Disease-relevant gene clusters derived by joint literature and gene expression analysis

Markus Bundschus, Simon Paradies, Matthias Siebert
Fabian Birzele, Katrin Fundel, Robert Küffner, Ralf Zimmer

Institut für Informatik, Ludwig-Maximilians-Universität München, Amalienstr. 17, 80333 München, Germany
E-mail: bundschu@cip.ifi.lmu.de, paradies@cip.ifi.lmu.de, siebertm@cip.ifi.lmu.de

Overview

Feature Construction

Identification of disease-related gene clusters

- Initial gene set
- Feature Vector Construction
- Hierarchical clustering approach

Cluster Ranking

- Goal
- Cluster ranking to assess the importance of clusters
- Overall cluster ranking using weighted rank sum from DT + ORA

Results

Conclusion

➢ Results
  ➢ Top ranking clusters contain OA relevant genes (collagens, MMPs, ADAMs and interleukins)

➢ Features
  ➢ Clusters might predict additional genes not measured on the chip
  ➢ Starting point for improved chip design
  ➢ Insights into underlying biological mechanisms

➢ Outlook
  ➢ Optimize clusters by iterative procedure
  ➢ Report publications relevant to clusters
  ➢ More in-depth evaluation of clusters
  ➢ Incorporate additional features in feature vectors

Acknowledgement: The authors wish to thank Bioinformatics Munich (BIM) for financial support