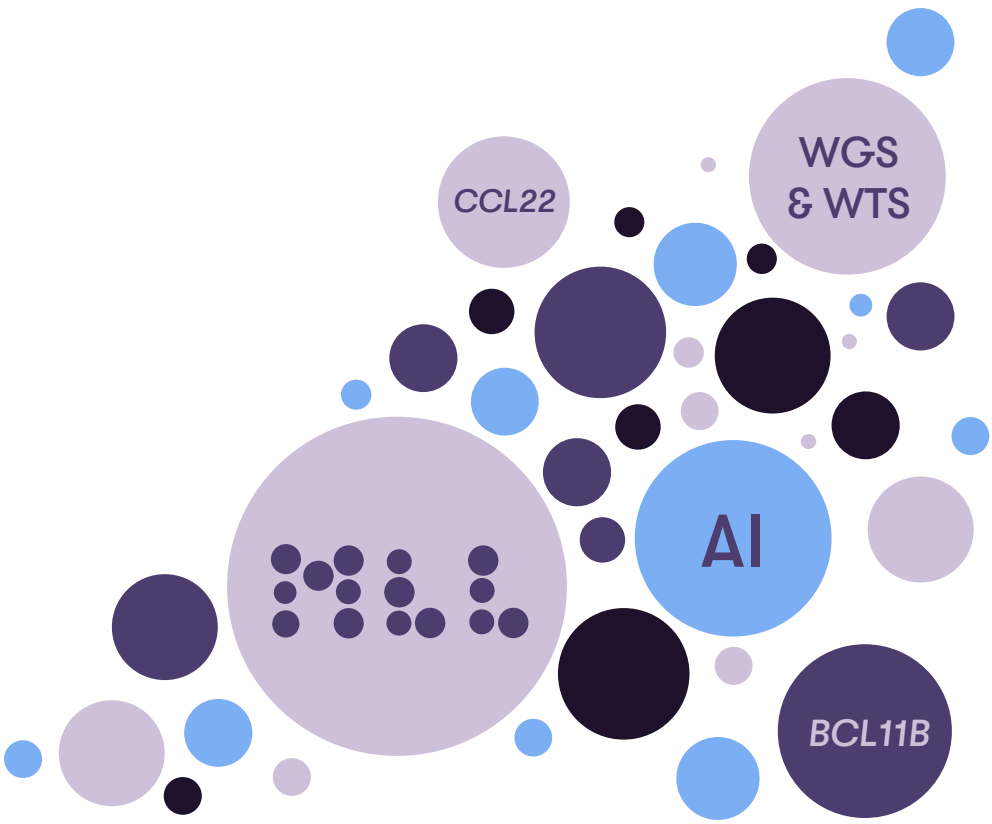
RESEARCH REPORT 2021

IN A TIME WHEN EVERYTHING SLOWS DOWN, RESEARCH CONTINUES TO THRIVE



Dear research interested reader,

after another scientifically successful year, we would like to take the opportunity to summarize our findings of 2021. 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

Even though the last two years have been dominated by the pandemic, we have kept science firmly in our mind. Providing the best possible care for our oncology patients remains our heartfelt mission. Scientifically, we have continued to focus on the evaluation of our 5,000 genome and transcriptome datasets (WGS/WTS) and the application of artificial intelligence (AI). While the WGS/WTS data contributes to the biological and genetic understanding of leukemias and lymphomas, the implementation of AI-based methods aims to optimize our diagnostic workflows.

One of the basic requirements for reliable and reproducible scientific research is the generation of standardized, high-quality data. This is especially true for next-generation sequencing (NGS) data. Over the years we have sequenced more than 5,000 human genomes, 4,700 whole transcriptomes, and 41,000 NGS gene panels, carefully optimizing every step of the workflows to meet our high quality standards and achieve a high degree of automation. This year we also obtained the CAP accreditation for our various diagnostic workflows and launched a new sequencing service (MLL SEQ) to offer our sequencing experience and capacities to everyone. We also introduced the 4-tier system for the classification of sequencing variants.

Here, we have learned how unique and therefore scientifically valuable our database of sequencing variants and their frequencies of occurrence are. Our knowledge in this area is further expanded by the extensive analysis of our WGS & WTS data. Excitingly, the explorative and comprehensive analysis of our genomic and transcriptomic data led to the discovery of new biomarkers and subtypes for ALL and NK cell neoplasms: the oncogenic role of *BCL11B* in ALL and *CCL22* mutations in NK cell neoplasms. Two exciting projects where we benefited greatly from the collaboration with St. Judes and Charles Mullighan's group. Without their know-how and the data from functional assays, such publications would not have been possible.

Over time, we have accumulated large data sets that allow us to train machine learning algorithms and to implement AI-based methodologies for various applications. In cytogenetics, for example, the routine diagnostic workflow already benefits from AI-based support for automatic karyotyping. The method was developed in close collaboration with MetaSystems. Moreover, collaborative projects with AWS in morphology for blood differentials and in immunophenotyping for diagnosing main entities of hematologic malignancies are nearly completed and are close to validation for the routine setting. Prognostic and mutation data of our large MDS cohort data also contributed to the training of a personalized MDS risk stratification model that outperformed current risk stratification scores.

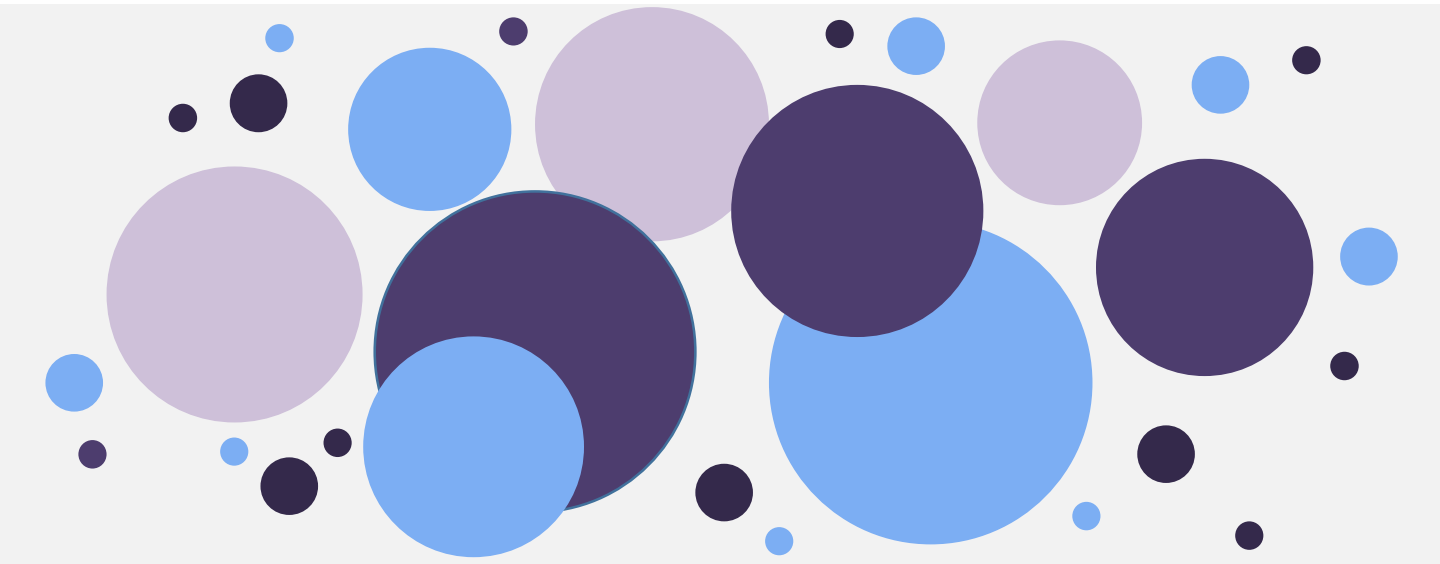
Unfortunately, we could not welcome any guest scientist (supported by the Torsten Haferlach Leukemia Foundation) at the MLL in the last year. However, the time was well spent to wrap up the analyses and data evaluation of their ongoing projects in a number of publications. Also, collaborations with large consortia like Harmony or ERIC did not stand still and made significant progress. In the field of Data Science, we have entered into a mutual scientific support with the Helmholtz Zentrum München to contribute to the education of young scientists here as well.

Despite pandemic circumstances, we continued to create knowledge through scientific research and made it accessible with currently 57 per-reviewed publications in 2021. Hence, we will start the new scientific year with joy, motivation and dedication.

M. Meggendorfer

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

Optimization of existing workflows



Standardization and rigorous quality control are key for diagnostic accuracy and reliable research

For more than 10 years, the MLL has ISO 15189 accreditation for all assays. Now in 2021, MLL Dx, MLL's affiliate, has been peer-inspected and successfully accredited by the College of American Pathologists (CAP). MLL Dx was founded in 2017 as a sister company of MLL, providing comprehensive diagnostics of hematological malignancies for international patients and in clinical studies. By combining the services and knowledge of MLL and MLL Dx, applying state-of-the-art methods, and adhering to a strict quality management plan, we aim to provide the best possible diagnosis for every patient. The CAP's laboratory accreditation program, with its annually updated checklists, has helped us to further optimize our workflows to ensure high accuracy and standardization of the various assays that form the basis for a high-quality laboratory. In addition to patient diagnostics, we are always interested in learning more about the characteristics of different disease subgroups and advancing our research. The CAP program has encouraged us to re-evaluate our established quality standards to bring them into line with international practices and to further expand our use of various controls to make our assays even more robust. Moreover, the CAP program did not only lead to further improvements of our wet-lab processes but also provided guidance to streamline our analysis pipelines, coding standards and guidelines, as well as the documentation procedure even more.

DAkkS Deutsche Akkreditierungsstelle

CAP ACCREDITED COLLEGE of AMERICAN PATHOLOGISTS

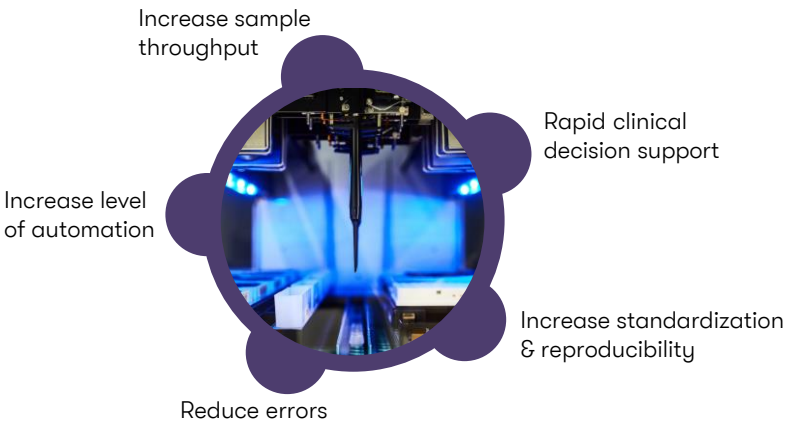


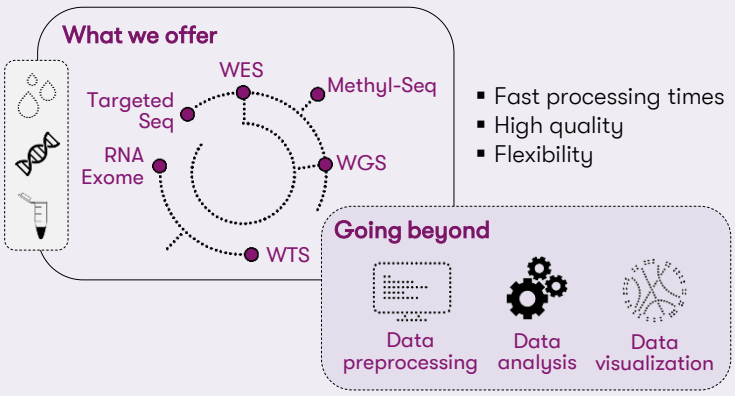
Figure 1: Overview of our aims for continuous workflow optimization.

We are continuously optimizing our workflows with existing resources and new technologies to improve their efficiency and the quality of the results. We always intent to automate as many steps of a workflow as possible to increase our efficiency, reproducibility of results, as well as prevent individual errors and reduce manual bias (Fig. 1). Only reproducible and homogenous results allow the comparison of different processing batches and the data integration from various sources.

High-quality data is a basic requirement for reliable and reproducible scientific research, which allows us to further explore the genetic characteristics of leukemias and lymphomas. Due to our high quality standards, we can guarantee the accuracy of the produced data and are confident in sharing our data for external scientific projects through various fruitful collaborations.

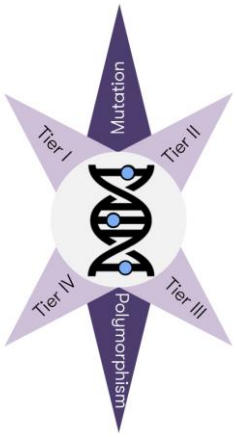
MLL SEQ

The 5,000 genome project allowed us to carefully optimize and automate every step of our NGS workflow, strengthened our expertise to sequence at the most qualitative level in high throughput with shortest turn-around times in a clinical setting. With MLL SEQ we are now offering our sequencing experience and capacities to everyone (mllseq.com/). However, our portfolio doesn't end with sequencing; we also offer data analysis and visualization services.



Introduction of the 4-tier system for the classification of sequence variants

Sequence alterations have classically been divided into two categories: "mutation" and "polymorphism", with only the former considered relevant or pathogenic. However, with the start of large-scale sequencing projects for rare diseases and various cancers, as well as the increase of available genetic profiles of healthy individuals, a broad spectrum between clearly disease-associated pathogenic alterations and nonpathogenic polymorphisms has been revealed. In the case of *BRAF*, for example, the well known V600E alteration allows the use of targeted therapies but this doesn't apply to atypical and less well studied alterations in the same gene. In the last years, genomic sequence data have been produced in an exponential way but the knowledge regarding the clinical interpretation and functional analysis of detected variants, doesn't grow equally fast. Hence, in the case of missing reference data a refined distinction might be necessary.



On the basis of a combined effort by various working groups a 4-tiered system to categorize somatic sequence variations based on their clinical significances has been proposed, which we adapted to central questions in hematological diagnostics:

1) Can a diagnosis, prognosis or therapy be currently derived?
2) Does the change prove the clonality of hematopoiesis?
3) Can the change be used as a progression marker?

Therefore, we integrated our internal variant data, public databases and *in silico* predictions into a classification that can support clinical decisions. Cancer genomics is a rapidly evolving field and hence we reevaluate the clinical significance of any variant in therapy, diagnosis, or prognosis periodically.

A large proportion of *RUNX1* or *TP53* alterations belong to this category, as they are associated with poor prognosis in many entities. Other examples in this category include alterations that allow targeted therapy, cause resistance to such therapy, or allow diagnosis according to WHO. However, the significance for diagnosis, prognosis, and therapy must always be evaluated in the overall context of all findings obtained, and the respective disease-specific interpretation of the molecular genetic findings must be considered.

Positive/Tier 1

- Known pathogenic/very likely pathogenic
- relevant to diagnosis, prognosis, and/or therapy

Positive/Tier 2

- likely somatic alteration
- no definite statement on pathogenicity can be made at the time of analysis

Even though compared to tier 1 the data status is often weaker, tier 2 alterations can guide the diagnostic process and prove the clonality of hematopoiesis. Tier 2 alterations can also be used as progression markers. It's important to consider the overall context for the interpretation, including additional findings. The significance of tier 2 alterations in genes, for which mainly hotspot mutations are known, is unclear.

Congenital alterations are present in healthy and malignant cells. In order to reliably differentiate between germ line and somatic variants, the sequencing of a matching control (e.g. buccal swab, fingernail, etc.) is necessary. Only acquired variants are suitable as progression markers, since they should no longer be present after successful therapy.

Variant/Tier 3

- currently unclear whether this is a rare congenital germ line alteration or an acquired somatic alteration

Polymorphism/Tier 4

- Benign/likely benign
- No known disease association

They are by far the most frequent alterations in our daily life, are documented internally but not listed in the report.

Extension of the routine molecular genetics workflow to detect CNV and CN-LOH events

Chromosome banding analysis (CBA) is the current gold standard for the detection of chromosomal aberrations. In recent years, it has been shown that broad next-generation sequencing (NGS) assays, such as whole exome sequencing (WES) and whole genome sequencing (WGS), are valid alternatives to CBA. However, both assays are still comparably expensive, hampering their broad application in routine diagnostics of hematological malignancies, for now. Our routine molecular genetics workflow applies target enrichment sequencing (TES) for molecular profiling of patients with hematologic neoplasms. The hybridization and capture assay is highly flexible due to the variety of available gene panels. Since April 2021, we extended the workflow to simultaneously assess copy number variations (CNV) and copy-neutral loss-of-heterozygosity (CN-LOH) by adding a CNV spike-in panel (Fig. 1).

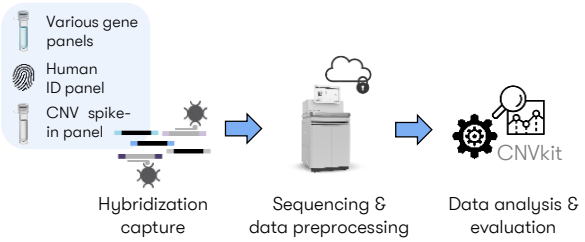


Figure 1: Schematic overview of the routine target enrichment workflow to detect CNV and CN-LOH events.

The xGen human CNV backbone panel (Integrated DNA Technologies) contains >9,000 oligonucleotide probes that span the entire genome with an approximate distance of ~0.34 Mb. The obtained coverage and variant calling data is 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. The panel allows the detection of chromosome/-arm gains and losses as well as most derivative chromosomes (Fig. 2).

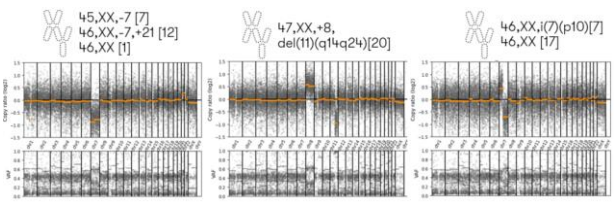


Figure 2: Examplary plots for patients with monosomy 7 (left), trisomy 8 (middle), and isochromosome 7 (right).

Copy-neutral loss-of-heterozygosity (CN-LOH) - not detectable by CBA - is gaining importance as an additional prognostic factor and can either cause the duplication of an activating mutation in an oncogene, the deletion of a tumor suppressor gene or the gain/loss of specific methylated regions, influencing gene expression (Fig. 3).

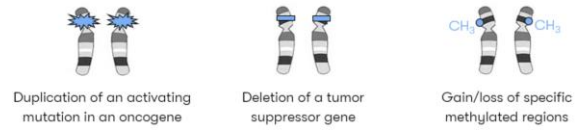


Figure 3: Simplified illustration of the effect of CN-LOH events.

A detailed analysis of 1196 patients, sent to the MLL between 04/2021–07/2021 for diagnostic work-up, revealed that ~10% of the patients harbor a CN-LOH event without any association to age or gender but a slightly higher incidence in myeloid compared to lymphoid neoplasms. CN-LOH occurred most frequently in 4q, 7q, 9p and 11q. 4q CN-LOH co-occurred with high variant allele frequencies (VAF) of *TET2*. 7q CN-LOH occurred nearly exclusively in myeloid neoplasms (95%) and was associated with high VAFs in *EZH2* and *CUX1* variants. 9p CN-LOH led to JAK2V617F homozygosity and 11q CN-LOH occurred more often in myeloid than lymphoid neoplasms (79% vs 21%), mostly associated with *CBL*, *KMT2A*-PTD and *ATM*.

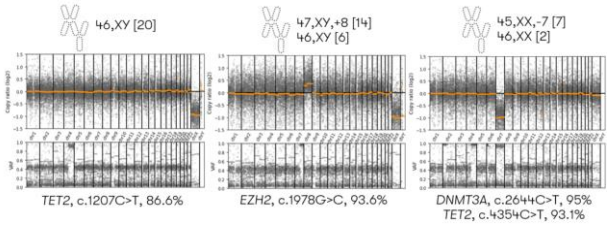


Figure 4: Examplary plots for patients with CN-LOH in 4q (left), 7q (middle), and 2p, 4q and 6p (right).

WGS data was used to reveal additional associations between high VAFs of various genes and CN-LOH events (Fig. 5).

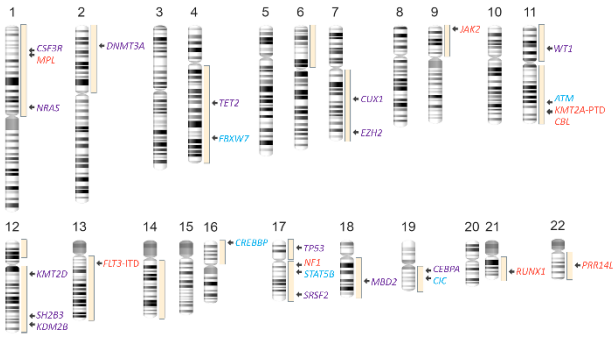
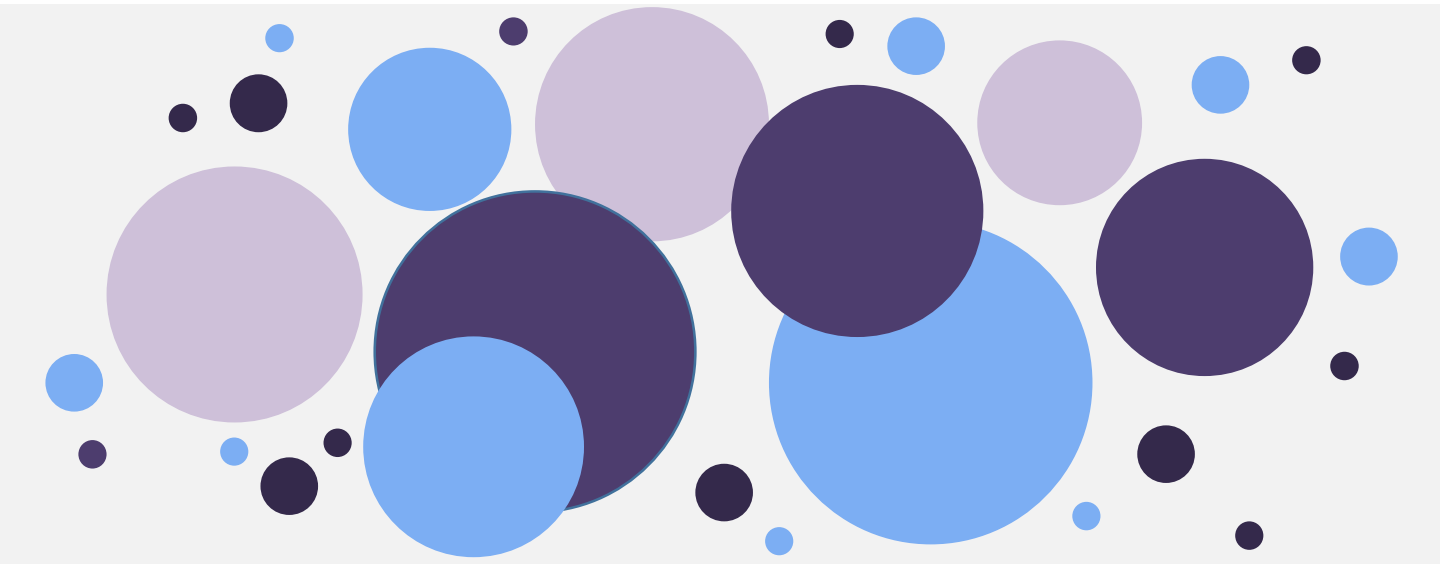


Figure 5: Schematic overview of the associations between CN-LOH events and high VAFs of various genes and their occurrence: red: only myeloid neoplasms, blue: only lymphoid neoplasms, purple: myeloid and lymphoid neoplasms.

By using a CNV spike-in panel, TES adds additional diagnostic and prognostic information by enabling simultaneous detection of selected gene mutations and genome-wide CNVs, as well as CN-LOH, without increase in sequencing costs and turn-around times.

Walter et al. 2021, *Blood*,
https://doi.org/10.1182/blood-2021-152195

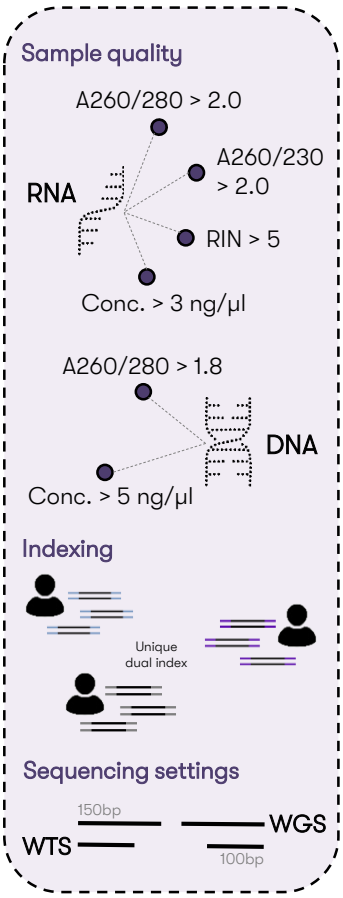
WGS & WTS for routine diagnostics



WGS & WTS in leukemia: reliable tools for diagnostics?

Motivation

The analysis of the tumor genome has undergone revolutionary developments over the past decade, with whole genome sequencing (WGS) providing an unprecedented insight into cancer biology and pathogenesis. In order to complete the molecular picture, whole transcriptome sequencing (WTS) has been used and is often applied to subtype classification in acute lymphoblastic leukemia patients. Comprehensive WGS and WTS data may potentially impact diagnostics, prognostication and therapy selection in the near future. However, there are still some pitfalls to recognize and challenges to overcome before WGS & WTS can advance clinical diagnostics and patient care.

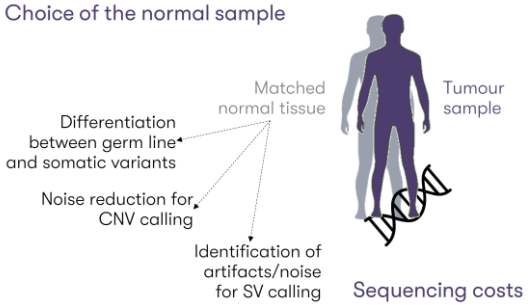


Pre-analytical considerations

- the amount of DNA input varies depending on the applied library preparation protocol
  - PCR-free methods require >1 µg of DNA input
  - Amplification-based methods require 10 ng-100 ng DNA
- RNA input for WTS ranges from ~10 pg to ~300 ng
- only high quality nucleic acids will result in clinically useful WGS & WTS results
  - DNA quality is stable for samples stored at room temperature but extended storage time will result in reduced DNA amounts
  - RNA is extremely labile and degrades fast in blood samples stored at room temperature
- the right control sample has to be chosen with care
  - buccal swap or sorted T-cells are commonly used as controls in haematology to reliably differentiate between germ line and somatic variants, but these tissues yield only low amounts of high-quality DNA (+ minor contamination)
  - due to the relative nature of RNA-Seq the choice of an adequate control and normalization method is of prime importance
- implementation of quality checks and standardization of the library prep workflows are essential to reduce manual bias and to increase homogeneity and reproducibility of WGS & WTS results
- selection of a solid indexing strategy (e.g. unique dual indices) for efficient multiplex sequencing
- selection of the adequate sequencing settings, depending on the insert size of the library fragments
- the sequencing depth directly influences the sensitivity of the results
- a fast and secure computational infrastructure has to be established to safely store the data and to allow a fast computation

Current challenges

- especially for WGS, limited sensitivity due to a trade-off between sequencing depth and costs
- the coverage is not uniformly distributed over the whole genome, with significantly lower coverage in GC-rich and repetitive regions
- comparably long turn-around times for WGS & WTS workflows
- limited knowledge of clinically relevant genomic variants
- despite the reduction in sequencing costs, WGS & WTS are still comparably expensive



Meggendorfer et al. 2020,
Best Practice & Research Clinical Haematology,
https://doi.org/10.1016/j.beha.2020.101190

Walter et al. 2020,
Seminars in Cancer Biology,
https://doi.org/10.1016/j.semcancer.2020.10.015

Meggendorfer et al. 2021,
Seminars in Cancer Biology,
https://doi.org/10.1016/j.semcancer.2021.06.009

Whole genome sequencing through the looking glass

It took 14 years to sequence the first human genome, involving thousands of scientists all over the world and ragging up the enormous sum of 3 billion dollars. The scientific description of the finished human genome sequence was finally presented in October 2004. Since then many things have changed. With the introduction of next-generation sequencing, the costs for sequencing a single human genome dropped rapidly over the years to ~$1000. After a brief period of stagnation, the costs per genome have recently fallen below the $100 mark. Hence, we have already come a long way in making whole genome sequencing (WGS) affordable for healthcare. However, the results also have to be rapidly available in order to maintain short turnaround times, providing the patient with a timely diagnosis. Starting in 2017, we have sequenced more than 5,000 genomes, carefully improving and optimizing every step of the workflow. So, how are human genomes sequenced at the MLL?

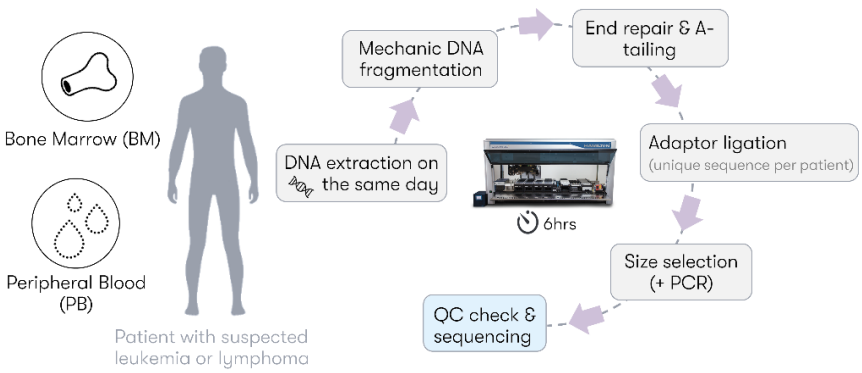


Figure 1: Schematic overview of the WGS library preparation workflows

As soon as a bone marrow biopsy or peripheral blood sample arrives at the MLL, the samples are pre-processed (cell lysis) and stored. The DNA is then extracted using the MagNA Pure 96 instrument. WGS libraries are prepared from 1µg of DNA with the TruSeq PCR free library prep kit within the next day(s). The WGS library preparation (Fig. 1) runs fully automated on the Hamilton NGS STAR system.

Every WGS library is rigorously checked for different quality parameters and the samples are then sequenced on the NovaSeq 6000 system, generating 150bp paired-end reads. Depending on the used flow cell, the sequencing itself takes between 25 and 48 hours. For the detection of somatic mutations and small clones, the WGS samples are sequenced with an average coverage of 90x. However, the sensitivity for variant detection is still lower compared to targeted sequencing. The sequencing data is directly streamed into our private instance of the Amazon Web Services (AWS) cloud and preprocessed in the BaseSpace sequence hub (Fig. 2).

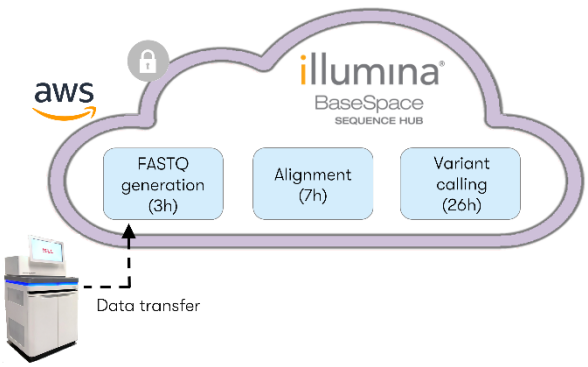


Figure 2: Schematic overview of the WGS data pre-processing.

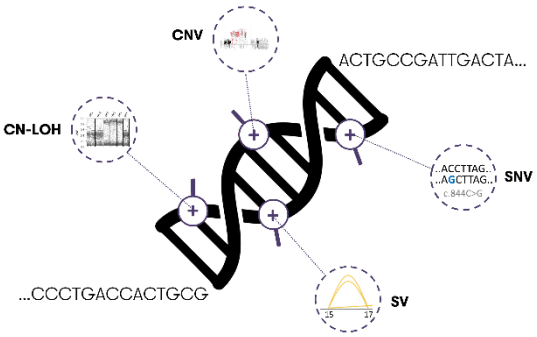


Figure 3: Schematic overview of the different data types that are analyzed from WGS samples.

Most of the times no sample specific normal tissue is available and a so-called Tumor/unmatched normal workflow is used for variant calling to reduce technical artefacts and germ line calls. The obtained calls are further analyzed and filtered by in-house pipelines (SNV, SV). Copy number variations are assessed with GATK4, applying a panel of normals (WGS data of patients with a normal karyotype as reported by cytogenetics) for the normalization of the profiles. Copy-neutral loss of heterozygosity is estimated applying Hadoop software. The combined data is analyzed for patient-specific reports, listing the identified aberrations and potential implications.

2.161
Genomes sequenced since 2019

Median turn-around time (days)
7-8

>3.4
Petabyte of sequencing data

Computation time to analyze a single WGS sample (hrs)
38-45

A comprehensive yet focused analysis of hematological neoplasms - the clinical cancer genome and transcriptome

The human genome consists of about 3 billion letters and all human beings are 99.9% identical in their genetic makeup. Important clues about the predisposition and causes of various diseases can be found in the remaining 0.1% of differences. With the decrease in sequencing costs whole genome sequencing becomes feasible opening the way to comprehensively access a patient's genetic profile. However, the clinical evaluation of the genome focuses solely on acquired mutations and the manual annotation and interpretation of identified genetic variants will be restricted to genomic regions associated with the suspected diagnosis to reduce interpretation complexity and to facilitate comprehensive reporting in a clinically-relevant timeframe. Genetic variants identified outside the predefined regions of interest will be ignored and won't be included in the manual evaluation process, effectively eliminating the possibility of incidental findings.

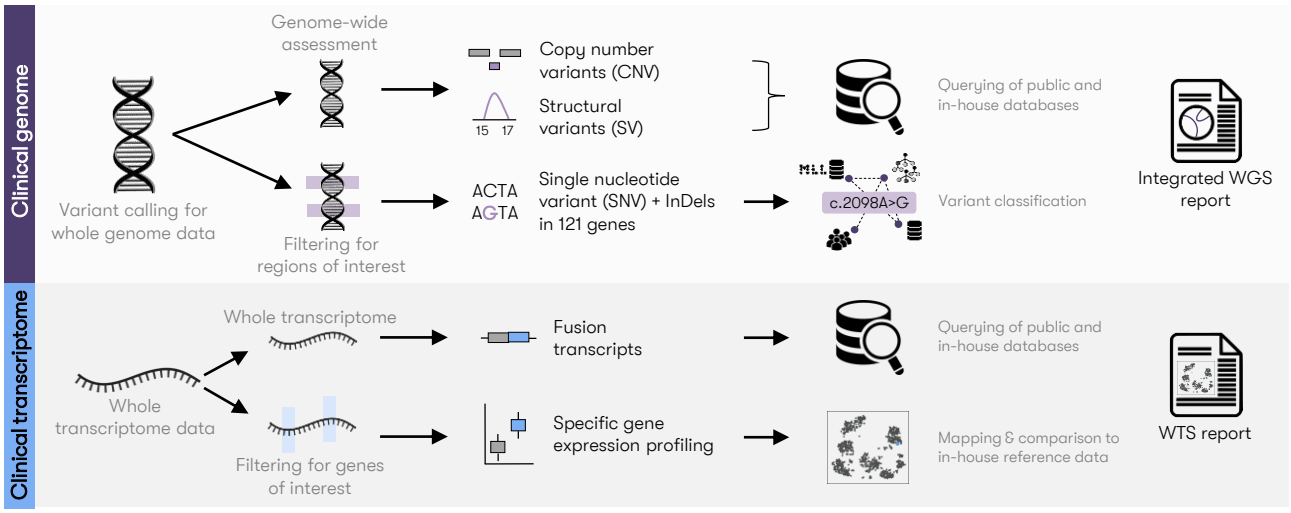


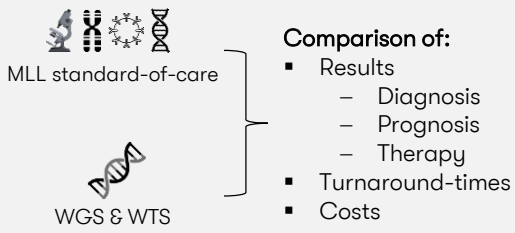
Figure 1: Pictographic overview of the clinical genome and transcriptome

For the evaluation of the somatic mutation profile we currently consider 121 genes with established associations to hematological malignancies. For these genes only protein-altering and splice-site variants that passed the quality filtering by the variant caller are considered. Furthermore, each variant is queried against the gnomAD database and variants with a global population frequency > 0.1% are excluded. For indels, additional 'non-PASS' variants are considered if they have a variant allele frequency (VAF) ≥ 0.1 and ≥ 5 read support. Variant classification is performed following a 4-tier system (p. 7) and the filtered and annotated list of variants is reported. Structural variants and copy number variations are assessed genome-wide, as currently done by chromosome banding analysis. Artefact filtering is performed based on our 5,000 in-house genome data set. A patient's transcriptome is analyzed for the occurrence of fusion transcripts and specific gene expression signatures or single gene expressions with diagnostic or prognostic value for different leukemias and lymphomas. The expression values are normalized and compared to either a control group of healthy individuals or within a disease entity. 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.

Solving Riddles Through Sequencing (SIRIUS)

ClinicalTrials.gov Identifier: NCT05046444

For ~90% of all patients with suspected hematologic malignancies a swift diagnosis can be made by the application of standard techniques. The other 10% remain a challenge. The objective of this trial is to test whether WGS & WTS can surpass the current gold standard regarding diagnostic precision and routine reliability for these challenging cases.



More data, more clinically relevant information?

Increasing the size of the genomic region of interest naturally increases the probability to encounter so-called 'incidental findings': the detection of clinically relevant genomic variants unrelated to the current diagnostic question. But do we also gain additional, clinically relevant information in relation to the actual diagnostic question by increasing the size of the sequenced region? We addressed this question by analyzing the sequencing data (deep sequencing & WGS) of 588 patients diagnosed with MDS by cytomorphology and cytogenetics according to the WHO classification. 85% of MDS patients carry at least one mutation at diagnosis but the median number of mutations per patient is low. So how much genetic information can be gained by increasing the number of analyzed genes?

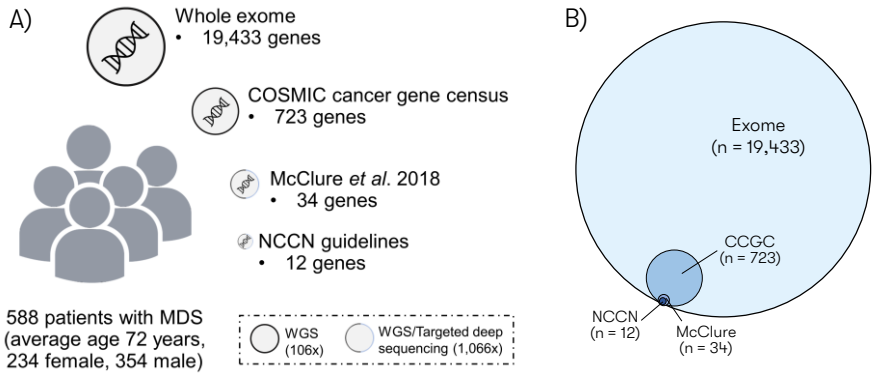


Figure 1: Schematic overview of the cohort, applied *in silico* panels and the sequencing depth of the respective methods (A), as well as the gene overlap between the different panels (B).

The targeted sequencing and WGS data was analyzed for 2 and 4 virtual gene panels respectively: the 12 genes as given in the MDS NCCN guidelines with a incidence >5% (Version 2.2020), the 34 gene panel summarized in *McClure et al.* 2018, the 723 COSMIC cancer gene census (CCGC, v91) panel and finally the exome (Fig. 1).

Variant interpretation of protein altering mutations was performed by considering ClinVar, Cosmic, functional impact predictors and a population frequency in gnomAD ≤0.00005. Variants with no or discordant information were declared as variants of uncertain significance (VUS). Jumping the size from just 12 analyzed genes (NCCN) to the whole exome, the percentage of patients carrying at least one mutation increased from 81% to 100% and the median number of mutations per patient increased from 2 to 10 (Fig. 2). For the smaller panels a slightly higher number of mutations could be detected in male patients (NCCN gene panel: 63/234, 27% vs. 55/354, 16%; p=0.001; 34 McClure gene panel: 52/234, 22% vs. 44/354, 12%; p=0.002).

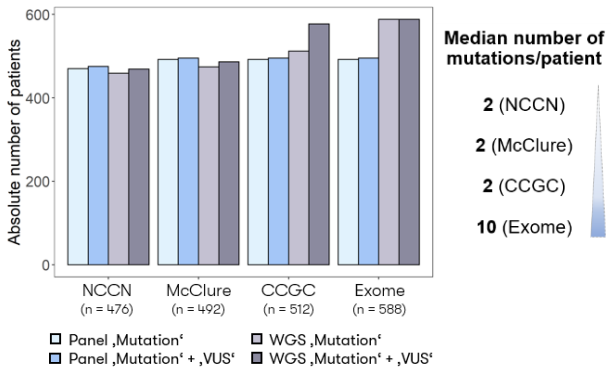


Figure 2: Number of patients that harbor at least one mutation (light blue, lavender) or at least one mutation + VUS (blue, purple) per panel.

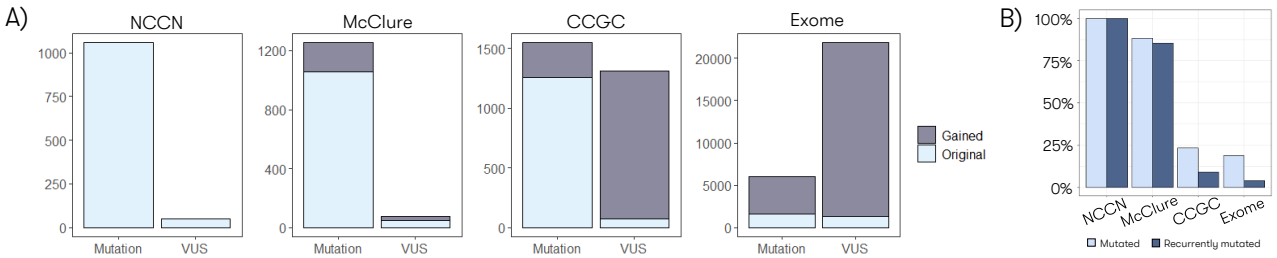


Figure 3: Overview of gained variants per panel. (A) Number of gained variants (purple) for each panel transition in comparison to the detected variants from the previous panel(s) (light blue). (B) Relative abundance of recurrent mutations among the gained mutations per panel.

As can be seen in Fig. 3A, the number of detected mutations decreased while the number of VUS rose with increasing panel size. This is most evident in the exome dataset with a remarkable increase of 20,540 VUS detected. Here, all patients harbored at least one mutation but most of the gene mutations occurred in only one individual (Fig. 3B).

As demonstrated by this study, for MDS patients, a large gene panel shows only small gain of clinically relevant information but many additional VUS. Therefore, for MDS, the choice of a small but well selected gene panel is preferable to large sequencing efforts to obtain the same amount of information with less cost and evaluation time.

Meggendorfer et al. 2020, *Blood*, https://doi.org/10.1182/blood-2020-139274

Meggendorfer et al. 2021, *EHA Library*, 06/09/21; 325650; EP892

WGS & WTS allow the identification of key genetic subtypes in acute lymphoblastic leukemia

Considering the clinical and genetic characteristics, acute lymphoblastic leukemia (ALL) is a rather heterogeneous hematological neoplasm for which current standard diagnostics require various analyses encompassing morphology, immunophenotyping, cytogenetics, and molecular analysis of gene fusions and mutations. Even with this plethora of methods, it is not possible to comprehensively cover all subtypes described so far. 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. Hence, we explored the clinical utility of WTS for the genetic characterization of ALL and subsequent patient stratification following the classification tree in Figure 1.

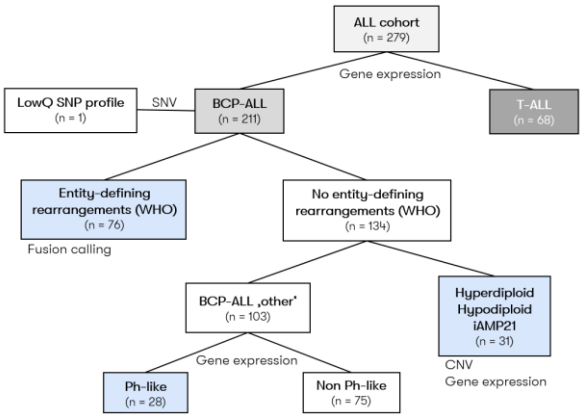


Figure 1: Classification tree. Design of the multi-method classification approach and distribution of the patients.

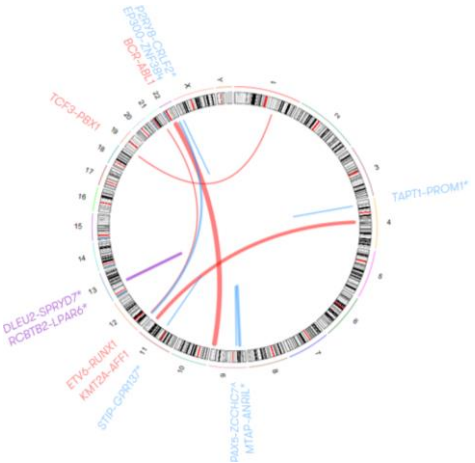


Figure 2: Illustration of identified recurrent fusion transcripts. Red - canonical, blue - known, purple - novel. *indicate deletion/read-through events, ^ inversions. The line width is proportional to the fusion transcript frequency in the cohort.

In the initial classification step we used the expression levels of described marker genes to reliably distinguish between BCP-ALL and T-ALL samples. Subsequently, the BCP-ALL samples were further subclassified by the identification of recurrent risk-stratifying gene fusions (Fig. 2). We used three different fusion callers (Arriba, STAR-Fusion, Manta) to detect fusion transcript candidates, relying only on those identified by at least two callers and not detected in control samples. Compared to the standard methods, 97% of recurrent risk-stratifying fusions could be identified by WTS. Additionally, read-through fusions indicative of *CDKN2A* and *RB1* gene deletions were recurrently detected in the cohort, along with 57 putative novel fusions with yet untouched diagnostic potentials. The CNVkit toolkit was used to identify patients with high hyperdiploidy or low hypo-diploidy/near-triploidy, correctly identifying 17 (94%) low hypodiploid/near-triploid and 12 (80%) high hyperdiploid cases. 103 BCP-ALL samples of the cohort had no established abnormalities and were further referred to as BCP-ALL 'other'.

However, the beauty of broad assays such as WGS and WTS is that they allow to go beyond the standard analytical spectrum to stratify patients even further. Thus, we used the gene expression profile to classify BCP-ALL 'other' cases into the '*BCR::ABL1* like' and 'non *BCR::ABL1* like' groups. We also identified cases harboring other translocations - beside the subtype defining rearrangements - involving *ZNF384*, *HLF*, and *NUTM1* that were all assigned to the 'non *BCR::ABL1* like' group and, hence, could be further subclassified based on these genetic alterations (Fig. 3). A group of samples with a homogenous gene expression profile distinct from the others could be assigned to the newly identified subgroup of BCP-ALLs with *DUX4* rearrangements. WGS data allowed the identification of cases with a) a complex karyotype and *TP53* mutations, b) a near haploid karyotype, c) *PAX5* mutations, d) MYC rearrangements and e) rearrangements involving IGH and members of the CEBP family.

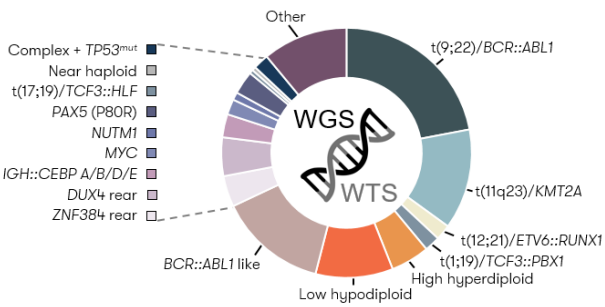


Figure 3: Genetic subtyping of BCP "other" ALL by thorough WGS and WTS data evaluation.

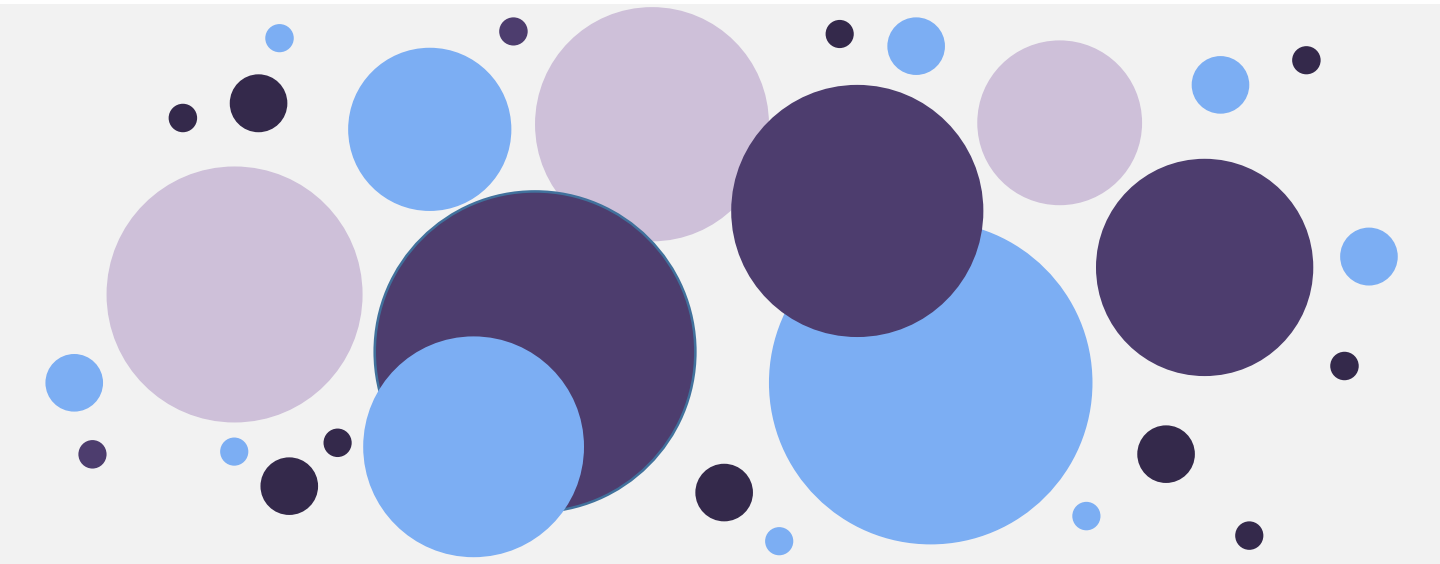
Hence, WTS can be used to classify ALL patients with a single assay and is superior to conventional methods in cases which lack entity-defining genetic abnormalities. The analysis can be further enriched by the integration of WGS data, adding relevant information for prognostication and therapy selection for subsets of patients.

Walter et al. 2021, *BMC Cancer*, https://doi.org/10.1186/s12885-021-08635-5

Haferlach et al. 2021, *Blood*, https://doi.org/10.1182/blood-2021-150260

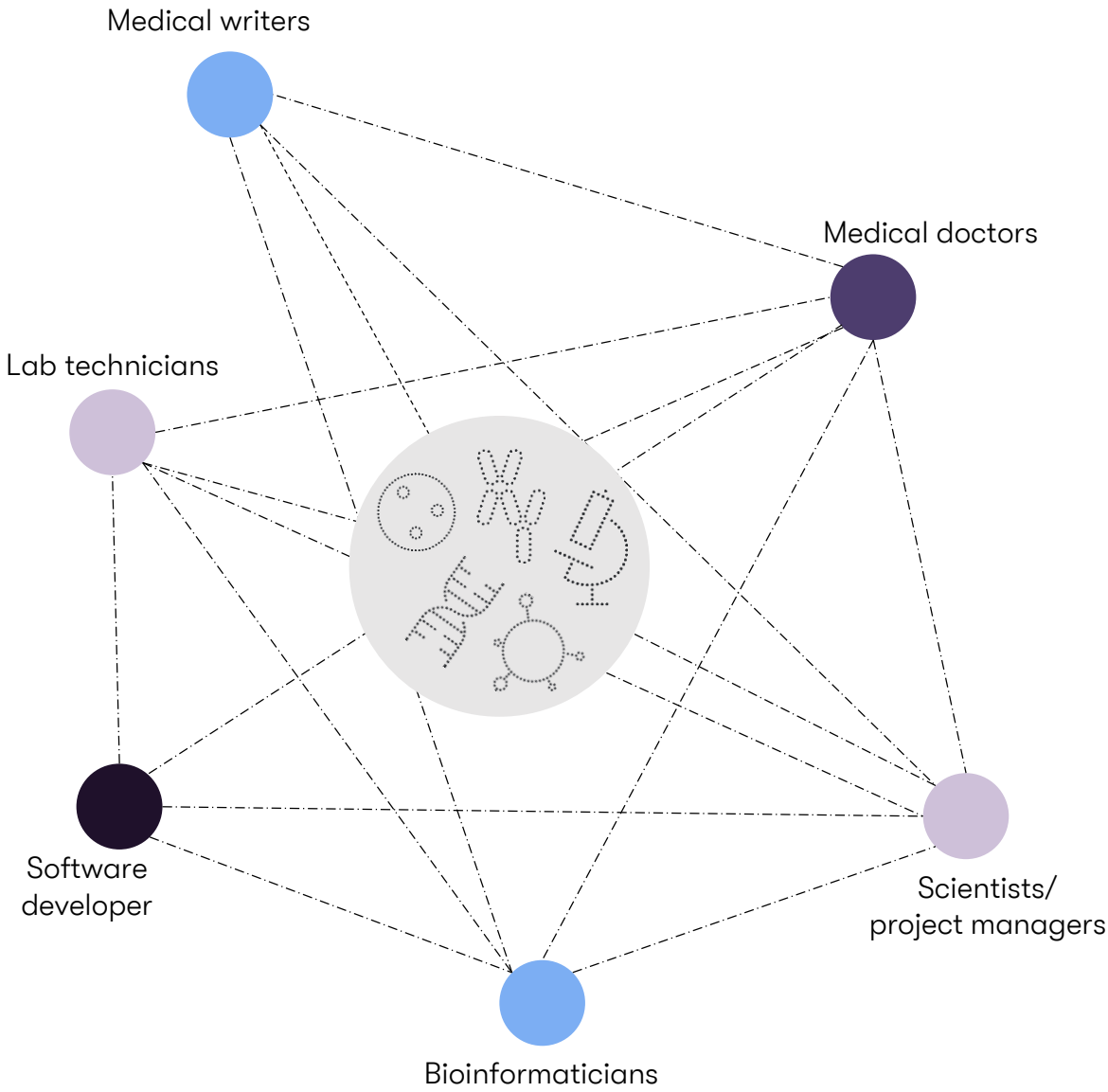
Research highlights, Manuscript speed dating,

Guest visiting scientists



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: Enhancer hijacking drives oncogenic *BCL11B* expression in lineage-ambiguous leukemia

BCL11B a locus with multiple functions

BCL11B is an important oncogene for early human acute leukemia with ambiguous lineage. The *BCL11B* locus can drive a variety of immature leukemias through different pathogenic mechanisms:

1) Recurrent chromosomal rearrangements between *BCL11B* and various oncogens (e.g. *TLX3*), leading to *BCL11B* enhancer hijacking that drives ectopic expression of the oncogenes.

2) The occurrence of a t(2;14)(q22;q32) translocation resulting in an in-frame *ZEB2-BCL11B* fusion transcript (often in combination with activating *FLT3* mutations) and aberrant *BCL11B* expression in early progenitor cells.

3) Repositioning of different regulatory sequences upstream or downstream of *BCL11B* resulting in high *BCL11B* expression (=*BCL11B*-activated acute leukemia) and a unique expression profile (incl JAK/STAT signature) further characterized by *FLT3* mutations in the absence of *NOTCH1*-activating mutations.

Acute leukemias of ambiguous lineage (ALAL) are notoriously challenging to classify and lack clear subtype-defining genomic alterations and appropriate treatment strategies. ALALs either show limited lineage differentiation or exhibit immunophenotypic features of multiple lineages and are often characterized by various gene mutations affecting myeloid maturation, kinase signaling, and chromatin modification among others. However, the oncogenic drivers of the different types of ALALs remain poorly understood. *Montefiori et al.* identified diverse, predominantly noncoding structural variants (SV) driving enhancer hijacking events leading to aberrant *BCL11B* expression and a specific gene expression signature in hematopoietic progenitor cells, defining a subtype of lineage-ambiguous leukemia.

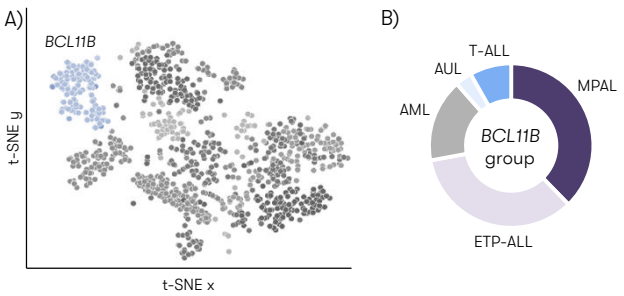


Figure 1: *BCL11B* deregulation in lineage-ambiguous leukemia. A) Exemplary representation of the specific gene expression profile of *BCL11B* deregulated cases loosely based on *Montefiori et al.* B) Relative frequency of the various types of ALAL within the *BCL11B* group. MPAL: mixed phenotype acute leukemia; ETP-ALL: early T-cell precursor acute lymphoblastic leukemia; AML: acute myeloid leukemia; AUL: acute undifferentiated leukemia; T-ALL: T-cell acute lymphoblastic leukaemia

The initial analysis of transcriptional profiles of 2,574 patient samples showed that various types of leukemias (Fig. 1B) with stem/myeloid/T-cell phenotypes clustered together, defying immunophenotypic and diagnostic boundaries (Fig. 1A).

The unifying feature of these 61 cases? Monoallelic expression of *BCL11B*, suggesting *BCL11B* deregulation as the driver event. Samples within the *BCL11B* group were characterized by high *FLT3* expression (Fig. 2) and enriched for either an activating internal tandem duplication (ITD) or D835Y mutation in *FLT3*, detected in 49/61 (80%) of the cases.

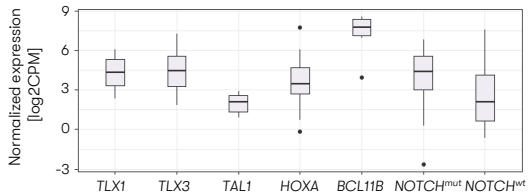


Figure 2: *FLT3* expression in various subtypes of T-ALLs (MLL data).

To determine the genomic mechanism for *BCL11B* deregulation, samples from the *BCL11B* group were subjected to HiChiP analysis, which revealed that the identified noncoding SVs led to ectopic chromatin interactions between *BCL11B* and various super-enhancers that were particularly active in uncommitted normal hematopoietic cells. Moreover, a new mode of oncogenic enhancer formation caused by tandem amplification (>15x) of an otherwise inconspicuous noncoding element in cis with *BCL11B* (= *BCL11B* enhancer tandem amplification, or BETA) was identified. Subsequent single-cell analysis supported the hypothesis that *BCL11B* drives, or reinforces, a stem cell gene expression program in *BCL11B*-group leukemia, possibly reflecting the cellular origin of this disease.

In summary, *BCL11B* can act as an oncogene when it's aberrantly activated in hematopoietic stem/progenitor cells. This contrasts with its well-known tumor suppressor roles in T-ALL.

Meijerink 2021, *Blood,*
https://doi.org/10.1182/blood.2021011856

Montefiori et al. 2021, Cancer Discovery,
https://doi.org/10.1158/2159-8290.CD-21-0145

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

FISH and WGS in newly diagnosed and relapsed/refractory multiple myeloma – WGS will affect future treatment decisions

Introduction
Multiple myeloma (MM) is a malignant plasma cell disorder with broadly varying clinical symptoms and outcome for which the diagnosis and risk stratification is mainly based on Fluorescence in situ hybridization (FISH). However, different studies have shown the clinical value and potential of WGS to gain new insights into MM genetics.

Patients & Methods
Parallel analysis of 100 patients (pts) by WGS FISH.

Results
- Comparison of FISH and WGS results showed 100% concordance for recurrent SVs: t(11;14) in 29%, t(4;14) in 17%, and t(14;16) in 5%
- Due to heterogeneous breakpoints FISH is less reliable for the detection of *MYC* rearrangements than WGS (57% vs. 95%)
- Rare IGH rearrangements could be completely resolved by WGS with a higher detection rate compared to FISH (11 pts vs. 8 pts)
- WGS detected 95% of copy number variations identified by CBA/FISH. Missed aberrations detected by FISH were all identified in small subclones with a median clone size of 10%
- WGS revealed additional clinically relevant information such as *TP53* and 1p deletions not detected by FISH, as well as potential drug targets like *BRAF* and *CRBN* mutations

Conclusion
WGS is superior to FISH for the identification of biallelic events and rearrangements with varying breakpoints and rare partner genes. Moreover, WGS provides a comprehensive assessment of the genetic profile for patients diagnosed with MM. However, the assay is limited in the detection of small clones (<15%) and variant ploidy levels. Here, as well as for patients with less than 10% plasma cell infiltration of the bone marrow, FISH is still the gold standard.

Truger et al. 2021, *Blood,* https://bit.ly/3GUlWFu

Single- and double-hit events in genes encoding immune targets before and after T cell-engaging antibody therapy in MM

Introduction
Novel T cell–engaging therapies such as chimeric antigen receptor (CAR) T cells or bispecific antibodies (BsAb) have entered the treatment for multiple myeloma (MM) and show exceptionally high response rates in relapsed/refractory (RR) patients. However, patients still continue to relapse due to various mechanisms of resistance to T cell–based therapies.

Patients & Methods
WGS and WTS analysis of a 56-year-old patient with CD3xBCMA BsAb as well as 100 patients with MM.

Results
- Detection of a homozygous deletion at 16p13.13 leading to antigen loss and subsequent failure of BCMA therapy in a 56-year-old patient
- Detection of heterozygous deletions in immunotherapy targets in 30% of T cell immunotherapy-naïve patients
- 15/21 targets were encoded on chromosomes involved in hyperdiploid karyotypes or on chromosome 1q and were consequently amplified in ~50% of these patients
- Single-nucleotide variants occurred only with low frequencies in genes encoding for immune targets
- No changes in gene expression could be detected for gains, heterozygous deletions, and SNVs in immunotherapy targets. Only biallelic aberrations resulted in a loss of expression of the respective gene

Conclusion
The study supports the use of immunotherapies earlier in the disease process, because the frequency of deletions and mutations in genes encoding immunotherapy targets was lower in patients with newly diagnosed MM vs those with relapsed/refractory MM.

Truger et al. 2021, *Blood Advances,* https://doi.org/10.1182/bloodadvances.2021004418

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

Mutational patterns and their correlation to CHIP-related mutations and age in hematological malignancies

Introduction

Molecular mutations play crucial roles in the pathogenesis and classification of many types of leukemias and lymphomas but the mutational landscapes of different hematological cancers and their relation to clonal hematopoiesis of indeterminate potential (CHIP) and age have not been evaluated comprehensively.

Patients & Methods

Mutational analysis of whole genome sequencing data from 3096 patients diagnosed with 28 different hematological malignancies.

Results

- Cases with at least one mutation in CHIP genes (*DNMT3A*, *TET2*, *ASXL1*) were characterized by high frequencies of mutations in *RUNX1*, *SRSF2*, *IDH2*, *NRAS* and *EZH2*
- *TP53*, *KRAS*, *WT1*, and *SF3B1* mutations occurred more frequently in cases without CHIP-associated gene mutations
- Older patients showed a significantly higher median number of mutations, with age-dependent mutational profiles for most entities

Conclusion

The data and obtained results indicate a potentially larger role for CHIP in various hematological malignancies, characterize the relationship between DTA mutations and other mutations in more detail and enlighten the different mutational landscapes of leukemias and lymphomas and their interrelation with aging.

Stengel et al. 2021, *Blood Advances,* https://doi.org/10.1182/bloodadvances.2021004668

The diverse landscape of fusion transcripts in 25 different hematological entities

Introduction

Genomic alterations, including structural variants (SV), are a hallmark of hematological malignancies. SVs lead to the co-localization of remote genomic material and, depending on the breakpoint locations, result either in the generation of a fusion transcript (breakpoints are located within two genes) or an aberrant expression of one gene (breakpoints are located outside of genes, placing one gene under the influence of the regulatory sequence of the partner).

Patients & Methods

WTS was performed in 3549 patients diagnosed with 25 different hematological neoplasms. Potential fusions were called with Arriba, STAR-Fusion, and Manta. Only fusion transcripts, called by at least two callers, not identified in control samples, validated by WGS and containing at least one protein-coding gene, were considered.

Results

- 1308 (806 distinct) fusion transcripts were identified in 932/3549 (26%) patients
- 541/932 (58%) patients harbored a minimum of one recurrent fusion transcripts
- Myeloid entities, except for CML and AML, showed low fusion frequencies
- Despite high fusion frequencies in lymphoid neoplasms, only few fusions occurred recurrently
- 24/50 (48%) of the recurrent fusions were specific for one entity
- 1,270 different genes were involved in the 806 fusion transcripts of which the majority (1,189 genes) were solely involved in unique fusions

Conclusion

The entities can be divided into three groups: (1) malignancies with a high proportion of cases harboring fusions, (2) entities with a large number of detected fusions, and (3) *BCR-ABL* negative chronic myeloid malignancies with few fusions.

Haferlach et al. 2021, *Leukemia & Lymphoma* https://doi.org/10.1080/10428194.2021.1953009

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

Diagnostic challenge of identifying cases with recurrent t(8;14)(q24.21;q32.2) Involving *BCL11B* in ALAL: an analysis of eight patients

Introduction
BCL11B (located at 14q32.2) encodes for a zinc finger transcription factor that is specifically expressed in T cells. Aberrant *BCL11B* expression can be the result of *BCL11B* mutations or structural rearrangements of *BCL11B*, including inversions and a recurrent translocation with 5q35 in T-ALL patients.

Patients & Methods
Genetic characterization of eight cases of acute leukemia of ambiguous lineage (ALAL) with rearrangements involving *BCL11B* and 8q24 as identified by whole genome sequencing (WGS).

Results
- Patients carrying the t(8;14)(q24.21;q32.2) were very difficult to detect by chromosome banding analysis (CBA) only
- A significant *BCL11B* overexpression was detected in patients harboring the translocation t(8;14)(q24.21;q32.2), whereas *MYC* expression was comparable to control levels
- All patients with a t(8;14)(q24.21;q32.2) translocation harbored a *FLT3* mutation (6 *FLT3*-ITD, 2 *FLT3*-TKD)

Conclusion
The identification of t(8;14)(q24.21;q32.2) in cytogenetic routine diagnostics is challenging, but not impossible. The presence of a *BCL11B* rearrangement should be considered in ALAL patients with conspicuous immunophenotype and *FLT3* mutation.

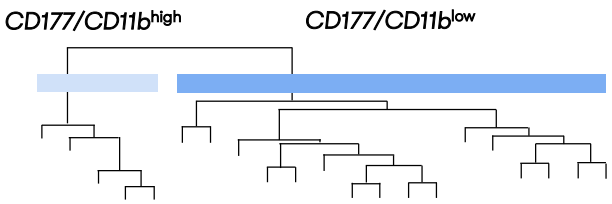
Bendig et al. 2021, *Leukemia & Lymphoma*, https://doi.org/10.1080/10428194.2021.1999436

Comprehensive analysis of the genetic landscape of 21 cases with BPDCN by whole genome and whole transcriptome sequencing

Introduction
Blastic plasmacytoid dendritic cell neoplasm (BPDCN) is a very rare but highly aggressive hematologic malignancy with poor prognosis, a limited understanding of its pathogenesis, and no entity-specific alterations.

Patients & Methods
Genetic profiling of 21 BPDCN cases by whole genome sequencing (WGS) and whole transcriptome sequencing (WTS) to reveal patterns of the molecular pathomechanisms of BPDCN.

Results
- WGS data indicated a complex karyotype in 14/17 (82%) cases
- Deletions of 5q, 9q, 12p, 13q and 15q were frequently observed in the cohort
- The most frequently detected mutations occurred in epigenetic or splicing factor genes (e.g. *TET2*, *ASXL1*, *SRSF2*), followed by mutations in DNA repair genes
- Unsupervised clustering of the gene expression profiles segregated the cohort into two distinct groups that mainly differed in the expression levels of *CD177* and *CD11b*
- Pathway enrichment analysis revealed an activation of immune response in the *CD177/CD11b* up-regulated group



- Increased immune response

Conclusion
The multi-level analysis of our study provides a detailed insight into molecular events underlying BPDCN and suggests novel treatment approaches for these patients.

Summerer et al. 2021, *Leukemia & Lymphoma*, https://doi.org/10.1080/10428194.2021.1924372

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

WGS demonstrates substantial patho-physiological differences of *MYC* rearrangements in patients with PCM and B-cell lymphoma

Introduction
Chromosomal translocations involving MYC are seen in various types of hematological malignancies (especially B-cell neoplasm), impacting disease progression and overall survival.

Patients & Methods
Whole genome (WGS) and whole transcriptome sequencing (WTS) were performed in 385 cases with plasma cell myeloma (PCM), Burkitt lymphoma (BL), diffuse large B-cell lymphoma (DLBCL) and high-grade B-cell lymphoma (HGBL).

Results
- For 150/385 (39%) patients a MYCr could be detected
- MYCr in B-cell lymphoma and PCM differ in their spectrum of breakpoints, range of partners and complexity
- 41 MYC mutations were detected in 27 MYCr cases with a significantly higher frequency in lymphoma cases compared to PCM
- FISH detected 69% (n = 94) of MYCr cases identified by WGS

Conclusion
MYCr features of cases with PCM showed distinct characteristics compared to B-cell lymphoma cases, resulting in poor MYCr detection rates by FISH of only 50% in PCM, compared to 94% in lymphoma.

Bendig et al. 2021, *Leukemia & Lymphoma,* https://doi.org/10.1080/10428194.2021.1964021

Detection of ABL1 kinase domain mutations in therapy naïve *BCR-ABL1* positive acute lymphoblastic leukemia

Introduction
Mutations in the ABL1 kinase domain are the main mechanism of resistance to tyrosine kinase inhibitors (TKI) in Philadelphia-positive (Ph+) leukemia. In acute lymphoblastic leukemia (ALL), a very early acquisition of mutations can be observed, which may even exist before TKI therapy is applied. However, *ABL1* mutation screening is not a standard procedure at initial diagnosis.

Patients & Methods
Resistance mutation screening of 91 *BCR-ABL1* positive ALL patients by NGS before TKI treatment and with a molecular follow-up of at least six month (n = 35).

Results
- Detection of known ABL1 kinase domain mutations in 5/91 (5.5%) patients at initial diagnosis but no difference in outcome was observed
- Longitudinal mutation testing detected one or more known resistance mutation in 15/19 (79%) Ph^+ ALL cases with relapse

Conclusion
The majority (93%, 14/15) of mutations found in Ph^+ ALL cases with relapse were most likely acquired under the selective pressure of (TKI-) treatment and were not present in therapy naïve patients. Therefore, testing at initial diagnosis should only be considered in addition to established mutation testing in refractory/relapsed disease.

Baer et al. 2021, *Haematologica,* https://doi.org/10.3324/haematol.2021.279807

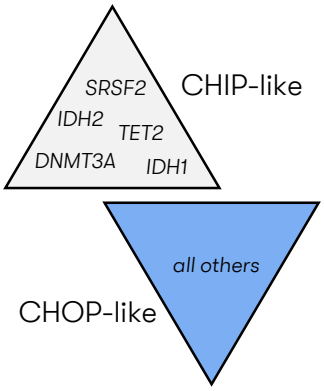
A potpourri of research and opinions – the prosperous work of our guest scientists in 2021

LUCA CAPPELLI

Indeterminate and oncogenic potential: CHIP vs CHOP mutations in AML with *NPM1* alteration

Patients & Methods: Analysis of the mutational profile of 150 consecutive patients (pts) with *NPM1* mutated AML at diagnosis, complete molecular remission (CMR) and relapse by NGS.

Results: In addition to *NPM1*, 301 mutations were detected across all 150 pts at diagnosis. At CMR, 69/150 pts carried at least one mutation and 27% of the pts had persisting DTA (*DNMT3A*, *TET2*, *ASXL1*) mutations. Patients with persisting non-DTA mutations at CMR had a significantly worse EFS and OS. Three clonal evolution pattern emerged: mutations mainly lost at CMR/relapse, persistent mutations at CMR/relapse (CHIP-like) and mutations gained at CMR/relapse (CHOP-like), which are also an adverse prognostic factor in *NPM1*^{mut} AML.



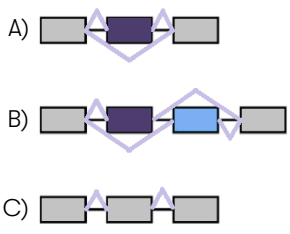
Cappelli et al. 2021, Leukemia, https://doi.org/10.1038/s41375-021-01368-1

ANNA WOJTUSZKIEWICZ

Maturation State-Specific Alternative Splicing in *FLT3*-ITD and *NPM1* Mutated AML

Patients & Methods: Exploration of differential splicing profiles of 382 primary AML samples with *FLT3*-ITD and *NPM1* mutations by WTS.

Results: The co-occurrence of *FLT3*-ITD and *NPM1* mutations is associated with differential splicing of FAB-type specific gene sets but is also characterized by a maturation state independent perturbation of cell cycle control and DNA damage response, albeit involving different genes. Differentially expressed genes in *FLT3*-ITD^{+}/*NPM1*$^{+}$ samples were also FAB-type dependent and included mostly regulators of hematopoietic differentiation. Indicating that differential expression and splicing complement each other in regulating two important aspects of oncogenesis: uncontrolled proliferation & impaired differentiation.



A) exon skipping
B) mutually exclusive exons
C) retained intron
...

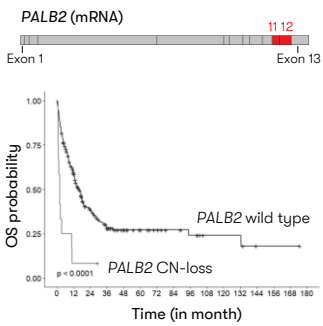
Wojtuszkiewicz et al. 2021, Cancers, https://doi.org/10.3390/cancers13163929

ANTONELLA PADELLA

Loss of *PALB2* predicts poor prognosis in AML and suggests novel therapeutic strategies targeting the DNA repair pathway

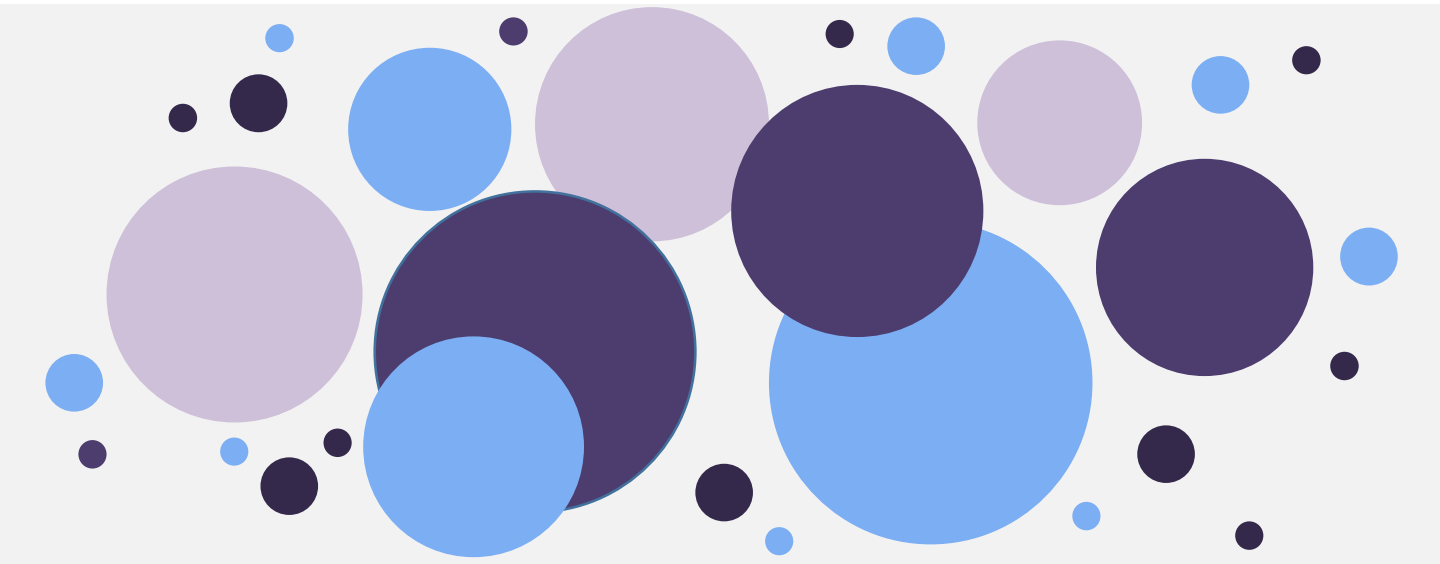
Patients & Methods: Analysis of single-nucleotide polymorphism arrays and whole-exome sequencing data from various data sources to study the genomic alterations of *PALB2* in AML.

Results: A partial deletion of *PALB2* (mainly exons 11 & 12) was detected in ~5% of the patients and was associated with the loss of genomic regions frequently altered in poor-prognosis AML, including chromosomes 5q and 17p13. Patients with a partial *PALB2* deletion frequently harbored *TP53* mutations, were older than wild-type cases and were characterized by a poor prognosis, independent of other negative prognostic factors.



Padella et al. 2021, Blood Cancer Journal, https://doi.org/10.1038/s41408-020-00396-x

The rise of artificial intelligence in routine diagnostics at MLL



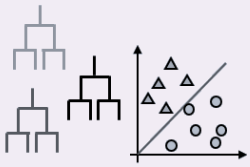
Artificial intelligence is about to make itself indispensable in hematology in the near future

Artificial intelligence (AI) is a term that is hard to escape these days. AI-based methods have revolutionized many aspects of everyday life. One of the most recent applications of AI-powered technology? The semi-automated offside system in soccer, which will be tested during the Arab Cup in the coming weeks. The tremendous success of AI-based methods, especially for image analysis, attracted a lot of attention in recent years. At last, AI is beginning to deliver on the promise of omnipotence made in the 1960s and 1970s when the first AI research was conducted. However, at that time, the high expectations could not be met. So what has changed? Huge increases in computing power have enabled faster training and evaluation cycles, allowing for efficient and accelerated tuning of hyper parameters. In addition, the availability of large, digitized data collections provides the necessary basis to comprehensively train a model.

Artificial intelligence (AI)

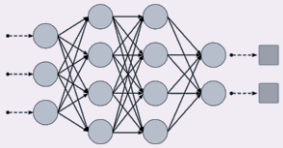
AI intents to mimic human intelligence and decision making, efficiently reflecting human behavior.

Machine Learning (ML)



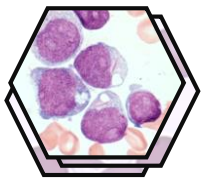
ML algorithms classify data based on knowledge gained from patterns and their statistical representation.

Deep Learning (DL)



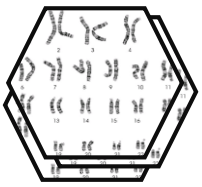
DL models are based on artificial neural networks to learn representations of data by multiple levels of abstraction.

In health care, medical image classification has benefitted the most from the introduction of ML methods so far. Hence, within the last year, we have implemented different ML-based methods at the MLL to successfully support decision-making in cytomorphology, cytogenetics, and immunophenotyping:



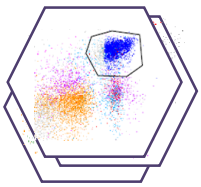
Cytomorphology

In cytomorphology the correctness of the obtained results depends primarily on the experience and capabilities of the personnel to accurately detect and classify aberrant cells. But intra- and inter-observer reproducibility is lower than we would like and, hence, an automated pre-processing and evaluation of the microscopic images could benefit the reproducibility of results and would allow the hematologists to focus on edge cases that do not fit the standard pattern, reducing the overall workload. How we did it? Take a look at the next page.



Cytogenetics

Cytogenetics provides diagnostic and prognostic information retrieved from the classification of chromosomes by size and banding as displayed in a karyogram. However, karyotyping is a very time-consuming and complex task. Multiple attempts have been made to develop (semi-) automatic analysis systems but embedding a ML-based classification system in a routine diagnostic workflow? It hasn't been done that often before, but we have. Let's take a look behind the scenes of the cytogenetics department at the MLL (p.27).



Immunophenotyping

Besides cytomorphology, multi-parameter flow cytometry is the central method for the diagnosis of leukemias and lymphomas. Here, the results also depend heavily on the expertise and knowledge of the person performing the evaluation, with inherent inter-observer variability. Thus, to reduce the dependency on expert knowledge and to increase reproducibility of data interpretation we implemented an automated procedure, that uses an ML-based algorithm to classify various neoplasms. Don't believe it? Then take a look at page 28.

Walter et al. 2021, *Oncogene*,
https://doi.org/10.1038/s41388-021-01861-y

Automated peripheral blood cell differentiation using artificial intelligence

For more than a century, the diagnosis of hematologic neoplasms has relied primarily on the results of cytomorphology, which provides an initial diagnosis and guides other diagnostic methods such as cytogenetics, immunophenotyping or molecular genetics to substantiate the result. However, the correctness of the obtained cytomorphologic results depends largely on the experience and capabilities of the personnel, and even among experienced and skilled hematopathologist, inter-/intra-observer-reproducibility is only 75 to 90%. In addition, manual evaluation can be rather tedious and time-consuming, limiting the number of cells that can be processed per sample and the sample throughput in general. In the last 5 years, the number of samples sent to the MLL has increased by almost 40%, and so has the workload of our hematopathologists. Fortunately, due to the advances in digital microscopic imaging and machine learning technologies, automated image processing and classification have become feasible. In order to streamline and standardize the process of peripheral blood cell differentiation, we set up a workflow to automatically record and digitalize microscopic images of blood-smears. In addition, we took advantage of our large collection of well annotated and digitalized single cell images of blood-smears to train an ML model to identify 21 predefined classes. The classes reflect different cell types and maturation states (Fig.1), including one garbage class for images that could not be assigned unequivocally to either one of the classes by the model.

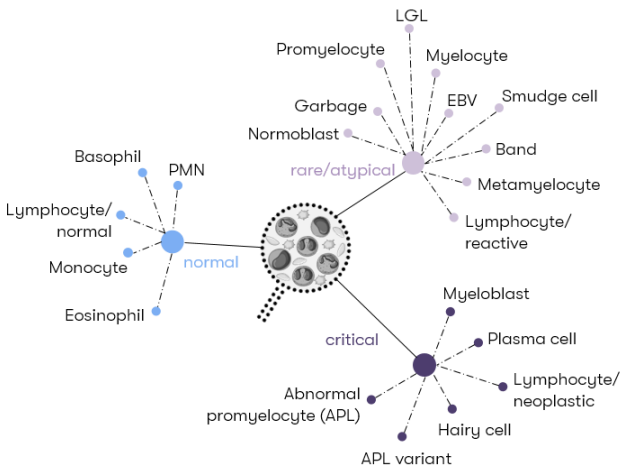


Figure 1: The 21 predefined classes for image classification and their grouping regarding diagnostic value.

The performance of the application in comparison to routine diagnostics is currently evaluated in our prospective clinical trial BELUGA (Better Leukemia Diagnostics Through AI; ClinicalTrial.gov Identifier: NCT04466059. The blood-smears are scanned at 10x magnification to define areas of interest and leukocyte position.

Afterwards, single cell images (144x144px) are generated by a high resolution scan in 40x and fed into the supervised ML model to produce class probability scores for each image/cell (Fig. 2). The ML model is based on ImageNet-pretrained Xception using Amazon Sagemaker.

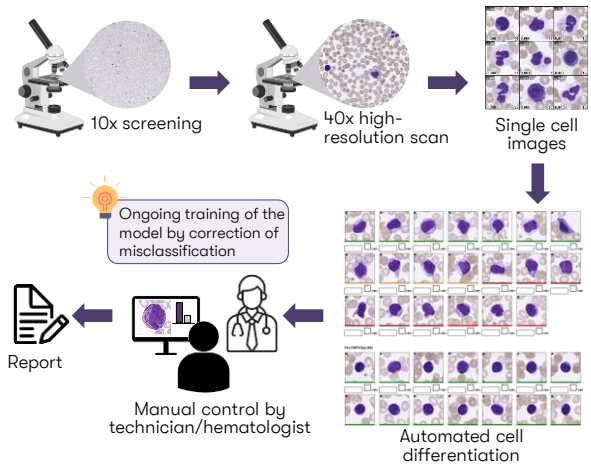


Figure 2: Pictographic display of the AI-supported diagnostic workflow.

In a first interim analysis, 10,082 patient samples, sent to the lab between January and July 2021 with a suspected hematologic neoplasm, were classified. In routine diagnostics, for each sample ~100 cells are analyzed to keep a healthy balance between evaluation time and diagnostic accuracy. The automated workflow has a significant speed advantage, allowing the assessment of ~500 cells per sample in a comparable time frame (on average 4:37 min per sample). Comparing the results for cell differentiation between humans and the ML model revealed a high consensus (see below).

52%	Segmented Neutrophils	53%
2.25%	Eosinophils	3.36%
0.72%	Basophils	0.72%
7.5%	Monocytes	6.64%
31.7%	Lymphocytes	24%
0.97 %	Pathogenic blasts	1.65%

Taking a closer look at the results for pathogenic cases, routine diagnostics identified 536 samples with blast cells and 2,323 samples with at least one atypical/malignant lymphocyte, of which the ML model identified 493 (91%) and 2.279 (98%) cases, respectively. The next step will be to transfer the knowledge to microscopic images from bone marrow samples – take a look at the next page.

A glimpse at what's coming next: AI-based differentiation of bone marrow cell morphologies

As demonstrated on the previous page, the automated cell differentiation from peripheral blood smears works quite well and the method is currently tested in parallel to routine diagnostics in a prospective study. However, in addition to peripheral blood we also receive bone marrow samples because for some hematological malignancies only bone marrow samples are informative for an accurate and comprehensive diagnosis. By comparing Fig. 1A and Fig. 1B the challenge is immediately obvious: the cell density is significantly higher in bone marrow smears compared to peripheral blood, hampering the detection of single cells.

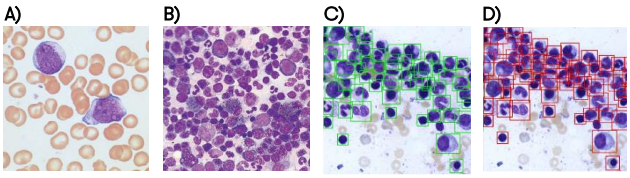


Figure 1: Exemplary digital microscopic images for A) peripheral blood smear, B) bone marrow smear, C) manual cell labeling (green rectangles), D) automated object (= cell, red rectangles) recognition supported by AI.

Therefore, a pre-screening of the microscopic image is necessary to identify areas of high quality and single cell resolution. Next, specialized software systems are applied to automatically locate the single cells in the respective image and to highlight them by drawing a quadratic region around each cell. The last part is only intended to facilitate manual correction and comparison and doesn't serve any diagnostic purpose. As can be seen in Fig. 1D, the automated detection of cells is already comparable to the manual labeling (Fig. 1C). However, it's also obvious that a quadratic frame around a circular object in a crowded space will almost always introduce some noise. Hence, we are currently evaluating, if a more accurate labeling of the cells, i.e. highlighting the exact borders of a cell (Fig. 2), rather than simply framing a cell, can reduce background noise to further increase classification accuracy.

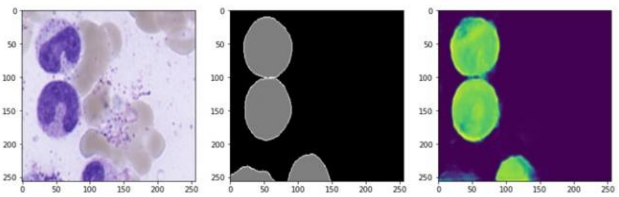


Figure 2: The process of manually outlining the different cells of a digitized microscopic image.

To get one step closer to the automated analysis of bone marrow smears, we teamed up with the Institute of AI for Health (Helmholtz Munich) to analyze the largest available expert-annotated pool of bone marrow cytology images.

In a joint effort, a dataset of 171,374 expert-annotated single-cell images from 945 patients diagnosed with a variety of hematological neoplasms was used to train a convolutional neural network (CNN) for the classification of bone marrow morphologies. The dataset was further extended by images displaying artefacts, cells that could not be identified, and cell types that were not part of the classification, to avoid biasing the annotation and to improve robustness and generalizability. The images were centered on the annotated cell but it was left to the algorithm to detect the main image content without any additional processing of the images. The images were fed into a ResNeXt-50 model and the highest prediction probability yielded the class per image. As expected, the results per class improved with an increase of available training data and maturation states closely related in the leukocyte differentiation were often misclassified (Fig. 3). However, some of the misclassifications can be considered tolerable, i.e. the confusion between segmented and band neutrophils, and, hence, tolerance classes were used to account for this uncertainty.

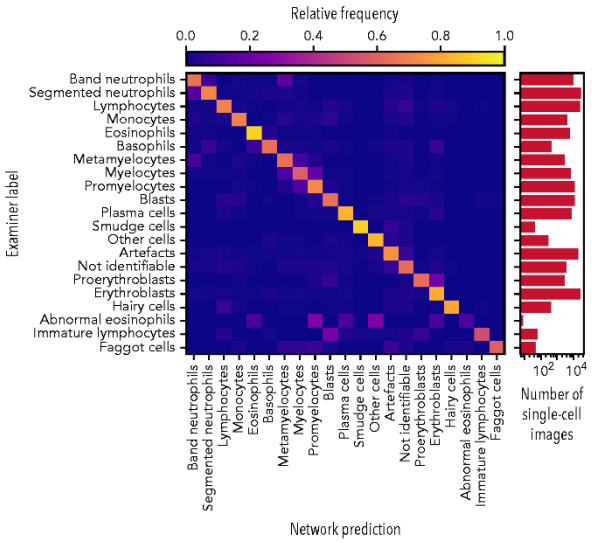


Figure 3: Confusion matrix of the predictions obtained by the ResNeXt classifier on the test database annotated by gold standard labels provided by human experts. Image taken from *Matek et al. 2021*

A detailed analysis of the network's classification decisions showed that the algorithm had learned to differentiate between the relevant input and background noise to focus on the leukocyte for the class prediction. A visual display of the extracted features confirmed the clear separation of the classes and reflected the continuous transition between the classes. The results of the two studies show, that AI-supported cell differentiation is feasible, independent of the sample material.

Haferlach et al. 2021, *Blood,*
https://doi.org/10.3324/haematol.2021.279807

Matek et al. 2021, *Blood,*
https://doi.org/10.1182/blood.2020010568

AI substantially supports chromosome banding analysis in hematologic diagnostics

Chromosome banding analysis (CBA) is the current gold standard to identify cytogenetic abnormalities that allow patient stratification for prognostic prediction in hematological malignancies. Patient-specific information is derived from the classification of chromosomes by size and banding, as displayed in a karyogram. However, the generation of a karyogram is a very time-consuming and complex process in which viable cells are cultured and arrested in the metaphase stage of cell division before the fixed cell suspension is dropped on the slides to perform chromosome banding and staining. Metaphases are then captured, chromosomes separated and karyotyping is performed (Fig. 1). Here, the banding pattern of the chromosomes is essential for highlighting diagnostically important details and distinguishing between normal and aberrant chromosomes.

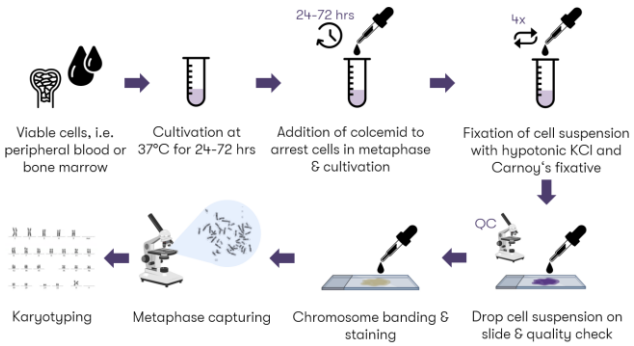


Figure 1: Simplified pictographic overview of the chromosome banding analysis workflow.

CBA also plays an important role for therapy selection and short turn-around times are desirable for rapid, safe and efficient treatment. In recent years, an increase in the degree of automation for the wet lab processes could be observed with the development of robots that handle every step of the workflow. But karyotyping is also a very complex and time-consuming task that would benefit from automated processing. Despite noticeable progress, many challenges in automatic karyotyping haven't yet been fully resolved. First, the identification and separation of individual chromosomes. As of Feb 2021, our routine CBA workflow includes automatic chromosome separation based on object recognition instead of contrast differences, reducing the need for manual intervention. It follows the classification of the individual chromosomes by assigning them to their respective position in the karyogram. To automate the procedure, we build an AI-based classifier for normal chromosomes. 100,000 manually arranged karyograms with normal karyotype were taken from our digital archive to train a deep neural network (DNN) that determines 1) chromosome class and 2) chromosome orientation. The initial DNN training took 16 days on a Nvidia RTX 2080 Ti graphic card with 4352 cores.

The model was subsequently validated on an independent set of 500 karyograms with normal karyotype, resulting in a classification agreement of 98.6% between the human experts and the DNN on a per chromosome basis (Fig. 2).

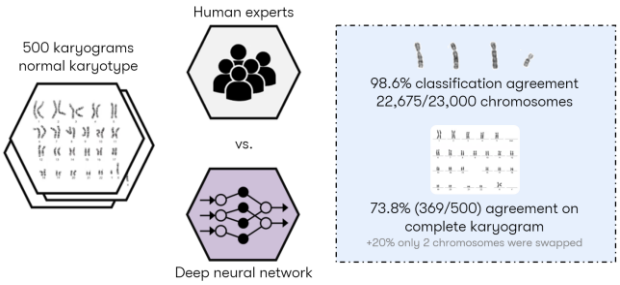


Figure 2: Depiction of the results from the model evaluation.

For complete karyograms, a higher discrepancy was detected. An evaluation of the results revealed, that the majority of misclassifications involved chromosomes very similar in size, shape, and appearance (chr 4 & chr 5; chr 14 & chr 15), which are also challenging to classify for humans, if chromosome quality is suboptimal. After the introduction of the first AI-based model, the number of cases reported <5 days increased compared to the manual workflow. Since July 2021, it's possible to classify all recorded metaphases per patient at once, further increasing the number of cases reported within 7 days (Fig 3.).

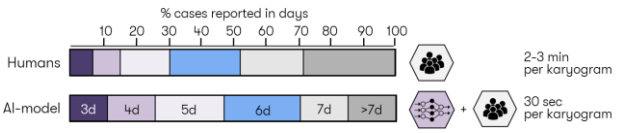


Figure 3: Comparison of turn-around time between the AI-supported and the manual workflow.

The current version of the AI-supported karyotyping also reports the probability of the class prediction, facilitating manual verification (Fig. 4).

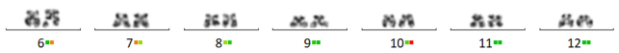


Figure 4: Section from a karyogram showing the color-coded class probability depicted beneath each chromosome: green 100%, yellowish 50%, and red close to 0% .

Most of the presented results are for normal karyotypes but the algorithm works also quite well for aberrant karyotypes. Here, numerical aberrations are easier to handle compared to structural aberrations because they involve normal chromosomes and the method could be easily extended for this task. Structural abnormalities are more challenging, mainly because the available training data is limited and not easily expandable. However, derivative chromosomes that are different from any normal chromosome are left out for manual classification, saving also time for cases with abnormal karyotype. Thus, step by step, we are getting closer to fully automate karyotyping.

Automated diagnostics of hematologic neoplasms by AI-based models using flow cytometric data

Multiparameter flow cytometry (MFC, immunophenotyping) is an integral part of routine diagnostics for hematologic malignancies. Due to the high dimensionality of the data, computational assistance is indispensable for the evaluation of MFC data because it's basically impossible for humans to draw any conclusion from the raw files generated by the cytometer. However, the provided software usually performs only relatively simple data processing and visualization tasks to facilitate data analysis for human experts. Quality management systems have been implemented worldwide to standardize the wet-lab processes such as sample preparation and measurement, but data analysis and interpretation are still the responsibility of humans, completely relying on expert knowledge applied individually to each patient sample. To reduce the dependency on expert knowledge and to potentially increase consistency of data interpretation, the implementation of automated processes is desirable (Fig. 1).

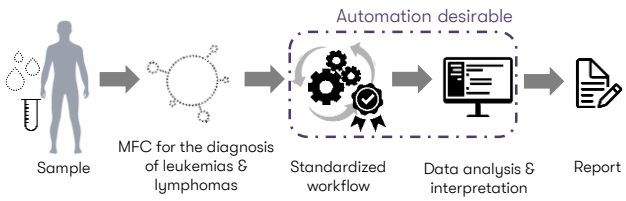


Figure 1: Simplified overview of the MFC workflow.

Our first approach to apply an AI-based model to solve a multiclass MFC problem with expert accuracy was conducted in collaboration with the university of Bonn. Here, the flow cytometric data was transformed into self organizing maps (SOM, Fig. 2) to allow an image analysis in 2D. The SOMs were then used to train a convolutional neural network to automatically classify seven B-cell non Hodgkin lymphoma (B-NHL) subtypes (e.g. chronic lymphocytic leukemia, follicular lymphoma, mantle cell lymphoma) and healthy individuals.

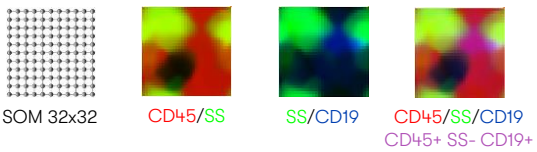


Figure 2: Exemplary SOMs for the visualization of various markers.

The model was able to reliably distinguish between B-NHL and healthy individuals with an average F1 score of 0.98 (max value = 1). B-NHL subtype classification yielded a weighted F1 score of 0.94 and a detailed analysis of the predictions showed that misclassifications occurred more often between clinically similar subtypes due to the close resemblance of their flow cytometric profiles.

Zhao et al. 2020, *Cytometry,*
https://doi.org/10.1002/cyto.a.24159

Although the first attempt at subtype classification worked reasonably well, the preceding image transformation has the potential disadvantage of data reduction. Hence, in a second attempt, we relied on the raw data from MFC for the classification of main entities of hematologic malignancies to take full advantage of the complex data.

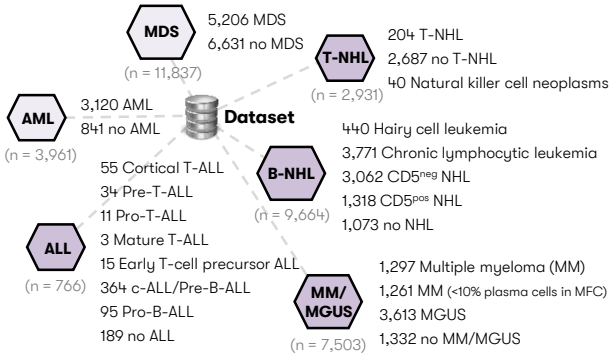


Figure 3: Overview of the used dataset.

In total 36,662 cases were used to train six ML-based models: one model per disease subtype (Fig. 3). The data was standardized, transformed and rescaled and feature engineering techniques were applied for each model to extract the most relevant features per entity. In addition, expert-based features, including focusing on cell populations of interest by applying clustering techniques, determining the marker distribution for subpopulations and calculating covariance between key markers, were applied for certain subtypes.

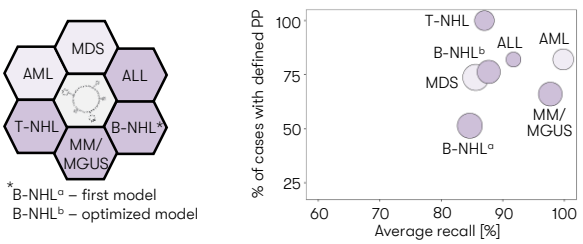


Figure 4: Performance of the different models for subtype classification. The size of the bubbles is proportional to the number of included cases.

The average recall values for the different models ranged from 84% to 99.8%, considering only cases with prediction probabilities above certain thresholds (Fig. 4). Thus, relying on raw MFC data to train AI models is feasible and leads to promising classification results, paving the way for implementing the models in routine diagnostic settings. We anticipate that the trained models will substitute up to 75% of routine MFC data analysis in the future. Our next steps will focus on the identification of additional entities, the application of transfer learning to achieve universal applicability, and the extension of the models to detect also measurable residual disease patterns.

Bellos et al. 2021, *Blood,*
https://doi.org/10.1182/blood-2021-150697

Not quite there yet, but on the right track: AI-based models in clinical molecular genetics

The previous pages have shown the enormous potential for implementing AI-based models to support diagnostics in hematology, with the exception of one prominent field: molecular genetics. As mentioned before, the recent success of AI-based models in diagnostics is mainly based on the availability of significantly improved image recognition algorithms, powered by tech giants such as Google, Amazon and Meta. However, data from molecular genetics, especially NGS, are usually not in the form of images, limiting the application of image classification software. Molecular genetic data matrices, such as mutation profiles, are also very sparse because each patient harbors only a small number of mutations, making it difficult for an algorithm to identify the relevant features. In addition, analyzed gene panels are often inconsistent among patients, depending on the suspected diagnosis and associated genes, as well as the institute/laboratory that performs the sequencing. In contrast to the other fields in hematology, that implement AI-based models to mimic human behavior, clinical molecular genetics intents to apply AI-supported models within precision medicine workflows to perform tasks that are impractical to do for humans, i.e. risk stratifying patients based on multiple factors, such as clinical variables and gene mutations (Fig. 1).

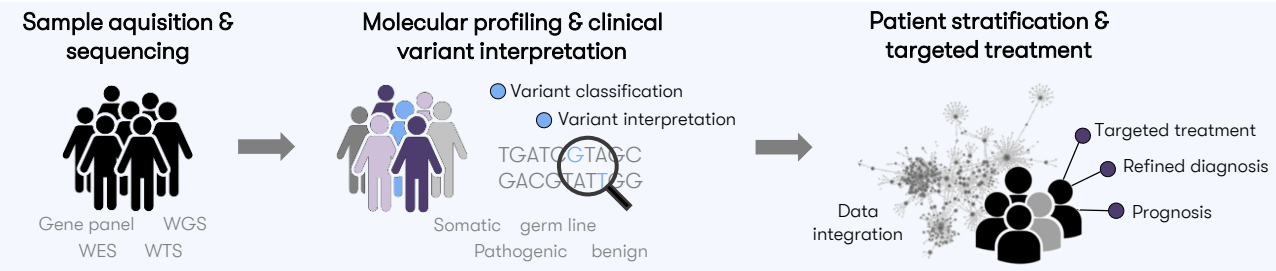


Figure 1: Pictographic overview of the general concept of precision medicine.

Recently, two attempts were made to improve risk stratification of MDS patients by applying ML techniques for which a large MLL data collection contributed significantly to the training of these models. *Nazha et al.* focused on the analysis of 1,471 MDS patients to build a personalized prediction model that can provide survival (OS) and leukemia transformation probabilities at different stages of a patient's disease course. Accurate outcome predictions and identification of relevant factors can help physicians determine appropriate therapies. The analysis was based on the mutation status of 24 genes with established clinical impact in MDS that were mutated in at least 30 patients in the cohort, cytogenetic information, as well as various clinical variables such as bone marrow blast count, hemoglobin, and platelets. Interestingly, the number of mutations per patient was an independent prognostic factor for OS and leukemia transformation. The final model outperformed the established risk scores (IPSS and IPSS-R) even when mutations were added to the scoring systems (Fig. 2).

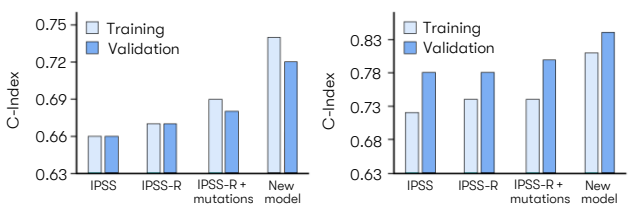


Figure 2: The concordance index of the new model compared with that of IPSS and IPSS-R for (A) OS and (B) leukemia transformation in the training and validation cohorts. Adapted from *Nazha et al.*

Nazha et al. 2021, Journal of Clinical Oncology, https://doi.org/10.1200/JCO.20.02810

Radakovich et al. developed a machine learning model for the diagnosis of myeloid malignancies independent of bone marrow biopsy data. 15 genomic/clinical variables (selected due to their importance for the prediction) were included in the final model to distinguish between MDS and other conditions (Fig. 3).

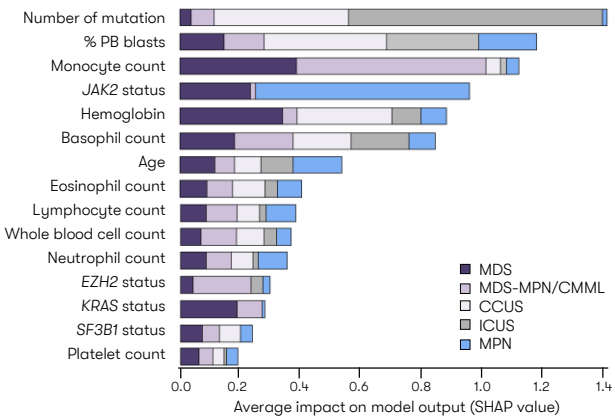
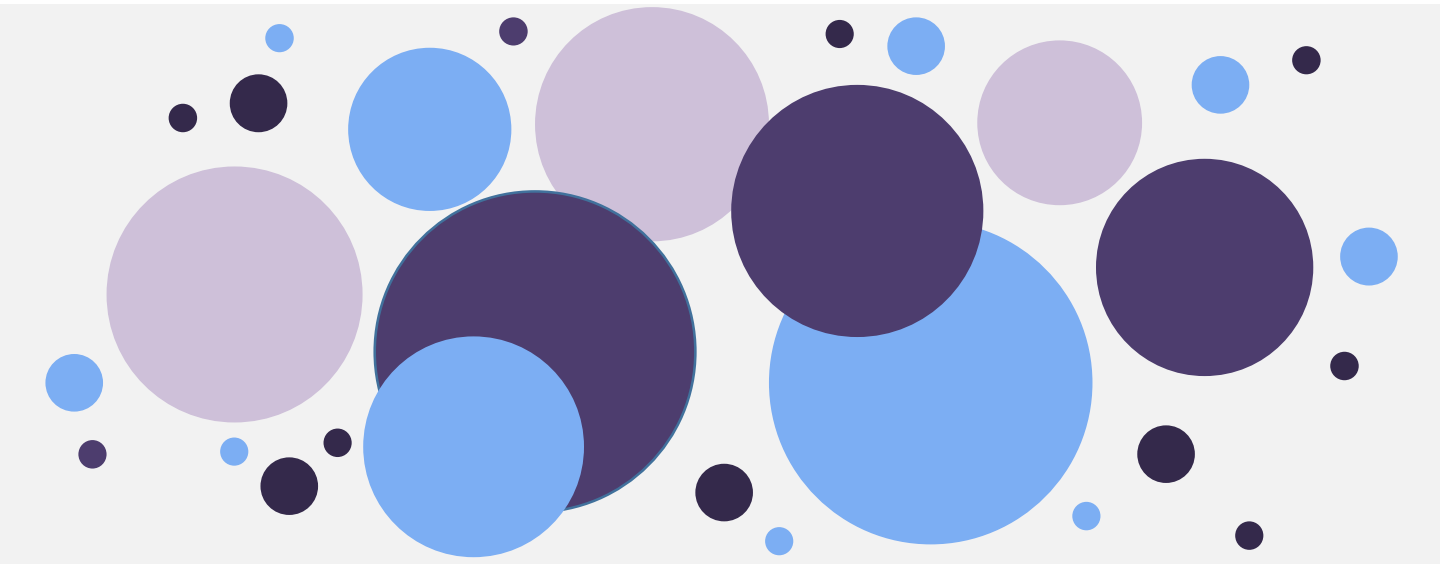


Figure 3: Global model feature importance. The larger the Shapley (SHAP) value, the greater the impact of the variable on the output. PB: peripheral blood. Adapted from *Radakovich et al.*

The model generated highly accurate predictions in the validation cohort with a AUROC (Area Under the Receiver Operating Characteristic curve) value of 0.93 (values > 0.9 are considered excellent). Although the results of these studies are not universally applicable, they still demonstrate the potential value of ML-based models for disease classification and risk stratification.

Radakovich et al. 2021, Blood Advances, https://doi.org/10.1182/bloodadvances.2021004755

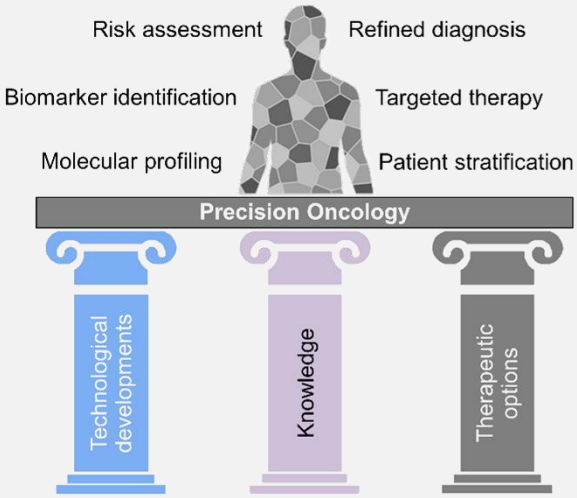
Outlook and future collaborations



Combining forces for the development of innovative methods and approaches to advance patient care

The rocky path towards precision medicine

Various assays and diagnostic methods have been developed to address the clinical challenges of delivering a rapid and accurate diagnosis for each cancer patient. However, to date, clinical practice has mainly focused on the one-genetic-alteration-one-drug approach and population averages, ignoring the unique characteristics and needs of the individual patient. Further technological and algorithmic improvements and the increase in knowledge through various large-scale molecular studies will lead to more personalized models that allow a tailored treatment and improved risk assessment for each patient. Furthermore, it is indispensable to increase the diversity of research participants to ensure that personalized treatments or prevention strategies are available to all.



To get one step closer to the final goal of personalized medicine, we will further expand our network of fruitful collaborations. It is particularly worth mentioning that for the first time we will be participating in two innovative training networks (ITN) as part of Horizon 2020, the biggest EU Research and Innovation program.

HARMONY

One of the key elements in shortening the process for developing new drugs or treatment approaches and creating comprehensive disease models is the collection and harmonization of high-quality data on the outcomes of existing treatments. However, the collection of real world data is often hampered by lack of data as well as variations in health care practice. The Harmony alliance is a European network of excellence made up of >80 organizations to collect data on as many patients with blood cancer as possible. The data is brought together and harmonized on one large platform.

https://www.harmony-alliance.eu/

GenoMed4All

As described on the previous pages, we are very interested in adopting AI-based methods to further improve diagnostics for the benefit of our patients. However, the high complexity of AI-based methods is both a curse and a blessing, because on the one hand it enables significant performance improvements, but on the other hand it limits the interpretability of the model, which leads to uncertainties about how they work. Therefore, we were happy to join GenoMed4All, which aims to provide reliable, interpretable, and trustworthy AI-services for clinical support in various hematological diseases.

https://genomed4all.eu/

INTERCEPT-MDS

The project aims to improve diagnostics and treatment for MDS by characterizing changes in cellular heterogeneity and identifying early disease cells through the analysis of multi-omics data. Unraveling the regulatory mechanisms that determine cell-to-cell heterogeneity will be essential to identify key regulators (e.g. epigenetic modifiers) as potential points of intervention. For our PhD student the work will start with a detailed analysis of in-house WGS & WTS data from MDS samples to generate hypotheses that will subsequently be tested and expanded by single-cell analysis.

https://intercept-mds.eu/

Outlook

Speculating about the future is always a risky exercise, especially since as scientist we normally deal with facts and not opinions. However, the past can always serve as a guide for future actions. Looking back on 2021, we close a successful scientific year that already paves the way for the extended application of new diagnostic technologies and methodologies in 2022.

Recent publications and the initiation of large-scale sequencing studies all around the world demonstrate the increasing scientific and clinical value of comprehensive genomic assays such as WGS and WTS. As demonstrated by Duncavage and colleagues, WGS provides a greater diagnostic yield than conventional cytogenetics and rapid and accurate genomic profiling for patients with AML or MDS (*Duncavage et al.* 2021, PMID: 33704937). Our own results of a comprehensive comparison for patients with ALL and AML between the results of gold standard techniques and WGS support the these conclusions. In addition, we could show that the use of WTS for patients with ALL provides additional diagnostic and prognostic information. Therefore, we are very supportive of this comprehensive approach and will focus our efforts next year on the gain of clinical information by a hematologic WGS/WTS for AML and ALL patients.

The new WHO is in progress and suggests that genetics, and in particular molecular genetics, will gain importance for classification, which has to be demonstrated and supported in routine settings. Here, too, we are keen to make our contribution.

Of course, we hope to welcome guest scientists again in the new year, which is already assured by our participation in various international research projects. Within the projects, we are looking forward to welcoming scientists at the MLL and to expand our biological and clinical knowledge about the different leukemias and lymphomas but also our technical knowhow by the optimization and implementation of new AI-based models. It's always our goal to provide state-of-the-art diagnostics and therefore we will also expand our NGS portfolio by establishing new assays and techniques, such as single cell sequencing.

We remain curious and motivated to further advance science in the field of hematologic neoplasms in the coming year.