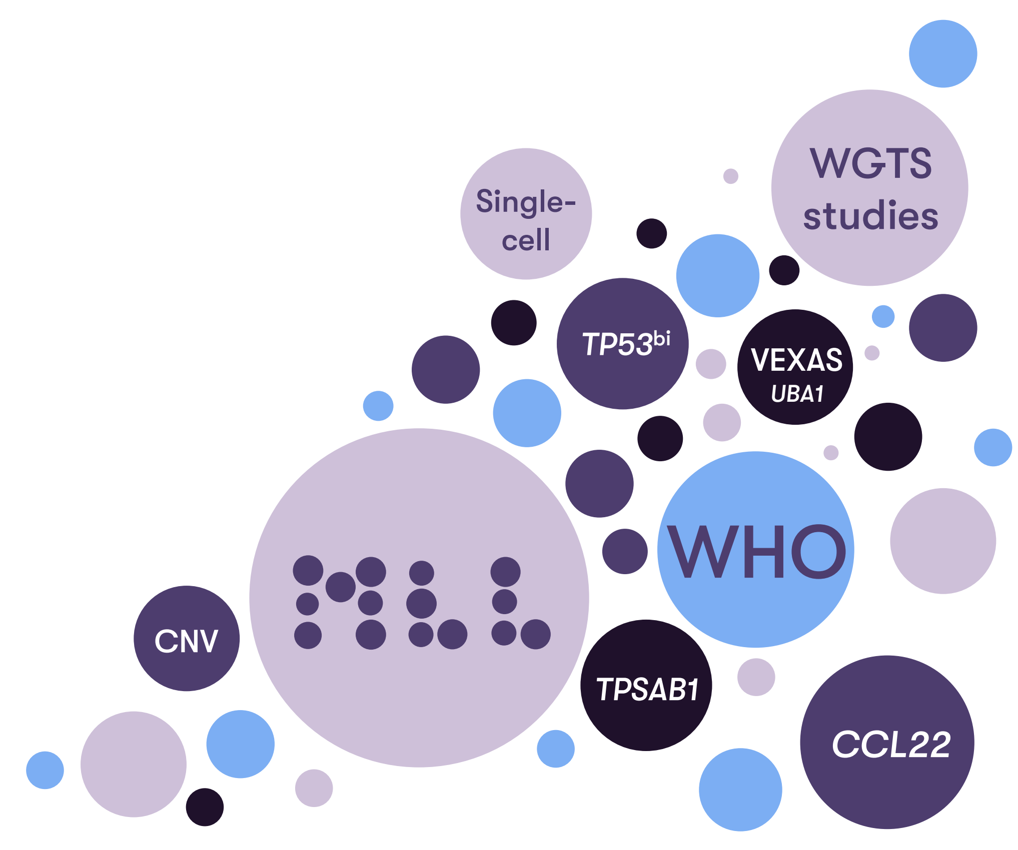
RESEARCH REPORT

2022

CHALLENGING CURRENT STANDARDS, BECAUSE OUR RESEARCH IS ALWAYS IN MOTION



Dear research interested reader,

after another scientifically successful year, we would like to take the opportunity to summarize our findings of 2022. 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 [Mediendienstestaatsvertrag] Prof. Dr. med. Wolfgang Kern

Table of contents

Foreword

Refinement of the MLL diagnostic workflows

- Implementation of the new WHO classification scheme in routine diagnostics (p. 6)
- CNV spike-in panel for the assessment of *TP53* allelic state (p. 7)
- Establishing the new molecular international prognostic scoring system (IPSS-M) for MDS (p. 8)
- Detection of hereditary alpha tryptasemia via ddPCR (p. 9)
- Detection of somatic mutations in *UBA1* associated with VEXAS syndrome (p. 10)

WGS & WTS for routine diagnostics

- The choice of the right control sample (p. 12)
- How WGTS can contribute to unsolved cases in hematology (p. 13-14)
- Detecting the unusual without compromising diagnostic accuracy – WGTS for acute leukemias (p. 15-16)

WTS for the detection of fusion transcripts in MLN-TK (p. 17)

WTS-based BCP-ALL diagnostics (p. 18)

Published research highlights

- *CCL22* as the hallmark of a distinct subset of CLPD-NK (p. 21)
- Manuscript speed dating (p. 22-25)

INTERCEPT-MDS (p. 26)

ASH recap

- Genetics as a mainstay of new classifications and to improve prognosis estimation (p. 28)
- The application of machine learning in hematology diagnostics (p. 29)

Outlook

- Genome network hematology (p. 31)
- Single-cell sequencing (p.32)
- Development gems (p. 33)

Foreword

Those who know us, know that we always look ahead, push boundaries, and rarely stand still. Nevertheless, we would like to look back at our scientific year 2022, review the highlights and share with you our efforts to generate and share knowledge. While the years before were marked by the pandemic, our world has faced other challenges in the past year or is still doing its best to overcome these crises. Nonetheless, our research and its integration into our patient care has not stagnated but remained in motion in 2022.

We continued to refine and optimize our diagnostic workflows. A major endeavor here was certainly the publication of the new WHO 2022, to which we contributed not only as scientific experts but also as authors. We also promptly introduced the new classification into our daily diagnostic routine in the laboratory. A newly defined entity of MDS with bi-allelic *TP53* inactivation has encouraged us to include the CNV spike-in panel in routine diagnostics to fully address *TP53* alterations. This has also a major impact on the published IPSS-M risk score, which we not only use in the laboratory, but also share real-life data with the scientific community. New genes like *UBA1* or the analysis of alpha tryptasemia are now part of our daily routine.

As announced at the end of 2021, we have continued to address holistic sequencing using whole-genome (WGS) and whole-transcriptome sequencing (WTS) and have demonstrated in a clinical setting what these methods mean in particular in AML and ALL. The algorithms for data interpretation are becoming more and more sophisticated and especially the grouping in ALL shows impressively what is possible using transcriptome analysis. What started as a research project is now diagnostic reality. But also unresolved diagnoses continue to tempt us like detectives to still find sequence changes that represent a possible diagnosis and therapy for the patient. Since we also want to contribute to this future patient care on a national level, you will stumble across the Genomnet - Genome Network Hematology - in the coming year.

With a proud 63 peer-reviewed publications, we have again shared our knowledge. While one quarter of the publications originate genuinely from our laboratory, the other three quarters highlight the international collaboration that science thrives on. Without great collaborations, such data, cohorts, and hypotheses would not be conceivable. This year, we again express our gratitude for our collaborators and partners who are equally committed to advancing leukemia diagnostics and working with us to bring the best to our patients. For example, the discovered and now high-ranking published *CCL22* as a mutation hotspot in CLPD-NK is certainly one of our highlights.

What has pleasantly made the scientific year 2022 different from the two previous years is the revived congress culture. As practical as virtual meetings can be, it is nice to meet face to face with colleagues again, discussions seem to be so much easier and lively. Thus, within the course of the Intercept-MDS, not only the mid-term meeting could take place in our premises, but we also participated in the Spetses Summer School 2022. But not only the young scientists were granted this, we were also back at the DGHO (Vienna), EHA (Vienna) and of course at the ASH (New Orleans) congresses. Judging by the contributions at ASH, it was a successful scientific year and meeting for us. In addition to 8 talks, we presented our data in 14 posters and covered a broad spectrum from changes in classifications to the application of artificial intelligence. However, not only the short-term exchange at congresses, but also our guest scientists energized the scientific exchange. We had three international guests from Japan, Australia and the USA, as well as a colleague from Berlin. It is always impressive what new ideas, hypotheses and results can emerge in such an atmosphere.

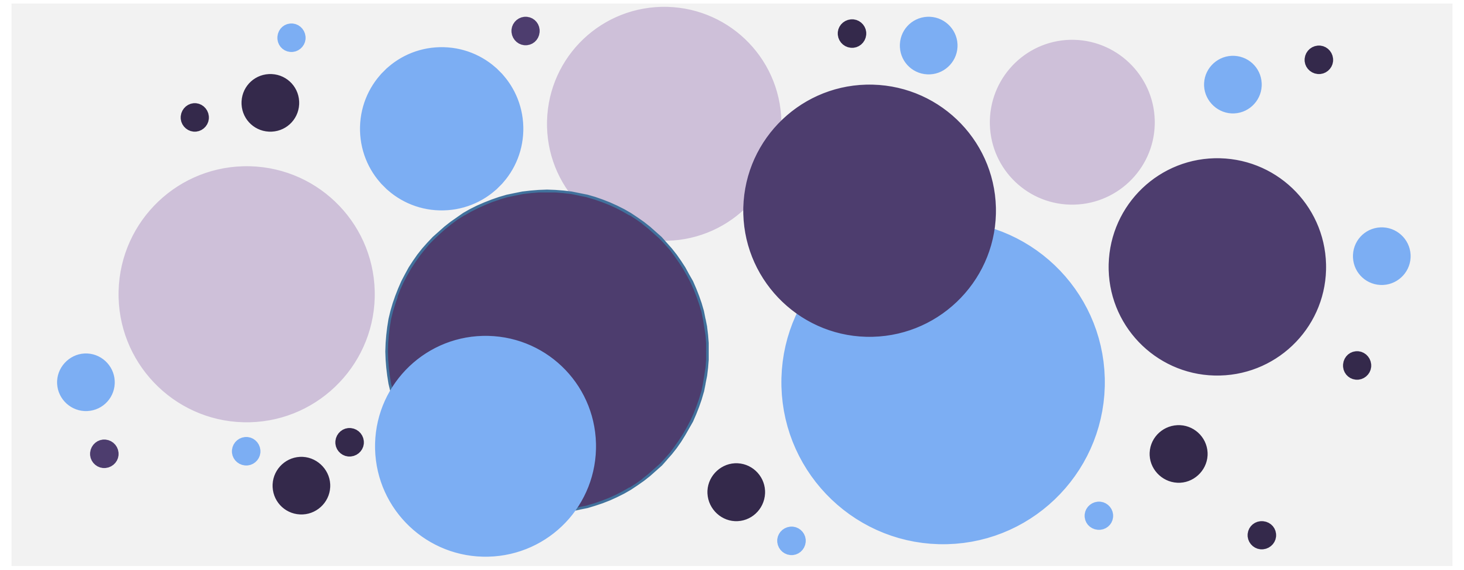
Even though it is risky to speculate about the future, topics are already in sight that we want to deal with in the new year 2023. In part, these are already taking shape, so we dare to give a somewhat more detailed outlook on the upcoming topics in this Research Report, of course also to pique your curiosity.

We remain committed to our passion for research, and this year we are once again starting a new scientific year with verve, enthusiasm and drive.

M. Meggendorfer

Dr. rer. nat. Manja Meggendorfer, MBA,
Head of Research & Development

Refinement of MLL diagnostic workflows



Implementation of the new WHO classification scheme in routine diagnostics

Since its first edition in 1981, the publication of new editions of the WHO classification has always taken many years. In order to keep up with the rapid growth of knowledge, the authors of WHO 2017 set the goal to publish the new edition within 5 years. Thus, in the summer of 2022, the first version of the 5th WHO edition „classification of haematolymphoid tumours“ was published as an online ß-version. The constantly growing knowledge and the increasingly complex genetic landscape has necessitated the collaboration of more than 420 experts as authors on the new classification. In addition, there is a shift from originally morphologically based diagnostics to a genetically driven classification system, which is particularly evident in the classification of myeloid diseases (Fig. 1).

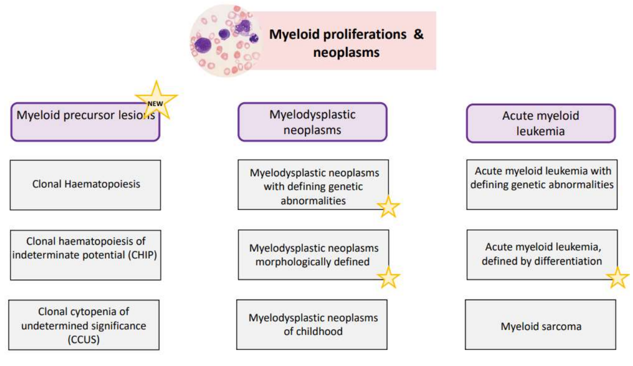


Figure 1: Changes/updates introduced in the 5th edition of the WHO Classification of Haematolymphoid Tumours: Myeloid and Histiocytic/Dendritic Neoplasms.

Thus, in MDS two genetic subgroups - „MDS with *SF3B1* mutation“ and „MDS with biallelic inactivation of *TP53*“ - have been introduced in addition to MDS with deletion 5q. Genetics also strongly impacts AML diagnostics: AML with genetically defined abnormalities are now ranked higher than AML classified by differentiation - effectively eliminating the blast cutoff for the vast majority of genetically defined AML cases. In contrast, the 20% blast cutoff still applies to distinguish between AML defined by differentiation and MDS.

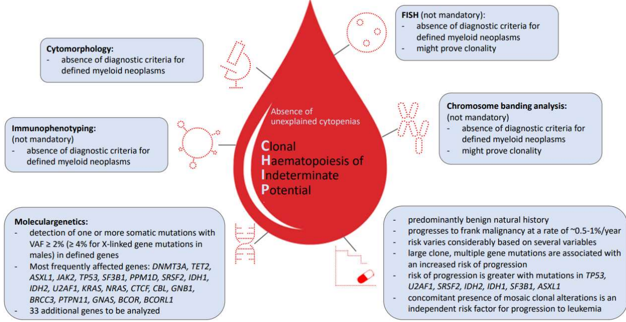


Figure 2: Formal definition of clonal hematopoiesis of indeterminate potential according to the new WHO guidelines.

Due to the increased understanding of the molecular genetics of clonal hematopoiesis (CH), the 5th WHO classification recognizes CH as a completely new category of precursor myeloid disease state, including clonal hematopoisis of indeterminate potential (Fig. 2) and cytopenia of undetermined significance (Fig. 3).

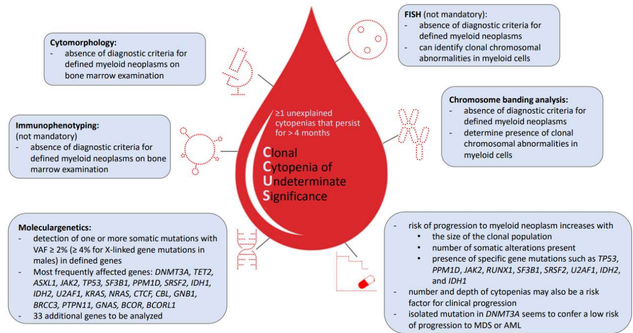


Figure 3: Formal definition of cytopenia of undetermined significance according to the new WHO guidelines.

An improved hierarchical taxonomy has been added to groups that do not leave any disease in a vacuum, allowing any alteration to be grouped into "other genetic alterations." With the new structure and a relational database behind it, we can hope for faster updates in the future. For early 2023, we plan to implement the new WHO diagnoses into our internal systems and workflows, and the new diagnoses will of course also be used in our reports. But not only in diagnostics, which we provide daily to our patients and their treating physicians, but also in research we have adopted the new WHO classification.

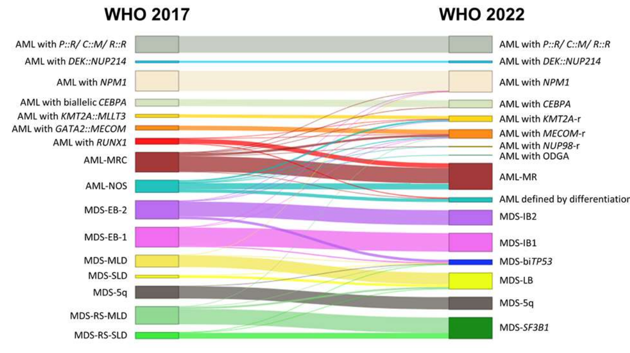


Figure 4: Sankey plot of the diagnoses of AML/MDS patients according to WHO 2017 on the left side and the new WHO on the right side.

At this year's ASH 2022 in New Orleans, the new WHO classification was one of the frequently discussed topics, along with the ICC (p. 28). We were able to quantify the changes in the new classification and demonstrate the changes for AML/MDS diagnostics (Fig. 4).

Introduction of CNV spike-in panel for *TP53* allelic state assessment

TP53 alterations are present in both myeloid and lymphoid malignancies and are associated with complex karyotype and dismal outcome. *TP53* alterations are poor prognostic markers, comprising gene mutations, as well as allelic imbalances, including *TP53* deletions (del) and regions of copy-neutral loss of heterozygosity (CN-LOH, Fig. 1).

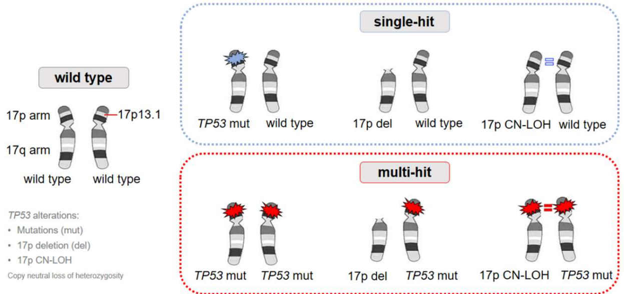


Figure 1: Schematic overview of *TP53* alterations.

Chromosome banding analysis (CBA) and FISH are the current gold standard for the detection of chromosomal aberrations, including 17p/*TP53* del. However, CN-LOH events are not detectable by either technique but play a pivotal role for *TP53* allelic state assessment and risk stratification in MDS patients (Meggendorfer et al. 2022). Moreover, MDS with biallelic *TP53* inactivation has recently been classified as a separate entity according to the WHO classification 2022 (previous page), as only the multi-hit state predicts risk of death and leukemic transformation. In recent years, it has been shown that whole exome sequencing and whole genome sequencing are valid alternatives to CBA. However, both assays are still quite expensive, hampering their broad application in routine diagnostics of hematological malignancies, for now. Hence, we extended our routine molecular genetics workflow with a CNV spike-in panel to simultaneously assess copy number variations (CNV) and CN-LOH (Fig. 2).

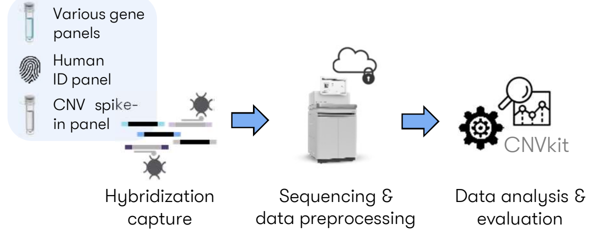


Figure 2: Schematic overview of the routine target enrichment workflow to detect CNV and CN-LOH events.

The xGen human CNV backbone panel (Integrated DNA Technologies) contains >9,000 oligonucleotide probes that span the entire genome with a distance of ~0.34 Mb. The obtained coverage and variant calling data is automatically analyzed with the CNVkit software toolkit (*Talevich* et al. 2016, https://doi.org/10.1371/journal.pcbi.1004873). The coverage tracks are normalized by a reference profile, compiled from >50 samples with a normal karyotype (Fig. 3).

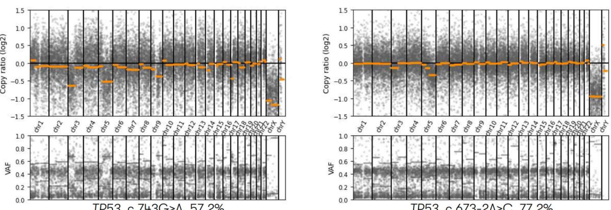


Figure 3: Examplary plots for patients with del(17p) (left), and CN-LOH in 17p (right).

Taking advantage of the optimized workflow and available WGS data we found that MDS cases with <5% blasts separate from MDS cases ≥5% blasts and from AML cases by the predominance of *TP53* single hit and by being the only subgroup in which a complex karyotype showed an independent adverse impact on overall survival (OS, Fig. 4).

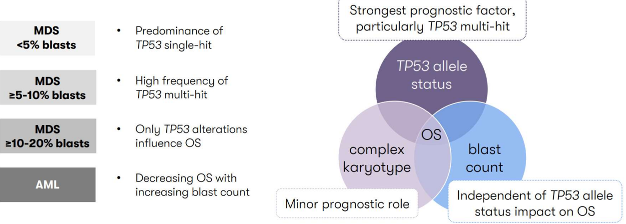


Figure 4: Graphic summary of the impact of complex karyotype, blast count, and *TP53* allelic state on OS.

Since CN-LOH is not routinely assessed in most diagnostic laboratories, it was proposed to use the variant allele frequency (VAF) of *TP53* mutations to estimate *TP53* allelic state. We assessed the *TP53* allelic state in 2,076 cases with suspected MDS and identified 185 cases with *TP53* mutation. 113/185 were classified multi-hit, whereof 21/113 (19%) would have been misclassified without CN-LOH information (Fig. 5). VAF was not sufficient to predict the *TP53* allelic status, due to its broad ranges (e.g. 15-95% for cases with CN-LOH).

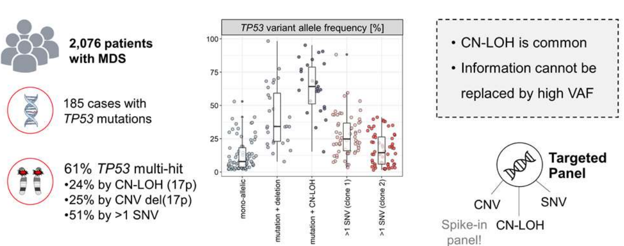
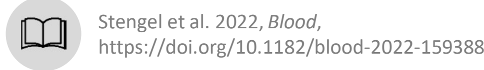


Figure 5: VAF distribution for various *TP53* alterations depending on the allelic status.

Meggendorfer et al 2022, DGHO, https://dgho2022.abstractserver.com/program/#/details/presentations/594



Stengel et al. 2022, *Blood*, https://doi.org/10.1182/blood-2022-159388

Establishing the new molecular international prognostic scoring system (IPSS-M) for MDS

The molecular international prognostic scoring system (IPSS-M) is a new MDS prognosis calculator that combines genomic profiling with hematologic and cytogenetic parameters, improving the risk stratification of MDS patients. The new scoring system supports clinical decision making by empowering diagnosis and therapeutic interventions to be tailored to each patient's molecular profile. The IPSS-M and an accompanied calculator/algorithm were introduced by Bernard et al. in June 2022 and quickly gained popularity. In Oct 2022 more than 5,000 unique users from 65 different countries had used the IPSS-M model to compute risk for 42,500 patient profiles (https://bit.ly/3VKwsWX). In contrast to the previously used IPSS-R, the IPSS-M incorporates the mutational status of 31 genes (Fig. 1).

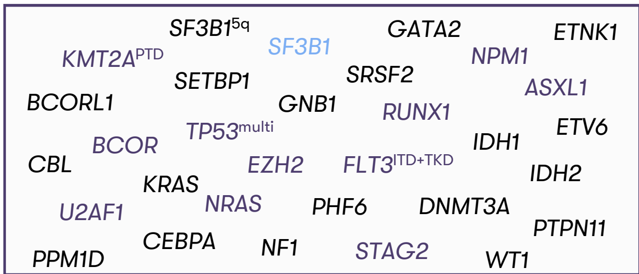


Figure 1: Overview of the 31 genes contributing to IPSS-M. The font color indicates the prognostic impact (purple = poor outcome; blue = good outcome; black = no significant association).

Most commonly used gene panels do not capture all the required genes of the IPSS-M model. However, the calculator accounts for missing values by calculating the IPSS-M under the best, average, and worst case scenarios. Moreover, part of the output of the study by Bernard et al. was to also inform panel design, to develop "a solution for the present and future, not for the past " (quote from Elli Papaemmanuil). The IPSS-M score can be assigned to one of six risk categories and is scaled so that a score value of 0 represents the average patient (Fig. 2). Values of -1, 1, or 2 correlate

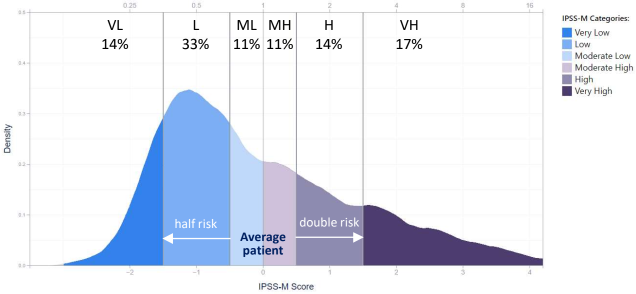


Figure 2: Density plot of IPSS-M risk scores calculated for 2,701 MDS patients with available data for hemoglobin, platelets, marrow blasts, gene mutations, and cytogenetics (original by Bernard et al. 2022).

to half, double, or fourfold risk compared with the average patient. In a recent study, Baer et al. independently validated the results of the IPSS-M in a cohort of 735 MDS cases (419 with <5% blasts) characterized in depth by WGS (Table below).

Table 1: IPSS-M risk score distribution in the validation cohort of 735 MDS cases.

IPSS-M categories	MLL	Bernard *et al.*
Very Low (VL)	15%	14%
Low (L)	41%	33%
Moderate Low (ML)	11%	11%
Moderate High (MH)	7%	11%
High (H)	12%	14%
Very High (VH)	14%	17%

The IPSS-M risk categories showed strong prognostic separations for OS, leukemia-free survival (LFS) and leukemic transformation (LT). However, in contrast to Bernard et al., the two intermediate risk categories - ML and MH - did not separate in our cohort (Fig. 3).

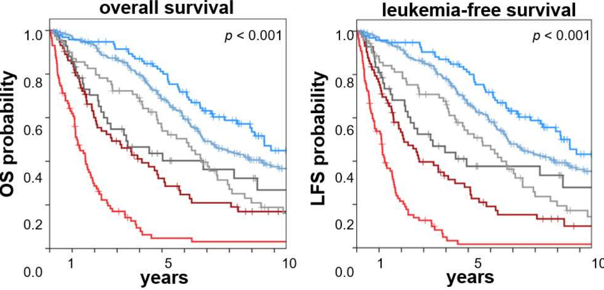
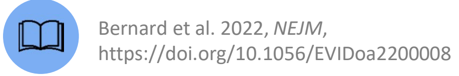


Figure 3: Overall survival and leukemia-free survival for 735 MDS cases sent to the MLL between 2006 and 2019.

At ASH 2022, there was a full session on the IPSS-M with presentations from 3 institutes (incl. the MLL) on validation of the risk score at different clinics and diverse real-world populations. In general, the obtained results matched the reported findings by Bernard et al. and it was often found that the IPSS-M improved MDS prognostication and may lead to better selection of patients for specific treatments and/or clinical trials.

Integrated Report

The IPSS-M can be calculated on the website of MDS foundation (https://mds-risk-model.com/). If the treating physician submits blood count values and orders cytomorphological, cytogenetic and molecular genetic diagnostics, the IPSS-M risk score of the patient is listed in our integrated report to support tailored diagnosis and treatment options.



Bernard et al. 2022, *NEJM*, https://doi.org/10.1056/EVIDoa2200008

Baer et al. 2022, *Blood*, https://doi.org/10.1182/blood-2022-159939

TPSAB1 digital droplet PCR CNV analysis for the detection of hereditary alpha tryptasemia

Background

Hereditary alpha tryptasemia (HαT) is a recently described autosomal dominant genetic trait caused by increased germline copies of the gene *TPSAB1* and occurs in approximately 5% of the general population. The tryptase locus on the short arm of chromosome 16 contains four tryptase encoding genes but only *TPSAB1* and *TPSB2* encode the isoforms of the enzyme tryptase (97% identical), a serine protease, which are secreted by activated mast cells. After mast cell activation or in the presence of a mastocytosis, higher serum tryptase basal levels can be measured. The *TPSAB1* gene can be expressed in the allelic constellations α or β (*TPSAB2* is always present as β), so that in normal cases 0, 1, or 2 α-alleles of *TPSAB1* are present in the germline. It has been shown that the basal serum tryptase levels and the severity of clinical symptoms display a gene dose relationship with the numbers of alpha-encoding *TPSAB1* copies.

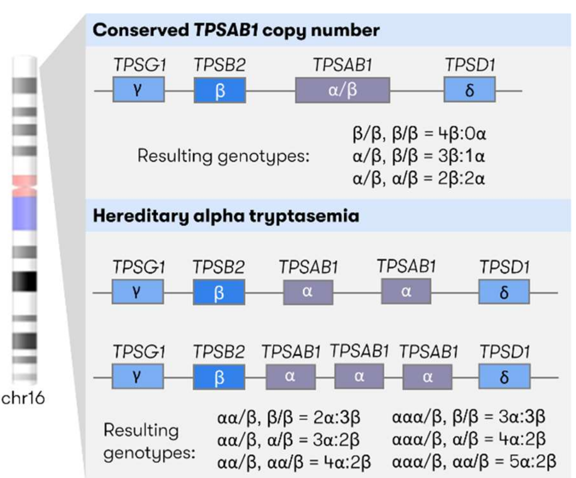


Figure 1: *TPSAB1* genotypes (modified from Lyons et al. 2016)

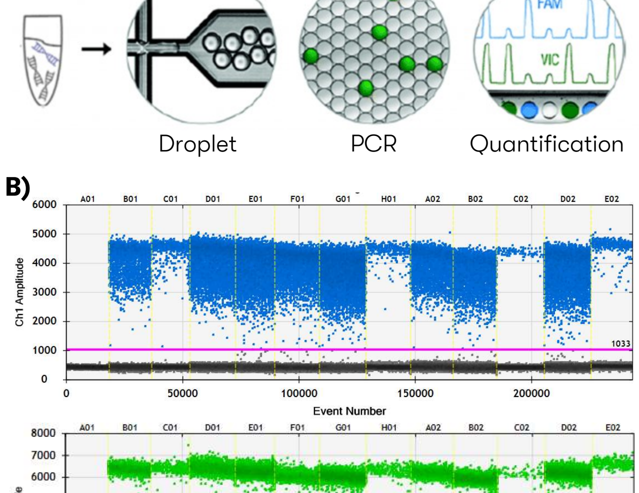


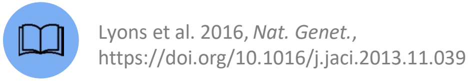
Figure 2: Digital droplet PCR. A) Schematic overview of the ddPCR workflow. B) Exemplary plots of a *TPSAB1* positive case (FAM-channel blue dots, HEX-channel green dots). Positive droplets are those above the threshold intensity indicated by the pink line.

Digital droplet PCR

The concept of ddPCR, which is often summarized as 'divide and conquer' was first described in 1992. The idea is to dilute the sample and divide it into separate wells so that each well contains one or no copies of the sequence of interest, which is then amplified. Then, only the positive and negative signals need to be counted to determine the copy number of a DNA molecule. Although the theory behind ddPCR is simple, reliable implementation has proven difficult. However, with advances in microfluidics technology, ddPCR can now be used to achieve precise target DNA quantification at high levels of sensitivity and specificity. In comparison to real-time qPCR, ddPCR is simpler, faster, and less error prone. The basic protocol can be modified for CNV analysis. CNV ddPCR uses duplex target (FAM) and reference (HEX) Taqman assays. To detect hereditary alpha tryptasemia, the FAM-labelled probes are specific for *TPSAB1α/β* and the HEX-labelled probe for *AP3B1*, which is used as reference locus. The obtained data can be visualized for each channel separately in a 1D plot (Fig. 2B). For *TPSAB1α/β* germline copy number determination, the technical copy numbers of the *TPSAB1α/β* allele are normalized against the copy numbers of the diploid *AP3B1* gene.

Droplet Digital PCR for Minimal Residual Disease Detection

Minimal residual disease (MRD) detection has prognostic relevance in hematological malignancies. By analyzing MRD levels, molecular response and the potential reoccurrence of cancer can be monitored. Due to the high sensitivity, precision, and absolute quantitation capabilities of the assay, ddPCR technology is an ideal tool for MRD monitoring. At the MLL we use ddPCR to quantify MRD levels of various genes and mutations, including *NPM1*, *KIT*, and *TP53*.



Lyons et al. 2016, *Nat. Genet.*,
https://doi.org/10.1016/j.jaci.2013.11.039

Detection of somatic mutations in *UBA1* associated with VEXAS syndrome

Background

Recently, VEXAS syndrome, a disease only discovered in 2020, was deemed the "year's best in hematology diagnosis" by the American Society of Hematology. The identification of a new hotspot mutation in *UBA1*, the causal gene of VEXAS syndrome, offered new insights into the known association of rheumatological diseases and hematological malignancies. *UBA1* is located on the X chromosome in humans, but is a known X chromosome inactivation escape gene and somatic *UBA1* loss of function variants have been associated with a severe hematoinflammatory disease, subsequently termed VEXAS (Fig. 1). Despite its young age the disorder has already made it to primetime TV: an episode of Chicago Med featured a patient who was diagnosed with VEXAS syndrome. Due to the raising awareness of VEXAS syndrome, hematologists are realizing that they have patients who likely have this disorder, and hence the demand for *UBA1* sequencing has increased in recent months.

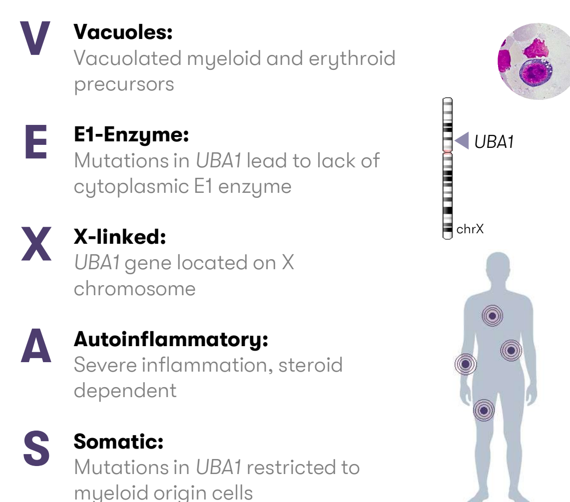


Figure 1: Schematic overview of the keypoints of VEXAS syndrome.

Detection of *UBA1* variants – sequencing of the whole gene or just a single exon?

The original work by Beck et al. identified three causative variants, affecting methionine-41 in *UBA1* (M41T, M41L, M41V), which altered the start codon of the cytoplasmic isoform in exon 3. Hence, subsequent screenings of large cohorts (> 1,000 patients) focused on *UBA1* exon 3, revealing another, potentially causative non-synonymous variant (S56F). A large study of 3,328 patients that looked at the prevalence and characteristics of *UBA1* hotspot mutations in MDS was presented at ASH 2022. They could show that a *UBA1* mutation alone or in combination with other mutations involved in myeloid pathogenesis accounts for a significant proportion (8%) of patients diagnosed and treated for MDS without disease classification. However, as part of the Intercept-MDS program (p. 26), our PhD student M. Sakuma, reviewed the whole genome data of 853 patients and identified *UBA1* variants other than those at codon M41 in MDS patients, which might also be relevant in VEXAS syndrome pathogenesis (Fig. 2).

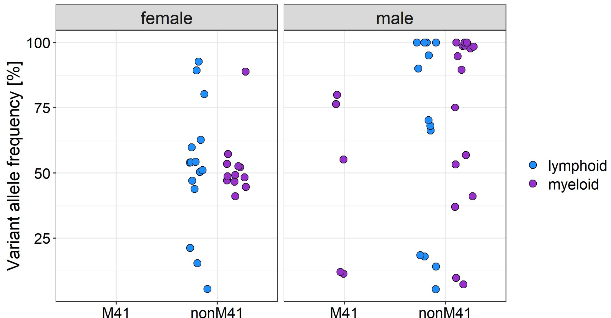


Figure 2: Variant allele frequency distribution of *UBA1*^{M41} and *UBA1*^{nonM41} variants for female and male patients from different lineages (blue = lymphoid, purple myeloid).

Further studies are needed to confirm and extent our results but our findings suggest that sequencing of the entire *UBA1* gene would be beneficial to better understand VEXAS syndrome and hematological malignancies.

Transcriptional profile of *UBA1*^{M41} cases

Interestingly, analysis of the transcriptome did not result in a specific expression signature of *UBA1*^{M41} patients in comparison with MDS patients but confirmed neutrophil activation in VEXAS patients compared to healthy controls (Fig. 3).

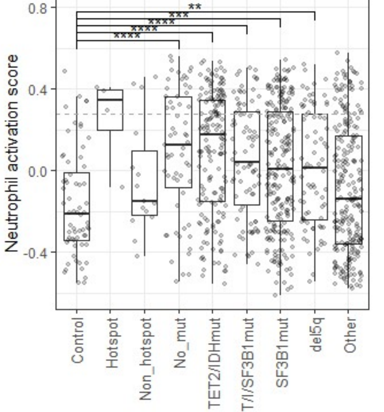
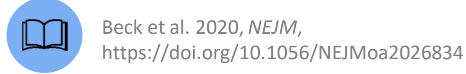


Figure 3: Neutrophile activation score for controls, different MDS subgroups and patients with *UBA1*^{M41} (Hotspot) and *UBA1* variants outside codon M41 (Non_hotspot).

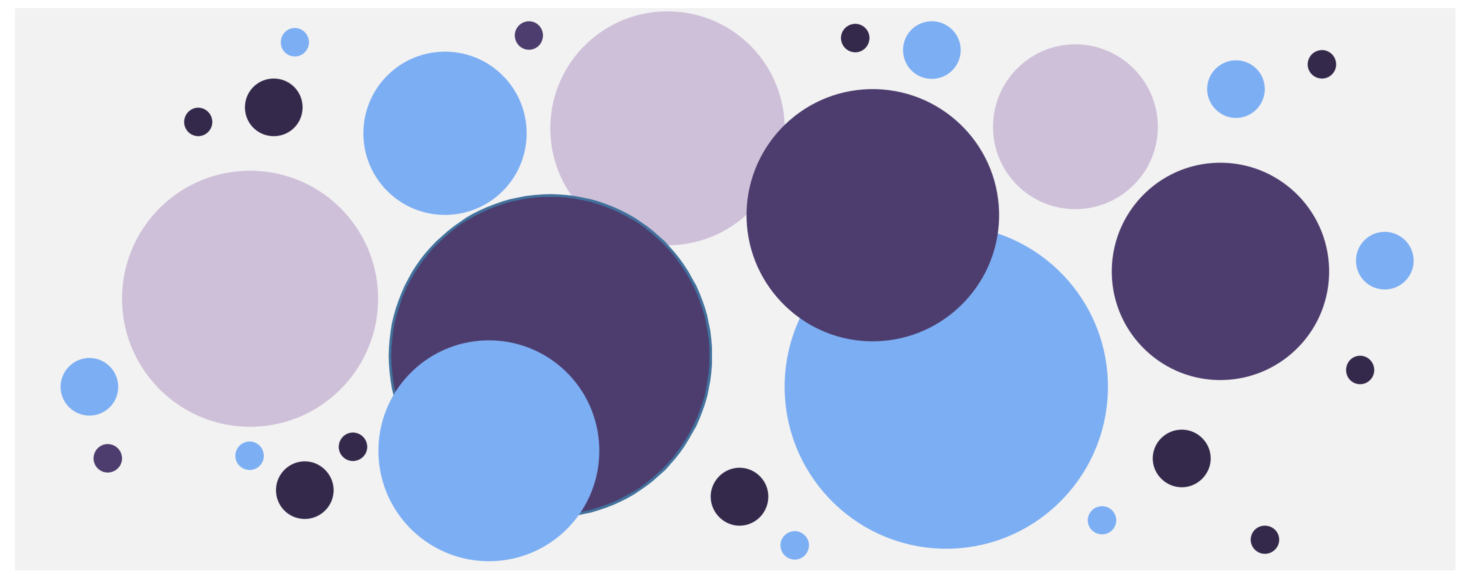
UBA1 sequencing at MLL

Our myeloid panel captures the entire gene sequence of *UBA1*, not only exon 3.



Beck et al. 2020, *NEJM*, https://doi.org/10.1056/NEJMoa2026834

WGS & WTS for routine diagnostics



The choice of the right control sample in hematology – tumour in normal contamination of CD3+ cells

The choice of the right control sample

In the last Research Report, we already discussed the current challenges in implementing WGS and WTS in routine diagnostics and briefly touched upon the problem of finding the right control for bone marrow and peripheral blood samples. In the last year, some progress has been made in the development of tumor-only pipelines, but a normal control is still necessary for reliable discrimination between germline and true somatic mutations (Fig. 1).

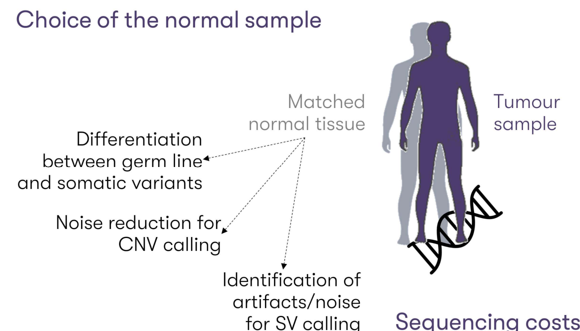


Figure 1: Matched normal tissue offers several advantages for variant identification but it should be considered that it also increases sequencing costs.

Tissues such as buccal swap, hair follicles, cultivated fibroblasts, skin punches or sorted T-cells are commonly used as control samples in hematology. For myeloid diseases it is also common practice to use CD3+ cells. T cells can be used as a normal control in myeloid disease because they are derived from a different progenitor cell and thus should not carry the somatic changes of the disease. Here, the purity of the CD3 sorted cells must be assessed prior to any library preparation to minimize the risk of contamination. At the MLL, CD3+ MAC sorted cells are used as normal controls for myeloid cases with sufficient material, requiring a CD3 purity >90% and <5% blasts (Fig. 2A).

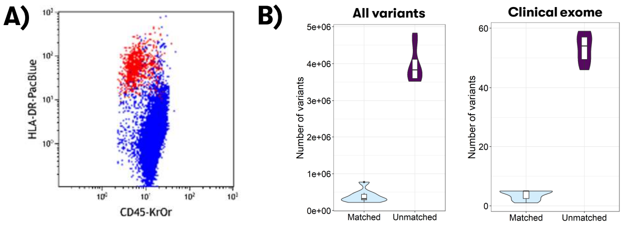


Figure 2: CD3+ cells as normal controls. A) Exemplary CD3+ sample with >90% T cells (blue) and <5% blasts (red). B) Number of all identified variants (left) and the clinical exome (right) with and without a matched control.

As expected, the total number of variants detected in an unmatched analysis is significantly higher than the number of variants detected in the presence of a matched control sample (Fig. 2B). This is true even when the region of interest is reduced to the clinical exome, as is often done for the clinical evaluation of WGS data.

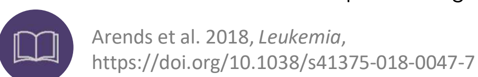
https://caravagnalab.github.io/TINC/

Tumor in normal contamination of CD3+ cells

Despite the advantages and feasibility of using CD3+ cells as control samples, a small fly in the ointment remains: the presence of somatic mutations in the CD3+ cells. We can identify novel and patient-specific variants by what is called tumor-normal subtraction – identifying variants that are present in the tumor but absent from the germline. The subtraction is done bioinformatically, but most algorithms expect that the germline is tumor-free. This is also the reason why the purity of the CD3 cells is checked so strictly - to minimize the risk of contamination and to reduce the background noise. However, even with high CD3 cell purity, it's virtually impossible to obtain a control sample that is completely free of cancer cells. With a sufficiently high contamination, tumor variants could therefore be falsely labeled as germline during subtraction and would thus no longer be visible to the biologist who evaluates the variant calls (Fig. 3).

Figure 3: Tumor-normal subtraction with contaminated control sample.

What we saw in our own small study (p. 15-16) was observed and followed up on a large scale by Genomics England. Their bioinformatics team developed the software package TINC – a bioinformatics approach to determine the contamination of tumor DNA (tumor read fraction) in a matched normal sample and to adjust the variant calling accordingly. However, tumor in normal contamination is not the sole reason for the presence of somatic variants in the germline sample. Arends et al. 2018 already showed that CHIP (clonal hematopoiesis of indeterminate potential) mutations could also be identified in B- and T-cells, albeit with a significantly lower variant allele frequency. Nevertheless, even a small number of reads can cause the variant to be removed during subtraction and no longer be visible to the evaluator. Future single-cell studies will show whether some somatic mutations might occur at an earlier maturation state than previously thought.


Arends et al. 2018, *Leukemia*, https://doi.org/10.1038/s41375-018-0047-7

How WGTS can contribute to unsolved cases in hematology – a prospective head-to-head study

The WHO classification of hematologic tumors serves as reference to diagnostic decision making, combining the results from cytomorphology, immunophenotyping, chromosome banding analysis (CBA) supplemented by FISH, and molecular genetics (including targeted panel sequencing). However, 5 to 10% of cases with unexplained blood count changes remain a challenge for clinicians and hemato-pathologists. With the advent of high throughput sequencing and the concomitant decline in sequencing costs, comprehensive and detailed genomic profiling has become a viable alternative to classical methods for patients with challenging cases. We set up the prospective study SIRIUS (see next page for more information) to test if whole genome and whole transcriptome sequencing (WGTS) can resolve cases with unexplained blood count changes and to compare WGTS to gold standard diagnostics with respect to clear-cut diagnosis and turn-around time.

Patients and Methods

Since 09/2021 31 patients with unsolved suspected blood cancers have been enrolled after informed consent (Fig. 1).

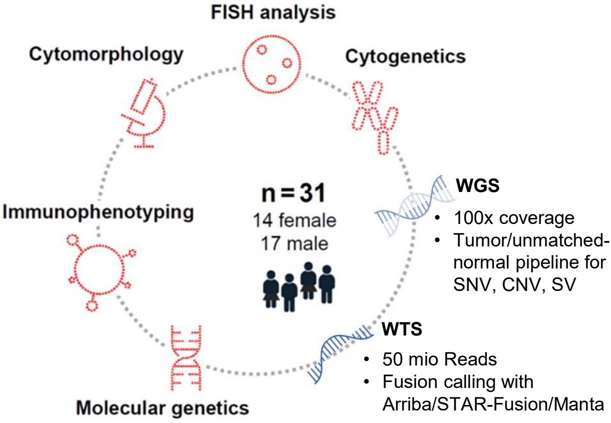


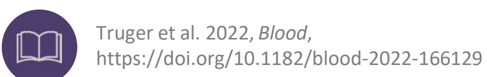
Figure 1: Enrolled patients and standard diagnostics.

Results

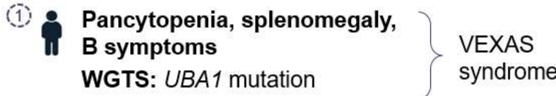
In comparison with gold standard diagnostics WGTS provided additional information in 10/31 (32%) patients within a median of 21 days after sample collection.

WGS identified rare SNVs

Variants detected by WGS included SNVs with association to Diamond-Blackfan anemia (*RPL11*) and VEXAS syndrome (*UBA1*) and a variant of uncertain significance with possible association to Pelger-Huët anomaly (*LBR*).



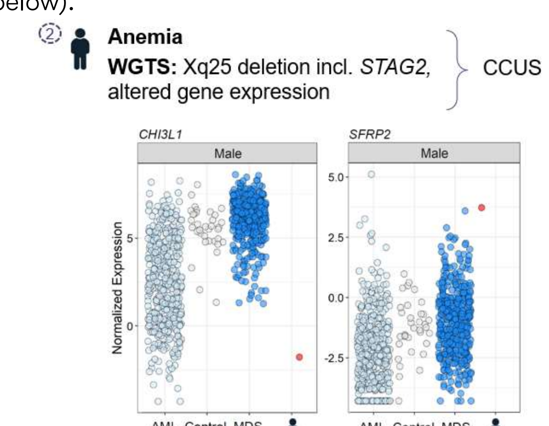
Truger et al. 2022, *Blood*, https://doi.org/10.1182/blood-2022-166129



The analysis of identified variants in a 46-year-old male patient with clinical features of VEXAS syndrome revealed two mutations in the *UBA1* gene, distinct from the known *UBA1* hotspot mutations associated with VEXAS syndrome. Neither of these variants was detected in CD3+ cells isolated from peripheral blood from the same patient consistent with their somatic origin.

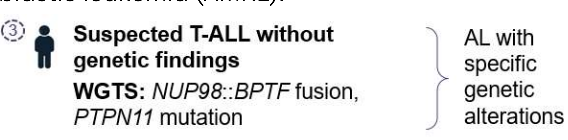
WGTS led to new findings

In one patient with unexplained persistent cytopenia WGS revealed an Xq25 deletion comprising *STAG2* and provided evidence of clonal hematopoiesis (CCUS). *STAG2* mutations are well known in myeloid neoplasms, but *STAG2* deletions are not. *STAG2* loss in an AML cell line resulted in altered gene expression (Smith J et al. 2020) and the patient's WTS data showed consistent up-/downregulation of specific genes (e.g. *SFRP2*, *CHI3L1*), compared to normal controls and to MDS/AML patients w/o *STAG2* alteration (see below).



WGTS identified clones with no proliferation in vitro

WGTS of a 26-year old man showed an SV of chromosomes 11 and 17 resulting in *NUP98::BPTF* fusion. Due to lack of proliferation in vitro the aberrant clone was not detected by CBA. *NUP98* rearrangements mostly occur in children with AML and in T-ALL. Specifically, the fusion *NUP98::BPTF* was found in a patient with acute megakaryoblastic leukemia (AMKL).



Smith et al. 2020, *Journal of Translational Medicine*, https://doi.org/10.1186/s12967-020-02500-y

How WGTS can contribute to unsolved cases in hematology – a prospective head-to-head study

WGS clarified SVs already detected by CBA

One patient with Waldenstroem macro-globulinemia and unusual clinical presentation showed a t(14;19) usually leading to an *IGH::BCL3* rearrangement. WGS identified *NECTIN2* instead of *BCL3* as translocation partner, which is recurrently trans-located in T-cell neoplasms and was found to be rearranged to IGH in DLBCL.

WGTS detected therapeutically relevant alterations

At relapse after BCMA-targeting CAR T-cell therapy WGTS identified biallelic deletion of *TNFRSF17* (*BCMA*, Fig. 2) and strongly reduced gene expression in a patient with multiple myeloma (MM).

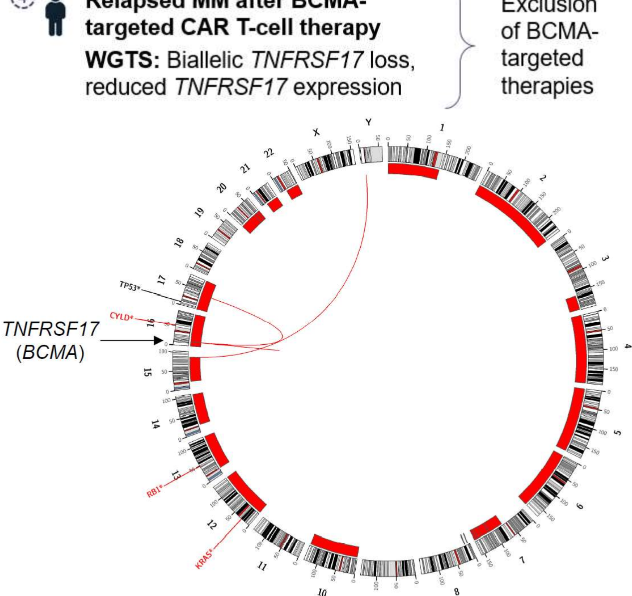
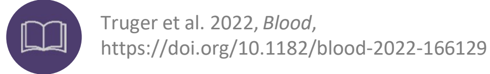


Figure 2: Circos plot, MM patient at relapse after BCMA-targeted CAR T-cell therapy. Inner track shows losses (red), red lines represent SVs, genes with mutations are depicted in red (Tier 1) and black (Tier 2)

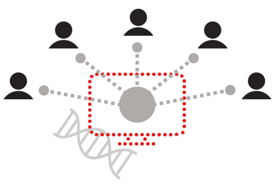
Such escape from immunotherapy by target inactivation has been shown to be irreversible (Da Viă M et al. 2021), guiding further treatment decision as other BCMA-targeted therapies have to be excluded. The study also showed that in 37/168 patients with MM heterozygous *TNFRSF17* loss or monosomy 16 can be identified at baseline and might be a risk factor for BCMA loss after immunotherapy. Hence, therapeutically targeting a combination of different MM targets may overcome the substantial genomic heterogeneity in MM, avoiding clonal selection based on the loss of one particular antigen.



Truger et al. 2022, *Blood*, https://doi.org/10.1182/blood-2022-166129

Conclusions

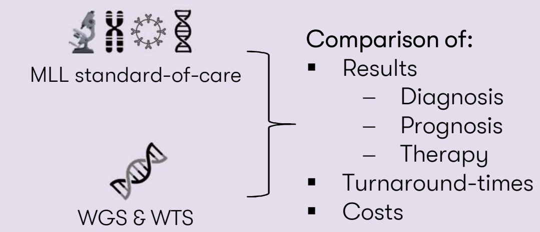
Initial real-world data of this ongoing study demonstrate the power of WGTS to add valuable information to gold standard diagnostics in about 1/3 of unclear cases within three weeks. By providing diagnostically and therapeutically relevant results, WGTS can save costs on further tests or inappropriate therapies. In the context of challenging cases, close cooperation of experts in hematology, pathology, human genetics and bioinformatics (tumor boards) allows best possible use of such comprehensive data and the most appropriate diagnoses for patients.



Solving Riddles Through Sequencing (SIRIUS)

ClinicalTrials.gov Identifier: NCT05046444

The objective of this trial is to test whether WGS & WTS can surpass the current gold standard regarding diagnostic precision and routine reliability for challenging cases.



We are still recruiting patients for this trial! If you have a patient for whom individual diagnostic findings and determined blood values do not fit together conclusively - or simply an uncertainty remains as to whether the patient is definitely clinically ill, then send us material from your patient and the signed consent form and we will do the rest.

What do we need?

- 5-10 ml blood or bone marrow aspirate
- Signed consent form

What do you get?

- Comprehensive genetic profiling of your patient

Detecting the unusual without compromising diagnostic accuracy – WGTS for hematological malignancies

Following WHO and ELN guidelines, standard diagnostics and risk stratification for acute leukemias (AML and ALL) are based on genetic information gathered by a combination of cytogenetics and targeted molecular profiling. With the introduction of comprehensive assays such as WGS and WTS and associated clinical bioinformatics workflows, the diagnosis of hematologic malignancies can be further refined and individualized. The significant reduction in sequencing costs makes it possible to obtain a comprehensive genetic profile for every patient in a reasonable time but is it really beneficial? We aimed to evaluate the clinical utility and added value of WGTS in a real-world setting in comparison to standard of care genetic diagnostics.

Patients and Methods

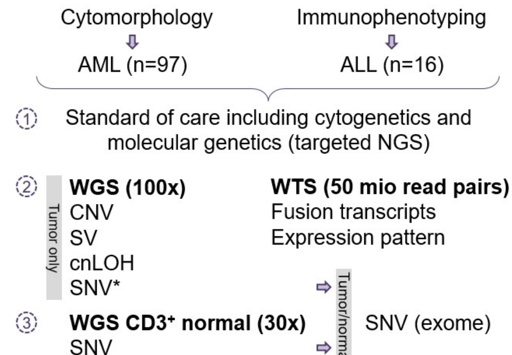
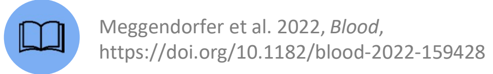


Figure 1: Workflow overview. *targeting 121 myeloid/lymphoid genes.

The study enrolled 113 patients, 97 with AML and 16 with ALL, diagnosed according to WHO criteria based on cytomorphology, immunophenotyping, cyto- and molecular genetics. WGS (100x) and WTS (50 Mio read pairs) were performed on NovaSeq instruments. For selected AML cases (n=39) CD3+ MAC sorted cells were used as normal controls (30x). Variants were called with Strelka2, Manta and GATK using a tumor w/o normal pipeline, fusions with Arriba, STAR-Fusion and Manta (Fig. 1).

Results

All patients were analyzed for structural variants (SV), copy number variations (CNV) and copy-neutral loss of heterozygosity (cnLOH) based on WGS data and the presence of fusion transcripts and expression patterns based on WTS data. The study showed high concordance between the gold standard techniques and WGTS. WGTS provided additional information for a high number of patients due to the higher resolution and unrestricted analytical scope (Fig. 2 & Fig. 3).


Meggendorfer et al. 2022, *Blood*, https://doi.org/10.1182/blood-2022-159428

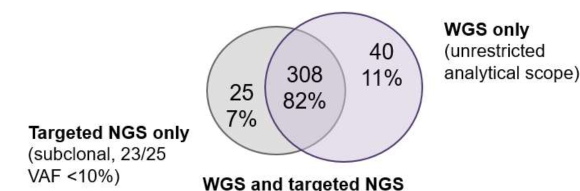


Figure 2: Venn diagram illustrating overall concordance of SNV detection for targeted panel NGS and WGS.

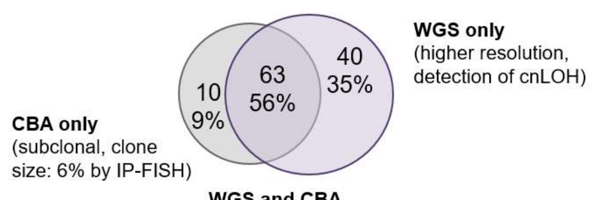


Figure 3: Venn diagram illustrating overall concordance by number of patients with detected abnormalities for CBA and WGS.

The use of normal controls

As discussed on page 12, proper use of a matched normal sample can dramatically reduce the number of variants identified by filtering out not only common but also individual SNPs and sequence variations (Fig. 4).

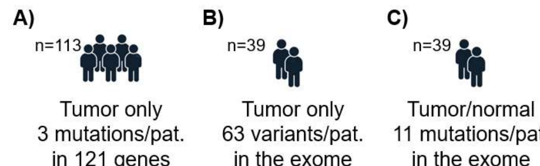
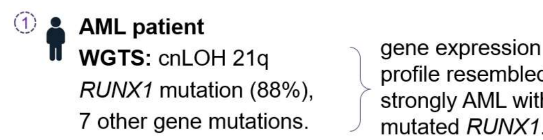


Figure 4: Mean of detected variants compared between (A) the targeted tumor only approach (121 genes), (B) exome tumor only approach, and (C) exome tumor/normal approach.

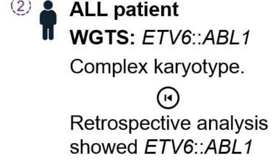
However, it should be kept in mind that somatic mutations in *IDH1/2*, *DNMT3A* and *TP53*, among others, can also be detected in non-leukemic immune cell types and therefore CD3+ cells should be used with caution as normal samples.

Case Reports

WGTS additionally identified alterations of unknown significance but some were of prognostic and/or therapeutic importance. Some patients are presented here:



Detecting the unusual without compromising diagnostic accuracy – WGTS for hematological malignancies

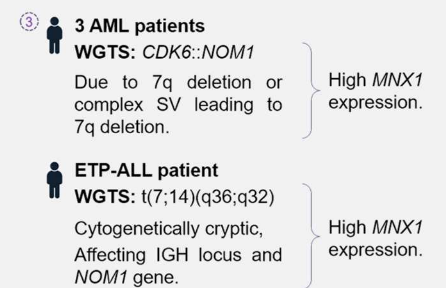


Case2

WGS also identified a cytogenetically cryptic SV leading to an *ETV6*::*ABL1* rearrangement in an ALL patient with complex karyotype (poor prognosis), which might be treated with tyrosine kinase inhibitors. Retrospective analysis showed that the aberration had already been present when the patient presented with an MPN and a normal karyotype (CBA) two years earlier.

Case 3

Of note, in 3 AML patients a 7q deletion or complex SV leading to a 7q deletion resulted in a *CDK6*::*NOM1* fusion transcript and high *MNX1* expression. In addition, WGS revealed in one patient with ETP-ALL a cytogenetically cryptic translocation t(7;14)(q36;q32) affecting the IGH locus and the *NOM1* gene, accompanied by high *MNX1* expression. Elevated *MNX1* expression as a result of t(7;12)(q36;p13) is a very rare event in pediatric AML, but our findings might suggest a more important role of *MNX1* expression in adult acute leukemia than previously known.



Case 4

Moreover, in some cases WGTS provided information about genetic alterations which led to the refinement of the diagnosis. For example, a 57 year-old male patient was initially diagnosed with low hypodiploid ALL based on FISH results but the diagnosis was subsequently refined to near haploid ALL. WGS data showed multiple chromosome losses and the transcriptional profile of the patient did not fit with other profiles of low hypodiploid ALL samples. Cytogenetics was not available due to insufficient in vitro proliferation.

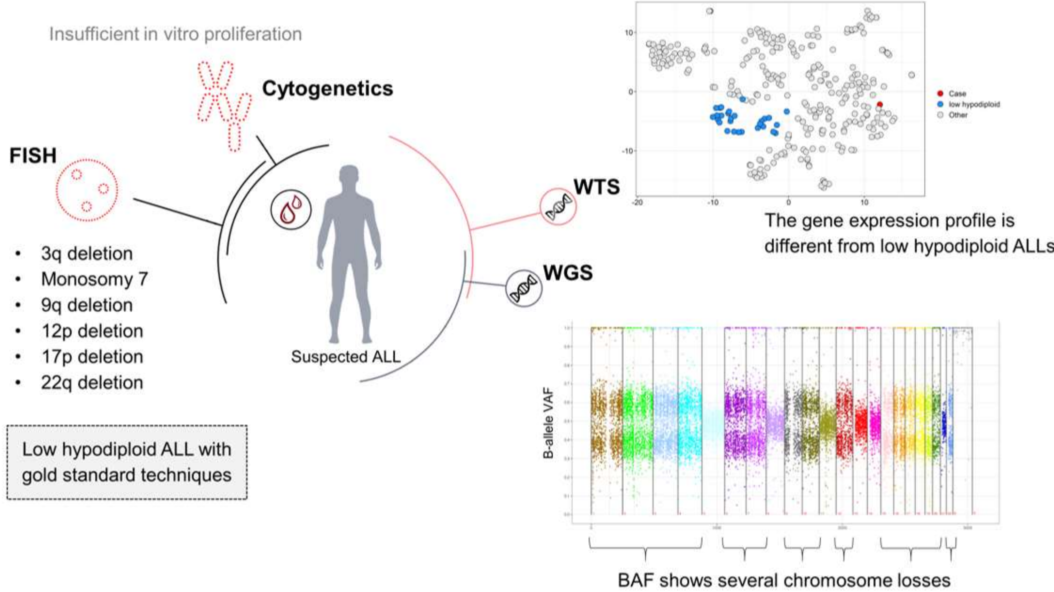
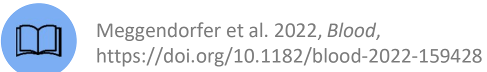


Figure 5: 57 year-old male patient with low hypodiploid ALL according to gold standard techniques which was subsequently refined to near haploid ALL based on WGTS results.

Conclusions

This prospective study in a real-world setting demonstrated the diagnostic power of WGTS with an accuracy comparable to gold standard techniques, added important prognostic information, identified potential new therapeutic targets and gave new insights into disease biology.


Meggendorfer et al. 2022, *Blood*,
https://doi.org/10.1182/blood-2022-159428

WTS for the detection of fusion transcripts in patients with hypereosinophilic syndrome

WTS is becoming increasingly important for the diagnosis of hematological malignancies. In addition to the comprehensive detection of fusion transcripts, transcriptome analysis also allows the determination of subtype-specific expression profiles to accurately stratify patients. Therefore, we have included WTS analysis in our routine workflows for two entities: acute lymphoblastic leukemias (ALL, B lineage) and myeloid/lymphoid neoplasia with eosinophilia and tyrosine kinase gene fusion (MLN-TK). For MLN-TK the diagnostic criterion is the detection of a tyrosine kinase gene fusion. *FIP1L1::PDGFRA*, *ETV6::PDGFRB*, *ZNF198::FGFR1*, and *PCM1::JAK2* are the most commonly detected fusions but other tyrosine kinase gene fusions may occur in rare cases. Due to their therapeutic relevance, reliable and comprehensive detection of tyrosine kinase gene fusions is of utmost importance to ensure the best possible treatment for each patient. If none of the common gene fusions can be detected in suspected MLN-TK cases, an extended, unrestricted screening is recommended as part of stepwise diagnostics (Fig. 1).

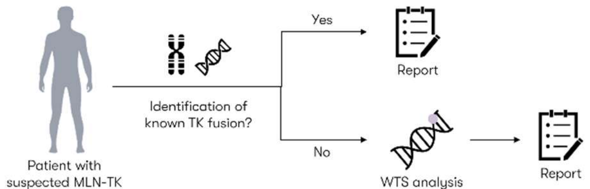


Figure 1: Recommended procedure for suspected MLN-TK cases.

Reliable detection of fusion transcripts

Transcriptome analysis offers the advantage of detecting not only known but also rare tyrosine kinase fusion transcripts. The sequence of a detected fusion can also be used to identify whether it occurred in the reading frame and thus leads to a functional tyrosine kinase, which then serves as a potential target for corresponding inhibitors. We use three different fusion callers (Arriba, STAR-Fusion, Manta) to detect fusion transcript candidates, relying only on those identified by at least two callers. Detected fusion transcripts are queried against public databases and only transcripts that have not been found in healthy individuals and are reported in public databases with relevant disease/subtype associations are reported (Fig. 2).

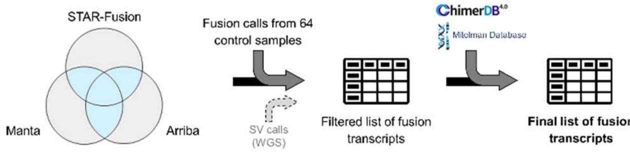


Figure 2: Workflow overview for the reliable detection of fusion transcripts with WTS.

Depending on the expression level of the fusion transcript and the obtained sequencing depth (target: 50 mio paired-reads/cluster per sample), some fusion transcripts can be challenging to detect. However, in these cases the expression values of the involved fusion genes can be used as surrogates to determine the presence of a known fusion transcript (Fig. 3). This approach obviously does not work for unknown fusion transcripts.

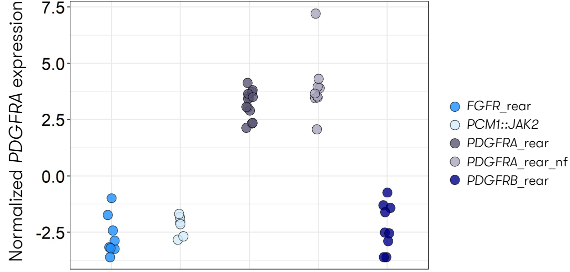


Figure 3: Normalized *PDGFRA* expression for different MLN-TK subgroups. Cases for which the *PDGFRA* fusion transcript was not detected are shown in light purple.

WTS can help to further characterize detected aberrations – a brief case report

Chromosome banding analysis of the bone marrow aspirate of a 56-year-old male patient with eosinophilia revealed a translocation involving the long arm of chr 5 and the short arm of chr 18 (46,XY[3]/46,XY,t(5;18)(q32;p11)[17]). WTS on the other hand detected a fusion transcript involving the chromosomes 5 and 14 (Fig. 4).

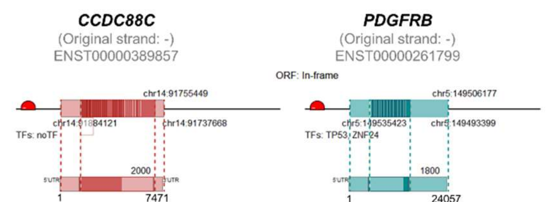


Figure 4: Schematic representation of the detected fusion transcript by WTS.

The subsequently performed chromosome painting revealed a three-way translocation (Fig. 5), which confirmed the WTS result and explained the different results of CBA and WTS.

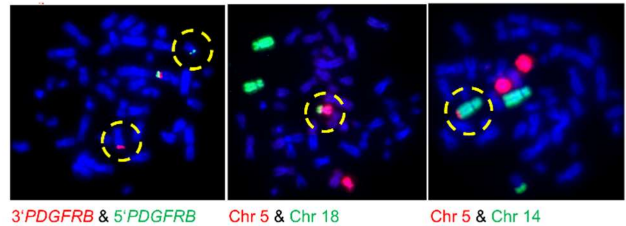


Figure 5: Chromosome painting refining the karyotype to 46,XY[3]/46,XY,t(5;18;14)(q32;p11;q32)[17]

WTS-based diagnosis of patients with acute lymphoblastic leukemia

Gene expression classifiers for the identification of B-Precursor ALL Subtypes

Considering the clinical and genetic characteristics, acute lymphoblastic leukemia (ALL) is a rather heterogeneous hematological neoplasm for which current standard diagnostics require various analyses. Classification of B-ALL into its subtypes is primarily performed genetically by determining copy number alterations of single chromosomes or whole chromosome sets and by detecting translocations and corresponding fusion transcripts. Especially in recent years, several subtypes have been identified that are characterized by specific molecular genetic markers that are best detected by larger assays such as WGS and/or WTS. Through the integrated analysis of different, large cohorts, more and more rare subtypes are detected. A recent example is the *CDX2/UBTF* subgroup, which is characterized by a distinct immunophenotype, expression profile, and high-risk features (p. 23). In recent years, several classifiers have been developed that stratify patients based solely on expression profiles. The different B-ALL subgroups are characterized by rather homogenous expression profiles that can be easily distinguished from other subgroups (Fig. 1).

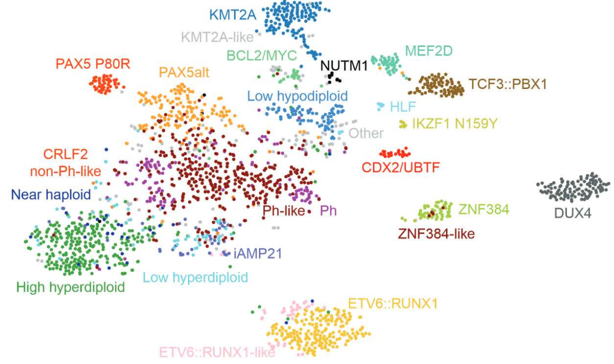
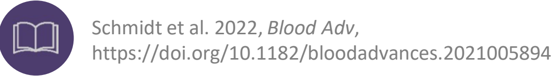


Figure 1: t-SNE plot of different B-ALL subgroups (Kimura et *al*. 2022, https://doi.org/10.1182/blood.2022015444).

For patient stratification, we apply the ALLSorts algorithm from Schmidt et *al*. 2022 that can distinguish and classify 18 different B-ALL subtypes. A small weakness of the algorithm is the reliable detection of subgroups with deviating numbers of chromosome numbers (hyper-/hypo-diploid). A new promising classifier is ALLCatchR, which can be used to assign patients to 21 established BCP-ALL subtypes with an accuracy >90%. The algorithm was developed by the University Hospital Schleswig-Holstein, Kiel, and validated on our cohort of B ALL samples. The algorithm was also presented at ASH 2022 and we plan to integrate ALLCatchR into our workflow in the first quarter of 2023.



Schmidt et al. 2022, *Blood Adv*, https://doi.org/10.1182/bloodadvances.2021005894

Reliable patient stratification by WTS, even for rare subgroups – a brief case report

Due to insufficient in vitro proliferation, chromosome banding analysis of the bone marrow aspirate of a 44-year-old male patient diagnosed with B-ALL did not yield any results. The ALL FISH panel came back negative. However, gene expression analysis assigned the case patient to the *ZEB2/CEBP* subgroup (Fig. 2).

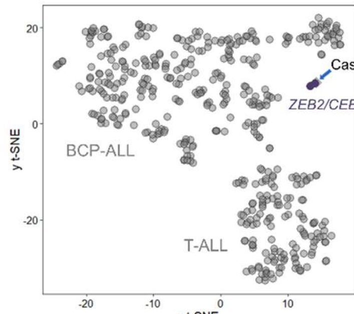


Figure 2: t-SNE plot of the gene expression profiles of B-ALL and T-ALL cases. The *ZEB2/CEBP* subgroup is highlighted in purple, the case patient clustered in close proximity.

Interestingly, mutation analysis revealed a somatic *ZEB2* variant (c.3113A>G) with high VAF (98.5%). The evaluation of the copy number profile - obtained by the spike-in CNV panel (p. 7) – revealed a CN-LOH event in the long arm of chromosome 2 (Fig. 3). *ZEB2* is located on 2q22.3, which explains the high VAF.

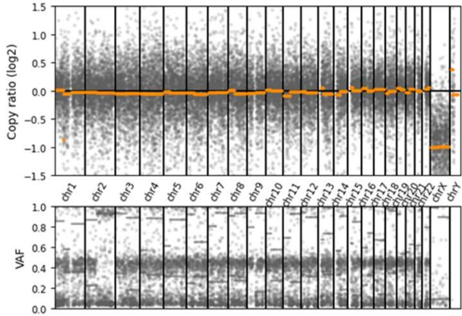


Figure 3: Copy number profile of the case patient. The 2q CN-LOH event is clearly visible in the VAF plot (bottom).

ZEB2 mutations (p.His1038Arg) have prognostic relevance, as demonstrated by Zaliova et *al*. 2021 (https://doi.org/10.3324/haematol.2020.249094).

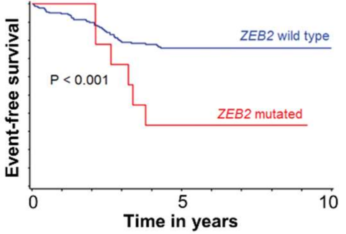
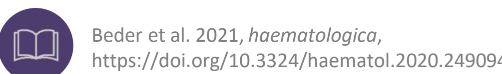


Figure 4: Prognostic impact of *ZEB2* mutations compared to *ZEB2* wild type. Plot modified of Zaliova et al. 2021.

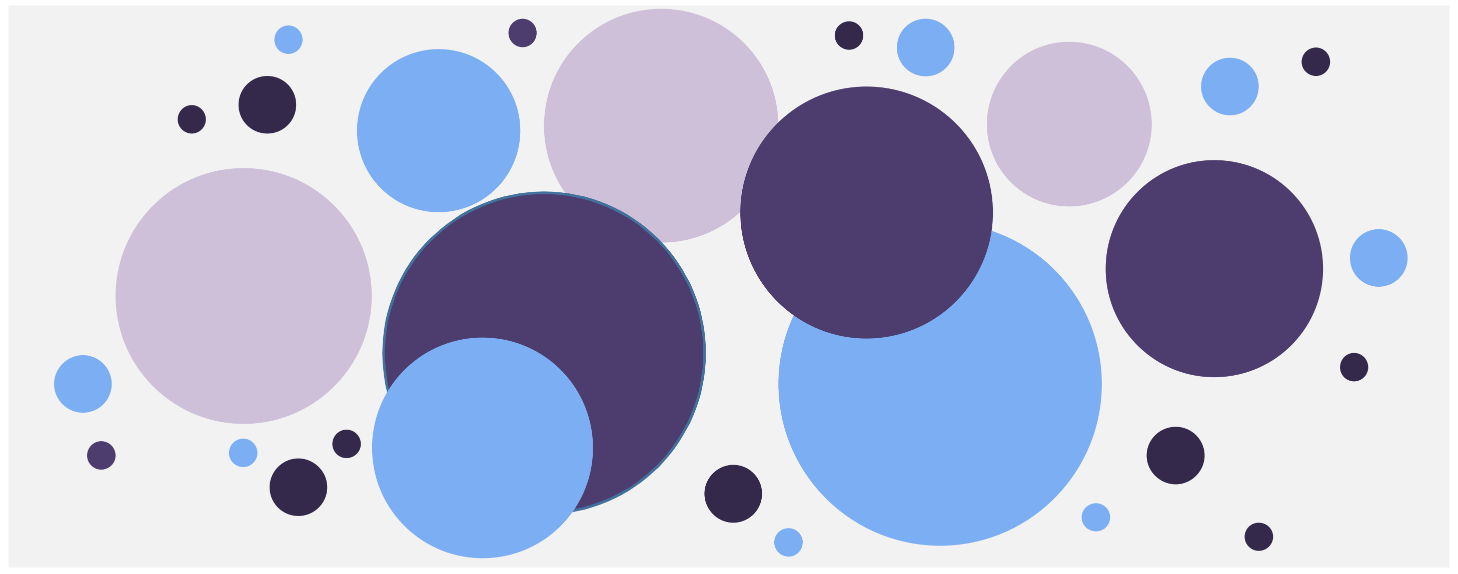
Transcriptome sequencing at MLL

WTS is listed on our request form as a diagnostic method for B-ALL as well as for MLN-TK.



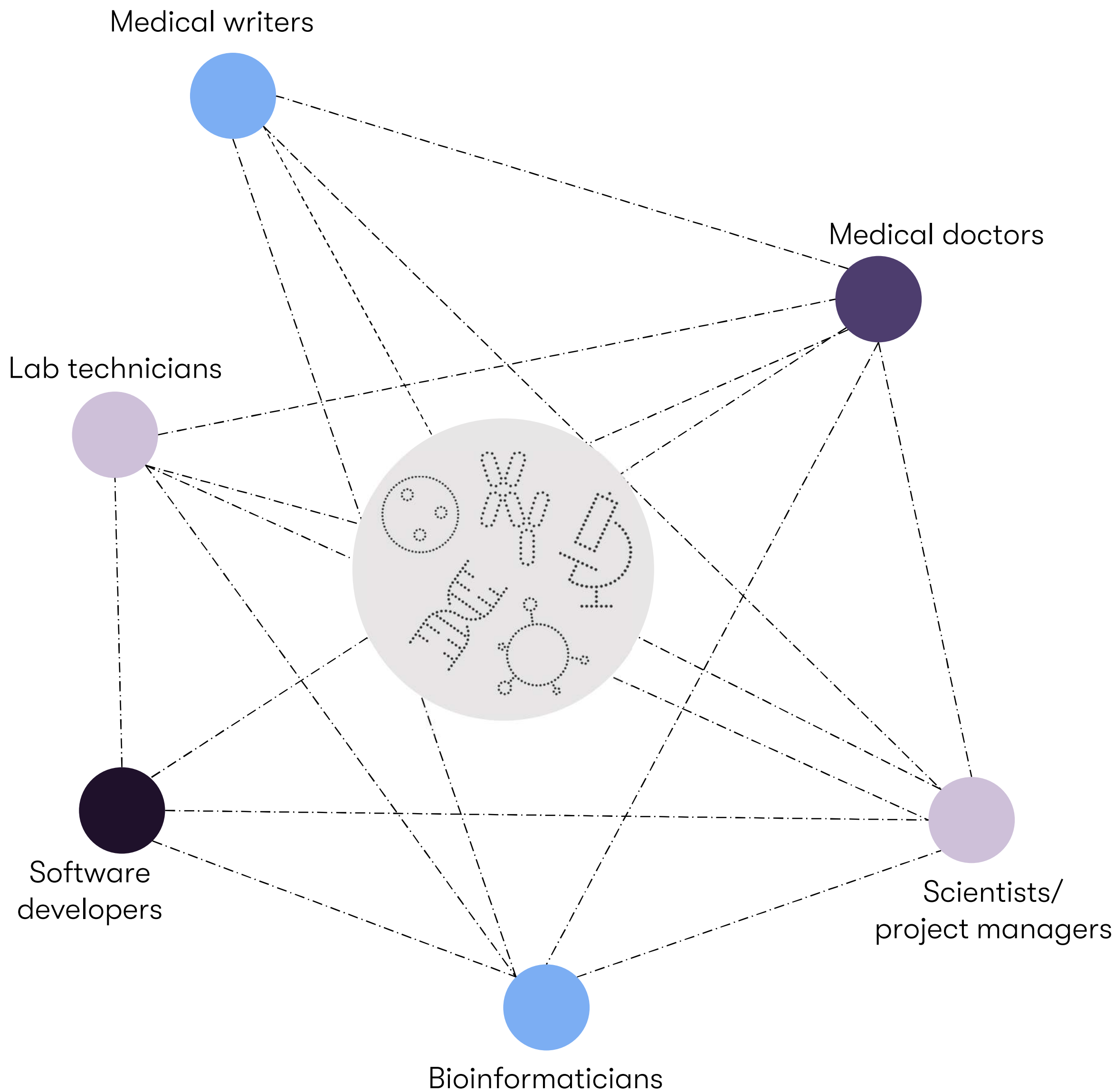
Beder et al. 2021, *haematologica*, https://doi.org/10.3324/haematol.2020.249094

Research highlights & Manuscript speed dating



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 (Gohar et al. 2019, https://doi.org/10.3389/fmed.2019.00035). 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

Research highlight: *CCL22* mutations drive natural killer cell lymphoproliferative disease

Background

Chronic lymphoproliferative disorders of natural killer cells (CLPD-NK) is a subset of large granular lymphocyte (LGL) and represents an exceptionally rare and heterogeneous indolent disorder. CLPD-NK, characterized by clonal expansion of CD3- NK cells, was included as a provisional entity in the WHO 2016 classification as a means of distinguishing it from EBV (Epstein-Barr-Virus) induced aggressive NK-LGL leukemia. Classification of CLPD-NK is mainly based on cell surface expression of CD16, CD56 and CD57, whereby the underlying genetic mechanisms are not fully understood. Mutations in the *JAK/STAT* pathway (especially *STAT3*) are found in 30% of patients, but the genetic or exogenous drivers responsible for other cases are unknown.

Patients & cohort

The CLPD-NK cohort consisted of a discovery cohort (n = 59 patients, MLL 5K) and a validation cohort (n = 62, University of Virginia). Samples from the discovery cohort were analyzed by WGTS and the identified variants were subsequently validated by targeted sequencing. For the validation cohort targeted sequencing of *STAT3* and *CCL22* was performed.

Results

Analysis of 59 CLPD-NK cases identified sequence mutations of the C-C chemokine-encoding gene *CCL22* in 16 (27%) cases (mt*CCL22*), which were mutually exclusive of *STAT3* and *STAT5B* mutations but not *TET2* mutations. WGS of 2,837 hematological malignancies, including T-LGL (n = 120) and other NK leukemias (n = 61), showed that mt*CCL22* were only found in CLPD-NK. mt*CCL22* were clustered at the conserved leucine 45, proline 46 and proline 79 residues (Fig. 1) and were clonal. The finding was confirmed in an independent validation cohort.

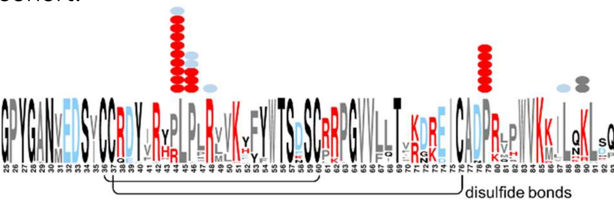


Figure 1: Sequence logo indicates conservation of CCL22 protein. Amino acid 36-92 display the CC chemokine structure, disulfide bonds are shown in black. Circles represent mutations: red - missense; blue - indel; gray - frameshift.

Sequencing of purified CD3+ T and CD56+ NK cells showed that the mt*CCL22* were somatic mutations acquired by the CD56+ NK population. No unique clinical characteristics could be identified for the different mt*CCL22* hotspots.

Differential expression analysis

WTS analysis of the discovery cohort identified 110 up-regulated and 197 down-regulated genes, enriched in cytokines and chemokines important for cell-to-cell communication, immune cell maturation, and transcription factors. The immunophenotype and the gene expression profile were similar to CD56^{bright} normal NK cells.

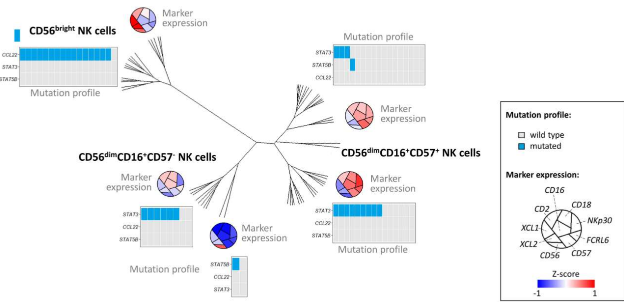


Figure 2: Clustering of CLPD-NK samples with indicated marker expression.

Deregulation of microenvironmental crosstalk

In contrast to other leukemias, *CCL22* mutations do not exert leukemogenic effects through direct cell instrinsic or autocrine effects. Data obtained by performed in vivo studies suggests that sequential induction of mutually tropic NK and hematopoietic microenvironment cells is required to induce CLPD-NK with mt*CCL22* (Fig. 3).

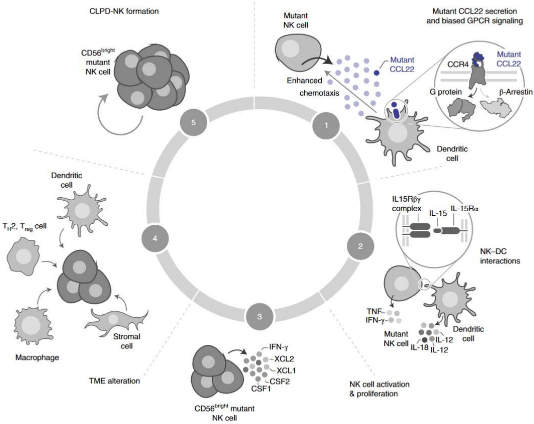
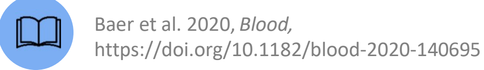
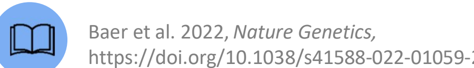


Figure 3: Schema of *CCL22*-mutant CLPD-NK leukemogenesis (Baer et al. 2022).

Conclusion

We have shown that somatic mt*CCL22* define a distinct CLPD-NK subgroup that resembles CD56^{bright} normal NK cells and is driven by deregulated cross-talk between mt*CCL22* NK cells and the NK-tropic hematopoietic microenvironment.


Baer et al. 2020, *Blood*, https://doi.org/10.1182/blood-2020-140695


Baer et al. 2022, *Nature Genetics*, https://doi.org/10.1038/s41588-022-01059-2

Manuscript speed dating – get the gist of a manuscript in less than 3 minutes

Aberrant somatic hypermutation of *CCND1* generates non-coding drivers of mantle cell lymphomagenesis

Introduction

Mantle cell lymphoma (MCL) is an aggressive B-cell neoplasm genetically characterized by the translocation t(11;14)(q13;q32), leading to *CCND1* overexpression. Aberrant somatic hypermutation (aSHM) can target proto-oncogenes and drive oncogenesis. One example is the widespread occurrence of *CCND1* mutations in MCL patients, although the biological significance of these mutations is not yet fully understood.

Patients & Methods

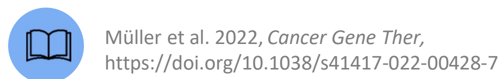
Analysis of the mutational landscape and gene expression profiles of 84 MCL patients by WGTS.

Results

- *CCND1* mutations are found mainly in samples without *ATM* mutations
- *CCND1* non-coding variants are enriched in a MCL specific manner in transcription factor-binding sites
- non-coding variants in *CCND1* are associated with increased *CCND1* mRNA expression
- coding variants in the first exon of *CCND1* in MCL are more often synonymous/benign than in other types of lymphomas carrying a t(11;14) translocation

Conclusion

The increased frequency of somatic variants due to aSHM could be a consequence of selection pressure manifested at the transcriptional level rather than a pure mechanistic consequence of aberrant activation-induced cytidine deaminase activity.



AML, NOS and AML-MRC as defined by multilineage dysplasia share a common mutation pattern

Introduction

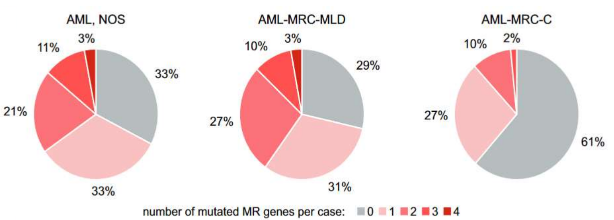
According to the WHO 2017 guidelines, AML cases without prior history of cytotoxic treatment and without recurrent genetic abnormalities are classified as either AML with myelodysplasia-related changes (AML-MRC) or AML, not otherwise specified (AML, NOS). However, a comprehensive genetic characterization of AML, NOS has not yet been performed.

Patients & Methods

Analysis of 2,188 de novo AML cases by targeted NGS. Excluding cases not classifiable according to WHO (n = 109, 5%), AML cases with recurrent genetic rearrangements (n = 273, 12%), with *NPM1* or bi*CEBPA* mutations (n = 476, 22%), and with other rare rearrangements (n = 90, 4%).

Results

- 70% of AML, NOS show a normal karyotype
- Trisomy 8 was the most frequent chromosomal aberration in AML, NOS cases with aberrant karyotype
- 67% of AML, NOS cases showed at least one MR mutation (*SRSF2*, *SF3B1*, *U2AF1*, *ZRSR2*, *ASXL1*, *EZH2*, *BCOR*, or *STAG2*)
- *RUNX1* mutations were positively correlated with the presence of MR mutations in AML, NOS



Conclusion

Patients with MR mutations constitute a substantial proportion of AML, NOS or AML-MRC-MLD cases and might be classified as a new subgroup of AML-MRC, likely reflecting the distinct biology of AML entities more accurately.

Fuhrmann et al. 2022, *Leukemia*,
https://doi.org/10.1038/s41375-022-01631-z

Manuscript speed dating – get the gist of a manuscript in less than 3 minutes

Enhancer retargeting of *CDX2* and *UBTF::ATXN7L3* define a subtype of high-risk B-progenitor ALL

Introduction

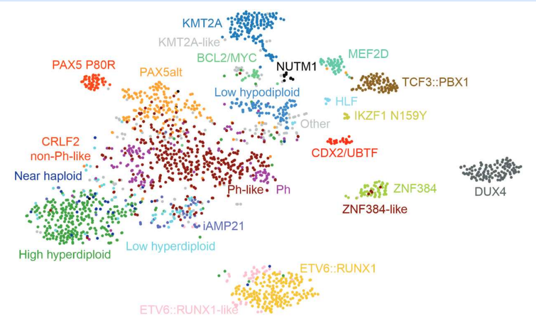
Transcriptome sequencing has identified multiple clinically relevant B-ALL subtypes but a minority of cases lack a known genetic driver, despite showing a distinct gene expression profile suggesting an unidentified, subtype-defining driver lesion.

Patients & Methods

Transcriptome and genome analysis of 3,221 newly diagnosed and 177 relapsed B-ALL cases by WGTS.

Results

- Identification of 17 newly diagnosed and 5 relapsed B-ALL cases with a distinct gene expression profile and two unique genomic alterations
- Immunophenotype was characterized by CD10 negativity and IgM positivity
- The patients were mostly adolescents and young adults (AYA) and female
- Gene expression analysis identified 1,216 differentially expressed genes, including up-regulation of *CDX2* and *NTRK3* and down-regulation of *FLT3* and *MME* (CD10)
- A genomic deletion of *UBTF* of ~10kb resulted in a chimeric fusion involving *UBTF* and *ATXN7L3*
- Somatic 13q12.2 deletions spanning *FLT3* promoter led to upregulation of *CDX2* through a mechanism of enhancer retargeting



Conclusion

CDX2 deregulation and *UBTF* fusion define a B-ALL subtype with distinct immunophenotype, expression profile, and high-risk feature.

Kimura et al. 2022, *Blood*, https://doi.org/10.1182/blood.2022015444

How T-lymphoblastic leukemia can be classified based on genetics using standard diagnostic techniques enhanced by whole genome sequencing

Introduction

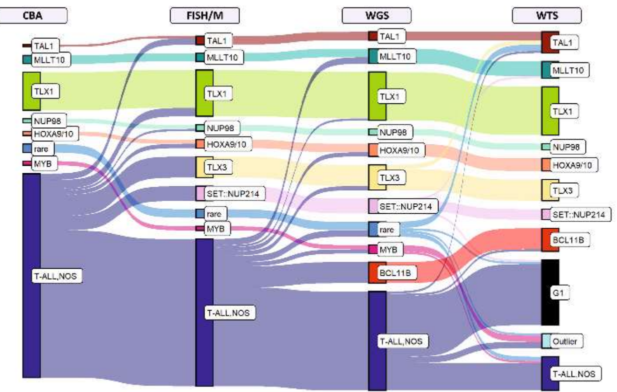
In contrast to the introduction of multiple genetically defined BCP-ALL entities in the 5th edition of the WHO classification, no genetically defined entity has been introduced in T-lymphoblastic leukemia (T-ALL). However, to move towards personalized medicine a distinct classification based on biomarkers assessable by routine diagnostic methods is essential.

Patients & Methods

Genetic analysis (CBA, FISH) supplemented by WGTS of 131 T-ALL cases.

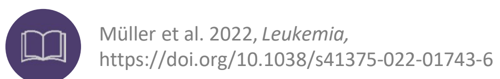
Results

- A combination of CBA, FISH and RT-PCR assigned 60/131 (45,8%) T-ALL cases to 9 genetically defined subgroups
- In roughly 20% (26/131) cases CBA was not sufficient for subtype classification due to the cytogenetically cryptic nature of the abnormality and/or insufficient in vitro proliferation
- WGS assigned an additional 17% to distinct genetic subgroups



Conclusion

A first step towards genetic classification of T-ALL cases in a routine setting should be based on commonly available methods such as CBA, FISH, RT-PCR. In a second step, unclassified cases can be resolved by WGTS or other novel techniques.



Müller et al. 2022, *Leukemia*, https://doi.org/10.1038/s41375-022-01743-6

Manuscript speed dating – get the gist of a manuscript in less than 3 minutes

SF3B1 mutated MDS: Blast count, genetic co-abnormalities and their impact on classification and prognosis

Introduction

About half of MDS patients harbor somatic mutations in splicing pathway genes, of which *SF3B1* is the most commonly mutated gene. MDS with SF3B1 mutations has been recognized as a distinct subgroup by various guidelines. However, incidence, genomic landscape, AML transformation rate and clinical outcome have not been analyzed in detail.

Patients & Methods

WGS analysis of 734 MDS cases. Therapy-related MDS were excluded.

Results

- *SF3B1* mutations (*SF3B1*mut) were identified in 231/734 (31%) MDS patients and were mainly found in MDS-RS
- *SF3B1*mut were associated with better OS, except in MDS -5q cases, where they were associated with a shorter OS
- *SF3B1* cases fulfilling the criteria of the *SF3B1* entity (IWG-PM, *SF3B1*ent) had a longer OS compared to *SF3B1*mut that did not fall into this category (*SF3B1*nent)
- 69/231 (30%) *SF3B1*mut samples showed aberrant karyotypes
- 47% (67/144) *SF3B1*ent cases did not harbor any additional mutation in 73 analyzed genes, while 53% (77/144) harbored 1-4 additional mutations
- in multivariate analysis only *RUNX1*mut and del(5q) were independent prognostic factors
- 7% of *SF3B1*mut cases progressed to AML compared to 15% of *SF3B1* wild-type patients, with a higher AML transformation rate in *SF3B1*nent

Conclusion

Our data suggest that the identification of the good prognostic subset within *SF3B1*mut patients can be achieved by excluding only cases with del(5q) and/or *RUNX1* mutations, independent of blast count.

Huber et al. 2022, *Leukemia*,
https://doi.org/10.1038/s41375-022-01728-5

SF3B1 mutations in AML are strongly associated with MECOM rear and may be indicative of an MDS pre-phase

Introduction

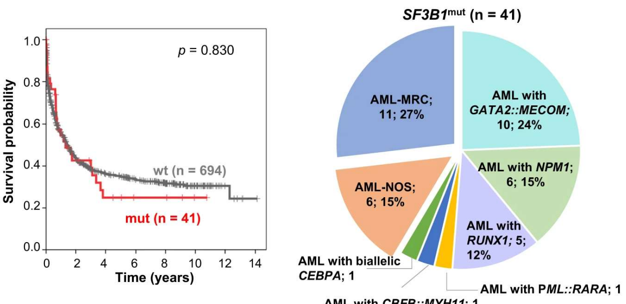
The presence of *SF3B1* mutations is considered as diagnostic criteria for AML-MR according to the 5th edition of the WHO classification. Here, we address the prognostic impact of *SF3B1* mutations in AML and evaluate the genetic landscape of *SF3B1* mutated patients at AML diagnosis and during follow-up.

Patients & Methods

WGS analysis of 735 AML cases (89% de novo AML, 11% s-AML). Therapy-related AML were excluded.

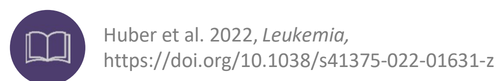
Results

- 6% (41/735) of AML patients harbored a somatic *SF3B1* mutation (*SF3B1*mut)
- *SF3B1*mut were detected in 8 different AML entities, of which AML with *GATA2::MECOM* was the most frequent one (10/36, 28%)
- *SF3B1*mut did not affect OS in the total AML cohort
- *SF3B1*mut patients harbored 3.3 mutations, with *RUNX1* and *NRAS* being the most frequent ones
- in 40/41 (98%) of *SF3B1*mut cases similar or higher *SF3B1* VAFs were observed compared to other co-mutations or aberrations



Conclusion

SF3B1 mutations are enriched in poor risk AML, are strongly associated with MECOM rearrangements and myelodysplasia-related changes and are most likely acquired early in a pre-leukemic clone, which may be indicative of an MDS pre-phase.



Huber et al. 2022, *Leukemia*,
https://doi.org/10.1038/s41375-022-01631-z

Manuscript speed dating – get the gist of a manuscript in less than 3 minutes

Identification of a specific immunophenotype associated with a consistent pattern of genetic mutations including *SRFS2* and gene expression profile in MDS

Introduction

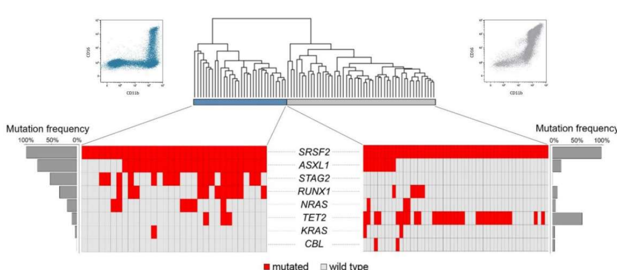
MDS comprise a heterogeneous group of diseases and the identification of homogeneous subgroups is desirable for improved prognostication and treatment. We identified a specific CD11b/CD16 expression pattern in granulocytes and reduced CD45 expression in myeloid progenitor cells (MPC) in a subgroup of MDS patients which was associated with *SRSF2* mutations.

Patients & Methods

WGTS analysis of 83 MDS cases: 32 *SRSF2*mut with the specific immunophenotype and 51 *SRSF2*mut without.

Results

- Unsupervised clustering of the transcriptomic profiles revealed 2 distinct clusters, consistent with the defined immunophenotypic groups
- 116 differently expressed genes were identified, including 13 transcription factors, 4 FDA approved and 6 potential drug targets
- The gene signature was significantly enriched for cellular developmental process, cell adhesion and extracellular matrix organization
- >50% of cases with the specific immunophenotype were characterized by *SRSF2/STAG2* co-mutations and *STAG2*mut was only found in this group



Conclusion

SRFS2 and *STAG2* mutations and the specific immunophenotype should be further addressed, as both mutations are associated with poor outcome and could be potential targets for therapy in MDS.

Weiß et al. 2022, *Cytometry B Clin Cytom*, https://doi.org/10.1002/cyto.b.22057

Mutations in spliceosome genes in myelodysplastic neoplasms and their association to ring sideroblasts

Introduction

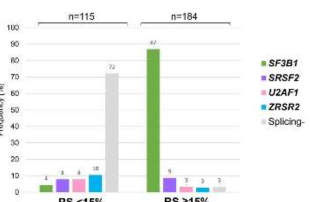
MDS with low blasts and *SF3B1* mutation (MDS-*SF3B1*) is considered a separate entity in both WHO 2022 and ICC classification guidelines. The entity is defined by the presence of an *SF3B1* mutation, a bone marrow blast count <5% and the absence of certain cytogenetic abnormalities and biallelic *TP53* inactivations.

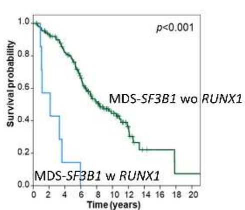
Patients & Methods

WGS analysis of 704 de novo MDS cases and mutational profiling of 1804 de novo MDS patients by targeted panel sequencing.

Results

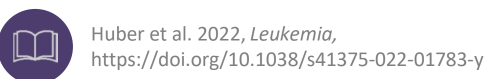
- 40% (262/660) of MDS cases with data on the presence of ring sideroblasts (RS, 660/704) showed RS ≥ 15%
- 299/660 patients had low blasts (LB)
- Splicing mutations were detected in 97% of LB cases with RS ≥ 15% compared to only 28% in LB and RS < 15%
- 22 cases qualified as "MDS with low blasts and ring sideroblasts" (MDS-LB-RS, WHO 2022)
- OS was significantly shorter in MDS-LB-RS vs MDS-*SF3B1* but comparable to other MDS-LB
- *RUNX1*mut cases were associated with shorter OS compared to *RUNX1* wild-type cases





Conclusion

*SF3B1*mut are associated with LB and increased RS, whereas other splicing mutations showed the reverse trend. MDS-LB-RS and wild-type *SF3B1* are suggested to be classified as MDS-LB. *RUNX1*mut cases should be excluded from the MDS-*SF3B1* entity due to their negative prognostic impact.


Huber et al. 2022, *Leukemia*, https://doi.org/10.1038/s41375-022-01783-y

Combining forces for the development of innovative methods to advance patient care – MLL & INTERCEPT-MDS

MLL maintains extensive international collaborations to advance scientific research in hematologic neoplasms. Since last year, MLL also participates in several EU funded research projects: GenoMed4All and INTERCEPT-MDS, which have already been briefly presented in the last Research Report. As part of the Intercept-MDS program, MLL is also mentoring a PhD student who is working on the analysis of WGTS data from MDS patients from the MLL 5K project. In the course of this, she has also studied the clinicogenomic characteristics of somatic and germline *UBA1* variants and discovered five additional *UBA1* variants as potential novel causes of VEXAS syndrome (p. 10). For this work she received the best poster award during the Summer School 2022 (see paragraph below).

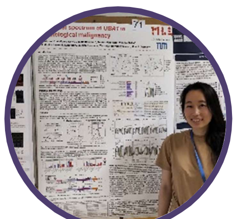


Figure 1: Maki Sakuma presented her work on *UBA1* variants in the context of VEXAS syndrome during the Summer School 2022 in Spetses. One of her key findings is the strong gender-bias of *UBA1* variants in myeloid diseases.

Mid-term meeting 2022

In the first week of May, MLL hosted the mid-term meeting of the innovative training network INTERCEPT-MDS. The meeting was organized by Wencke Walter with strong support of the administration office and the communication department. The network includes 11 PhD students and 13 project leaders from ten different countries (Fig. 2).

Figure 2: Overview of the institutes participating in the INTERCEPT-MDS program.

The meeting was held in a hybrid format to allow all project leaders to participate. INTERCEPT-MDS is financed by the European Union's Horizon 2020 research and innovation program and relies on a multidisciplinary approach to identify biomarkers and molecular processes using "omics" technologies on the individual cell level to enable early detection and possible treatment of clonal myeloid diseases. The huge amount of data generated requires thorough bioinformatics and statistical analysis.

In order to optimally prepare the students for their upcoming work, they were taught bioinformatics and statistical basics by Dr. Ana Conesa of the Institute for Integrative Systems Biology (I2SysBio) in Valencia in a three-day workshop, which specifically apply to transcriptome and gene expression analysis.

Figure 3: Dr. Carsten Marr during the mid-term meeting.

In various scientific lectures, the students also gained insights into the clinical application of "multi-omics" methods, big data analyses, and artificial intelligence. Prof. Torsten Haferlach spoke about the application of artificial intelligence in hematological diagnostics and the integration of such models into MLL workflows. Dr. Carsten Marr, group leader at the Institute for Computational Biology and director at the Institute for AI in Healthcare, provided exciting and detailed insights into the application of machine learning methods for classifying myeloid diseases (Fig. 3). Prof. Claudia Haferlach brought the series of lectures to a satisfying conclusion with her presentation on clinical diagnostics for MDS.

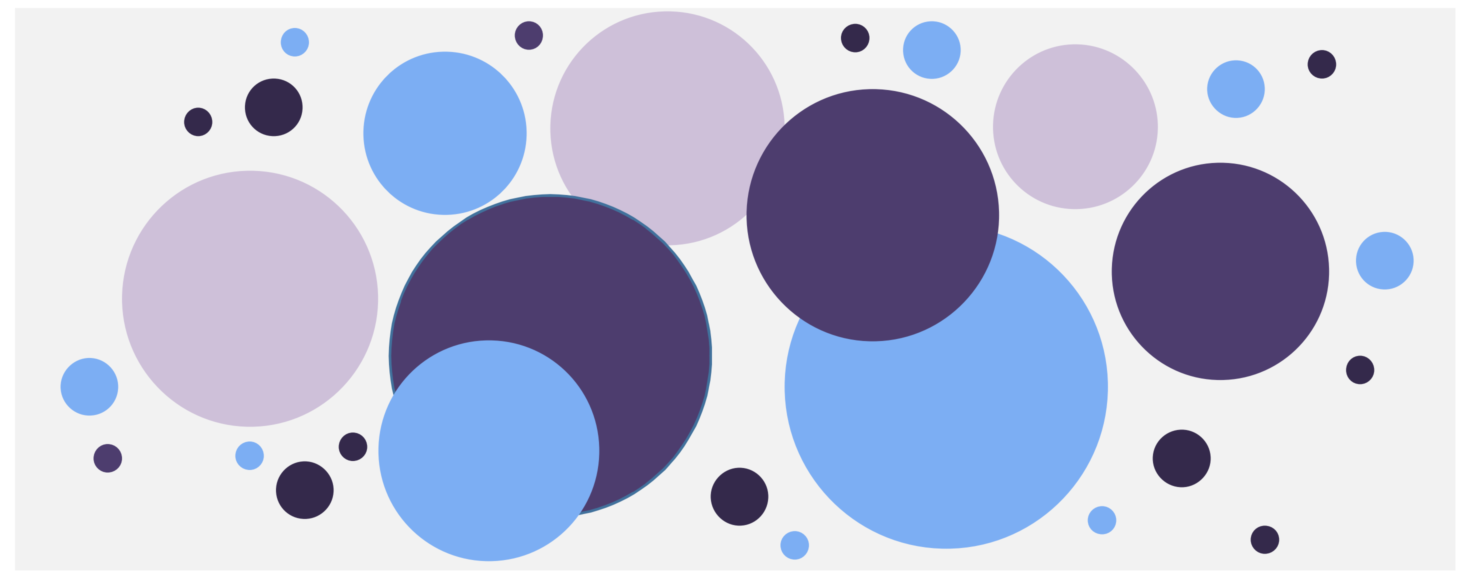
Spetses Summer School 2022

In Sept 2022, the INTERCEPT-MDS network reunited for the Spetses Summer School. Here, the students presented their preliminary results during two poster sessions and got the chance to directly interact with different supervisors and experts. In one of the lectures the MLL presented the current work on WGTS for leukemia diagnostics. The highlight of the program were two keynote lectures of Edith Heard on X-chromosome inactivation and Douglas Higgs describing the cell model of hematopoiesis and erythropoiesis. All the lectures took place in the main Hall of A.K.S.S, where James Watson and Francis Crick presented their discovery of the DNA double helix structure in 1966.

Outlook

Over the next two years the MLL will host other students from the program who will either work on the in house WGTS data or get an insight into different workflows, like single-cell sequencing (p. 32). Moreover, the MLL will support the network by performing library preparations and sequencing of single-cell libraries.

Brief ASH 2022 recap

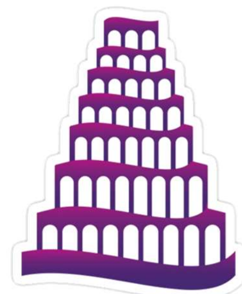


Genetics as a mainstay of new classifications and to improve prognosis estimation

CLASSIFICATION

Classification according to WHO 2022 and ICC: Do we invent a Babylonian confusion of languages?

The parallel publication of two new classifications has led to some confusion, making diagnosis more difficult for physicians and patients. We evaluated the impact of the 5th WHO 2022 guideline on the classification of AML and MDS patients and the differences in disease categorization compared to ICC. As presented by Gregor Hörmann, the new classifications consistently follow the idea of a more genetics-based definition but a small number of patients (~ 1%) will be differently classified as AML or MDS. According to WHO and ICC, cytomorphology – a very subjective method - is irreplaceable for MDS diagnostics. However, Prof. Claudia Haferlach showed that based on karyotype and mutation status of ~ 30 genes we identified 9 biologically distinct subgroups, reflecting biology better than blast counts. Hence, a purely genetic classification of MDS is feasible.



PROGNOSTICATION

Genetics as a mainstay to improve prognostic estimation

This year's publication of the IPSS-M (p. 8) represents an important innovation in the prognostic assessment of MDS. In her talk, Constance Bär shared data on an independent validation of the predictive superiority of the IPSS-M over the IPSS-R and also compared the IPSS-M to two other personalized prognostic models that incorporate molecular genetic factors. However, the three models did not only differ in their feature set but also showed different strength to predict OS and LFS. Both the WHO 2022 classification of MDS and molecular prognostic scores for MDS take *TP53* alterations into account. In a cohort of 1,520 patients with MDS and AML, Stengel et al. identified *TP53* double-hit as the most important prognostic factor, whereby the incidence of *TP53* double-hit seems to be dependent on the proportion of blasts. The work was presented by Manja Meggendorfer in a special session on *TP53* alterations.

A research potpourri – some poster highlights from ASH 2022

In addition to the 8 lectures, we presented our scientific work in 14 posters covering a broad spectrum from prospective WGTS studies to IGH rearrangement calling and various mutational analyses in different myeloid and lymphoid entities. Piers Blombery, a guest visiting scientist from Australia, presented his work on novel non-coding, coding and structural variants in hairy cell leukemia and - together with Anna Maierhofer from the MLL – novel germline multi-exon deletions in patients with *DDX41* familial predisposition to myeloid malignancies. Anna Stengel shed more light on the genetic landscape of AML and MDS, focusing on *IDH1/IDH2* and *RUNX1* mutations. Isolde Summerer identified a rare myeloid subtype of MYC-positive double minutes with an APL-like cytomorphology. Heiko Müller presented his work on IGH rearrangement calling from WGTS data and could also show that *KMT2A*::*AFF1* rearrangements lock the cell of origin at a point of ongoing V(D)J recombination. Martha-Lena Müller presented data on the evaluation of cyto-plasmic lineage-associated markers in MPAL as devised by WHO 2022 and argued in favor of considering multiple markers instead of single markers. Elisabeth Weiß could show that a reduced core set of markers is already sufficient for MDS diagnosis by flow cytometry without loss of accuracy. Constance Bär shared her work on the genetic landscape of ICUS/CCUS and compared the evolution pattern to MDS.

The application of machine learning in hematology diagnostics

Due to the increasing acceptance of ML-based algorithms and models in hematology, there were two sessions with presentations on emerging tools, techniques and artificial intelligence in hematology at ASH 2022. Prof. Torsten Haferlach presented an update on the implementation of a ML-based algorithm for cell type classification of blood smears. ML-based algorithms are now also being used in molecular biology to interpret large data sets such as those obtained from WGTS. Niroshan Nadarajah presented one such algorithm that was developed in collaboration with AWS.

Machine learning algorithm correctly identifies 95% of cells in differential count of blood smears: A prospective study on >29,000 cases and >17 million single cells

The aim of this study was to test the value of AI driven differential counts in a specialized hematology laboratory by comparison to humans and create a cloud based web platform. The head-to-head study revealed that the classifier performed comparable to routine methods with the added benefit of total reproducibility, regardless of the experience of a single diagnostician. Additionally, the automated upload and computation time for 500 images is <30 sec and multiple samples can run in parallel and hence, reduce the turn-around time.

Evaluation of a transparent Artificial intelligence (AI) disease classification system with Whole genome sequencing (WGS) and whole transcriptome sequencing (WTS) data in a prospective study with 325 cases

The aim of this study was to validate an AI algorithm predicting the disease entity based on WGS and WTS data only, while depicting relevant features for a decision and thus making its results transparent and verifiable by humans. The results of the head-to-head study showed a high concordance (81%) between the model and human experts, although for overlapping disease entities continued calibration is needed to improve accuracy. The talk was also highlighted during “Best of ASH”.

AI and machine learning: a new frontier in hematology

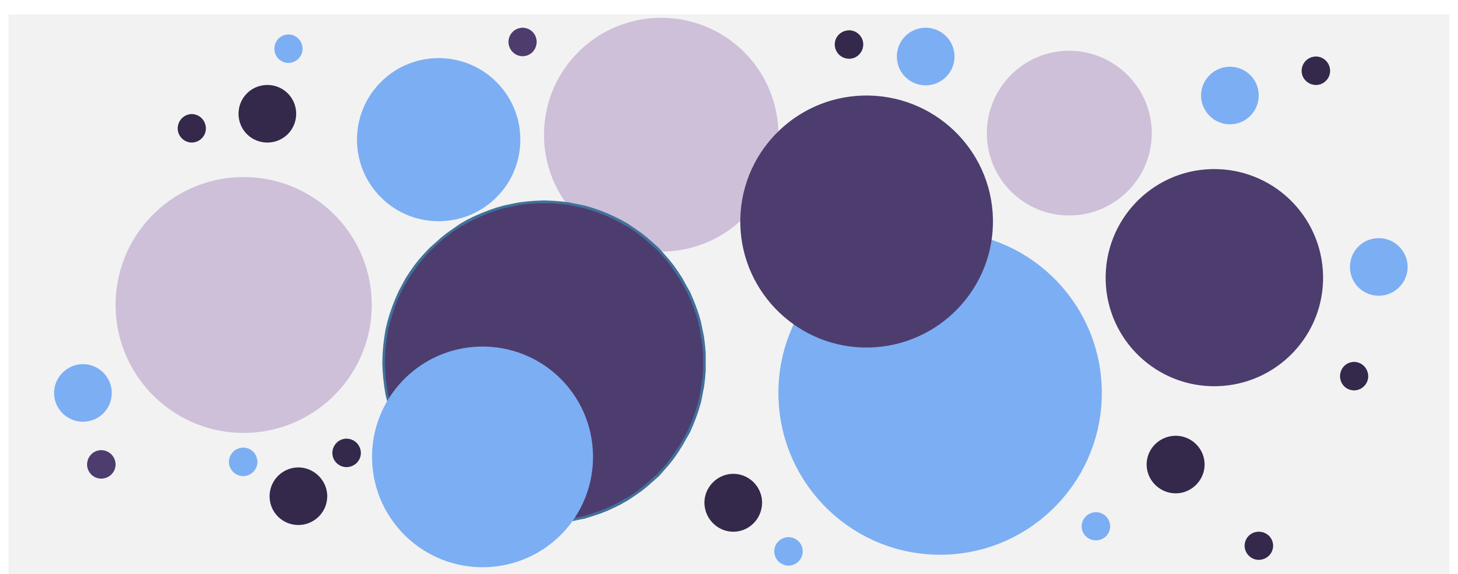
The success and acceptance of AI-based technologies in hematology will also depend on establishing regulations and guidelines that delineate and foster the step-wise integration of ML models in the clinic to ensure that patients are neither barred from the access to valuable innovations nor exposed to erroneous devices. Here, comprehensive training is necessary to equip the next generation of hematologists with the necessary background knowledge to thoroughly judge new developments. In a special session, four 20-minute presentations covered various aspects of machine learning and the potential application of such models in clinical diagnostics. In addition, the presentations also highlighted funding sources, and initiated a critical discussion regarding data security and ethical challenges.

Shannon McWeeney, Manja Meggendorfer, Ilana Goldberg, Aziz Nazha

Haferlach et al. 2022, *Blood*, https://doi.org/10.1182/blood-2022-165863

Nadarajah et al. 2022, *Blood*, https://doi.org/10.1182/blood-2022-169093

Future projects & outlook



Genome network hematology - the platform for the generation and analysis of a patient's molecular profiles

Mission

We put the latest findings in the field of hematological neoplasms into practice. In this way, we enable all patients to have access to optimal diagnostics and the best possible therapy. This can best be realized through a network of clinical care and diagnostics experts, patient organizations, and the pharmaceutical industry, interested in research.

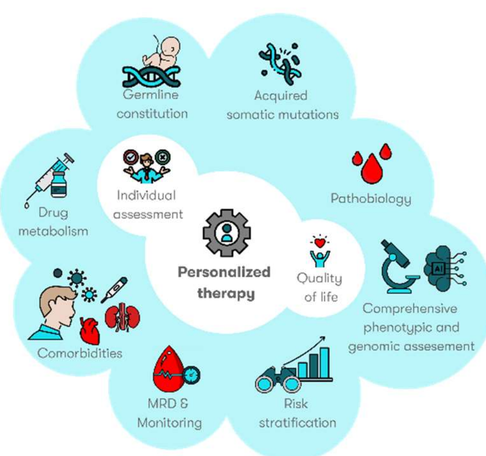
Vision

Personalized medicine will soon become reality. Evidence-based algorithms are used to perform cost-effective step-wise diagnostics that enable selection of the optimal therapy tailored to the individual patient. Therapy results are reviewed periodically and adjusted as needed. The goal of this coordinated concept is to improve the chances of recovery.

Goals

Establishing knowledge-generating patient care: Data from research projects and patient care are brought together. Hence, new findings from daily patient care drive new research projects, the results of which in turn improve patient care.

In recent years, comprehensive assays such as WGS and WTS have gained attention and several studies have demonstrated the clinical utility and added value of these assays. However, widespread clinical acceptance has not yet been achieved. But that will soon change, as individualized molecular profiling opens up the possibility of more tailored and effective treatment - the goal of precision medicine.

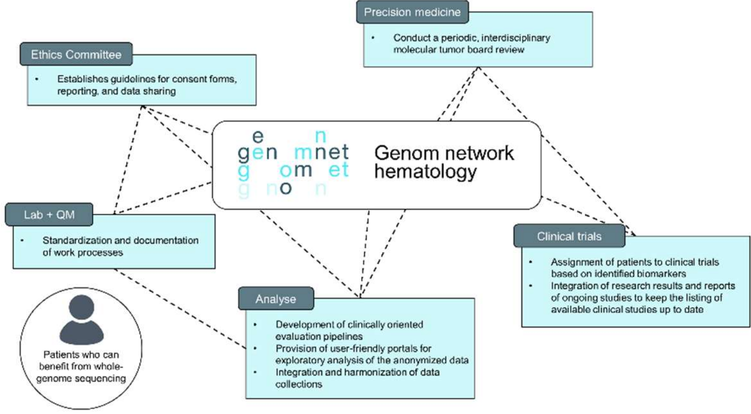


Personalized medicine – the dream of every physician

Personalized medicine is the desired goal for an individualized therapy with fewer side effects and higher chances of success. This has already been implemented for individual cases and subgroups. However, for a widespread application and implementation in clinical practice - irrespective of the disease - reliable diagnostics, comprehensive molecular characterization of the patient and access to targeted substances and the knowledge of the treating physicians and patients about the corresponding possibilities are required. WGS contributes to this goal by providing the most comprehensive and clinically relevant molecular profile of a patient with a single test. In combination with electronic health records and targeted therapies, personalized patient care becomes the new reality.

Studies, projects and molecular tumor boards

At MLL, there is a great deal of experience in the evaluation of WGTS data in hematologic neoplasms. This expertise will now be made available to others in the context of studies. The genome network hematology supports studies and projects that evaluate, for individual patients, the impact of the results of WGS/WTS analyses on potential treatment decisions. In addition, it is envisioned that WGTS results will be discussed in molecular tumor boards to select the best possible therapy for the patient.



Single-cell sequencing – insights into cell diversity, cell lineage and evolution

Single-cell sequencing captures the diversity of cell states within a (tumor) sample by simulteanously measuring thousands of individual cells. The technique includes several different assays like single-cell gene expression, single-cell ATAC, single-cell immune profiling, which can also be combined for single-cell multi-omics analysis by integration of the different data layers. As with any NGS assay, the better the quality of the input material, the better the results.

Cell sample quality assessment

In addition to the large quantity of patient material in the form of preserved cells for nucleic acid extraction, we also store patients' vital cells (cryo-preserved cells), which can be used for single-cell analyses. Hence, in the summer of 2022, we started to establish a single-cell workflow, focusing initially on the multiome assay (10xGenomics). For this assay nuclei are isolated, tagged/barcoded, and then the libraries of the celluar components of interest (DNA & RNA) are prepared for sequencing. For most steps demonstrated protocols can be used. However, the isolation of nuclei from the cryopreserved samples requires additional optimization. For the determination of cell viability, number of cells and nuclei concentration, we acquired a LUNA-FX7 instrument (logos biosystems) that allows the parallel processing of up to 8 samples (Fig1.)

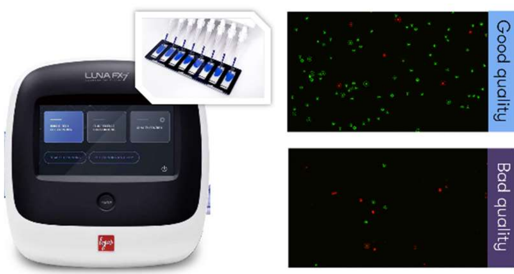


Figure 1: With the LUNA-FX7 live cells appear green fluorescence by Acridian Orange staining, whereas dead cells and isolated nuclei turn red fluorescence by Propidium Iodide uptake. A cell viability > 90% is desirable.

Library preparation with the multiome kit

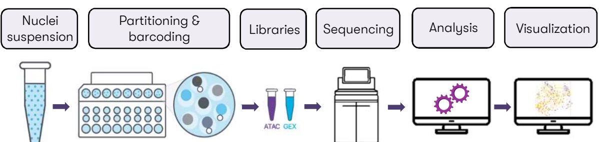
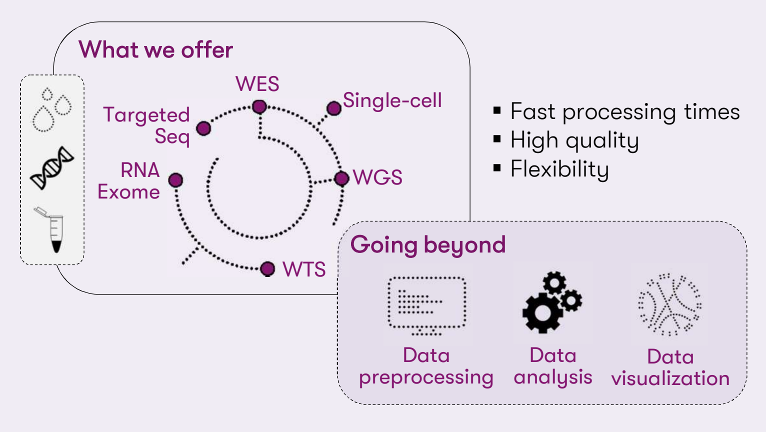


Figure 2: Schematic overview of the multiome library preparation (10xGenomics) and further processing.

In order to ensure reliable and publishable results, the input material must be of high quality. Hence, a cell viability >90% and a minimum of 1-2 mio living cells are required. In cases with low cell viability, a dead cell removal kit is applied to get rid of cell debris and dead cells, which would negatively impact the library preparation. In some cases a high number of erythrocytes are still present in the sample. The erythrocytes do not negatively affect library preparation, but result in higher sequencing costs because their mRNAs also contribute to the library and hence must also be sequenced even though they do not providing relevant information. For the multiome kit, the isolation of intact nuclei and the subsequent accurate quantification of these is key. The isolated nuclei are obtained by lysing mononuclear cells. Here, the composition of the lysis buffer and the lysis time are the key parameters to obtain intact nuclei and to avoid overlysing. Overlysis can lead to the leakage of nuclear content leading to a high level of background in the data. Once the nuclei are isolated, typical NGS library preparations are performed.

Last year we already offered sequencing of single-cell libraries. Here, the required sequencing settings depend on the chosen assay and we follow the recommendations by the respective kit manufacturer to ensure high quality sequencing data. This year, we have gone one step further and have set up a workflow that enables us to generate multiome libraries (10xGenomics) to simultaneously asses gene expression and open chromatin from the same cell.



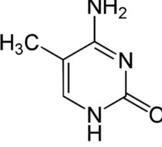
- Fast processing times
- High quality
- Flexibility

https://mllseq.com/single-cell-seq

https://mllseq.com/glossary-modal

Always forward, never backwards – a brief outlook on our planned methodological and technical developments in 2023

NovaSeq™ X Plus Sequencing System
After > 5 years of development Illumina announced the new NovaSeq X Plus system during the Innovation Roadmap session in Sep 2022. The system has up to 8/16 addressable lanes per run and with its ability to sequence more than 128 genomes (30x coverage) per run will most likely revolutionize human WGS. The system comes with an integrated secondary analysis, effectively removing the bioinformatics bottlenecks. Hence, sequencing of the human genome will cost ~200$. The XLEAP-SBS reagents can be shipped over a range of temps of -20°C to 50°C and hence dry ice is no longer needed. In 2023, we will extend our sequencing facility by adding a NovaSeq X Plus.



Hemoglobinopathies
In the coming year, we want to expand our diagnostics and not only take care of malignant diseases, but also dive into a part of benign hematology. Our initial focus will be the haemoglobinopathies, a group of recessively inherited genetic conditions affecting the hemoglobin component of blood. Here, genetic alterations can cause protein dysfunction, resulting in hemoglobins with increased or decreased oxygen affinity or abnormal polymerization. Thalassemias are manifested by altered synthesis of globin genes, which in turn is also caused by mutations. In addition to classical PCR methods (gap-PCR and MLPA) to detect the mutations, we can also use our experience in NGS, hence we are currently designing a new panel to detect mutations in the hemoglobin genes. In addition, completely new methods will also find their way into our laboratory. Ferritin status and hemoglobin electropheresis will also become part of our daily work. Besides the diagnostic approach, there will certainly be a lot to explore here. We will report more on this in the coming year.

AML-MRD panel
Monitoring treated neoplasms plays an important role, in order to measure the success of therapy and, if necessary, to adjust therapy in time. So far, the focus has been on quantitative PCRs that can detect fusion transcripts with very high sensitivity. At the same time, however, increasing knowledge that includes other markers, such as those identified by NGS and panel diagnostics at diagnosis, is also bringing into focus a diverse set of mutations in a growing number of genes. Classical NGS analysis with de-duplicated reads bottoms out at a detection limit of 1% allelic load. The difference to sequencing background noise is blurring. Techniques such as the use of UMIs (unique molecular identifier) can increase the sensitivity. The new guidelines for monitoring of AML markers published by the ELN therefore request a sensitivity of 0.1% allelic load. Hence, we will set up a new sensitive AML-MRD panel with UMIs or even duplex sequencing technology to tackle the challenge. We will keep you updated on our progress.

DNA methylation analyses
While we have so far focused on changes and abnormalities in the coding DNA sequence, we want to look at the epigenetic level by DNA methylation analyses. DNA methylation regulates gene expression by suppressing the transcription of genes through hyper-methylated CpG islands in promoter regions. This plays an essential role in normal cell development, but can also lead to altered gene expression in neoplasms. We are not only interested in the interplay of mutations in chromatin modulators on the methylation pattern, but also in the impact of methylation inhibitors during therapy. Last but not least, studies in other cancers show that even a classification of tumors can be performed by the methylation pattern, as these often reflect the degree of maturity of the degenerated cell. We want to explore all of this for leukemias and are starting with a whole-genome bisulfite sequencing (WGBS) approach, but also integrating single cell ATAC-Seq results. We are curious to see what else the DNA of the leukemia cell hides.