

Exposé zur Studienarbeit

## Visualisierung von Expressionsdaten in Pathways

Oliver Arnold

Betreuer: Ulf Leser, Silke Meiners

### 1 Hintergrund

Mechanische, chemische oder physikalische Reize können bei Endothelzellen zu Schädigungen führen, die eine endotheliale Dysfunktion zur Folge haben. Diese bildet die Grundlage für die Ausbildung kardiovaskulärer Erkrankungen.

Das Ubiquitin-Proteasom System ist ein zentraler Mechanismus zum Abbau von falsch gefalteten Proteinen und zur Regulierung von Proteinkonzentrationen von wichtigen regulatorischen Proteinen des Stoffwechsels in Zellen. Proteine, die zum Abbau vorgesehen sind, werden in einem mehrstufigen enzymatischen Prozess mit einer Polyubiquitin-Kette markiert und vom Proteasom gebunden, das das Protein dann in Peptide zerlegen kann. Die Peptide lassen sich in einzelne Aminosäuren zerlegen, die wieder zur Synthese von Proteinen verwendet werden können [COS00]. Die partielle Hemmung des Ubiquitin-Proteasom Systems mit einer niedrigen, nicht-toxischen Dosierung von Proteasominhibitoren hat eine positive Auswirkung auf Endothelzellen, wodurch eine verbesserte endotheliale Funktion und ein besserer Schutz vor oxidativem Stress erreicht wird. Die Nutzung von Proteasominhibitoren könnte somit als neuer therapeutischer Ansatz zur Behandlung von endothelialen Dysfunktionen bei kardiovaskulären Krankheiten geeignet sein. [MLLDBSS06]

Um die Wirkung der Proteasominhibitoren genauer zu untersuchen, werden an der Medizinischen Klinik und Poliklinik für Kardiologie und Angiologie der Charite Oligonukleotid-Microarrays verwendet, mit deren Hilfe man feststellen kann, wie unterschiedliche Dosierungen von Proteasominhibitoren die Expression von Genen beeinflussen. Für einen Vergleich der Genexpressionen zwischen einer Referenzzelle und einer behandelten Zelle werden mRNAs aus den Zellen extrahiert und daraus mittels reverser Transkription cDNA erzeugt. Durch die Markierung mit Farbstoffen können die Hybridisierungssignale der cDNA auf dem DNA-Chip mit Laserscannern quantitativ ausgewertet werden. [BH02]

Pathways sind Serien von chemischen Reaktionen innerhalb einer Zelle, die für die Verarbeitung von Nährstoffen und die Synthese von chemischen Verbindungen, wie z.B. Nukleinsäuren und Proteinen, notwendig sind. Die einzelnen Reaktionen werden jeweils durch ein hochspezifisches Enzym katalysiert. Da die Anwendung von Proteasominhibitoren in Endothelzellen Auswirkungen auf die Expression von Genen hat und damit auf die Aktivität der Enzyme, sind auch Reaktionsketten in den Pathways betroffen.

## 2 Ziele

In dieser Studienarbeit geht es darum die Ergebnisse von Genexpressionsexperimenten in Pathways zu visualisieren. Dazu soll eine geeignet sortierte Liste von Pathways berechnet werden, in der jeder Pathway von einer Menge von erhöhten oder verminderten Enzymaktivitäten betroffen ist. Die durch die veränderten Genexpressionen betroffenen Enzyme sollen in den einzelnen Pathways hervorgehoben und gemeinsam mit dem Pathway visualisiert werden.

Eine Vertiefung der Arbeit wird sich mit der Abschätzung der statistischen Signifikanz der Ergebnisse beschäftigen, wobei die Anzahl der Enzyme in einem Pathway und wie viele davon im Experiment reguliert sind mit einfließen sollen. Desweiteren soll noch eingehen, wie viele der Enzyme einer Affymetrix-ID zugeordnet werden konnten.

## 3 Vorgehen

Als Eingabe werden Daten erwartet, die pro Gen eine Affymetrix-ID und die Information über die Expressionänderung des Gens in einem Experiment enthalten. Die Affymetrix-ID kann mit Hilfe einer Tabelle<sup>1</sup> in eine Unigene-Nummer gemappt werden. Unigene-Nummern fassen Expressed Sequence Tags (ESTs), die einem Gen zugeordnet sind, zu einem Cluster zusammen. Danach lässt sich über die NCBI-Webseite<sup>2</sup> mittels der Unigene-Nummer das entsprechende Gen und das dazugehörige Enzym ermitteln. In der KEGG-Datenbank<sup>3</sup> kann dann festgestellt werden, in welchen Pathways das Enzym verwendet wird. Um die Pathways darstellen zu können, ist das Parsen der KEGG-Datenbank für Pathways<sup>4</sup> notwendig. Zur Visualisierung soll Prefuse<sup>5</sup> benutzt werden. Sollte sich die Darstellung der Pathways dadurch zu unübersichtlich gestalten, können auch die Koordinaten der Objekte aus den KEGG XML Maps verwendet werden.

Dem Benutzer wird eine geeignet sortierte Liste bereitgestellt, in der einzelne Pathways zur Visualisierung ausgewählt werden können. Die Sortierung der Liste richtet sich nach dem Ausmaß der Beeinflussung eines Pathways durch veränderte Enzymaktivitäten. Besonders betroffene Pathways stehen weiter oben in der Liste als weniger betroffene Pathways. Neben den Pathways soll mindestens eine weitere Information, z.B. die Anzahl der betroffenen Enzyme, angegeben werden.

---

<sup>1</sup><http://bioinformatics.bioen.uiuc.edu/gosurfer/download/HG-U133AB%20GoSufer.xls>

<sup>2</sup><http://www.ncbi.nlm.nih.gov/UniGene/>

<sup>3</sup><http://www.genome.jp/kegg/>

<sup>4</sup><http://www.genome.jp/kegg/xml/>

<sup>5</sup><http://prefuse.org/>

## Literatur

- [BH02] Baldi P, and Hatfield GW. *DNA Microarray and Gene Expression: From Experiments to Data Analysis and Modeling*. Cambridge University Press. 2002.
- [COS00] Ciechanover A, Orian A, Schwartz AL. *Ubiquitin-mediated proteolysis: biological regulation via destruction*. *Bioessays*. 2000; 22:442-451
- [MLLDBSS06] Meiners S, Ludwig A, Lorenz M, Dreger H, Baumann G, Stangl V, Stangl K. *Nontoxic proteasome inhibition activates a protective antioxidant defense response in endothelial cells*. *Free Radic Biol Med*. 2006; 40:2232-2241.