

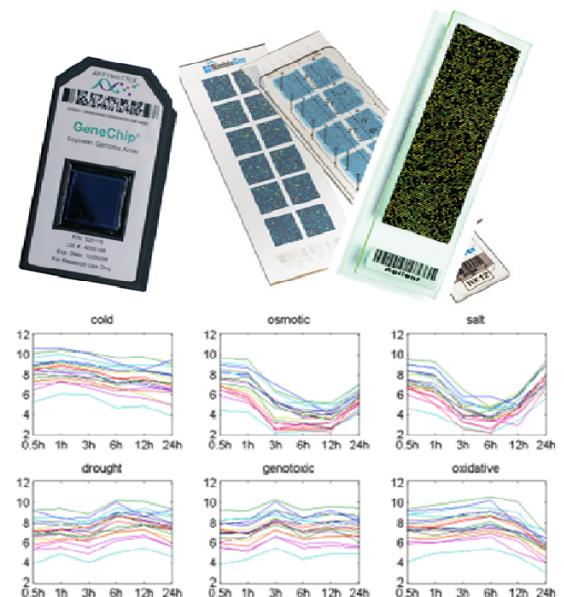


Ein plattformübergreifendes Clustering-verfahren in Java

„High-throughput“ Daten gibt es in der Bioinformatik bereits seit ungefähr zehn Jahren. Für mRNA Expressionsdaten gibt es bereits sehr viele gut funktionierende Clustering-Verfahren. Seit einiger Zeit kommen jedoch immer mehr auch weitere Datentypen hinzu, z.B. die Expression von micro RNAs, Expression von Proteinen, Phosphorylierungsgrad von Proteinen, Histon-Modifikationen, DNA-Methylierung, und vieles. Am Lehrstuhl Kognitive Systeme gibt es einen Datensatz, welcher mRNA, miRNA, DNA-Methylierung und (Phospho-)Protein Expressionsdaten enthält.

Ziel dieser Arbeit ist es, ein plattformübergreifendes Clustering-Verfahren in Java zu implementieren, welches folgenden Anforderungen genügt:

Neben einem normalen Einteilen in Gruppen co-exprimierter Gene sollte es möglich sein, Gruppen von Genen zu identifizieren, welche definierten Ansprüchen genügen. Ein Beispiel wäre Gruppen von Genen, welche eine starke Demethylierung im Promotorbereich aufweisen und eine erhöhte mRNA-Expression besitzen. Allgemeiner ausgedrückt sollte der Algorithmus automatisch Gene, unter Berücksichtigung von Plattform, Zeit und Bedingung in Gruppen einteilen. Dieser Algorithmus sollte als benutzerfreundliches Java-Programm implementiert werden, in welchem der Benutzer gezielt die Clustering-Anforderungen auswählen kann.



gewünschten Anforderungen des

Kontakt

Clemens Wrzodek
Sand 1, Raum 306
Tel. (07071) 29- 70436
Clemens.Wrzodek@uni-tuebingen.de