# The Craft of Phylogeny Reconstruction

Fabian Klötzl & Bernhard Haubold

Aquavit 2018

MPI für Evolutionsbiologie

# GDPR — Privacy Policy

Feel free to share pictures of my slides on social media.

# GDPR — Privacy Policy

Feel free to share pictures of my slides on social media.
Do not take pictures of me.

# GDPR — Privacy Policy

Feel free to share pictures of my slides on social media.
Do not take pictures of me.

`kloetzl.info#downloads`

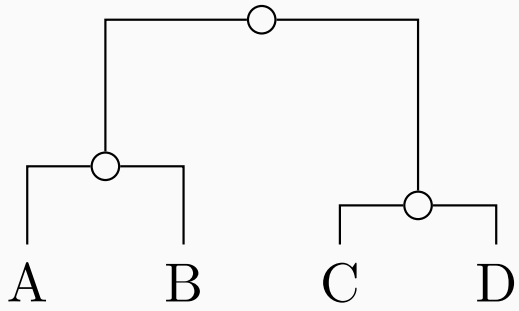
# Phylogeny Reconstruction

```
>A
AACGTTGTGCA
>B
CACGTTGTG
>C
AACGATGCGC
>D
ACCGGTGTGCT
```

$\Rightarrow$

$$\begin{pmatrix} 0 & 0.1 & 0.25 & 0.3 \\ 0.1 & 0 & 0.3 & 0.3 \\ 0.25 & 0.3 & 0 & 0.05 \\ 0.3 & 0.3 & 0.05 & 0 \end{pmatrix}$$

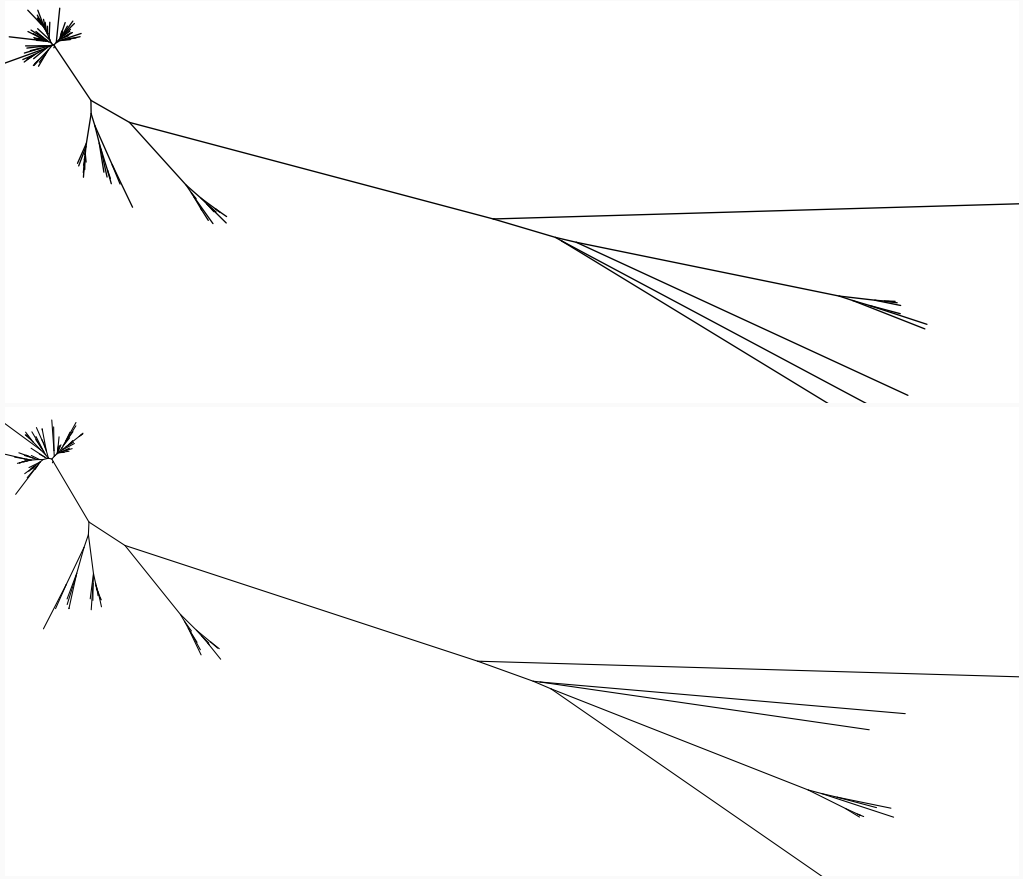
$\Rightarrow$



# 109 E. coli Genomes

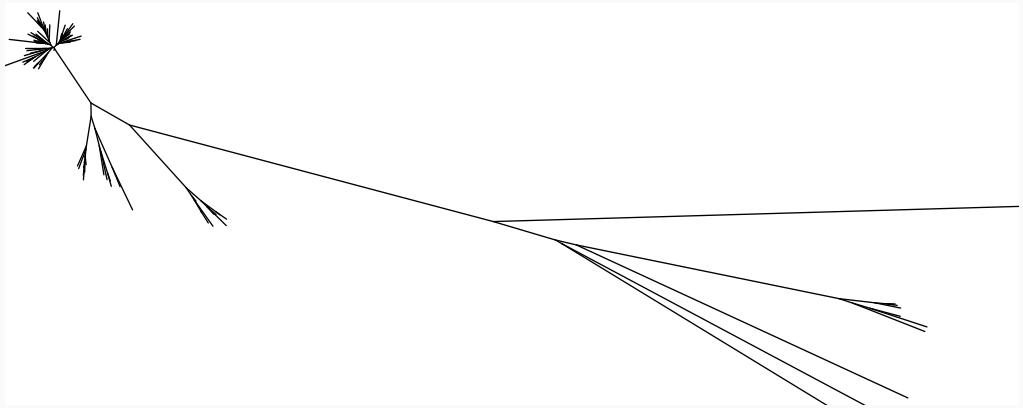
mugsy (alignment):
2 days

andi: 137s
RF distance: 127
cone distance: $1.43 \cdot 10^{-3}$

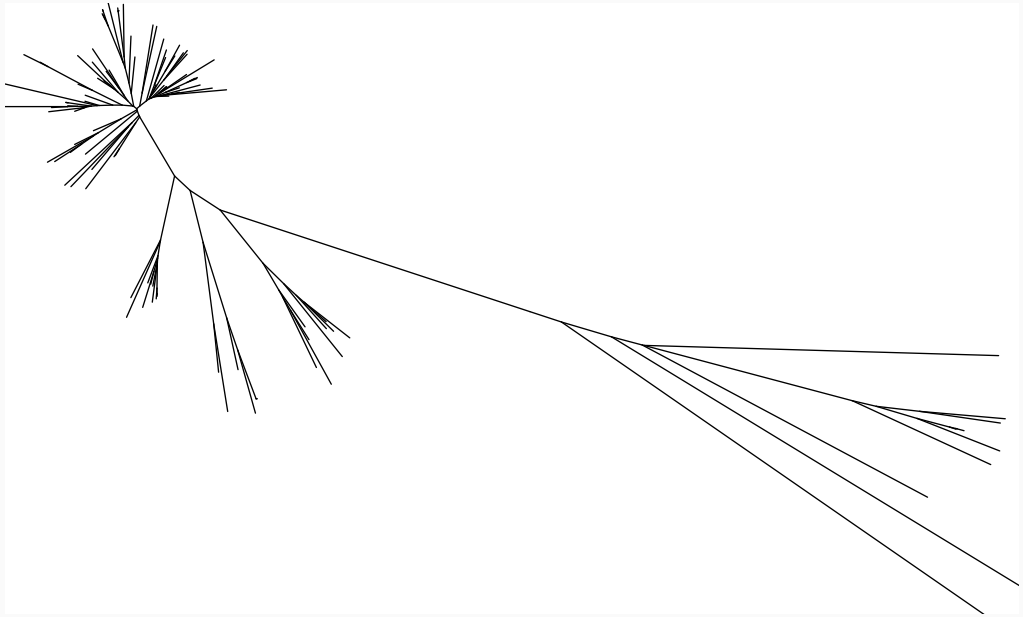


# 109 E. coli Genomes

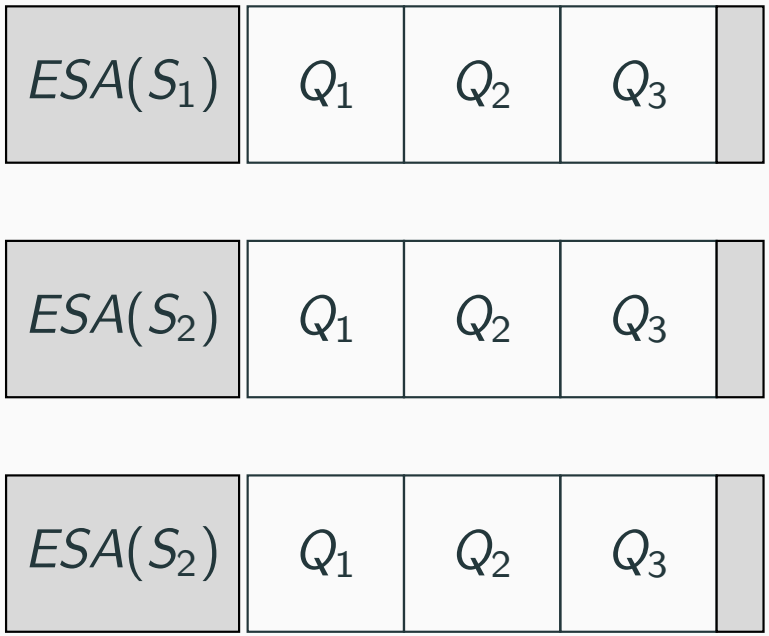
mugsy (alignment):
2 days



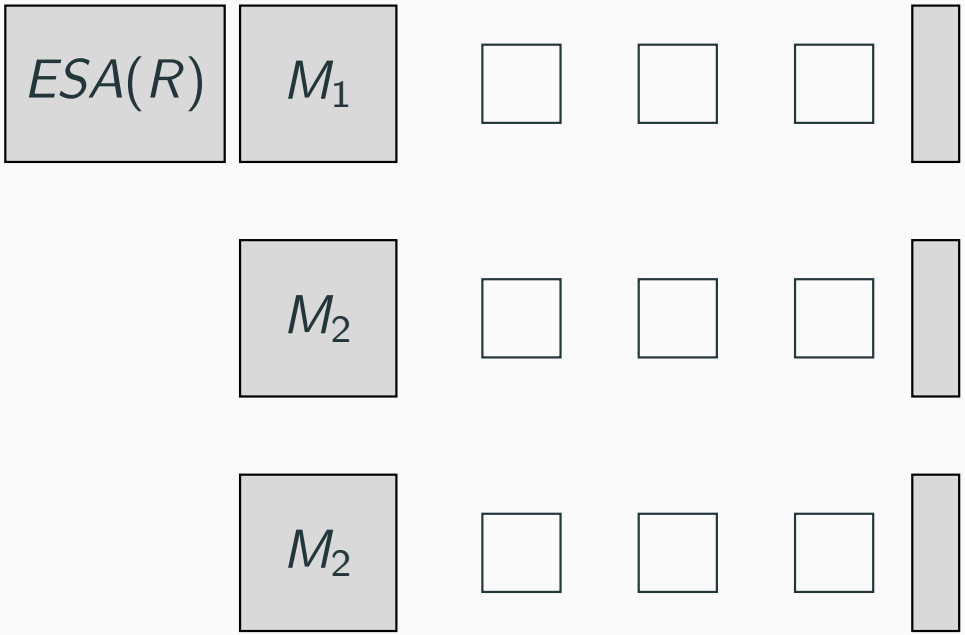
mash: 20s
RF distance: 161
cone distance: $7.91 \cdot 10^{-3}$



# andi

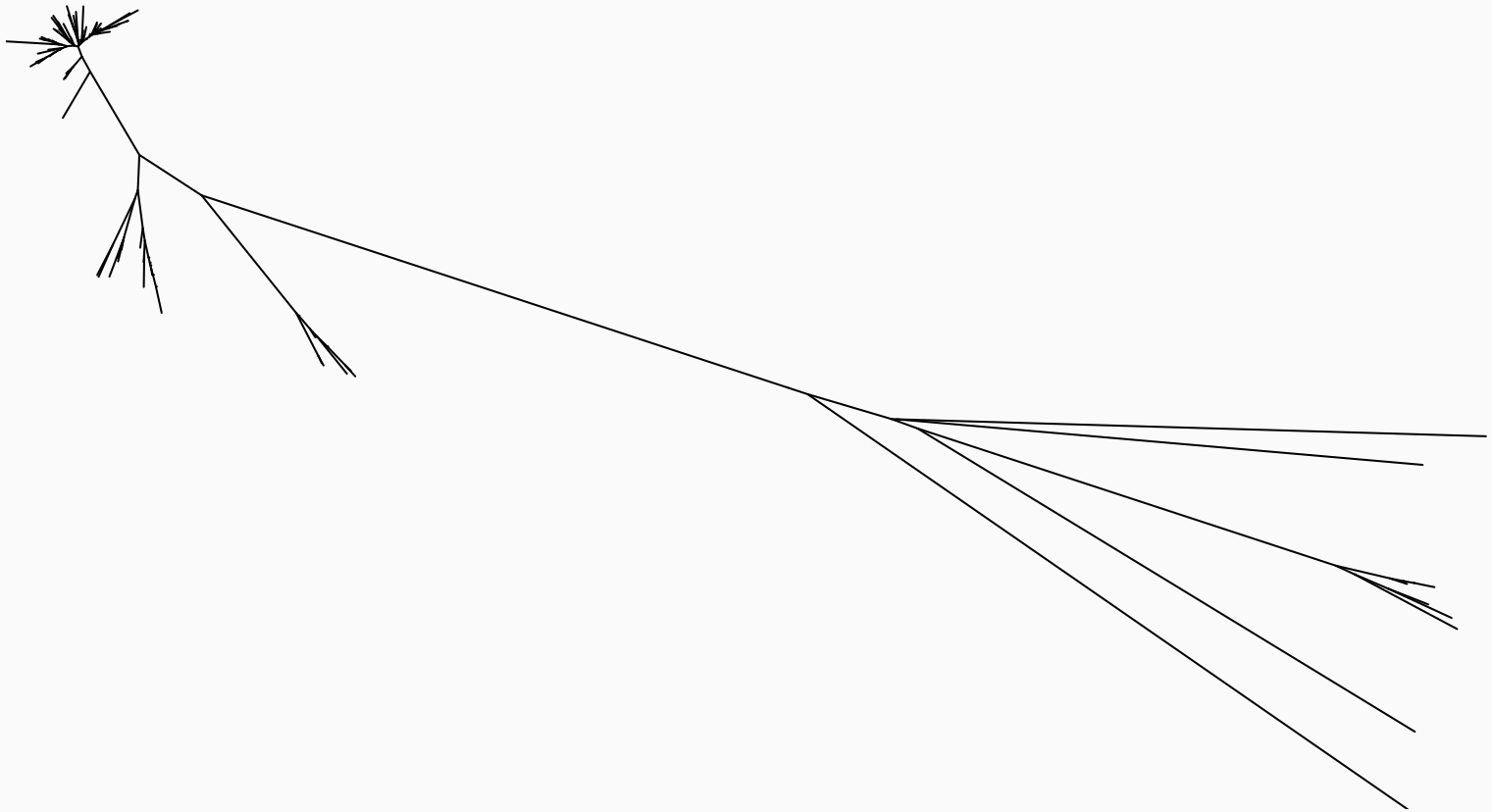


# Linearizing andi $\rightarrow$ phylonium

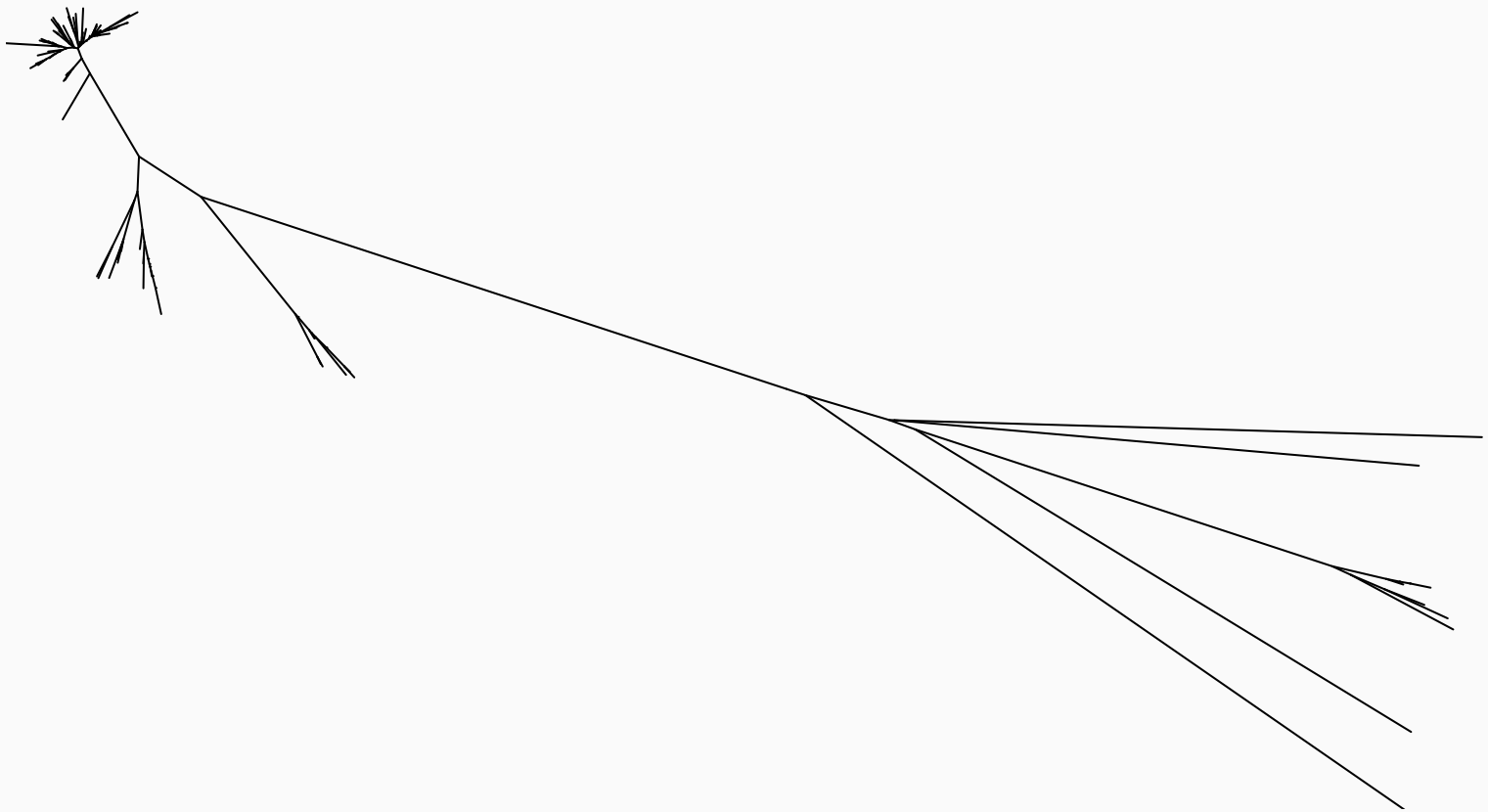



- Minimize pairwise work
- $O(n^2) \rightarrow O(n)$
- Reconstruction quality depends on the reference $R$

# Reference Choice



# Reference Choice

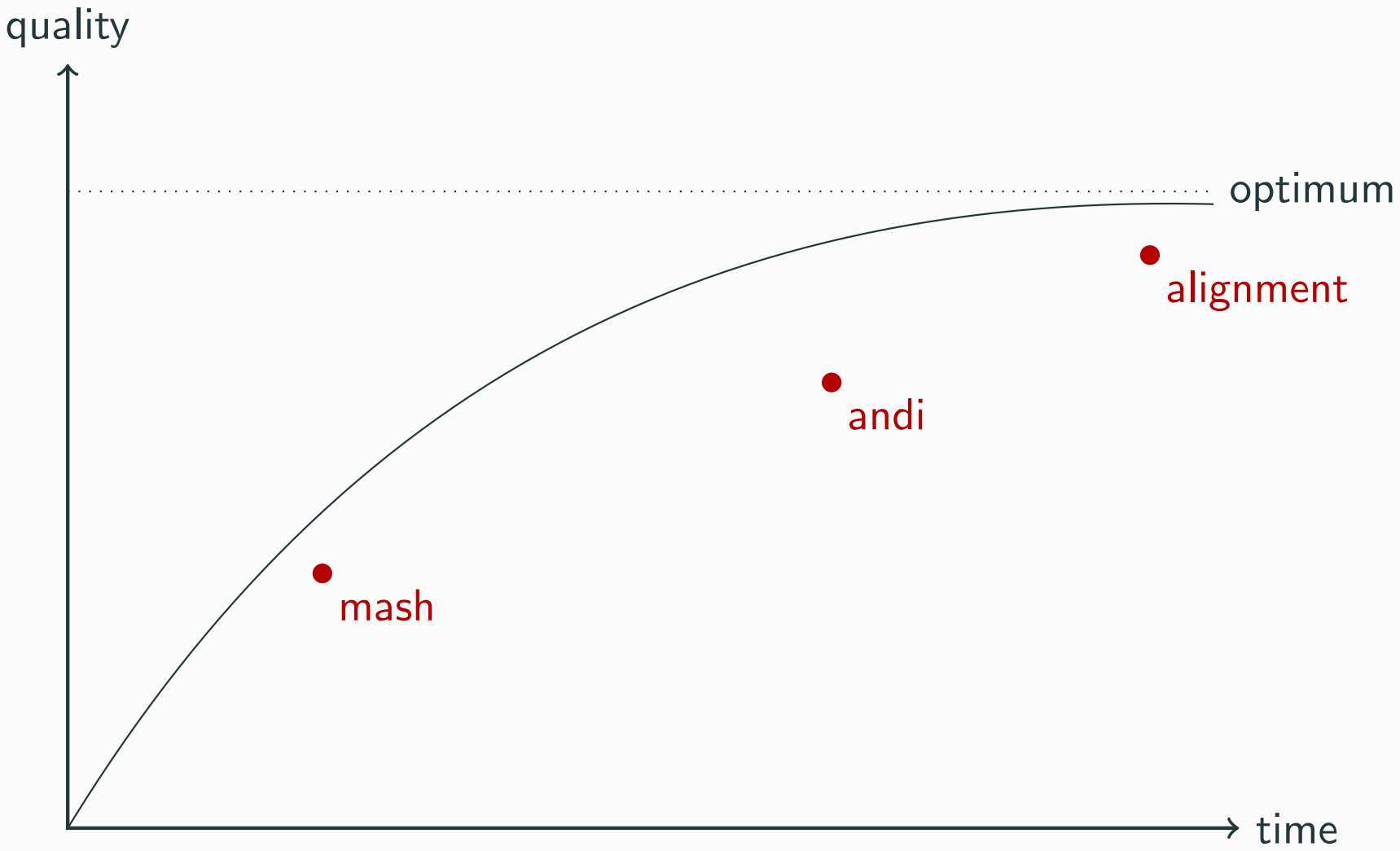


2pass strategy: 1. any reference, 2. ref from main clade

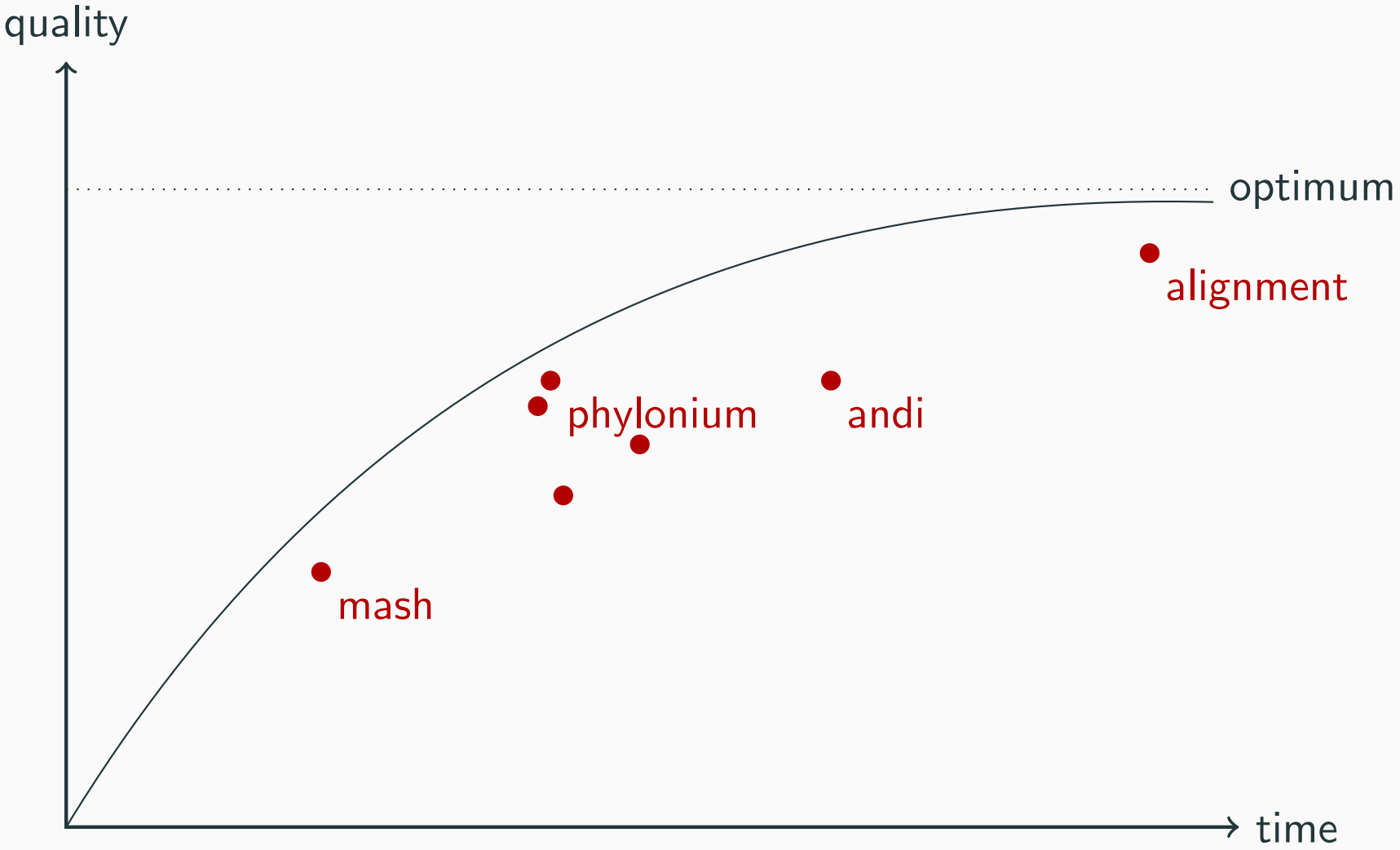
# 109 E. coli Genomes

| | RF distance | cone distance | time |
|---|---|---|---|
| mugsy | 0 | 0 | 2 days |
| andi | 127 | $1.43 \cdot 10^{-3}$ | 137s |
| mash | 161 | $7.91 \cdot 10^{-3}$ | 20s |
| phylonium | 151 | $2.21 \cdot 10^{-3}$ | 125s |

# Phylogeny Reconstruction



# Phylogeny Reconstruction



# Acknowledgements

Bernhard Haubold

andi:

- apt-get install andi
- brew install science/andi
- github.com/evolbioinf/andi

phylonium:

- coming soon