



## ***Test und Vergleich mehrerer Short Read Assembler***

Bei der Sequenzierung eines Genoms wird seine DNA-Sequenz aufgeschlüsselt. Man erhält kleine, sich überlappende DNA-Abschnitte des Genoms, hat aber wenig Informationen wie sie miteinander verbunden sind. Diese Abschnitte (Reads) können von einem Assembler zusammengesetzt werden.

Um die Kosten einer Sequenzierung zu senken, wurden einige neue Technologien entwickelt. Der Nachteil dieser neuen Technologien ist, dass die resultierenden DNA-Abschnitte deutlich kürzer sind als das bei früheren Verfahren der Fall war. Dadurch sind auch die Assemblierungsprobleme schwieriger geworden und neue Assembler wurden entwickelt, die mit diesen Short Reads umgehen können.

Im Rahmen einer Bachelorarbeit sollen einige dieser Short Read Assembler getestet und miteinander verglichen werden.

### **Aufgaben:**

- Installation von 2 bis 3 aktuellen Short Read Assemblern
- Zusammenstellung eines Datensets auf dem die Assembler getestet werden sollen
- Durchführung und Auswertung der Tests
- Vergleich der Ergebnisse

### **Voraussetzungen:**

- Programmierkenntnisse sind von Vorteil, sowie Erfahrungen in der Installation von Fremdsoftware
- Grundlegende Kenntnisse oder Interesse an Assemblyproblemen sind wünschenswert

**Dauer der Arbeit:** 4 Monate

### **Links:**

<http://www.ebi.ac.uk/~zerbino/velvet/>  
<http://www.cse.ucsd.edu/~ppevzner/software.html#EULER-short>  
<http://sharcgs.molgen.mpg.de/software/sharcgs.shtml>

### **Kontakt:**

Juliane Damaris Klein  
Algorithms in Bioinformatics  
WSI/ZBIT, Sand 14  
Room: C324b  
kleinj@informatik.uni-tuebingen.de

