SFI Centre for Research Training in GENOMICS DATA SCIENCE

Centre for Research Training

Science Foundation Ireland For what's next

Changes in gene expression among myeloid cells during AML progression

Sarah Ennis[1,2,3], Alessandra Conforte[1], Vishvesh Karthik[1], Pilib Ó Broin[2,3], Eva Szegezdi[1,3]

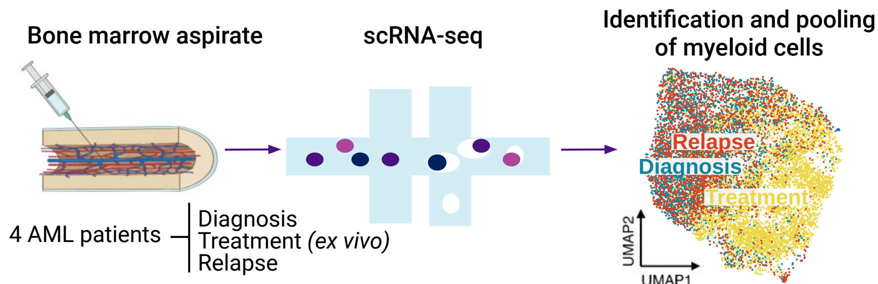
[1] Discipline of Biochemistry, School of Natural Sciences, National University of Ireland, Galway, Ireland
[2] School of Mathemathics, Statistics and Applied Mathematics, National University of Ireland, Galway, Ireland
[3] The SFI Centre for Research Training in Genomics Data Science

Introduction

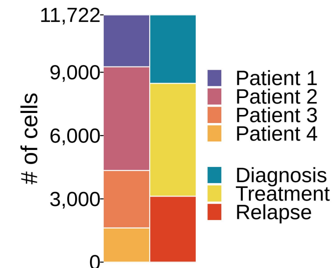
Acute myeloid leukemia (AML) is an aggressive blood cancer which causes an accumulation of immature myeloid cells in the bone marrow[1]. Recent studies have used single-cell technologies to demonstrate the heterogeneity present among AML cells at a genomic and transcriptomic level, showing that malignant cells co-exist at multiple stages along the myeloid cell differentiation trajectory[2,3]. To further investigate this heterogeneity and how it evolves during disease progression, we performed scRNA-seq of bone marrow aspirates from 4 AML patients at diagnosis, treatment (*ex vivo*) and relapse. We focused our analysis on the myeloid cell population, and investigated changes in gene expression among these cells across timepoints that could shed light on why some cells survive treatment and persist at relapse.

Results

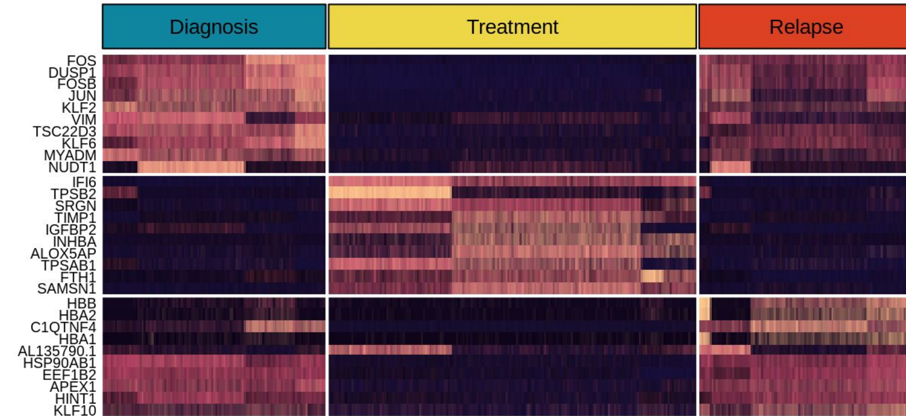
A total of 11,722 myeloid cells were identified. Comparing cells at each timepoint to all other cells; 3,449 genes were found to be differentially expressed. Genes upregulated at treatment are associated with interferon signalling, while genes upregulated at diagnosis and relapse are associated with regulation of splicing and mRNA processing. Gene Ontology (GO) terms related to neutrophil-mediated immunity were found to be enriched at all 3 timepoints.



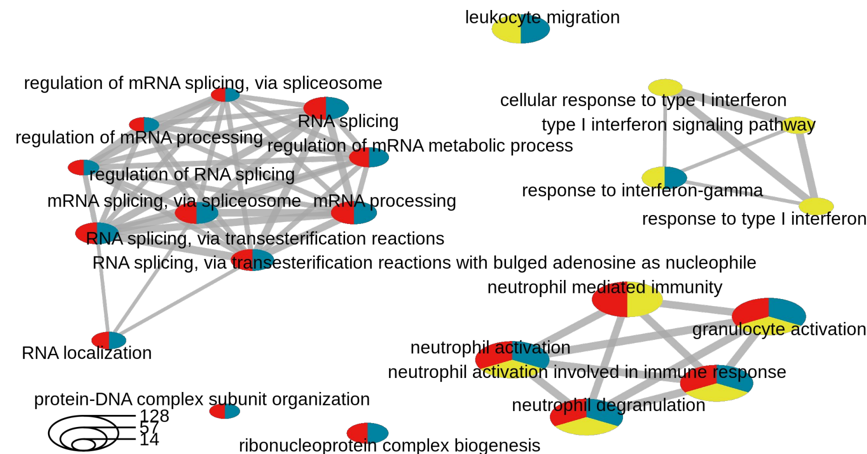
A. Summary of the workflow for this study. Bone marrow aspirates were obtained from AML patients at diagnosis, treatment and relapse, and subjected to droplet-based scRNA-seq. Myeloid cells in each sample were identified based on expression of marker genes and pooled into a single dataset.



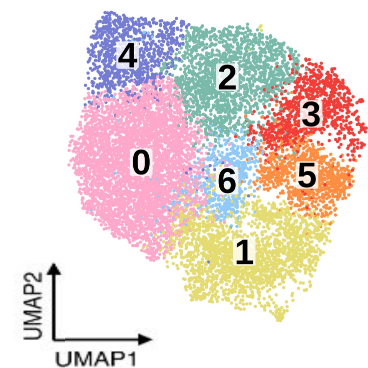
B. Number of myeloid cells coming from each patient (left) and timepoint (right).



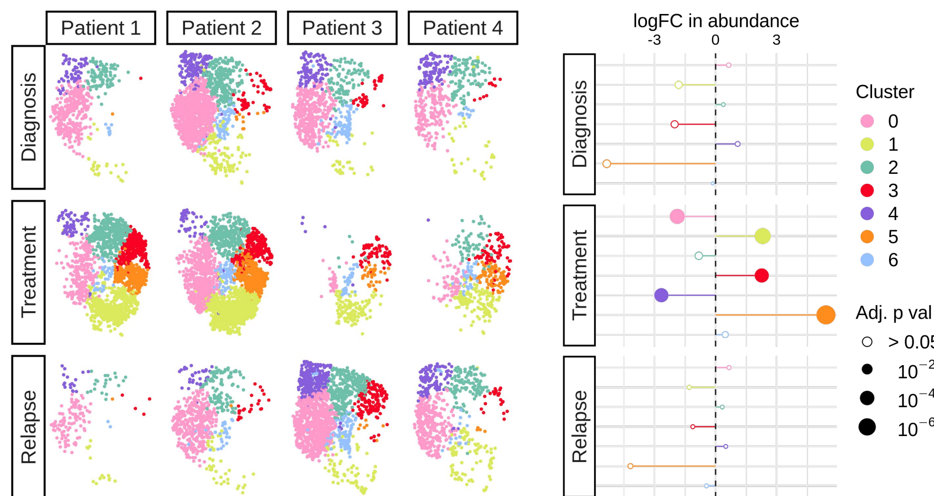
C. Expression of the top 10 upregulated genes at each timepoint. Differentially expressed genes (DEGs) were identified by comparing cells at each timepoint to all other cells. Genes with logFC >= 0.25, adj. p value <= 0.05 were considered to be upregulated.



D. GO terms enriched among upregulated genes at each timepoint. Nodes represent the 30 most enriched terms and edges represent overlapping genes.



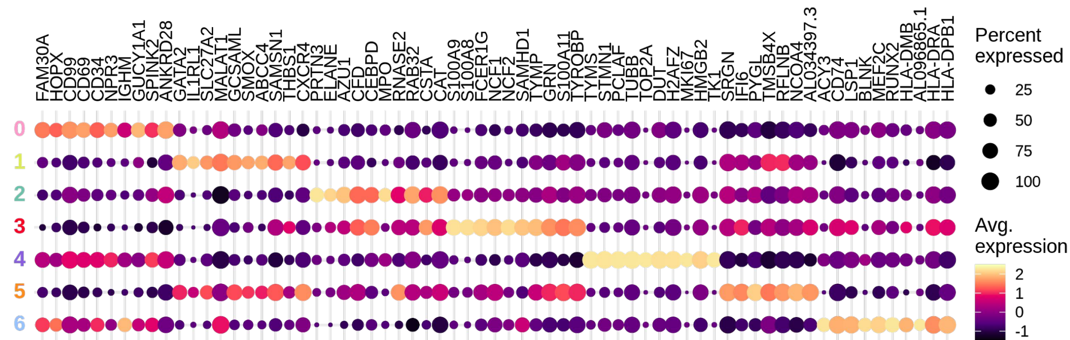
E. Clusters of cells within the myeloid population. 7 distinct clusters were identified by a community detection algorithm.



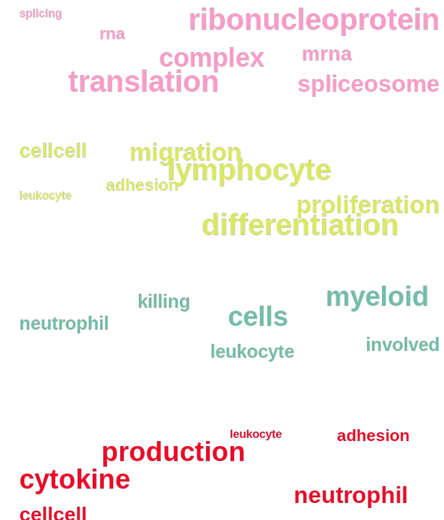
F. Differential abundance of clusters. Left: prevalence of each cluster across patients/timepoints. Right: logFC in abundance of clusters at different timepoints.

Results

Clusters 1, 3 and 5 are significantly more abundant at treatment while clusters 0 and 4 are significantly less abundant. Clusters 1, 3 and 5 are associated with cell-cell adhesion and leukocyte differentiation. Cluster 0 shows increased expression of genes related to RNA splicing, and cluster 4 is associated with mitotic processes.



G. Average expression of the top 10 upregulated genes in each cluster. DEGs were identified by comparing cells in each cluster to all other cells.



H. Most frequent words occurring among the 30 most enriched GO terms for each cluster.

Conclusions

- We performed scRNA-seq on bone marrow aspirates from patients with AML at diagnosis, treatment and relapse. To focus the analysis on the potentially malignant cell population, we **identified myeloid cells in each sample and pooled them** into a single dataset.
- In the population of cells surviving treatment, genes related to interferon signalling were upregulated, potentially implicating this pathway in the protection of malignant cells from treatment.
- Certain subsets of myeloid cells became more abundant at treatment. They showed higher expression of genes related to cell-cell adhesion and leukocyte differentiation and proliferation, indicating that **cells persisting after treatment are further along the myeloid differentiation trajectory**.
- Clusters of cells that were depleted at treatment were associated with translation and replication, indicating that **cells actively translating mRNA or undergoing mitosis are less likely to survive treatment**
- No significant differences in cluster abundance were observed for the diagnosis or relapse timepoints and only subtle gene expression changes could be found between them. This suggests that **even though diagnosis and relapse cells may differ in their clonal makeup, they are very similar at the transcriptomic level**.

Acknowledgements

This presentation has emanated from research conducted with the financial support of Science Foundation Ireland under Grant number [18/CRT/6214]. This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 777995

Methods (summary shown in **Fig A**)

Samples

Bone marrow aspirates were obtained for 4 AML patients at diagnosis and relapse from The Finnish Hematology Registry and Clinical Biobank. A portion of cells were treated with standard of care chemotherapy in an *ex-vivo* system[4]. Libraries were prepared using the 10X Genomics Chromium 3' Gene Expression kit (v3) and sequenced on an Illumina NovaSeq 6000.

Data Analysis

Reads were aligned and quantified using the kallisto|bustools kb-python wrapper[5]. Cell quality control, normalisation, graph-based clustering and visualisation were performed using the ScanPy python module according to current best-practices[6]. Clusters of cells expressing canonical myeloid marker genes (CD34, KIT, CD38, ANPEP) were isolated from each sample and merged into a single dataset. Harmony was used to generate a batch-corrected UMAP graph and the Leiden community detection algorithm (resolution = 0.5) was used to identify clusters of cells. The *FindAllMarkers* function from the Seurat R package was used to identify differentially expressed genes (DEGs) and the *enrichGO* function from the clusterProfiler R package was used to find enriched gene ontology (GO) terms among DEGs. Differential abundance analysis was carried out using the edgeR R package. Code is available on GitHub: Sarah145/myeloid_analysis.

References

1. http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/leukaemia-aml
2. van Galen P, Hovestadt V, Wadsworth Ii MH, et al. Single-Cell RNA-Seq Reveals AML Hierarchies Relevant to Disease Progression and Immunity. Cell. 2019;176(6):1265-1281.e24. doi:10.1016/j.cell.2019.01.031
3. Miles LA, Bowman RL, Merlinsky TR, et al. Single-cell mutation analysis of clonal evolution in myeloid malignancies. Nature. 2020;10.1038/s41586-020-2864-x. doi:10.1038/s41586-020-2864-x
4. Dhami SPS, Tirincsi A, Baev D, et al. Theranostic drug test incorporating the bone-marrow microenvironment can predict the clinical response of acute myeloid leukaemia to chemotherapy. Br J Haematol. 2020;189(6):e254-e258. doi:10.1111/bjh.16684
5. Páll Melsted, A. Sina Booeshaghi, et al. Modular and efficient pre-processing of single-cell RNA-seq. bioRxiv 673285; doi: https://doi.org/10.1101/673285
6. Luecken MD, Theis FJ. Current best practices in single-cell RNA-seq analysis: a tutorial. Mol Syst Biol. 2019;15(6):e8746.

HOST INSTITUTION

PARTNER INSTITUTIONS

University College Dublin
An Coláiste Ollscoile, Baile Átha Cliath

Trinity College Dublin
Coláiste na Tríonóide, Baile Átha Cliath
The University of Dublin