





Gliederung

Concept

Algorithms

KSEARCH

SANS

Implementation

Our implementation

Comparison to developers implementation



Implementation of KSEARCH and SANS Algorithms in the genomtools-environment

**Meike Bruhns, Florian Markowsky,
Michael Spohn**

University of Hamburg
Center for Bioinformatics
Bundesstraße 43
20146 Hamburg, Germany

<http://www.zbh.uni-hamburg.de>
info@zbh.uni-hamburg.de





Outline

Concept

Algorithms

KSEARCH

SANS

Implementation

Our implementation

Comparison to developers implementation



Sequence comparison

- ▶ Determining similarity of sequences very important task in sequence analysis
- ▶ Most commonly used: BLAST



Idea

- ▶ Count frequencies of kmers in query sequence
- ▶ Compare to kmer frequencies of database entries



Analysis

- ▶ Runtime:
- ▶ Space requirements:



Idea

- ▶ Construct suffix tree for query sequence q
- ▶ Construct suffix tree for database sequences
- ▶ Merge suffix trees:
 - ▶ for each suffix of q



Analysis



