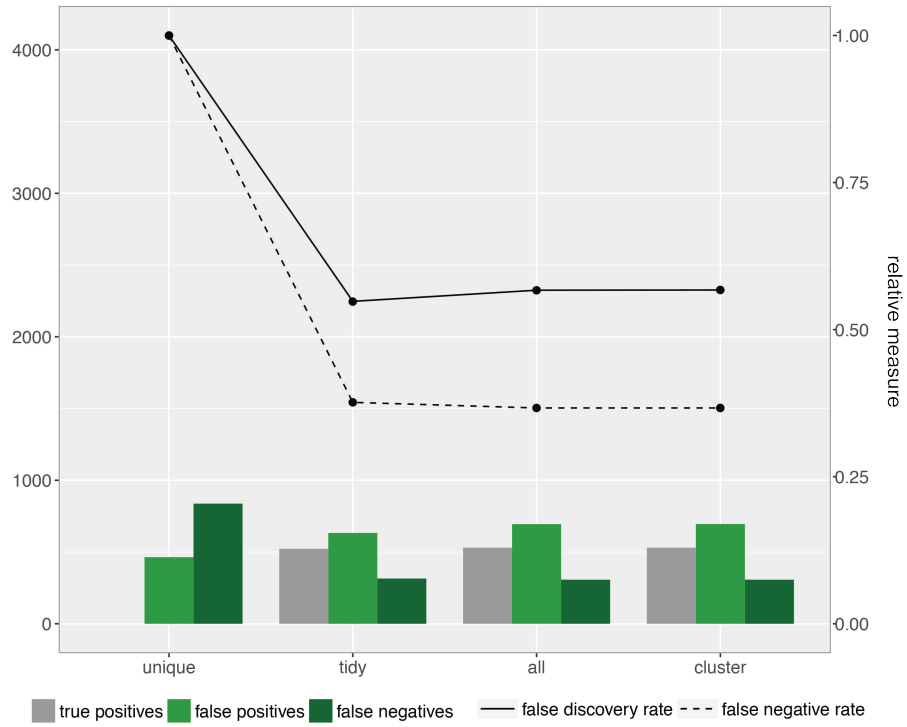
- to call misincorporations
- calculate expected error rate for each substitution as background
- weight expected rates considering multimappers
- check for every genomic position the likelihood to see a certain substitution by chance given the background rate
- merge the p-value likelihoods for all substitutions with Fisher's method

Simulation Testing - Analysis

→ comparison of different query mapping filter methods vs. cluster method

→ tRNA background vs. 100% clustered tRNA background

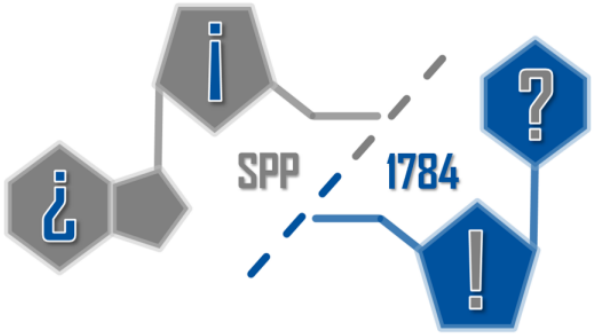




Work in progress

- validate the pipeline by testing it with bigger simulated NGS data
- analyze different data sets of different species
  - ↪ understand correlation with temperature adaption
  - ↪ analyze the evolution, conservation
  - ↪ analyze the variation across different cell types and tissues

Many thanks to...



- Peter F. Stadler
- Mario Mörl
- Fabian Amman
- Christian Lorenz

...and thanks to you for your attention!