

Abstract

It has been shown that the analysis of microarray data is a very promising approach in experimental molecular biology. Various clustering techniques can be used to analyze gene expression patterns from microarray data at a genomic scale.

A web application as an implementation of a web based client server architecture has been developed to analyze data from microarray experiments. The software is based on the Java 2 Enterprise Edition (J2EE) technology from Sun Inc including Java Server Pages (JSP), Servlets, Enterprise Java Beans (EJB) and on the open source framework struts from the apache software foundation. To provide a user interaction with the data and increase the user-friendliness JavaScript and image maps have been used. The microarray data can be analyzed with different clustering algorithms including 1) Hierarchical Clustering (HCL), 2) K-Means Clustering (KMC), 3) Self Organizing Maps (SOM), 4) Support Vector Machines (SVM), 5) Principal Component Analysis (PCA) and 6) Correspondence Analysis (CA). Various distance measurements can be used to measure the similarity between genes.

The platform independent and scalable web application makes it possible to analyze large data sets with most web browser also from low resource workstations without having to install any additional software packages. All user data will be saved on the server so that it can be accessed from everywhere in the network. Therefore the web application is a useful tool to analyze large microarray data also in networks with low hardware resources.

Keywords: microarray, cluster analysis, web application, J2EE, bioinformatics

Kurzfassung

Die Microarray-Technologie hat sich in den letzten Jahren für die Analyse von Gen Expressionen auf Genom Ebene etabliert. In den Microarraydaten wird nach Genen mit gleichen Expressionsmustern gesucht, da bei diesen von einer ähnlichen Funktion bzw. einer Coregulation ausgegangen wird. Für diese Suche können verschiedene Clusteranalyse-Verfahren verwendet werden.

Im Rahmen der Diplomarbeit wurde eine Web-Applikation basierend auf der Microarray-Analyse-Suite Genesis als Implementation einer Web basierten Client Server Architektur entwickelt. Die Software basiert auf der Java 2 Enterprise Edition (J2EE) Architektur von Sun Inc und dem open source Framework struts der apache software foundation. Daten aus Microarray-Experimenten können mit Hilfe der Web-Applikation dargestellt und durch die verschiedenen Clusterverfahren 1) Hierarchical Clustering (HCL), 2) K-Means Clustering (KMC), 3) Self Organizing Maps (SOM), 4) Support Vector Machines (SVM), 5) Principal Component Analysis (PCA) und 6) Correspondence Analysis (CA) analysiert werden. Hierbei können für die Ähnlichkeitsmessung der Gene verschiedenste Distanzmessverfahren angewandt werden. Durch die Verwendung einer Image Map und JavaScript wird eine Interaktion des Benutzers mit den dargestellten Bildern und Daten ermöglicht. Mit Hilfe eines Navigationsbaumes wird eine übersichtliche, intuitive und einfache Navigation und Bedienung der Web-Applikation ermöglicht.

Der große Vorteil der Web-Applikation liegt in der Entlastung der Client Rechner, da rechenintensive Analysen durch einen hierfür ausgelegten Server durchgeführt werden. Der Benutzer benötigt nur einen Webbrowser um Microarray-Analysen durchführen zu können. Somit wird eine maximale Plattformunabhängigkeit garantiert und es entfällt die Installation eines eigenen Programmes. Die Daten werden zentral abgespeichert wodurch der Benutzer von jedem Rechner aus auf seine Analysen zugreifen kann. Die Web-Applikation stellt somit ein hilfreiches Werkzeug zur Analyse von Microarray-Daten vor allem in Netzwerken mit beschränkten Hardwareresourcen dar.

Schlüsselwörter: Microarray,Clusteranalyse,Web-Applikation,J2EE, Bioinformatik