

Liquid biopsy

WHO 2022

ICC 2022

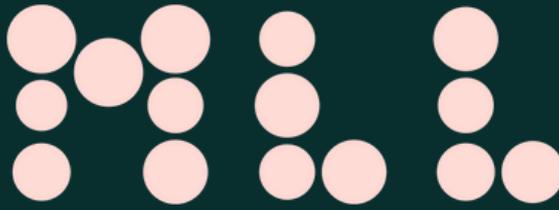
Genomnetzwerk

Multiples Myelom

Single-cell Sequencing

CHIP register

AI



IT TAKES ALL OUR INNOVATIVE
POWER, TO KEEP IN THE SAME PLACE

2023

RESEARCH REPORT

Dear research interested reader,

after another scientifically successful year, we would like to take the opportunity to summarize our findings of 2023. To further facilitate reading and to better convey our data, we have included various graphical representations and summaries and we hope you enjoy the content.

Also, let me briefly explain the cover and the chosen color scheme: Blue is commonly regarded as the color of innovation and thus symbolizes our constant striving for improvement and new developments for the benefit of our patients. Purple is believed to exist at the edge of the imagination, representing thinking that's outside average and stands symbolically for our scientific attempts to unravel the underlying mechanisms that drive and manifest the various leukemias and lymphomas.

Enjoy the reading!



Dr. Wencke Walter
Scientist Bioinformatics & author of this report

MLL Münchner Leukämielabor GmbH
Max-Lebsche-Platz 31
81377 München

T: +49 (0)89 99017-0
F: +49 (0)89 99017-111
info@mll.com
www.mll.com

Manager entitled to represent the company:
Prof. Dr. med. Dr. phil. Torsten Haferlach
Prof. Dr. med. Wolfgang Kern
Prof. Dr. med. Claudia Haferlach

Registry court: Munich District Court
Commercial registry number: HRB 155656
VAT identification number pursuant to section 27 a
VAT Act: DE 240985463

Person responsible for content pursuant to section 10 paragraph 3
Media Services Treaty [Mediendienstaatsvertrag] Prof. Dr. med. Wolfgang Kern

TABLE OF CONTENTS

04

Foreword

05

Refinement of the MLL
diagnostic workflows

12

Machine learning in
hematology

18

Scientific hot topics
in 2023

23

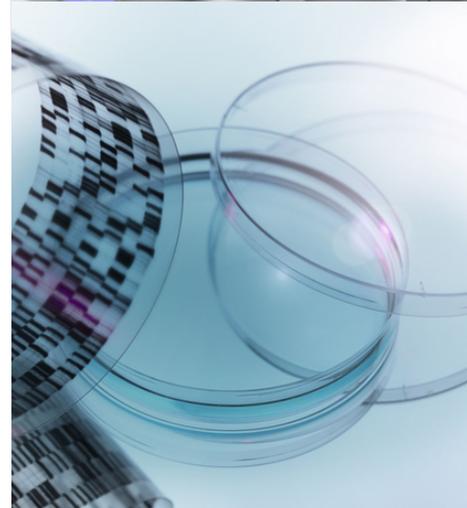
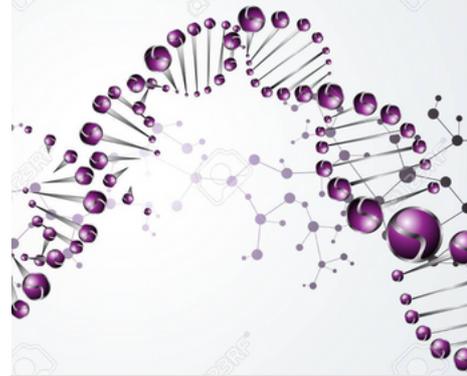
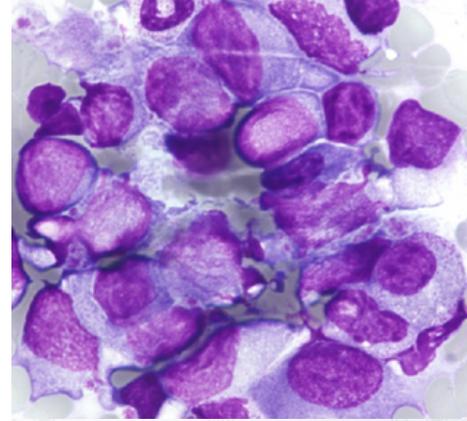
MLL Research &
Manuscript speed
dating

29

MLL's collaborative
projects

39

Future projects &
outlook



F O R E W O R D

It is with great pleasure that we present to you the culmination of our scientific endeavors in the form of this comprehensive research report for the year 2023. This year was marked by various challenges, including war, natural disasters, and inflation, yet we managed to make significant strides in research. The report reflects the tireless dedication and innovative spirit of MLL's team as we continue to push the boundaries of knowledge in the field of hematology.

Our diagnostic workflows continued to evolve and expand, enabling us to provide now the full range of diagnostics for the identification and characterization of hemoglobinopathies. MLL's step-by-step diagnostic workup and immediate processing of samples have proven integral to handling cases with unclear cytopenias, facilitating efficient and targeted diagnostic procedures for these patients. Additionally, our implementation of flow cytometry-based cell sorting has proven successful, particularly for enriching small populations of B-cell lymphomas for improved molecular genetic diagnostics. Moreover, a NovaSeq X series device was added to our sequencing facility, enabling faster and more cost-effective sequencing.

We also delve into the application of machine learning-based models in hematology diagnostics, providing valuable insights into the integration of AI algorithms to improve the accuracy, efficiency, and reproducibility of clinical diagnostics. This report emphasizes the transformative impact of AI on hematological diagnostics and the need for continuous research, validation, and collaboration to ensure the reliability of AI models in real-world situations.

Our report also covers the latest trends in hematology, including the adoption of liquid biopsies for early cancer detection, the refinement of diagnostic workflows through machine learning, the utilization of single-cell sequencing for precision oncology, and the importance of long-read sequencing technology in identifying somatic structural variations and its potential application in cancer epigenomics. We also discuss the ongoing debate about the two classification guidelines for hematologic neoplasms and their potential implications for patients and physicians.

Our team has been fortunate enough to participate in various national and international research projects this year. By sharing our extensive knowledge of hematologic neoplasms, biobank and data collection with other research groups, we have been able to facilitate a range of fascinating research projects. One of the major questions that drove several of our 2023 projects was whether clonal hematopoiesis of indeterminate potential (CHIP) had an impact on diseases beyond blood cancer. We were pleased to have contributed to a wide range of diagnostic fields, including cardiology, neuropathology, COVID-19, and lung fibrosis.

As we look to the future, we are excited about the new technologies and the promising prospects for their integration into clinical practice. This report not only chronicles our journey thus far but also illuminates the path ahead, inspiring continued exploration and discovery in the pursuit of clinical excellence in hematology.

We extend our gratitude to all our collaborators, partners, and supporters who have contributed to our scientific endeavors, and we invite you to join us in our pursuit of scientific excellence and advancement in the field of hematology.



A handwritten signature in black ink that reads "M. Meggendorfer".

Dr. rer. nat. Manja Meggendorfer, MBA,
Head of Molecular Genetics & co-author of this report

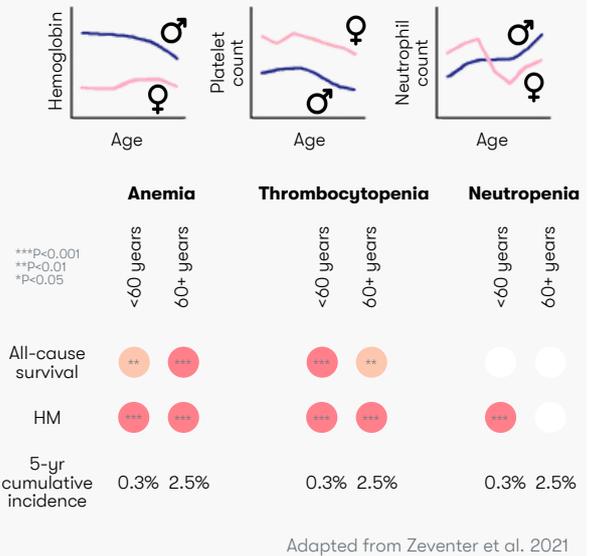


REFINEMENT OF
MLL DIAGNOSTIC
WORKFLOWS

MLL STAGING DIAGNOSTICS FOR UNCLEAR CYTOPENIAS

Background

One of the most frequent questions in hematological practices and hospitals is the clarification of cytopenias. Peripheral blood cytopenias are a reduction in blood cell counts, manifesting as anemia, neutropenia, and/or thrombocytopenia and may precede the development of hematological malignancies. Different studies have aimed to assess the prevalence of (unexplained) cytopenia in the general population and associated hematological outcomes in the context of aging. The 5-year cumulative incidence of hematological malignancies was comparably low (0.6%) for individuals aged >60 years, with higher incidences among those with anemia or thrombocytopenia but not neutropenia. Particularly with the advent of a group of indolent hematopoietic disorders such as idiopathic cytopenia of unknown significance (ICUS), clonal hematopoiesis of indeterminate potential (CHIP), and clonal cytopenia of unknown significance (CCUS) the clarification of cytopenias is required to distinguish these cases from MDS.



Diagnostic Workup of Unclear Cytopenias

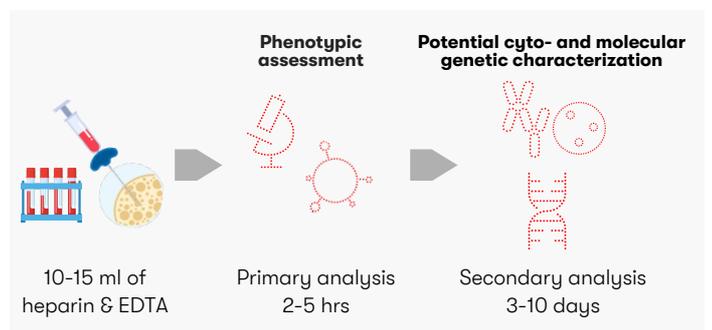
Cytopenias can be caused by a variety of factors, both benign and malignant. While laboratory testing and patient history may provide some insight, bone marrow aspiration is often required to paint a complete picture. Even after excluding reactive conditions, myelodysplastic neoplasia is only one potential malignant cause of cytopenia. Many other diseases, including acute leukemias, lymphomas and plasma cell myelomas, may be present in the bone marrow samples of corresponding patients.



Following a bone marrow collection (aspiration and punching recommended), it's often unclear which diagnostic methods will be most effective in characterizing the potential underlying bone marrow disease. MLL offers a "staged diagnostic workup" to address this uncertainty. This approach involves a stepwise examination of the bone marrow aspirate or blood sample. In the first step, disease phenotype is determined using suitable methods such as cytomorphology and/or immunophenotyping, or other neoplasms are excluded. In the second step, suitable methodological additions are decided upon, especially with regard to the cyto- and molecular genetic characterization of a defined neoplasia or the extended exclusion of a clonal event. By adopting this staged approach, patients receive more efficient diagnostics and targeted treatment, and the healthcare system is spared additional costs.

Efficient Implementation: A Prerequisite for Prompt Processing of Samples

Immediate processing of relevant samples upon their arrival at the laboratory is critical for efficient implementation. While phenotype diagnosis by cytomorphology and immunophenotyping at MLL takes less than 24 hours, prioritization is necessary considering the volume of up to 700 samples/day. As part of a plausibility check, hematologists and biologists perform an immediate assessment of each sample upon its arrival at MLL. Cases with indications for step diagnostics are then assigned to a prioritized workflow, ensuring that cytomorphology and immunophenotyping assessments are conducted within 2-3 and 4-5 hours, respectively, after sample arrival. If necessary, further diagnostics can then be performed on the fresh sample material on the same day, which ensures high quality of the downstream analyses (Figure below). In addition to clinical data, we require a sufficient amount of sample material (10-15 ml of heparin and EDTA each) from the clinicians sending in the samples. Our examination order allows for the specification of the methods used for a step-by-step diagnosis, thus limiting costs for inpatients.



DIAGNOSTIC POTENTIAL OF IMMUNOPHENOTYPIC ABERRATIONS IN MDS

For a long time, the diagnosis of MDS and its classification into subgroups was based solely on cytomorphological features. However, the WHO 2022 classification of tumors of hematopoietic and lymphoid tissue introduced cytogenetic aberrations as part of the diagnostic criteria. Nowadays, integrated diagnostics are known to be of great significance and include immunophenotyping and molecular genetics. Immunophenotyping can be used to detect anomalous antigen expression patterns in suspected or confirmed MDS, providing valuable information for diagnosis and prognosis. Recently, the International MDS-Flow (iMDS-Flow) working group of the European LeukemiaNet (ELN) conducted a prospective and multicenter study to validate the importance of different immunophenotypic markers and identify parameters that contribute to the clinical utility of immunophenotyping in MDS diagnostics.

Study Outline

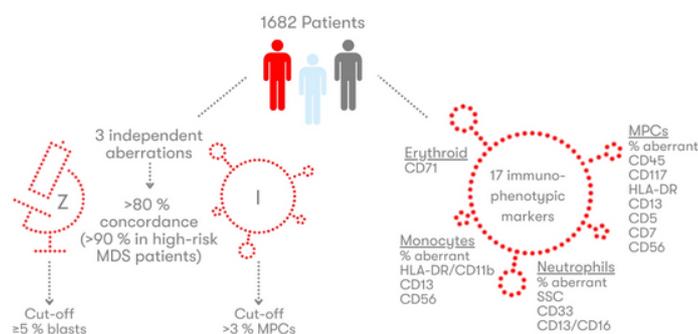
The aim of the prospective study, which was conducted in 19 laboratories, was to examine patients who underwent diagnostic work-up in bone marrow aspirates for suspected MDS due to unclear cytopenia. The study was outlined as follows:

- Patients with non-MDS malignancies causing cytopenia or patients diagnosed with myelodysplastic/myeloproliferative neoplasm (MDS/MPN) were excluded from the study, with the exception of both chronic myelomonocytic leukemia (CMML) and MDS/MPN with ring sideroblasts and thrombocytosis (MDS/MPN-RS-T).
- The results of flow cytometric evaluation were scored in each patient for each single parameter as “normal” or “abnormal” per the ELN 2012 recommendation.
- The category of ‘no signs of MDS’ was defined by no aberrant antigen expression, ‘limited signs of MDS’ by one or two aberrantly expressed antigens, or more than two aberrantly expressed antigens but confined to one cell lineage, and ‘in agreement with MDS’ by more than two aberrantly expressed antigens distributed to more than one cell lineage.
- The categories of ‘no signs of MDS’ and ‘limited signs of MDS’ were combined, when applicable, to ‘non-MDS’.
- Depending on the MDS subtype, patients were grouped into three different cytomorphological categories: ‘low-risk MDS’, ‘high-risk MDS’ and ‘CMML’.
- The Revised International Prognostic Scoring System (IPSS-R) data was collected along with data on age, gender, and ring sideroblasts.

The study included 1,682 patients, with a median age of 72 years. The male:female ratio was 1.6:1. MDS or CMML were confirmed in 1,029 cases, while 653 were diagnosed non-MDS. Ring sideroblasts were recorded in 824 cases and found elevated above 5% in 340. There were no cases diagnosed as normal but genetically diagnosed as MDS/CMML. Data on IPSS-R was available in 857 of 926 MDS cases.

Identification of Immunophenotypic Parameters Associated with MDS or CMML

This study identified 17 distinct immunophenotypic parameters that are independently associated with MDS or CMML in at least one of the subgroups (low-risk MDS, high-risk MDS, and CMML), as well as the overall cohort. Flow cytometric results demonstrated an 80% concordance with cytomorphology when using a cut-off value of three independently occurring MDS/CMML-related aberrations, regardless of the number of cell compartments affected. In high-risk MDS patients, the agreement was even above 90%.



Summary of the study results

Furthermore, the study found that the percentage of myeloid progenitor cells (MPCs) is a strong indicator for MDS and can be used in the diagnostic decisions. A cut-off of >3% MPCs, as detected by flow cytometry, was strongly associated with MDS/CMML. In fact, 98% of cases with >3% MPCs were found to have MDS by cytomorphology. Therefore, the cut-off value, similar to the value of 5% blasts in cytomorphology, can be used for diagnosis.

Impact on diagnostic workflow

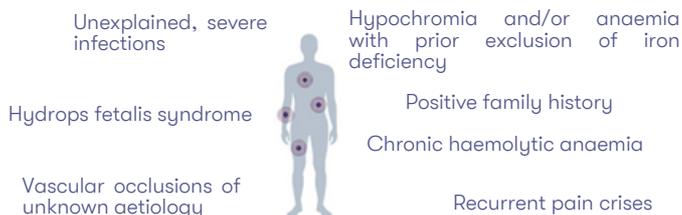
The iMDS Flow Working Group recommends that the 17 identified parameters should be integrated into routine diagnostics as a core set of markers for the immunophenotypic evaluation of suspected MDS cases. Wherever possible, it is suggested to investigate any additional markers listed in the ELN recommendations.



HAEMOGLOBINOPATHIES - THALASSAEMIAS AND ABNORMAL HAEMOGLOBINS

The Prevalence of Hemoglobinopathies

Hemoglobinopathies are a common form of monogenetic hereditary diseases that are increasing in global prevalence each year. Approximately 5% of the world's population carries these diseases, with the highest prevalence found in the Mediterranean regions, parts of Asia, the Middle East, and West Africa. However, due to migration, hemoglobinopathies are also becoming increasingly significant in Germany. For instance, thalassaemia is now a common diagnosis in patients with a corresponding genetic background when identifying the cause of microcytic, hypochromic anemia. In Germany, there are nearly 5,000 patients diagnosed with severe hemoglobinopathies and around 600,000 carriers. To provide clinicians with a central point of contact for this increasingly prevalent disease, we have expanded our range of examinations at MLL to include the full range of diagnostics for the identification and characterization of hemoglobinopathies. Hemoglobinopathy diagnostics are recommended for patients with various indications:

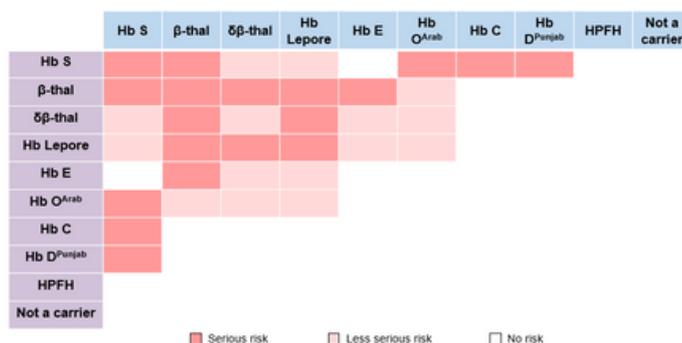


Brief Overview of the Main Hemoglobinopathies

Hemoglobinopathies include congenital quantitative (thalassaemias) or qualitative (abnormal hemoglobins) disorders of the synthesis of the globin chains of hemoglobin. The most common forms of thalassaemia are α -thalassaemia and β -thalassaemia. The clinical spectrum ranges from blood count changes without clinical symptoms to transfusion dependency. Of particular importance among the abnormal hemoglobins is HbS, which underlies sickle cell disease. This is associated with chronic hemolytic anaemia, recurrent vasoocclusive pain crises and the resulting organ damage. Other regularly occurring hemoglobin structure variants are HbC, HbE and HbD.

Stages in Diagnosing Hemoglobinopathies

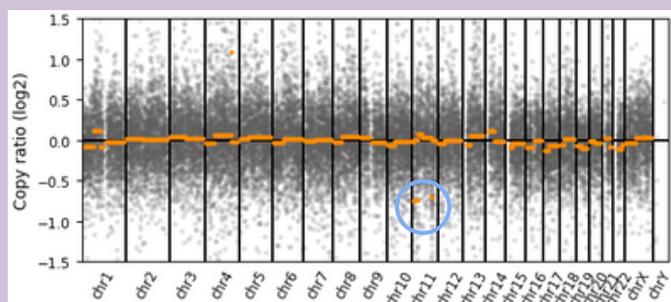
Hemoglobinopathies are genetic conditions that are inherited recessively, where more than a thousand mutations have been identified resulting in either hemoglobin variants or thalassaemias. Those who carry hemoglobin variants are generally asymptomatic, and may not be aware of their status unless screened. However, if prospective parents are identified as carriers of thalassaemia, they should receive prenatal genetic testing with counseling due to the increased risk for the child:



Diagnosing hemoglobinopathies is a process that occurs in two stages. The first stage involves evaluating red blood cell counts and differentiating hemoglobin. The second stage requires molecular genetic analysis to confirm any suspicious findings and provide further detail. Detecting α -thalassaemia is achieved through deletion-specific PCR, while abnormal hemoglobins or β -thalassaemia can be confirmed by NGS. Large deletions that cause rare forms like $(\gamma)\delta\beta$ -thalassaemia can be detected using multiplex ligation-dependent probe amplification (MLPA). In 2023, MLL introduced diagnostic methods that have since been used to examine 831 patients' red blood cell counts and hemoglobin fractions. An additional 444 patients underwent MLPA and NGS analysis, and a handful of cases were further analyzed through WGS/WTS to provide additional genetic abnormality details. In one instance, chromosomal regions 11p15, 11p14, and 11q22 had structural changes and deletions. These changes were identified as acquired after being compared to control material, indicating the presence of clonal hematopoiesis (see box below).

Case report

Blood from a 34-year-old woman with iron deficiency was sent to the MLL for diagnostic workup. The cytomorphological findings partly indicated the presence of a hemoglobinopathy and further tests were requested. Using the CNV spike-in panel, two large deletions involving the HBB gene cluster were detected in chromosome 11. WGS confirmed the deletions and provided detailed information regarding the breakpoints and affected genes.

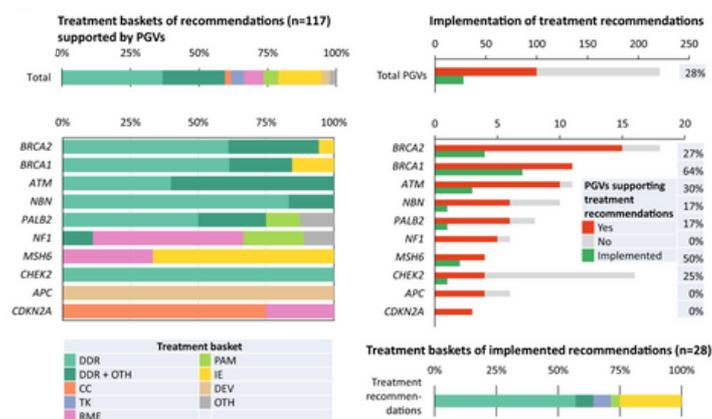


MYELOID NEOPLASMS WITH FAMILIAL PREDISPOSITION

Millions of people are diagnosed with cancer every year and at times several members of the same family are affected. This may be due to the influence of the shared environmental factors and a similar lifestyle, or it may be a coincidence. On the other hand, cancer can also be caused by heredity, which leads to a higher risk of tumour diseases and is present in approx. 5-10% of all tumours. In addition, genetic tumour risk syndromes are characterised by an early age of onset and a higher probability of multiple tumour diseases in one person. Genetic testing can reveal the cause of a tumour risk syndrome in such cases and plays an important role in counselling relatives and assessing their tumour risk, recommending cancer prevention measures or prophylactic surgery and treatment options.

Germ Line Variant Interpretation

In February, we welcomed Dr. Arne Jahn, group leader for genetic tumour risk syndromes at the Institute of Clinical Genetics, University Hospital Carl Gustav Carus, TU Dresden, to our scientists' meeting, who presented his work on comprehensive cancer predisposition testing within the prospective MASTER trial and provided insights into germ line variant interpretation. Interestingly, analysis in the MASTER study showed that 10% of patients carried a pathogenic germ line variant (PVG) in 35 genes associated with autosomal dominant cancer predisposition, up to 75% of which were unknown prior to the study.



Treatment recommendations supported by PGVs & their implementation .

Hence, genetic germ line testing in patients with rare cancers can identify the very first patient in a hereditary cancer family and lead to clinical benefit across a broad spectrum of entities. The group has also advocated that the routine introduction of germ line genetic testing in precision oncology, combined with the harmonisation of workflows to evaluate germ line variants, will increase clinical utility and advance research.

Myeloid Neoplasms with Germ Line Predisposition

In hematological malignancies, most efforts have focused on the identification of acquired genetic alterations, while the role of germ line aberrations has long been underestimated. Due to the broad heterogeneity in penetrance, clinical symptoms and age of onset, as well as the fact that somatic disease-associated mutations also occur in most predisposition genes, the existence of a predisposition is often not recognised.

Table 1: WHO 2022 classification of myeloid neoplasms with germ line predisposition

Myeloid neoplasm classification	Genes involved
Myeloid neoplasms without a preexisting disorder or organ dysfunction	<i>CEBPA</i> <i>DDX41</i>
Myeloid neoplasms & preexisting disorder or organ dysfunction	<i>RUNX1</i> , <i>ETV6</i> , <i>ANKRD26</i>
Myeloid neoplasms & other organ dysfunction	<ul style="list-style-type: none"> Germ line <i>GATA2</i> mutation BM failure syndrome Telomere biology disorders JMML associated with neurofibromatosis, Noonan syndrome/-like disorders, Down syndrome
	<i>GATA2</i> Multiple <i>TERT</i> , <i>TERK</i> <i>NF1</i> , <i>PTPN11</i> , <i>CBL</i> , <i>KRAS</i>

With the inclusion of the category “Myeloid neoplasms with germ line predisposition” in the WHO guidelines (table 1), genetic predispositions in hematological neoplasms have received more attention and it is becoming increasingly clear that these observations can now be extended to lymphoid malignancies, as demonstrated by ICC 2022.

Genomic Landscape of *DDX41* Variants

Germ line mutations in *DDX41*, which are detected in up to 4% of adult patients with AML and MDS in routine diagnostics at MLL, are clinically associated with an increased lifetime risk of myeloid neoplasms. Interestingly, *DDX41* germ line mutations are often identified in patients without a family history of hematological neoplasia. In a recent joint paper published by MLL and the Peter MacCallum Cancer Centre, the known clinical and genomic landscape of *DDX41* variants has been expanded through analysis of a large cohort and several new recurrent pathogenic germ line and somatic variants have been described (see page 27).

The clinical importance of detecting variants in predisposition genes has led to their rapid inclusion in many diagnostic gene panels (e.g. MLL myeloid panel), as they play a role not only in the treatment of patients, but also in potential predictive testing.

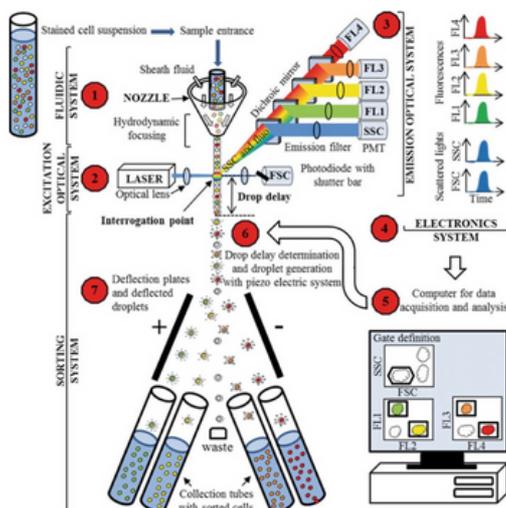


LYMPHOMA CELL SORTING TO ENRICH LOW DEGREE INFILTRATION CLONAL B-CELLS

Exploring the Benefits of Flow Cytometry-Based Cell Sorting

Flow cytometry-based cell sorting is a widely used method for enriching cell populations, either to enhance the detection limit of genetic changes in diagnostics or to analyze specific cell populations in a targeted manner. While most molecular genetic tests are either analyzed from bulk samples or in research by single-cell sequencing, cell sorting based on flow cytometry offers an efficient method by enriching or isolating the cells as a cost-effective alternative.

Our immunophenotyping department has established cell sorting as an effective method, initially testing it in research projects, and subsequently incorporating it into routine diagnostics for enrichment of small cell populations. Technically, cell sorting is a highly sophisticated method. During cell sorting, energy is supplied to a piezoelectric crystal, which vibrates continuously, breaking the stream of cells into droplets. These droplets detach from the stream a few millimeters below the nozzle. When a particle is detected and meets the predefined sorting criteria, an electric charge is applied to the stream as soon as the droplet containing that particle breaks away. After detaching from the stream, the charged droplet passes through highly charged deflection plates. Electrostatic attraction and repulsion guide each charged droplet into a predefined collection device, such as a plate or tube. In the standard mode, uncharged droplets fall into the waste bin. In summary, cell sorting utilizes vibrations, charges, and clever physics to separate and collect specific cells. It's like sorting marbles by color, but on a microscopic scale!



Picot et al. 2012. *Cytotechnology* 64, 109–130.

Several challenges arose, necessitating optimization in the process. After cell sorting, DNA or RNA must be isolated from the cells for use in molecular genetic assays. It is

crucial that the cell count is not too low, as insufficient starting material would hinder subsequent analysis. Additionally, cell integrity is essential - only intact cells allow successful DNA/RNA isolation. RNA, on the other hand, rapidly degrades in stressed cells or undergoes programmatic changes, affecting its transcriptional profile and therefore might skew the results. Material age and sorting duration significantly influence this process. Faster sorting may yield more vital cells, but purity may suffer. Finding the right balance is very challenging, which we have successfully realised in the enrichment of small populations of B-cell lymphomas and DNA extraction.

Enrichment of Small B-Cell Lymphomas for Improved Molecular Genetic Diagnostics

Traditionally, B-cell lymphomas are classified based on histology and immunophenotype. However, genetics increasingly intersect with these classifications to provide prognostic information. Unfortunately, the application of NGS is limited by the sensitivity of small B-cell clones (<5%). A study presented by V. Ecker at ASH 2023 aimed to overcome this limitation by enriching small clones (>15%) via cell sorting to enable robust NGS panel sequencing for genetic characterization and risk stratification.

In total 43 samples with <5% clonal B-cells were applied for cell sorting and clonal B-cells were enriched to a median of 17%. 13/43 samples were used in a pilot study to optimize the sorting procedure and the extracted DNA from all samples underwent NGS panel sequencing (66 genes, ~1,500x coverage). For technical and biological validation, native sample material was sequenced in parallel for a subset of patients (n=30).

In 19/38 samples, we identified mutations associated with B-cell lymphoma, including *MYD88*, *CXCR4*, and *NOTCH1*. All cases with negative NGS results (n=15) after sorting, were also negative when tested by NGS on matched native sample material without prior enrichment of clonal B-cells. Among the 15 cases with positive NGS results after sorting, only four (27%) exhibited mutations when tested on native sample material. The median variant allele frequency (VAF) in the 15 cases positive after sorting was 8.8%. As expected, the VAF was lower in native sample material compared to sorted sample material in the four cases where mutations were detected in both tests (3.1% vs 8.9%).

Our results show that cell sorting, combined with NGS, can unveil the intricate genetic landscape of B-cell lymphomas for improved molecular genetic diagnostics.



ON THE WAY TO HIGH-THROUGHPUT WGS WITH THE NOVASEQ X

The Human Genome Project created the basis for the comprehensive annotation and analysis of each individual's genetic information. In subsequent years, however, it has become clear that it is not enough to simply read the information, but that in the case of cancer and other diseases, clinically relevant gene-disease associations must also be identified that can be examined for genetic abnormalities. The genetic information of two people is 99.9% identical, but the remaining 0.1% contains a high degree of diversity, so that large cohorts of different age groups and ethnicities are necessary in order to reliably distinguish individual variability from disease-associated changes and to be able to develop targeted therapies.

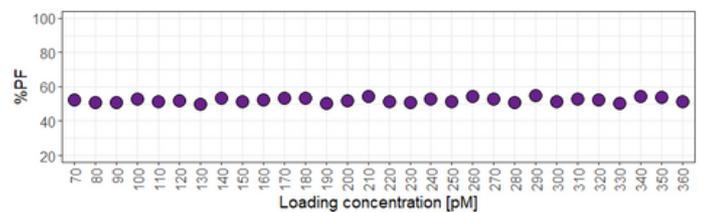
Implementation & Optimization of the NovaSeqX

In recent years, many large-scale sequencing projects have been carried out and identified variants, as well as clinical and demographic metadata, have been collated in comprehensive databases. In addition to the still existing problem of incomplete variant annotation, which is currently being tackled with various ML-based methods, two challenges still stood in the way of the introduction of WGS into routine diagnostics: the comparatively high costs and the relatively long turnaround times. Both points have been addressed and drastically reduced with the NovaSeqX series from Illumina.



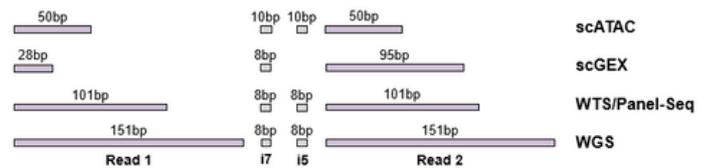
We already gave an outlook on the NovaSeqX in the last research report and in July 2023 the time had come: the first NovaSeqX+ was installed and launched at the MLL. Initially equipped only with the 10B FC, we set about validating the device for the Illumina TruSeq PCR-free libraries. However, optimising the loading concentration for the NovaSeqX turned out to be quite challenging.

PCR-free libraries cluster less efficiently and bias can be introduced during cluster amplification. Also, compared to other library types, the sequencing of PCR-free libraries requires a higher loading concentration. Based on the recommendations by Illumina, 90pM should be loaded to achieve ~60-65% PF. However, when we tested loading concentrations ranging from 70pM up to 360pM the %PF never surpassed ~55% PF, a less than optimal result (Figure below). Fortunately, we are not the only ones who encountered this problem so we are confident that the issue can be resolved in the near future.



Flexible Run Setups

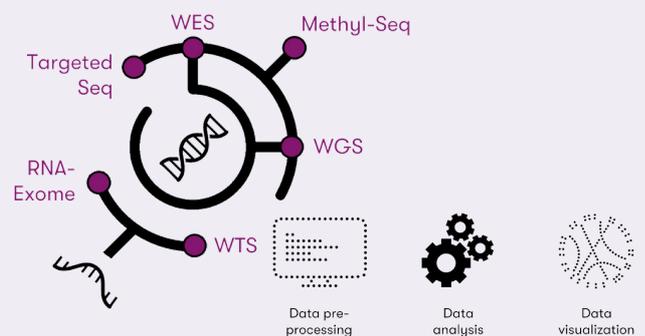
One of the great advantages of the NovaSeqX series is the possibility to load the 8 lanes with completely different libraries and to adapt the connected downstream analysis directly to the assay-specific sequencing settings. Single-cell sequencing, for example, often has very specific read lengths, as the first read only contains the barcode information for the individual cells and is therefore difficult to combine with other DNA-based library preps (Figure below).

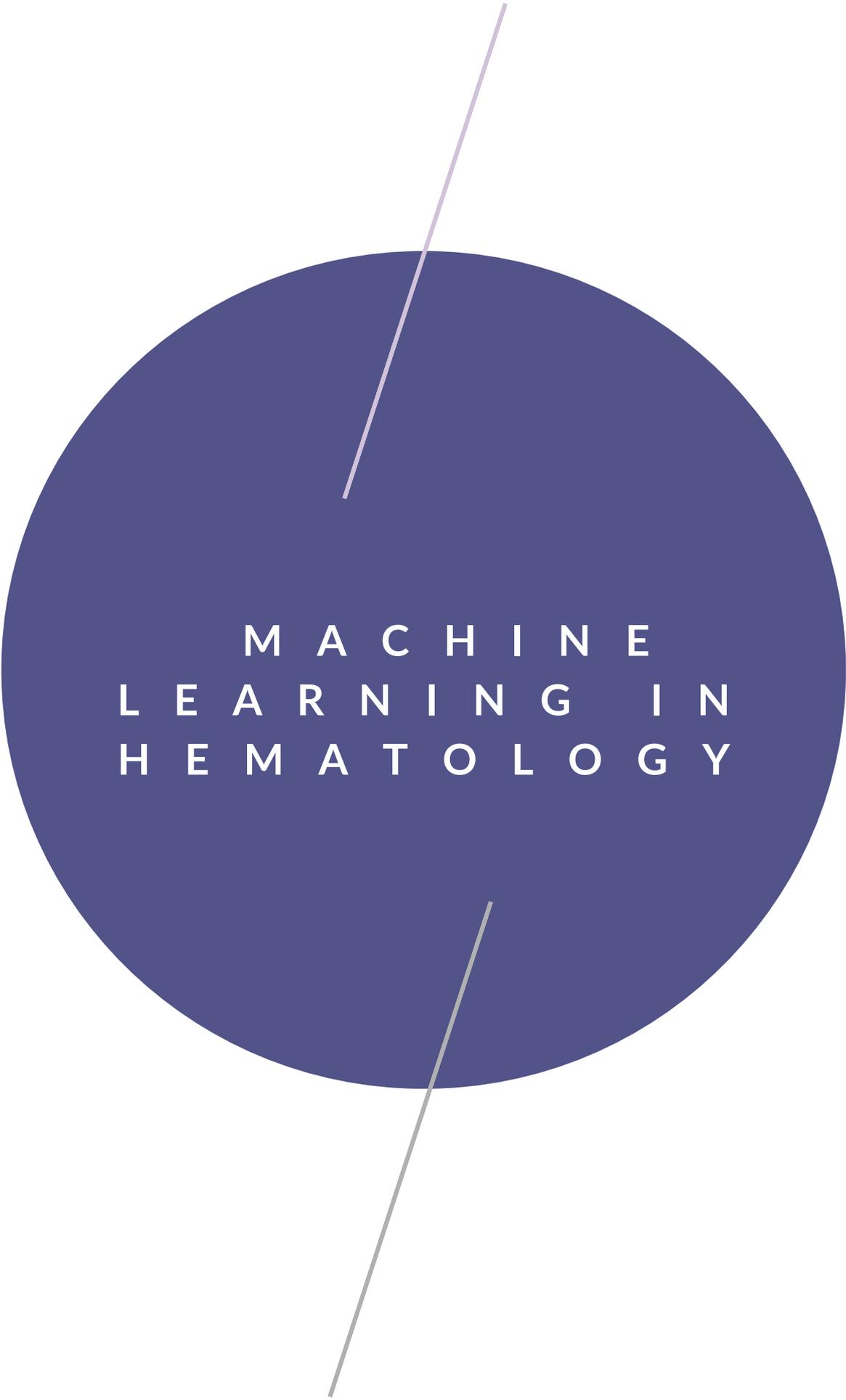


In addition, some manufacturers combine their unique molecular identifiers (UMIs), which are relevant for low-frequency variant calling, with the i7 index, which means that the number of cycles for this read also deviates from standard settings. However, in combination with the Illumina DRAGEN Bio-IT platform, all of these scenarios can be perfectly handled and we are looking forward to a faster and more flexible sequencing plan in the future.



This year, we have added one NovaSeq X series device to our sequencing facility in order to be able to offer faster and more cost-effective sequencing. It is possible to order the loading of either an entire flow cell or individual lanes. If we also carry out the library preparation, we take care of the optimum loading ourselves.





MACHINE
LEARNING IN
HEMATOLOGY

HOW TO READ ARTICLES THAT USE MACHINE LEARNING METHODS

New clinical diagnostic tools use complex ML methods, which can be difficult to understand for those unfamiliar with the field. However, with a clear understanding of basic concepts and critical thinking, valuable insights can be gained from scientific papers that utilize these methodologies. Moreover, a comprehensive understanding will enable the reader to distinguish between poor scientific studies, comprehend the strengths and limitations of the presented algorithms, and learn how to overcome them.

Step 1 Understand the Problem Being Addressed

When delving into an ML paper, it is crucial to comprehend the clinical or scientific impact of the problem being addressed. If the goal is to solve a clinical problem, it is important to assess how the algorithm's outcome or recommendation can aid physicians or researchers in their practice, and whether it is advanced enough to be integrated into clinical workflows. It's worth noting that significant clinical problems in healthcare often have an impact on patient outcomes or operations.

Assess the Quality of the Data

The size of the training, validation, and test sets should be sufficient to build a reproducible and generalizable model. The data must be appropriate, relevant, and representative of the population being studied. Also, it is essential to evaluate bias in the data, including disparities or overrepresentation of certain characteristics, such as race, gender, or socioeconomic status, to avoid producing a biased model.

Step 2

Step 3 Familiarize Yourself with the ML Methods Used

Many authors provide a brief method overview, but it's important to have a good understanding of the methods used and underlying concepts, as well as the key issues in building ML models and how they were addressed in the respective paper.

Evaluate the Results

For the evaluation of the results, different factors should be considered: it is important to divide the cohort appropriately and report matrices such as precision, recall, and F1 score, along with the entire confusion matrix. Evaluation should be done on the test cohort, not on the training cohort. To detect overfitting or underfitting, one can examine training and validation accuracy, learning curves, cross-validation results, and test set performance. Explainability of the ML model is also crucial in healthcare to ensure that the model does not use irrelevant patterns in the data set to make predictions.

Step 4

Step 5 Critically Evaluate the Conclusions and Implications

When evaluating the conclusions and implications, several key factors should be considered. These include limitations, generalizability, and potential impact on the field. Furthermore, if the study aims to develop a new ML model, it is crucial to consider practical deployment options. As such, authors need to outline their plans for making the model available to the public and identifying the necessary steps to deploy the model.



AI IDENTIFIES DIAGNOSTIC CELLS OF GENETIC AML SUBTYPES

Summary of the Study

Blood and bone marrow smear microscopy is a crucial diagnostic tool, but the process is time-consuming, tedious, and lacks standardization. SCEMILA is an algorithm trained to distinguish blood smears from healthy stem cell donors and four types of AML. It classifies a patient's blood sample based on roughly 400 single cell images and highlights the most relevant cells for the decision making. Hence, the results can be cross-checked with human expertise. SCEMILA aims to contribute to a future where machine learning and human experts collaborate for high-performance blood cancer diagnosis.

? What is explainable AI?

Many of the current machine learning methods are so intricate that it can be challenging for humans to comprehend how a particular result was achieved. Enter Explainable AI (xAI) – an initiative that aims to develop strategies to boost the transparency of these "black box" models and promote trust and acceptance of AI, particularly in fields that impact human life, such as medicine. Additionally, xAI strives to shed light on the features and underlying factors that explain why and how an AI decision was made.

Requires expert knowledge



Tedious

Subjective

Time-consuming

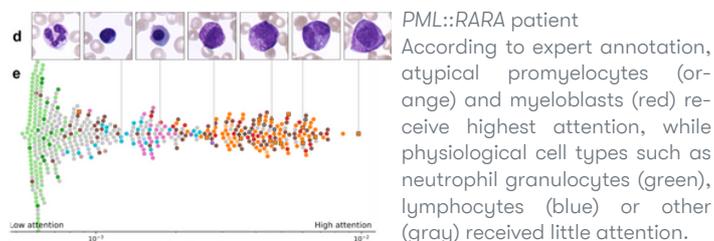
Low level of reproducibility

Background

Deep neural networks can extract information from large image data sets with single cell resolution, such as discriminating cancer types, predicting patient survival, and classifying blood and bone marrow cells. Towards explainability in a tissue-to-cell level, attention mechanisms have been applied to identify relevant regions in histological scans, but a quantitative comparison to pathologists' evaluations is lacking. Genetic subtype discrimination of AML derived blood smears is an ideal use case for xAI because of established morpho-genetic correlations for some subtypes and the availability of well annotated training data. AML diagnosis is based on the identification of >20% blasts and specific cell anomalies characteristic for certain subgroups. However, due to low abundances of these cells, reliable detection and classification can be challenging.

Accurate Classification of AML Subtypes by Single-cell Based Multiple Instance learning

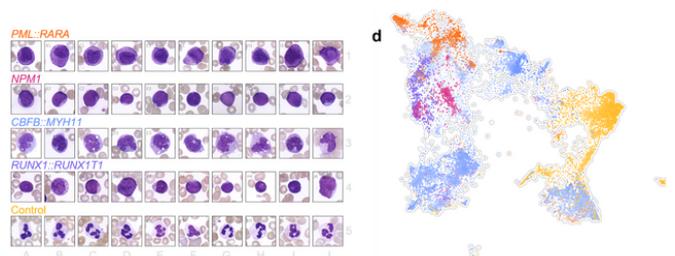
High classification performance with F1 scores of >0.69 showed that single-cell blood smear morphology can be used to automatically discriminate between AML genetic subtypes and that the model learned relevant morpho-genetic correlations. Moreover, cells that are known to be specific for the different genetic subtypes received the highest attention, whereas healthy cells and cells not related to AML received low attention (Figure below).



Subsequently, individual cell images were passed to the trained SCEMILA model to deconvolute each patient's blood smear composition. For 21/24 APL patients the subtype was correctly predicted. The three misclassified cases were all predicted as *RUNX1::RUNX1T1* and in one of these cases no AML diagnosis would be inferred from the single-cell images alone in a clinical setting.

Identification of Subtype Specific Single-cell Features

PML::RARA cells were mostly large and granulated with occasional bilobed shapes. *NPM1* cells were small with little cytoplasm and a cup-like morphology. *CBFB::MYH11* cells had myelomonocytic features, while *RUNX1::RUNX1T1* showed the smallest cells with the least amount of cytoplasm and occasional larger myeloid progenitor cells. The study also used a UMAP embedding to map single cells, revealing AML subtype-specific cell clusters. *PML::RARA* cells formed two promyeloid clusters (orange), while two types of *NPM1* blasts appeared (red) that partly overlapped with *RUNX1::RUNX1T1* blasts (blue). Cells predicted to be *CBFB::MYH11* populated a confined region associated with monocytes and monoblasts, but surprisingly, also the lymphocyte region.



Hence, SCEMILA can support routine diagnostics by identifying rare but diagnostically relevant cells.



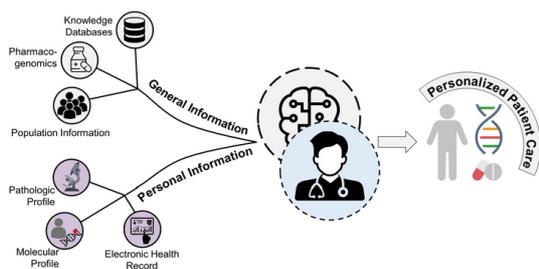
AI IN HEMATOLOGICAL DIAGNOSTICS: GAME CHANGER OR GADGET?

The Emergence of AI in Hematology

Despite the increasing awareness of the importance of genetic context, hematological diagnostics are still largely based on the evaluation of phenotypic features. Traditionally, diagnostic studies have been conducted using disease entities defined by microscopy, according to the guidelines of the World Health Organization (WHO) classification. However, this method has its limitations, as diagnostic ambiguity is still a common issue, and the quality of the results is reliant on the experience and skills of the operator. To address these concerns, there has been a growing interest in the implementation of automated processes that can generate standardized and structured data, thereby supporting a variety of downstream AI development projects.

In hematology, AI-based methods have been developed to automatically detect and classify single-cell images or superordinate patterns from digital microscopy. They have also been used to support the evaluation of flow cytometric data, to aid cytogeneticists in karyotyping, to develop personalized models integrating multiple data sources, to estimate the response to certain therapies, and to accurately predict the prognosis of various leukemias.

The technology is rapidly evolving and ready for clinical use, as demonstrated by the FDA's approval of 38 AI/ML-enabled medical devices in 2021 alone. At the annual meeting of the American Society of Hematology (ASH), the largest conference for hematologists, AI-based systems and sophisticated machine learning models are playing an increasingly important role. In 2021, 69 abstracts were accepted for publication that reported either AI-based systems or sophisticated machine learning models with an impact on diagnosis, prognosis, or treatment decisions, an increase of 40% from the previous two years.



AI can aid in advancing hematology by collecting and analyzing data from various sources to provide specialized systems that support clinicians in diagnosis and treatment guidance.

Understanding the Different AI methods in Hematology

In medicine, the terms “artificial intelligence” and “machine learning” are often used interchangeably. However, it's important to distinguish between the two. ML refers to the automatic detection of patterns and associations within the data, while AI strives to simulate human behavior and intelligence. Present-day medical ML-based algorithms fall under the category of weak or narrow AI. These algorithms are designed to perform a single task based on specific datasets within a predefined range. For example, an algorithm may be used to automatically classify cell types or chromosomal aberrations. In contrast, strong AI is a type of system that can find workable solutions to various problems without the need for human intervention. At present, however, none of the models used in medicine can reliably perform multiple tasks.

Various ML techniques are used in hematology, which can be divided in three major methods:

- Supervised learning is the most commonly used technique to build clinical models and requires well-curated datasets with labeled observations to learn a function that maps the input data (e.g. images of different cells) to the intended output (e.g. cell type classification).
- Unsupervised learning can be used for the exploratory analysis of unlabeled data to infer underlying factors and their interaction (e.g. refinement of disease subgroups).
- In reinforcement learning, a learner or agent performs different operations to interact with an environment with the goal of maximizing its expected payoff, similar to the situations when clinicians have to adjust their action (= treatment) according to a patient's conditions (e.g. genetic profile).

The performance of the methods can be significantly improved by effectively combining them with deep learning.

Conclusion

AI-based hematological diagnostics can improve diagnosis by assisting hematologists in the decision making process. While AI has limitations in generalization and performance on underrepresented or divergent populations, it can expertly perform tedious tasks. However, multi-center studies are necessary to validate AI models for reliable performance in real-world scenarios. Moreover, clinicians should always check and challenge all AI-based results.



INTEGRATION OF AI APPLICATIONS IN MLL'S DIAGNOSTIC WORKFLOWS



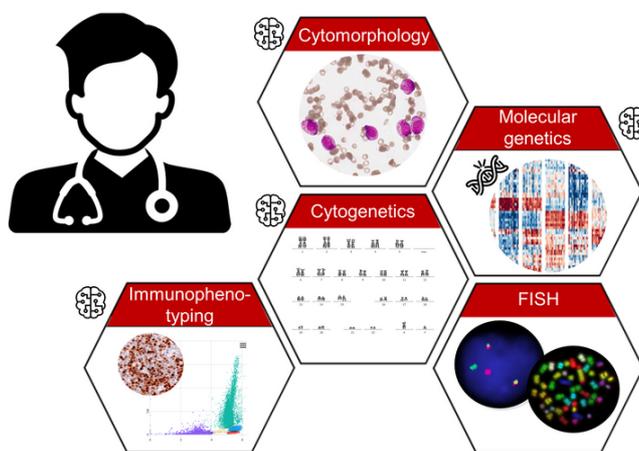
Prof. Torsten Haferlach recently chaired a captivating special interest session at ASH in San Diego, titled "AI in Haematology: Where Do You Stand in 2023?". The event featured talks from various professionals in the field, including Dr. Weida Tong, Dr. Yates Coley, and Prof. Haferlach himself. The talks focused on how AI can be used to enhance drug discovery and in the Hematology Clinic for diagnostics and therapy decisions. They also discussed safety assessments and ethical issues. During the session, Prof. Haferlach presented MLL's diagnostic workflows and showcased how various AI applications are being integrated.

Cytomorphology

A two-stage convolutional neural network (CNN) system is used to classify multiple cell types of varying lineages and maturation states with an average accuracy of 98.36%. Misclassifications occurred between cell types with similar morphological characteristics. The system also detects dysmorphic features of peripheral blood cells, aiding in identification of malignant cells and hematologic neoplasms. However, bone marrow smears and histologic approaches to trephine biopsies, which have significantly higher cell density, are more difficult to evaluate. Nevertheless, new models have been developed with high precision and recall values for diagnostically relevant classes.

Molecular Genetics

A new web-based multi-classifier, which uses Amazon Sagemaker and LightGBM, has been developed and validated to evaluate and interpret WGS and WTS data. By extracting single nucleotide variants, structural variants, and copy number alterations from WGS data, in addition to assessing gene fusions and gene expression using WTS, a patient-specific feature vector for disease classification is created and used as a model input. Relevant features that contribute the most to the final predicted diagnosis are then visualized and the sample is projected onto a map of more than 4500 samples from different entities.



Immunophenotyping

We developed an AI model that uses flow cytometry data to classify B-NHL. By leveraging XGBoost and raw cytometric data, the model was able to surpass human-expert diagnosis through careful feature engineering. The process involves multiple steps, including data transformation, scaling, and cluster formation, as well as light chain restriction definition. AI proved to be a viable solution for flow cytometric data analysis, resulting in a 12.5% increase in cases accessible for analysis. Our xAI strategy enabled us to track feature weight and actual values down to the events in the dot plots. Similar models are under development for other entities and to reliably identify small populations of pathogenic cells for MRD assessment.

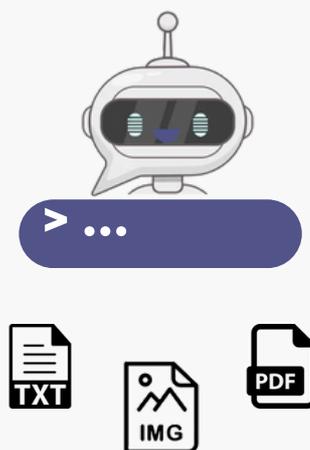
Chromosome Banding Analysis

For the past three years, a sophisticated CNN has been developed for automated karyotyping. The model takes into consideration various characteristics such as chromosome shape and size, centromere location, and banding profile. Aligning and orienting chromosomes vertically as a preprocessing step further improved the classification accuracy. Misclassification usually involved chromosomes very similar in size, shape, and appearance, which are also challenging to differentiate for human professionals. The model is highly effective in identifying normal karyotypes and numerical abnormalities, reducing processing time significantly, and can even identify common structural aberrations.

AI CHATBOT FOR MEDICINE - BENEFITS, LIMITS, AND RISKS

? What is the AI chatbot technology?

A chatbot generally consists of two main components: the system in the background that drives the conversation and an interface for making requests. To use a chatbot, one starts a "session" by entering a query in plain natural language. These are usually questions, but can also be a directive to perform a specific task (e.g. "Please read and summarize the following manuscript") or files (e.g. images, PDFs, etc.) can be uploaded and linked to a request (e.g. an image can be uploaded and asked what can be seen on it). The AI system in the background then tries to give as concrete an answer as possible and if there is a clear answer to the question posed, then the answer from the system is correct in most cases. Even though new studies have shown that chatbots invent information in at least 3% of the time and as high as 27%; even when summarizing articles and facts. These AI-generated responses which contain false or misleading information presented as facts are called "hallucination" or "delusion". So, as usual, every AI-generated response should be taken with a grain of salt.



How could such an interaction play out..?

A newly diagnosed CML patient is to receive 400mg Imatinib daily. He currently takes: Simvastatin 20mg daily and Triazolam 0.25mg at night. What should I bear in mind when dosing this patient with Imatinib?



When initiating treatment with Imatinib, a tyrosine kinase inhibitor commonly used for Chronic Myeloid Leukemia (CML) ...



Stop. Could you please summarize the important points?



Certainly. When prescribing Imatinib for CML, note potential interactions with Simvastatin and Triazolam, both metabolized by CYP3A4, which Imatinib inhibits.

Dosage and Monitoring Adjustments:

1. Simvastatin: Consider dose reduction or switch to a non-CYP3A4-metabolized statin; monitor for myopathy signs.
2. Triazolam: Evaluate the need and possibly adjust the dose; monitor for excessive sedation.
3. Liver Function: Regular testing is advised.

Adjustments should be made in collaboration with an oncologist and pharmacologist.





**SCIENTIFIC HOT
TOPICS IN 2023**

CANCER DIAGNOSTICS TRENDS IN 2023 - IS IT TIME FOR ANOTHER CHANGE?

Due to the rapid development of technology over the last few decades, healthcare trends have been subject to change every couple of years. The last decade started with clinical trial results centered on the use of checkpoint inhibitors, included the FDA's approval of the first CAR T-cell therapy and ended with discoveries into the mechanism that enables cells to sense and adapt to changes in oxygen abundance, already resulting in exciting new treatments. In addition, more than 10 drugs had been approved for acute leukemia in the late 2010s, whereas there had been very few agents in the previous 25 years. Since then large sequencing efforts have further demystified the genetics of cancer, providing the oncological field with great opportunities for precision prevention and early interception. Technological advances are often the basis for progress in cancer research - but which methods were in vogue in 2023 and how can they potentially be used in leukemia diagnostics?



Liquid Biopsies as Early Detection Systems

While biopsies remain the primary diagnostic tool for cancer, they are invasive and require the removal of tissue samples, often through surgery. However, liquid biopsies have drastically revolutionized the field of clinical oncology, leveraging circulating tumor cells (CTC) and cell products to achieve minimally invasive and comprehensive disease detection, as well as to monitor treatment response and disease progression through molecular alterations. First results in the early diagnosis of several types of cancer are promising.

MLL's Perspective on This Topic

We are currently working closely with the University Hospital of Würzburg to evaluate the use of liquid biopsy in multiple myeloma. So far, we have tested various library preparations to find the most suitable method and the preliminary results are promising, as evidenced by similar copy number profile between the liquid biopsy and bone marrow profiles. However, detecting structural variations has proven to be a more challenging task. Moving forward, we plan to continue our efforts in 2024 and will provide regular updates.



MLL's Perspective on This Topic

If you want to work at the forefront of clinical diagnostics and offer every patient the best possible diagnosis, then there is no way around the use of AI-based algorithms to improve reproducibility, accuracy and efficiency. We always endeavour to implement new and safe innovations as quickly as possible for the benefit of the patient. We have already integrated AI-based methods into our workflows, but this is just the beginning and there is more to come.

Artificial Intelligence to Fight Cancer

The use of AI in medicine has increased in many areas, including analysing medical images, identifying at-risk patients and coding medical records. AI can detect 20% more cases of breast cancer than experienced radiologists, it can help to understand the relationship between the effects of different drugs in relation to the disease context, and it can provide appropriate advice when needed - a new synergy between man and machine is emerging.



Large Language Models & Medical Chatbots

Since the public release of OpenAI's ChatGPT in November 2022, LLMs have gained substantial attention. These models can answer questions, summarize, and translate text at a level that is almost indistinguishable from human capabilities. While LLMs have the potential to democratize medical knowledge and improve access to healthcare, they may also spread misinformation and worsen scientific misconduct due to a lack of transparency and accountability.

MLL's Perspective on This Topic

Our team is always in search of new advancements and innovations to improve our work processes. However, we are diligent in evaluating the practicality and dependability of each method. We have successfully utilized LLMs in several applications and prioritize transparency to allow our scientists and clinicians to understand the decision-making process. LLMs encourage critical thinking and provide significant benefits, especially when processing medical documents and communicating scientific findings.



CANCER DIAGNOSTICS TRENDS IN 2023- IS IT TIME FOR ANOTHER CHANGE?



Long-read Sequencing Technology

Long-read sequencing technology has been named “Method of the year 2023” by Nature Methods and has revealed a surge in the number of confidently identified somatic structural variations (SVs) in repetitive areas. It also exhibits improved sensitivity in detecting SVs less than 1 kbp in size, and is beneficial in the exploration of complex SV patterns by enabling haplotype-resolved genomic sequence assembly. Moreover, the technique is highly promising in cancer epigenomics, as it can provide both genetic and DNA methylome data from the same long reads.

MLL’s Perspective on This Topic

Our team is currently investigating the use of long-read sequencing (LRS) for the diagnosis of hematological neoplasms. This technology has shown significant improvements in identifying SV and single nucleotide variations (SNV) in highly repetitive genomic regions. Diagnosing hemoglobinopathies by the detection of α - and/or β -globin gene variants, can be a challenge with short-read sequencing. To overcome this limitation, we are developing a LRS assay for the HBA gene locus and associated genes to ensure accurate and reliable detection.



MLL’s Perspective on This Topic

Our knowledge base continues to expand as our data pool and genomic information stored in the cloud increase in volume. Analyzing this data not only advances diagnostics, but also provides crucial insights into cancer therapies, offering significant support for modern precision medicine. This year, we introduced single-cell sequencing to our laboratory, leading to various successful research projects. Our single-cell ATAC data was particularly useful for understanding the epigenetic regulation of BCP-ALL *BCR::ABL1* subgroups (p. 31). We aim to set up additional single-cell assays and maintain our focus on analyzing mutations at the single-cell level in the future.

Precision Oncology Goes Down to the Individual Cell

The goal of cancer precision medicine is to personalize treatment according to the unique genomic background of each patient’s tumor. Unfortunately, current prognostic and drug response biomarkers rely heavily on bulk ‘omic’ data. This approach inadequately captures intratumor heterogeneity and muddles signals from normal versus tumor cells, which can affect the accuracy of clinical predictions. Fortunately, single-cell technologies are emerging as a promising solution to these challenges. By providing a highly detailed map of genetic and phenotypic diversity within tumors and their microenvironment, these technologies enable a more precise understanding of subclonal tumor populations.



The Role of Patient Data Digitalization in Improving Healthcare Quality

Digitalizing patient data and results, including diagnoses, therapeutic measures, and treatment reports, is considered a significant factor in improving the quality of care. European countries have implemented this requirement to varying degrees, with smaller nations like Estonia already achieving full digitalization. Larger countries such as England and Sweden have made significant progress and have already digitized multiple processes. In Germany, there is a strong desire to expedite the digitalization of the healthcare system, as AI-based methods necessitate digital and structured data.

MLL’s Perspective on This Topic

From its inception, our laboratory has prioritized the digitization of data and work processes. We have implemented an in-house LIMS system to ensure that all patient information is collected, stored, and easily retrievable at any moment. While most of the data is already structured, there are still some areas where the data is less organized, primarily due to input in free text fields. However, we are continuously working to improve the structure of our databases by primarily processing older data sets. Newly acquired data is already included in a structured form thanks to improved frameworks and the integration of AI-based methods.



TWO CLASSIFICATIONS FOR HEMATOLOGIC NEOPLASMS: THE DEBATE CONTINUES

Background

In the scientific community, it is common practice for established concepts and methods to undergo scrutiny and modification as new breakthroughs are made. This is also true in the field of hematologic oncology, where questions persist until the next discovery provides an answer. In 2022, the WHO released its fifth edition of updated classifications of tumors of hematopoietic and lymphoid tissue. The lymphoma classification held few surprises, but major changes were made to the myeloid neoplasms classification. Soon after, an independent group published the International Consensus Classification (ICC) of Myeloid Neoplasms and Acute Leukemias, which introduced new terminology for MDS/AML. The hematologic oncology field is now grappling with reconciling these conflicting guidelines. Recent articles have been published comparing the two classifications, highlighting the current uncertainty and lack of consensus in the field. This two-track approach is also causing a lack of understanding and a need for discussion among scientists.



Daniel Arber, MD, lead author of the ICC paper and Professor and Chair of Pathology at the University of Chicago

Dr. Arber believes that there are two classifications for a reason. He argues that the WHO classification of myeloid neoplasms contains some inaccuracies, particularly in the classification of MDS. "Most of the differences [in the MDS classification] are entities added in the WHO classification that really are not supported by the literature," Dr. Arber explained. "Then some nomenclature changes, which were arbitrarily made."

Uwe Platzbecker, MD, a co-author of the WHO 2022 and Director of the Clinic and Policlinic for Hematology at the Leipzig University Hospital



We also examined the effect of using two different classifications on patient diagnosis and discovered that in AML patients, the diagnosis was consistent between WHO2022 and ICC2022 in 86% of cases. However, there are notable differences that may have a significant impact on individual patients (see page 25). The differences can be perplexing for patients, as Dr. Platzbecker, noted: "I think patients are also bewildered. We don't speak a common language anymore. Not within our scientific community, nor with our patients. Hematologists make an argument now and create a lot of pressure...to harmonize the classification."



Robert Hasserjian, MD, lead author of the ICC MDS section and Director of the Hematopathology Fellowship Program at Massachusetts General Hospital

Although there have been disagreements between some individuals from the WHO and ICC, it's important to recognize that both classifications have more similarities than differences. The differences highlight the various interpretations of data and suggest that more careful consideration should be given to assessing and defining patient subgroups. Dr. Hasserjian agreed that differences could arise from differing interpretations of data: "Creating a classification is a scientific process. There was very little overlap in authorship, and I think the good news is that both the ICC and WHO had similar changes. It's expected that two trials or two studies that look at a relatively similar group of patients might sometimes yield different results."

Dr. Khoury has explained that the WHO has begun integrating ICC terminology into its classification as a "goodwill" effort to harmonize the two systems. Protecting the WHO's system as the standard for categorizing all cancers in every organ system is paramount, according to Dr. Khoury. The goal is to simplify the system for patients, healthcare providers, and researchers, which is why the international working group, comprising representatives from the WHO and ICC, is developing international guidelines and designing a web-based calculator to support the use of both classifications.

"We're not waiting for the sixth edition. We're doing it now, by the WHO adopting many of the ICC terms as acceptable terminology within the WHO classification. Each entry has a subsection that says, 'acceptable terminology' or 'not recommended terminology.' Basically, what we are doing is creating 90% convergence."

Joseph Khoury, MD, lead author of the WHO paper and Professor of the Department of Pathology and Microbiology at the University of Nebraska Medical Center



ENVIRONMENTALLY SUSTAINABLE CANCER CARE



Reducing Plastic Waste in Laboratories

Although laboratories worldwide produce ~2% of global plastic waste, a large majority of this waste comes from disposable items such as packaging and labware, including pipetting tips, tubes, and plates. Unfortunately, the frequent replacement of labware is necessary to avoid contamination, making it difficult to reduce waste in this area. However, for NGS, there are "one-tube methods" that reduce the amount of plasticware required without sacrificing quality, making it crucial to stay up to date with new developments. Due to the high sample throughput at the MLL and the resulting increased consumption of materials, the deliveries alone result in a significant amount of plastic waste. To combat this issue, we are working with suppliers to minimize packaging waste, and reagent suppliers are also committed to reducing packaging weight and disposal costs through the use of more recyclable plastics and limited dry ice or ice packs. The new NovaSeqX series reagents, for example, support ambient-temperature shipping, eliminating the need for dry ice and reducing waste.



The Environmental Impact of Scientific Conferences

Large scientific conferences with 25,000-30,000 participants, such as the annual ASH meeting, contribute significantly to the carbon footprint generated by the scientific community. It's estimated that these events produce roughly 22,000 metric tons CO₂e, which is equivalent to the annual carbon footprint of 1,000 medium-sized laboratories. To raise awareness of the impact of scientific conferences on the environment, some researchers have taken to cycling to these events. In 2023, Prof Ana Conesa from the Intercept MDS project and her team, which is based in Valencia, Spain, cycled to two conferences in Lyon and Barcelona under the hashtag #Cycling4GreenScience. Although biking may not be practical for most conferences, attending virtual meetings and opting for train travel instead of flying can help reduce the carbon footprint caused by scientific conferences. We therefore endeavour to make our trips as climate-neutral as possible.

GREEN



SCIENCE



The Carbon Footprint of Bioinformatics and Data Analysis in Research

Hematological malignancy diagnosis and research are shifting towards genetic-based classifications with the aid of large-scale studies and comprehensive analyses, such as whole genome sequencing. However, the bioinformatics analysis of these complex datasets requires large-scale computational infrastructures, which have a non-zero carbon footprint, contributing to greenhouse gas emissions. Studies show that the yearly electricity usage of data centers and high-performance computing facilities already surpasses that of smaller countries like Ireland or Denmark (Jones et al. 2018, Nature 561(7722): 163-166). To raise awareness and provide recommendations for greener bioinformatics, the University of Cambridge developed the "Green Algorithm." At MLL, we switched to cloud computing for computationally intensive processes in 2017. As a result, not only do the analyses of NGS/WGS data run in the cloud, but also most of the AI-based models to support routine diagnostics. AWS provides a customer carbon footprint tool that offers a detailed breakdown of the carbon emission of individual services. For example, from January to August 2023, we saved 42 MTCO₂e through cloud computing and AWS renewable energy purchases and only produced 0.132 MTCO₂e. Despite the progress, there is still a long way to go, with AWS aiming to achieve net-zero carbon emission by 2040.

0.132 MTCO₂e

Estimated AWS emissions

42.299 MTCO₂e

Emissions saved on AWS

0 MTCO₂e

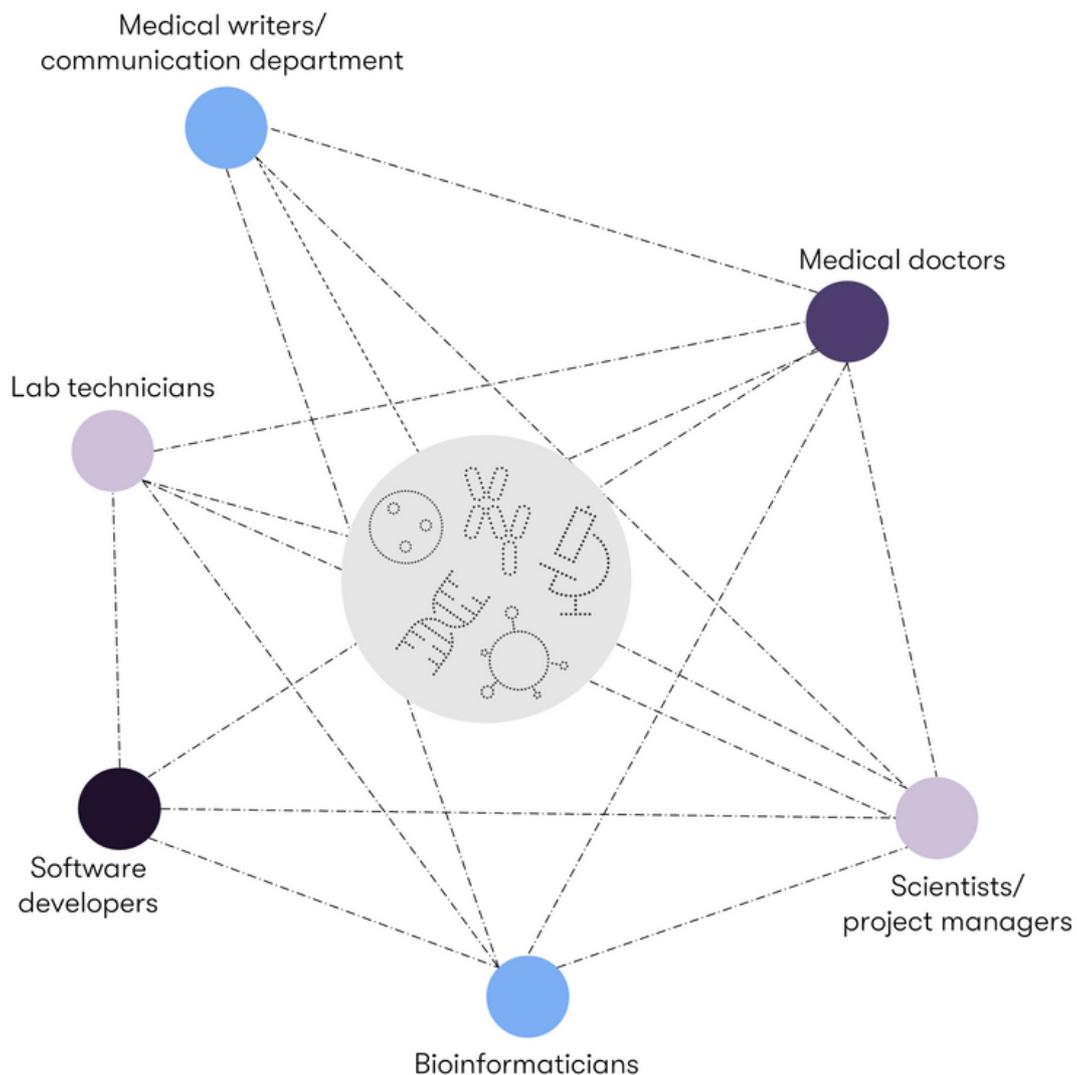
In 2024 our computations on AWS will be powered by 100% renewable energy



M L L R E S E A R C H &
M A N U S C R I P T
S P E E D D A T I N G

RESEARCH AT THE MLL - EVERYONE IS A PART OF THE WHOLE

It has long been recognized that successful medical innovations require a team of experts from multiple disciplines. However, most of the times the different analyses and evaluations are performed in parallel and are never fully integrated. Being aware of this fact and bringing together experts from different fields under one roof, we try to work as closely as possible to achieve the best possible outcome for our patients and to advance our research in a cohesive manner.



„The constant exchange between our various departments and the close cooperation between the individual specialists makes comprehensive and high-quality research work possible.“

Dr. Manja Meggendorfer

MANUSCRIPT SPEED DATING - GET THE GIST OF A MANUSCRIPT IN LESS THAN 3 MINUTES

Risk Prediction in MDS: Independent Validation of the IPSS-M—ready for Routine?

Introduction

MDS are clonal disorders of hematopoietic cells characterized by peripheral cytopenias, morphologic dysplasia, ineffective hematopoiesis and risk of leukemic transformation. Biology and clinical outcome within MDS are extremely heterogeneous, making individual risk prediction key for management and therapy decision.

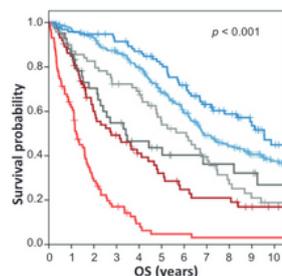
Patients & Methods

626 de novo MDS patients referred to MLL between 09/2005 and 01/2020 with a median follow-up of 9.5 years were genetically characterized by PCR-free WGS and routine work-up to assign patients to the six IPSS-M risk categories.

Results

- The risk distribution for the 626 showed a skewing towards low-risk categories (Figure), comparable with the original IPSS-M publication
- Clear prognostic separation for overall survival, leukemia free survival and leukemic transformation according to IPSS-M categories
- Compared to IPSS-R 111 patients (25%) were up- and 87 patients (19%) down-staged according to IPSS-M
- Analysis of patients up- or down-staged more than one level suggests that the IPSS-M category better reflects the individual survival
- A high *TP53* VAF (>55%) can be used as surrogate for *TP53* multi-hit, whereas the correct assessment of *KMT2A*-PTD status is essential

IPSS-M categories	
Very Low (VL)	15%
Low (L)	41%
Moderate Low (ML)	11%
Moderate High (MH)	7%
High (H)	12%
Very High (VH)	14%



Conclusion

We independently confirm the increased predictive power of the IPSS-M in MDS as compared to IPSS-R.

Risk Assessment According to IPSS-M is Superior to AML ELN Risk Classification in MDS/AML Overlap Patients Defined by ICC

Introduction

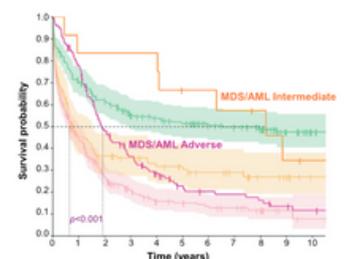
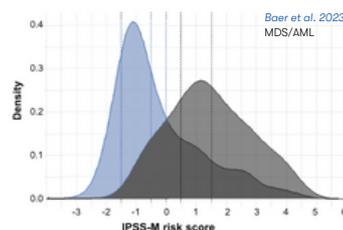
MDS/AML overlap was recently introduced by the ICC as a new entity for myeloid diseases with 10-19% blasts in the absence of AML-defining recurrent genetic abnormalities. The entity was introduced primarily to give these patients a better chance of being enrolled in clinical trials for either MDS or AML. However, there are no guidelines for risk stratification of MDS/AML patients.

Patients & Methods

137 patients with MDS/AML according to ICC were analyzed by WGS & WTS. Bona fide MDS and AML patients were used for comparison.

Results

- MDS/AML subgroups (MDS/AML *TP53*mut, MDS/AML with myelodysplasia-related (MR) gene mutations or MR cytogenetic abnormalities, MDS/AML not otherwise specified) showed a significant differences in OS
- IPSS-M risk categories showed a clear skewing towards high-risk groups for the MDS/AML cohort
- The OS of the respective risk groups was well comparable to the MDS cohort
- No MDS/AML patient fulfilled criteria for the favorable ELN risk group per definition; 9% (12/137) were classified intermediate risk, and the vast majority (91%, 125/137) as adverse risk
- The OS of ELN risk groups for MDS/AML patients substantially differed from the AML control cohort



Conclusion

MDS/AML patients are best stratified by MDS-based risk assessment according to IPSS-M.



MANUSCRIPT SPEED DATING - GET THE GIST OF A MANUSCRIPT IN LESS THAN 3 MINUTES

AML Classification in the Year 2023: How to Avoid a Babylonian Confusion of Languages

Introduction

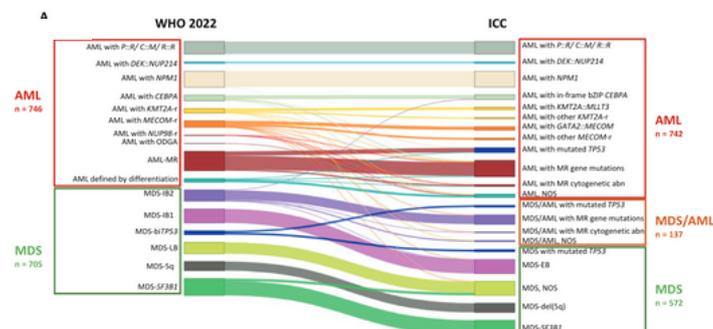
In parallel to the WHO 2022 guidelines, an alternative ICC has been proposed. With regard to AML classification, the two guidelines differ in various respects; for example, the ICC assigns patients with 10-19% blasts and without defining genetic abnormalities to the new MDS/AML category. In addition, the ICC includes the category AML with *TP53* and there are also differences in the *CEBPA* mutated AMLs, to name a few. The changes described by ICC are also mainly reflected within ELN 2022 recommendations and the impact of the new classifications on AML diagnoses and ELN-based risk classification has not been evaluated to far.

Patients & Methods

717 MDS and 734 AML referred to MLL between 09/2005 and 01/2020 were genetically characterized by whole genome and transcriptome sequencing.

Results

- The classification of the disease subtype differed in 14% of AML cases between ICC and WHO 2022
- Overall survival differed significantly between AML entities classified according to WHO 2017, WHO 2022, and ICC
- In relation to the ELN 2022 guidelines, there was a decrease in the number of patients in the favorable and intermediate risk groups and an increase in the adverse risk group



Conclusion

WHO 2022 and ICC classification were concordant in 86% AML cases. The differences are due either to biological aspects/subgroup definitions or to divergent inclusion criteria of the corresponding subgroups.

MDS Subclassification - Do We Still Have to Count Blasts?

Introduction

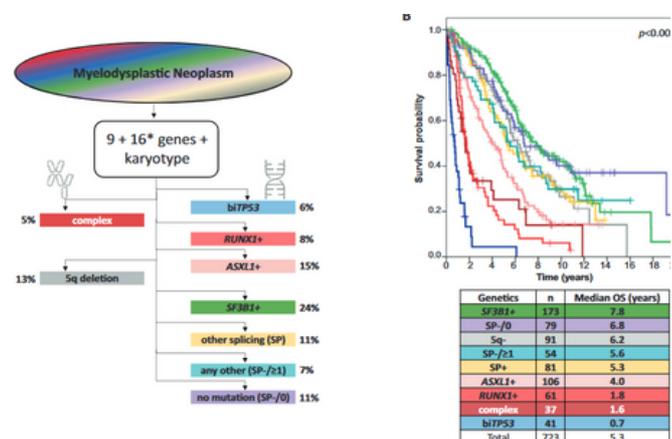
Since their definition, MDS classification relied on cytomorphology determining dysplasia, cytopenia, and blast count. WHO 2022 and the ICC use different blast count cutoffs to define myeloid diseases, and although both guidelines address the importance of genetic abnormalities, only a handful of genetically defined entities are included.

Patients & Methods

735 de novo MDS samples were genetically characterized by whole genome sequencing.

Results

- Number of mutations positively correlated with blast counts and negatively with OS
- *SF3B1* mutations and *del(5q)* were associated with low blast count <5%, mutations in *ASXL1*, *RUNX1*, *SRSF2*, *U2AF1*, *ZRSR2*, and *biTP53* were associated with increased blast count ≥5%
- 19% of cases remained unclassified by genetics only because neither *del(5q)*, complex karyotype, *biTP53* nor a mutation in *SF3B1*, *SRSF2*, *U2AF1*, *ZRSR2*, *RUNX1* or *ASXL1* were detected
- Nine genetically defined non-overlapping hierarchical subgroups could be identified (Figure)



Conclusion

Based on karyotype and mutation status of 9 genes only, we demonstrated that MDS can be separated into distinct subgroups, reflecting biology better than blast counts.



MANUSCRIPT SPEED DATING - GET THE GIST OF A MANUSCRIPT IN LESS THAN 3 MINUTES

The Clinical and Genomic Landscape of Patients with *DDX41* Variants Identified During Diagnostic Sequencing

Introduction

Deleterious germ line variants in the DEAD-box helicase 41 (*DDX41*) gene are the most frequent monogenic cause of germ line predisposition to hematologic malignancy, being present in 3% of patients with MDS or AML. Variants in *DDX41* present challenges to the diagnostic laboratory in terms of variant curation, given their incomplete penetrance, multiple founder mutations, and relatively poorly understood biology.

Patients & Methods

5,737 patients were analyzed by NGS for *DDX41* as well as genes recurrently somatically mutated in myeloid diseases.

Results

- 152 different *DDX41* variants were detected in 5.1% of patients over the entire cohort, with variants detected in 137/2,157 patients with AML, 145/2,865 patients with MDS, and 12/715 patients with classical MPN
- *DDX41* variants were observed in patients in 4 main patterns:
 - multiple *DDX41* variants; min 1 variant >35-40% VAF and another variant detected at <35-40%
 - 2 *DDX41* variants detected at <35-40% VAF
 - only putative germ line variant(s) without putative somatic *DDX41* variant
 - only 1 putative somatic *DDX41* variant without a putative germ line variant
- The most frequently co-mutated genes in AML and MDS were *ASXL1*, *SRSF2*, *DNMT3A*, *NPM1*, *SF3B1*, and *TET2*
- A general male predominance for *DDX41* variants was observed in the cohort

Conclusion

We have further expanded the landscape of genomic variation in *DDX41* and our proposed curation framework modifications may be of use to diagnostic laboratories to provide optimal advice to clinical teams managing these patients and families.

Novel Causative Variants of VEXAS in *UBA1* Detected Through Whole Genome Transcriptome Sequencing

Introduction

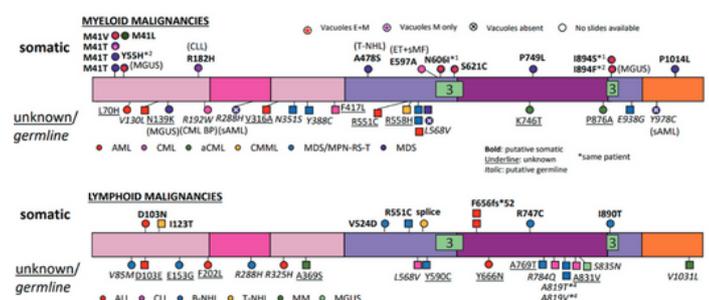
In 2020 the somatic loss of function variants in *UBA1* in the hematopoietic stem and progenitor cell compartment was described as a cause of a severe hematoinflammatory disease termed VEXAS (Vacuoles, E1 enzyme, X-linked, Autoinflammatory, Somatic) Syndrome. Initially, causative variants were only identified in *UBA1* exon 3, but subsequent reports have already expanded the variant spectrum, although it is far from complete.

Patients & Methods

4,168 patients were genetically characterized by whole genome and transcriptome sequencing.

Results

- The cohort comprised 5 patients with *UBA1* M41 variants, 16 patients with putative somatic *UBA1* variants, 20 patients with putative germline *UBA1* variants and 16 patients with variants of unknown origin
- A clear sex bias could be detected in myeloid malignancies, as potential somatic variants occurred only in males
- Immunodysregulatory symptoms are common among patients with myeloid malignancies carrying *UBA1* non-M41 variants



Conclusion

WGTS data analyses identified several potentially clinically relevant *UBA1* non-M41 variants that contribute to a more detailed and exhaustive description of the landscape of *UBA1* variants in hematologic malignancies.



MANUSCRIPT SPEED DATING - GET THE GIST OF A MANUSCRIPT IN LESS THAN 3 MINUTES

Specific Subtype Distribution with Impact on Prognosis of *TP53* Single hit and Double Hit Events in AML and MDS

Introduction

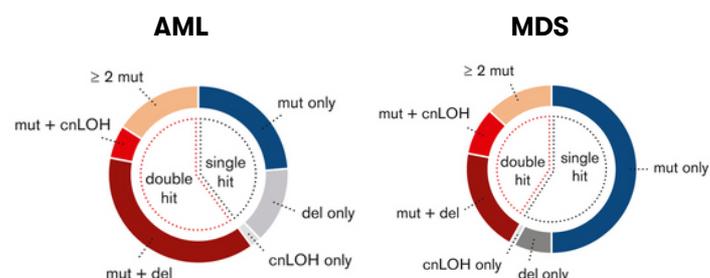
TP53 is the most frequently mutated gene in cancer and alterations can be detected in ~5% to 15% of de novo AML and MDS patients. *TP53* mutations are generally associated with advanced stages of the disease, a complex karyotype, resistance to conventional (chemo-)therapies, and dismal prognosis.

Patients & Methods

772 AML and 747 MDS patients were genetically characterized by WGS and cases were categorized into (1) one *TP53*mut without accompanying del or cnLOH, (2) *TP53* del-only, (3) cnLOH-only, (4) *TP53*mut and *TP53*del, (5) *TP53*mut and cnLOH, and (6) ≥ 2 *TP53*mut.

Results

- At least 1 *TP53*alt was detected in 11% of AML and in 13% of MDS patients
- The presence of a single *TP53*mut already influences OS, but a *TP53*dh worsens OS dramatically in both patients with AML and MDS
- The presence of *TP53*alt and influence on prognosis differs markedly within the analyzed AML and MDS subtypes, some depicting low frequencies and/or only minor or no prognostic relevance
- A significant negative impact of *TP53*alt on OS was found for all subgroups, except for the MDS 5q-subgroup



Conclusion

*TP53*alt play an important role in prognosis and risk stratification of patients with AML and MDS.

Interplay of *TP53* Allelic State, Blast Count and Complex Karyotype on Survival of Patients with AML and MDS

Introduction

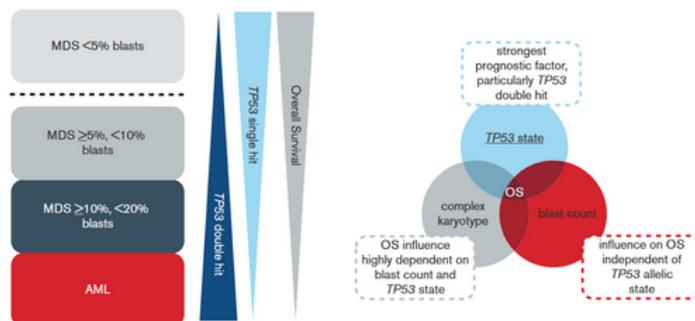
The role of the %blast has been controversially discussed regarding its impact on survival in MDS: *TP53* double hit was more frequently detected in cases with increased blasts, but it was also shown that the prognosis of MDS (with increased blasts) with *TP53* double hit does not appear to be dependent on the blast percentage.

Patients & Methods

772 AML and 747 MDS patients were genetically characterized by WGS and MDS cases and subdivided into cases with <5% blasts (n = 419), $\geq 5\%$ but <10% blasts (n = 175), and $\geq 10\%$ but <20% blasts (n = 153).

Results

- TP53*alt frequencies are comparable between MDS subgroups with different blast counts
- MDS cases with <5% blasts clearly differed from the other subgroups by showing predominantly *TP53* single-hit event
- A complex karyotype was found to be significantly associated with the presence of a *TP53* double hit
- MDS with <5% blasts showed the best outcome, followed by MDS with $\geq 5\%$ but <10% blasts, MDS with $\geq 10\%$, <20% blasts, and AML
- TP53* double hit was associated with dismal outcome in all subgroups



Conclusion

MDS with biallelic *TP53*alt and $\geq 5\%$ blasts should be regarded AML-equivalent regardless of the diagnosis of MDS or AML



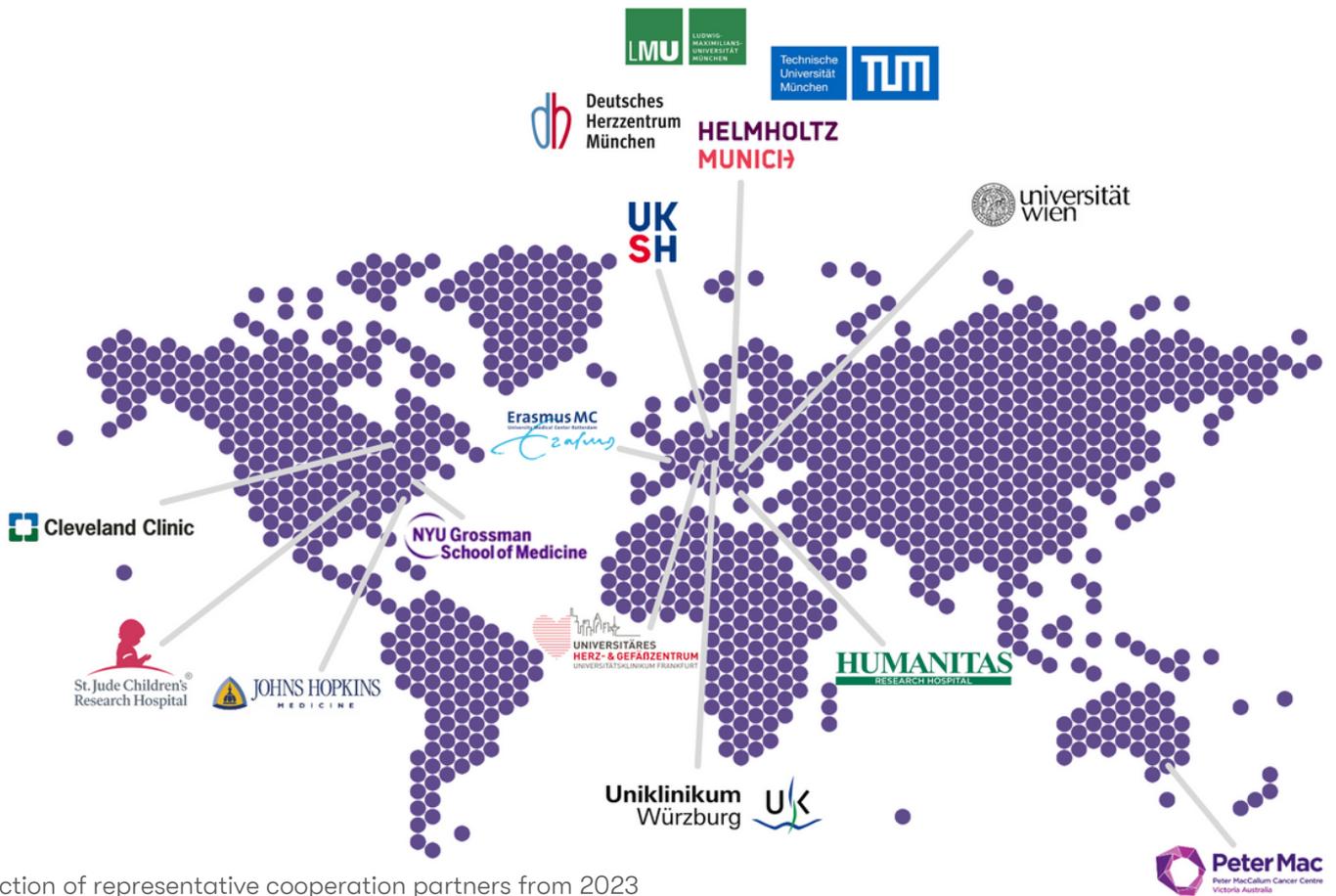


M L L ' S
C O L L A B O R A T I V E
P R O J E C T S

POOLING RESOURCES & JOINING FORCES FOR THE BENEFIT OF THE PATIENT

MLL's Contribution to National and International Research Projects

We have again participated in many national and international collaborative projects this year and made our broad knowledge in the field of hematologic neoplasms, as well as our extensive biobank and data collection, available to other research groups. Our team is always thrilled to take part in these fascinating research projects and contribute to basic leukemia and lymphoma research. Given the current emphasis on multi-omics and data integration, the work often exceeds the capacity of a single institution, requiring large-scale projects to pool resources and generate clinically significant knowledge for the benefit of patients. The following pages highlight a few collaborative studies in more detail.



Selection of representative cooperation partners from 2023

8,613

Within 8 months, 8,613 samples were sequenced for a collaboration project in addition to the daily routine.

38

In 2023, we contributed to 38 collaborative scientific manuscripts with our expertise and data.

16

In 2023, we started or continued 16 collaborative research projects.

GENE EXPRESSION BASED CLASSIFICATION OF BCP-ALL CASES

The Gene Expression Classifier ALLCatchR Identifies B-cell Precursor ALL Subtypes

Introduction

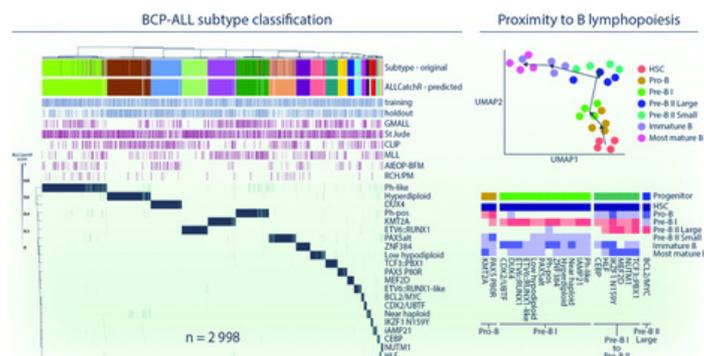
Effective risk stratification and target-specific treatments have led to improved outcomes in both pediatric and adult patients with B-cell precursor acute lymphoblastic leukemia (BCP-ALL). With RNA-Seq, all subtypes of BCP-ALL can be identified using a single method, which sets a new standard for diagnosis. However, to make RNA-Seq a routine clinical diagnostic, unified analysis methods are needed. Hence, this collaborative project aimed to develop a gene expression-based classifier that could accurately allocate BCP-ALL subtypes and enable RNA-Seq to be used for routine BCP-ALL diagnostics.

Patients & Methods

WGS and WTS analysis of 2,998 BCP-ALL samples (64.5% pediatric, 35.5% adult) from 6 independent datasets, including 266 samples from the MLL, which were used to validate the classifier.

Results

- Reliable prediction of 21 gene expression-defined molecular subtypes, samples blast counts, patient's sex, and immunophenotype
- Definition of subtype-specific cutoffs based on the comparison of scores from samples belonging to the corresponding subtype and all remaining samples of the cohort
- Generation of a RNA-Seq reference of human B-lymphopoiesis



Conclusion

High-confidence predictions were achieved in 83.7% of samples with 98.9% accuracy. Only 1.2% of samples remained unclassified.

Developmental Trajectories and Cooperating Genomic Events Define Molecular Subtypes of BCR::ABL1-positive ALL

Introduction

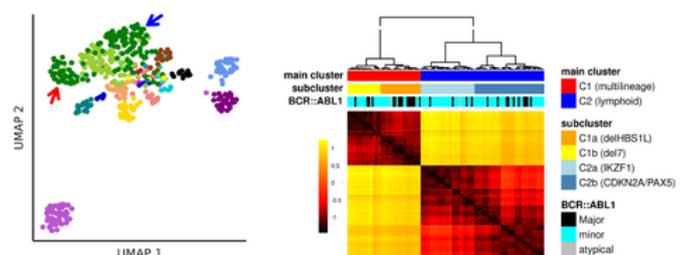
BCR::ABL1-positive ALL is considered a high-risk disease subtype in both children and adults. Treatment typically involves tyrosine-kinase inhibitors combined with dose reduced chemotherapy regimens. However, dependency on lineage-specific targets can lead to lineage infidelity and resistance mechanisms, especially for lineage-restricted immunotherapies. BCR::ABL1-positive ALL show either a lymphoid or stem cell origin and are classified as 'lymphoid-only' or 'multilineage' accordingly.

Patients & Methods

WGS and WTS analysis of 327 BCR::ABL1-positive ALL patient samples from four cohorts, including 61 samples from the MLL.

Results

- Unsupervised analysis of subtype specific gene expression revealed two distinct main clusters (C1, C2) and two subclusters (C1a, C1b, C2a, C2b) within BCR::ABL1-positive ALL
- FISH analysis on FACS-sorted hematopoietic compartments identified BCR::ABL1 in 28%-99% of myeloid cells in all C1 samples, whereas C2 samples harbored BCR::ABL1 exclusively in lymphoid precursors or mature B cells
- C1a was characterized by focal deletions in HBS1L and C1b was enriched for monosomy 7
- C2a was enriched for homozygous IKZF1 deletions and C2b for CDKN2A/B and PAX5 deletions



Conclusion

The work provides a novel framework to subclassify BCR::ABL1-positive ALL into distinct biological and clinically relevant entities.



MOLECULAR PATTERNS IN MYELOID NEOPLASIA

The E592K Variant of *SF3B1* Creates Unique RNA Missplicing and Associates with High-risk MDS without Ring Sideroblasts

Introduction

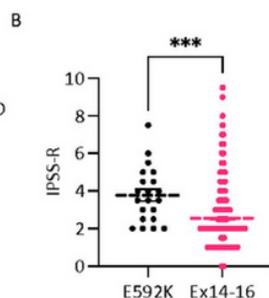
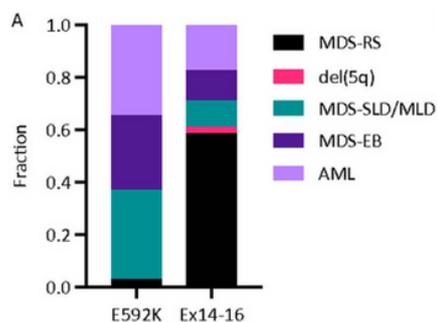
SF3B1 is the most frequently mutated spliceosome gene in MDS, with a frequency of over 30%. In MDS treatment, *SF3B1* mutations are used to determine eligibility for luspatercept and are targeted by investigational therapies. However, it is still unclear how distinct *SF3B1* mutation hotspots impact disease features and the RNA missplicing events that cause them.

Patients & Methods

Mutational profiling and gene expression analysis of 2,288 patients with *SF3B1*-mutant MDS or AML. The cohort included 2 E592K, 10 E622D, 6 K666N, 9 K666R, 12 K700E, and 12 WT samples from the MLL 5K dataset.

Results

- Consistent with previous reports K666N was enriched in higher-risk disease types, whereas K666R and E622D were decreased in higher-risk disease
- The pattern of missplicing by K666N was found to be different from that of other MDS/AML-associated hotspots and also K666R, demonstrating that although these variants affect the same starting amino acid, they are not functionally equivalent
- E592K cases also showed an enrichment in higher-risk classifications, a distinct splicing pattern, a lack of ringsideroblasts and a divergent co-mutation profile, with *ASXL1* being the most frequent partner (83%)



Conclusion

These findings regarding cases with E592K variants add to our understanding of the functional diversity of spliceosome mutations

Molecular Patterns Identify Distinct Subclasses of Myeloid Neoplasia

Introduction

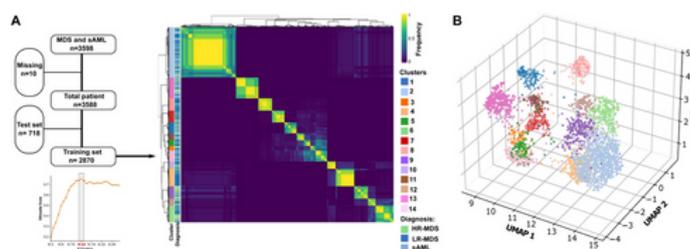
Although morphological and clinical features have been the traditional basis for diagnosis and classification of MDS and AML, the inclusion of molecular data can provide insight into the functional pathobiology of these conditions. This study tested the hypothesis that related molecular patterns can be analyzed in an unbiased/unsupervised fashion to characterize molecularly defined configurations of MDS/sAML.

Patients & Methods

Molecular profiling by targeted NGS of 3,588 MDS, CMML and sAML patients, including 1,275 samples from the MLL.

Results

- Unsupervised analysis identified 14 molecular clusters (MC1-MC14) according to distinct genomic features, like karyotype and the mutation status of *SF3B1*, *TET2*, *DNMT3A*, *EZH2*, *ASXL1*, *JAK2* and *RAS* pathway genes
- LR-MDS patients comprised most of MC8 (78%), MC10 (78%), and MC5 (72%), whereas HR-MDS and sAML cases comprised more than 30% of MC3, MC6, MC9, and MC12
- Very high risk and high risk IPSS-M groups were mainly enriched within MC1, MC9, MC11, MC12, and MC13
- By grouping MCs according to overall survival, 5 risk categories could be distinguished



Conclusion

ML analytics were used to classify patients with MDS and sAML based on functional and pathological similarities, rather than clinical features. This classification can help identify patients who may respond to targeted treatments and aid in testing the efficacy of drugs.



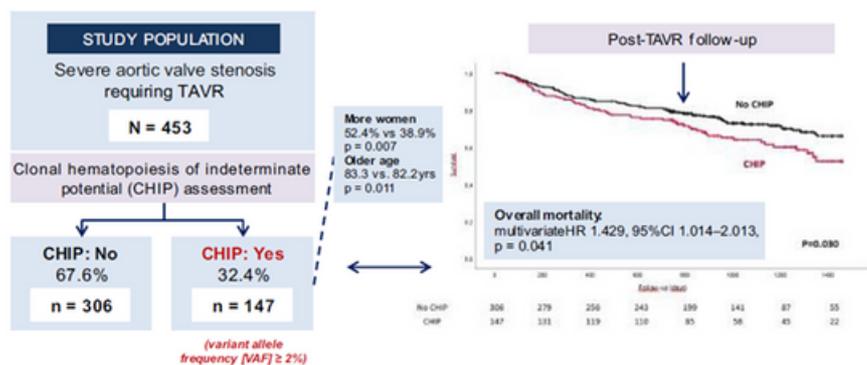
CLONAL HAEMATOPOIESIS OF INDETERMINATE POTENTIAL



In 2023 a number of projects were driven by the question if CHIP - clonal hematopoiesis of indeterminate potential - influences the outcome of other diseases than blood cancer. We were happy to contribute to a large variety of diagnostic fields, including cardiology, neuropathology, but also COVID-19 or lung fibrosis. We have a great pleasure in combining the fields of knowledge and generating new research results. We would therefore like to thank our cooperation partners for the exciting insight into other medical fields and are delighted to be able to contribute with our knowledge of CHIP and our sequencing experience. Many thanks for the co-authorships!

Long-Term Risk Associated With Clonal Hematopoiesis in Patients With Severe Aortic Valve Stenosis Undergoing TAVR

This study by Silvia Mas-Peiro examines the impact of *DNMT3A* and *TET2*-CHIP-driver mutations on the long-term outcomes of patients with severe aortic valve stenosis (AVS) who underwent Transcatheter Aortic Valve Replacement (TAVR). The study found *DNMT3A*-/*TET2*-CHIP-driver mutations in 32.4% of the patients, with a higher prevalence in women and older individuals. These mutations were significantly associated with a higher all-cause mortality up to 4



years after TAVR, both in univariate and multivariate analyses. This association was even stronger in patients who never smoked, indicating the mutations' impact was independent of DNA methylation changes caused by smoking. This finding suggests that the presence of these mutations should be considered in the clinical management and prognosis of patients undergoing TAVR for severe aortic stenosis.

Increased Prevalence of CHIP in Hospitalized Patients With COVID-19

This study by Judith Schenz et al. focuses on understanding the impact of Clonal Hematopoiesis of Indeterminate Potential (CHIP) on COVID-19 pathophysiology. The study found a significantly higher prevalence of CHIP-driver mutations among hospitalized COVID-19 patients (37.8%) compared to the expected prevalence based on the median age of the cohort (17%). CHIP was associated with an increased risk of hospitalization during the course of COVID-19 but did not independently affect the outcome among the hospitalized group. Younger patients (45-65 years) with CHIP exhibited persistent lymphopenia, while older patients (>65 years) developed neutrophilia over time. At admission, CHIP-positive and CHIP-negative patients had similar lung injury and infection-related parameters. However, significant differences were observed in cardiac biomarkers (NT-proBNP and Troponin T), which were higher in CHIP-

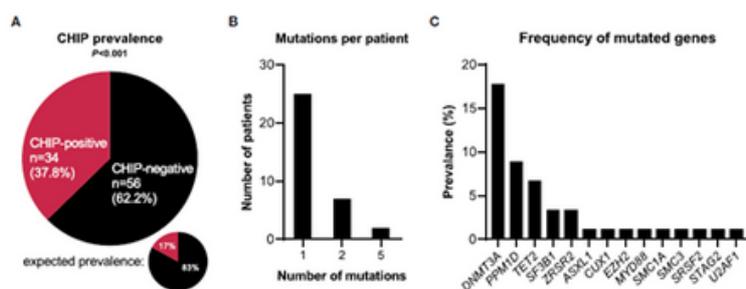


FIGURE 1
CHIP-associated mutations in hospitalized patients with COVID-19. (A) Prevalence of clonal hematopoiesis in hospitalized patients with COVID-19. CHIP-positive = patients carrying at least one CHIP-driver mutation with a VAF ≥ 0.01. Expected prevalence according to Watson et al. (4). Observed and expected prevalence were compared using binomial test. Statistically significant results ($P \leq 0.05$) are highlighted by bold print. (B) Number of individual mutations per CHIP-positive patient. (C) Prevalence broken down by affected genes.

positiv patients, suggesting a possible impact on cardiac function. No significant differences were found between CHIP-positive and CHIP-negative patients in terms of lymphocytic populations, monocyte subpopulations, or cytokine levels at admission. However, at discharge, CHIP-positive patients showed higher neutrophil counts and lower lymphocyte counts, indicating an altered immune response in the presence of CHIP.



CLONAL HAEMATOPOIESIS OF INDETERMINATE POTENTIAL

Prevalence and Therapeutic Implications of Clonal Hematopoiesis of Indeterminate Potential in Young Patients With Stroke

This study by Ernst Mayerhofer et al. performs a genetic screening for Clonal Hematopoiesis of Indeterminate Potential (CHIP) in young stroke patients, exploring its prevalence and potential as a novel risk factor for stroke. The CHIP prevalence in the stroke cohort was 21%, and with that significantly higher than in the general population. Patients with CHIP were generally older and had a higher burden of atherosclerosis. Patients with CHIP showed higher carotid intima-media thickness and a higher incidence of complex aortic plaques, indicating an association with atherosclerosis. Interestingly, out of the patients with CHIP, 11 were identified for further hematologic evaluation,

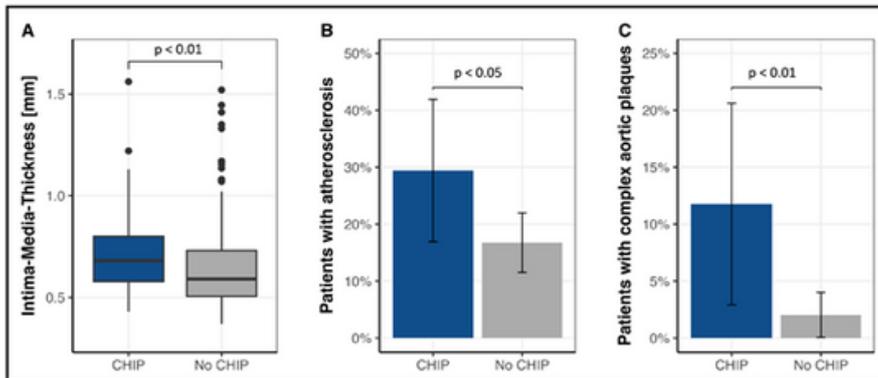


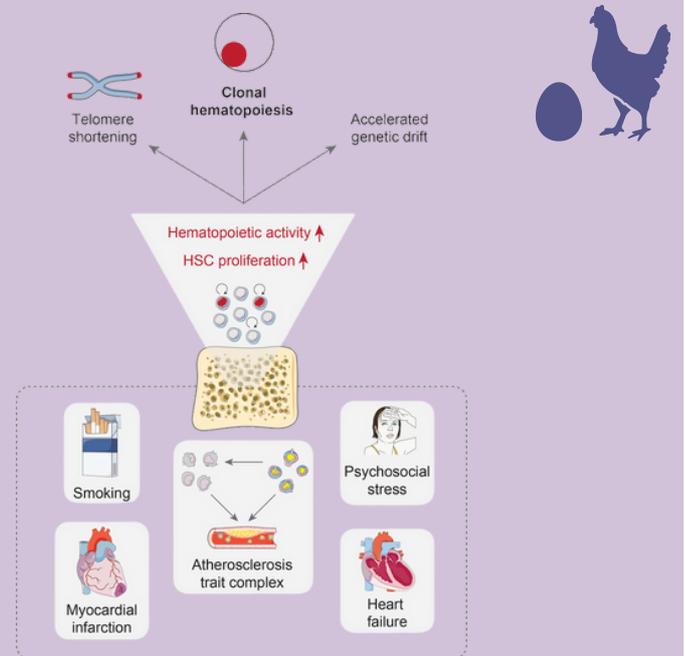
Figure 3. Association of clonal hematopoiesis of indeterminate potential (CHIP) with atherosclerosis. Carotid intima-media thickness was significantly higher (+0.09 mm) in patients with CHIP (A). Atherosclerosis (B) including complex aortic plaques (C) was significantly more prevalent in patients with CHIP.

leading to the diagnosis of myeloproliferative neoplasms in two cases. The study underscores the diagnostic and therapeutic value of genetic screening in young stroke patients, particularly those with undetermined etiology. CHIP may contribute to stroke risk through atherosclerotic or thromboembolic mechanisms, though the exact causality remains unclear.

The Chicken-or-Egg Question of Clonal Hematopoiesis

The contribution of clonal hematopoiesis to the risk increase of cardiovascular disease is a hot research topic. Interestingly, the question which came first, the chicken or the egg, is not finally answered. While the creation of an inflammatory environment by clonal hematopoiesis promoting the development of atherosclerosis has been well documented in a number of preclinical models, also the other way round cannot be ruled out, namely that the inflammatory environment might also promote hematopoietic stem cell (HSC) proliferation and thus the accumulation of mutations.

The study "Increased Stem Cell Proliferation in Atherosclerosis Accelerates Clonal Hematopoiesis" by Alexander Heyde et al. uses mathematical models to further explore the relationship between clonal hematopoiesis and cardiovascular disease. The paper demonstrates that HSC division rates are notably increased in mice and humans with atherosclerosis. This increased proliferation expedites somatic evolution and the expansion of clones with driver mutations. Increased HSC proliferation not only accelerates driver clone expansion but also speeds up the effects of neutral drift, potentially leading to clonal hematopoiesis emergence.



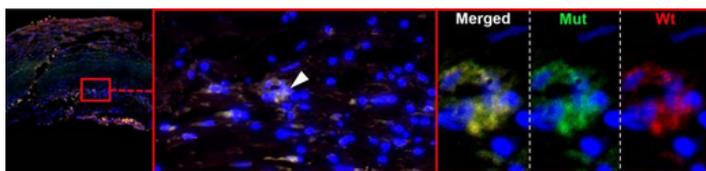
In summary, the research suggests that atherosclerosis can cause clonal hematopoiesis through chronic promotion of HSC proliferation. Combined with the body of evidence on CHIP as a causal factor for the development of atherosclerosis, a vicious circle of inflammation with a mutual building of CHIP and atherosclerosis can be proposed.



CLONAL HAEMATOPOIESIS AND CARDIOVASCULAR DISEASE

Researching CHIP at the Technical University of Munich's German Heart Center

The Technical University's German Heart Center in Munich is responsible for treating nearly 20,000 patients with heart disease every year. Genetic research into coronary heart disease and heart attacks remains a significant focus. Since 2017, Professor Dr. Heribert Schunkert (Director of Cardiology) and Dr. Moritz von Scheidt (Cardiologist and Junior Group Leader for Somatic Mutations) have been leading a research group that specifically studies clonal hematopoiesis of indeterminate potential (CHIP). The group also established one of the most extensive cardiovascular tissue databases in the world, with samples from over 1,100 individuals. This offers exclusive access to human coronary arteries, from those who are healthy to those who are severely diseased. In collaboration with the MLL, more than 250 individuals in the group were found to have CHIP. The CHIP-mutated macrophages were identified and visualized for the first time at the single-cell level in human atherosclerotic plaques (see figure). CHIP mutations are found to lead to a pro-inflammatory phenotype, which suggests that it can contribute to significant local damage in addition to its systemic effects. Currently, the hypothesis is being tested to determine whether CHIP contributes to increased vulnerability of atherosclerotic plaques, which makes them more prone to rupture. This increases the risk of heart attack and sudden cardiac death.



Left: atherosclerotic plaque. Middle: CHIP-affected leukocytes. Right: DNMT3A mutation c.2333G>T in green, DNMT3A wild type in red, cell nuclei in blue.

On top of the basic scientific work, the German Heart Center Munich also screens patients with established coronary heart disease for CHIP in collaboration with the MLL. In 2023, more than 8,500 patients were tested for CHIP, and one-third of them were found to have it. Hematological integration is conducted at one of the first CHIP clinics in Germany, which is headed by Prof. Florian Bassermann (Director of Hematology and Oncology) and Dr. Judith Hecker (Head of CHIP Outpatient Clinic). The prevalence of CHIP in patients with coronary heart disease is significantly higher than in the general population. Patients with CHIP are at a higher risk of complex vascular calcification, myocardial infarction, stroke, and premature death.

Introducing the German CHIP Register: A Comprehensive Resource for Patients and Researchers

The German CHIP Register is a non-profit organization established in October 2023 by the TUM to advance understanding of CHIP, provide support for patients, and explore new therapeutic options. The organization provides a user-friendly overview of CHIP for patients and aims to raise awareness among patients, families, and society. Despite its significant impact on cardiovascular risk, routine CHIP screening is not yet conducted in Germany and guidelines do not provide treatment recommendations. Supported by institutions like the German Center for Cardiovascular Research (DZHK) and the MLL, the German CHIP Register conducts ongoing research into cardiovascular and hematological aspects of the disease and current clinical studies are investigating the benefits of different anti-inflammatory medications for CHIP.

“The German CHIP Register is grateful for the expertise and excellent partnership with the entire MLL team, as we work together to make CHIP more visible and treatable in the future.”



Dr. med. Dr. med. univ. Moritz von Scheidt
Medical specialist for internal medicine and
cardiology

High-Throughput Screening for CHIP register samples

The CHIP study used high-throughput screening with Illumina DNA Prep and IDT enrichment (panel of 11 genes), automated overnight for maximum efficiency. NovaSeq 6000 system with S2 Flow Cell was used for sequencing, allowing up to 530 patients/libraries to be sequenced together. A target coverage of 2,000x was used to ensure high sensitivity. The bioinformatics pipeline automatically triggered the SNV calling (Pisces, DRAGEN) and the subsequent automatic deletion of sequencing artifacts and addition of known variants to our patient database. Unknown variants with a VAF >5% required manual interpretation, whereas those with a VAF <5% were automatically rescheduled for validation. This allowed the analysis of up to 530 patients per run in just a few (3-4) hours, with the results delivered in a structured format.



GENOMNET - WORKING TOGETHER FOR MORE COMPREHENSIVE DIAGNOSTICS



Mission: Advancing Hematological Neoplasm Care

Our mission is to apply the latest discoveries in hematological neoplasms to benefit patients by providing optimal diagnostics and therapy. Achieving this goal requires a collaboration among clinical care and diagnostics experts, patient organizations, and the pharmaceutical industry, all of whom share a keen interest in research.



<https://www.genomnet.de/>



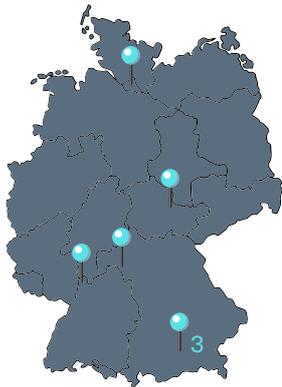
Envisioning the Future of Personalized Medicine

In the near future, the dream of personalized medicine will be realized. Utilizing evidence-based algorithms, cost-effective step-wise diagnostics will be performed, enabling doctors to select the ideal therapy for each patient. Treatment outcomes will be reviewed regularly and adjusted as necessary. The ultimate aim of this coordinated concept is to enhance the chances of recovery.



Goals: Establishing Knowledge-Generating Patient Care through Research

The primary objective is to integrate data from both research projects and patient care to create a knowledge-generating patient care system. This approach ensures that new findings from daily patient care inform and inspire new research projects. Ultimately, the results of these projects are implemented to improve patient care.



Current members of the network:

- MLL Munich Leukemia Laboratory, Munich
- Comprehensive Cancer Center, University Hospital Rechts der Isar TUM, Munich
- LMU München, LMU University Hospital, Medical Clinic III, Munich
- Comprehensive Cancer Center Mainfranken, Würzburg
- University Medical Centre Mannheim, Medical Clinic III, Mannheim
- Comprehensive Cancer Center Central Germany, University Hospital Jena, Medical Clinic II, Jena
- UKE Hamburg, Medical Clinic II, Hamburg
- new interested members are welcome...

Ongoing Projects:

The Unsolved Case

This study seeks to determine if whole genome sequencing (WGS) and whole transcriptome sequencing (WTS) can offer additional information for patients whose standard diagnostic tests have yielded inconclusive results. The research will compare current diagnostic methods with potential future ones, assessing factors such as turnaround time and cost structure. The aim is to identify potential research questions for future studies.

MDS post Multiple Myeloma Therapy

After CAR T-cell therapy for multiple myeloma, some individual patients have been observed to develop myelodysplastic neoplasia. This is also the case with other therapies for multiple myeloma. To identify whether a particular genomic profile exists in these patients, the study aims to characterize myelodysplastic neoplasms using WGS and WTS in patients with a corresponding clinical constellation.

Mechanisms of Resistance to Targeted Therapies in Multiple Myeloma

Patients with multiple myeloma who are scheduled for a targeted therapy, such as BCMA- or GPRC5D-targeted bispecific antibodies or CAR T-cells, or have already undergone this treatment and are non-responsive, are eligible to participate in this study. The objective is to introduce WGS and WTS for patients in similar clinical circumstances in order to determine the effect of this information on treatment selection and clinical outcomes.

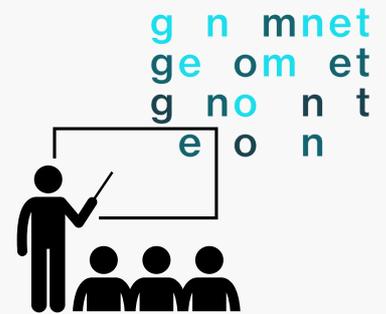
Genetic Alterations and Expression Profiles in iHES Patients with Cardiac Involvement

This study focuses on a comprehensive analysis of genetic alterations and expression profiles of iHES patients with cardiac involvement. WGS and WTS will be utilized to achieve this aim. The goal is to gain a better understanding of relevant pathomechanisms for cardiac involvement in iHES and to identify potential genetic signatures that may have a correlation with the severity of cardiac involvement.

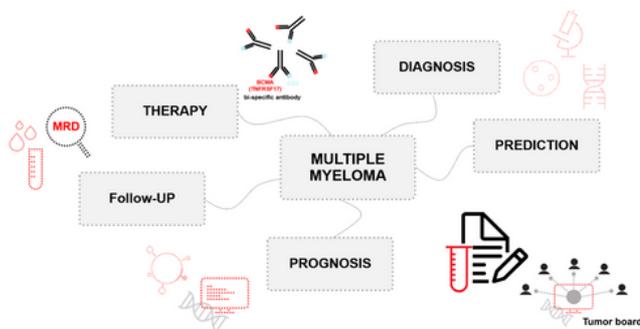
GENOMNET - WORKING TOGETHER FOR MORE COMPREHENSIVE DIAGNOSTICS

Report on the First Training Event of the Hematology Genome Network

The Hematology Genome Network at the MLL kicked off its first training event on June 22, 2023. The event aimed to facilitate knowledge transfer and networking between diagnostics and the clinic. Three experts presented and discussed various topics pertaining to the field. Ms. Marietta Truger from the MLL discussed current diagnostics for multiple myeloma (MM). Dr. Leo Rasche from the Medical Clinic and Polyclinic II in Würzburg provided an overview of current therapeutic options for MM, outlined the progress, and future challenges, and made the connection to diagnostics. Prof. Katharina Götze from the Clinic and Polyclinic for Internal Medicine III, Klinikum Rechts der Isar, discussed the link between MDS and MM and possible causal relationships.



What do we need?



Diagnostics for multiple myeloma

Ms. Truger's presentation focused on MM diagnostics, emphasizing the importance of genetic characterization and risk assessment. The talk discussed the use of Interphase-FISH as the gold standard for genetic analysis and identified high-risk and standard-risk genetic markers. It also explored the role of clonal plasma cells in myeloma initiation and progression and highlighted new diagnostic tools like whole genome and RNA sequencing. The presentation considered measurable residual disease (MRD) as a crucial prognostic factor.

New opportunities and new challenges in myeloma therapy in 2023

The presentation by Dr. Rasche discussed advances and challenges in MM therapy. It covered various drugs and immunotherapies, when to initiate treatment, high-risk MM identification and treatment, insights into second-line therapies, BCMA therapies, and curative strategies. It also analyzed CRAB criteria and the role of molecular data in MM treatment. Double-hit myeloma and its impact on treatment strategies were also discussed.

Wann beginnt man mit der Behandlung?
CRAB? SLiMM CRAB? HR-SMM?

Was ist Hochrisiko MM und sollte man die HR anders behandeln?

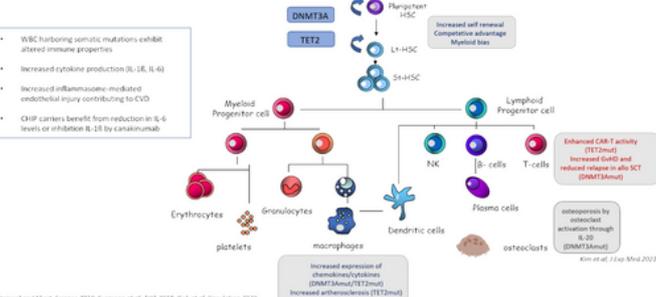
Zweitlinientherapie?

Wann kommen die BCMA Therapien?

In welcher Sequenz setzen wir die neuen Immuntherapien ein?

Was wäre eine (kurative) Strategie beim MM?

CH and Inflammation



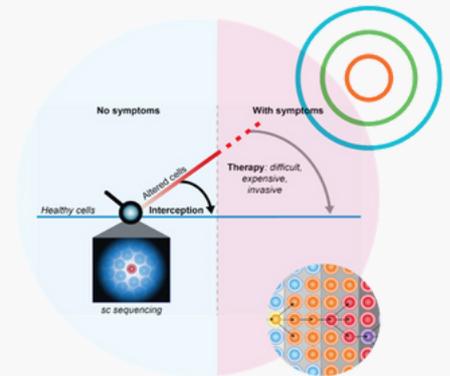
MDS after MM

Katharina Götze's presentation covered the challenges of managing MDS in patients with MM. It discussed the occurrence of secondary MDS in MM, the risk presence in untreated monoclonal gammopathy, and improved survival due to novel therapies. The talk also explored open questions regarding non-treatment related risk factors in MGUS/MM, the potential role of clonal hematopoiesis in MDS/AML development, and the influence of the bone marrow microenvironment in these processes.

EXPLORING CELL-TO-CELL HETEROGENEITY IN MYELOID DISEASES

? What is INTERCEPT-MDS

INTERCEPT-MDS is an Innovative Training Network funded by the European Union's Horizon 2020 research and innovation programme. Born from the LifeTime Initiative, the aim of INTERCEPT-MDS is to train Europe's first experts at the doctorate level in the novel field of disease interception: treating a disease before it fully develops by removing altered cells. INTERCEPT-MDS brings together 10 European public and private institutions in a European network of experts in leukemia, epigenetics and single-cell genomics. Through a multidisciplinary and multisectoral approach, INTERCEPT-MDS studies disease interception in the context of clonal myeloid diseases, using single-cell genomic methods to identify altered cells and test epigenetic perturbations to specifically remove them.



From left to right: Maki Sakuma, Stephani Schmitz, Shubhra Bhattacharya.

As announced in our last research report, this year we hosted various INTERCEPT-MDS students at MLL. As part of their training, students spend several months at other institutes to expand their method portfolio and to access internal data sets. Moreover, these secondments provide the opportunity to also gain insight into the work of companies and the organizational structure of institutes outside of academia.

The experiences shall help the students to thoughtfully develop their career plans. In April, Stephani Schmitz from Erasmus MC in Rotterdam, the Netherlands, joined the MLL research group for eight weeks to mine the MLL 5K dataset and explore the genomic landscape of MDS patients. Stephani's work focuses on exploring how $\text{del}(5q)$ mutated hematopoietic stem and progenitor cells with CSNK1A1 mutations gain a clonal advantage in the bone marrow. Back in Rotterdam, she is now in the process of generating a comprehensive single-cell data set of MDS patients with $\text{del}(5q)$.

The second student was Shubhra Bhattacharya from the Josep Carreras Leukemia Research Institute, Badalona, Spain. As a passionate bioinformatician, her work focuses on mutations in chromatin regulators and the application of computational models to evaluate how these mutations contribute to MDS progression. During her three-month stay at MLL, she worked on the analysis of MLL-internal single-cell data from patients with hematological neoplasms, focusing primarily on the detection of CNVs

from scGEX data. She will now apply her gained knowledge to her own single-cell data set of MDS patients harboring STAG2 mutations, which was generated by MLL.

From August to the end of September, our student Maki Sakuma was a guest at BioBam, Valencia, Spain. The company specialises in the development of bioinformatics tools, so Maki received extensive training in the analysis of single-cell data, especially cell type annotation and trajectory analysis.



The focus of Maki's PhD work is on computationally inferring the expression status of different clones from MLL 5K bulk sequencing data to characterise MDS patients at the single-cell level and identify potential therapeutic targets and diagnostic biomarkers. Hence, her stay in Valencia has given her a solid basis for the comprehensive data analysis and she will be working with our MDS single-cell data in the coming months. As the entire network focuses on the analysis of single-cell data, the students received a specialised bioinformatics course for data analysis from BioBam in September. With the newly acquired knowledge, they can start analysing and evaluating their own data.

Outlook

In 2024, the network will meet for the last time to evaluate the progress of the project and discuss the final steps for the PhD students. To this end, the network will take part in the EHA Congress in Madrid. Moreover, we expect our student, Maki Sakuma, to complete her work on the project at the end of 2024 and we are looking forward to the results.



FUTURE
PROJECTS
& OUTLOOK

SCIENTIFIC ENDEAVORS FOR THE YEAR 2024

Conferences



Attending conferences is a vital part of our research strategy as it allows us to network with other professionals in our field, share our knowledge with a broader audience and engage with other research groups. By taking part in these events, we can remain up-to-date with the latest scientific advancements and gain valuable insights to potentially improve our workflows.

AI-Enabled Data Management

This year we have expanded our use of AI-based models for collecting and analyzing clinical data. Our systems can now automatically retrieve vital information such as dates and causes of death, along with provided therapies. In the coming year, we aim to further improve these AI models while also continuing to work on automating the generation of reports that integrate various test results.

Methylation Analysis



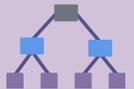
This year, we've experimented with several methods for methylation analysis. In 2024, we aim to optimize the workflow and data evaluation further. This will add an extra layer of information to our research projects and could even have diagnostic applications in the future. As part of this, we're considering integrating our findings with the GEX-based classifier.

WGS in Routine Diagnosis of Patients with Suspected Hematological Neoplasms



The establishment of the genome network for hematological malignancies marks a significant milestone in the integration and utilization of WGS in the diagnosis of leukemia and lymphoma. However, it is crucial to continue pursuing this approach vigorously, disseminating knowledge on its possibilities and benefits to both doctors and patients. This will enable the rapid translation of new findings into knowledge-based care.

Developing a GEX-Based Disease Classifier for Hematologic Neoplasms



Various research and collaboration projects have shown us that the transcriptome of patients with hematologic neoplasms contains clinically relevant information that aids in the diagnosis of various entities. However, the analysis and evaluation of gene expression values is always relative and the success of a potential model strongly depends on what is compared to what. For this reason, we have chosen a hierarchical design for a GEX-based disease classifier that provides broader group classifications if more specific subtypes cannot be identified.

Long-read Sequencing

Apart from establishing an LRS assay to diagnose hemoglobinopathies, we plan to combine it with our GEX-based classifier for a significantly reduced turnaround time. LRS is capable of generating high-resolution transcriptomic data in real-time and at a low cost. Therefore, we aim to incorporate real-time transcriptomic profiling by processing RNA sequencing data from patients with suspected hematological malignancies on-the-fly with our trained GEX-based classifier. In the best case scenario, this can result in a TAT of only 4 hours, which would provide a valuable gain in time and information, especially for the diagnosis of acute leukaemias.

Liquid Biopsy in Multiple Myeloma

Our team is currently in close collaboration with the CCC Mainfranken to assess the effectiveness of liquid biopsy in multiple myeloma cases. We plan to further improve our DNA extraction, library preparation, and data analysis methods to provide clinicians with a comprehensive overview of patient molecular profiles, allowing for less invasive disease monitoring.

Enhancing MRD Detection



Our team is continuously working towards improving our workflows. In the upcoming year, we have planned to incorporate error correction using unique molecular identifiers to enable NGS-based measurable residual disease (MRD) detection. Additionally, we will evaluate TwinStrand Duplex sequencing for DLBCL cases and keep up with the advances in single-cell MRD.

Optimizing Workflows with AI Models to Deliver Quality Results



We utilize ML-based models to ensure our workflows meet the highest quality standards, resulting in faster and more reliable work. This enables every MLL employee to explore new methods, undertake other projects, and improve existing workflows. Despite already achieving great success, with the latest developments in AI, we can further increase our capabilities and support more processes through ML. Our AI team is dedicated to pushing the boundaries and providing MLL employees with the best possible support in their daily practice and to offer patients the best possible diagnosis.

