Two Splice Factor Mutant Leukemia Subgroups Uncovered at the Boundaries of MDS and AML using Combined Gene Expression and DNA-Methylation Profiling

Taskesen, Erdogan; Havermans, Marije; van Lom, Kirsten; Sanders, Mathijs; van Norden, Yvette; Bindels, Eric; Hoogenboezem, Remco; Reinders, Marcel; Valk, Peter; More Authors

Publication date
2011

Document Version
Final published version

Citation (APA)
Taskesen, E., Havermans, M., van Lom, K., Sanders, M., van Norden, Y., Bindels, E., Hoogenboezem, R., Reinders, M., Valk, P., & More Authors (2011). *Two Splice Factor Mutant Leukemia Subgroups Uncovered at the Boundaries of MDS and AML using Combined Gene Expression and DNA-Methylation Profiling*. 1.

Two Splice Factor Mutant Leukemia Subgroups Uncovered at the Boundaries of MDS and AML using Combined Gene Expression and DNA-Methylation Profiling

Erdogan Taskesen, Marije Havermans, Kirsten van Lom, Mathijs Sanders, Yvette van Norden, Eric Bindels, Remco Hoogenboezem, Marcel J.T. Reinders, Maria E. Figueroa, Peter J.M. Valk, Bob Löwenberg, Ari Melnick and Ruud Delwel

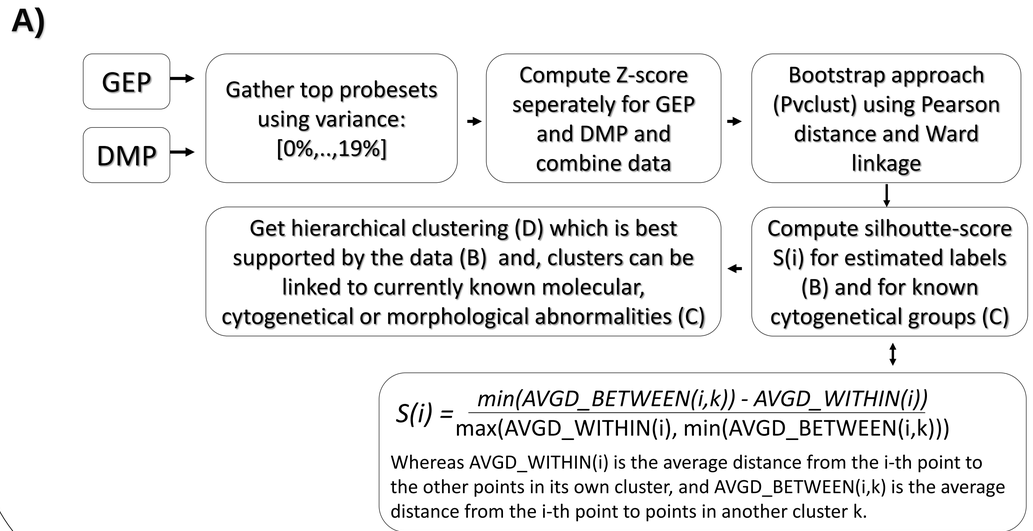
Summary

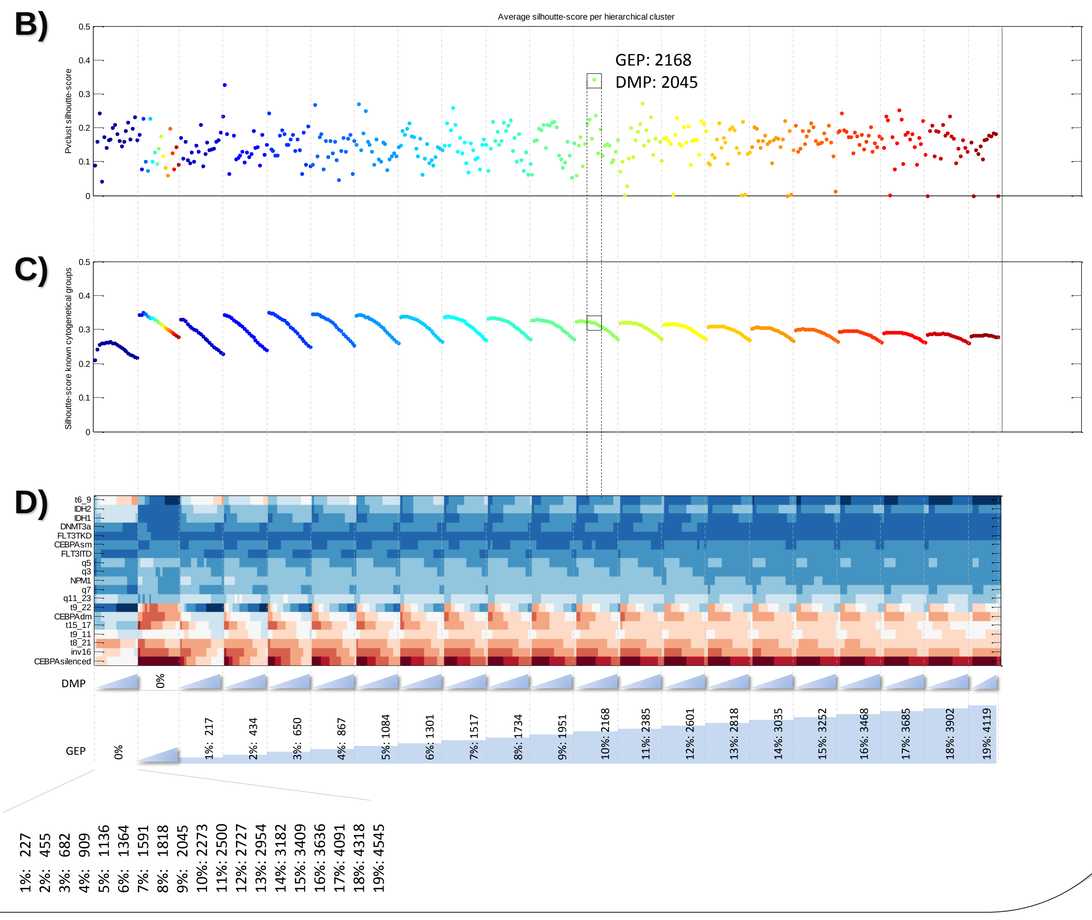
Acute Myeloid Leukemia is a highly diverse disease containing many cytogenetic and molecular abnormalities. We analyzed the DNA methylation (DMP) and gene expression profiles (GEP) of 344 AML patients using an unsupervised and supervised approach. We hypothesized to better characterize the disease phenotype by combing these features as these may result in specific patterns in cancer cells which reflect biological differences. The unsupervised approach segregates patients into 18 clusters, among them six clusters that are defined by the World Health Organization, such as inv16, t(15;17), t(8;21) and CEBPA double mutants. In addition we identified four novel AML subtypes that could not be explained by the enrichment of any currently known recurrent cytogenetic, molecular, morphological or clinical feature. Two of these clusters are categorized with good stability. One of these cluster could be characterized with pathways that are involved in the accumulation of red blood cells and highly predictable using 21 GEP and 3 DMP features, whereas the other cluster is characterized with T-cell related pathways and highly predictable with 9 GEP and 4 DMP features.

Framework to identify an robust hierarchical clustering

A major disadvantage of a hierarchical clustering approach is the uncertainty of the derived clusters, e.g. the use of different number of features may result in different patient-clusters. We therefore created a framework (A) to derive a robust hierarchical clustering using both gene expression and DNA methylation profiles.

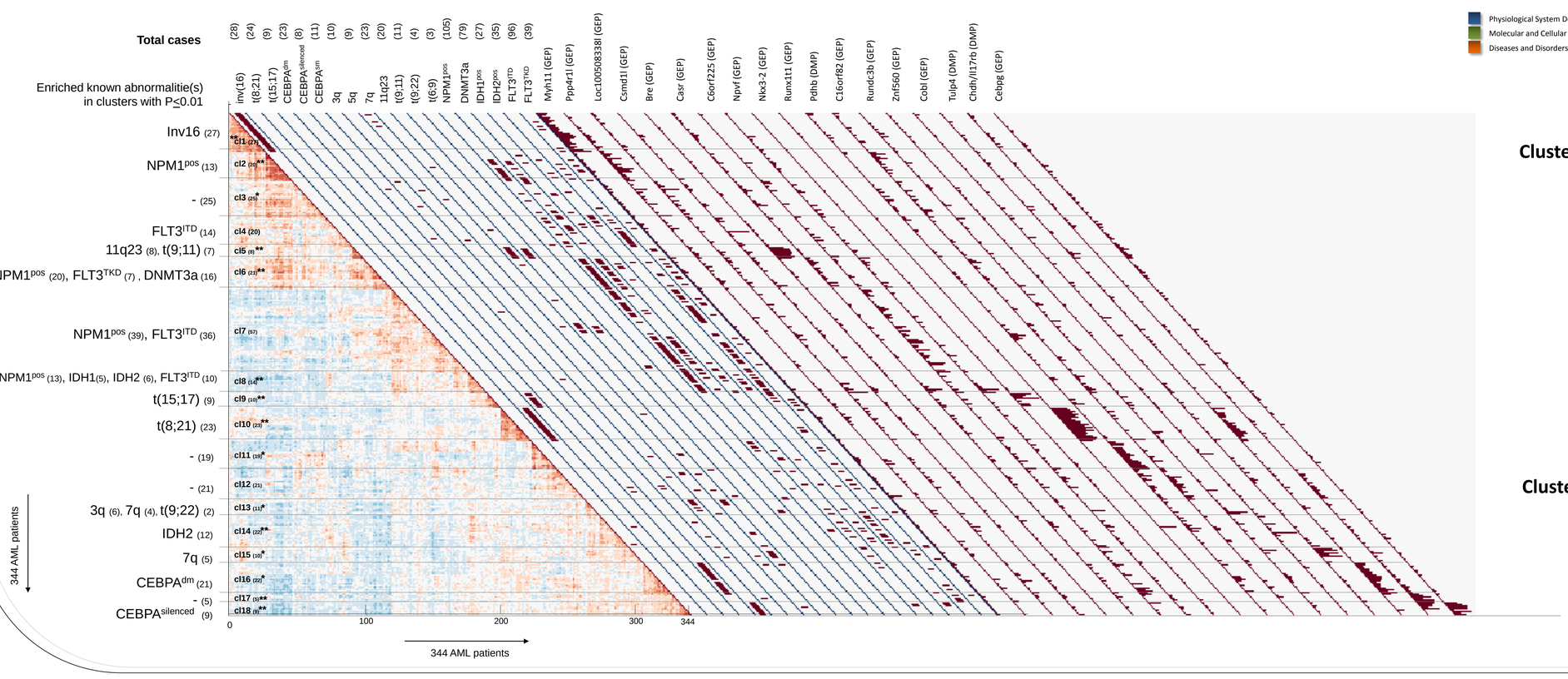
A)

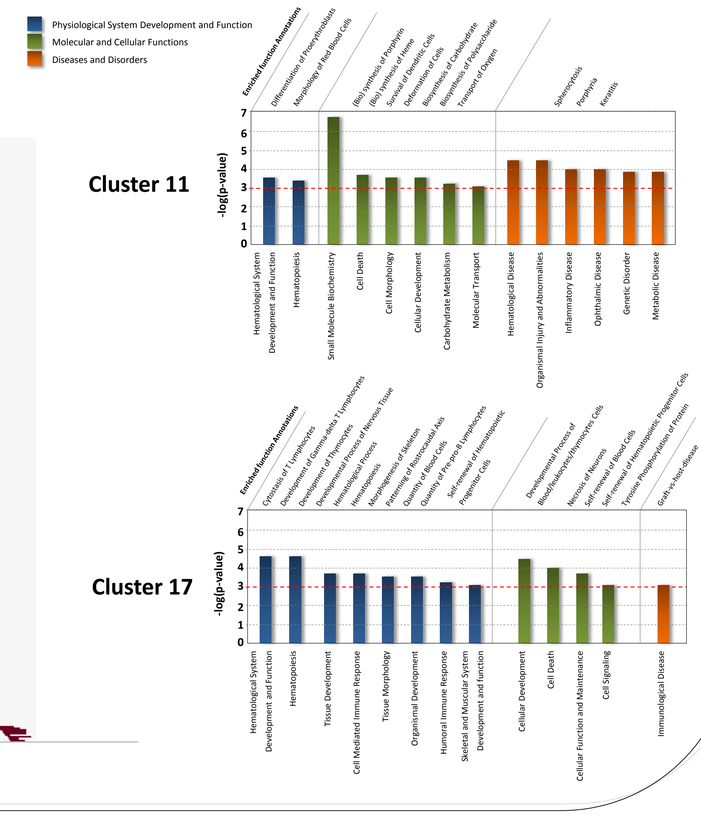




Identification of novel and known clusters

Eighteen clusters are identified with the use of 2168 gene expression and 2045 DNA methylation probe sets. Six clusters (1, 9, 10, 15, 16 and 18) are previously defined by the World Health Organization and, also detected by using solely GEP. Four clusters (cluster 3, 11, 12 and 17) could not be explained by the enrichment of any currently known recurrent cytogenetic, molecular, morphological or clinical feature.





Conclusion

The final hierarchical clustering, using both gene expression and DNA methylation patterns, is superior above the hierarchical clustering using solely GEP or DMP features. Our clustering is strongly supported by the data and thereby better resembling the disease phenotype.

Acknowledgement: This research is supported by the Dutch Cancer Society (KWF Kankerbestrijding) and the Center for Translational Molecular Medicine (BioCHIP)